Journal of Experimental Psychology:
Human Perception and Performance
2002, Vol. 28, No. 2, 431–446

0096-1523/02/$5.00 DOI: 10.1037//0096-1523.28.2.431

Unraveling Mechanisms for Expert Object Recognition: Bridging Brain Activity and Behavior

Isabel Gauthier
Vanderbilt University

Michael J. Tarr
Brown University

Behavioral sensitivity to object transformations and the response to novel objects (Greebles) in the fusiform face area (FFA) was measured several times during expertise training. Sensitivity to 3 transformations increased with expertise: (a) configural changes in which halves of objects were misaligned, (b) configural changes in which some of the object parts were moved, and (c) the substitution of an object part with a part from a different object. The authors found that *holistic–configural* effects can arise from object representations that are differentiated in terms of features or parts. Moreover, a *holistic–inclusive* effect was correlated with changes in the right FFA. Face recognition may not be unique in its reliance on holistic processing, measured in terms of both behavior and brain activation.

Considerable neuropsychological and neuroimaging evidence exists for an area in the ventral temporal cortex that is selective for faces, or at least more active for face processing as compared with the processing of most other object categories (De Renzi, 1986; Farah, 1990; Farah, Levinson, & Klein, 1995; Haxby et al., 1994; Kanwisher, Chun, McDermott, & Ledden, 1996; Kanwisher, McDermott, & Chun, 1997; McNeil & Warrington, 1993; Moscovitch, Winocur, & Behrmann, 1997; Puce, Allison, Gore, & McCarthy, 1995; Sergent, Ohta, & MacDonald, 1992; Sergent & Signoret, 1992). Neuroimaging studies have begun to investigate the computational role of this area in object recognition (e.g., Gauthier, Tarr, Moylan, Skudlarski, et al., 2000; Kanwisher, Tong, & Nakayama, 1998). However, a variety of factors limit the number of possible manipulations that can be included in neuroimaging studies. Moreover, functional magnetic resonance imaging (fMRI) methods may not be sufficient by themselves to resolve what computations are taking place in any given neural substrate. In contrast, there is an extensive psychophysical tradition of numerous subtle manipulations on face recognition that address critical computational issues (e.g., Bruce, 1982; Diamond & Carey, 1986; Farah, 1996; Galper, 1970; Yin, 1969). We rely on this tradition in this article, reporting the results of multiple psychophysical experiments during the acquisition of expertise with novel objects. To leverage these methods, our approach combines psychophysical assessment with neuroimaging techniques in two ways. First, the psychophysical procedure we used to train participants to expertise was the same as that used in a recent neuroimaging study, in which activity for Greebles in the fusiform face area (FFA) increased with expertise (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; half of the 10 participants in the present study also participated in this earlier fMRI study). Second, we correlated behavioral measures of expertise acquisition with concurrent neural changes in these same participants.

A good deal of research has been devoted to understanding the factors that determine category selectivity in the FFA (Chao, Martin, & Haxby, 1999; Gauthier, Anderson, Tarr, Skudlarski, & Gore, 1997; Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier et al., 1999; Gauthier, Tarr, Moylan, Anderson, et al., 2000; Gauthier, Tarr, Moylan, Skudlarski, et al., 2000; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Kanwisher et al., 1997; Tanaka & Gauthier, 1997). An emerging hypothesis is that the neural specificity seen for face processing is due to the particular constraints of humans' extensive experience with faces (Gauthier, 2000). A correlate is that given similar constraints, nonface stimuli will recruit the same neural substrate. In other words, the FFA may not be specific for faces per se, but rather only for the operations we typically, and by default, perform when perceiving faces.

Faces consistently elicit more activity than a variety of control stimuli in the FFA, a finding that is not disputed even when authors disagree as to what it means (Kanwisher, 2000; Tarr & Gauthier, 2000). One interpretation is that faces appear special because they are recognized at the individual level far more often than other objects. Moreover, almost all humans have much greater experience in face recognition as compared with the recognition of individual exemplars from other categories (e.g., birds or cars). The role of these two factors in the specialization of the FFA is supported by recent neuroimaging studies. Nonface common objects elicit more activation in face-selective brain regions when recognized at the subordinate level as compared with the basic level (e.g., *robin* rather than *bird*; Gauthier et al., 1997; Gauthier,

---

Isabel Gauthier, Department of Psychology, Vanderbilt University; Michael J. Tarr, Department of Cognitive and Linguistic Sciences, Brown University.

This work was supported by National Science Foundation Award SBR-961589 and by funding from the James S. McDonnell Foundation to the Perceptual Expertise Network (PEN). For more information about PEN, see www.psy.vanderbilt.edu/faculty/gauthier/PEN. We thank Brad Duchaine and two anonymous reviewers for comments on a previous version of this article, and the members of PEN for stimulating discussions on these issues.

Correspondence concerning this article should be addressed to Isabel Gauthier, Department of Psychology, Vanderbilt University, 301 Wilson Hall, Nashville, Tennessee 37203. E-mail: isabel.gauthier@vanderbilt.edu

Skudlarski, et al., 2000; Gauthier, Tarr, Moylan, Anderson, et al., 2000). In addition, expertise with novel nonface objects (Greebles) or familiar objects (birds and cars) can lead to increased recruitment of the FFA (Gauthier, Skudlarski, et al., 2000; Gauthier et al., 1999).

In light of these neuroimaging data, face recognition can be seen as the most common case of expert subordinate-level recognition—a conclusion also reached on the basis of psychophysical studies (Diamond & Carey, 1986; Tanaka, 2001; Tanaka & Gauthier, 1997). At the same time, a great deal of effort has been devoted to describing the computational basis of this type of recognition behavior, in particular, how it differs from other types of object recognition. Some of the strongest insights come from the work of Farah and colleagues (Farah, Wilson, Drain, & Tanaka, 1998). These authors have suggested that the recognition of faces differs from that of other objects in its reliance on *holistic* processing. They hypothesize that faces are not represented as a collection of individual features or parts (or only to a lesser extent than other objects) but rather as undifferentiated wholes. According to this model, putatively holistic behavioral effects are the signature of one of two distinct visual representation systems: a *holistic system,* which is thought to be essential for face recognition and useful for object recognition, or a *part- or feature-based system,* which is thought to be essential for word recognition and useful for object recognition. An impressive amount of evidence supports the idea that face recognition is processed in a more holistic manner as compared with common object recognition (e.g., Diamond & Carey, 1986; Farah, 1996; Hole, 1994; Moscovitch et al., 1997; Rhodes, Brake, & Atkinson, 1993; Tanaka & Farah, 1993; Tanaka & Sengco, 1997; Yin, 1969; Young, Hellawell, & Hay, 1987).

Defining *Holistic*

Many authors have noted that one person’s definition of *holistic* processing may not be another’s; terms such as *configural* and *holistic* are sometimes contrasted and sometimes used synonymously, generating theoretical confusion. One reason is that some authors suggest a pluralistic picture of holistic mechanisms (Carey & Diamond, 1994; Gauthier, Williams, Tarr, & Tanaka, 1998), whereas others treat the many different holistic effects as measures of the same underlying phenomenon. (For reviews of different hypotheses regarding the holistic nature of face recognition, see Farah et al., 1998, and Moscovitch et al., 1997). For the purpose of this article, we will treat the word *holistic* as a superordinate term that cannot be attached to a single specific mechanism or representational format. The term is used in too many manners, and certainly we are unable to claim to know what it truly means or impose a new definition not already generally accepted by the field.

Because the purpose of our study was to question some assumptions about the nature of representations underlying so-called *holistic* effects, we define three terms that relate to this concept in an operational manner (using a framework consistent with Tanaka & Gauthier, 1997, and Gauthier et al., 1998). We will also highlight some differences between effects that have all been referred to as *holistic* at some point, leaving open the possibility that despite these differences, some of these effects may simply constitute different ways of measuring the same underlying process.

1. *Holistic–configural* effects arise when individual object parts are placed in the context of the other individual parts from the same object.[1] Each individual part is better recognized in the original learned configuration than in the context of these same object parts in a new configuration. For instance, the nose of a familiar face, *Bob,* may be easier to pick out in a forced-choice test when the eyes of Bob are in the original configuration than when the eyes are moved apart (e.g., Tanaka & Sengco, 1997). This effect focuses on the unique contribution of configural processing to part identification.

2. *Holistic–inclusive* effects arise when individual object parts are better recognized in the context of other individual parts from the same object and instance as compared with the context of parts from other objects and instances. Holistic–configural effects represent a sensitivity to changes in the configuration of the original parts, whereas holistic–inclusive effects reflect a sensitivity to changes in the parts themselves but not restricted to the original configuration. In other words, holistic–inclusive effects appear to reflect obligatory processing of all of the features of an object, even when instructions direct the observer to focus on only a single part. Holistic–inclusive effects are holistic in the sense that a single part cannot be processed alone, and they are not configural in that the processing of the different object parts is not sensitive to their location in the image. (It is thus analogous to an obligatory large attentional window or trying to take a picture of a small object with a wide-angle lens.)

3. *Holistic–contextual* effects arise when individual object parts are better recognized in the context of other parts than in isolation (e.g., Tanaka & Farah, 1993).[2] It may seem that holistic–contextual effects can be accounted for by a combination of holistic–configural and holistic–inclusive effects (an advantage for the correct parts in addition to an advantage of the correct configuration). However, the results of prior studies suggest that there may be a third and independent phenomenon. In particular, holistic–contextual effects appear to occur independently of an observer’s expertise, whereas holistic–configural and holistic–inclusive effects appear to increase with expertise (Gauthier & Tarr, 1997; Gauthier et al., 1998).

Although this has not been tested, it is possible that holistic–contextual effects arise simply from the presence of the correct type of parts (even if they are not from the same object) or simply from the presence of any kind of parts arranged in approximately the correct configuration for an object (e.g., as in the fruit faces of the painter Arcimboldo (1537–1593), in which one fruit may be perceived as a nose only when other fruits are positioned in the correct configuration). Although more thorough experimentation could answer these questions and other questions regarding holistic–contextual effects, this article focuses on distinguishing this phenomenon from holistic–configural and holistic–inclusive processing, both of which are of greater interest because of their hypothesized role in expertise.

---

[1] Although we use the term *parts,* we do not mean to imply that object representations are necessarily composed of parts in the most commonly used sense—3-D volumes. Rather, here *parts* refers to salient and identifiable features of the object, including many features that might not be considered parts in part-based models of object recognition.

[2] This is referred to as *holistic–relative* by Tanaka and Gauthier (1997).

The contributions of these different processes are often confounded in experimental designs. For instance, Farah et al. (1998) found that matching of face parts was influenced by the amount of similarity between the other parts (a holistic–inclusive effect). However, there was no way to know whether this effect depended on the features being in the correct configuration. (In fact, the same effect was marginally significant with inverted faces.) The results of the study we present lead to the following conclusions regarding these three different, putatively holistic effects:

First, holistic–configural effects do not necessarily reflect the use of a representation that is undifferentiated in terms of parts. Farah et al. (1998) argued that such an undifferentiated holistic representation is more important for faces than for objects as evidenced by holistic–configural effects being stronger for faces than objects. Although these authors hypothesize that faces may also be represented by their parts to some extent, the claim that we dispute is that holistic–configural effects (minimally) indicate the existence of a system representing objects as wholes undifferentiated in terms of parts. In the discussions that follow, we refer to the idea that there exists such an undifferentiated holistic representation as the *undifferentiated-template hypothesis*. The alternative would be that the same system may initially represent an object in terms of many salient parts but that this representation may be refined with further experience, with the representational units growing to include additional information (larger parts) that covaries in a manner diagnostic for the recognition goal.

Second, as with holistic–configural effects, holistic–inclusive effects increase with expertise, but they appear to reflect a different underlying mechanism. Holistic–configural and holistic–inclusive effects, as assessed behaviorally, show different learning patterns and do not correlate equally with changes in neural activation that we observe in the FFA during the acquisition of expertise.

Third, holistic–contextual effects (in contrast to holistic–configural and holistic–inclusive effects) are obtained with nonface objects, as well as with faces, and are not stronger in experts than novices. Both of these conclusions are supported by the results of previous research (Gauthier & Tarr, 1997; Gauthier et al., 1998; Tanaka & Gauthier, 1997), but this is the first time that the impact of expertise on holistic–contextual effects has been assessed using a within-subject design.

In our study, holistic–configural effects were measured by comparing part recognition in the context of an old versus new configuration of the other parts (either by reorienting the top parts of a Greeble or by misaligning the two halves of a composite Greeble). Holistic–inclusive effects were measured by comparing part recognition in the presence of other parts from the same versus a different Greeble (regardless of whether the parts were aligned). Holistic–contextual effects were measured by comparing part recognition for isolated parts versus parts in the context of other parts.

Acquisition of Expertise

If our conclusions are valid, one may ask why these three effects have often been assumed to reflect a common mechanism—one thought to be more important for faces than for objects? One possible reason has been the emphasis on examining recognition abilities in their end state—that is, after they are in place as a consequence of experience. In contrast, studying these processes during acquisition may provide a better understanding of the differences that develop between object and face recognition (i.e., between novice and expert recognition). Supporting our argument is that large differences between novice recognition and expert recognition in its final form are not all that surprising. Thus, even if these differences arose from gradual quantitative changes in a single system, they ultimately might be sufficiently dramatic so as to be mistaken for qualitative differences. A second point in favor of studying the acquisition of expertise is that the various components of expertise may also be misinterpreted as a unitary process if investigated only post acquisition (as when faces are used exclusively as stimuli). Thus, obtaining data on the onset of different aspects of perceptual expertise may reveal that although many specific behavioral or neural effects are present concurrently in experts, these effects can be dissociated at other times during the acquisition process and, consequently, in a model of face or expert object recognition.

In the present study, we used a design in which participants were trained to become perceptual experts with novel objects (Greebles). We included experimental manipulations designed to examine whether expert processing is supported by undifferentiated object representations (Farah et al., 1998). This issue was assessed in part by using a behavioral effect diagnostic of holistic–configural processing—the old–new configuration advantage. Tanaka and Sengco (1997) found that the recognition of one part of a face is impaired by a change in the configuration of the other parts of the face. Gauthier and Tarr (1997) obtained the same effect for Greebles, but only for expert participants. In both cases, the effect was obtained on all of the parts tested (e.g., moving the eyes apart had an effect on the recognition of the eyes, the nose, and the mouth). Two conclusions have been made on the basis of these results. First, the acquisition of expertise may lead experts to process nonface objects using holistic–configural processing, much as all of us appear to do with faces. Second, the object representations that mediate such effects appear to be undifferentiated templates (Farah et al., 1998).

These conclusions, however, are mitigated by a second study of Greeble expertise, in which experts showed more sensitivity to configural changes than did novices in the same paradigm, but only for one of the three parts tested—the middle part of the Greebles (Gauthier et al., 1998). The reason for our failure to completely replicate the holistic–configural effect with all three parts could be any one of many subtle differences in training or test parameters between the first and second Greeble studies. In any case, the finding of configural sensitivity on only one of the Greeble parts weakens the relationship between holistic–configural effects and an inferred undifferentiated template. Of course, one possibility is that the participants in the second Greeble study did not actually shift to configural processing for any of the parts and that the effect obtained for the middle part is spurious (although the obtained effect was in the predicted direction, and the statistics used to verify the presence of the effect took the number of different parts used in the analysis into account). Given this ambiguity, the present study explores this issue in more depth, examining the unfolding pattern of expertise acquisition and examining whether the middle part configural processing effect obtained in Gauthier et al. (1998) replicates at some point during learning. If configural effects for a single part can be replicated during the acquisition of expertise, we believe this poses a counterexample to the argument that holistic–configural effects (minimally) demonstrate the existence of an undifferentiated template.

Interrogating Object Processing During Expertise Acquisition

Given two representational systems, one undifferentiated and one part-based (Farah, 1990), it is possible that the acquisition of expertise might prompt the recognition of objects from the trained category to rely increasingly on the part-based system. Our logic is as follows: a specific pattern of acquisition for the old–new configuration advantage would be inconsistent with increasing reliance on undifferentiated object representations—that is, configural sensitivity arising for different parts according to different time courses. Such a finding would indicate that holistic–configural processing that may seem to be occurring in an undifferentiated fashion in experts is actually the result of many intermediate steps. In other words, holistic–configural processing may be part- or feature-based.

On the basis of this logic, we investigated the time course of acquisition of holistic–configural and holistic–inclusive effects, using part-matching tasks in which the to-be-ignored part was either from the original object or was not and was in the correct configuration or was not. In particular, we tested whether holistic–inclusive effects could be obtained regardless of the configuration of parts. Such a result would support the hypothesis that these two effects actually reflect distinct mechanisms.

We also included transformations that are known to influence object perception for both novices and experts and for which no qualitative change with expertise was predicted (or for which a prediction was unclear). Some of these transformations had previously been shown to elicit holistic–contextual effects, in which parts of objects are better recognized in the context of the other parts than in isolation, for both novices and experts (Gauthier & Tarr, 1997; Gauthier et al., 1998; Tanaka & Gauthier, 1997). Viewpoint changes known to affect both face (Bruce, 1982; Hill & Bruce, 1996) and object recognition (e.g., Tarr, 1995; Tarr, Williams, Hayward, & Gauthier, 1998) were used as a second type of transformation. Because we rotated stimuli around the vertical axis, we did not expect to observe any change in viewpoint dependency with the onset of expertise. (Changes that disrupt the top-down relationships between features and parts might be expected to have a stronger effect on expert rather than novice recognition, because experts typically acquire expertise for only upright versions of the objects in question; Diamond & Carey, 1986; Yin, 1969.)

Our interrogation of the process of expertise acquisition also extends previous results (Gauthier & Tarr, 1997; Gauthier et al., 1998) obtained in a naming paradigm to a more perceptually based task. In prior studies of Greeble recognition, novices and experts learned names for unfamiliar Greebles and were tested on their ability to name these particular Greebles (or match them to a prompt name) when transformed in various ways. The sensitivity to various transformations measured using this procedure could be mediated, at least in part, by naming, semantic processes, or memory processes somewhat independently of perceptual expertise. Here, we measure sensitivity to image and object transformations using a task that does not require access to long-term memory nor learning to associate specific names with specific Greebles. In each test trial, participants viewed an individual Greeble for 1,500 ms, followed by a brief pattern mask and a prompt identifying a single Greeble part or the entire Greeble (not an individual Greeble), and then a second image of a Greeble, in its original or a transformed version. The task was simply to judge whether the specified part (or the entire Greeble) specified by the prompt was the same in the two Greeble images. Because the study Greeble was presented before the cue, and trials for the different transformation types were randomized, participants could not focus on a single part of a Greeble without compromising accuracy. Our efforts here to emphasize perceptual processing during expertise acquisition parallel those of Farah et al. (1998) in their study of face recognition. Those authors used perceptually based selective attention and masking paradigms to obtain evidence suggesting that face perception is special in its reliance on holistic processing.

Five of the 10 participants for whom we collected behavioral data also participated in an fMRI longitudinal study (Gauthier et al., 1999) so that changes in the neural substrates thought to mediate perceptual expertise could be monitored. In our published results of this neuroimaging study, we found increases in activation in the right FFA of the temporal lobe with increasing expertise for upright Greebles as compared with inverted Greebles or with familiar common objects. In the General Discussion, we explore specific correlations between our behavioral results and these neurally based expertise effects. The fMRI result validates our use of the acquisition of Greeble expertise as a useful model of the development of face expertise (in addition to the behavioral effects obtained during Greeble training, which model closely analogous effects obtained with faces; Gauthier & Tarr, 1997; Gauthier et al., 1998). The relationship between this learning process and the developmental processes involved in the acquisition of the same effects for faces is clearly an analogy. However, to better understand the role of learning in the development of face expertise, one must first elucidate the components of the learning process—one goal of the present study.

Method

Participants

Ten participants from the Yale University community gave informed consent and took part in the experiment in return for pay.

Materials

The stimuli were 70 Greebles (Gauthier & Tarr, 1997; Gauthier et al., 1998; available at http://www.cog.brown.edu/~tarr/stimuli.html), photorealistically rendered 3-D objects that all share a common configuration. (Examples of Greebles are shown in Figures 2–6, which appear later in this article, and in Gauthier & Tarr, 1997, and Gauthier et al., 1998). Each Greeble consists of a vertically oriented body with four protruding appendages (from top to bottom): two *boges,* a *quiff,* and a *dunth.* Greebles can be categorized into five different families on the basis of the shape of the main body. Each individual Greeble is distinguishable from other members of its family by the shapes of its appendages. Every appendage is unique in the set, although some pairs are more similar than others. In the experiment, the Greebles were all rendered with the same gray shade, stippled texture, and overhead lighting direction. Screen images were about 6.5 cm high × 3.25 cm inclusive, and when viewed from about 60 cm, subtended approximately 6.2° × 3.1° of visual angle. Experiments were conducted on Macintosh computers equipped with color monitors (72 pixels per inch) using rapid serial visual presentation (RSVP) software (Williams & Tarr, 1998).

Procedure

Training. Participants were trained using a procedure modified from Gauthier and Tarr (1997) and Gauthier et al. (1998) to categorize 30 grayscale Greebles at the family and individual levels. Each participant was trained for approximately 7 hr, during four to five 1.5-hr sessions spread out over 2 weeks. The five family names were introduced in the first session and individual names for five Greebles were learned in each of the first four sessions. The procedure used a combination of several different tasks: When learning family or individual names for the first time, participants were presented with inspection trials in which Greebles were shown on the screen with their corresponding names, with no response required. Participants subsequently practiced pressing the response keys associated with the names of specific Greebles, while the correct name appeared on the screen in conjunction with the appropriate Greeble or with no name on the screen but feedback as to the correct response on each trial. Eventually, two tasks formed the body of the training: Blocks of a naming task were alternated with a verification task. During naming, participants saw a Greeble on the screen and had to press the first letter of its individual name. (A correct response for Greebles without an individual name at a given point in the training was to press the spacebar.) During verification, participants judged whether a label (family, individual, or *NIL*) shown for 1,000 ms matched a Greeble presented 200 ms later—regardless of the specificity of the label, the response was always *same* or *different.* When Greebles without an individual name were seen in the verification task, participants were to respond *same* if the Greeble was preceded by a *NIL* label or *different* if the Greeble was preceded by the name of another Greeble.

There were 60 trials per block for naming and 120 trials per block for verification, with two blocks of each task in Session 1 and three naming and four verification blocks in Sessions 2–4. After the end of Session 4, response times (RTs) for hits at the family and individual levels in the verification task were compared after each block. Participants were considered to have reached the criterion for expertise with Greebles when these two RT measures were no longer statistically different; at this point, training was stopped. This criterion models the phenomenon observed in real-world experts of a shift in the entry level, for instance, bird experts being as fast recognizing birds at the species level (e.g., *robin*) as at the category level (e.g., *bird*; Tanaka & Taylor, 1991).

Testing. In each testing session, participants performed a total of 288 same–different trials using a set of eight Greebles (a different set for each session). Each same–different trial consisted of the following sequence of events: A fixation cross was displayed for 500 ms, a study Greeble was displayed for 1,500 ms, a pattern mask was displayed for 200 ms, a cue (either *Greeble,* the name of a specific part, or *top*) was displayed for 100 ms, and lastly, a test Greeble was presented until the participant responded *same* or *different.* The test Greeble was shown in its canonical viewpoint (defined as the 0° view illustrated in Figure 1) or transformed in one of several ways. The different tests could thus be randomized and interleaved. If the cue was *Greeble,* the test Greeble was either in its canonical orientation, rotated in depth 25°, 50°, or 75° around its vertical axis, or

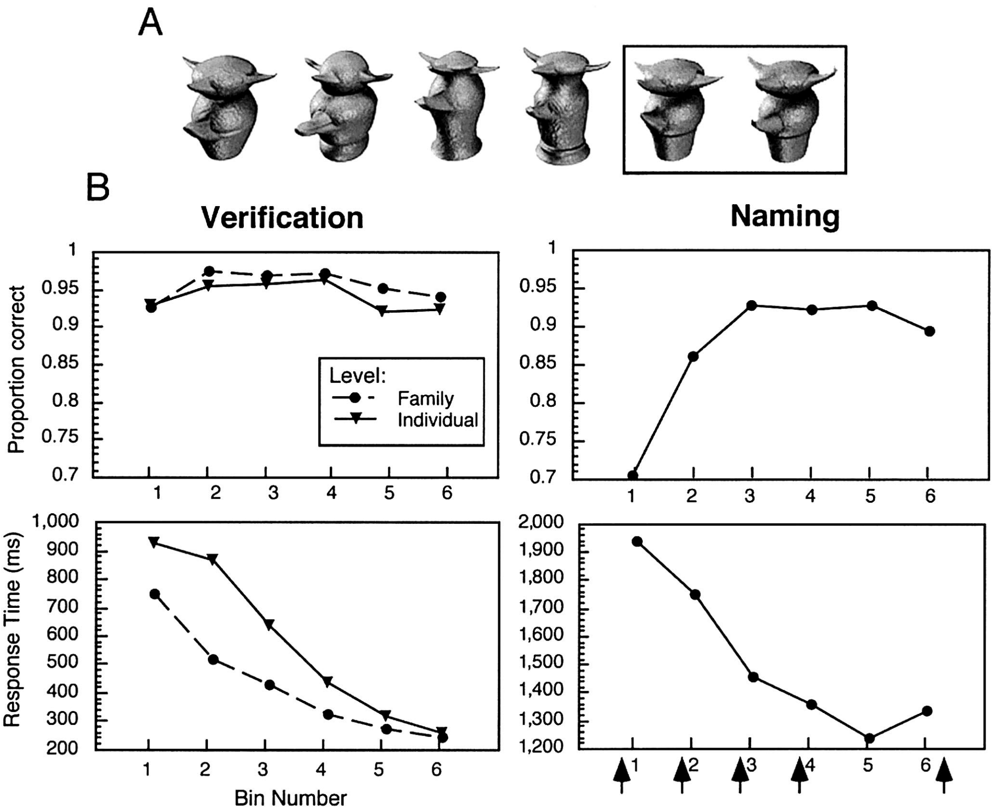

Figure 1. A: Examples of Greeble stimuli. The first five Greebles are from different families, and the two Greebles in the boxed area are two individuals from the same family. B: Performance during the naming and verification training tasks with Greebles. Arrows indicate in what bins the five testing sessions occurred.

again in its canonical orientation, but with inverted luminance.[3] If the cue was the name of a part (*boges, quiff,* or *dunth,* from the top pair to the bottom part; see Figure 1), the test image was either a Greeble in the canonical orientation, a Greeble in the canonical orientation with a transformed configuration of parts (each boge moved 15° upward), or the designated part shown in isolation. If the cue was *top,* the test Greeble was composed of two halves, split along the horizontal midline, in either an aligned or misaligned configuration. The two halves were both from the original Greeble or only the bottom half was that of a distractor Greeble from the same family, thereby creating a composite condition (Young et al., 1987). Thus, there were a total of four distinct manipulations: aligned–original, misaligned–original, aligned–composite, and misaligned–composite. Each of the aforementioned transformations is illustrated, with the results for each comparison in Figures 2–5, which appear later.

FMRI testing. Before each test session, each participant was tested in one shorter session consisting of sequential matching trials. Five of the participants were tested in an fMRI study (Gauthier et al., 1999), with the same Greeble set as in the subsequent test session, whereas the other 5 participants were tested with the same behavioral tests, but not in an fMRI context (i.e., in a standard psychophysical laboratory setting). (These tests were performed outside the scanner to ensure an equal amount of experience with Greebles in the group not participating in the fMRI study.) During these sessions, participants performed 192 sequential matching trials with pairs of upright or inverted Greebles. The number of appearances for individual Greebles was counterbalanced so that each participant received the same amount of experience with each test Greeble. Thus, we could be confident that because subsequent testing sessions used a different set of Greebles not used during training, we were not investigating the effect of practice with particular stimuli.

Statistical methods. To investigate different acquisition patterns for the effects tested here, we used the following procedure (in addition to examining the main effects and interactions in analyses of variance [ANOVAs], useful to quantify effects that do not change with training). For each effect of interest, a difference between conditions was calculated (e.g., the whole–part effect is defined by the difference of sensitivity in the whole vs. parts conditions). In the case of RTs, the ratio of the difference between conditions over the baseline condition was calculated to take into account the large overall increase in speed taking place during training. Three different contrast tests for the session factor were then evaluated: a linear contrast (-2, -1, 0, 1, 2), an expert contrast (-1, -1, -1, -1, 4) that described a sudden change occurring only after reaching expertise criterion (verification was being as fast at the individual level as at the family level), and a familiarity contrast (-4, 1, 1, 1, 1) that described a sudden change occurring immediately after the first training session. We consider both linear and expert effects as expertise effects; however, if a change occurred following the familiarity pattern, it was not interpreted as an expertise effect.

The dependent variables used throughout the study were sensitivity—as measured by d' (Green & Swets, 1966; MacMillan & Creelman, 1991)—and the geometric mean of RTs for hits. (The geometric mean is the antilog of the means of logs for each participant in each cell—a measure that is less susceptible to outliers than the arithmetic mean.) It was possible for a particular task to yield a constant difference between conditions throughout training in one dependent measure (e.g., sensitivity) and to yield an interaction between expertise and conditions in a second dependent measure (e.g., RT). In this case, we assume that the first measure reflects mostly task characteristics that do not change over the course of training, whereas the second one reflects changes due to the training (on top of any stable effects due to the task). This model is surely imperfect, but the relationships between RT and sensitivity are not completely understood. Here, we use both measures to characterize expertise acquisition, but take a conservative stance and consider possible trade-offs between them.

Results and Discussion

Training

Data from the verification and naming training tasks are shown in Figure 1. Results are grouped into bins in the following manner: Bins 1–4 included test trials in which 5, 10, 15, and 20 individual Greebles, respectively, were known and practiced. Data from all participants are included in these bins. Bin 5 includes trials of additional practice for 7 of the 10 participants, and Bin 6 includes additional trials for only 1 participant (because participants were trained to criterion starting from Session 4). Naming responses are considerably slower than verification responses: The verification task used only two responses, whereas during naming, up to 20 responses were possible.

The training results obtained here are similar to those obtained in prior studies (Gauthier & Tarr, 1997; Gauthier et al., 1998): by Bin 5, 9 out of 10 participants had reached our criterion; that is, their RTs for verification trials at the individual level were not statistically different from those at the family level.

Holistic–Contextual Processing

The whole–part paradigm has been used as a test of holistic processing in several studies (Tanaka & Farah, 1993; Tanaka et al., 1996; Tanaka & Sengco, 1997). The basic finding is that face parts are better recognized when placed in the context of a whole face than when presented in isolation. Initially, it was reported that the whole–part advantage was specific to upright faces, as it could not be found with inverted or scrambled faces, or other control stimuli such as houses (Tanaka & Farah, 1993; Tanaka & Sengco, 1997). This pattern of results suggested that face parts, unlike parts of most other objects, are encoded in a holistic fashion. Supporting this interpretation, Farah (1996) reported that prosopagnosic patient L.H. did not benefit from parts of a face being presented together, suggesting that he could no longer see faces as wholes. These and other findings led to the hypothesis that the whole–part advantage is diagnostic for the kind of holistic processing specific to upright faces and perhaps also to other cases of expert recognition.

This conclusion, however, was challenged by Gauthier and Tarr (1997), who suggested that the whole–part advantage was not specific to faces and, moreover, that it was not mediated by expertise: They obtained the effect with Greebles and found no difference between Greeble novices and experts. In a similar manner, Tanaka et al. (1996) had previously obtained evidence supporting both of these conclusions: They obtained a whole–part advantage with biological cells and cars as stimuli, with both novices and experts. Thus, the whole–part advantage would seem to reflect a (holistic–contextual) process, separable from other effects that consistently show sensitivity to expertise manipulations—in particular, the old–new configuration advantage that we describe in the next section.

Given this background, we expected to find an advantage of whole Greebles relative to isolated parts throughout our entire

---

[3] The results for the luminance transformation are not reported here because they were ambiguous. There was a numerical but nonsignificant increase in sensitivity to this transformation in the last session.

study, with little influence of expertise. For each of the three Greeble parts, there were 16 trials showing the target part in the old configuration and 16 trials showing the part in isolation. Half the trials showed a test image with the target part in the old configuration (i.e., the entire Greeble was then the same as the study Greeble), and the other half were distractor trials, in which only the designated part was replaced by a foil part from another Greeble in the same test set. (In the isolated part condition, only a foil part was presented alone on the screen.)

RTs were not considered for this test: As discussed in Gauthier and Tarr (1997), isolated parts are likely to lead to faster responses simply because of the smaller amount of visual information that must be parsed. For the dependent measure of sensitivity, our results confirm findings from previous studies in which the whole–part advantage was not different between novices and experts (Gauthier & Tarr, 1997; Gauthier et al., 1998; Tanaka et al., 1996). This is supported by an ANOVA performed with sensitivity (*d'*) as a dependent variable and transformation (old vs. isolated), part (boges, quiff, and dunth), and session (1–5) as within-subject factors. There were main effects of transformation, *F*(1, 9) = 100.8, $p < .01$ (with *d'* for old higher than that for isolation), part, *F*(2, 18) = 47.0, $p < .01$ (with dunth and boges giving higher *d'* than quiff), and session, *F*(4, 36) = 3.7, $p < .02$. There was a significant interaction of Task × Part, *F*(2, 18) = 13.3, $p < .01$. Scheffé tests revealed that although all parts showed the whole–part advantage, the effect was more dramatic with the quiff, which participants were especially poor at recognizing in isolation. Sensitivity for individual parts is shown in Table 1. Contrasts over sessions were performed on the sensitivity difference between the isolated parts and old configuration conditions (shown in Figure 2).

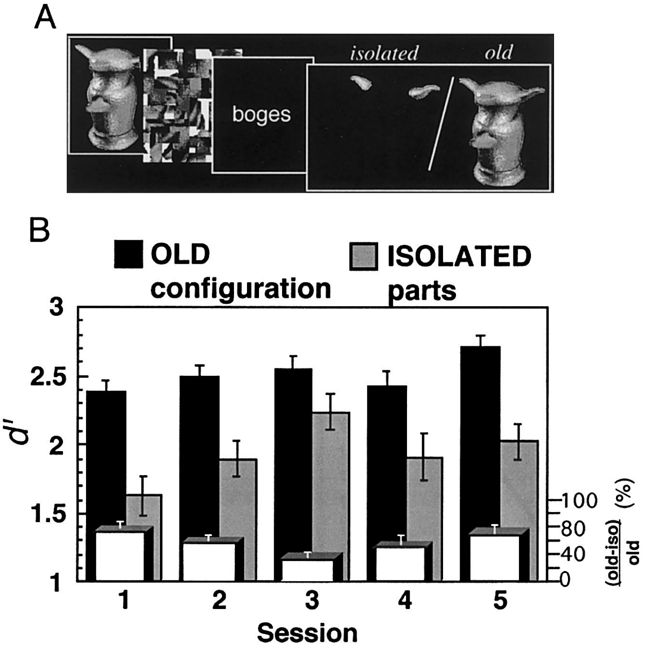

Figure 2. A: Example of a trial in the old and isolated conditions. B: Performance for Greeble part matching in the old and isolated (iso) parts conditions. Three-dimensional bars illustrate the ratio of the sensitivity difference between the conditions over sensitivity in the old configuration condition. Error bars represent the standard error of the mean for each condition.

Holistic–Configural Processing

As discussed in the introduction, holistic–configural effects for one part of an object and not another part of the same object are incompatible with the hypothesis that holistic–configural effects are mediated by an undifferentiated template. Here we investigated the learning pattern for the old–new configuration advantage to further test the assumption that holistic–configural effects reflect undifferentiated object representations.

For each of the three Greeble parts, there were 16 trials showing the part in the old configuration and 16 trials showing the part in a transformed configuration, in which the boges were each rotated 15° upward (Figure 3). Half the trials showed a test image with the same part as the study Greeble, and the other half were distractor trials, in which only the designated part was replaced by a foil part from another object from the current test set.

Table 1
Sensitivity for Part Recognition in Isolation or in the Old Configuration Throughout Training

	Boges		Quiff		Dunth	
Session	Old	Isolated	Old	Isolated	Old	Isolated
1	2.62	2.07	2.10	1.10	2.40	1.70
2	2.59	2.18	2.20	1.47	2.69	2.05
3	2.65	2.59	2.36	1.64	2.65	2.49
4	2.81	2.33	2.21	1.01	2.25	2.39
5	2.55	2.43	2.73	1.48	2.86	2.19

ANOVAs on *d'* and RTs for hits as dependent variables were performed with transformation (old vs. new), part (boges, quiff, and dunth), and session (1–5) as within-subject factors. The results for 1 participant were dropped from the ANOVA on RTs because there were no hits in three cells of the design.

The ANOVAs on *d'* revealed main effects of transformation, *F*(1, 9) = 58.3, $p < .01$ (with old configurations yielding higher *d'* than new configurations), and session, *F*(4, 36) = 3.3, $p < .02$. There was a significant interaction of Transformation × Part, *F*(2, 18) = 8.6, $p < .01$. Scheffé tests revealed that boges showed the strongest transformation effect (Figure 4), most probably because they were the moved part in the new configuration condition.

The contrasts on the sensitivity difference revealed a significant linear effect for the dunth, *F*(1, 9) = 8.7, $p = .02$. A near-significant linear effect was also obtained for the boges, *F*(1, 9) = 4.7, $p = .06$. However, these two effects were opposite in that the boges showed an increase in the old–new advantage with training, whereas the dunth showed a decrease in the effect. This is supported by an ANOVA on sensitivity difference for these two parts, which revealed a significant interaction of session with part, *F*(4, 9) = 5.5, $p < .01$. None of the contrasts on sensitivity were significant for the quiff.

An ANOVA on RTs revealed main effects of session, *F*(4, 32) = 33.0, $p < .01$, and transformation, *F*(1, 8) = 42.0, $p < .01$ (with old being faster than new). This last effect was much stronger with boges than with the other parts, as revealed in a

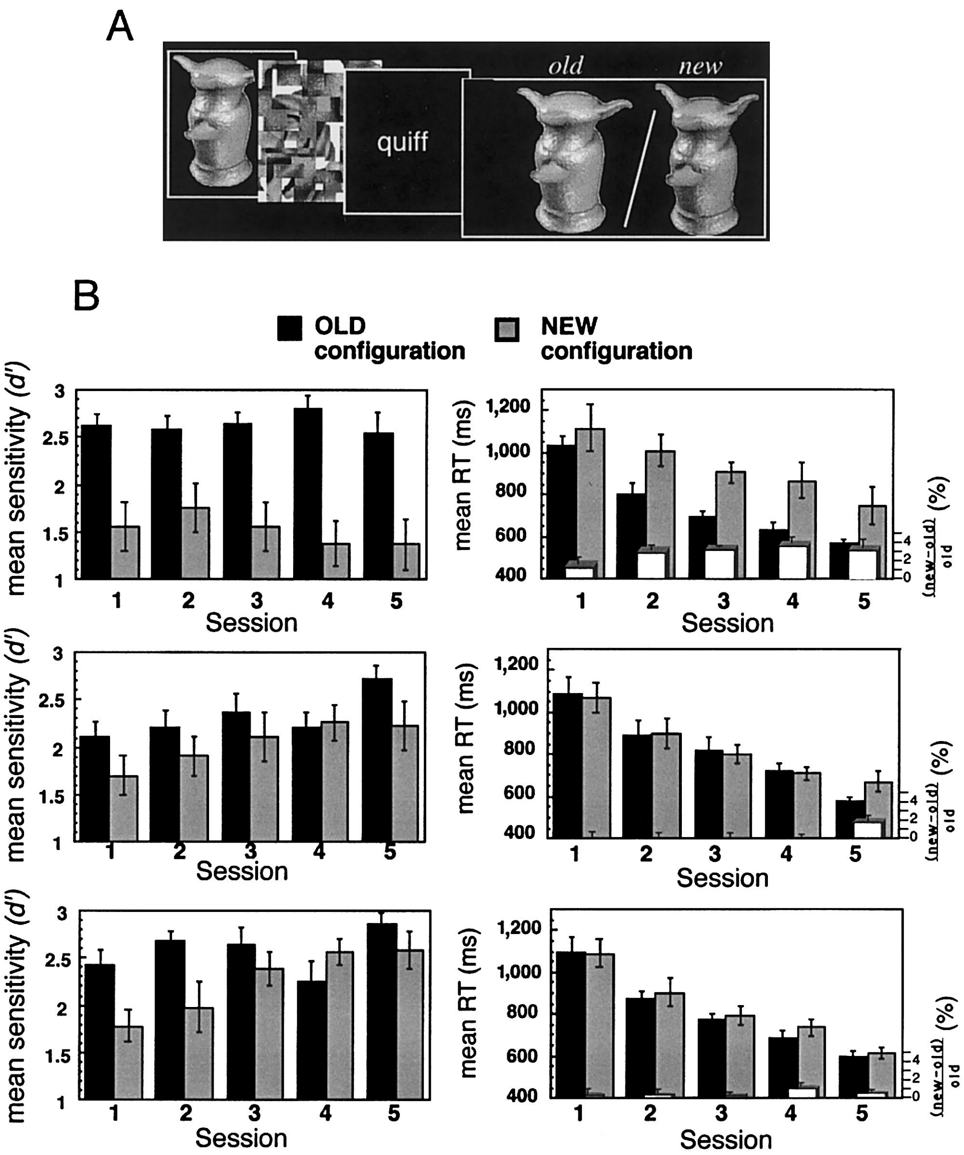

Figure 3. A: Example of a trial in the old and new conditions. B: Performance for Greeble part matching in the old and new configuration conditions. Top row of graphs: Boges (top parts). Middle row of graphs: Quiff (middle part). Bottom row of graphs: Dunth (lower part). Three-dimensional bars in the response time (RT) graphs illustrate the ratio of the RT difference between conditions over the RT in the old condition. Error bars represent the standard error of the mean for each condition.

significant Transformation × Part interaction, $F(2,16) = 16.9$, $p < .01$, and subsequent post hoc tests. The dramatic improvement in speed of response during training impeded the detection of an increasing difference between the two conditions. For this reason, we ran contrast tests on the ratio of the difference between conditions over RTs in the old condition [(new − old)/old], as illustrated in Figure 4. The only significant contrast in the RT ratio was an expert effect for the quiff, $F(1, 9) = 6.6$, $p < .05$. This effect did not approach significance for the other parts ($p > .90$). The interaction of Part × Session was not significant.

The present results replicate those of Gauthier et al. (1998) in that the old–new configuration advantage shows an expertise effect for only the middle part (quiff). The quiff may be particularly sensitive to the old–new configuration advantage as implemented here for several reasons: (a) its central location on the Greebles, (b) its proximity to the parts that are moved, and (c) during testing, a larger number of trials relied on judgments for the upper than for the lower part of the Greebles (because of the composite trials), which may have encouraged participants to focus on this region. Our results also indicate that the acquisition pattern of the old–new configuration advantage can vary for different parts of the same object: In this study, the advantage was robust throughout training for boges, it showed an abrupt increase after passing the expertise criterion for the quiff, and it showed a gradual decrease with time

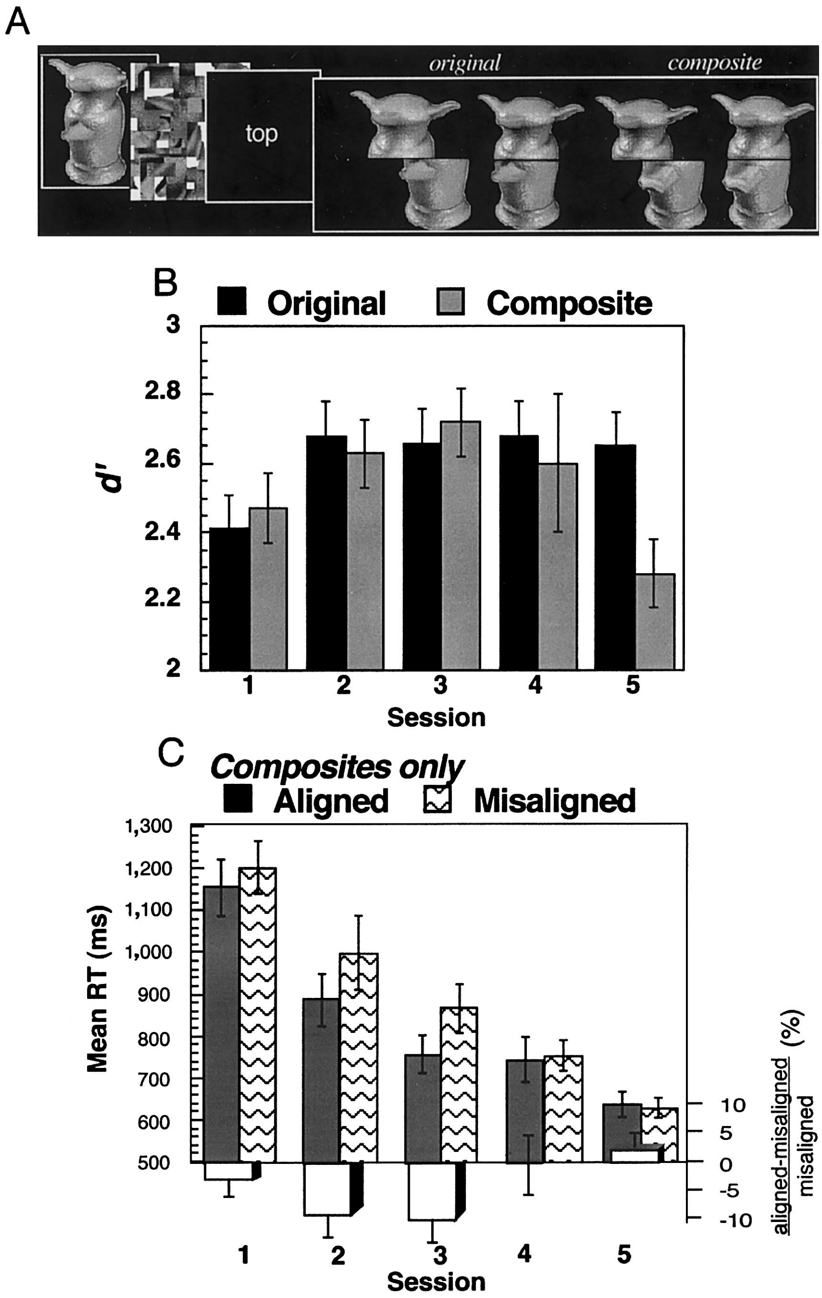

Figure 4. A: Example of a trial in the original and composite conditions, with aligned and misaligned Greebles. B: Sensitivity for the matching of Greeble top halves in the original and composite conditions, collapsed across alignment. C: Response times (RTs) in the aligned and misaligned conditions in only the composite condition. Three-dimensional bars illustrate the ratio of the RT difference between conditions over the RT in the misaligned condition. Error bars represent the standard error of the mean for each condition.

for the dunth. In Gauthier and Tarr (1997), Greeble experts showed a substantial holistic–configural effect on the dunth with training that was somewhat longer on average. Thus, although holistic–configural effects on the dunth can be obtained, they may take longer to arise because the part is the farthest from the transformed boges. Finally, it is not very surprising that the boges showed a very large sensitivity to their own transformation throughout training. Although the shape of the boges per se was relatively unaltered by the transformation, their orientation was changed, which is known to impact performance for at least entire objects (Tarr et al., 1998).

The replication of a quiff-only expertise effect as well as the learning differences between parts are relevant to Farah and colleagues' (Farah et al., 1998) hypothesis. Again, these authors suggested that the old–new configuration advantage for faces reflects a reliance on undifferentiated wholes. Our results reveal that it is possible for only some of the parts of an object to develop strong configural binding (e.g., it is possible for the recognition of the quiff to depend on the configuration of the boges). Thus, even when significant holistic–configural effects are obtained for all parts tested (Gauthier & Tarr, 1997; Tanaka & Sengco, 1997), this may not necessarily indicate the existence of an undifferentiated object representation.

Treating the isolated parts condition as a baseline, we found some evidence that the presence of a Greeble, regardless of the configuration of parts, facilitates part recognition. (Participants performed worst with isolated parts and best in the old configuration condition.) However, it is not clear that the isolated parts condition should be considered a baseline in this context, in that it constitutes an abnormal presentation of a Greeble part. Therefore, it is also possible that the difference in conditions is due to impaired, rather than facilitated, processing.

Composites

Young et al. (1987) reported that the recognition of a composite face's top or bottom half was slower when the other half (from a different individual) was placed in an aligned as opposed to misaligned configuration. This effect was thought to reflect the observers' bias to process faces holistically when face parts were in a correct configuration (e.g., aligned, leading to a salient percept of an entirely new person).

The difficulty of recognizing an individual part in the presence of a distractor part has been investigated in many ways. Young et al. (1987) used misaligned versions as their control condition for aligned composites of famous faces. Carey and Diamond (1994) used a similar paradigm in children, using classmates' faces. In an attempt to extend the phenomenon to unfamiliar faces, Hole (1994) used a simultaneous matching paradigm, with only aligned composites in upright and inverted orientations. The composite effect has not been studied with many categories of objects, which makes it difficult to evaluate its specificity for faces (however, see Gauthier et al., 1998). It is also difficult to make predictions on the basis of the old–new configuration advantage. Although in both paradigms a single part must be recognized in the presence of a configural change, the old–new advantage effect reflects the interference from a configural change of an original part, whereas the composite effect appears to measure the release from the interference of a distractor part when the canonical correct configuration is disrupted.

One issue yet to be resolved is whether manipulations such as inversion or misalignment remove all interference from the distractor part. The problem is that composite tests typically do not include a comparison in which both halves are from the original face. Thus, whether there is any interference at all from the presence of a distractor part in an unusual configuration is not typically considered. Gauthier et al. (1998) introduced an original condition in which the two halves of a Greeble came from the original object and were either aligned or misaligned. An effect of alignment was found, as well as an advantage of original versions versus composites, regardless of alignment. However, only expert participants were tested, so the role of expertise per se could not be assessed.

In this study, we first considered possible changes with expertise for the effect of the presence of a distractor versus an original part in the image (a composite–original effect) and whether this depended on configuration. This effect measures holistic–inclusive processing (i.e., all the parts of an object, regardless of their configuration, influence part recognition). On the basis of prior results (Gauthier et al., 1998), our prediction was that with expertise, the top of a Greeble will be better recognized when the bottom part comes from the same individual, regardless of alignment. Second, we assessed whether with increasing expertise, interference from a distractor part becomes more dependent on the configural relations between the distractor and target parts (a composite alignment effect, or what is generally called the *composite effect*). In this case, prior work (Gauthier et al., 1998) found that experts recognized misaligned halves from different Greebles faster than aligned halves. We assumed that this effect measures holistic–configural processing (i.e., the influence of the distractor part on the recognition of the target part depends on their configuration). On the basis of the results presented in the previous section and prior studies, we expected that this effect would also increase with expertise.

There were 16 trials in each of the four conditions (aligned–original, misaligned–original, aligned–composite, and misaligned–composite). Half the trials presented a top half that matched the target and the other half of the trials presented the top half of a distractor.

ANOVAs on d' and RTs for hits were performed with configuration (aligned vs. misaligned), transformation (original vs. composite), and session (1–5) as within-subject factors. None of the effects were significant for sensitivity (all ps $>.14$). The ANOVA for RTs revealed only a significant main effect of session, $F(4, 36) = 38.2$, $p < .01$.

To investigate possible learning changes for a composite–original effect (reflecting holistic–inclusive processing), we collapsed the results across configuration (Figure 4) and performed contrasts on the sensitivity difference between composites and originals, as well as on the ratio of the difference in RTs for composites and originals over RTs for originals. Only the expert contrast for the sensitivity difference came close to significance, $F(1, 9) = 3.5$, $p = .09$. The effect of condition was significant in Session 5, $F(1, 9) = 6.4$, $p = .03$, but not in any of the other sessions (smallest p value was greater than .50). There was no significant effect of alignment, nor any interactions with this factor (Fs < 1). Thus, only experts showed impaired recognition of a Greeble's top half when the bottom half was from a distractor Greeble (relative to when both parts were from the same Greeble). Gauthier et al. (1998) also found a large (greater than 10% accuracy) and significant composite–original effect in Greeble experts. Combined with our current results, there is good support for the idea that this effect is found in experts and not in novices. In both studies, the effect was found for both aligned and misaligned Greebles (see Table 2 and Gauthier et al., 1998, Figure 11), suggesting that holistic–inclusive processing is somewhat independent from configural effects. Gauthier et al. obtained this composite–original effect with Greeble experts using familiar Greebles. (Novices were not tested in that study.) However, the present results are the first direct evidence that this effect depends on expertise.

The composite alignment effect, a measure of holistic–configural processing, was investigated using the results for only

Table 2
Sensitivity (d′) and Geometric Mean Response Times (in Milliseconds) Throughout Training For Part Recognition, as a Function of Alignment and Pairing of the Target and Distractor Part

	Aligned		Misaligned	
Session	Composites	Originals	Composites	Originals
	Sensitivity			
1	2.47 (0.18)	2.35 (0.23)	2.47 (0.09)	2.47 (0.15)
2	2.71 (0.14)	2.65 (0.17)	2.55 (0.14)	2.71 (0.16)
3	2.55 (0.18)	2.57 (0.14)	2.89 (0.12)	2.74 (0.14)
4	2.64 (0.25)	2.65 (0.16)	2.56 (0.19)	2.71 (0.17)
5	2.34 (0.19)	2.65 (0.17)	2.21 (0.21)	2.65 (0.20)
	Geometric mean response time			
1	1,164 (68)	1,158 (61)	1,206 (63)	1,171 (82)
2	896 (64)	906 (65)	1,003 (88)	933 (63)
3	766 (46)	764 (45)	871 (59)	825 (66)
4	753 (53)	750 (31)	756 (37)	762 (40)
5	643 (30)	644 (39)	631 (23)	664 (24)

Note. Standard errors are in parentheses.

composite Greebles. An ANOVA on sensitivity revealed no significant effect, but the same analysis on RTs yielded a significant effect of session, $F(4, 36) = 30.2$, $p < .01$, a marginal effect of configuration, $F(1, 9) = 4.9$, $p = .05$, and a marginally significant Session × Configuration interaction, $F(4, 36) = 2.4$, $p = .07$. To investigate possible learning changes for this effect, we performed contrasts on the ratio of the difference in RTs for aligned and misaligned composites over RTs for misaligned composites (Figure 4). The composite alignment effect has been previously reported in RTs (Gauthier et al., 1998; Hole, 1994; Young et al., 1987), whereas the composite–original effect has been obtained in sensitivity here and in Gauthier et al. (1998). The linear contrast was significant, $F(1, 9) = 8.8$, $p < .02$. In addition, there was an unexpected significant quadratic effect, $F(1, 9) = 6.0$, $p < .05$. Sensitivity for composite versions, regardless of alignment, also shows a U-shaped pattern. One hypothesis is that this U-shaped effect reveals a common process underlying the whole–part advantage and the composite–original effect. (In both cases, matching of Greeble parts out of their original context showed a significant quadratic effect.) The linear change in holistic–configural processing as measured by RTs reflects an effect size of $r_{\text{contrast}} = .70$, conventionally considered a large effect (Rosenthal, Rosnow, & Rubin, 2000).[4] This is more relevant to the hypothesis that holistic–configural processing can be recruited by expertise than is the effect size in any single session. In addition, it may not be meaningful to compare effect sizes across the tasks used here and those used with faces in prior studies. The mean results for sensitivity and RTs for matching the top of Greebles as a function of the identity and alignment of the bottom part are shown in Table 2.

One prior study considered the development of the composite effect. Carey and Diamond (1994) tested children of ages 6 and 10 years, as well as adults, in a composite task with upright and inverted stimuli. They found that participants of all ages demonstrated a face composite effect with little change in development, whereas the inversion effect increased monotonically with age. They concluded that the inversion effect—found before to increase with expertise (Diamond & Carey, 1986)—may not result from an increased reliance on holistic encoding (as measured by the classical composite effect). A correlate is that their results provide no evidence for an increase in holistic encoding with age and/or expertise.

By age 6, the processes used by children to perceive and recognize faces may have already been altered in a significant manner in response to experience. Thus, the Diamond and Carey (1986) study would be blind to earlier occurring developmental effects with regard to holistic encoding. In contrast, we were able to test our participants when they were complete novices with Greebles. Our results suggest that expertise indeed leads to a change in holistic–configural processing: We found an increase with expertise in the composite alignment effect, as defined by faster recognition of misaligned than aligned composites.

Viewpoint Effects

For each of the eight Greebles used in each testing session, there were 32 trials showing the same Greeble as the second stimulus at each of four orientations (0°, 25°, 50°, and 75°) and 32 trials showing a distractor Greeble in the same four orientations.

As shown in Figure 5, although Greeble matching became faster with training, it remained strongly viewpoint dependent throughout the experiment, with little change in the slope for RTs as a function of viewpoint. ANOVAs on *d’* and RTs for hits as dependent variables were performed with orientation (0°, 25°, 50°, and 75°) and session (1–5) as within-subject factors. We found a main effect of orientation (but no effect of session) on *d’*, $F(3, 27) = 22.2$, $p < .01$; participants performed better with Greebles in the canonical 0° orientation than in all other orientations. For RTs, there were main effects of orientation, $F(3, 27) = 28.8$, $p < .01$, and session, $F(4, 36) = 18.5$, $p < .01$, with no interaction.

To test for training effects, contrasts were performed on both the RT difference between 0° and 75° and the ratio of this difference over RTs for 0°. None of the contrast tests were significant (all $ps > .15$). The absence of a reduction in the viewpoint effect may appear surprising, especially because such effects have been found to diminish with practice in other studies (Jolicoeur & Milliken, 1989; Tarr & Pinker, 1989), even when participants are tested with unfamiliar exemplars that are visually similar to those that were learned (Tarr & Gauthier, 1998). However, the absence of an interaction between orientation and expertise is less surprising once we consider that if expertise reduced viewpoint effects, face recognition would not be expected to show the viewpoint effects obtained in several studies (Bruce, 1982; Hill & Bruce, 1996; Hill, Schyns, & Akamatsu, 1997). Whereas the improvement on Greeble matching appears to be view invariant (i.e., the trained view does not seem to benefit more from experience than other views), Greeble matching itself seems to rely on viewpoint-dependent representations throughout training. In addition, this does not reveal whether other changes (e.g., any potential holistic effects) that occur with experience would also become viewpoint

---

[4] In comparison, r_{contrast} was equal to .53 for the expert contrast in the composite–original effect.

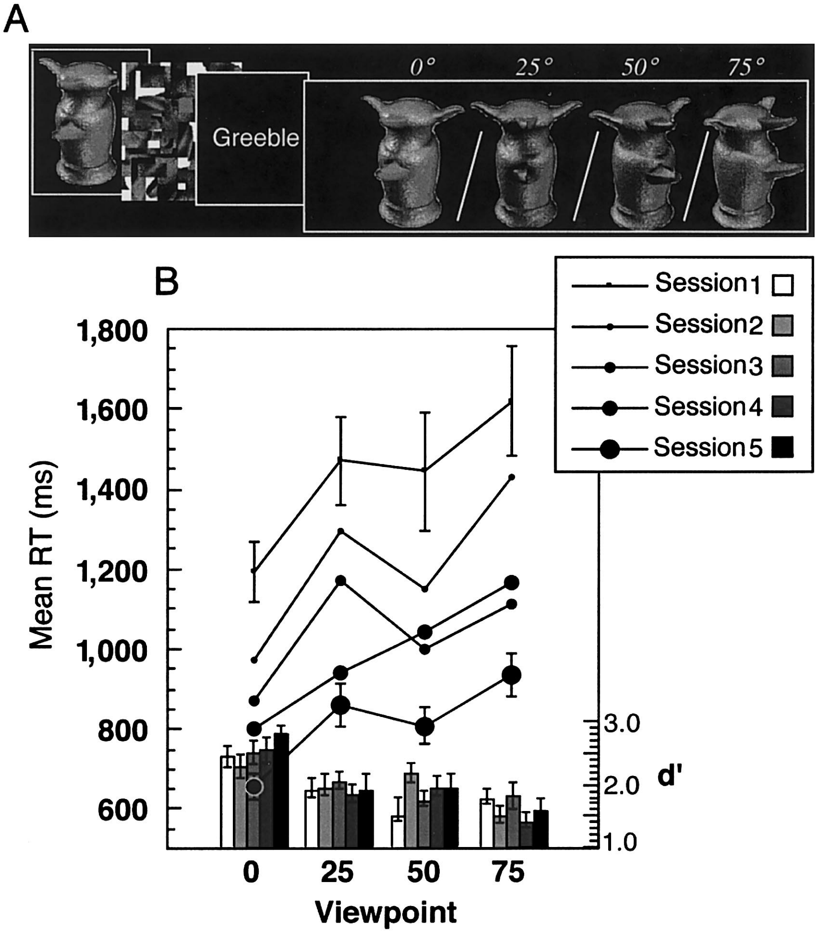

Figure 5. A: Example of a trial in the original viewpoint and rotated conditions. B: Mean response times (RTs), shown as lines, and sensitivity, shown as bar graphs, for Greeble matching with the second stimulus in various viewpoints. Error bars represent the standard error of the mean for each condition.

invariant. Finally, the limited impact of expertise on viewpoint changes that do not affect the top-bottom relations between parts (such as a rotation in depth around the vertical axis) can be contrasted with changes in orientation. For example, the effects of inversion in the picture plane are strongly modulated by expertise level (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Gauthier et al., 1998).

Correlations With Activation in the FFA

Multiple sessions of blood oxygen level dependent (BOLD) measurements using fMRI during tasks with upright and inverted Greebles and faces were obtained for 5 of the 10 participants. The method and results of this fMRI study have been published separately (Gauthier et al., 1999). They reveal that the middle FFA, predominantly in the right hemisphere, is recruited by expertise with upright Greebles. In this section, we explore correlations between the signal changes in this area and the behavioral measures discussed in the present article. Although our sample size is relatively small (5 participants tested in 5 sessions), we think such an analysis offers a good exploratory value. Any functional brain imaging study is very limited in the number of possible experimental manipulations and it may gain significant explanatory power from correlations with multiple tasks performed outside the scanner.

Details of the fMRI experiments and analyses can be found in Gauthier et al. (1999): Participants were scanned once before they had any experience with the Greebles, three times during the training procedure and once after they passed the criterion for expertise. The task was sequential matching (same–different) with Greebles and faces, upright and inverted. A new set of eight Greebles and a new set of eight faces were used in each fMRI session. (Each set was then used in the following behavioral session.) A region of interest (10.4 mm [y] × 13.6 mm [x]; see Gauthier et al., 1999) within the ventral temporal cortex was centered on the area that was more active for upright than for inverted faces in each session, for each participant. The Talairach coordinates for this region were almost identical to those reported in other studies for the FFA (Kanwisher et al., 1996, Kanwisher et al., 1998). Independent localizer tests consisting of passive viewing of faces versus familiar common objects were also used to identify the location of the FFA in individual participants. (The FFA was essentially identical whether defined through the orientation or category effect.)

The hypothesis considered here was that there should be a strong correlation during training between the activation obtained in the right FFA and a behavioral effect that reflects computational aspects of the process mediating neural changes occurring during expertise acquisition. Moreover, this correlation should be present for upright but not for inverted Greebles. Although any such correlation is intriguing, recognition behavior of any sort is almost certainly not the product of only a small number of brain voxels—we strongly contend that recognition behavior depends on the activity across many regions of the visual system.

We first computed the correlation between the percentage of signal change for upright minus inverted Greebles, in the right FFA, with three different indices corresponding to only the behavioral expertise effects obtained in this study, using each session of each participant as an independent data point. The behavioral indices used for this correlation analysis are those showing a significant learning or expertise effect: (a) the composite–original effect (sensitivity), (b) the composite alignment effect (speed ratio), and (c) the old–new configuration advantage for the quiff (speed ratio). The first measure should reflect holistic–inclusive processing, whereas the latter two should reflect holistic–configural processing.

The only behavioral measure that showed a significant correlation with the inversion effect for Greebles across sessions was the composite–original effect, $r^2 = .19$, $p = .03$. For the other two measures, r^2 values were minuscule (0 and .01) and did not reach statistical significance (both ps $> .50$). To ascertain that useful information is gained by using each individual's behavioral data as opposed to a theoretical learning function to predict fMRI activity, we also considered the correlation of the fMRI results (for activation in the upright inverted Greebles conditions) with linear (-2, -1, 0, 1, 2) and expert (-1, -1, -1, -1, 4) contrast weights for each session. The correlation with these theoretical functions were small and nonsignificant (both r^2s $< .02$; both ps $> .50$), indicating a clear advantage of using each individual's behavioral data on a test of holistic–inclusive processing in predicting activity in the brain during Greeble expertise acquisition. This correlation is even more impressive given that the behavioral and fMRI tasks are quite different.

As shown in Figure 6, the composite–original effect correlated specifically with the activation obtained for upright Greebles minus fixation ($r^2 = .22$) and showed no relation with the activation for Greebles inverted or faces at both orientations (each relative to fixation). None of the four behavioral measures showed a significant correlation with the activation obtained for Greebles in the left FFA, consistent with the suggestion that Greeble expertise seems to be mainly a right hemisphere process (Gauthier et al., 1999).

General Discussion

In this article, we distinguish holistic–configural and holistic–inclusive processing as two independent effects of expertise acquisition, both different from holistic–contextual processing, which does not depend on expertise. To summarize, we have found evidence that:

1. The type of holistic processing measured by the whole–part advantage (holistic–contextual processing) is neither specific to faces nor depends on expertise level. We replicated findings by Gauthier and Tarr (1997) and Gauthier et al. (1998) that holistic–contextual processing can be obtained with Greebles even in novice participants and that the whole–part advantage is not larger for experts than for novices. This is the first time that this is demonstrated in a within-subject design. Tanaka and Gauthier (1997) also reported on an unpublished study by Tanaka and colleagues (Tanaka et al., 1996), in which the whole–part advantage was obtained for images of cells and cars in novice participants.

The only category that shows a robust absence of holistic–contextual processing appears to be inverted faces (Tanaka & Farah, 1993; Tanaka & Sengco, 1997). To the extent that processing of inverted faces differs qualitatively from that of nonface objects, it is unlikely to constitute novice-level processing. The absence of an holistic–contextual effect with inverted faces possibly could be associated with long-term expertise within a category, such as all humans have with upright faces. For example, highly automatic holistic–inclusive and holistic–configural strategies associated with faces may be invoked for inverted faces (but not used efficiently) and lead to suboptimal use of both holistic and parts information.

2. Neither holistic–configural nor holistic–inclusive processing is specific to faces (see also Diamond & Carey, 1986; Gauthier & Tarr, 1997; Gauthier et al., 1998), as both may be recruited during the acquisition of expertise with novel objects. This study confirmed prior findings that the composite–original effect and the old–new configural effect could be obtained with nonface objects and showed a significant increase with expertise (Gauthier & Tarr, 1997; Gauthier et al., 1998). This is the first time this is demonstrated in a within-subject design.

Although the magnitude of these effects in Greeble experts with a few hours of training may be less than that which can be found in face experts, the effect size of the expertise effect for these phenomena was quite large (for only the quiff expert contrast, $r_{\text{contrast}} = .65$; for the linear contrast on the composite alignment effect, $r_{\text{contrast}} = .70$; for the expert contrast on the composite–original effect, $r_{\text{contrast}} = .53$).[5] Thus, both holistic–configural and holistic–inclusive processing appear to be very strongly related to amount of expertise. Unless it can be demonstrated that larger effects with faces than with objects are not due to greater expertise with faces, it becomes very difficult to argue that this state of affairs reflects category-specific processing over and above the difference in expertise.

3. Holistic–inclusive and holistic–configural processing can be distinguished in that they can show different acquisition patterns and are possibly supported by different neural substrates within the visual system. We observed that holistic–inclusive and holistic–configural effects reflect different processes in that only the composite–original effect was found to be correlated with changes taking place during the expertise training in the right FFA. These results should be interpreted cautiously as they are based on a relatively small number of participants and the specific conditions of this study. However, our findings suggest that functional neuroimaging techniques can be integrated with behavioral measures and, as such, may offer a powerful means of investigating the

---

[5] An r value of .50 is equivalent to Cohen's d of 1.15 and is considered a large effect (Rosenthal et al., 2000).

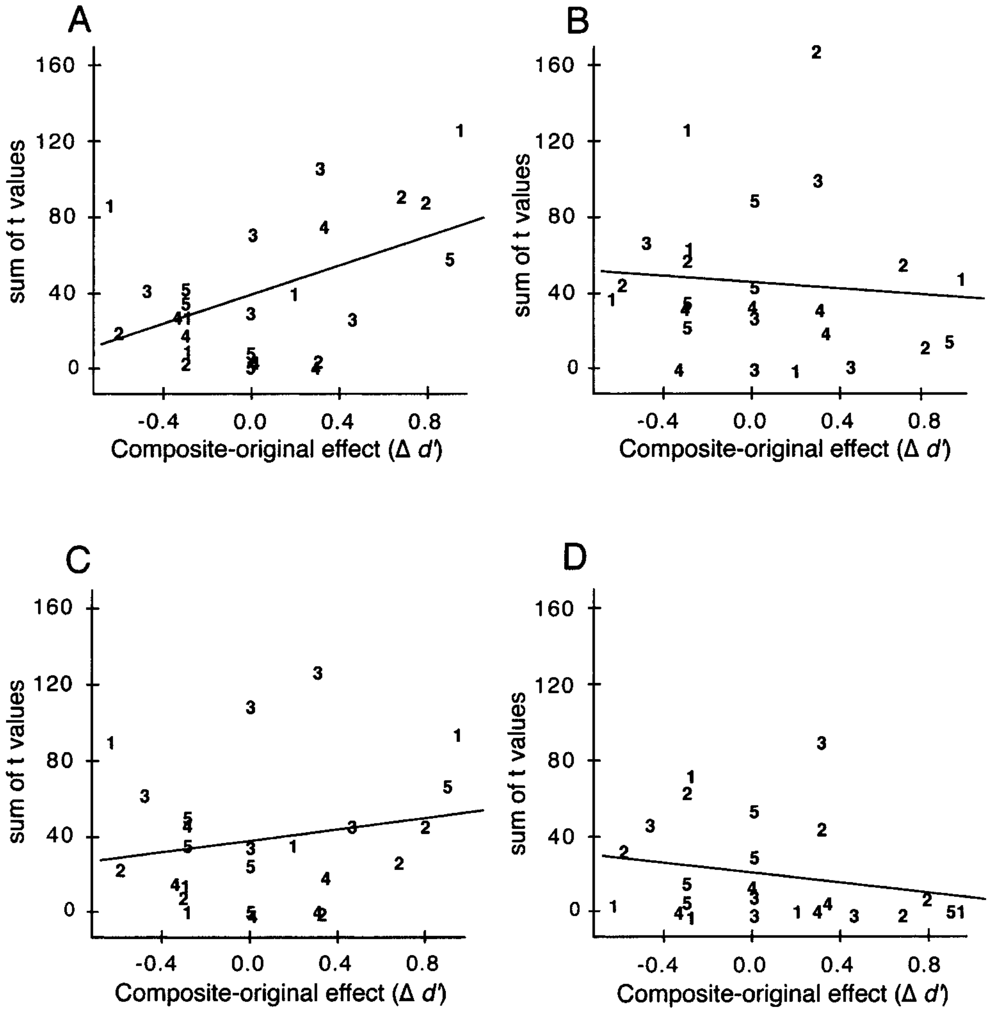

Figure 6. Correlations of the activation in the right middle fusiform face area during tasks with Greebles and faces in upright and inverted orientations, with the composite–original effect. The different numbers indicate the 5 different participants, with 5 sessions per participant. A: Correlation of brain activity for Greebles upright minus fixation with the composite–original effect (R^2 = .22, *p* = .02). B: Correlation of brain activity for faces upright minus fixation with the composite–original effect (R^2 = .09, *p* = .65). C: Correlation of brain activity for Greebles inverted minus fixation with the composite–original effect (R^2 = .03, *p* = .38). D: Correlation of brain activity for faces inverted minus fixation with the composite–original effect (R^2 = .05, *p* = .29).

computational bases of changes taking place in the visual cortex with learning.

Perhaps the clearest evidence that these two effects are independent is that the sudden increase in sensitivity to the identity of the distractor part in the composite test was obtained regardless of whether Greeble halves were aligned (see Table 2). This new distinction for two processes that both increase with expertise suggests a plurality of changing mechanisms that cannot be easily captured by a switch between a part-based and a whole-based representational system.

4. Holistic–configural processing is not supported by undifferentiated object representations, rather part-based and whole-based effects may stem from a common flexible representational system. Proponents of what we call the undifferentiated-template hypothesis of face processing (Farah et al., 1998) do not claim that holistic processing is the only manner in which faces are recognized. However, they do argue for the existence of a holistic representation undifferentiated in terms of parts for faces on the basic of holistic–configural effects.

Our findings on the acquisition of such effects for different parts of the same object provide an empirical refute of this logic. The amount of interference from a configural change in one part was not the same for all other Greeble parts and did not follow the same learning pattern for different parts. It is clear that an old–new configuration advantage cannot be a proof of existence for a representation that is undifferentiated in terms of parts if at times it can be obtained for only some parts of an object. This illustrates a specific advantage in studying the acquisition over time of configural and holistic effects in an experimental situation: In long-term expertise, such as is found with faces, all parts may appear to be processed in the same manner, suggesting an undifferentiated representation. However, this may simply be because the separate interactions between different parts all had the time to grow strong enough. In general, our results point to the possibility that the shift from part-based to object-based (or from the types of effects associated with objects vs. faces) may be a gradual quantitative change rather than a sudden qualitative shift between two different modes of processing. This is consistent with Leder and colleagues's (Leder & Bruce, 2000; Leder, Candrian, Huber, & Bruce, 2001) work on the face inversion effect that suggests configural effects are due to local processing of the relations between facial features.

5. Greeble matching was supported by viewpoint-dependent representations in both novices and experts, suggestive of a common underlying representational format that may support both part-based and whole-based effects (see Tarr, in press).

Together these results question conclusions made on the basis of studies overemphasizing one extreme case of holistic–configural and holistic–inclusive processing—that of face recognition. Relying only on extremes can make the end states of domain-general acquisition processes appear to be domain-specific and modular. We suggest that both face and object processing can be studied best under conditions in which we can witness the visual system adapting over time to the demands of complex recognition tasks.

References

Bruce, V. (1982). Changing faces: Visual and nonvisual coding in face recognition. *British Journal of Psychology, 73,* 105–116.

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

Chao, L. L., Martin, A., & Haxby, J. V. (1999). Are face-responsive regions selective only for faces? *NeuroReport, 10,* 2945–2950.

De Renzi, E. (1986). Current issues on prosopagnosia. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 243–252). Dordrecht, the Netherlands: Nijhoff.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117.

Farah, M. J. (1990). Visual agnosia: Disorders of object recognition and what they tell us about normal vision. Cambridge, MA: MIT Press.

Farah, M. J. (1996). Is face recognition "special"? Evidence from neuropsychology. *Behavioural Brain Research, 76,* 181–189.

Farah, M. J., Levinson, K. L., & Klein, K. (1995). Face perception and within-category discrimination in prosopagnosia, *Neuropsychologia, 33,* 661–674.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. W. (1998). What is "special" about face perception? *Psychological Review, 105,* 482–498.

Galper, R. E. (1970). Recognition of faces in the photographic negative. *Psychonomic Science, 19,* 207–208.

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

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

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

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

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., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). Does visual subordinate-level categorisation engage the functionally-defined fusiform face area? In N. Kanwisher & M. Moscovitch (Eds.), *Journal of Cognitive Neuropsychology: Special Issue on Face Recognition, 17,* 143–163.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12,* 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,* 2401–2428.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Haxby, J. V., Horwitz, B., Ungerleiderm, L. B., Maisog., J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: A PET-RCBF study of selective attention to faces and locations. *Journal of Neuroscience, 14,* 6336–6353.

Hill, H., & Bruce, V. (1996). The effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 986–1004.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62,* 201–222.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23,* 65–74.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, USA, 96,* 9379–9384.

Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 200–210.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3,* 759–763.

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research, 5,* 55–67.

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.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effects of face inversion on the human fusiform face area. *Cognition, 68,* 1–11.

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

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30,* 73–83.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge, England: Cambridge University Press.

McNeil, J. E., & Warrington, E. K. (1993). Prosopagnosia: A face-specific disorder. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 46*(A), 1–10.

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

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology, 74,*1192–1199.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47,* 25–57.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research.* New York: Cambridge University Press.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. *Brain, 115,* 15–36.

Sergent, J., & Signoret, J. L. (1992). Functional and anatomical decomposition of face processing: Evidence from prosopagnosia and PET of normal subjects. *Philosophical Transactions of the Royal Society of London, Series B,* 335, 55–56.

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

Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. In R. L. Goldstone, D. L. Medin, & P. G. Schyns (Eds.), *Mechanisms of perceptual learning* (Vol. 36, pp. 83–125). San Diego, CA: Academic Press.

Tanaka, J. W., Giles, M., Szechter, L., Lantz, J. A., Stone, A., Franks L., & Vastine, K. (1996). *Measuring parts and wholes recognition of cell, car, and dog experts: A test of the expertise hypothesis.* Unpublished manuscript, Oberlin College, Oberlin, OH.

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

Tarr, M. J. (in press). Visual object recognition: Can a single mechanism suffice? In M. A. Peterson & G. Rhodes (Eds.), *Analytic and holistic processes in the perception of faces, objects, and scenes.* New York: JAI/Ablex.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2,* 55–82.

Tarr, M. J., & Gauthier, I. (1998). Do viewpoint-dependent mechanisms generalize across members of a class? *Cognition: Image-based recognition in man, monkey, and machine, 67,* 71–108.

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

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233–282.

Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint-dependent. *Nature Neuroscience, 1,* 275–277.

Williams, P., & Tarr, M. J. (1998). *RSVP: Experimental control software for MacOS* (Version 4.0) [Computer software]. Retrieved January 10, 1995, from http://www.cog.brown.edu/~tarr/rsvp/

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

Young, A. W., Hellawell, D., & Hay, D. (1987). Configural information in face perception. *Perception, 10,* 747–759.

Received May 5, 2000
Revision received June 8, 2001
Accepted August 7, 2001 ■