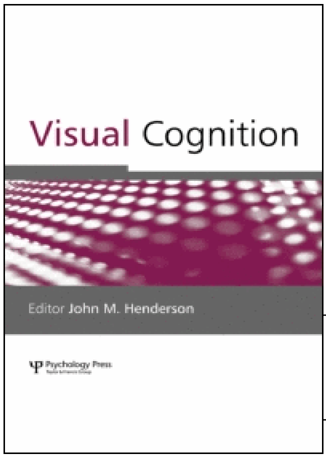

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683696

An analysis of letter expertise in a levels-of-categorization framework

Alan C. -N. Wong [a]; Isabel Gauthier [a]
[a] Vanderbilt University, Nashville, TN, USA

First Published on: 01 December 2006

To cite this Article: Wong, Alan C. -N. and Gauthier, Isabel (2006) 'An analysis of letter expertise in a levels-of-categorization framework', Visual Cognition, 15:7, 854 — 879

To link to this article: DOI: 10.1080/13506280600948350
URL: http://dx.doi.org/10.1080/13506280600948350

VISUAL COGNITION, 2007, 15 (7), 854–879

Psychology Press
Taylor & Francis Group

An analysis of letter expertise in a levels-of-categorization framework

Alan C.-N. Wong and Isabel Gauthier
Vanderbilt University, Nashville, TN, USA

While there has been increasing effort in dissociating the neural substrates recruited by perception of different objects, the theoretical and behavioural work needed to understand such dissociation lags behind. In an attempt to compare expertise in letter and face perception, we outline a theoretical framework that characterizes different types of object expertise based on the task demand (level of abstraction) required in object categorization. Face perception requires categorization at a subordinate level, whereas letter perception involves mainly basic-level categorization. Accordingly face and letter perception should represent two different types of expertise and display different neural and behavioural markers. Results from three behavioural experiments supported the predictions of the framework in that letter expertise is characterized by an enhancement of the basic-level advantage, instead of its attenuation as typically found for face perception. We compare this framework with Farah's taxonomy of visual abilities based on cooccurrence of deficits in visual agnosias.

Object, letter, and face processing have often been considered different domains of study in Psychology. For instance, neuropsychologists classify visual recognition deficits in the broad categories of visual agnosia, alexia, and prosopagnosia (Farah, 2004). In recent years, a growing number of studies have highlighted neural differences between the processing of these different object categories. Electrophysiological and brain imaging studies have revealed ventral occipitotemporal regions (with a left preponderance) that are more recruited by the processing of letters and letterstrings relative to that of other categories (Beauregard et al., 1997; Bentin et al., 1999; Cohen et al., 2000; Flowers et al., 2004; Garrett et al., 2000; Gauthier et al., 2000; Howard et al., 1992; James, James, Jobard, Wong, & Gauthier, 2005; Petersen et al., 1990; Puce, Allison, Asgari, Gore, & McCarthy, 1996;

Please address all correspondence to Alan Wong, Psychology Department, Vanderbilt University, 301 Wilson Hall, Nashville, TN 37203, USA. E-mail: alan.wong@vanderbilt.edu

This work was supported by a grant from the James S. McDonnell Foundation to the Perceptual Expertise Network and NEI grant EV13441-01 to IG.

http://www.psypress.com/viscog DOI: 10.1080/13506280600948350

Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999; Wong et al., 2006). Similarly, faces and other objects have also been dissociated, with higher activation for faces mainly in the right ventral occipitotemporal region including parts of the fusiform gyrus (Allison et al., 1994; Kanwisher, McDermott, & Chun, 1997; Puce et al., 1996; Xu, 2005). Apart from the left and right preponderance for letter- and face-selective regions, a few studies have also provided more subtle dissociations within hemispheres, with the letter-selective regions more lateral than the face-selective ones (Gauthier et al., 2000; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Wong et al., 2006). Different neural substrates suggest that different mechanisms may be involved in the processing of faces, objects, and letters.

The distinction between the processing of faces and that of other objects has been aided by behavioural work establishing qualitative differences between these tasks, such as studies of the face-inversion effect (Yin, 1969) and part–whole effect (Tanaka & Farah, 1993). Theoretically, a framework that describes object perception as a process of categorization at different levels of abstraction (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) has been used to account for the differences between face and object processing in terms of expertise (Tanaka, 2001; Tanaka & Taylor, 1991). For instance, people can identify a face (e.g., as "Elvis Presley") as fast as they can say it is a face (Tanaka, 2001). In contrast, objects are generally categorized faster at the basic level (e.g., bird) than at a subordinate level (e.g., robin). However, expert observers such as bird watchers can accomplish both tasks equally fast (Tanaka & Taylor, 1991). Based on the object categorization framework, recognition of most objects can be characterized as basic-level categorization by default, whereas face perception (and expert perception of other objects) generally requires categorization at a subordinate level. The prolonged experience in fulfilling the different task demands for faces and other objects (i.e., subordinate- vs. basic-level categorization) leads to different behavioural phenomena and neural substrates associated with them (Bukach, Gauthier, & Tarr, 2006; Gauthier, 2000). This framework has also facilitated later studies of perceptual expertise with novel objects by providing an operational definition of expert performance, namely, that categorization at a subordinate level is as fast and accurate as basic-level categorization (Gauthier & Tarr, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Gauthier, Williams, Tarr, & Tanaka, 1998).

Much less is known about the nature of the differences between letter and object perception. Letters have rarely been explicitly contrasted with other object domains in theoretical and behavioural work. Instead, letter perception appears so integral to the problem of word recognition that letters are almost exclusively studied from a psycholinguistic perspective. For example, studies have revealed how letters are better recognized in the context of words (e.g., McClelland, 1976; Reicher, 1969). Theories of letter and word recognition also put a lot of emphasis on linguistic (orthographic, phonological, semantic) processing (e.g., Johnson & Pugh, 1994; Perfetti, Liu, & Tan, 2005). Even studies of letter perception that contrast it with the processing of other shapes like digits do not directly address whether letter perception is supported by the same mechanisms that mediate common object recognition (Polk & Farah, 1998). Apparently, there are so many differences between letter perception and object perception that it might not be meaningful to compare these domains. Yet some theories of letter perception (e.g., Dehaene, Cohen, Sigman, & Vinckier, 2005; McClelland & Rumelhart, 1981; Oden, 1979) rely on processes resembling those in some theories of object perception (e.g., Riesenhuber & Poggio, 1999), suggesting possible links between these fields of research.

In many ways, letter perception can be considered a kind of expertise. Most of us devote a considerable amount of our waking time reading, and most modern life activities, like driving or watching television, rely on recognition of letters in some form. We also recognize letters with amazing speed and efficiency (LaBerge & Samuels, 1974). Neuroimaging studies reveal specialization in extrastriate cortex not only for words and pseudowords (Cohen et al., 2000; Peterson et al., 1990; Pugh et al., 1996) but also for letterstrings (Peterson et al., 1990; Polk & Farah, 1998; Puce et al., 1996), and even for single letters (Flowers et al., 2004; James et al., 2005; Longcamp, Anton, Roth, & Velay, 2003; Wong et al., 2006). But how is our expertise with letter perception related to face expertise? As mentioned above, letter and face perception recruit different neural substrates within the visual system, but there lacks an account of why and how such differences in cortical specialization arise. Here, we extend the categorization framework that has already proven useful for understanding object perception (Rosch et al., 1976; Tanaka & Taylor, 1991) and expert processing of faces and objects (Gauthier, Anderson, Tarr, Skudlarski, & Gore, 1997; Gauthier et al., 1998; Tanaka, 2001) to the processing of single letters. By comparing letter and face expertise within this common framework, we can begin to establish a taxonomy of expertise. This approach should result in a better understanding of why visual expertise with letters recruit different mechanisms from visual expertise with dogs, birds, cars, and faces. Also, it should provide a basis for predicting the behavioural and neural characteristics accompanying types of visual expertise yet to be studied in detail.

THE STRUCTURE-BASED FRAMEWORK OF OBJECT TAXONOMY

The framework we propose regards object perception as the process of categorizing objects at different levels of abstraction. For example, the same object can be correctly labelled an animal, a bird, a parrot, or Uncle Wong's parrot. Rosch and her colleagues (1976) proposed a framework of concrete object taxonomy based on their visual structure. They argued that "there is generally one level of abstraction at which the most basic category cuts can be made" (p. 383). This level, termed the basic level, contains the most inclusive categories in which members are similar to each other but dissimilar to members of other categories (definition of similarity is discussed below). All levels more general than the basic level are termed superordinate, whereas all levels more specific than the basic level are termed subordinate (Figure 1).

Before explaining why the levels are defined in this way, it is important to note that the concept of the basic level has been used not only in perception but also in nonobject domains such as concept categorization, and its definition thus involves factors other than visual structure (e.g., knowledge, experience). Here, however, for the purpose of differentiating the task demands of visual object categorization, we focus only on the visual aspects. We ask: What would be the basic level for letters and faces if only structure was considered? To differentiate the concept of the basic level used here with the more general connotation in the literature, we term it the "structural basic level" in Figure 1.

There are numerous possible ways to define object structures, depending on the knowledge used (Murphy & Medin, 1985). Here, we focus on an a priori form of structure free from the effects of knowledge and experience.

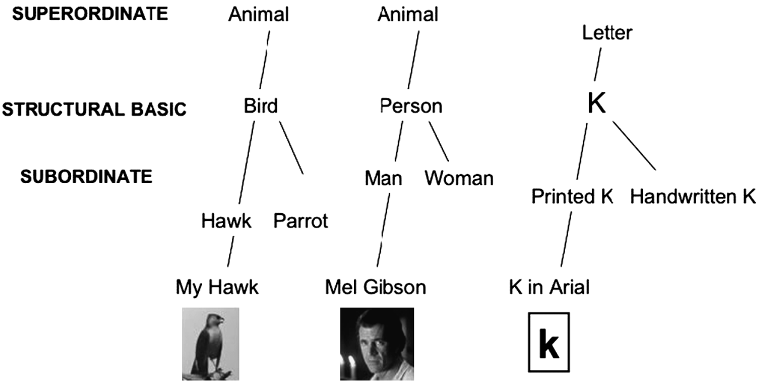

Figure 1. Examples of categorization at different levels of abstraction for common objects, faces, and letters.

Two concepts are essential in our definition of visual similarity: qualitative and quantitative properties. Qualitative properties refer to the parts of an object and the coarse spatial relations (e.g., above, below, beside) among these parts. For example, that faces have two eyes above a nose forms a description of qualitative properties. Quantitative properties refer to the metric properties of an object (e.g., exact shape of each part) and more detailed relations among the parts, such as the distance between two eyes (Carey & Diamond, 1994). We define the structural basic level as the most general level at which members of a category share the same qualitative properties and differ only in quantitative properties. The result is that morphing members of a basic-level category can result in a shape recognizable as a member of that category, as described below.

Let us use human faces and letters to illustrate the point. Human faces can be regarded as a basic-level category, as all faces share the same parts (eyes, nose, mouth, ears) as well as coarse relations among parts (e.g., the eyes are above the nose), but differ from other object categories qualitatively (e.g., birds have wings but human faces do not). Faces generally differ from each other only quantitatively (e.g., curvature of the lips and the distance between the eyes). Therefore, morphing of different human faces results in an average image recognizable as a human face (one that tends towards the central tendency and is generally considered attractive; Langlois & Roggman, 1990). The same idea applies to the letter category "M". The "M" category is a basic-level category because this is the most general level where members share with each other the same parts (two vertical and two oblique strokes) in the same qualitative relations, while differing from members of other categories (e.g., Z) in qualitative properties. Within the "M" category members differ in metric properties (e.g., length of the slant strokes and the angles at which the vertical strokes connect to the oblique strokes). Therefore, different "M"s can be averaged together to form a shape recognizable as a member of the "M" category.

With the definition of the basic level specified, we can now look at how objects and letters are organized in this structural taxonomy (Figure 1). A distinction between a "human" and a "bird", for example, would be made at the basic level. Finer discrimination among different persons (e.g., identifying "Mel Gibson") or different bird species (e.g., recognizing a "hawk") requires categorization at a subordinate level. Distinction between a living thing and nonliving thing, or between an animal and a letter, requires categorization at a superordinate level. As discussed before, identifying letters (e.g., as a "k") corresponds to categorization at the basic level. More specific instances of a certain letter (e.g., "k" in different fonts) represent categories at different subordinate levels. Categories such as "Roman letters" and "letters" belong to superordinate-level categories. Given this, a key difference between face and letter perception already emerges: Whereas face identification requires subordinate-level categorization, letter identification requires categorization at a basic level.

It is worth noting that, in this framework, discriminating between the upper and lower cases of a letter can be a basic-level or subordinate-level task. It is a basic-level task for letters with qualitative differences between the upper-case and lower-case versions (A, B, D, E, F, G, H, I, J, L, M, N, Q, R, T, Y). For other letters, different cases share qualitative structures and differ mainly in size (C, K, O, P, S, V, W, X, Z), and discrimination among cases works at a subordinate level. We choose to regard letter perception as mainly a basic-level task because, most of the time, discrimination among letters is much more important than discrimination between cases.

Some influential theories of letter perception have focused on the format of letter representations. Posner and colleagues (Posner, 1978; Posner & Boies, 1971; Posner, Boies, Eichelman, & Taylor, 1969; Posner & Mitchell, 1967) suggested that visual presentation of a letter activates stored representations of the form and name of the letter (visual and phonetic codes). Different versions of a letter (e.g., A and a) have different visual codes while sharing the same phonetic code (/a/). A number of studies questioned the phonetic format of the code, and provided support for the existence of an abstract letter identity, a representation invariant from the size, case, orientation, etc., of a letter (Besner, Coltheart, & Davelaar, 1984; Bigsby, 1988; Coltheart, 1981; Evett & Humphreys, 1981). Although we are interested in letter and object categorization at different levels of abstraction, we make no assumption regarding whether there are separate representations of an object at different levels of abstraction. A rough correspondence could be drawn though, in that the visual code represents information at or below the basic level, while the phonetic code or the abstract letter identity can be regarded as representations containing details at the superordinate level, because they are invariant to case differences.

DIFFERENT TASK DEMANDS FOR LETTER AND FACE PERCEPTION

Perceptual performance is determined by the interaction between the perceiver, the stimuli, and the task demands (Bukach et al., 2006; Schyns, 1998). Perceptual expertise illustrates how these factors interact. Whereas letter perception requires distinguishing different letters and thus basic-level categorization, face perception (as well as dog perception by dog experts, for example) requires subordinate-level categorization. The prolonged experience of fulfilling different task demands for letters and faces results in two different types of expertise that can later be engaged automatically by the presentation of stimuli from these categories. Recent computational findings are consistent with this characterization of the task demands for letter and face perception. For example, it has been shown that, a network trained for discriminating among letters was not as good as a face-discrimination network in performing a fine-grained discrimination task on blob patterns (each with four blobs forming a Y-shape-like configuration) that differ in small shifts in the blob locations (Zhang & Cottrell, 2004). They concluded that letter perception requires less discriminating power than face perception.

The comparison between expertise with letters and faces has been hard to make before due to the lack of a common framework for different object categories. The framework outlined above allows us to compare empirically letter and face expertise, using the difference in performance for basic- versus subordinate-level categorization as a basis for comparison. In three experiments we measured performance for basic-level categorization with judgement tasks based on letter identity. Judging letter identity is highly common and familiar among readers and provides probably the best example of basic-level categorization of letters. Superordinate-level categorization was assessed with discrimination between letters and digits (Experiments 1 and 2) or between Roman letters and Chinese characters (Experiment 3). For the subordinate level, participants were asked to discriminate between different font types of the same letter. For type designers/typographers and those working in advertising, web design, etc., fine discrimination between different renditions of the same letters is a natural and useful skill.

In Experiments 1 and 2 we examined categorization performance at different levels of abstraction for letters to see if the advantage of the basic level also disappears with expertise for letters, as it does with expertise for faces. In Experiment 3 we compared categorization performance between experts and novices to understand at which level expertise improves categorization performance.

EXPERIMENT 1

We examined letter categorization at different levels of abstraction using a category verification task. This task involved auditory presentation of either a superordinate-level label (e.g., LETTER), a basic-level label (e.g., B), or a subordinate-level label (e.g., Handwritten B), followed by a visual target, which was a letter or a digit (e.g., b). Participants had to decide whether the label matched the target by key pressing. Using a category verification task, it was found that people are faster at categorizing common objects at the basic level than at the other levels (Rosch et al., 1976). In contrast, dog and bird experts categorize the objects of their expert domains at the subordinate level as fast as at the basic level (Tanaka & Taylor, 1991). Similar attenuation of the basic-level advantage was also found for faces (Tanaka, 2001), and can be obtained with novel objects after extensive training (Gauthier et al., 1998). We hypothesized that if letter perception represents a similar type of expertise as that for faces and the other objects just mentioned, then categorization performance should be similar at both basic and subordinate levels for letters. Alternatively, if expert letter perception is constrained by the task demand of a basic-level categorization task and thus represents a different type of expertise, then the rate of verification at the basic level should be faster than at the other levels.

We also manipulate target presentation time, with regard to the finding that shortening stimulus presentation time hinders verification at the subordinate level more than at the basic level for common objects (Jolicoeur, Gluck, & Kosslyn, 1984). It was suggested that additional perceptual processing is required for subordinate-level categorization, and with brief stimulus presentation there is insufficient time for this additional perceptual processing to occur. This selective disruption of subordinate-level categorization was later shown only for common objects but not familiar faces or objects in a domain of expertise (Tanaka, 2001; Tanaka & Taylor, 1991). Here we examined, for letters and digits, which of the superordinate, basic, and subordinate levels would be more affected by reducing stimulus presentation time. If letters are like faces then shortened presentation time (to the same extent as past studies) should not affect judgement at both basic and subordinate levels. Otherwise, if letter perception represents a type of basic-level expertise, which does not lead to improvement at the subordinate level in the way face perception does, then the subordinate level should be the first level affected by using a short stimulus presentation time.

The perception of letters and that of digits have been regarded as partially distinct. In visual search experiments the rate of searching for letters among digits is faster than for letters among other letters (Jonides & Gleitman, 1972). Higher fMRI activity has also been found on or near to the left fusiform gyrus for letters relative to digits (Polk et al., 2002), although digits also tend to recruit letter-selective areas, albeit to a lesser degree (James et al., 2005). It was suggested that letters occur more frequently with each other compared with digits, leading to segregated neural substrates for letters but not digits. Within the proposed object taxonomy framework and consideration of task demands, however, we predicted similar patterns of categorization performance for letters and digits. Perception of digits, like letters, also requires basic-level categorization. This overlap in task demand between letter and object perception should thus lead to the basic-level advantage observed for both categories.

Method

Participant. Fourteen undergraduates at Vanderbilt University participated in the experiment for course credit. All of them had English as their first language, and had normal or corrected-to-normal vision.

Apparatus and stimulus. Four iMac computers (15-inch CRT) each with a monitor set to a 1024 × 768 pixel resolution were used. Presentation of stimuli was controlled by RSVP software (Williams & Tarr, n.d.).

Eight lower-case letters (f, h, I j, k, r, w, y) and eight digits (2, 3, 4, 5, 6, 7, 8, 9) in printed and handwritten forms were used (Figure 2A). A mask was formed by combining fragments from several scrambled letters and digits. The letters, digits, and the mask were about 0.6 cm wide and 0.9 cm high

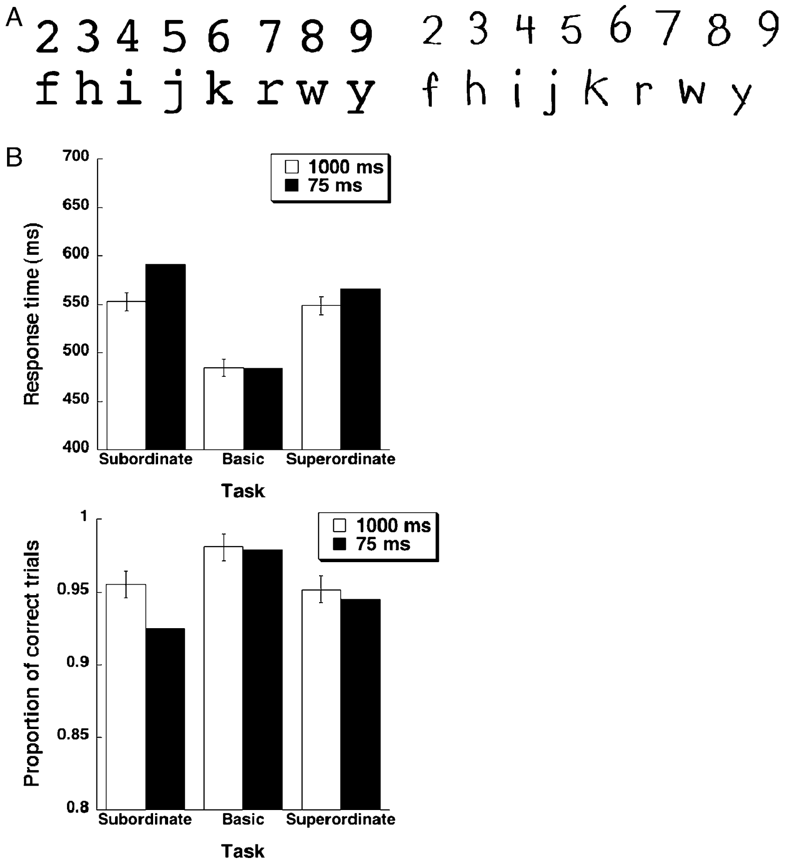

Figure 2. The stimuli used (A), and performance in the "yes" trials (B), in Experiment 1. Error bars represent the standard errors for the Task × Duration interaction contrasts.

each on the screen, subtending a visual angle of 0.28° and 0.43° respectively, with a viewing distance of about 70 cm from the monitor.

Design and procedure. There were three within-subject factors: stimulus type (letters, digits), task (subordinate, basic, superordinate), and stimulus duration (1000 ms, 75 ms).

Each trial began with a 500 ms blank followed by a question presented auditorily (e.g., "Is it a letter?") lasting for about 1 s. After that a 500 ms fixation cross and a 200 ms blank appeared, followed by a target (a letter or a digit) presented for either 1000 ms or 75 ms. At the end of the trial a pattern mask was presented. After the onset of the target the participant had 2 s to respond, and a new trial would begin after the response or when 2 s had passed.

The same letters and digits were used for all verification tasks. Participants had to press the key "1" (for "yes") when the target matched the auditory label and press the key "2" (for "no") otherwise. For the superordinate-level task, the question was either "Is it a LETTER?" or "Is it a DIGIT?". For the basic-level task, the question was "Is it a K?", "Is it a TWO?", etc. In a "no" trial, the label in the question was randomly chosen from the names of the other seven letters (when a letter was presented as target) or digits (when a digit was presented). For the subordinate-level task, the question was about both the name of a letter/digit and its form (e.g., "Is it a HANDWRITTEN K?"). In a "no" trial, the label in the question and the target matched in the identity but not in form for half of the trials (e.g., "Is it a HANDWRITTEN J?" followed by a printed j) and mismatched in identity for the other half (e.g., "Is it a HANDWRITTEN J?" followed by a printed k). The eight letters and eight digits in two forms, the short and long stimulus durations, the three verification tasks, and the two required responses (yes, no), formed a total of 384 trials. The trials were presented in random order over six blocks, and participants could take a rest in between blocks.

To ensure that the participants had enough knowledge about the subordinate-level information of the stimuli, they were allowed to refer at any time to a sheet with all letters and digits printed and labelled in printed and handwritten forms. In addition, before the experiment, they were trained to perform the verification task (with long stimulus presentation) only at the subordinate level until they got 31 out of 32 trials correct in a block (96.8% accuracy).

Results

Out of the correct trials only those with a response time above 200 ms were included in the analysis of response times, resulting in less than 1% of trials discarded. Figure 2B and Table 1 show the performance for all three tasks with short and long stimulus durations. Results for letters and digits were qualitatively the same (there was no main effect or interaction involving stimulus type) and were thus collapsed in the figure and table.

As seen in Figure 2B, performance in the "yes" trials was generally better at the basic level than the other levels. The advantage of the basic level over the subordinate level was enlarged with a shorter stimulus presentation time. Analyses of variances (ANOVAs) with a 2 × 3 × 2 design (Stimulus Type × Task × Duration) confirmed the above observations. Response time data revealed a significant Task × Duration interaction, $F(2, 26) = 4.53$, $p < .05$. Scheffé tests ($p < .05$) showed that shortening stimulus duration impaired performance only for the subordinate-level task but not for the other two tasks. Both response time and accuracy data revealed a main effect of task: RT, $F(2, 26) = 33.78$, $p < .0001$; accuracy: $F(2, 26) = 15.39$, $p < .0001$. Scheffé tests ($p < .05$) showed that performance was better for the basic level than the subordinate or superordinate level.

The "no" trials showed only a main effect of task: RT, $F(2, 26) = 65.37$, $p < .0001$; accuracy: $F(2, 26) = 23.21$, $p < .0001$. Scheffé tests indicated shorter response time at the basic level than both subordinate and superordinate levels, and lower accuracy at the subordinate levels than the other two levels.

Discussion

The results suggest two differences between letters/digits and faces. First, while perception of faces and other objects of expertise often involve a disappearance of the basic-level advantage, such advantage exists in experts for letters and digits. Even with a long presentation time, categorization at the subordinate level was worse than that at the basic level, similar to what was found for objects in prior work (Jolicoeur et al., 1984). This reflects the extra perceptual processing required to obtain information for

TABLE 1
Performance in "no" trials at long and short durations in Experiment 1

	Long		*Short*	
Task	*RT*	*Accuracy*	*RT*	*Accuracy*
Subordinate	585	0.910	588	0.914
Basic	483	0.983	499	0.989
Superordinate	593	0.956	598	0.961

The standard errors for the Duration × Task interaction were 6.413 ms and 0.009 for RT and proportion of correct trials, respectively.

subordinate-level judgement. Second, a shorter presentation time selectively disrupted performance at the subordinate level, whereas performance at the basic level remained intact. This is consistent with Jolicoeur et al.'s (1984) finding with common objects, but differs from Tanaka's finding of intact subordinate-level performance at short presentation times for faces and objects of expertise (Tanaka, 2001; Tanaka & Taylor, 1991). These two differences suggest that letter perception may not be regarded as the same type of expertise as face perception.

Verification at the superordinate level was slower than at the basic level to a similar extent with both long and short presentation times. The additional time required for superordinate-level verification suggests that participants may have recognized the stimulus at the basic level first (e.g., "G"), before accessing the superordinate level it belongs to ("LETTER"). Presentation time had little effect on superordinate-level verification, presumably because no extra visual processing is required at the superordinate level relative to the basic level.

EXPERIMENT 2

In Experiment 1 performance for the basic level was found to be superior to those for the superordinate and subordinate levels. One issue, however, concerns the choice of fonts for the subordinate-level judgement. The handwritten and printed versions differed drastically in terms of line thickness, or typographic weight. Participants could have just focused on this salient differences and ignored all other shape features, rendering the task trivial (like a task of discriminating between the same letter in different ink colours).

To address this issue, and to generalize the results to other fonts, we conducted the same experiment using a different pair of fonts (Snell Roundhand and Arial Italics). These two fonts, though dissimilar from each other (see in Figure 3), do not differ drastically in just a single salient feature. We chose the italic version of Arial so that the slant difference between the two fonts was not very salient. It should also be noted that a label-image matching task was used in both Experiments 1 and 2, rather than simultaneous matching of two letters, where one might be encouraged to compare two letters presented side by side and extract differences in a single salient feature.

Method

Participant. Seventeen undergraduates at Vanderbilt University participated in the experiment for course credit. All of them had English as their first language, and had normal or corrected-to-normal vision.

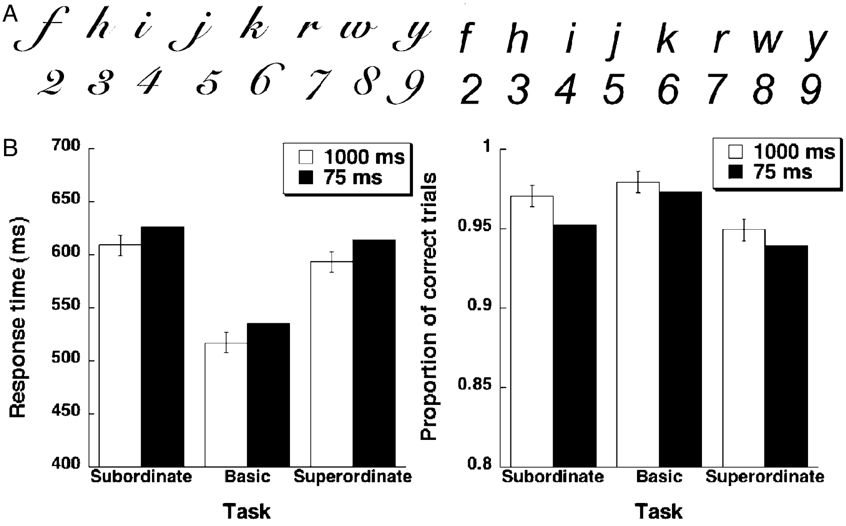

Figure 3. The stimuli used (A), and performance in the "yes" trials (B), in Experiment 2. Error bars represent the standard errors for the Task × Duration interaction contrasts.

Apparatus and stimulus. The same apparatus and stimuli as Experiment 1 were used, except that the Snell Roundhand and Arial fonts were used instead of the printed and handwritten forms (Figure 3A).

Design and procedure. The design and procedure were the same as those in Experiment 1.

Results

Out of the correct trials only those with a response time above 200 ms were included in the analysis of response times, resulting in less than 1% of trials discarded. Figure 3B and Table 2 show the performance for all three tasks with short and long stimulus durations. Results for letters and digits were qualitatively the same (except for minor differences in "no" trials), and were thus collapsed when presented in the figure and the table.

As seen in Figure 3B, performance in the "yes" trials was generally better at the basic level than the other levels. Analyses of variances (ANOVAs) with a 2 × 3 × 2 design (Stimulus Type × Task × Duration) confirmed the above observations. Both response time and accuracy data revealed a main effect of task: RT, $F(2, 32) = 25.85$, $p < .0001$; accuracy: $F(2, 32) = 12.11$, $p < .0001$. Scheffé tests ($p < .05$) showed that response time was better for the basic than subordinate level, better for the basic than superordinate level, and not different between the subordinate and superordinate levels. Accuracy was

TABLE 2
Performance in "no" trials at long and short durations in Experiment 2

	Long		*Short*	
Task	*RT*	*Accuracy*	*RT*	*Accuracy*
Subordinate	675	0.926	672	0.905
Basic	545	0.985	541	0.982
Superordinate	640	0.947	639	0.950

The standard errors for the Duration × Task interaction were 17.4 ms and 0.011 for RT and proportion of correct trials, respectively.

better for the basic than superordinate level, better for subordinate than superordinate level, and not different between the basic and the subordinate levels. Also, there was a main effect of duration for response time, $F(1, 16) = 7.78$, $p < .05$, with faster responses for long than short duration. There was no other significant main effect or interaction.

The "no" trials showed a main effect of task: RT, $F(2, 32) = 29.96$, $p < .0001$; accuracy: $F(2, 32) = 21.34$, $p < .0001$. Scheffé tests indicated faster and more accurate responses at the basic level than at both subordinate and superordinate levels. Accuracy was also higher for the superordinate than subordinate level. Response time showed a main effect of stimulus type, $F(2, 32) = 4.63$, $p < .05$, with faster responses for digits (625 ms) than letters (612 ms). Accuracy also showed a significant interaction between stimulus type and task, $F(2, 32) = 3.51$, $p < .05$. Scheffé tests showed that for digits, the basic-level trials were more accurate than the superordinate-level trials, which were more accurate than the subordinate-level trials. For letters, accuracy was lower for the subordinate level than for both basic and superordinate levels, which were not different from each other.

Discussion

Similar to Experiment 1, the basic-level advantage was shown for letters and digits compared with the superordinate and subordinate levels, for both long and short presentation times. One difference is that shortening presentation time disrupted performance only for the subordinate level in Experiment 1, but for all levels in this experiment. A speculation is that the fonts used in this experiment made the tasks harder, as indicated by the higher response times in general compared with Experiment 1 (although it should be noted that different participants were tested in the two experiments). Shortening presentation time to 75 ms in this experiment may thus be sufficient to impair performance at all levels. With slightly longer presentation times (e.g., 150 ms) we may be able to observe disruption for the subordinate level only. In any case, the critical finding of the basic-level advantage over the subordinate level was replicated.

EXPERIMENT 3

In Experiments 1 and 2 we did not find for letters the same attenuation of the basic-over-subordinate-level advantage as shown in other cases of face-like expertise (e.g., faces, dogs, birds, or Greebles). Although we suggest that this is caused by the extensive experience in recognizing letters specifically at the basic level (and not at other levels), there are other possible accounts for this result. Perhaps expertise with letters causes a decrease in the basic-level advantage, just not as much as face expertise does. Or perhaps letter expertise improves categorization of letters in general (across all levels), and therefore the basic-level advantage remains.

In Experiment 3, we attempted to address more directly the effect of expertise on letter categorization. Tanaka and Taylor (1991) showed that the basic-level advantage for birds and dogs shown in novices is reduced in experts. This is consistent with the idea that subordinate-level expertise results in a reduction of the basic-level advantage. However, a similar framework led us to regard letter perception as a type of basic-level expertise. With the development of this type of expertise, improvement should occur for categorization at the basic level but not for the subordinate level. Thus, it is the advantage of the basic level over the subordinate level that should increase in letter experts.

To test if letter perception demonstrates an opposite behavioural marker compared with face perception and bird and dog perception for experts, an expertise manipulation was introduced, using a comparison between Chinese-English bilinguals and non-Chinese readers (who were fluent in English but had no Chinese-learning experience). On each trial, participants were shown two characters (two Roman, two Chinese, or one of each) simultaneously and asked to judge whether the two characters were in the same category, in terms of (a) the subordinate level (e.g., whether they were both handwritten "H"s); (b) the basic level (e.g., whether they were both "B"); or (c) the superordinate level (e.g., whether they were both Roman or both Chinese characters). Since Chinese-English bilinguals have greater experience in perceiving Chinese characters, they should show faster verification of Chinese characters at the basic level compared with the non-Chinese readers. Nevertheless, the two groups were both experts with Roman letters and thus should not display any difference in verification speed at the basic level for Roman letters. In other words, the bilinguals should demonstrate a greater basic-level advantage for Chinese characters (but not Roman letters) compared with non-Chinese readers. Another prediction was that, when viewing the Chinese characters, the bilinguals should be better than the non-Chinese readers only at the basic level but not at the other levels. In other words, expertise should have a selective effect at the basic level only.

The use of Chinese characters not only allowed manipulation of expertise level, but should also help generalize our results to characters of other expert writing systems. Neural studies have identified temporal and spatial overlap of neural selectivity for Roman letters and Chinese characters (Wong, Gauthier, Woroch, Debuse, & Curran, 2005; Wong et al., 2006), suggesting overlap of processes involved in the perception of the two character types. Another note about the design is that a simultaneous matching task was used instead of a category verification task, such that the English readers could also perform the task on Chinese characters with minimum difficulty even though they did not have labels for the characters.

Method

Participant. Twelve Chinese-English bilinguals at the Chinese University of Hong Kong and twelve non-Chinese readers at Vanderbilt University participated in the experiment for monetary rewards. The Chinese-English bilinguals had Chinese as their first language and all had over 15 years of experience in reading both Chinese and English texts. The non-Chinese readers did not know or learn any Chinese before the experiment. All of them had normal or corrected-to-normal vision.

Apparatus and stimulus. Six iMac computers (15-inch CRT) each with monitors set to a 1024 × 768-pixel resolution were used. Presentation of stimuli was controlled by RSVP software (Williams & Tarr, n.d.).

Twenty lower-case Roman letters and twenty Chinese characters were used (see Figure 4A). Each character could appear in one of two font types (printed, handwritten) and one of three different sizes: large, medium, or small (Roman letters: 1.3 × 1 cm, 1.1 × 0.85 cm, 0.9 × 0.7 cm for letters with ascendants or descendants, 1 × 1 cm, 0.85 × 0.85 cm, and 0.7 × 0.7 cm for other letters; Chinese characters: 1.2 × 1.2 cm, 1 × 1 cm, 0.8 × 0.8 cm). For a viewing distance of 70 cm the letters subtended a vertical visual angle from 0.73° to 1.06°.

Design and procedure. The design included one between-subject factor [Group (Chinese-English bilinguals, non-Chinese readers)] and two within-subject factors [character type (Roman, Chinese) and task (subordinate, basic, superordinate)]. Participants had to make a "yes/no" judgement concerning whether two simultaneously presented letters belonged to the

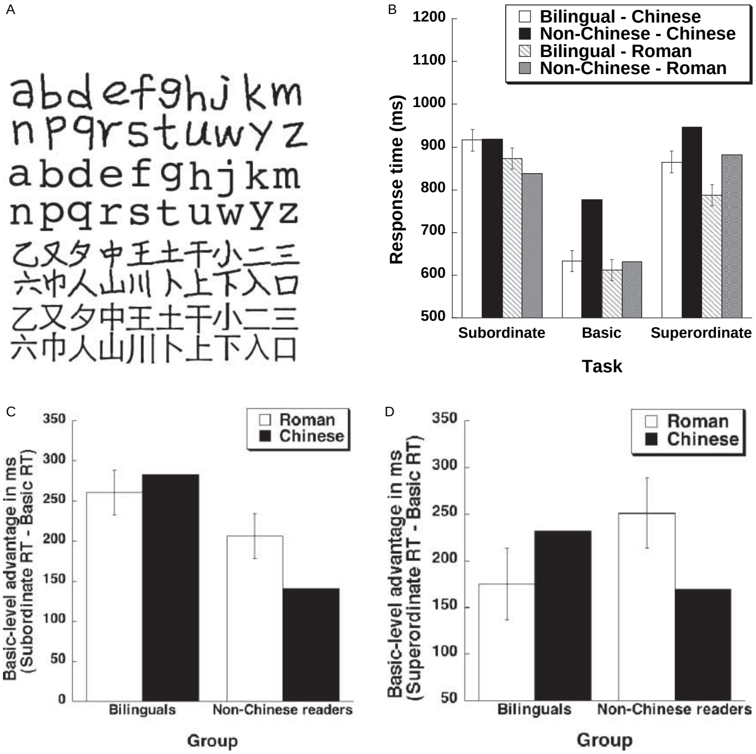

Figure 4. The stimuli used (A), response times (B), and the basic-level advantage compared with the subordinate level (C) and with the superordinate level (D) in Experiment 3. Error bars in (B) represent the standard errors for the Group × Stimulus Type × Task interaction contrasts. Error bars in (C) and (D) represent the standard errors for the Group × Stimulus Type interaction contrasts.

same subordinate-level category (e.g., both are the handwritten "m"s), to the same basic-level category (e.g., both are "b"s), or to the same superordinate category (e.g., both are Chinese letters, or both are Roman letters). During each trial, a fixation cross appeared at the middle of the screen for 500 ms followed by two letters each presented on either side of the fixation (the centre of the letter was 3 cm away from the fixation). The maximum time allowed for making a response was 3 s. The words of "same letter same font", "same letter", or "same category" were shown on the top left-hand corner to remind the subject of the task for a particular block of trials. The two letters presented were never matched in terms of size (e.g., a large-sized letter was paired with a middle-sized or small-sized letter, but not another large-sized letter).

Each task consisted of 160 trials. In all tasks half of the trials involved Chinese characters and half Roman letters. Half were "yes" trials and half were “no” trials. For the subordinate-level task, there were 40 “yes” trials with handwritten characters and 40 “yes” trials with printed characters. The “no’ trials were created by taking the “yes” trials and replacing them either with the same character of the other font (for 40 trials) or with another randomly chosen character of the other font (40 trials). For the basic-level task, half of the “yes” trials and half of the “no” trials involved characters of the same font. The “no” trials were formed by duplicating the “same” trials and replacing one of the characters with one randomly chosen from the same set (Roman or Chinese). For the superordinate-level task, the “yes” trials involved different characters of the same set, while the “no” trials involved one character from the Roman set and one from the Chinese set. The positions (left, right) of the different character types were counterbalanced across trials. There were three blocks for each task, and they alternated with each other with an order counterbalanced across participants.

Results

Only “yes” trials were analysed, since there was no meaningful distinction in the “no” trials between the superordinate-level conditions for Roman and Chinese characters. Out of the correct trials, only those with a response time above 200 ms were included in the analysis of response times, resulting in less than 1% of trials discarded.

Figure 4B and Table 3 show the performance of the two groups in different conditions. Although the bilinguals and non-Chinese readers performed at similar levels for the basic-level task with Roman letters, the bilinguals were faster than the non-Chinese readers for the basic-level task with Chinese characters. This group advantage was specific to the basic-level task with the Chinese characters. The two groups did not differ for the subordinate-level tasks with both character types. In addition, the bilinguals’ advantage over the non-Chinese readers for the superordinate-level task was not any greater with Chinese than Roman characters.

TABLE 3
Proportion of correct “yes” trials in Experiment 3

	Chinese characters		*Roman letters*	
	Bilingual	*Non-Chinese*	*Bilingual*	*Non-Chinese*
Subordinate	0.927	0.958	0.958	0.968
Basic	0.971	0.960	0.990	0.983
Superordinate	0.906	0.871	0.941	0.959

The standard error for the three-way interaction was 0.017.

A 2 × 2 × 3 (Group × Character Type × Task) ANOVA on response times confirmed these observations. The three-way interaction was significant, $F(2,\ 44) = 4.04$, $p < .05$. Scheffé tests ($p < .05$) revealed that bilinguals responded faster than non-Chinese readers for the basic-level task with Chinese characters but not with Roman letters. For the subordinate-level task, both groups performed similarly for both Chinese and Roman characters. For the superordinate-level task, the bilinguals were faster than the non-Chinese readers with both Chinese and Roman characters. An interaction contrast showed no difference in the bilingual's advantage over the non-Chinese readers in the superordinate-level task across the two types of characters ($t < 1$).

Figure 4C shows the amount of the basic-level advantage (as a response time difference between the basic and subordinate levels) for the two groups viewing the two character types. An ANOVA showed a significant interaction between group and character type, $F(1,\ 22) = 5.23$, $p < .05$. Scheffé tests ($p < .05$) revealed that whereas bilinguals experienced a basic-level advantage for Roman and Chinese characters of a similar magnitude, non-Chinese readers had a greater basic-level advantage for Roman than Chinese characters.

Figure 4D shows another measure of the basic-level advantage (as a response time difference between the basic and superordinate levels). Similarly, an ANOVA showed a significant Group × Stimulus Type interaction, $F(1, 22) = 6.59, p < .05$. Scheffé tests ($p < .05$) revealed that while the basic-level advantage was not significantly different between Roman and Chinese characters for the bilinguals, the non-Chinese readers had a greater basic-level advantage for Roman than Chinese characters.

The accuracy data (Table 3) showed no main effect or interaction involving Group. There was a significant Character Type × Task interaction, $F(2, 44) = 4.08$ $p < .05$. Scheffé tests ($p < .05$) indicated that for both Roman and Chinese characters the basic level was more accurate than the superordinate level and no different from the subordinate level. However, the subordinate level was better than the superordinate level only for Chinese characters but not Roman letters.

Discussion

The increase in the basic-level advantage found in conditions of letter expertise clearly dissociates it from face-like expertise, which is typically characterized with an attenuation of the basic-level advantage (Tanaka, 2001; Tanaka & Taylor, 1991). Also, the expertise effect was specific for the basic level. This suggests that the results of Experiments 1 and 2 were not likely to be due to letter expertise merely improving letter categorization at all levels. The basic-level advantage for the Roman and Chinese characters revealed by the bilinguals was statistically indistinguishable. Despite the visual and linguistic differences between the two writing systems, the two types of characters led to similar results.

As discussed before, superordinate-level categorization is sometimes thought to follow categorization at the basic level. One might then expect that an advantage in basic-level performance for Chinese characters in bilinguals should also be seen at the superordinate level. However, the group difference in performance was about the same for Roman and Chinese characters. Although this is a separate issue from the basic- versus subordinate-level comparison, it may be worth future investigation.

GENERAL DISCUSSION

Thinking of face and letter recognition in the context of a common framework (where we can anchor them both in terms of the structural basic level) helped us to ask whether perceptual expertise with letters leads to similar changes in categorization performance than have previously been described for face expertise. According to this analysis, letter and face perception represent two different types of expertise, with the former leading to improvements mainly at the basic level and the latter leading to improvements mainly in subordinate-level categorization.

If expertise for faces and letters differs in a fundamental way, one might expect other differences to exist. Specifically, the strategies adopted by experts in the two domains may differ, and the neural substrates that may become specialized to support these strategies should also differ. For face perception (and similar types of object expertise) at the subordinate level, expertise has been found to rely on mechanisms that are "holistic", in the sense that experts tend to process all the parts of a face, even when instructed to selectively attend to only one part (Bruce & Young, 1986; Gauthier & Tarr, 2002). In addition, they appear to rely on configural or relational information (e.g., the distance between parts) more than they generally do for letters and other objects (Ge, Wang, McCleery, & Lee, 2006; Tanaka & Sengco, 1997). In the case of letter perception, we have argued that the goal is basic-level categorization. However, letter perception appears to differ in at least one way from the recognition of most objects, even though in both cases the goal is to access the basic-level identity (e.g., chair vs table; "h" vs "m"). The difference comes not from the recognition goal, but rather lies in the regularities in subordinate-level information that are typically available for letters but not for objects. We generally recognize letters in rapid succession, and these letters are presented at a common subordinate-level, that is, the same size and font. Such regularities in font can facilitate the recognition of letters at the basic level (Sanocki, 1987, 1988, 1992) in a similar way as voice regularity can facilitate the recognition of speech (Dupoux & Green, 1997; Mullennix, Pisoni, & Martin, 1989). In recent work, this "font-tuning" effect was shown to be associated with expert but not novice perception of Roman letters and Chinese characters (Gauthier, Wong, Hayward, & Cheung, 2006). Therefore, while expertise in discriminating objects that differ in metric properties appears to recruit holistic and configural strategies, expertise in recognizing sequences of objects at the basic level can cause one to rely on regularities in style, if such regularities are present. In support of these differences in processing, the neural substrates mediating face and letter expertise have been found to recruit different sets of neural areas within the ventral occipitotemporal system, as discussed earlier.

Theoretical attempts to directly compare face, object, and letter processing are rare in the literature. One exception is the model suggested by Farah (2004) to account for the patterns of co-occurrence of different types of associative agnosia for faces, objects, or words. Ideally, the same theory should be able to account for category specialization in the normal brain, and the patterns of category-selective impairments following brain damage. Farah's model attempts to explain all three disorders using a single continuum describing how object parts are processed. The ability to represent complex objects without decomposing them into parts is thought to underlie face perception, and, if this ability is impaired, then face agnosia (prosopagnosia) is observed. The ability to rapidly encode multiple parts is argued to support word perception, so its impairment would lead to word agnosia (pure alexia). The recognition of different objects may rely more or less on either of these abilities, so if either ability is severely affected, then some degree of object agnosia will also be observed. Farah's two-factor model led to specific predictions about co-occurrence among types of agnosia, the most important being that there could never be a case of object agnosia without either prosopagnosia or pure alexia. This elegant model was, at the time it was proposed, supported by a review of 99 case reports published between 1966 and 1990. However, since then, examples of agnosia without alexia or prosopagnosia have been described (Humphreys & Rumiati, 1998).

A difference between Farah's (2004) model and the account proposed here is that whereas Farah's model has been applied to perception of words and letter sequences in general, our focus is on individual letter perception. We chose to compare single letters rather than words to faces for one reason: The neuroimaging literature has so far failed to provide good evidence for a visual area that is strictly specialized for words. The visual word form area in the left fusiform gyrus responds more to words and pseudowords than to consonant strings (Cohen & Dehaene, 2004; Cohen et al., 2000; Cohen et al., 2002); however, the same area also responds to several nonvisual contrasts (Price & Devlin, 2003, 2004), and it is not more active in response to consonant strings than digits strings or strings of unfamiliar Chinese characters (James et al., 2005). In contrast, areas of the left ventral cortex that are selective for single letters show a stronger response for them than for single digits or unfamiliar Chinese characters (James et al., 2005), and also respond more to letters than to objects and faces (Wong et al., 2006). While expertise for print likely exists at several levels, visual effects at the level of single characters are easier to study independently of phonological and semantic influences.

In contrast to Farah's (2004) model, the account proposed here dissociates both face and letter expertise from common object recognition: face and object recognition differ in terms of the target level of categorization (subordinate vs basic), whereas object and letter perception differ (at least) in terms of the presence of regularities in subordinate-level information (shared style) for items that are processed in rapid succession. Another difference with Farah's model is that our account does not expect all object categories to fall along a single continuum (e.g., from holistic to part-based processing). Rather, any new case of expertise (e.g., sight-reading, x-ray diagnosis, handwriting expertise) should be analysed in terms of the recognition goal of expert observers as well as the stimulus properties available to support this goal. Only then can predictions be made of whether a given skill should be supported by the same mechanisms/neural substrates as face perception or letter perception, or rather would be expected to rely on different strategies and specialize yet another part of the visual system.

REFERENCES

Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., et al. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology*, *71*(2), 821–825.

Beauregard, M., Chertkow, H., Bub, D., Murtha, S., Dixon, R., & Evans, A. (1997). The neural substrate for concrete, abstract, and emotional word lexica: A positron emission tomography study. *Journal of Cognitive Neuroscience*, *9*, 441–461.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). Erp manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260.

Besner, D., Coltheart, M., & Davelaar, E. (1984). Basic processes in reading: Computation of abstract letter identities. *Canadian Journal of Psychology*, *38*(1), 126–134.

Bigsby, P. (1988). The visual processor module and normal adult readers. *British Journal of Psychology*, *79*, 455–469.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*(4), 159–166.

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition*, *1*(2/3), 253–274.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage*, *22*, 466–476.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291–307.

Cohen, L., Lehericy, S., Cohochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, *125*, 1054–1069.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, *15*, 245–286.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341.

Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 914–927.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *33A*, 325–350.

Farah, M. J. (2004). *Visual agnosia* (2nd ed.). Cambridge, MA: MIT Press.

Flowers, D. L., Jones, K., Noble, K., VanMeter, J., Zeffiro, T. A., Wood, F. B., et al. (2004). Attention to single letters activates left extrastriate cortex. *Neuroimage*, *21*(3), 829–839.

Garrett, A. S., Flowers, D. L., Absher, J. R., Fahey, F. H., Gage, H. D., Keyes, J. W., et al. (2000). Cortical activity related to accuracy of letter recognition. *Neuroimage*, *11*(2), 111–123.

Gauthier, I. (2000). What constrains the organization of the ventral temporal cortex? *Trends in Cognitive Science*, *4*(1), 1–2.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., & Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Current Biology*, *7*(9), 645–651.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, *37*(12), 1673–1682.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431–446.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), 568–573.

Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*(15/16), 2401–2428.

Gauthier, I., Wong, A. C.-N., Hayward, W. G., & Cheung, O. S.-C. (2006). Font-tuning associated with expertise in letter perception. *Perception*, *35*(4), 541–559.

Ge, L., Wang, Z., McCleery, J. P., & Lee, K. (2006). Activation of face expertise and the inversion effect. *Psychological Science*, *17*, 12–16.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, M. (2002). Eccentricity bias as an organizing principle for human high order object areas. *Neuron*, *34*, 479–490.

Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K., Weiller, C., et al. (1992). The cortical localization of the lexicons: Positron emission tomography evidence. *Brain*, *115*(6), 1769–1782.

Humphreys, G. W., & Rumiati, R. I. (1998). Agnosia without prosopagnosia or alexia: Evidence for stored visual memories specific to objects. *Cognitive Neuropsychology*, *15*(3), 243–277.

James, K. H., James, T. W., Jobard, G., Wong, A. C.-N., & Gauthier, I. (2005). Letter processing in the visual system: Different activation patterns for single letters and strings. *Cognitive, Affective, and Behavioral Neuroscience*, *5*(4), 452–466.

Johnson, N. F., & Pugh, K. R. (1994). A cohort model of visual word recognition. *Cognitive Psychology*, *26*, 240–346.

Jolicoeur, P., Gluck, M., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*, 243–275.

Jonides, J., & Gleitman, H. (1972). A conceptual category in visual search: O as letter or as digit. *Perception and Psychophysics*, *12*, 457–460.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*, 115–121.

Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage*, *19*(4), 1492–1500.

McClelland, J. L. (1976). Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 80–91.

McClelland, J. L., & Rumelhart, D. L. (1981). An interactive activation model of context effects in letter perception. Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*, 365–377.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Oden, G. C. (1979). A fuzzy logical model of letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 336–352.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, *112*(1), 43–59.

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, *249*(4972), 1041–1044.

Polk, T. A., & Farah, M. J. (1998). The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling, and behavioral studies. *Proceedings of the National Academy of Sciences of the USA*, *95*(3), 847–852.

Polk, T. A., Stallcup, M., Aguirre, G. K., Alsop, D. C., D'Esposito, M., Detre, J. A., et al. (2002). Neural specialization for letter recognition. *Journal of Cognitive Neuroscience*, *14*, 145–159.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, *78*, 391–408.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology: Monographs*, *79*(1, Pt. 2).

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review*, *74*(5), 392–409.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, *19*, 473–481.

Price, C. J., & Devlin, J. T. (2004). Reply to letter to the editor. *NeuroImage*, *22*, 477–479.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, *16*(16), 5205–5215.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain*, *119*(4), 1221–1238.

Reicher, G. M. (1969). Perceptual recognition as a function of the meaningfulness of the stimulus material. *Journal of Experimental Psychology*, *81*, 275–280.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Sanocki, T. (1987). Visual knowledge underlying letter perception: Font-specific, schematic tuning. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 267–278.

Sanocki, T. (1988). Font regularity constraints on the process of letter recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 472–480.

Sanocki, T. (1992). Effects of font- and letter-specific experience on the perceptual processing of letters. *American Journal of Psychology*, *105*(3), 435–458.

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition*, *67*(1–2), 147–179.

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General*, *130*(3), 534–543.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*(5), 583–592.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457–482.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, *122*(11), 2119–2132.

Williams, P., & Tarr, M. J. (No date). RSVP: Experimental control software for MacOS [Online].

Wong, A. C.-N., Gauthier, I., Woroch, B., Debuse, C., & Curran, T. (2005). An early electrophysiological response associated with expertise in letter perception. *Cognitive, Affective, and Behavioral Neuroscience*, *5*(3), 306–318.

Wong, A. C.-N., Jobard, G., James, K. H., James, T. W., & Gauthier, I. (2006). Expertise with characters in alphabetic and non-alphabetic writing systems engage the same occipito-temporal area. *Manuscript under review*.

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*(8), 1234–1242.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145.

Zhang, L., & Cottrell, G. W. (2004). *Seeing blobs as faces or letters: Modeling effects on discrimination*. Paper presented at the 2004 international conference on Development and Learning, La Jalla, CA.

Manuscript received November 2005
Manuscript accepted July 2006
First published online December 2006