Expertise in object and face recognition

James Tanaka
Oberlin College

Isabel Gauthier
Yale University

in Psychology of Learning and Motivation, *Mechanisms of Perceptual Learning* (eds. Goldstone, Schyns & Medin), vol. 36: 83-125, 1997.

Expertise in object and face recognition

A considerable amount of psychological research has been devoted to understanding conceptual changes that occur as a result of domain-specific expertise. This research includes de Groot's early work (1966) on the memory abilities of chess experts and Chi and colleagues' experiments (1981) examining the problem solving strategies of physics experts. Less research has focused on perceptual changes that occur as a result of expertise. If expertise influences higher cognitive functions, such as those involved in memory and problem solving, how might it influence the more fundamental processes involved in object perception and recognition?

In this chapter, we explore the role that experience plays in shaping the object recognition processes by examining the recognition processes of object experts (i.e., people who quickly identify objects at subordinate levels of abstraction). By definition, experts have more experience with and are more knowledgeable about objects in their domain of expertise than are novices. However, a provocative question is whether qualitative differences exist between the recognition processes of experts and of novices.

We begin with the assumption that the name by which an object is identified can provided important insights into the mental representation mediating recognition. In his seminal paper, "How shall a thing be called?," Roger Brown (1958) argues that the object name reflects the level at which the object referent needs to be categorized for the community's nonlinguistic purposes or, to use his term, for the level of *usual utility*. As Brown points out, the level of usual utility changes according to the demands of the linguistic community and this is especially true for expert populations. So, for example, while it is quite acceptable for most of us to refer to the object outside our office window as a "bird," if we were among a group of bird watchers, it would be important to specify whether the object was a "white-throated" or "white-crown sparrow." Generally, experts prefer to identify objects in their domain of expertise more specifically than novices do.

While few would argue that experts identify objects in their domain at a more specific level than novices, a separate question is whether experts initially recognize objects at this more specific level. In Section I of the chapter, we define object expertise as the ability to quickly and accurately recognize objects at specific or subordinate levels of abstraction. In this section, two kinds of object expertise are discussed: a specialized kind of expertise involving the recognition of objects from a particular object class (i.e., birds, dogs) and a ubiquitous form of expertise involving the recognition of faces. Object experts and "face" experts are similar in that both kinds of recognition involve identification at subordinate levels of abstraction. In the case of object experts, this is the level of "robin" or "beagle," and in the case of face experts, this is the level of "Bill Clinton" or "Margaret Thatcher."

However, despite the subordinate shift in recognition demonstrated by both kinds of experts, it is controversial as to whether the processes of face recognition and expert object recognition are mediated by the same or different mechanisms. As we discuss in Section II, on one side of the issue, the modularity hypothesis maintains that faces are recognized by a separate face-specific module that is biologically endowed and relatively unaffected by experience. In contrast, the expertise hypothesis claims that face recognition is not "special" in a biological or computational sense, but, like other forms of object expertise, is acquired through extensive training and experience. .

One approach to the modularity/expertise question is take a putative face-specific computation and see how it might apply to expert object recognition. It has been suggested that, unlike other objects, faces are recognized holistically (i.e., memory for a face feature is better when tested in the whole face than when tested in isolation). Is holistic recognition specific to faces or does it emerge as a consequence of expertise? In Section III, we discuss two recent empirical studies in which the holistic processes of real-world and laboratory experts were compared to the processes of novices. Our general conclusions are that, first, contrary to the predictions of a strict modularity account, holistic recognition is not unique to faces in light of evidence demonstrating the holistic recognition of other non-face objects. Second, experts differ from novices in their enhanced sensitivity to the configural properties of a stimulus. Thus, while novices and experts can recognize non-face objects holistically, only the experts seem to be acutely aware of configural properties that distinguish one object from another. Before discussing the cognitive processes of expert object and face recognition, we begin by defining what it means to be an "object expert."

Section I: Object and Face Experts

A single object can be categorized at many levels of abstraction. For instance, the object *white crown sparrow* can be categorized as an "animal," "bird," "sparrow," or "white crown sparrow." While an object may be categorized at many levels of abstraction, it is recognized at only one level of abstraction. In recognition, the initial level at which the perceived object triggers its representation in memory has been referred to as its *entry point* [1] (Jolicoeur, Gluck & Kosslyn, 1984).

<u>Entry points and basic levels</u>

Rosch (1978) believed that entry points are not arbitrary, but are determined by the structure found in the stimulus environment. She argued that "in the perceived world, information-rich bundles of perceptual and functional attributes occur that form natural discontinuities and that basic cuts in categorization are made at these discontinuities" (p.31). Accordingly, Rosch felt that the fundamental or *basic* level of categorization would be determined by the level at which objects showed the most increase in shape similarity.

To test this claim, Rosch, Mervis, Gray, Johnson, & Boyes-Braem (1976) calculated the amount of overlap that objects share with their category cohorts at different levels of abstraction (e.g., the amount of shape similarity for objects that are members of the category "furniture" versus "chair" versus "kitchen chair"). It was found that objects showed the largest increase in

---

[1] The term entry point is equivalent to Biederman's notion of primal access.

perceptual similarity at the level of "shirt," "car," "dog," or "chair," and therefore, should be considered basic. Rosch et al. tested the primacy of this level of abstraction across a variety of object recognition tasks. In a speeded naming task, they found that subjects preferred to use basic level terms to identify objects (e.g., chair, dog) rather than superordinate level terms (e.g., furniture, animal) or subordinate level terms (e.g., easy chair, beagle). In a category verification task, subjects were faster to verify objects at the basic level than at the superordinate or at the subordinate category levels. In an identity priming task, subjects were faster to judge whether pictures of two simultaneously presented objects were identical or different when primed with a basic level name than with a superordinate level name whereas subordinate level names provided no additional priming over basic level names. Based on this evidence, Rosch et al. argued that the basic level of categorization is the level at which objects are first recognized[2] (i.e., the entry point in recognition).

What kind of information distinguishes objects at the basic level? Basic level objects share a greater number of part features relative to members of superordinate and subordinate categories (Rosch et al. 1976, Tversky & Hemenway, 1984). Part features are defined as divisible, local components of an object that are segmented at points of discontinuity, are perceptually salient, and are usually identified by a linguistic label. For example, attributes such as "wings," "legs," "beak," and "head" are the part features shared by members of the basic level category "bird."[3]

Given the predominance of distinguishing part features at the basic level, it is not by coincidence that current models in vision (Biederman, 1987; Hoffman & Richards, 1985) have taken a parts-based approach to entry point recognition. These models share the common processes of extracting primitive part features from the object image and comparing their part descriptions to object representations stored in memory. Most part-based models of object recognition postulate a finite alphabet of primitive parts which can be organized together to represent objects. One example of this is Biederman's Recognition-By-Components (RBC) theory which proposes that objects are represented by a small number of simplified parts (taken from a finite set of 32 "geons") and abstract relationships between them such as "on top of" or "to the right of" (Biederman & Gerhardstein, 1993). While such models are designed to account for basic-level recognition, they are not sufficiently robust to support subordinate level identification of objects that share a common part description (e.g., different kinds of cars or birds).

---

[2] Other studies (Jolicoeur, Gluck & Kosslyn, 1984; Murphy & Brownell, 1984) have shown that for atypical objects (e.g., penguin, rocking chair), the subordinate level category, not the basic level category, is the entry point of recognition. However, consistent with the shape similarity position, atypical objects are structurally distinctive to typical category members.

[3] In contrast, superordinate categories tend to contain functional features (e.g., "eats," "breathes") whereas subordinate categories tend to contain modified parts (e.g., "large beak") or surface (e.g. "red""") features.

Despite this shortcoming, the appeal of part-based models is that they provide a perceptual or bottom-up solution to the problem of object recognition. However, other categorization studies have suggested that it is not the identification of part information *per se*, but the diagnosticity of the part information that is the most important in recognition. Schyns and Murphy (1994) used complex 3D shaded objects (Martian rocks) which could be parsed in different ways to test this hypothesis. Subjects studied objects from one of two categories of Martian rocks: Category A always included Part a, Category B always included Part B, all other parts varied randomly. Subjects' parsing of Martian rocks including either Part A or B was measured in a part delineation task before, during and after exposure to one of the two categories. Before learning a given category, subjects almost never selected the target part, while they almost always did so after learning. Subjects of both groups and a new group of naive subjects were then exposed to a new group of rocks (Category C) which all had Parts A and B adjacent to each other. Subjects previously exposed to categories A and B delineated Parts A and B as two distinct units, whereas naive subjects grouped Parts A and B together as a single part. Although Parts A and B could clearly be segmented as individual units, subjects did not do so unless they had learned that those parts were diagnostic for categorization.

Murphy (1991) has also shown that nonpart information, such as size, color and texture, can play a role in basic level categorization. In a series of experiments, he found that categories that were optimal in their distinctiveness and informativeness showed a basic-level advantage, regardless of whether distinctiveness and informativeness were based on part information. For example, objects that were the same color and texture as other category members and differed from contrast category members on the color and texture dimensions were the fastest to be categorized in a verification task. These results point out one of the potential problems associated with part-based models of entry level recognition. In some cases, part information might not be essential to the basic level representation and thus, play no role in entry point recognition (Murphy, 1991). Murphy acknowledged that in the real world, basic level objects might very well be centered about part information. His findings merely emphasized that parts are not necessary nor sufficient to the basic level.

<u>Defining expertise in object and face recognition</u>

If basic level categories in the natural world are typically distinctive with respect to their parts, object categories subordinate to the basic level are distinguishable on the basis of "modified" parts and color attributes (Tversky & Hemenway, 1984). For example, "beak" is a part feature that differentiates the basic level "bird" from its contrast categories, whereas "long beak" is a modified part feature that distinguishes the subordinate "woodpecker" from its contrast categories. There is a trade-off between informativeness and distinctiveness with regard to subordinate level categories. On one hand, subordinate categories have the advantage of being more visually informative than basic level categories so that, for example, more is known about the visual appearance of an object if it is categorized as a "sparrow" versus as a "bird." On the other hand, subordinate level objects are more visually similar to their contrast categories than basic level objects are and thus, additional analysis is required to differentiate the subordinate "sparrow" from its contrasting subordinates, "wren" and "finch."

The processing cost incurred by subordinate level categorization was demonstrated in the study by Jolicoeur et al. (1984). In their study, they asked subjects to categorize pictures of common objects at superordinate ("furniture," "tool," "animal,"), basic ("chair," "saw", "bird"), and subordinate ("kitchen chair," "cross-cut saw," "robin") levels of abstraction. Pictures were presented in either a long exposure (250 msec.) or short exposure (75 msec.) condition. Their main finding was that for basic level categorizations, reaction times were the same in the short and long exposure conditions. Presumably, 75 milliseconds was a sufficient amount of time for subjects to abstract the part features necessary for the basic level categorizations. In contrast, subordinate level categorizations were disrupted by the brief exposure duration suggesting that additional visual analysis was required for subordinate level categorizations beyond the amount required for basic level categorizations.

For most tasks, categorization at the generic basic level is adequate. However, for some tasks, specific, subordinate categorization is required. For instance, basic-level access is sufficient to find a chair to sit on or to run away from a lion, but subordinate-level access is necessary to find a friend in a crowd or to identify the type of fish one just caught. When forced to perform subordinate recognition of a large class of objects on a repeated basis, people can improve dramatically to the point of performing with such ease that they appear to use an entirely different mechanism. Thus, because subordinate recognition is so taxing perceptually but can be strongly influenced by experience, it is the most relevant factor in object recognition expertise. Taking the example of bird watching, the level of interest for experts is that of differentiating between similar species birds and not that of spotting birds in the forest. Typical measures of bird expertise would be both the speed and the accuracy with which one can identify an object at the subordinate species level of "white-throated sparrow" or "yellow warbler." Although experts clearly have greater accessibility to subordinate-level representations than novices do, it is not clear whether the subordinate level plays a direct role in entry level recognition. It is plausible, for instance, that experts, like novices, initially recognize objects at the basic level and, once the basic-level representation is activated, are simply faster to access the subordinate-level concepts. According to the basic-first hypothesis, expertise doesn't directly influence the entry point in recognition, but facilitates post-entry level categorization. Alternatively, it is possible that expertise could affect the entry point recognition such that initial recognition might occur at the more subordinate level of abstraction.

Object expertise. As a test of the two competing hypotheses, Tanaka and Taylor (1991) investigated the entry level recognition in two kinds of experts:

dog experts and bird experts. They found that in a speeded naming task, experts identified pictures of objects from their domain of expertise using subordinate level labels rather than basic level labels (e.g., bird experts identified a picture of a robin with the label of "robin" rather than "bird"). However, when identifying objects outside their domain of expertise, the expert subjects used basic level terms (e.g., bird experts identified a picture of a beagle with the label of "dog"). The role of subordinate level representations in entry point recognition was more directly shown in a category verification experiment. When categorizing pictures of objects from their domain of expertise, Tanaka and Taylor found that experts were just as fast to categorize objects at the subordinate level as at the basic level. Based on the naming and verification measures, it can be inferred that the experts' subordinate level representations were equally as accessible as their basic level representations

This research suggests that entry point recognition at the basic level is not mandatory. Depending on the task demands and the amount of experience of the categorizer, the entry point can shift downward to subordinate levels. In view of the finer grain of visual analysis needed for the subordinate level categorizations relative to basic level decisions, downward shifts in recognition demonstrate the influence that task demands and experience have on the recognition process. What are the limits of this influence? The task demands of face recognition and our experience with faces presents the ideal test case for this question.

Face expertise. Whereas relatively few people in the population qualify as experts in the recognition of dogs and birds, it has been claimed that virtually everyone is an expert in the recognition of faces (Carey 1992, Diamond & Carey, 1986). Like other kinds of object expertise, face expertise requires the identification of people at specific levels of abstraction. Indeed, face recognition signifies the most specific kind of *unique identity* categorization in which the category label (i.e., proper name) contains a single exemplar. Objects categorized at the level of unique identity are maximally similar to their category cohorts because structural similarity and category level are highly correlated. Therefore, categorization at the level of unique identity is the most taxing level of categorization for the visual recognition system. From a perceptual standpoint then, there are good reasons to think that faces are *not* first recognized at the level of unique identity. Instead, the level at which members begin to bear a structural resemblance to one another, the level of "person" or "human," might serve as the basic level.

Following the recognition tasks used by Rosch et al. (1976), Tanaka (1996) tested the entry point of face recognition. In a speeded naming task, subjects were asked to label quickly and accurately a stack of pictures depicting common objects and familiar faces (e.g., "Bill Clinton," "Margaret Thatcher"). The main finding was that subjects preferred to identify faces with proper name labels rather than basic level terms (i.e., person or face) whereas they identified nonface objects with basic level labels. However, naming faces with their proper names might not necessarily reflect enhanced accessibility to the unique identity level inasmuch as showing a social bias for using proper names to identify familiar people.

The accessibility of the unique identity level was tested in category verification. In this experiment, subjects verified category membership of familiar faces at the superordinate level of "living thing," basic level of "human," and subordinate level of unique identity (e.g., "Bill Clinton"). Subjects were as fast to verify faces at the subordinate, unique identity level as at the basic level. Hence, the unique identity representation was as accessible as the basic level representation[4]. Similar to the bird and dog experts study, equal accessibility of basic and subordinate levels for faces suggests that the subordinate level does not "replace" the basic level as the entry point of recognition. Rather, it is possible that the two levels are accessed in a parallel. The issue of single or multiple entry points in recognition is discussed below.

As a final test, the underlying face representation was probed in an identity priming task. In this task, subjects were presented with a superordinate (e.g., "living thing"), basic (e.g., "human"), subordinate (e.g., "Bill Clinton"), or neutral word primes followed by two simultaneously shown pictures. The subjects' task was to decide whether the two pictures were visually identical or different. Facilitation was measured by the difference in reaction time between primed trials and neutral trials for the same matching picture stimuli. The priming paradigm assumed that the word prime activates the subjects' visual representation, which in turn is used to enhance the matching response (Posner & Mitchell, 1967; Rosch 1975; Rosch et al., 1976). Results revealed no difference in priming between basic level (e.g., "dog") and subordinate level (e.g., "beagle") word primes for non-face objects, whereas for faces, subordinate level words (e.g., "Bill Clinton") produced greater priming effects than basic level primes (e.g., "person"). Presumably, the unique identity prime elicited a highly detailed face representation that was utilized in the matching paradigm. While the notion of multiple entry points remains an open issue, the behavioral evidence indicates that one of the entry points in face recognition at least is the level of unique identity.

Whether the level of unique identity is preferentially accessed over the basic level or whether both are accessed in parallel could be tested in a repetition priming task. Ellis, Young and Flude (1990) found that prior exposure to a face produced repetition priming of faces in a familiarity decision task, but not in an expression or sex judgment task. They argued that a familiar face automatically activates identity level representations and tasks (e.g., familiarity decision tasks) which access the same identity representations should show facilitative priming effects. According to this logic, prior exposure to a face should facilitate identity level judgments (e.g., "Bill Clinton), but not basic level judgments (e.g., "human"). Alternatively, if both levels are activated by the face stimulus, identity and basic level responses

---

[4] Verification times to the subordinate level of gender (i.e., male, female) were no faster than unique identity judgments.

should shown the facilitative effects of repetition priming for identity and basic level judgments.

Finally, it is important to emphasize the type of structural information that distinguishes faces at the level of unique identity. Obviously, all faces share a common description in terms of generic parts and their organization which is why a structural description in terms of simplified parts (as in Biederman's RBC theory) cannot differentiate faces at the level of unique identity. In principle, more precise (metric) representations of each part of a face could support face recognition. In face recognition, metric differences have been referred to as second-order relational properties (Diamond & Carey, 1986). Second-order relational properties are derived by encoding metric differences in part and configural information between a given face and the average face in the population. For example, a particular face might be differentiable from other faces on the basis of how large a nose is relative to the average nose and how narrow the set of eyes are relative to the average inter-eye distance. Therefore, second-order relational properties are vital to the successful recognition of faces at the level of unique identity.

Summary. The hallmark of expertise is the ability to recognize objects at subordinate levels of abstraction quickly and accurately. This form of expert recognition was demonstrated in a specialized domain, as in the case of bird and dog recognition, or a general domain, as in the case of face recognition. Downward shifts in recognition produced by expertise suggests that a strict structuralist approach to recognition advocated by Rosch might not be entirely accurate. Although Rosch was correct in asserting that world is rich in information about "basic cuts in categorization," research in expert recognition suggests that the fundamental divisions in object categories are modifiable as a consequence of expertise. In one sense, expert object and face recognition are similar in that both demonstrate entry point recognition at a level that is subordinate to the basic level. However, what is not clear is whether subordinate recognition in these two domains is achieved via the same mechanisms. While there is little doubt that expert recognition in specialized domains is obtained through years of experience and training, as discussed in the next section, there is considerable debate as to whether face recognition is the product of experience or biology.

Section II. Face Recognition as a Model of Perceptual Expertise.

Recognition expertise has been discussed mainly in the context of face recognition (Carey, 1992; Cohen-Levine, Banich & Koch-Weser, 1988; Diamond & Carey, 1986; Rhodes, Tan, Brake, & Taylor, 1989), for both practical and theoretical reasons. First, because it is usually assumed that all normal humans are experts at face recognition, face experts are the most available subjects! Second, the role of expertise in face recognition has been discussed by several authors (Cohen-Levine, 1989; Diamond & Carey, 1986) as an alternative to the idea that face processing is mediated by a specialized module (Farah, Levinson & Klein, 1995; Nachson, 1995; Yin, 1969). However, if face recognition turns out to be modular, its use as the conventional model for the study of expertise could be problematic. Thus, although the modularity of face recognition has been discussed by several authors (Farah et al., 1995; Nachson, 1995), we feel it is appropriate to ask what evidence supports the idea that face recognition is predetermined and what are the possible contributions of face recognition to the study of expertise. (Note that in the following discussion, face recognition will be used interchangeably with "individual face recognition".) Behavioral effects occurring when faces are categorized at other levels, such as the cross race classification advantage in which subjects classify the race of other-race faces faster than same-race faces, will both be discussed here (see Levin, in press, for a discussion of whether the cross race advantage can be accounted for by differential expertise).

The expertise hypothesis

While it is often taken for granted that face recognition is a universal domain of expertise, it may be important to clarify this idea. There is no doubt that individual face recognition from a perceptual standpoint is a difficult task (as discussed in Section I) and that humans perform it efficiently and almost effortlessly. In this sense, face recognition is a universal domain of expertise because we are better at it than would be predicted by our ability with other tasks of comparable difficulty. In fact, defining face recognition as a domain of expertise often is a relative statement: humans are experts at face recognition relative to object recognition and adults are experts face perceivers compared to children. However, this view is still ambiguous in that it does not specify how adults come to perform at this level with faces and no other category. In other words, face recognition could distinguish itself during development because: (a) a face module was predetermined in the organism, or b) the organism has had particular experience with faces[5]. Thus, the first meaning (a) of face expertise is simply an observation that there seems to be something different in the way we process faces but remains agnostic as to its cause. The second meaning (b), which we will use here, consists of a particular hypothesis about how face recognition eventually becomes different from the recognition of objects in other subordinate classes.

It may be that the modularity and the expertise hypotheses cannot be successfully tested against each other using faces as stimuli. First, adult novices are practically impossible to find for face recognition. Perhaps one case may be the identification of cross-race faces, which is typically worse than own-race face recognition and perhaps not recognized in the same

---

[5]We do not wish to ignore the fact that most behaviors probably are a product of an interaction between biology and environment. Even the strongest expertise claims would have to acknowledge the genetic influences on the organization of the visual system. However, in our discussion, any effect which can be obtained with various non-face object categories will not be considered a modular effect. If an effect (or a brain area) does not appear to be specific to faces, changes correlated with development must then be due either to the maturation of general recognition systems or to learning.

holistic fashion (Rhodes, Tan, Brake & Taylor; 1989). However, it may be argued that faces of any kind cannot really be used as control stimuli. We are so proficient at extracting several types of information from faces, be it gender, emotion or race, that several automatic processes may interact together. For instance, while same-race faces are better identified at the subordinate level, the reverse advantage is true when faces are classified by race (Levin, in press) and how the two may interact is not yet understood. Second, changes in face processing in a child's development could be due to her experience, to maturation of predetermined systems, or to some kind of interaction between the two. In this context, comparisons between face and non-face object recognition in adults with different levels of expertise are key in determining whether the expertise hypothesis can provide a complete account of adult face recognition. Nonetheless, the study of face recognition can be pursued differently depending on whether it is conducted within the expertise or the modularity framework. For this reason, we will first describe the expertise and modularity views as two different ways to approach and analyze the question of face recognition and will later discuss some relevant empirical data.

To understand how the expertise hypothesis contrasts with the modularity hypothesis, it may be useful to describe what an analysis of expertise must include. Unlike the case of modularity which has been abundantly discussed in general terms, there is no clearly defined framework for the study of expertise. Any expertise problem can be analyzed in terms of four components, the first being a description of the initial problem to be solved. As described earlier, a task description solely in terms of the stimuli (e.g., a "face task" or a "letter task") is ill-defined since objects can be categorized at different levels of abstraction. An adequate description of the problem should therefore include an analysis of the stimulus set as well as a task analysis in terms of what operations are required from the subjects. The initial problem of interest for face processing is individual recognition of faces at the level of unique identity.

The second component of an expertise analysis is a description of how the novice and the expert solutions to the problem differ. Indeed, if the study of expertise phenomena is justified over and above the general study of learning processes, it is because such phenomena at least appear to reflect more than a mere increase in performance. (We acknowledge that expertise phenomena can eventually be explained by continuous learning mechanisms. However, the most interesting expertise phenomena may be those where the novice and expert strategies appear to be qualitatively different.) In the case of face recognition, several researchers have suggested that novices (children) use a feature-based strategy and that experts (adults) use a more holistic strategy (Carey, 1992). Moreover, we have suggested earlier that expertise is accompanied by a shift in the hierarchical organization of categorization levels.

A third component consists of specifying the training conditions that are necessary and sufficient for expertise to develop. The *sufficient* training conditions of expertise for face recognition are obvious: they exist in our common everyday experience with faces, which are omnipresent stimuli in our visual environment starting in infancy. It is less clear however what are the *necessary* constraints without which we would not become as efficient as we are at individuating faces. For instance, would passive exposure to faces be sufficient? Due to the difficulty of doing experimental studies of expertise, this has been the least well studied component in the expertise account of face recognition.

A final component in the description of an expertise phenomenon pertains to the specific conditions under which subjects can perform at an expert level (or the transfer conditions). Simply put, we are testing experts only to the extent that we are testing them in conditions in which they can use their expert abilities. Expertise is not merely a property of the subject but is defined by a match between a subject's training experience and the testing conditions. In the case of face recognition, expertise seems to be very specific to presentation conditions which resemble our everyday experience with faces. This is supported by evidence that certain test conditions, such as inversion and a number of configural transformations (Tanaka & Farah, 1993; Yin, 1969), appear to reduce differences between face and general non-face (non-expert) object processing.

The modularity hypothesis

As already mentioned, it has been suggested that face processing is mediated by a specialized module (Farah et al., 1995b; Fodor, 1983; Nachson, 1995; Yin, 1969). According to Fodor, a module can be recognized as a mandatory, domain-specific, hardwired input system which performs innately determined operations. Nachson (1995) reviewed evidence in favor of the modularity of face recognition which includes the claim that face preference may be innate (Goren, Sarty & Yu, 1975), that there are cortical cells responding specifically to faces (Rolls, 1992), that face recognition is disproportionately impaired by inversion (Yin, 1969) and that there is a face selective neuropsychological deficit (prosopagnosia- see Farah, 1990). As we will briefly review, none of these claims stands unchallenged. More particularly, we wish to discuss some reasons why trying to confirm a modular model is vulnerable to alternative explanations even in the face of a large amount of supporting evidence.

Unlike the expertise hypothesis, a modular hypothesis of face recognition proposes that an object's category membership rather than the particular task demands is the most important factor in engaging the face module. The importance of category-membership relative to other stimulus or task properties in the modular framework can be illustrated by the range of evidence which is cited as supporting the modularity hypothesis. The use of faces may in fact be the only common task factor underlying findings such as the existence of cells responding to the presentation of faces (Perrett et al., 1991), the innate preference for facial configuration (Goren et al., 1975) and behavioral effects with face identification (Yin, 1969) or face-parts recognition (Tanaka & Farah, 1993). An assumption of the modularity framework is that all of these effects are signs of an underlying face module: any dissociation between face and non-face stimuli can be used to support this view.

Complicating the debate on the modularity of face recognition, different authors have used the concept of modularity in different ways. Attempting to clarify this problem, Hay and Young (1982) dissociated the question of the existence of a specific part of the brain processing faces (specificity) from the question of whether or not faces are recognized in a qualitatively different way (uniqueness). Nonetheless, the use of terms such as specificity, uniqueness and modularity remains confusing in the discussion of face recognition (Nachson, 1995). Specificity and uniqueness are theoretically independent and both have been invoked as evidence for modularity, although even obtaining evidence supporting both criteria does not logically disconfirm an expertise account. However, a strong modularist position such as the one described by Fodor (1983) cannot do without either of them. The confusion may arise from an important difference between the expertise and the modularity hypotheses. The first focuses on a changing phenomenon that improves with learning. The second considers the adult performance as a steady state in which mechanisms are relatively unaffected by experience. Both models predict that performance improves and changes during development, but they disagree on the causes of this change. Thus, while it may be difficult to resolve the question by studying the development of face recognition, a more powerful strategy may be to investigate the plasticity of the adult visual system. For example, the modularity model rejects the possibility that what happens to face recognition in development could be produced with non-face stimuli in adults, given enough experience.

One difficulty of the modularity hypothesis is that it can't be directly confirmed, which is why its tests are often attempts to disconfirm a competing hypothesis, such as the expertise account. In contrast, the expertise hypothesis for any particular effect can be confirmed experimentally, in demonstrations that expertise with non-face objects can lead to results which have been claimed to be face-specific. This has been done in cases of the inversion effect (Diamond & Carey, 1986), the caricature effect (Rhodes & McLean, 1990), and the sensitivity to configural changes (Gauthier & Tarr, in press-a). As we will show in Section III, the two hypotheses are not exhaustive and it is possible that some effects cannot be entirely accounted for either by the modularity or the expertise account.

<u>Tests of uniqueness and the use of control stimuli</u>

Of all the criteria for modularity, most attention has been given to domain-specificity or uniqueness of face recognition (Nachson, 1995). While a criterion such as innateness is difficult to test experimentally, it is possible to compare faces with control stimuli in order to test whether a neuropsychological deficit, a physiological response, or a behavioral effect is specific to face recognition. One drawback to this approach is that results attributed to face-specific processes might be based on experiments lacking some critical control. For instance, despite the numerous studies showing a disproportionate inversion effect for faces relative to various control stimuli (e.g., Yin, 1969; Scapinello & Yarmey, 1970; Yarmey, 1971; Valentine & Bruce, 1986), Diamond and Carey (1986) demonstrated that the inversion effect can be obtained from dog experts when recognizing dogs. Thus, the face inversion effect can be eliminated from the list of face-specific effects.

The difficulty in interpretation of most face-specific results resides in the many possible explanations for any given dissociation between faces and control stimuli. Face recognition falls at the extreme of many continua and such factors cannot be controlled completely in an experimental situation. One reason why this lack of evidence replicating a face effect with non-face objects may not be compelling is because the richness and complexity of face recognition renders the task of developing adequate control measures challenging. Imagine the multidimensional space composed of all the factors which can be manipulated in a visual object recognition experiment (schematically represented in Figure 1). Some of these factors are stimulus-class membership, categorization level, and expertise level and they could possibly interact together to produce a large number of recognition situations which differ in crucial ways. A considerable amount of work has been done comparing face recognition to basic-level recognition of non-face objects (Kanwisher et al., 1996; Rumiati et al., 1994; Sergent, Otha & MacDonald, 1992; Takahashi et al., 1995) as well as to recognition of non-face objects at the exemplar level (Farah, Levinson & Klein, 1995; Tanaka & Farah, 1993; Tanaka & Sengco, in press). The general finding in such experiments is a behavioral or neurophysiological dissociation between faces and control stimuli. Note that even in the simplified framework shown in Figure 1A, any difference between face recognition and novice basic-level object recognition could be explained by one of three factors (stimulus-class membership, categorization level or expertise level). The usual approach to this problem, illustrated in Figure 1B, is to try to equate all dimensions between a face and a non-face object task so that only stimulus-class membership could explain the results obtained. The difficulty with this approach is that, on the one hand, equating factors perfectly is always difficult and, on the other hand, we do not even know how many relevant factors there are to equate. Not only do face and non-face tasks differ in categorization level and subject expertise, but social importance, stimulus homogeneity, stimulus complexity etc. could also be important dimensions. Thus, a more cautious approach is to investigate whether some of these dimensions could explain differences between face and non-face tasks is to manipulating them one at a time in the domain of non-face objects, so that if particular face effects are replicated, they cannot be explained by stimulus-category[6](Figure 1C).

---

[6] Of course, while this approach is more cautious when testing the hypothesis that a particular effect is due to stimulus-class membership, it does not permit very strong claims that the effect is actually due to the manipulated dimension, again because an unknown number of possible confounds apply.

Although subjects' expertise for control stimuli and faces is rarely matched, the few studies which have manipulated expertise have shown its importance. As described earlier, Tanaka and Taylor (1991) have found that expertise leads to a downward shift in entry point recognition. Diamond and Carey (1986) found that dog experts' recognition of dogs was as disrupted by inversion as face recognition. Rhodes and McLean (1990) obtained a caricature advantage previously found with faces (Rhodes, Brennan & Carey, 1987) in bird experts identifying birds from a very homogeneous class. Using an ingenious manipulation of expertise level with face stimuli, Rhodes et al. (1989) reported that the effect of inversion was larger for faces of the subjects' own race than for different race faces. Together, such results provide evidence that expertise at least as important as stimulus-class membership to explain behavioral dissociations between objects and faces.

<u>Tests of neural specificity</u>

Behavioral measures tend to show that although face recognition can be dissociated from non-face object recognition, those dissociations might be a result of the subjects' expertise in a particular object domain. Perhaps for this reason, discussion of modularity has recently shifted towards Hay and Young's second criterion (1982) of specificity at the neural substrate level (Davidoff, 1986; 1988; Nachson, 1995; Newcombe, Mehta, & DeHaan, 1994). The idea is that the face recognition system is not qualitatively different from the regular object recognition system but that it is nonetheless separate and engaged in a mandatory fashion by the presentation of faces. Evidence for this view comes mainly from two sources: findings of cells in the temporal cortex that respond selectively to faces (Perrett, Oram, Harries, Bevan, Hietanen, Beason & Thomas, 1991; Rolls & Baylis, 1986) and the existence of a selective deficit for face recognition called prosopagnosia (Farah, 1990; 1992). Additional support comes from intracranial recording (Allison, McCarthy, Nobre, Puce, & Belgeret, 1994) and functional neuroimaging (Haxby, Grady, Horwitz, Ungerleider, Maisog, Pietrini, & Grady, 1994; Kanwisher, McDermott, & Chun, 1996; Sergent, Ohta, & MacDonald, 1992) studies in humans, which all indicate activation of a region of the ventral temporal cortex (including the fusiform gyrus) to face stimuli. However, as discussed here, none of these lines of research has produce conclusive evidence on the face specificity question.

Logothetis, Pauls and Poggio (1995) have conducted single-cell experiments in monkeys which suggest that the selectivity of neurons can be tuned by experience with artificial objects similar to face cell selectivity. They extensively trained monkeys to identify novel 3D objects (~600 000 trials for a given object category) and found a population of IT neurons that responded selectively to novel views of previously unfamiliar objects (wire and spheroidal shapes). Moreover, the response pattern of those IT cells points to a direct role in recognition as monkeys showed no selective response for objects they failed to recognize. Logothetis et al. suggested that IT neurons may be "tunable" to accommodate changes in the recognition requirements of the animal and that neurons selective for particular views of objects in an homogeneous class may be mediating configural (holistic) representations. These results also appear to go against the idea that non-face cells show less selectivity than face cells (Bayliss, Rolls, & Leonard, 1985). Importantly, the tuning properties of cells to novel objects can hardly be explained by an innate predisposition and are more consistent with the type of plasticity postulated by the expertise hypothesis. Of course, it could still be the case that face cells are segregated from general recognition cells but this remains to be shown. First, face-cells of various types of selectivity are found in several parts of the monkey cortex such as IT and the superior temporal sulcus. Second, selectivity to a certain class is never found for all cells tested in a given area (Logothetis et al, 1995; Perrett et al., 1991). Third, experimenters rarely test the same cells with both faces and non-face objects so that it is difficult to assess the extent of overlap between cell populations tuned to face and non-face objects.

Another way to test the idea of a segregated face-specific neural substrate is to consider the possibility that it can be selectively damaged by brain lesions. There may not be a more debated question in the neuropsychology of object recognition than the one of whether a "pure" case of impaired face recognition exists. For example, many cases of prosopagnosia have been observed to be accompanied by more general deficits in object recognition as well as perceptual deficits (Kosslyn, Hamilton, & Bernstein, 1995; Levine & Calvanio, 1989, Rentschler, Treutwein, & Landis, 1994). However, correlated deficits need not be functionally related - in at least a subset of the patient population, face processing may actually be selectively damaged (e.g. De Renzi, 1986; De Renzi, Faglioni Grossi, & Nichelli, 1991). Pure cases of prosopagnosia are rare and most prosopagnosics show some level of impairment for perceiving and/or recognizing non-face objects, especially for within-class recognition of visually similar stimuli. This observation led to the hypothesis that faces could simply be the most difficult objects to recognize and that prosopagnosia could be explained by a mild form of agnosia which would only be detectable for the most difficult recognition task (Damasio, Damasio & Van Hoesen, 1982). Contrary to this account, there is at least one prosopagnosic, WJ, studied by McNeil and Warrington (1993) who was able to recognize sheep faces much better than human faces, even though normal subjects found the sheep faces more difficult. In another study designed to show that prosopagnosics can be disproportionally impaired at face processing, Farah et al. (1995a) tested the patient LH and normal controls for old/new recognition of faces and eye glass frames. While normal subjects achieved scores of ~85% for faces and ~70% for eye glass frames, LH did not show the same superiority for faces and scored ~65% in both conditions. It was concluded that LH's performance shows that prosopagnosia is a face-specific impairment.

However, it is not evident from accuracy measures alone whether a patient’s performance is qualitatively the same or different from normal controls. A complete picture of a patient's impairment cannot be provided without signal detection measures (which control for the possibility that patients may differ from normals on bias rather than sensitivity) and response time measures (which can reveal speed-accuracy tradeoffs - see Kosslyn, Hamilton, & Bernstein, 1995). Along these lines, Gauthier, Behrmann, Tarr, Anderson, Gore and McClelland (1996b) tested two prosopagnosics and normal subjects with shaded images of common and novel non-face objects at different levels of categorization and recorded both accuracy and response times. They found evidence that even though prosopagnosics could produce accuracy levels in the normal range for both common and novel objects, the patients' reaction times showed a disproportionate increase with each categorization level. These results illustrate how accuracy alone cannot provide an adequate measure of the relative performance of two groups of subjects.

Techniques such as PET and fMRI may offer a more flexible alternative than lesion studies to address the question of a specific neural substrate for face recognition. Several functional neuroimaging studies have found selective activation for faces in the fusiform gyri (Haxby et al., 1994; Kanwisher et al., 1996; Puce, Allison, Gore, & McCarthy, 1996; Sergent et al., 1992). However, none of these studies has attempted to control for factors such as level of categorization or expertise. One exception is a recent study that found selective activation in the fusiform gyri during subordinate category verification of non-face familiar objects (Gauthier, Anderson, Tarr & Gore, 1996). While a direct comparison with faces is required, such results indicate that the level of visual categorization may account for at least part of the dissociations found between face and object tasks in the ventral pathway. Similar experiments manipulating expertise could inform us on the plasticity of the adult extra-striate cortex.

What can be learned about face recognition using the expertise model and about expertise using the face model?

In view of the numerous factors (among which level of categorization and expertise) confounded in face processing, one question is: can an expertise analysis us help to understand the relative contributions of these factors. Figure 2 presents a list of factors which may constrain the expert solution to the face recognition problem, such as the symmetry of the objects, their shared configuration and their social importance, and a list of effects and measures, such as the inversion and composite effects, which have produced dissociations between face and non-face processing. One challenge resides in understanding which combination of constraints leads to any of these face effects. For instance, is expertise with objects that have a strong canonical orientation sufficient to produce an inversion effect, or is a shared configuration also necessary? Similarly, it is not entirely clear which of the listed effects should be grouped together as properties of the same underlying phenomenon. For example, are the composite and the whole/part advantage really two measures of a same holistic recognition mechanism, or can they be dissociated? These issues are particularly difficult to resolve because of the impossibility of performing experimental studies of face expertise.

Developmental studies of face recognition have indicated that measures which have often been considered equivalent may in fact reflect different underlying components. For example, Carey and Diamond (1994) compared the composite and the inversion effects in adults as well as children of 6 and 10 years of age. The composite effect (Young, Hellawell & Hay, 1987) is demonstrated by the disruption of the recognition for a part of a face (i.e., the top half) when it is aligned with a complementary part from another face (i.e., the bottom half). When presented upright, the parts seem to fuse together to produce an entirely novel face and disrupts part recognition. Carey and Diamond obtained a strong composite effect at all ages. However, when the fused faces were turned upside down, inversion was more disruptive for the older children than the younger children as measured by reaction time (but see Carey, 1981 and Flin, 1985). This suggests that there are at least two sources to the configural sensitivity in face processing: one seen throughout development and the other appearing with experience.

The expertise model offers a valuable tool to investigate the source of different face effects. Expertise studies count among the most successful methods to replicate face effects with non-face objects and such work plays a major role in bridging the gap between the face and object recognition literatures. That is, if behaviors such as configural encoding are not exclusive to faces, they may turn out to be very informative about general object recognition mechanisms. However, there is one drawback to face recognition being the prototypical expertise model. Because the study of expertise in object recognition has been set up around the face issue, all the non-face recognition tasks studied with experts are constrained to be as similar as possible to the face problem. In some sense, we have been exploring a very small part of the multidimensional space composed of all the possible combinations of constraints on visual object recognition (Figure 1A). For instance, Diamond and Carey (1986) proposed three conditions for a face-like inversion effect to emerge. First, the exemplars of the stimulus-class must share a configuration. Second, it must be possible to individuate the exemplars by using second-order relational features. Third, subjects must have the expertise to exploit such features. This proposal has shaped the study of expertise in that these conditions are often assumed to be necessary for the emergence of any of the face-specific effects. However, note that while these three conditions may produce *sufficient* circumstances for the inversion effect, it is still unknown whether they are *necessary* for the inversion effect, even less so for other effects. Although Diamond and Carey imply that only expertise with a set of homogeneous objects would lead to the inversion effect, this claim is only supported by their failure to find an inversion effect for landscapes. Contrary to this claim, Tanaka and Farah (1991) found that the magnitude of the inversion effect was the same for dot patterns that differed in first-order relational properties and second-order relational properties.

Another frequent assumption seems to be that sets of objects which share a configuration of features are the most challenging for the visual recognition system (Rhodes & Tremewan, 1994). However, there are a few examples in which sets of objects composed of the same or similar features in different spatial organizations lead to very difficult recognition tasks and to performance that shows dramatic orientation sensitivity, not unlike face recognition (Edelman & Bülthoff, 1992; Rock & DiVita, 1987; Tarr & Pinker, 1989). One example is a simple set of 7 stick figure objects (Tarr & Pinker, 1989; Gauthier & Tarr, in press-b) all composed of a small number of the same features in different configurations. Subjects learning to associate names with four of these objects may take as long as 30 minutes of practice, which speaks of the difficulty of the discrimination for such a small set of fairly simple stimuli. The face literature would suggest that such a stimulus category does not posses the characteristics necessary to lead to the inversion effect with expertise, as they do not share a configuration of features. However, because expertise can produce surprising qualitative shifts in behavior, it may be presumptuous to predict expertise performance with such objects from experiments with novices. The expert recognition of such a class has not been tested but it could turn out to produce some of the face effects. Therefore, even though the understanding of face processing may profit largely from experiments using tasks matched with face recognition, it can be argued that the study of expertise in object recognition could benefit from moving beyond the particular set of constraints which are combined in the face recognition problem.

Summary. The study of expertise in visual recognition has been shaped by the problem of face recognition. Even when experiments are not conducted with face stimuli, researchers have matched the constraints of the recognition task to that of face recognition in the hope of replicating face-like effects with non-face stimuli. This approach has led to a body of evidence supporting the idea that face recognition is not unique because similar mechanisms can be engaged with non-face objects given appropriate constraints and subject expertise. Although the face recognition mechanisms may not be unique, there may nonetheless be a face-dedicated neural substrate. Converging evidence from several techniques indicate that a region of the human inferior temporal cortex plays a crucial role in face recognition. However, the same region is also implicated in subordinate-level recognition of non-face objects and the existence of a strictly face-specific cortical area is still an open research question in cognitive neuroscience.

Apart from what it can contribute to the understanding of face recognition, the experimental study of expertise could be a valuable tool to study general object recognition mechanisms. Recognition and categorization tasks can be constrained in many different ways and most current models may fall short of accounting for this diversity. As a partial solution, the expertise model attempts to identify the relevant factors that influence recognition and how experience might shape recognition processes in a particular task domain.

SECTION III: An Operational Test of the Expertise Hypothesis

One experimental approach to the expertise question is to identify a putative face recognition operation and test whether that same operation is involved in other forms of expert object recognition. For a long time, face recognition researchers have claimed that faces are recognized more holistically than other kinds of objects. While this claim is not controversial, good operational tests of holistic processes were missing in the literature.

Recently, Tanaka & Farah (1993) have operationalized the concept of holistic recognition in the following task: In the learning phase of the experiment, subjects memorized a set of normal faces and a set of contrast stimuli; scrambled faces, inverted faces, and houses. After learning, in a two-choice test of recognition, subjects identified the face parts (eyes, nose, mouth) and the house parts (small window, large window, door) presented in isolation and in the whole object. For whole object test items, the targets and foils were constructed such that they differed only with respect to the critical feature under test. For example, if the nose feature was being tested, the other features of the target and foil faces (e.g., face outline, hair, eyes, mouth features) were identical. Therefore, the difference in subjects' performance between isolated part and the whole object test conditions should reflect the extent to which the object is represented as a whole. The central finding was that for normal faces, subjects recognized parts better in the whole face than in isolation. In contrast, subjects were no better at recognizing parts of scrambled faces, parts of inverted faces, and parts of houses when shown in the whole object condition than when shown in isolation. Thus, whereas recognition of a part from a normal face was facilitated by presence of the whole face, recognition of parts from the contrast stimuli (i.e., inverted faces, scrambled faces, houses) was not facilitated by whole object information. From these data, Tanaka and Farah suggested that faces are represented holistically.

Assuming that holistic processing is a specialized operation for the recognition of faces, the part/whole recognition paradigm can be used to address the following questions relevant to the expertise hypothesis:

1.) Can the part/whole advantage be found for recognition of other objects besides faces?

2.) Does expertise affect the magnitude of the part/whole advantage?

In an effort to examine these questions, two studies were carried out. In the first study, current object experts were tested for their holistic recognition of domain-specific objects and faces. The second study was a training experiment in which novice subjects were trained in their recognition of artificial "Greeble" stimuli. After training, their holistic recognition of old and new Greebles was tested.

Experiment 1: Testing holistic recognition in experts

Tanaka and Farah's parts/wholes experiment (1993) showed that whereas normal upright faces are recognized holistically, other stimuli (i.e., inverted faces, scrambled faces and houses) are recognized on the basis of their features. While results from this study are informative about normal face recognition processes, they do not directly test the predictions of the expertise hypothesis. Given the ubiquity of faces in the environment, subjects obviously had more exposure to normal, intact faces than to inverted faces, scrambled faces, and houses. Differential amounts of pre-experimental exposure might produce changes in encoding strategies such that, for example, more familiar objects might be encoded holistically relative to less familiar objects. Therefore, a fairer test would be to assess holistic recognition of expert subjects who had extensive exposure to contrast stimuli.

Taking the expertise factor into account, Tanaka, Giles, Szechter, Lantz, Stone, Franks, and Vastine (1996) recently tested holistic recognition of real-world experts who specialize in the recognition of particular classes of objects. Tanaka et al. reasoned that if holistic recognition emerges as result of extensive experience in making fine visual discriminations, then individuals who have significant experience in other object domains should show similar effects of holistic recognition for those objects. To test this prediction, they recruited expert subjects who had extensive experience in the recognition of biological cells, automobiles or Rottweiler dogs. Cells, cars and Rottweiler dogs were selected as the appropriate contrast stimuli for faces because these objects, like faces, had local identifiable features that were organized in a prototypical configuration.

In this study, all experts had a minimum of five years experience (most had over ten years experience) and were currently active in their field of expertise. Previous studies have shown that children as young as five years of age demonstrate holistic face recognition; hence, five years of expertise should be sufficient for holistic recognition to develop in other object domains. It is important to point out that given the vast amounts of experience that humans have with faces, it was not expected that experts would show the same levels of holistic processing for recognition of objects in their domain of expertise as faces, only that the experts show *more* holistic processing for these objects relative to novice controls.

In the training phase of the experiment, expert and novice subjects learned to associate names to six face stimuli and six non-face objects (i.e., cells, automobiles, Rottweiler dogs). Training continued until subjects could identify all twelve stimuli without error. This procedure was followed to help insure that any differences found between the expert and novice groups in holistic processing were not attributable to general attention or memory effects. That is, it is possible that experts might demonstrate more holistic recognition for domain-specific objects simply because experts remember these objects better than novices. Therefore, once overall recognition performance between the expert and novice groups were equated, the parts and wholes task could be employed to assess the recognition strategies that govern performance.

In this study, the recognition strategies of three types of experts were examined: biology experts, car experts, and dog experts. Biology expertise was tested because biological cells are similar in morphology to faces in that cells have identifiable internal features (i.e., nucleus, nucleolus, mitochondria) that can be manipulated in the frontal plane. Also, similar to face recognition, expert biologists must differentiate individual cells on the basis of metric differences in cell parts and their configuration.

Car expertise was also considered an appropriate expertise domain because automobiles when viewed from the frontal perspective are similar to faces in that they have discernible internal features (i.e., headlights, grill, and bumper) that are arranged in a prototypical configuration (i.e., the two headlights flank the grill and these features are above the bumper). Although there are many varieties of car experts (e.g., people who specialize in repairing cars, people who specialize in racing cars, etc.), the car experts recruited for this study specialized in identifying particular makes and models of cars. These recognition experts typically are members of car clubs, frequent auto shows, and subscribe to publications that review the most current car models (e.g., Car and Driver, Road and Track). For these individuals, identifying cars at subordinate levels of abstraction is a critical element of their expertise.

The final group of experts were dog experts who specialized in the breeding or judging of Rottweiler dogs. Because Diamond and Carey (1986) found that dog experts demonstrate an inversion effect for the recognition of dogs, it is also possible that they might be recognizing dogs holistically. While Diamond and Carey tested the recognition of body profiles, recognition of dog faces was tested in the current experiment. According to the standards specified by the American Kennel Club, the facial qualities of a dog are as crucial to successful dog breeding as its body posture. Therefore, sensitivity to facial structure seems to be an important component of dog expertise.

According to the expertise position, if holistic processing is a product of experience then experts should show an increased amount of holistic recognition (i.e., larger part/whole difference) for objects in their domain relative to novice subjects. The aim of this experiment is to address whether the holistic operations implicated in face recognition are also involved in the recognition of other objects and how expertise might affect holistic processing.

<u>Method</u>

The biology experts were faculty members of either the departments of Biology or Neuroscience at Oberlin College. Car experts were recruited from a local advertisement or were recommended to the experimenters by other car experts The dog experts were members of Rottweiler dog organizations in the Oberlin/Cleveland area with extensive experience in either training, showing, breeding, or judging Rottweiler dogs.

Stimuli were generated according to the procedure described by Tanaka and Farah (1993). Six composite target faces, cells, cars and dogs were constructed with the Adobe Photoshop graphics package (see Figure 3 for examples). For the human face stimuli, the eyes, the nose, the mouth features were taken from yearbook pictures of different individuals. For the cell stimuli, the nucleus, nucleolus, and mitochondria features were taken from an introductory textbook in cell biology. For the car stimuli, the headlights, bumpers, and grill features were constructed from pictures of various late model automobiles (e.g., Honda Accord, Ford Taurus). For the dog stimuli, the eyes, nose, and mouth features of different Rottweiler dogs were taken from various dog books. Foil objects were generated by replacing the feature of the target object with the feature from a different target. Isolated part versions of target and foil features were made by removing the feature from the face and placing it on a white background. Stimulus items were individually mounted on white card stock.

The experiment was divided into a learning and a recognition test phase. In the learning phase, subjects were informed that their task was to associate a name to a given face or object. Subjects were introduced to the face (object) with the appropriate name. After the initial introductions, subjects were shown a face (object) and asked to identify the stimulus with the correct name. If the subject responded incorrectly, the experimenter provided the correct name. Learning continued until the subject identified each item twice without error, at which time the procedure was repeated for the other stimulus set. Once subjects identified both face and non-face sets, a final learning test was administered and learning continued until the subject was able to identify all face and non-face items without error. Because we were interested in *processing* differences between experts and novices, we felt that it was important to train each group to the same level of competency.

Immediately following learning, a two-choice recognition test was administered. In the isolated feature condition, subjects saw a target and foil exemplar presented in isolation. The subjects were asked to identify the correct target exemplar (e.g., "Tom's nose," "Cell A's nucleus"). In the equivalent full object condition, subjects saw the target and foil exemplars embedded in a whole face and asked to identify the correct target face (object). The full face (object) stimuli were constructed such that they differed only with respect to critical feature under test; the other features were held constant. The left and right positions of the correct isolated part and full target stimuli were counterbalanced across test items. Experimental trials were randomly presented with the restriction that items from the same face (object) could not be tested on consecutive trials.

<u>Results and Discussion</u>

Biology experts versus novices. Overall, subjects correctly recognized face parts presented in isolation and in the whole face on 67% and 91% of the trials, respectively. Cell parts were correctly recognized on the 54% of the trials when presented in isolation and 68% when presented in the context of the whole cell. While cells demonstrated a reliable part/whole difference, faces demonstrated a larger part/whole difference than cells. As shown in Figure 4a, experts and novices did not differ in their part and whole recognition of cells and faces. Specifically, there was a greater part/whole difference for faces than cells and this pattern was essentially the same for experts and novices.

While faces demonstrated more holistic processing than cells in absolute terms, it could be argued that cell recognition demonstrated even greater holistic recognition than faces when chance performance is taken into account. Specifically, whereas recognition of cell parts improved from 4% above chance in the isolated part condition to 18% above chance in the whole cell condition (a greater than four-fold improvement), recognition of the face parts increased from 17% in isolation to 41% in the whole face condition (less than a three-fold improvement). Hence, interpretation of the relative differences in holistic processing between cells and faces is somewhat clouded by the disparate baselines between cell and face parts recognition in isolation. However, an important control used in this study was that both experts and novices were trained to the same performance criterion. Once the same level of competence is achieved, it is a fair question to ask to what degree was identification based on the recognition of a single part versus recognition of the whole object. Data from the isolated test condition suggest that subjects did not remember individual cell parts (i.e., mitochondria, nucleus, nucleous) and only some of the face parts. In the whole object condition, cell recognition improved, but not to the same degree as face recognition. Therefore, while all subjects learned to identify the cells and faces perfectly, a more holistic route was taken in face recognition than in cell recognition.

Importantly, counter to the predictions of the expertise hypothesis, experts did not demonstrate greater holistic effects for recognition of objects in their knowledge domain relative to novices. However, as mentioned in Section II, there are numerous reasons why any experiment might fail to find an expertise effect. First, an expert biologist possesses many skills and abilities in which the recognition of cells might play only a minor role. Unlike face recognition where expertise is defined by the ability to recognize faces quickly and at specific levels of abstraction, this trait might not be a defining characteristic of biological expertise. Additionally, although biological cells, like faces, have internal features, unlike faces, these features are not found in a standard configuration (e.g., the mitochondria can be found above or below the nucleus). Objects that lack a prototypical configuration might not engender holistic recognition. Thus, failure to find an expertise effect with biology experts might be due to the inappropriate contrast population and object domain.

Car experts versus novices. As mentioned previously, car fronts might serve as a better contrast stimulus than cells because the features of a car, like the features of a face, are arranged in a prototypical configuration. Also, recognition might play a larger role in car expertise than cell expertise.

Subjects correctly recognized face parts presented in isolation and in the whole face on 67% and 87% of the trials, respectively. Car parts were correctly recognized in isolation and in the whole car on the 67% and 75% of the trials, respectively. Faces demonstrated a larger part/whole difference than the contrast stimulus set of cars. However, as shown in Figure 4b, car experts, like the cell experts, did not show a larger part/whole difference for cars than novices.

Results with the car experts mirrored the results with the cell experts. Overall, faces were recognized more holistically than the contrast stimulus set of cars. However, unsupportive of the expertise hypothesis, car experts did not recognize domain-specific objects (cars) more holistically than novices. Therefore, increased holistic recognition must require something more than just expert identification of objects that share a prototypical configuration.

Dog experts versus novices. Given that cell and car expertise failed to promote holistic recognition, it is reasonable to ask what kind of expertise might be expected to encourage whole object processing. Diamond and Carey (1986) found that dog experts were disproportionately impaired in their recognition of dog pictures when the stimuli were inverted relative to novices. Interestingly, inversion effects were restricted to the particular breed of dog in which the experts specialized. If dog expertise produced an inversion effect for recognition of dogs, dog experts might also demonstrate greater holistic recognition for their specialized breed of dog.

Subjects correctly recognized human facial features tested in isolation and in the whole face on 64% and 87% of the trials, respectively. Dog facial features were correctly recognized in isolation and in the whole dog face on 62% and 65% of the trials, respectively. Once again whereas human faces demonstrated a larger part/whole difference, the contrast stimulus set of dog faces did not. As shown in Figure 4c, dog experts failed to show a larger part/whole difference for Rottweiler dog faces than novices.

The dog expert results do not support the expertise hypothesis. Both dog experts and novices failed to show a part/whole difference. Because the there was no evidence that dog faces were recognized holistically, holistic recognition of human faces seems not to generalize to the recognition of dog faces. This finding is consistent with earlier work showing that recognition of dog faces, unlike human faces, failed to show an inversion effect (Scapinello & Yarmey, 1970). Similarly, holistic face recognition appears to be a species-specific rather than a face-specific effect.

Two main findings emerged from Tanaka et al.'s expert study. First, there was no evidence to suggest that expertise influenced the holistic recognition of cells, cars, or dogs. Cell, car, and dog experts demonstrated the same level of holistic processing for domain-specific objects as novice subjects. Do these results argue against the expertise explanation of face recognition? Not necessarily. As pointed out in the earlier section, there are various reasons why recognition of domain-specific objects might differ from face recognition. Although cell, car, and dog experts were selected because it was believed the cognitive operations required for the recognition of these objects were similar to the operations necessary for face recognition, it appears that their skills were not tapped in the part/whole paradigm. Of course, there is always the possibility that some other, as yet untested, group of experts might demonstrate greater holistic recognition for objects in their domain of expertise than novices. However, after testing the expertise hypothesis using a representative cross-section of experts, this claim currently remains unsupported.

The second main finding was that while faces produced the largest amount of holistic processing, other contrast stimuli, cells and cars, were also recognized holistically. This result indicates that holistic processing is not exclusive to faces, but that recognition of other objects can rely on holistic representations, albeit to a lesser extent than faces. Therefore, differences in holistic processing of face and non-face objects appears to be a quantitative rather than a qualitative distinction. Consistent with this idea, Farah (1990, 1992) has argued that objects vary along a continuum to the degree that they are represented as parts and wholes. In Farah's model, faces are located at one extreme of the continuum in that they are represented more as wholes than other non-face objects. As shown in Figure 5, faces, cells, cars, and dogs did not differ with respect to the recognition of their parts. It was only when the parts were reinserted in the context of the whole object that faces dissociated themselves from the other objects. Thus, face recognition, if not "special", is at least distinctive in its dependence on whole object representations.

Experiment 2: Training Greeble experts

The results of the foregoing study with cell, car, and dog experts failed to support the expertise hypothesis because holistic recognition did not increase with experience. One caveat to such expertise studies is that there is limited control over the degree and quality of the expertise of the subjects. In this sense, studies using an experimental manipulation of expertise could offer more control over these variables. While it has been suggested that it takes several years to achieve the level of expertise that adults have with faces (Diamond & Carey, 1986), Gauthier and Tarr (in press) demonstrated that it is possible to create "experts" at novel object recognition in a relatively brief period (7-10 hours) of extensive training.

Two goals of the Gauthier and Tarr (in press) study were:

1.) to test whether experts with artificial, non-face objects (Greebles) could demonstrate a part/whole effect.

2.) to test whether Greeble experts could show sensitivity to configural transformations.

Method

Expertise with the novel objects ("Greebles") was attained by training subjects at three levels of categorization with a set of 30 objects not used in the test phases. During training, subjects saw examples of the two genders, the five families, and ten specific objects and learned novel names for these categories and objects (Figure 6). They then performed repeated runs of 60 trials of a yes/no paradigm at each of these levels of categorization. On any trial, one gender, one family, or one individual name was displayed followed by a single object. The task is to verify ("yes/no") whether the object is consistent with the label. The training continued until subjects reached a pre-established criterion: subjects' recognition of Greebles (as measured by sensitivity and response time) at the level of the individual was not reliably different from recognition at the level of gender or family. This is the "categorization shift" proposed in Section I as a characteristic of expert recognition (Tanaka & Taylor, 1991). This took approximately 8 hours of practice (over many days) or close to 5500 trials.

In the test phase of the experiment, a Greeble version of the part/whole recognition paradigm was used. Expert and novice subjects learned names for 6 Greebles and the generic nonsense names of the individual parts of each. They were tested on part recognition in three different conditions: isolated parts, parts in a new configuration (each of the top parts reoriented 15° towards the front), and parts in the trained configuration (Figure 7). On each trial one part of a particular target object is specified by a prompt (e.g., the equivalent of "Bob's nose"), followed by two pictures side-by-side on the screen. Subjects then selected whether the right or left image contained the specified part (2AFC design). Trials from each of the three conditions were randomly intermixed. Gauthier and Tarr (in press) tested 16 experts and 16 novices (not submitted to the training procedure) with sets of upright and inverted untrained Greebles.

<u>Results and Discussion</u>

The results demonstrated that experts were overall reliably faster, $F(1,30) = 8.21$, $p < .01$, and marginally more accurate, $F(1,30) = 3.65$, $p = .06$, than novices. This confirms that the part recognition task tapped into the type of experience for these particular experts, even though they were only trained at whole object recognition. Moreover, the part/whole advantage was obtained with Greeble stimuli. Greeble parts were better recognized in the context of intact Greebles relative to the recognition of the same parts in isolation. Although this advantage was only reliable (Scheffé's test, $p<.05$) for experts with upright Greebles, it was no different for experts as compared to novices for inverted Greebles (although the effect was smaller in magnitude, see Table 1). Moreover, there was no reliable interaction between Expertise (novice/expert) and Condition (studied/isolated). This led Gauthier and Tarr (1996) to argue that the visual properties of the objects and/or the task, rather than the level of expertise, may have been responsible for the part/whole advantage. This is consistent with results of the Tanaka et al. (1996) expertise study in suggesting that the part-whole advantage is *not* face-specific, but does *not* appear to depend on expertise since it can be found for novices.

		UPRIGHT GREEBLES	INVERTED GREEBLES
NOVICES			
	Transformed	3560/80	2919/76
	Studied	3853/76	3129/80
	Isolated parts	2923/70	2234/78
EXPERTS			
	Transformed	2695/86	2204/85
	Studied	2306/87	2382/83
	Isolated parts	1991/76	1717/79

Table 1. Response times (ms) and percent correct for part recognition of upright and inverted Greebles by novices and experts.

Gauthier and Tarr also found that experts recognized Greeble parts faster in the studied configuration as compared to Greeble parts in the transformed configuration. The Studied-Transformed comparison showed a significant effect of expertise $F(1.30) = 10.8$, $p<.005$, and a near-reliable interaction with expertise $F(1.30) = 3.85$, $p=.059$. Consistent with a recent study by Tanaka and Sengco (in press) with faces, the Studied-Transformed difference was found across the three types of Greeble parts even though only the top parts were moved. This strongly suggests that Greeble experts were no longer representing Greeble parts independently of each other. Unlike the part/whole advantage, the sensitivity to configural changes appears to measure a shift in recognition behavior produced by the expertise training. As discussed in the next section, perhaps the part/whole advantage and the sensitivity to configural changes are indicative of different types of holistic processing, one which appears with expertise and one may be seen as well in novices (just as Carey, 1992, has proposed for the inversion and composite effects). Further support for this conclusion could be gathered by testing other types of experts as well as children's face recognition abilities to see whether the part/whole advantage indeed occurs earlier than sensitivity to configural changes.

Several conclusions can be made from considering the expertise studies just described together with previous findings about the part/whole advantage in face recognition. First, all evidence indicates that the part/whole advantage and sensitivity to configural changes are not face-specific as was originally proposed. Second, the second of these two effects only appeared with expertise but the first showed no difference between novices and experts. We are still far from understanding the stimulus characteristics which lead to these effects. At this point, there does not seem to be a simple factor which distinguishes the type of stimuli for which the part/whole advantage has been obtained (upright faces, cells, cars and Greebles) from those stimuli which fail to produce it (scrambled and inverted faces, houses and dog faces). Third, although both novices and experts showed a part/whole effect, experts nonetheless showed a larger effect (11% difference) than novices (6%) with upright Greebles. Thus, the possibility that the whole/part advantage may be potentiated by expertise remains open.

It may be wise not to generalize the results obtained with a single measure to all tasks measuring configural processing. As described earlier, several authors have recently suggested that tasks which were once thought to be measuring a single "configural processing" factor may in fact be independent (Carey & Diamond, 1994; Rhodes and Tremewan, 1994). This is supported in the Greeble study by the finding of a cross-over interaction with expertise for the test measuring sensitivity to a configural change. Finally, although replicating a face effect with non-face objects may provide a test of the face-modularity hypothesis, it cannot be taken as direct evidence for the expertise alternative. In the case of the part/whole advantage, experiments with different types of control stimuli indicate that neither class-stimulus membership nor expertise holds a complete answer.

Section IV: Conclusions

In this chapter, we began by defining object expertise as the ability to recognize objects at the subordinate level of abstraction. Following this criterion, object experts, such as bird and dog experts, demonstrated a downward shift in recognition when identifying objects from their domain of expertise and normal adults when identifying faces. However, the functional equivalence of expert object recognition and face recognition does not imply that they are mediated by a similar mechanism. Indeed, there has been considerable debate in the literature as to whether face recognition is a form of expert object recognition or whether it is a special (i.e., distinct) kind of recognition.

In this chapter, we have adopted a computational approach to the *specialness* question. Face recognition researchers have suggested that faces, unlike other objects, are recognized holistically - a process that can be operationalized as difference in recognition of part when presented in isolation and in the whole object. Thus, we asked whether holistic recognition is face-specific computation or whether it is a general form of expert object recognition. In the following discussion, we address the separate questions of uniqueness and expertise as they relate to our studies with real-world and laboratory experts.

Is holistic recognition unique to faces? Because some non-face objects were recognized holistically in these studies, we cannot conclude that holistic recognition is the exclusive domain of face processes. Specifically, we found that real-world experts and novices recognized real, non-face objects (i.e., cells, cars) holistically as did Greeble experts and novices when recognizing Greebles. While the strong version of holistic recognition as face-specific was not supported by these results, faces, nevertheless, demonstrated the greatest degree of holistic processing relative to cell, car, and dog stimuli. These findings suggest that the distinction between featural processing versus holistic processing is a continuum rather than a dichotomy. On one end of the continuum are objects that are recognized exclusively with respect to their parts and, on the other end, are objects that are recognized exclusively on the basis of their wholes. Given that faces are recognized more holistically than other tested contrast stimuli, it would appear that they lie on the extreme end of the featural/holistic processing continuum Thus, while face recognition may not be unique in its reliance on holistic representation, faces might be exceptional with respect to the degree that they are processed holistically.

Does expertise require holistic processing? The test of real-world experts revealed that cell, car, and dog experts demonstrated no more holistic processing than novices as measured by the parts and wholes test. Similarly, Greeble novices and Greeble experts recognized Greeble parts better when tested in the original Greeble than when tested in isolation. These findings demonstrate that both novices and experts recognize objects holistically, but that experts did not demonstrate greater holistic processing than novices. However, in a test of configural sensitivity, it was found that Greeble experts, unlike Greeble novices, recognized parts reliably faster when those same parts were shown in an original configuration than when viewed in a new configuration. Moreover, changing the spatial position of one Greeble part disrupted the recognition of the other parts whose spatial positions were unchanged. These findings are consistent with the results of Tanaka and Sengco (in press) in which they found that changing the inter-eye distance in a face disrupted recognition of the nose and mouth features. Thus, for people who are expert in the recognition of faces or Greebles, the distinction between featural and configural information is blurred in the sense that changing one type of information (i.e., configural) affects recognition of the other type (i.e., featural).

Our work with experts and novices indicates that there are multiple routes to object recognition along the featural/holistic continuum,. As shown in Figure 8, there are objects, such as houses, scrambled faces, inverted faces, that are represented solely on the basis of their features, independent of the whole object context (Tanaka & Farah, 1993, Tanaka & Sengco, in press). Given that recognition of a dog part (i.e., eye, nose, mouth) was no better when tested in isolation than when tested in the whole object, dog novices and experts seemed to have adopted a featural approach to recognition.

If holistic processing is defined as the difference in part/whole recognition, our expertise studies indicate that it may be important to distinguish between two types of holistic representation. In the first kind of holistic representation, which we will refer to as a *relative* holistic representation, recognition of a part benefits from the presence of general or first-order configural properties of the object. Greeble novices, for example, demonstrated better performance when the part was tested in the whole object than in isolation. However, there was *no* difference in performance between the transformed configuration and the studied configuration conditions. This latter result indicates that Greeble novices were sensitive to the first-order relational properties of the Greeble (e.g., the quiff was located between the bogus), but not the more subtle second-order relational properties (e.g., the angle of the bogus).

In contrast, face recognition and expert Greeble recognition seems to rely on the encoding of second-order relational properties. We found that face recognition and expert Greeble recognition is reliably better if the parts are tested in the studied configuration than if tested in a different configuration. Moreover, recognition of *all* object features is disrupted if second-order relational information is altered. For this reason, we refer to this kind of holistic processing as *integral* (Garner, 1974) given that changes in configural information produces changes in featural information. While Greeble experts form integral, holistic representations of Greebles, it is an open question whether the holistic representations of cell and car novices and experts are relative or integral.

In summary, there seem to be many routes to visual expertise. Real-world experts do not necessarily generate holistic representations to identify domain-specific objects and in fact, recognition of diagnostic features may be what distinguishes an expert from a novice. However, in other cases where there is a high degree of homogeneity between object features, individuation of an exemplar must be based on several features and their configuration rather than on a single distinctive feature. Greebles and faces are good examples of homogenous object categories and therefore, learning to identify these kinds of objects would be expected to promote holistic processing. Because featural and configural distinctions between exemplars in a homogenous object class are subtle, it is reasonable to assume an extensive amount of experience and perceptual learning is necessary in order to be proficient in recognizing these kinds of objects. Hence, integral, holistic processes are recruited under the conditions where people have extensive experience identifying objects from a homogeneous shape class at subordinate levels of abstraction. This particular combination of stimulus factors and task demands is relatively rare in the natural world and might only occur when people recognize faces. However, our findings suggest that it is possible to simulate these conditions in the laboratory when subjects are trained to recognize Greebles. In this sense, we can say that the processes underlying face recognition may be ecologically special, but not computationally unique.

References

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. Cerebral Cortex, 4, 544-554.

Bayliss, G.C., Rolls, E.T. & Leonard, C.M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex of the superior temporal sulcus of the macaque monkey. Brain Research, 342, 91-102.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. Psychological Review, 94, 115-147.

Biederman, I. & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. Journal of Experimental Psychology: Human Perception & Performance, 19, 1162-1182.

Brown, R. (1958). How shall a thing be called? Psychological Review, 65, 14-21.

Carey, S. (1981). The development of face perception. In G. Davies, H. Ellis, & J. Shepherd (Eds.), Perceiving and Remembering Faces (pp. 9-38). New York: Academic Press.

Carey, S. (1992). Becoming a face expert. Philosophical transactions of the Royal Society of London, B335, 95-103.

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

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Cohen-Levine, S. (1989). The question of faces: Special is in the brain of the beholder. In A. W. Young & H.D. Ellis (Eds.), Handbook of research on face processing (pp. 37-48). Amsterdam: North-Holland.

Cohen-Levine, S., Banich, M.T., & Koch-Weser, M. P. (1988). Face recognition: A general or specific right hemisphere capacity? Brain and Cognition, 8, 303-325.

Damasio, A.R., Damasio, H., & Van Hoesen, G.W. (1982). Prosopagnosia: Anatomic basic and behavioral mechanisms. Neurology, 32, 331-341.

Davidoff, J.B. (1986). The specificity of face perception: Evidence from psychological investigations. In R. Bruyer (Ed.), The neuropsychology of face perception and facial expression (pp. 147-166). Hillsdale, NJ: Erlbaum.

Davidoff, J.B. (1988). Prosopagnosia: A disorder of rapid spatial integration. In G.Denes, C. Semenza, & P. Bisiachi (Eds.), Perspectives on cognitive neuropsychology (pp. 297-309). Hillsdale, NJ: Erlbaum.

de Groot, A.D. (1966). Perception and memory versus thought. In B. Kleinmuntz (Ed.), Problem Solving Research, Methods and Theory (pp. 19-50). New York: Wiley.

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

De Renzi, E., Faglioni, P., Grossi, D., & Nichelli, P. (1991). Appreciative and associative forms of prosopagnosia. Cortex, 27, 213-221.

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

Edelman, S., & Bülthoff, H.H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vision, 32, 2385-2400.

Ellis, A.W., Young, A.W., & Flude, B.M. (1990). Repetition priming and face processing: Priming occurs within the system that responds to the identity of a face, Quarterly Journal of Experimental Psychology, 42A, 495-512.

Farah, M. J. (1990). Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision. Cambridge, MA: The MIT Press.

Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain-specificity in visual object recognition. Current Directions in Psychological Science, 1, 164-169.

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

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995b). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. Vision Research, 35, 2089-2093.

Flin, R.H. (1985). Development of face recognition: An encoding switch? British Journal of Psychology, 76, 123-134.

Fodor, J. (1983). The modularity of mind. Cambridge: MIT Press.

Garner, W. R. (1974). The processing of information and structure. New York: Lawrence Erlbaum.

Gauthier, I., & Tarr, M.J. (in press-a). Becoming a Greeble Expert: Exploring mechanisms for face recognition, Vision Research.

Gauthier, I., & Tarr, M.J. (in press-b). Orientation Priming of Novel Shapes in the Context of Viewpoint-Dependent Recognition, Perception.

Gauthier, I, Anderson, A. W., Tarr, M. J., & Gore, J. C. (1996a). Level of categorization in visual recognition studied with function MRI. Presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.

Gauthier, I., Behrmann, M., Tarr, M. J., Anderson, A. W., Gore, J., & McClelland, J. (1996b). Subordinate-level categorization in human inferior temporal cortex: Converging evidence from neuropsychology and brain imaging. Neuroscience Abstracts.

Goren, C.C. Sarty, M., and Yu, P.W.K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. Pediatrics, 56, 544-549.

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

Hay, D.C., & Young, A.W. (1982). The human face. In A.W. Ellis (Ed.), Normality and pathology in cognitive function (pp. 173-202). New York: Academic Press.

Hoffman, D.D., & Richards, W. (1985). Parts in recognition, Cognition, 18, 65-96.

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

Kanwisher, N., McDermott, J., & Chun, M. M. (1996). A module for the visual representation of faces. NeuroImage, 3, S361.

Kosslyn, S.M., Hamilton, S.E., & Bernstein, J.H. (1995). The perception of curvature can be selectively disrupted in prosopagnosia. Brain & Cognition, 27, 36-58.

Levin, D.T. (in press). Classifying faces by race: the structure of face categories. Journal of Experimental Psychology: Learning, Memory, & Cognition.

Levine, D.N., & Calvanio, R. (1989). Prosopagnosia: A defect in visual configural processing. Brain & Cognition, 10, 149-170.

Logothetis, N.K., Pauls, J. & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. Current Biology, 5, 552-563.

Logothetis, N.K., Pauls, J. Bülthoff, H.H. & Poggio, T. (1994). View-dependent object recognition by monkeys. Current Biology, 4, 401-414.

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

Murphy, G.L. (1991). Parts in object concepts: Experiments with artificial categories. Memory and Cognition, 19, 423-438.

Nachson, I. (1995). On the modularity of face recognition: The riddle of domain specificity. Journal of Clinical and Experimental Neuropsychology, 17, 256-275.

Newcombe, F., Mehta, Z., & DeHaan, E.D.F. (1994). Category specificity in visual recognition. In M.J. Farah and G. Ratcliff (Eds.), The neuropsychology of high-level vision (pp. 103-132). Lawrence Erlbaum Associates: Hillsdale.

Perrett, D. I., & Oram, M. W., Harries, M.H., Bevan, R., Hietanen, I.K., Beason, P.J., & Thomas, S. (1991). Viewer-centered and object-centered coding of heads in the Macaque temporal cortex. Experimental Brain Research, 86, 150-175.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. Nature, 343, 263-266.

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

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.

Rentschler, I., Treutwein, B., & Landis, T. (1994). Dissociation of local and global processing in visual agnosia. Vision Research, 34, 963-971.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. Cognitive Psychology, 19, 473-497.

Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogeneous stimuli: Towards a model of configural coding. Perception, 19, 773-794.

Rhodes, G., Tan, S., Brake, S. & Taylor, K. (1989). Expertise and configural coding in face recognition. British Journal of Psychology, 80, 313-331.

Rhodes, G., & Tremewan, T. (1994). Understanding face recognition: Caricature effects, inversion and the homogeneity problem. Visual Cognition, I, 2/3, 275-311.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. Cognitive Psychology, 19, 280-293.

Rolls, E. (1992). The processing of face information in the primate temporal lobe. In V.Bruce & M.Burton (Eds.), Processing images of faces (pp. 41-68). Norwood, NJ: Ablex.

Rolls, E, & Bayliss, C.G. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the macaque monkey, Experimental Brain Res, 65, 38-48.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B.. Lloyd (Eds.) Cognition and categorization (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rosch. E. (1975). The nature of mental codes for color categories. Journal of Experimental Psychology: Human Perception and Performance, 1, 303-322.

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

Rumiati, R.I., Humphreys, G.H, Riddoch, M.J. & Bateman, A. (1994). Visual Cognition, I,181-225.

Scapinello, K.F., & Yarmey, A.D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli, Psychonomic Science, 21, 329-331.

Schyns, P. G., & Murphy, G. L. (1994). The ontongeny of part representation in object concepts. In D. Medin (Eds.), The Psychology of Learning and Motivation Vol. 31 (pp. 305-354). San Diego, CA: Academic Press.

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

Takahashi, N., Kawamura, M., Hirayama, K., Shiota, J., & Isono, O. (1995). Prosopagnosia, a clinical and anatomical study of four patients. Cortex, 31, 317-329.

Tanaka, J. W., & Farah, M. J. (1991). Second-order relational properties and the inversion effect: testing a theory of face perception. Perception & Psychophysics, 50, 367-372.

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

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, Ohio.

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

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.

Tanaka, K. (1996). Inferotemporal cortex and object vision. Annual Review of Neuroscience, 19, 109-139.

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

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. Journal of Experimental Psychology: General, 113, 169-193.

Valentine, T., & Bruce, V. (1986). The effect of race inversion and encoding activity upon face recognition, Acta Psychologica, 61, 259-273.

Yarmey, A.D. (1971). Recognition memory for familiar "public" faces: Effects of orientation and delay, Psychonomic Science, 24, 286-288.

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

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

Figure Captions

Figure 1. Simplified representation of the multidimensional space arising from all possible combinations of factors constraining object recognition, such as stimulus-class membership, expertise level and categorization level.

Figure 2. Different constraints over face recognition and several effects which distinguish face recognition. Little is known about the causal relationships between them.

Figure 3. Sample of parts and wholes test items for face, cell, car, and dog stimuli used in Experiment 1.

Figure 4. Recognition of face, cell, car, and dog stimuli by experts and novices.

Figure 5. Combined parts and wholes recognition for face, cell, car, and dog stimuli. Note that faces show the largest in whole object recognition.

Figure 6. Sample objects chosen from a set of 60 control stimuli for faces. Each object can be categorized at the Greeble, family, gender, and individual levels. The Greebles were created by Scott Yu using Alias Sketch! modelling software.

Figure 7. Example of the forced-choice recognition paradigm used in Gauthier and Tarr (1996).

Figure 8. The type of information found in featural and holistic representations as demonstrated by the studies with real-word experts and Greeble experts.

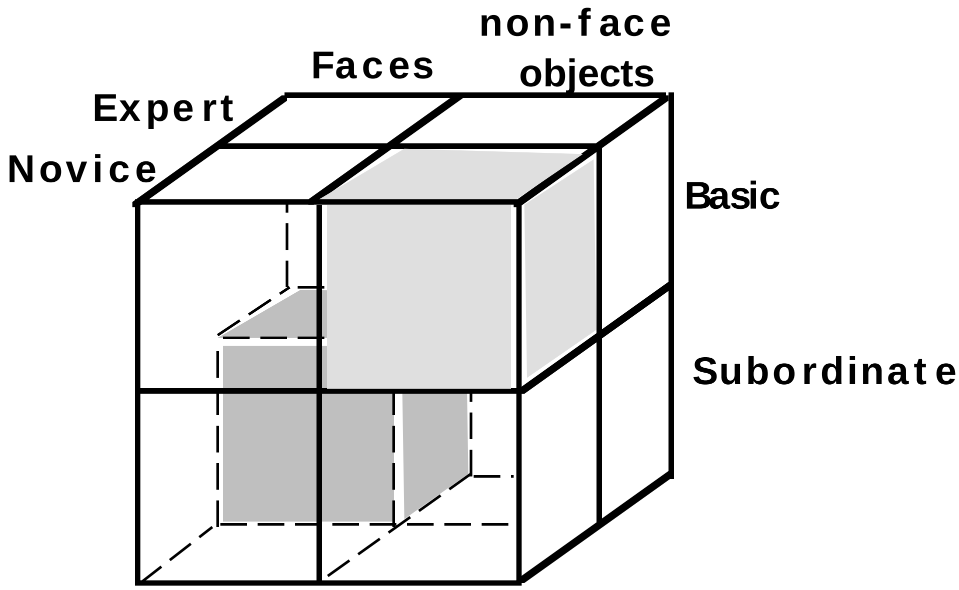
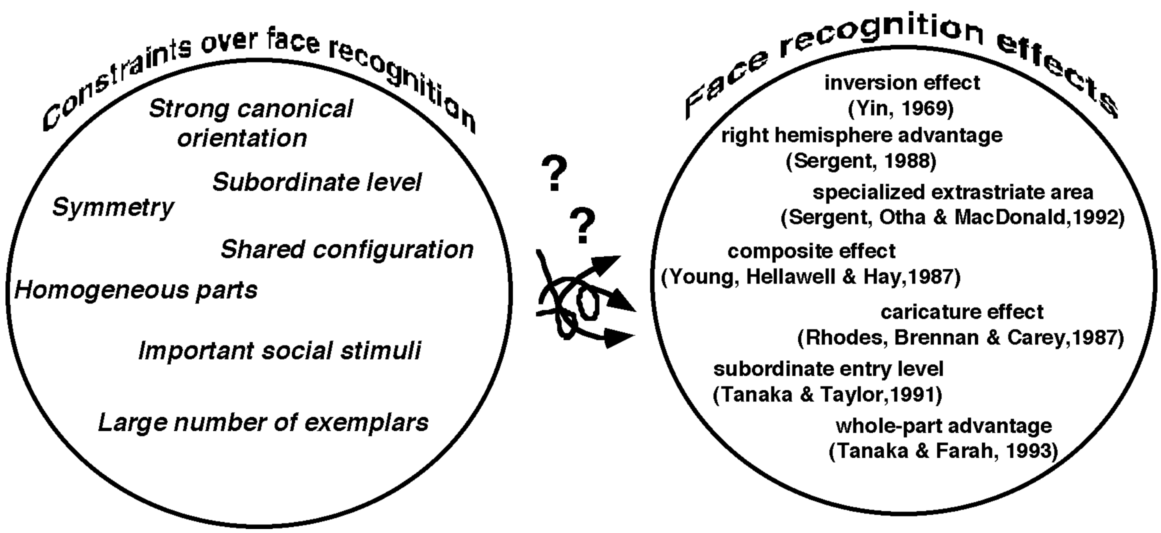


Figure 1



Constraints over face recognition
Strong canonical orientation
Subordinate level
Symmetry
Shared configuration
Homogeneous parts
Important social stimuli
Large number of exemplars
?
?
Face recognition effects
inversion effect
(Yin, 1969)
right hemisphere advantage
(Sergent, 1988)
specialized extrastriate area
(Sergent, Otha & MacDonald,1992)
composite effect
(Young, Hellawell & Hay,1987)
caricature effect
(Rhodes, Brennan & Carey,1987)
subordinate entry level
(Tanaka & Taylor,1991)
whole-part advantage
(Tanaka & Farah, 1993)

Figure 2

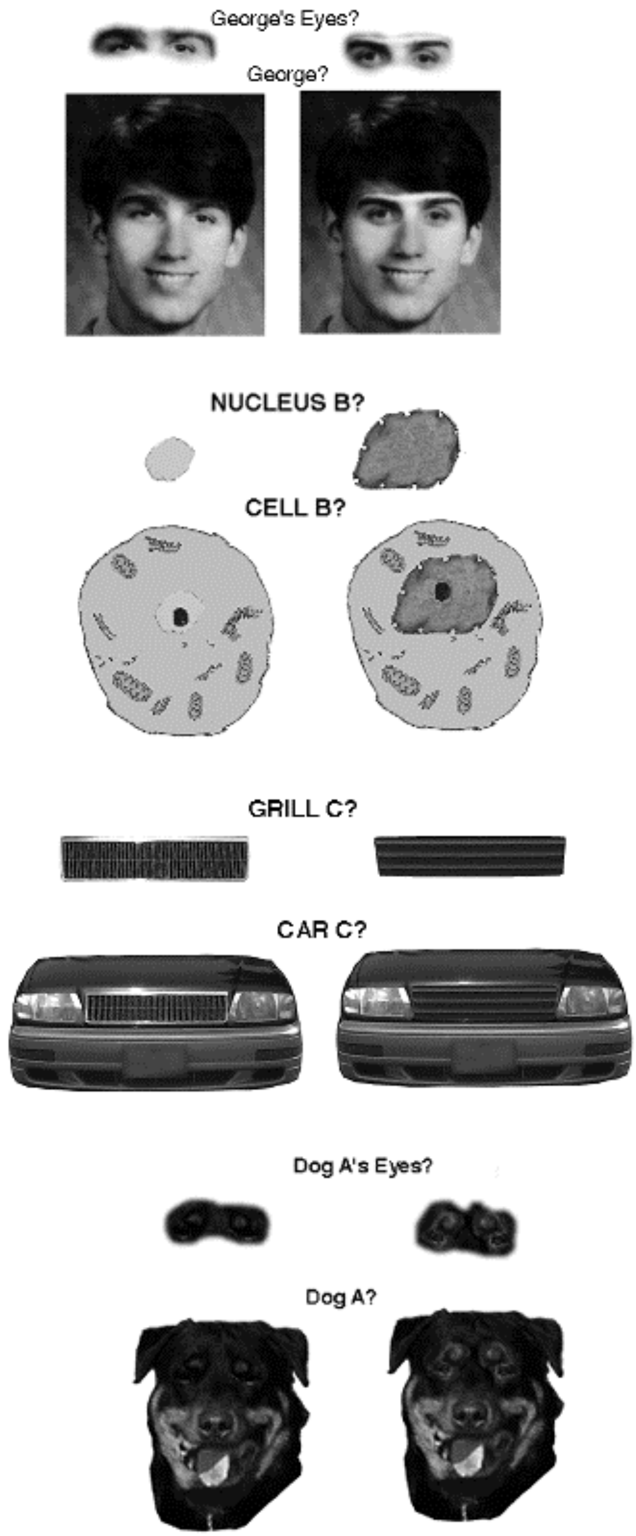

George's Eyes?
George?
NUCLEUS B?
CELL B?
GRILL C?
CAR C?
Dog A's Eyes?
Dog A?

Figure 3

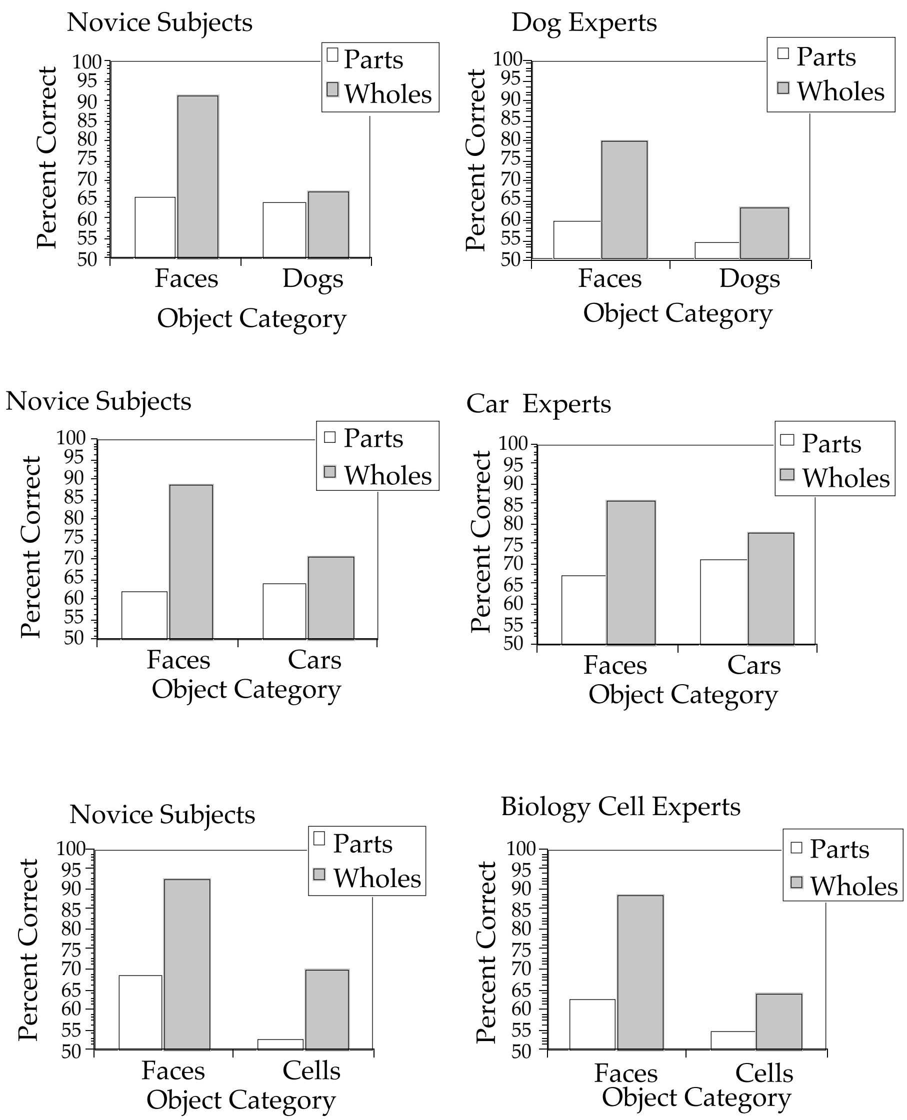

Novice Subjects
Parts
Wholes
Percent Correct
100
95
90
85
80
75
70
65
60
55
50
Faces
Dogs
Object Category
Dog Experts
Parts
Wholes
Percent Correct
Faces
Dogs
Object Category
Novice Subjects
Parts
Wholes
Percent Correct
Faces
Cars
Object Category
Car Experts
Parts
Wholes
Percent Correct
Faces
Cars
Object Category
Novice Subjects
Parts
Wholes
Percent Correct
Faces
Cells
Object Category
Biology Cell Experts
Parts
Wholes
Percent Correct
Faces
Cells
Object Category

Figure 4

Novice and Expert: Parts and Wholes Recognition

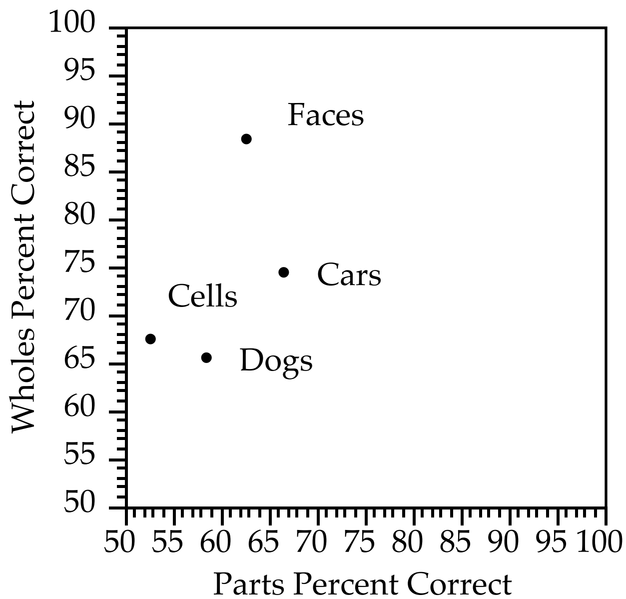


Figure 5

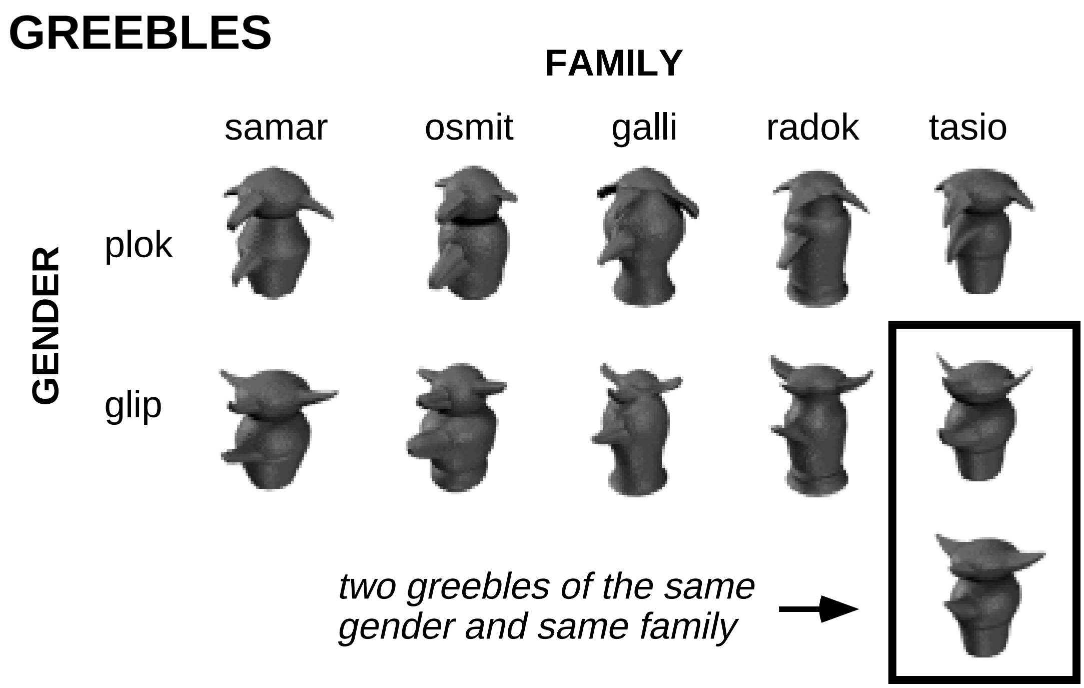

GREEBLES
FAMILY
samar
osmit
galli
radok
tasio
GENDER
plok
glip
two greebles of the same
gender and same family

Figure 6

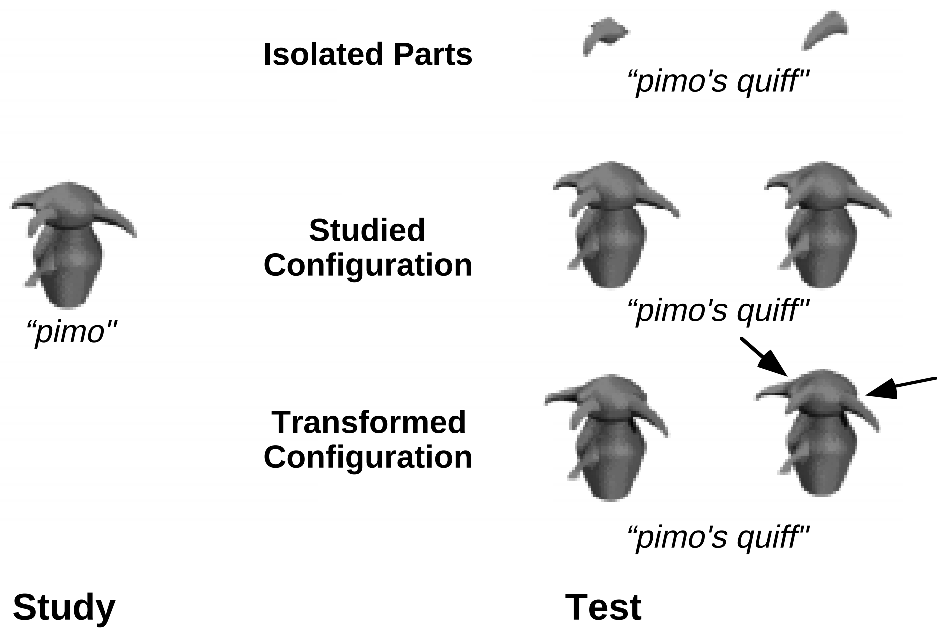

Isolated Parts
"pimo's quiff"
Studied Configuration
"pimo"
"pimo's quiff"
Transformed Configuration
"pimo's quiff"
Study
Test

Figure 7

Representation	***Configural Information***	***Stimuli***	***Novice and Expert Population***
featural	none	houses, scrambled & inverted faces	dog novices & experts
holistic - relative	categorical (1st order relational)	cars, cells	Greeble novices, car & cell novices (?), car & cell experts (?)
holistic - integral	metric (2nd order relational)	faces, Greebles	normals w/ faces, Greeble experts

Figure 8