Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 45 (2005) 431–448

Vision Research

www.elsevier.com/locate/visres

Behavioral and electrophysiological evidence for configural processing in fingerprint experts

Thomas A. Busey [a,*], John R. Vanderkolk [b]

[a] *Department of Psychology, Indiana University, Bloomington, 1101 E. 10th Street, Bloomington, IN 47405, United States*
[b] *Indiana State Police Laboratory, Fort Wayne, IN 46804, United States*

Received 14 May 2004; received in revised form 20 August 2004

Abstract

Visual expertise in fingerprint examiners was addressed in one behavioral and one electrophysiological experiment. In an X-AB matching task with fingerprint fragments, experts demonstrated better overall performance, immunity to longer delays, and evidence of configural processing when fragments were presented in noise. Novices were affected by longer delays and showed no evidence of configural processing. In Experiment 2, upright and inverted faces and fingerprints were shown to experts and novices. The N170 EEG component was reliably delayed over the right parietal/temporal regions when faces were inverted, replicating an effect that in the literature has been interpreted as a signature of configural processing. The inverted fingerprints showed a similar delay of the N170 over the right parietal/temporal region, but only in experts, providing converging evidence for configural processing when experts view fingerprints. Together the results of both experiments point to the role configural processing in the development of visual expertise, possibly supported by idiosyncratic relational information among fingerprint features.

Keywords: Fingerprints; Faces; Configural; N170; Expertise; Inversion

1. Introduction

The training and exposure that fingerprint examiners undergo as part of their profession represents an extreme case of perceptual learning. These experts receive extensive training in the fingerprint identification process with competency testing under an accomplished professional. In addition, the penalty for incorrect identifications is quite high: lives or careers could be ruined and labs shut down because of inappropriate accusations or exonerations. As a result, fingerprint examiners take their jobs very seriously and spend a great deal of time studying prints. This situation produces an intensive study of a stimulus set that may lead to profound changes to the perceptual systems of fingerprint examiners. Given this pool of expertise, it is somewhat surprising to find that very few if any empirical studies have addressed how long-term exposure to fingerprints might alter the perceptual processing of latent and inked prints by examiners. The goal of this article is to characterize the differences between fingerprint experts and novices, and address the nature of the strategies and visual skills that experts may have developed during training. The results not only bear on the nature of skill development with examiners, but help constrain models of perceptual learning as well, in particular the role and nature of configural processing in visual expertise.

While relatively little work has been done with fingerprint examiners, we draw upon several related studies of expertise that have identified behavioral and neural correlates of expertise (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002; Shiffrin & Lightfoot, 1997; Tanaka & Curran, 2001), since fingerprints share some

* Corresponding author. Tel.: +1 812 334 2940.
E-mail address: busey@indiana.edu (T.A. Busey).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2004.08.021

characteristics with faces and other stimuli that exhibit perceptual learning. Goldstone (1998) identified four general mechanisms that might support the development of perceptual expertise. For stimuli that can be represented along different psychological dimensions, attention weighting allows more emphasis to be placed on relevant dimensions, and differentiation allows increased separation between objects in psychological space. In addition to these manipulations of dimensional representations, new features can be created, either through imprinting, which creates new receptors specific to the to-be-learned features (Schyns & Murphy, 1994; Schyns & Rodet, 1997), or unitization, which creates complex configurations out of single features (Shiffrin & Lightfoot, 1997). For more naturalistic stimuli without clear psychological dimensions, much of the emphasis of expertise research has addressed the role of relational information and context-related effect in which the perception of one feature is influenced by the presence or absence of other features. Both of the mechanisms can be subsumed under the general category of configural processing. Configural effects have long been studied in faces (Yin, 1969), and more recently these effects have been extended to other types of objects. Perhaps the most comprehensive look at training effects with novel stimuli is work with Greeble stimuli by Gauthier and Tarr (1997) and Gauthier, Williams, Tarr, and Tanaka (1998), who described configural benefits for single features when surrounded by the appropriate context, but only after training and only for upright stimuli. Later work has suggested that this form of configural processing is supported by the gradual development of relational information between features throughout the course of learning (Gauthier & Tarr, 2002).

The neural basis for expertise has been addressed in imaging experiments (Gauthier et al., 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Tarr & Gauthier, 2000), electrophysiological studies (Gauthier, Curran, Curby, & Collins, 2003; Rossion, Gauthier, et al., 2002; Tanaka & Curran, 2001) and single-cell recording (Baker, Behrmann, & Olson, 2002; Logothetis, 2000). It appears that brain regions that initially are highly responsive to complex visual objects such as faces are also activated by learned stimuli after training, suggesting a recruitment of face-responsive areas to support expertise for other complex objects (although see Carmel & Bentin (2002) for a defense of a modular account of face processing). At the level of single cells, configural processing seems to occur via increasing specialization of responses to conjunction stimuli, rather than increased firing rates (Baker et al., 2002).

Fingerprint matching shares some similarity with a radiological screening process, and several articles have documented expertise effects with radiologists. Sowden, Davies, and Roling (2000) found that experts could better detect low-contrast dots embedded in simulated X-rays, and Myles-Worsley, Johnston, and Simons (1988) reported that experts had better memory performance for abnormal X-rays while exhibiting worse performance for normal X-rays.

Fingerprint examinations are somewhat unique as a task. Unlike tumor detection, which is essentially a categorization task, latent fingerprints are compared with a very specific candidate sample. While this task shares some of the characteristics of an identification process, both samples are present simultaneously. In addition, fingerprints share a very small set of features, some of which, such as ridge endings and bifurcations, are distributed in fairly random locations from one print to another. This makes relational information important. However, unlike faces, the feature locations are much less constrained on a fingerprint, and relatively little work has been done with analogous stimuli in the literature. Thus it remains to be seen whether configural processes can develop for fingerprints. If so, this will suggest the conditions under which configural processing can develop.

Given that relatively little literature exists on fingerprint examiners, our first aim is to identify whether experts do indeed differ from novices on tasks related to fingerprint examinations, and then determine whether performance differences might be tied to the mechanisms that have been identified that support perceptual learning. The results of our first experiment will point to the suggestion of configural processing in experts, and we follow this up with a second experiment designed to look for neurophysiological evidence of configural processing.

2. Experiment 1

Although some elements of initial triage and screening might be handled via a computer, virtually all evidence presented in court is based on a visual match made by an examiner. Fingerprints contain characteristic features such as general ridge paths of loop, whorl, or arch, as well as idiosyncratic features of specific ridge paths with ridge endings or bifurcations, and texture and pore positions on ridges. This provides a very consistent visual diet for examiners, which may enable their visual system to adopt strategies that enhance information acquisition from one fingerprint. The training may also enhance maintenance of visual information during an eyemovement, and thus Experiment 1 includes an element of visual working memory.

Fingerprints are somewhat like faces in that they have certain features that tend to occur in similar locations across exemplars, and thus may exhibit properties in experts similar to those seen with faces, most notably configural processing and superior subordinate-level categorization performance (e.g. Tanaka, 2001). Thus some of our stimulus manipulations follow from the face processing literature. However, latent fingerprints recovered from crime scenes tend to be moderately to severely degraded, may represent only part of the fingerprint, and are contaminated by visual noise deriving from dust, surface texture, pressure, and many other sources of variability. As one fingerprint expert described it, their job is to 'see through the noise' in order to pick out particular features that enable a match. Thus we also included manipulations designed to capture elements of expertise that have evolved to work under these conditions.

The typical fingerprint matching process involves a latent print placed next to a candidate inked print taken from another source. An expert examines the two prints, either enlarged on a screen or through a magnifying lens, to end up at one of three conclusions: (1) there is sufficient detail to reject the two prints as coming from the same source, (2) there is sufficient detail to conclude that the two prints come from the same source, and (3) there is insufficient detail to make a determination (usually due to a poor quality print). During the examination process, the expert must make eyemovements between the latent and inked prints in an attempt to visually match features. The matched features take on up to three levels of detail, which include Level 1 detail, which is area that is visibly just the general direction of ridge flow of a fingerprint; Level 2 detail, which is clear enough to specific individual ridge paths with ridge endings and bifurcations; and Level 3 detail, which is area clear enough to reveal the texture and pore position detail within a ridge. After an open-ended examination that can take minutes to hours, the expert makes one of the three conclusions described above.

To assess the possible mechanisms for expertise, we designed an experiment that abstracted what we believed to be some essential skills of the fingerprint matching process, but does not require lengthy examinations. We settled on an X-AB matching task, in which an observer is presented with a section of a fingerprint for study, and is then tested with two prints in a forced-choice test. The two test prints were sometimes degraded with visual noise or partially masked to simulate some of the characteristics of latent prints. Noise has the property that it makes local information variable, and McKone, Martini, and Nakayama (2001, 2003) argued that added noise was one method to isolate configural processing in faces. To assess the visual memories of experts and novices, a visual mask was inserted between the study print and the test prints, which remained visible for either 200 ms or 5200 ms. In addition, because our experts are scattered across the country and in Europe, we designed our experiment as a Java applet that runs in any browser and collected data over the web.

2.1. Method

2.1.1. Participants

Eleven fingerprint experts were recruited by the second author to participate. These experts were all active and had completed training required to practice in the field. Eleven novices were also recruited to participate, who included students at Indiana University, as well as older participants from the Bloomington, Indiana community. Care was taken to recruit observers who were equally motivated in each group. None of the participants received monetary compensation, but agreed to participate out of interest in the topic and a desire to assist law enforcement officials, however indirectly. The novices represented a somewhat younger group of observers, with many in their early 20 s although one was in her early 30 s and two were in their late 50 s. The practicing experts were typically mid-career professionals, with ages that ranged from early 40 s to late 50 s. We judged it too impolite to request exact ages of our experts since we were relying on their voluntary participation. All observers reported normal or corrected-to-normal vision. All observers were naïve as to the purpose of the study, and gave informed consent according to Indiana University guidelines.

2.1.2. Stimuli

We deemed it too difficult for novice observers to match latent to inked prints, and so we instead constructed a database of individual features cropped from inked prints from the NIST 27 database (National Institute of Standards and Technology). Fig. 1 shows pairs of examples from each of the six types of features used, which measured approximately 150 pixels in diameter. Because data was collected via the web, participants were asked to adopt what for them was a comfortable viewing distance. At a normal viewing distance of 27 in. on a 17 in. monitor at a resolution of 1024×768 pixels, the stimuli encompassed approximately 7° in diameter. [1]

Sixteen fragments of each type were included in the study. These were organized into pairs that were as similar as possible. Since the same image is used at both study and test in the X-AB task, care was taken to remove obvious artifacts such as lint or hairs that would

[1] Note that because we have no control over the testing conditions, we could not perform gamma correction or equate luminance and brightness across observers. While this may produce individual variability across subjects, it is unlikely to contribute to differences across groups unless one group used systematically inferior equipment. We designed our software to run on even 5 year old computers, and thus this explanation is unlikely to account for our group differences. In addition, differences in equipment are unlikely to result in the interactions between conditions that we observe for some but not other groups. We discuss the issues raised by a lack of gamma correction in a later section.

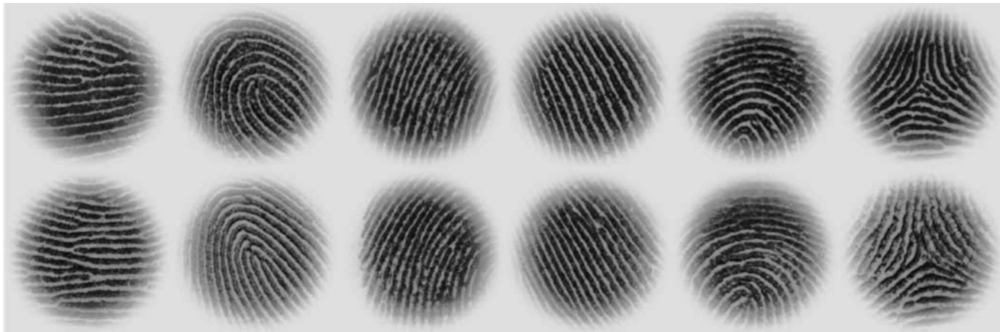

Fig. 1. Example fingerprint fragments used in Experiment 1 from the six different types of fragments. Pairs are grouped vertically in this figure. Each fragment was paired with a close match to reduce reliance on categorical information such as 'loop', 'whorl' or 'arch'.

make a match trivial. In addition, we introduced two manipulations at study designed to reduce the reliance on low-level features. First, we introduced a brightness jitter to the study image, randomly making it up to 20% brighter on each trial. We noted that the two test images in a pair sometimes differed in brightness and this jitter reduced brightness as a cue to the correct answer. We also introduced orientation jitter at study, such that the feature could be rotated up to 30° left or right of its original orientation. This also makes orientation a poor cue to identity. Our goal was to force observers to rely on the structure of the fingerprint features rather than specific feature anomalies that are not related to the structure of fingerprints, since inked and latent prints cannot be matched on the basis of brightness or the presence of lint or some other idiosyncratic feature.

2.1.3. Procedure

Fig. 2 diagrams a typical trial. A single study feature appeared for 1000 ms, which was immediately replaced by a visual mask. This mask remained on the screen for either 200 ms or 5200 ms. This was then replaced by two test features, one of which was the study feature (no longer perturbed by orientation or brightness jitter) and the other was a matched foil. The participant then made an unspeeded forced-choice response indicating the feature they believed was presented at study. Reaction times were not measured and speed was not stressed since actual fingerprint examinations are open-ended and we felt that expert examiners may feel uncomfortable with speeded responses. After the response, visual feedback on the accuracy of their decision was then provided, which served to motivate both groups of observers. We address any learning issues this may have raised in the results and discussion section. Observers clicked on a button for the next trial.

There were 48 pairs used for all 144 trials of the experiment, and either feature could be shown at study. Thus a pair was seen 3 times at test throughout the experiment and each feature was seen on average 1.5 times at study. By using 96 fragments and matched pairs at test, we tried to minimize the likelihood that subjects would learn the fingerprint fragments throughout the course of the experiment.

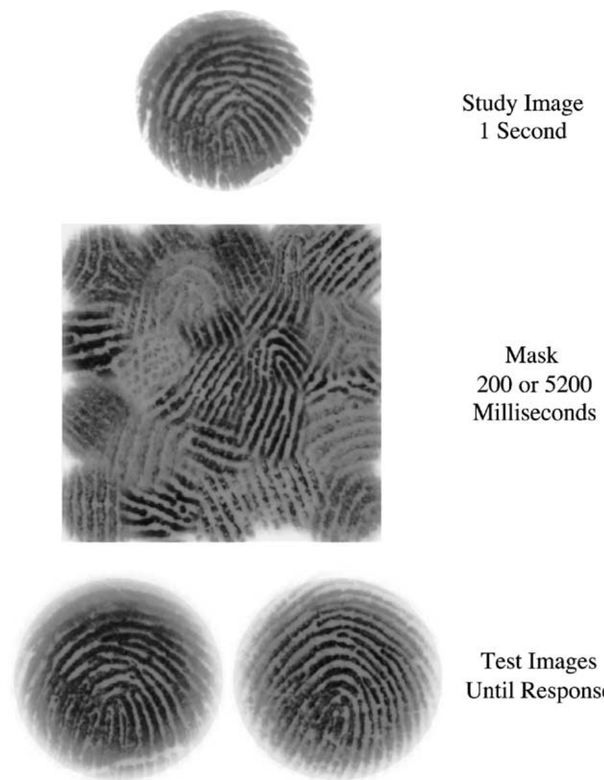

Fig. 2. Sequence of events in Experiment 1. Note that the study image has a different orientation and is slightly brighter to reduce reliance on low-level cues.

As shown in Fig. 3, there were two manipulations that could be applied to the test images. First, the test features could be presented in broadband visual noise, which served to obscure some of the visual features and approximates some of the noise that latent images contain. The spatial characteristics of this noise are not identical to the naturally occurring visual noise, but this manipulation may still tap whatever skills experts have developed to deal with noisy images.

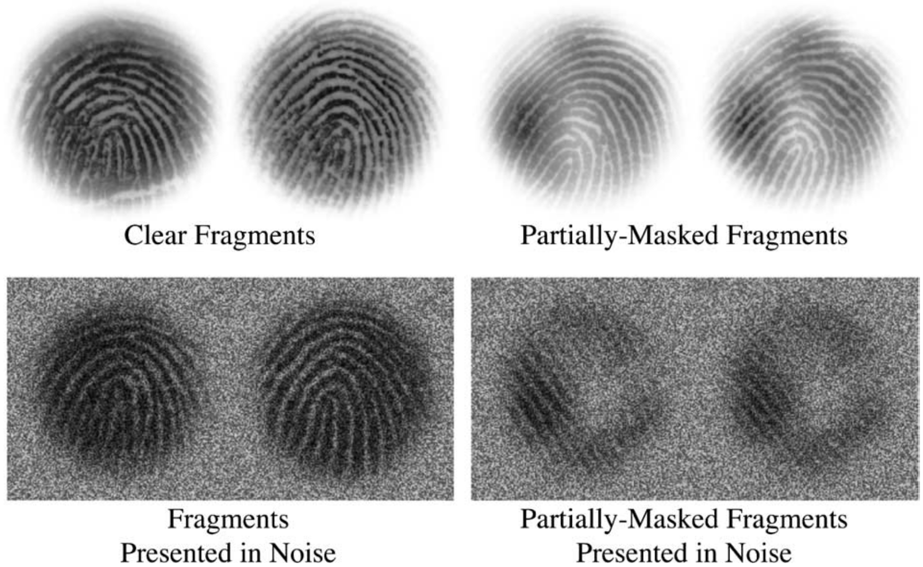

Fig. 3. Four types of test trials.

The second manipulation involved partially masking the print. This has two purposes. First, this manipulation is designed to simulate the fact that latent prints rarely have the amount of detail as the inked print, and tend to be patchy due to many factors, such as the texture of the surface they are found on. Second, this manipulation allows a particular statistical analysis to address configural processing as described in a later section. The construction of partial masks is illustrated in Fig. 4. First, visual noise was generated in Matlab (Mathworks Software) and severely low-pass filtered to produce the image in the upper left. This image was then treated as a semi-transparent mask, such that when multiplied with the fingerprint produces the partial print

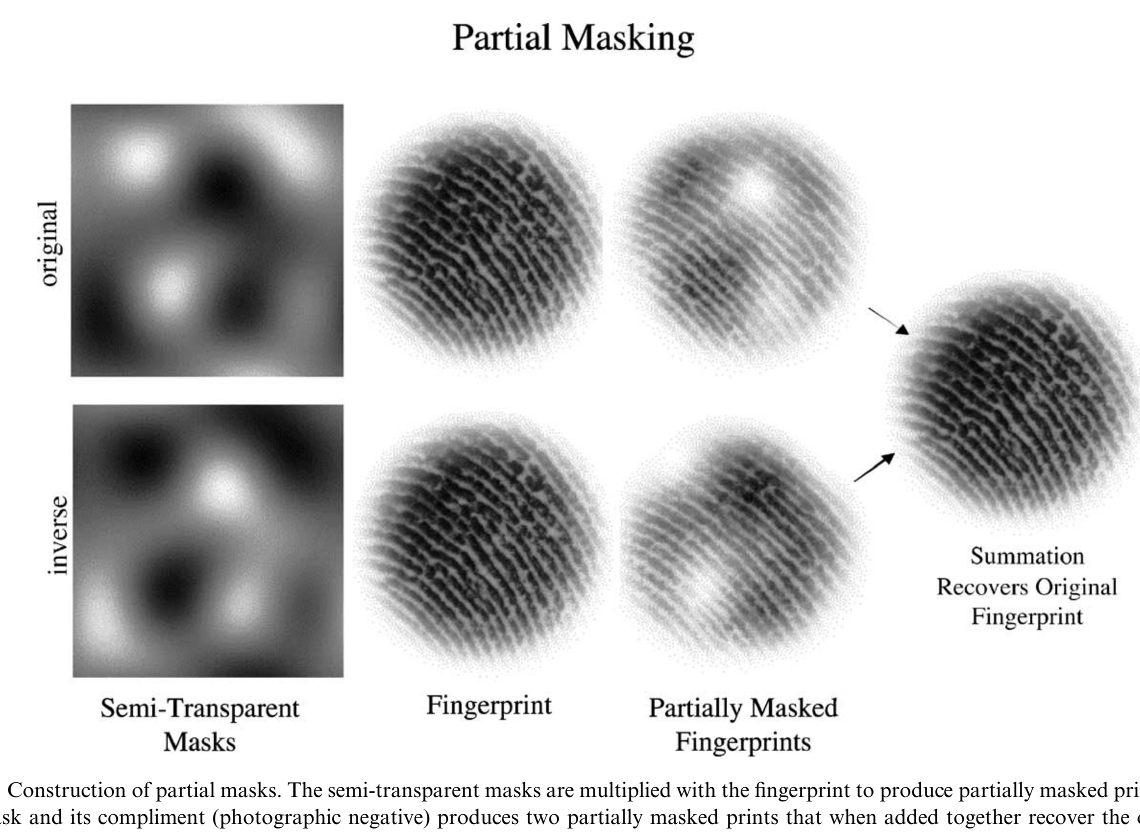

Fig. 4. Construction of partial masks. The semi-transparent masks are multiplied with the fingerprint to produce partially masked prints. Using both the mask and its compliment (photographic negative) produces two partially masked prints that when added together recover the original print.

to the right. Thirty masks were generated in Matlab and then applied in real time during the experiment to both images at test.

The compliment of each mask was also used, as illustrated in the lower left of Fig. 4. This was done by taking the inverse of the mask prior to applying it to the fingerprint feature. When applied to the fingerprint as a mask, this reveals areas previously masked, and hides regions previously revealed. The summation of the two partial prints yields the original fingerprint feature, as show on the right side of Fig. 4. The use of low-pass filtered noise as masks reduces the problems that might occur if sharp edges were used in the masks, which could produce new features at sharp boundaries. This also approximates the patchiness of latent prints when recovered from irregular surfaces.

As shown in Fig. 3, both noise and partial masking could be applied to a test image. There were two levels of each of the three manipulations (short and long mask, clear or noisy features, and full or partially masked prints) that were fully crossed to give eight conditions.

Participants completed 144 forced-choice trials which took approximately 30 min.

2.2. Results and discussion

Results for experts and novices are shown in Fig. 5, which present the proportion of correctly matched features in each condition. The data were submitted to a repeated measures analysis of variance, with expert/novice as a between subject factor and delay, noise and partial masking as within-subject variables. The results in Fig. 5 are graphed separately for short and long mask delays, since neither group showed an interaction between delay and either of the other two variables.

First, consider the experts, which are shown in the top panels of Fig. 5. Surprisingly, the experts show no effect of delay ($F(1,10) < 1$), and no interactions between delay and either noise or partial masking (both F values less than 1). However, experts show effects of both adding noise ($F(1,10) = 197.0$; $p < 0.05$) and partial masking ($F(1,10) = 79.7$; $p < 0.05$). Of particular interest is the strong interaction between partial masking and added noise ($F(1,10) = 151.9$; $p < 0.05$). This interaction may come from several possible sources, but, as discussed below, one intriguing suggestion is that performance on the full prints when presented in noise is higher than one would expect based on partial-image performance.

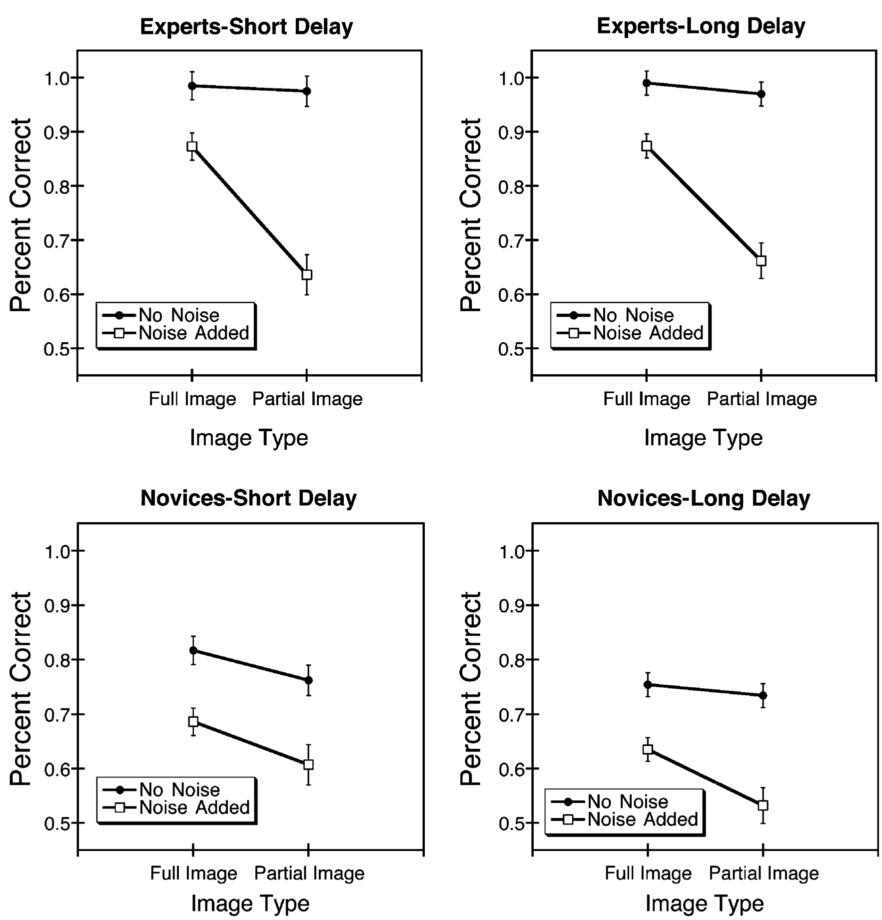

Fig. 5. Experiment 1 data. Error bars represent one standard error of the mean (SEM).

The three-way interaction between delay, noise and partial masking was not significant ($F(1,10) < 1$).

The novices show a different pattern of results. First, performance overall is significantly lower for novices compared with the experts ($F(1,20) = 11{,}204$, $p < 0.05$). In addition, novices show an effect of delay, which can be seen in the lower two panels of Fig. 5 ($F(1,10) = 7.89$, $p < 0.05$), although the effect of delay did not interact with either noise or partial masking (both F-values less than 1). While the novices also show effects of noise ($F(1,10) = 49.6$, $p < 0.05$) and partial masking ($F(1,10) = 15.4$, $p < 0.05$), they fail to exhibit an interaction between the two ($F(1,10) < 1$), which stands in contrast to the strong interaction seen with the experts. As with the experts, the three-way interaction between delay, noise and partial masking was not significant ($F(1,10) < 1$). It is important to note that interpretations of the interaction between added noise and partial masking for the experts is subject to scale dependency issues (e.g. Bogartz, 1976; Loftus & Bamber, 1990), which we address in a later modeling section.

These differences observed between experts and novices separately are confirmed by addressing the interactions with subtype (expert or novice). All three variables interact with subject type (delay × subtype: $F(1,20) = 5.3$; $p < 0.05$; noise × subtype: $F(1,20) = 6.4$; $p < 0.05$; partial masking × subtype: $F(1,20) = 6.9$; $p < 0.05$). In addition, the three-way interaction between subtype, partial masking and added noise was significant ($F(1,20) = 6.46$; $p < 0.05$), which confirms the interaction found between partial masking and added noise found only for experts. This three-way interaction makes the two-way interactions between subtype and the three within-subjects variables somewhat difficult to interpret, with the exception of the delay × subtype interaction, which results from the fact that the experts are unaffected by delay but the novices are.

The inclusion of feedback raises the possibility of learning, which may have affected the results. We used 96 different features in the 144 trials, and so each feature was only seen on average 1.5 times at study and three times at test, and most of the test trials were either partially masked, presented in noise, or both. This minimized the possibility of learning effects, but to test this we split the data for each subject into first and second halves of the experiment and included this as a factor in the original ANOVA described above. The first half/second half factor did not show a main effect, nor did it interact with any other factor or combination of factors (all $p > 0.05$). Thus learning does not appear to be a major factor in Experiment 1.

Of all of these results, most interesting are the effects of delay found only for novices, and the interaction between noise and partial masking found only for experts. These results demonstrate clear differences between the two groups, and the pattern of results suggests the nature of the processing differences. The strong performance observed in experts and their resistance to the longer delay suggests that they encode feature information into more durable storage, such as verbal re-descriptions. In addition, they may possess better visual memory that is robust against the mask, which may facilitate matching images across eyemovements.

Most intriguing, however, is the strong performance with the full image is embedded in noise, relative to the partial image in noise. Consider the values converted to d' values, which tends to linearize percent correct. In the forced-choice task, d' is equal to $(1/\sqrt{2})\,[z\mathrm{Inv}(\mathrm{pc}) - z\mathrm{Inv}(1 - \mathrm{pc})]$, where zInv is the inverse cumulative normal distribution and pc is percent correct for that condition. For experts in the short delay condition, the d' value for the partial image is 0.56, while the d' value for the full image is more than twice that (1.66). The novices showthe opposite pattern: d' for the partial image is 0.39, while the d' for the full image is less than double that value (0.72). Given that the full image can be construed as two partial images, it is somewhat surprising that the d' value more than doubles for the full image in the expert data.

2.3. Evidence for configural processing

What might account for this more than doubling of d' values when the second half of the image is added to create the full image? One possibility that we explore in this section is that of configural processing, where the presence of one part of the image influences the processing of the second part. Configural processing has received much support within the face perception literature (see Maurer, Le Grand, & Mondloch, 2002, and Rossion & Gauthier, 2002, for recent reviews), and is seen as one mechanism supporting perceptual expertise. The interaction between partial masking and subject type (expert vs. novice) suggests that there are differences between the groups, but this interaction is only partially informative. First, it is scale dependent, and the novices may be near floor. Second, an interaction may exist, but experts may show only additive summation, while novices might be sub-additive. Neither would be consistent with configural processing. Thus to address evidence for configural processing we implemented a multinomial model, as described below, which used a probability summation prediction to suggest what performance on the whole image should be given the partial image performance. Actual performance greater than that prediction is consistent with configural processing. Note that this modeling deals with scaling issues as well, since the probability summation prediction builds in the scale into the model. Thus a set of data could be near floor or ceiling and still demonstrate evidence for configural processing.

Recall that for each of the 26 partial masks used in the experiment, its converse mask was also used. Since adding together two partially masked images recovers the full image, each partially masked image contains exactly half of the information of the full image (barring non-linearities, discussed in a later section). Because of this design, we can use performance in the partial mask condition to make a prediction for performance in the full image condition. This was accomplished using a multinomial model, which is illustrated in Fig. 6. The model makes several assumptions:

(1) The observer recovers enough information from one half of an image in order to make a correct response with probability d.
(2) If insufficient information has been recovered in order to make a correct response (which happens with probability $(1 - d)$), the observer can still make the correct response via guessing with probability g, which was set to 0.5.

Full images can be construed as an image consisting of two halves, and therefore contain twice as much information as partially masked images (which contain only one half). The multinomial trees show in Fig. 6 reflect this. For a full image, the model provides two opportunities to recover enough information from half of an image in order to make the correct decision (each with probability d_b). For a partial image, the model provides only one opportunity (with probability d_h).

This model structure allows statistical tests of a configurality hypothesis. If the presence of one half influences the processing of the second half, then d_b should be different than d_h. If $d_b < d_h$ then the presence of the second half reduces the information acquired from the first half (and vice versa). If the two are equal then a form of independence holds and the two halves do not influence each other. The most interesting case is when $d_b > d_h$, which implies that *more* information is acquired from one half when the second half is present (and vice versa). This finding would be consistent with configural processing.

The relation between d_b and d_h was addressed by first testing a reduced model in which $d_b = d_h$. We used the GPT software (Hu & Phillips, 1999), which provides a χ^2 statistic that tests whether this model is rejected by the data. If we reject the reduced model, we have statistical evidence that $d_b \neq d_h$. The next step is to then fit the full model with separate estimates of d_b and d_h. This model is fully saturated, but the estimates of d_b and d_h reveal the directionality of the relation between d_b and d_h.

We saw no interaction between delay and the other two variables for either subject group, so we collapsed across delay when generating model predictions. We fit the multinomial trees separately for the no-noise and noise conditions. The input to the model is the number of correct and incorrect trials in the different conditions, combined for experts and novices separately. The program fits a single set of parameters for all experts and a separate set for all novices.

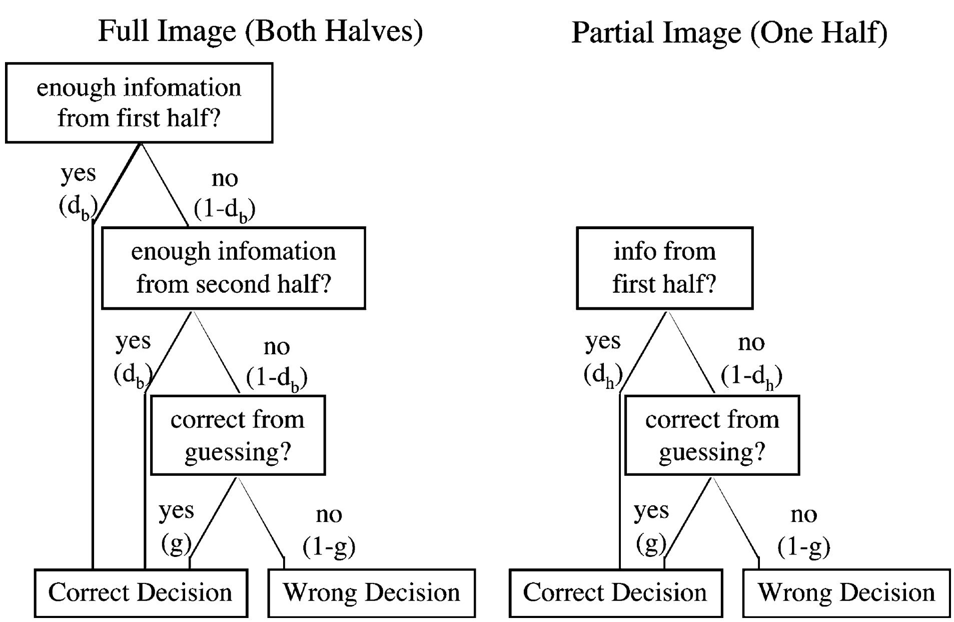

Fig. 6. Multinomial models for full and partial images. The parameters d_b and d_h represent the probability of obtaining enough information from half of an image when that half is in a full image (d_b) or a partial image (d_h). The relation between d_b and d_h determines whether the data contain evidence for configural processing, as described in the text. The parameter g is a guessing parameter set to 0.5, appropriate for the two alternative forced choice paradigm.

For the experts in the no-noise condition, we reject the reduced model ($\chi^2 = 7.75$; $p < 0.05$) and the full model estimated $d_b = 0.841$ and $d_h = 0.944$. This relation between d_b and d_h is opposite that predicted by a configurality hypothesis, and suggests that experts acquire *less* information from each half when the full image is shown. This conclusion may be affected by one of several non-linearities, as discussed later.

A different pattern emerges from the expert data when the images are presented in noise. As with the noise-free data, we reject the reduced model ($\chi^2 = 12.1$; $p < 0.05$). The full model estimates are $d_b = 0.497$ and $d_h = 0.298$. This finding of d_b significantly above d_h is consistent with configural processing, since the results are interpreted as the expert acquiring more information from one half of an image when the other half is present than when the other half is absent.

The novice data for the no-noise conditions are similar to that of the experts: we again reject the reduced model ($\chi^2 = 7.47$; $p < 0.05$) and find that $d_b < d_h$ ($d_b = 0.395$ and $d_h = 0.544$). This is once again the opposite pattern from configural processing. In noise, we find that we cannot reject the reduced model ($\chi^2 = 0.53$; $p > 0.05$), which implies that $d_b \approx d_h$ (the two values are very similar: $d_b = 0.18$ and $d_h = 0.14$). This last result is consistent with a form of independence between the two halves and provides no evidence for configural processing in novices. Thus only experts show evidence of configural processing, and only in the presence of noise.

There is an alternative way to look at these data that also deals with the scaling issues. Consider the data presented in noise for the experts, which shows a steep decline in performance when partially masked. The data for the novices when not in noise covers the same range of performance, and shows a much shallower drop. Thus looking at the data like this visually demonstrates a scale invariant interaction between partial masking and subject type. However, the modeling is still necessary in order to demonstrate configural processing.

The evidence for configural processing in noise with experts is perhaps not surprising when one considers the fact that experts are used to examining latent prints in noise and may have developed abilities such as configural processing to overcome the noise. McKone et al. (2001, 2003) argued that the addition of noise reduces reliance on individual features, since they become unpredictable when noise is added, thus pushing observers to use configural processing if possible (in this case only for our experts). What is perhaps more surprising is that we find evidence for configural processing with singleton fingerprint elements. The analogous experiment with faces (where configural processing effects are often described) would demonstrate configural effects when just single features such as an eye or a mouth is used. Thus in some sense our singleton features worked against a configural processing mechanism, and positive evidence for configural processing should be viewed in this light. This also suggest that configural processing can occur with stimuli that have relevant features in relatively unconstrained locations.

A critical assumption with the multinomial modeling above is that the partially masked stimuli contain exactly half of the information that the full image contains. This is true for a linear system such as an ideal observer, but there are two sources of non-linearities that could break this assumption. The first non-linearity is found in the gamma of monitors. Typically when these experiments are done the monitor is calibrated so that there is a linear relation between the internal scale (pixel values) and the amount of light coming from the monitor (luminance). Because these experiments were done over the web we were unable to calibrate the monitors of the participants. We did consider several systems that might provide rough calibration via gray-level matching, but these were judged to be too cumbersome to explain to our users and lacked any testing that would insure accurate calibration. As a result of this lack of calibration, there may be a non-linear relation between pixel values and luminance which may have made regions that were masked easier to see than they should have been in the no-noise condition, thus inadvertently providing more than half of the information in a partially masked image. This may underlie the sub-linear performance seen with both experts and novices in the no-noise condition. Adding noise to an image alleviates some of these concerns, because now not all of the information in the partially masked areas resides at one set of luminance levels. Thus in some sense the data from the added-noise conditions is perhaps more trustworthy since monitor non-linearities affect it less.

The second source of non-linearities is in the observer's visual system. The transfer function between luminance and the visual system's response may be non-linear, and this may also lead to better performance on the partially masked images than the full image performance can account for.

Neither form of non-linearity can account for the fact that we observe data consistent with configural processing only for the experts when they view images in noise, unless experts have different transfer functions than novices. This indeed could be one possible mechanism: perhaps in the process of examining thousands of fingerprints over the course of a career the experts have altered their sensitivity at low brightness values. Sowden et al. (2000) found better detection performance for radiologists, which suggests we might find similar low-level discriminability for fingerprint experts. The fact that we observe this in noise makes this explanation less likely, but different transfer functions remain a candidate difference between novices and experts. Thus it is desirable to look for converging evidence for configural processing from other domains to support the Experiment 1 conclusions. To explore this finding further, in Experiment 2 we consider evidence from EEG recordings.

3. Experiment 2

3.1. Converging evidence for configural processing from neurophysiology

Within the face perception literature, several different techniques have been used to demonstrate configural processing. From early illustrations of the Thatcher Effect (Thompson, 1980) to the influential Tanaka and Farah (1993) paper, evidence has mounted that upright faces are processed to some degree holistically. While this has taken on different meanings for different authors (see Rossion & Gauthier (2002) and Maurer et al. (2002) for reviews), the general consensus seems to be that for holistic processing, the perception of an individual feature is affected by the context in which it is presented, and that relational information plays an important role. Inverting the face reduces or eliminates these effects (e.g. Farah, Wilson, Drain, & Tanaka, 1998; McKone et al., 2001, 2003; see Rossion & Gauthier, 2002 for a summary of face inversion effects). Electrophysiological evidence of configural processing has been described as a delay of a particular component associated with faces, termed the N170, which is thought to represent "the late structural encoding stages of complex visual information processing" (Eimer, 2000). The N170 component is particularly large when faces are presented, and is thought to originate in parietal/temporal brain regions, primarily on the right side but also on the left (Henson et al., 2003; Horovitz, Rossion, Skudlarski, & Gore, 2004). Several papers have linked the N170 to expertise effects, including Tanaka and Curran (2001), who found evidence for expertise effects with bird and dog experts. A more recent paper by Gauthier et al. (2003) used an interference paradigm to demonstrate that car experts were more likely to automatically encode an irrelevant half of a picture despite instructions to the contrary. By intermixing car and face trials, the authors demonstrated that car and face perception regions interfered, suggesting that some of the same brain areas responsible for face recognition were also recruited for car identification. These behavioral results were found to be correlated with activity in the N170 component in the right hemisphere, suggesting that the expertise effects (in this case emerging through interference effects) were perceptual in nature, rather than strategic or decisional.

The N170 component is reliably delayed for inverted faces, often in both right and left hemispheres (Rossion et al., 2000), but not for other types of stimuli. Rossion et al. (2000) tested a host of non-face stimuli such as houses and greebles in both upright and inverted presentations and found a delayed N170 only for faces. A later training study (Rossion, Gauthier, et al., 2002) using greeble experts found a delayed N170, but primarily in the left hemisphere, while delayed N170 effects for faces tend to be stronger in the right hemisphere. More recent work found evidence for a delayed N170 to inverted cars (Rossion, Joyce, Cottrell, & Tarr, 2003) which was somewhat unexpected given the previous findings. However, it may be that the 3/4 views of cars used in the study may tap some elements of expertise that we gain via our everyday exposure to vehicles. Thus while the delayed N170 component is not specific to inverted faces, it does seem to represent a marker for expertise and possibly a signature of configural processing.

While fingerprints do not have the strikingly different features that eyes, mouths and noses represent, they do have readily identifiable features such as general ridge flow, specific ridge paths with ridge endings and bifurcations, and texture and pore positions. In addition, fingerprints do have an upright orientation and experts almost always orient a print prior to a comparison if the top is possible to determine from the print (sometimes latent prints are difficult to orient). If there exists a common structure to fingerprints, and fingerprint experts learn this structure primarily from upright prints, then this suggests that we might observe configural processing with upright fingerprints.

For novices, fingerprints represent a much more unfamiliar stimulus set than cars or houses that have been used as comparison stimuli. In addition, fingerprint experts receive much more training and exposure than that typically provided by psychology experiments, which may more dramatically alter the cortical representation of these stimuli. Thus upright and inverted fingerprints, when used with experts, provide a good test of the relation between expertise and the delayed N170.

The evidence from Experiment 1 that experts might use configural processing as part of their perceptual analysis, and the delayed N170 component seen with inverted faces has been interpreted as evidence for configural processing. These two lines of evidence suggest an obvious experiment: test fingerprint experts with upright and inverted fingerprints in an EEG experiment. If experts process upright fingerprints in a configurable manner, we should see a delayed N170 with the inverted fingerprints. Experiment 2 tests this prediction. We included upright and inverted faces to replicate the delayed N170 in our experts, and also ran novice observers as a control.

Several authors have made a distinction between category level (or entry level) and subordinate level tasks (Carmel & Bentin, 2002; Rossion, Curran, & Gauthier, 2002; Tanaka, Luu, Weisbrod, & Kiefer, 1999). Tanaka et al. (1999) experiments revealed that the basic and subordinate level categorizations can produce differences in brain activity as early as 130 ms after stimulus onset. While faces may be automatically categorized at the individual level, fingerprints, at least to novices, are likely not. To address this issue, we had our participants first perform an identification task for 400 trials, and then a categorization task for 400 trials. We are primarily interested in latency differences for the N170 component. To anticipate our results, the two tasks produce very similar patterns of results for the latency data, and thus the two tasks may be viewed as a replication in the present context.

3.2. Method

3.2.1. Apparatus

The EEG was sampled at 1000 Hz and amplified by a factor of 20,000 (Grass amps model P511K) and bandpass filtered at 0.1–100 Hz (notch at 60 Hz). Signals were recorded from sites F3, F4, Cz, T5, and T6, with a nose reference and forehead ground; all channels had below 5 kΩ impedance. Recording was done inside a Faraday cage. Eyeblink trials were identified from a characteristic signal in channels F3 and F4 and removed from the analysis with the help of blink calibration trials. Images were shown on a 21 in. (53.34 cm) Macintosh color monitor approximately 44 in. (112 cm) from participants. The data was digitally low-pass filtered below 30 Hz prior to estimation of the N170 latencies for the four conditions.

3.2.2. Observers

We recruited four experts from the United States. All had expertise similar to the experts in Experiment 1. One expert was the second author, although he was naïve as to the purposes of the EEG experiment prior to his participation. Four novice observers were recruited from the Indiana University community who did not have experience with fingerprint stimuli. As in Experiment 1, the experts were mid-career professionals, while the novices were advanced undergraduates.

3.2.3. Stimuli

The entire stimulus set appears in Fig. 7. Face stimuli consisted of grayscale frontal views of eight bald men. Fingerprint stimuli were chosen from the NIST 27 database of fingerprint stimuli, and were fully rolled standard prints for 10-print records rather than latent prints. We used fully rolled standard fingerprints rather than fragments in order to make orientation quickly apparent to the participants.

Faces subtended a visual angle of $5.8° \times 6.8°$. Fingerprints subtended a visual angle of approximately $7.0° \times 7.0°$. The images were shown at full contrast on a monitor with background set to 19.2 cd/m^2, black set to. 76 cd/m^2 and white set to 61.8 cd/m^2.

3.2.4. Procedure

Observers completed two halves of the experiment. In the first half they were to identify which of the eight faces or fingerprints was presented on each trial. Each image appeared for 1000 ms, which was followed by the observer's response of 1–8 on a numeric keypad. They were given a sheet with the 16 stimuli along with numbers assigned to each stimulus. All observers were asked to hold the sheet upright even if the stimulus on a particular trial appeared upside down. In the second

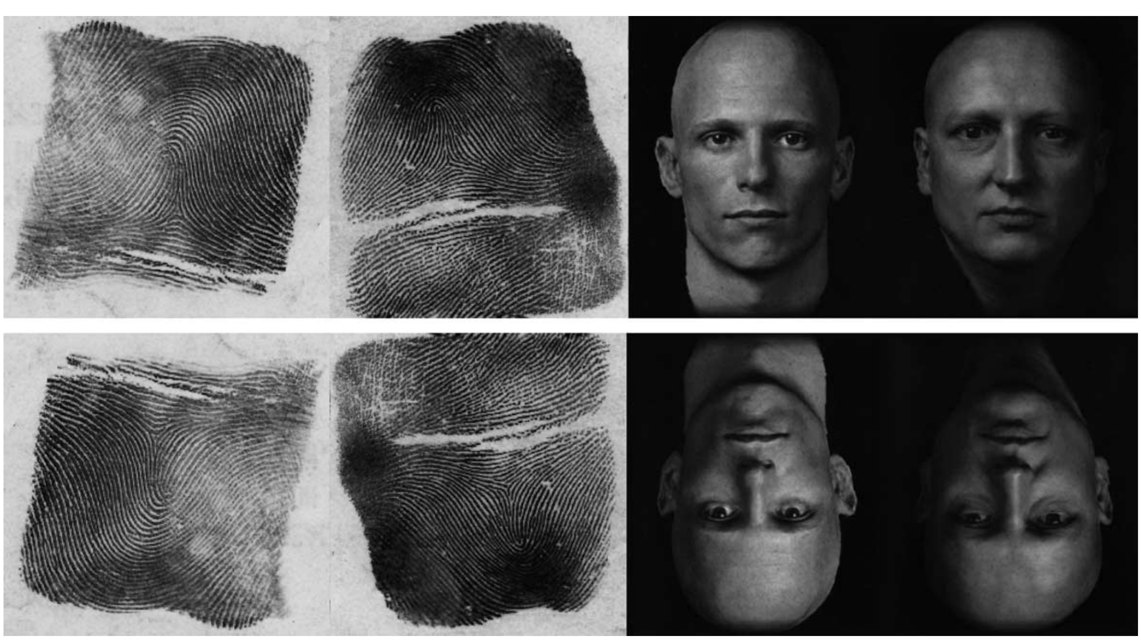

Fig. 7. Examples of Experiment 2 stimuli.

half of the experiment observers made a face/fingerprint categorization task, pressing one of two keys. No feedback was provided.

Stimuli were presented for 1000ms. EEG was recorded from 100ms prior to stimulus onset to 1100ms post-stimulus onset. The stimuli appeared in randomized order. Observers completed 200 trials (100 in each task) of each of the four main stimulus types (upright and inverted faces and upright and inverted fingerprints) for a total of 800 trials. All observers completed 400 trials of the identification task followed by 400 trials of the categorization task. We deliberately did not counterbalance the order of task across subjects because we judged the identification task to be more important and we did not want it to be influenced by any fatigue effects. As it turned out the effects are qualitatively similar for the two tasks and thus they represent a replication of the effects within the experiment. The lack of counterbalancing makes direct amplitude comparisons problematic should one wish to compare identification and categorization brain responses.

The inter-trial interval was set by the observer since they initiated the next trial with the response to the previous trial. After their response the next trial appeared with a delay ranging from 1700 to 1800ms. While this delay was random within this interval, the EEG signal may be contaminated with slow anticipatory waves (e.g. Vogel & Luck, 2000). While these cannot contribute directly to our condition differs due to the random order of the stimuli, we filtered our data using both a 1Hz highpass filter and also a linear drift correction algorithm. Neither signal processing technique altered the pattern of latency results significantly, although the highpass filtering produced noticeably cleaner data and thus we present the results from the filtered data. The unfiltered data produced very similar patterns of results.

3.3. Results

Behaviorally, accuracy was reduced for both inverted fingerprints and inverted faces relative to their upright versions, but this reached significance only for faces for both groups (novices: $t(3) = 5.1$; $p < 0.05$; experts: $t(3) = 4.8$; $p < 0.05$). There were no significant differences between novices and experts on any of the four stimulus conditions. Note, however, that behavioral performance is based on viewing the entire 1000ms presentation, whereas the N170 differences discussed below are based only on the initial percept of the stimulus. In addition, accuracy was very high (the lowest was 77% where chance is 12.5%). Thus there may be processing differences between experts and novices as revealed by EEG that are not evident in behavioral data.

We now turn to the EEG data. The data from Experiment 2 is shown in Figs. 8 and 9 for the electrodes of interest, in this case T5 and T6, which are located in the left and right parietal/temporal regions. Data from experts is shown in the top panels, while data from novices is shown in the bottom panels. Vertical lines are the computer-based estimates of the latency of the N170, which finds the minimum value in a window that includes the N170 component, in our case the window between 125 and 200ms.

3.3.1. Analysis of variance

The prior results derived from the literature provide a clear prediction regarding the latencies of the N170 component for upright and inverted stimuli. Before addressing this specific comparison, we first report the results of an overall analysis of variance, which has four within subject factors: Task (identification and categorization), Channel (T5 and T6), Stimulus (faces and fingerprints) and Orientation (upright and inverted). Subject type (expert or novice) was a between-subject factor. Readers mainly interested in the delayed N170 predictions may wish to skip to the next section and come back to the ANOVA results.

Given the predictions from the literature, the most interesting comparison is the interaction between stimulus, orientation and subject group. The Stimulus × Orientation × Subject Type interaction was significant ($F(1,6) = 10.3$; $p < 0.05$). The related four-way interaction that includes task was marginally significant ($F(1,6) = 4.1$; $p = 0.089$) but the four-way interaction between Stimulus, Orientation, Channel and Subject Type did not reach significance ($F(1,6) = 1.5$; $p > 0.05$). This latter result suggests that while the effects of inversion are larger in the right hemisphere, there may be enough differences in the left hemisphere to make this interaction non-significant. Thus both hemispheres may contain effects consistent with configural processing. The five way interaction that includes all factors was marginally significant ($F(1,6) = 4.1$; $p = 0.09$). There was also a significant Task × Stimulus × Subject Type interaction ($F(1,6) = 6.6$; $p < 0.05$). Finally, the main effects of Stimulus ($F(1,6) = 7.4$; $p < 0.05$) and Orientation ($F(1,6) = 59.8$; $p < 0.05$) were both significant, as was the interaction between Task, Stimulus and Orientation ($F(1,6) = 6.8$; $p < 0.05$). Overall, experts show faster N170 latencies ($F(1,6) = 7.94$; $p < 0.05$).

3.3.1.1. Testing specific predictions from the literature. We now turn to the specific predictions provided by prior work. Based on the existing literature we have a clear a priori prediction: inverted faces should produce an N170 that is delayed relative to upright faces. If experts process fingerprints configurally, they should show a delayed N170 for inverted fingerprints relative to the upright fingerprints. Given this prediction from the literature we conducted paired one-tailed t-tests comparing latencies for upright vs. inverted stimuli, with alpha set

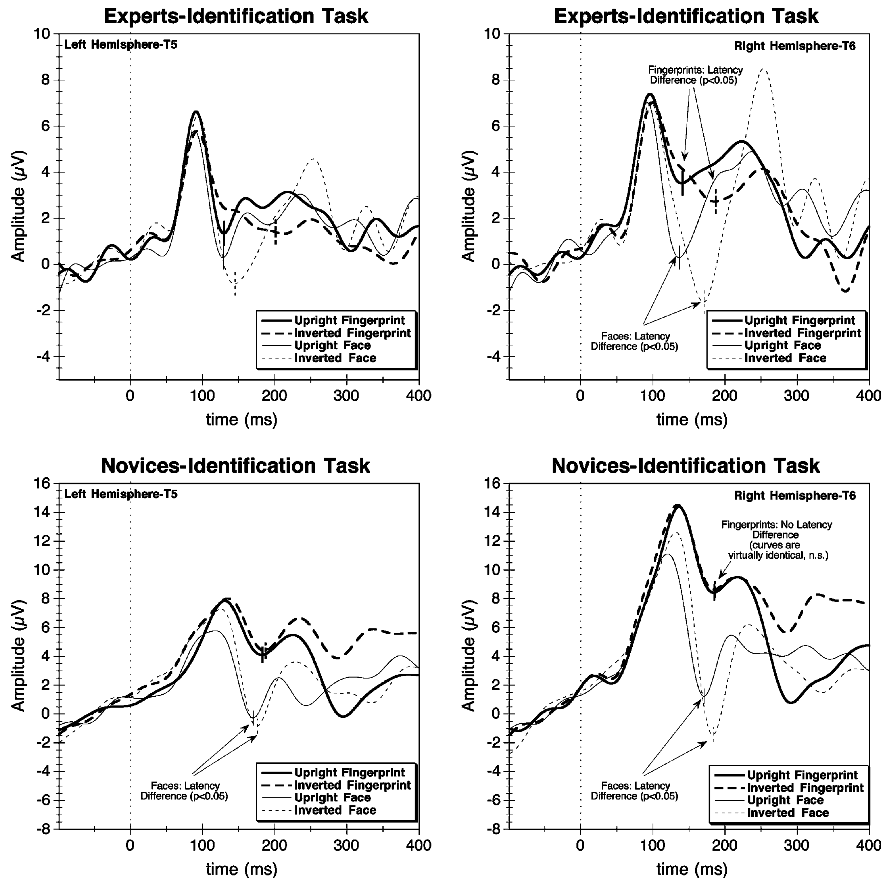

Fig. 8. Experiment 2 data—identification task. Vertical bars represent computer estimates of latency. Note different vertical scales for novices and experts.

to 0.05. Note that data from the identification and categorization tasks comes from different halves of the experiment, and we discuss the data separately below for each task. We focus primarily on latency effects, since while the N170 is also sometimes enhanced as well as delayed when stimuli are inverted, this effect does not always obtain in inversion experiments.

3.3.1.2. Data from the identification task. Consider the data from the face stimuli, shown as thin curves in Fig. 8. The largest latency differences are found in the right hemisphere (channel T6) of both groups, for both faces and fingerprints. With regard to faces, the inverted faces produced a delayed N170 relative to upright faces, in both novices ($t(3) = 7.64$; $p < 0.05$) and experts ($t(3) = 3.58$; $p < 0.05$). This replicates the existing literature and is consistent with configural processing of faces by both sets of subjects. The left hemisphere (channel T5) produced a significant latency difference only for novices for faces ($t(3) = 3.30$; $p < 0.05$).

We now turn to fingerprints, which show a similar pattern for experts but not novices. As shown by the thick curves in Fig. 8, the experts show a reliable difference between the upright and inverted fingerprints that begins as early as 130–140 ms after stimulus onset. In the right hemisphere (Channel T6), the inverted fingerprints produce an N170 that is systematically delayed ($t(3) = 3.54$; $p < 0.05$) relative to the upright fingerprints. The left hemisphere produced a similar delay, but did not reach significance ($t(3) = 2.56$; $p = 0.083$). The upright and inverted latencies for the novices are almost identical, and there was no significant difference between the latencies of upright and inverted N170 components ($t(3) < 1$) for these observers. Indeed their two curves follow the same trajectory until about 230 ms after stimulus onset. Thus the inversion effect with fingerprints is limited to expert examiners.

3.3.1.3. Data from the categorization task. The data from the categorization task mirrors that of the identification task in almost every respect, as shown in Fig. 9. In the right hemisphere, we find a latency difference for upright vs. inverted faces for the experts ($t(3) = 4.21$; $p < 0.05$), as well as for novices ($t(3) = 5.41$; $p < 0.05$).

The fingerprint data also replicates that from the Identification task, as shown in Fig. 9. In the right hemisphere, the experts show a delayed N170 component for the inverted fingerprints relative to upright fingerprints ($t(3) = 3.23$; $p < 0.05$). The results in the left hemisphere

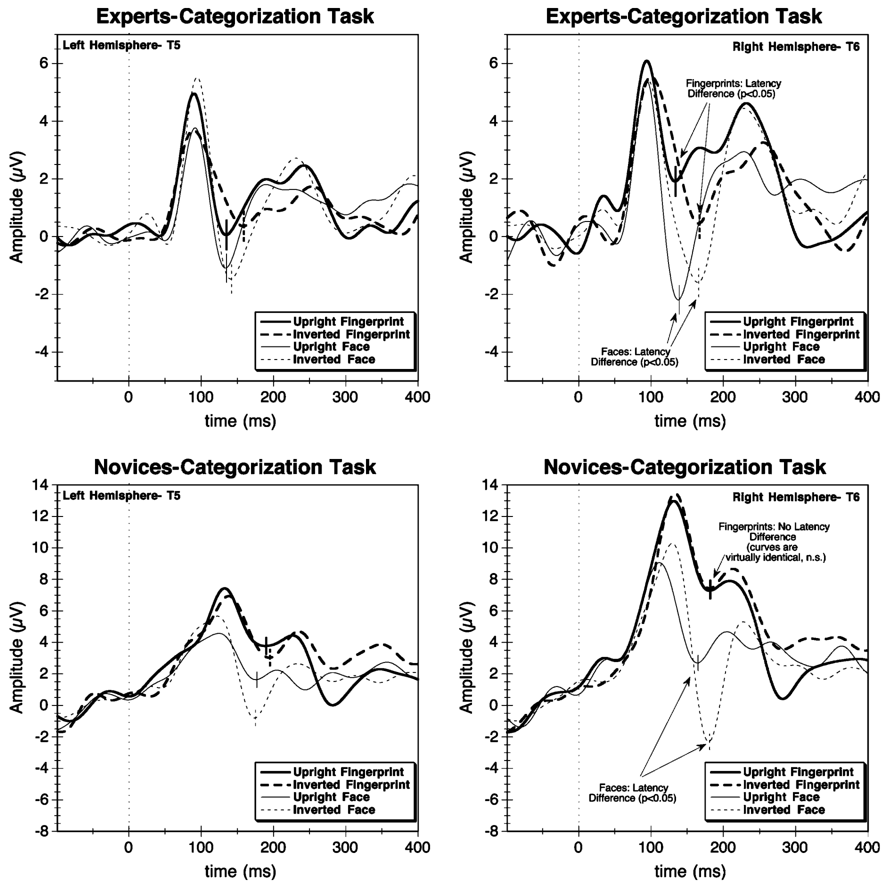

Fig. 9. Experiment 2 data—categorization task. Vertical bars represent computer estimates of the latency of the N170 component. Note different vertical scales.

produced a similar delay, but did not reach significance ($t(3) = 2.71$; $p = 0.073$). The novices show virtually identical latencies for upright and inverted fingerprints in both the right ($t(3) = 1.6$, n.s.) and left hemispheres ($t(3) = 1.97$, n.s.). The data from the experts begins to separate as early as 120 ms, while the novice data remains together until about 200 ms.

No significant effects were found in any of the amplitude data for either group of subjects. This is not surprising given the inconsistency of amplitude effects in the literature.

We chose not to analyze data at later time intervals. The N170 component is thought to reflect elements of perceptual processing, and establishing differences between experts and novices at the level of this component demonstrates that at least part of the elements of expertise lie in perception. This is an important conclusions since the results of Experiment 1 could be linked to better memories or strategies on the part of experts. While later components also may show differences, it will likely require experiments with additional conditions to fully identify the nature of the differences, and thus the later components are beyond the scope of the present article.

3.4. Discussion

The delayed N170 components for inverted fingerprints seen only in experts are consistent with configural processing of upright fingerprints. In particular, we find that the delay effects with fingerprints is found in the same general EEG components as those found with faces, suggesting that some of the same neural processes involved in expertise with faces may be recruited for fingerprints. These effects occurred mainly in the right hemisphere (where our face effects were also largest), and is consisting with other cognitive neuroscience findings that suggest that holistic processing and expertise effects are often larger in the right hemisphere (Gauthier et al., 2003; Gauthier & Tarr, 2002; Gauthier et al., 1999, but see Rossion et al., 2002 for left-hemisphere training effects). In conjunction with the mathematical modeling results of Experiment 1, this research provides converging evidence to suggest that experts have adopted a different form of processing when viewing upright fingerprints. This representation appears to include at least local relational information of the kind that produces dependencies between individual features such that they begin to be affected by the context in which they are presented.

Several authors have proposed different models of the nature of the representation that supports configural effects. Maurer et al. (2002) suggests that both first order and second order relations are important, with the metric information provided by spacing especially relevant for faces. A slightly different view was proposed by Rossion and Gauthier (2002), who stress distinctive local relational information or an alternative model that relies on local overlapping holistic templates. A single holistic template that does not represent individual features seems implausible.

The present study suggests that local feature relations are sufficient to produce configural processing, since Experiment 1 used fingerprint fragments rather than whole fingerprints, and found evidence for configural effects with observers. More importantly, many of the important features such as bifurcations and ridge endings are found in idiosyncratic locations in the fragments. Thus experts cannot have pre-manufactured templates that code relational information as in template match models. Instead, experts seem to possess the ability to quickly encode novel relational information once the features have been identified. This process must occur relatively quickly, since the features were only visible for 1 s prior to the mask. In addition, the experts have the ability to maintain this relational information in memory, which is hard since metric information resists verbal redescriptions.

The delayed N170 to inverted stimuli has been argued as a signature of configural processing in the literature (see Rossion & Gauthier, 2002 for review), but this raises the question of why the absence of configural processing should produce a delayed N170. A more appropriate interpretation might be that an upright stimulus produces an *earlier* N170 component, through either faster propagation of signals or more likely more neurons become active simultaneously. Distinguishing between advancement or delay hypotheses will require additional data. However, regardless of the neural mechanism, the latency differences for fingerprints with our experts demonstrate that they process upright and inverted stimuli differently, and the evidence from the face literature suggest that experts process upright fingerprints in a *qualitatively* different fashion, one in which the image is viewed as a gestalt rather than as collections of individual features. The present findings continue the theme of the development of expertise through acquisition of configural processes.

Evidence for inversion effects producing qualitatively different processing is based in part on interpretations of the delayed N170 effect as described in the literature, and links to behavioral work that suggests that upright faces are processed configurally (Farah et al., 1998; Freire, Lee, & Symons, 2000; Leder & Bruce, 2000; McKone et al., 2001, 2003). Work with patients has also suggested an isolated configural mechanism (Moscovitch, Winocur, & Behrmann, 1997). While a wide consensus is evident in the literature in support of configural processing in upright faces and the evidence is viewed by many as overwhelming, two contradictory views have recently been raised and should be noted. Sekuler, Gaspar, Gold, and Bennett (2004) measured classification images for upright and inverted faces in an identification task, which treats individual pixels as features and computes the regions that are most affected by adding noise. These regions are then inferred to be the features that are used by observers when making identifications. This process assumes a linear template that has no dependencies between features (pixels), and therefore acts as a null hypothesis for non-linear templates (i.e. configural processing). While performance was worse for inverted faces, the authors used the relation between efficiency and classification images to conclude that these differences were quantitative rather than qualitative. Thus if there are configural effects, they appear to be the same for upright and inverted faces. These results may be a function of the specific paradigm used in classification images, which involves only two faces and tens of thousands of trials, during which the observers may begin to process the images differently. For instance, they may begin to treat both the upright and inverted images as templates, processing both holistically. This suggestion remains speculative, however.

A second critique of at least one of the paradigms that have addressed holistic effects is by Wenger and Ingvalson (2002, 2003). Using a multidimensional extension of signal detection theory, they were able to link superior performance for upright faces reported by Farah et al. (1998) to criterion shifts rather than improvements in sensitivity when a feature is presented in its correct context. This suggests that subjects are not processing faces at a holistic level perceptually, but instead are using information from other features to alter their decision. Our delayed N170 component is thought to represent a visual processing stage, and thus it is less clear how decision biases might affect this process. However, there remains the intriguing possibility that the biases may have a perceptual locus.

It is not clear how far either criticism generalizes to the broad range of configural effects discussed in the literature, and clearly converging evidence may help delineate the processes that underlie configural effects. One issue that the two preceding examples illustrate is that configural effects may depend on the task used to assess them. Fingerprints may represent a fairly clean stimulus set because novices lack familiarity while experts have extensive training at the subordinate level. In addition, our Experiment 1 paradigm had no contradicting cues as in the Farah et al. (1998) designs and no inverted stimuli, and our Experiment 2 data showed latency differences with both tasks with very early time-course differences that are difficult to explain via decision mechanisms as in Wenger and Ingvalson (2002, 2003) unless some form of rapid feedback from higher cortical areas is involved (e.g. Ahissar & Hochstein, 2000). Thus the criticisms raised by the above examples cannot directly explain our results as less efficiency for inverted stimuli or interactions at decision stages.

A recent debate has emerged in the literature as to whether the delayed N170 effect supports a domain specific account of faces (Bentin & Carmel, 2002; Carmel & Bentin, 2002) or represents subordinate-level expertise and that any well-learned stimulus will produce the effect (Rossion, Curran, et al., 2002). The central issue seems to be whether the same neural substrates that are used to process faces are also recruited for other stimuli. Face-like responses have been observed in response to training (e.g. Gauthier et al., 2000; Tanaka & Curran, 2001) are sometimes found in similar but not precisely the same regions that produce robust face responses (Rossion et al., 2002). The most direct evidence comes from Gauthier et al. (2003), who showed interference between a face and a car task for experts, suggesting that the same neural areas were subserved by both. The Experiment 2 results lack the large number of channels that are required to do precise localization, but the fact that we observe N170 delayed response in the same channel for both faces and fingerprints in experts suggests that similar neural mechanisms are at work. It will probably require the specificity of single-cell recording to fully resolve the issue of whether face neurons can begin to represent other stimuli through training. However, we find similar effects in the same channel with faces and fingerprints (which are very unlike faces), and we interpret this as more consistent with a subordinate-level account. We are currently exploring this issue with a larger EEG recording setup to identify the precise locus of learning and the present conclusions about domain specificity must be tentative.

A strength of EEG recording is that it has excellent temporal acuity, and thus we have at least an upper bound on when a process has completed. Given the data shown in Figs. 8 and 9 for experts, we see that their voltage data begins to show differences as early as 130–150 ms and certainly by 180 ms. This timecourse places constraints on what kind of cognitive processes might contribute to the N170 component. Surprising events that require context updating often show large effects only at later components such as the P300 or N400. This latency information, paired with the spatial localization of the N170 to the parietal/temporal region (Henson et al., 2003; Horovitz et al., 2004; Itier, Taylor, & Lobaugh, 2004; Rossion et al., 1999) suggests that the N170 represents perceptual processing rather than later decision mechanisms. This is particularly relevant to our expert data, since it seems unlikely that our latency differences result just from the experts noticing that the stimuli are inverted (this would have produced latency differences for our novices as well) or demand characteristics on the part of our experts.

4. General discussion

The results from these two experiments provide some initial evidence for the nature of expertise among fingerprint experts. First, as one might expect, experts perform much better than novices in the behavioral task in Experiment 1. More surprising is their robustness against the delay, and this suggests that one reason they perform so well is that they are better able to encode the visual stimulus into a more durable storage that resists decay over time. Part of this facility may arise from a knowledge of which aspects of the features are relevant. Note that our stimuli are constructed in a way that paired similar features at test. Experts might focus on that information that distinguishes individual exemplars of a feature type, while novices might try to remember the kind of feature presented (lacking the ability to discriminate between features within a type). Thus even though the experiment procedures were described and the construction of test pairs was explained to both groups of subjects, experts may have been performing a within-class identification procedure while novices may have inadvertently been performing a between-feature classification task. The ability to encode and discriminate within-class exemplars is one form of expertise and this principle may apply here as well.

The evidence for configural processing as revealed by multinomial modeling in Experiment 1 was supported by the ERP results of Experiment 2, which also suggests a perception component to the fingerprint expertise, rather than just better memory or strategies on the part of experts. While the data from both experiments is in agreement with prior studies of expertise (Rossion et al., 2002; Tanaka & Curran, 2001), the present work suggests that if relational information subserves the configural effects, then this relational information must be quickly computed from features at idiosyncratic locations on the print. Within the literature there is some question of whether the delayed N170 represent a signature of configural processing of the upright, well-learned stimulus, or some other kind of expertise such as faster processing of familiar stimuli. That is, well-learned stimuli need not be processed in a configural manner, and there are some recent suggestions that the interactions are taking place at the decision stage (Wenger & Ingvalson, 2002, 2003). The combined evidence from Experiments 1 and 2 point to a configural process that underlies the delayed N170 effect, at least in our fingerprint experts.

The data in support of configural processing in experts comes only from the noise-added conditions of Experiment 1, while full clear images were used in Experiment 2. These methodological changes were made as a result of the procedural requirements of the inversion test in EEG recordings. It is perhaps not surprising, however, that experts show configural processing when presented with noisy images, given that latent prints are often corrupted by noise not associated with the fingerprint texture. The fact that we find configural processing even with fingerprint fragments suggests that this is a robust phenomenon among experts.

Our tasks were deliberately abstracted from the actual task of latent/tenprint matching, and care must be taken to specify exactly how far we wish to generalize our results to the actual task of fingerprint examinations. One topic that is under current discussion in legal settings is whether testimony from fingerprint examiners should be allowed in court, or whether juries should simply be given fingerprint evidence for evaluation and allowed to draw their own conclusions (e.g. US vs. Byron Mitchell). Central to this debate is whether fingerprint experts possess perceptual abilities not found in novices. The results of both experiments suggest that on tasks that are related to actual examinations, experts do in fact show qualitative differences as well as overall better performance on behavioral tasks. These differences extend to the neural signature of the perceptual processing of the stimuli. Whether these differences translate to better accuracy when it comes to actual fingerprint identifications is an open question, since we did not test latent/tenprint matches. However, it seems *unlikely* that the qualitatively different processes exhibited by experts would make them *less* accurate than novices at identifications, and there are lots of plausible perceptual mechanisms that suggest that experts would show improved identification performance as a result of the differences revealed by these experiments.

This research still leaves open the question of what features, detail or information experts use when matching fingerprints. That we have not addressed this may not be surprising given that researchers in the face recognition literature are still working on what constitutes a feature or the basis functions that describe a face. In some sense the finding of configural processing complicates matters, because configural processing implies dependencies between features that must be considered and a simple model that assumes independence among features will likely fail at some point. A complete approach will likely involve a combination of behavioral testing, eyemovement recording, cognitive neuroscience experiments and mathematical modeling. These approaches remain active topics in our research program.

Acknowledgments

The authors would like to thank Bethany Knapp for her assistance with the Multinomial Processing Tree modeling and Robert Goldstone for comments on an earlier draft. TAB was supported by a grant from NIH.

References

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

Baker, C. I., Behrmann, M., & Olson, C. R. (2002). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience, 5*, 1210–1216.

Bentin, S., & Carmel, D. (2002). Accounts for the N170 face-effect: A reply to Rossion, Curran, & Gauthier. *Cognition, 85*, 197–202.

Bogartz, R. S. (1976). On the meaning of statistical interactions. *Journal of Experimental Child Psychology, 22*, 178–183.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition, 83*, 1–29.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuro Report, 11*, 2319–2324.

Farah, M. J., Wilson, K. D., Drain, H. M, & Tanaka, J. R. (1998). What is "special" about face perception? *Psychological Review, 105*, 298–482.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*, 159–170.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*, 428–432.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*, 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a 'Greeble' expert: Exploring mechanisms for face recognition. *Vision Research, 37*, 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*, 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*, 568–573.

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

Goldstone, R. (1998). Perceptual learning. *Annual Reviews Psychology, 49*, 585–612.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*, 1047–3211.

Horovitz, S. G., Rossion, B., Skudlarski, P., & Gore, J. C. (2004). Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *NeuroImage, 22*, 1587–1595.

Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods. Instruments and Computers, 31*, 220–234.

Itier, R. J., Taylor, M. J., & Lobaugh, N. J. (2004). Spatiotemporal analysis of event-related potentials to upright, inverted, and contrast-reversed faces: Effects on encoding and recognition. *Psychophysiology, 41*, 643–653.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology, 53A*, 513–536.

Loftus, G. R., & Bamber, D. (1990). Learning-forgetting independence: Unidimensional memory models and feature models: Comments on Bogartz (1990). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 916–926.

Logothetis, N. K. (2000). Object recognition: Holistic representations in the monkey brain. *Spatial Vision, 13*, 165–178.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *TRENDS in Cognitive Sciences, 6*, 255–260.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 573–599.

McKone, E., Martini, P., & Nakayama, K. (2003). Isolating holistic processing in faces (and perhaps objects). In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes. Advances in visual cognition* (vol. viii, pp. 92–117). London: Oxford University Press, 393.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9*, 555–604.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning and Memory Cognition, 14*, 553–5557.

NIST 27: Fingerprint minutiae from latent and matching tenprint images. National Institute of Standards and Technology, Gaithersburg, MD.

Rossion, B., Curran, T., & Gauthier, I. (2002). A defense of the subordinate level-expertise account for the N170 component. *Cognition, 85*, 189–196.

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guerit, J. M. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*, 173–189.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews, 1*, 62–74.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science, 13*, 250–257.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuro Report, 11*, 69–74.

Rossion, B., Joyce, C., Cottrell, G., & Tarr, M. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage, 20*, 1609–1624.

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representation in object concepts. In D. L. Medin (Ed.), *The psychology of learning and motivation, Vol. 31* (pp. 305–354).

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory and Cognitiona, 23*, 681–696.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*, 391–396.

Shiffrin, R. M., & Lightfoot, N. (1997). Perceptual learning of alphanumeric-like characters. *The psychology of learning and motivation* (Vol. 36, pp. 45–82). San Diego, CA: Academic Press.

Sowden, P. T., Davies, I. R. L., & Roling, P. (2000). Perceptual learning of the detection of features in X-ray images: A functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance, 26*, 379–390.

Tanaka, J., Luu, P., Weisbrod, M., & Kiefer, M. (1999). Tracking the time course of object categorization using event-related potentials. *NeuroReport, 10*, 829–835.

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

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12*, 43–47.

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

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*, 764–769.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9*, 483–484.

US v. Byron Mitchell, July 7–13, 1999, 96-407, US District Court for the Eastern District of Pennsylvania.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*, 190–203.

Wenger, X., & Ingvalson, X. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 872–892.

Wenger, M. J., & Ingvalson, E. M. (2003). Preserving informational separability and violating decisional separability in facial perception and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1106–1118.

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