Vision Research
1997, Vol. 37, No. 12, 1673-1682

BECOMING A "GREEBLE" EXPERT: EXPLORING MECHANISMS FOR FACE RECOGNITION

Isabel Gauthier
Yale University

Michael J. Tarr
Brown University

Sensitivity to configural changes in face processing has been cited as evidence for face-exclusive mechanisms. Alternatively, general mechanisms could be fine-tuned by experience with homogeneous stimuli. We tested sensitivity to configural transformations for novices and experts with non-face stimuli ("Greebles"). Parts of transformed Greebles were identified via forced-choice recognition. Regardless of expertise level, the recognition of parts in the Studied configuration was better than in isolation, suggesting an object advantage. For experts, recognizing Greeble parts in a Transformed configuration was slower than in the Studied configuration, but only at upright. Thus, expertise with visually similar objects, not faces *per se*, may produce configural sensitivity.

Several researchers have proposed that configural information, i.e., the relations between parts, is especially important in the way faces are visually represented (Diamond & Carey, 1986; Farah, 1990; Rhodes, 1987; Sergent, 1988). If this is the case, face processing, as compared to the processing of non-face objects, should be particularly disrupted by changes in the configuration of parts. Tanaka and colleagues tested this hypothesis by examining whether configural transformations influenced the recognition of individual features (Tanaka & Farah, 1993; Tanaka & Sengco, 1996). In several studies Tanaka tested the forced-choice recognition of individual parts of faces (e.g., "Jim's nose") or control stimuli (houses, inverted faces, or scrambled faces). For each stimulus class three conditions were used: 1) parts in isolation (e.g., Jim's nose alone); 2) parts in the context of the studied object with some transformation in configuration (e.g., Jim's nose in Jim's face with the eyes moved slightly apart); 3) parts in the context of the studied object (e.g., Jim's nose in Jim's face). Crucially, the target and distractor parts were *exactly the same* in all three conditions and within each condition the context for both the target and distractor parts was identical. Thus, if subjects are using independent part representations, there should be no difference in the diagnostic information available between the three conditions. Nonetheless, parts of faces were most readily recognized in the Studied configuration, less readily in a Transformed configuration, and most poorly in isolation, suggesting that the parts of faces are not represented independently (a so-called "holistic representation"). In contrast, none of the tests with control stimuli -- scrambled faces, inverted faces, or houses -- revealed any advantage for recognizing parts embedded in the intact configuration of the studied object.

Whenever a particular effect, such as that just described, is obtained with faces and not control stimuli, the question arises as to whether this implicates a face-specific mechanism. From our perspective it is prudent to consider specialized mechanisms only after the best possible control conditions have failed to replicate a given effect. In the case of faces, this means using non-face stimuli that adequately match many of the visual and categorical constraints found for faces. For instance, one of the most famous phenomena associated with faces, the inversion effect, in which there is a disproportionate cost for recognizing inverted faces (Farah et al., 1995; Yin, 1969), has been obtained with a homogeneous set of non-face objects (dogs of the same breed), but only for expert participants (Diamond & Carey, 1986). Similarly, Rhodes and McLean (1990) obtained the caricature advantage, that is, caricatures of faces are recognized more quickly than the actual faces, with bird experts who identified members of a highly homogenous class of birds. Such demonstrations, however, do not necessarily rule out face-specific mechanisms in all phenomena associated with face recognition -- it is certainly possible that some of the effects which are considered to be face specific are mediated by a special mechanism. Therefore, each putative face-specific phenomenon should be tested using experimental conditions that are matched as carefully as is possible, including specifically, equivalent levels of visual homogeneity, categorical level of recognition, and degree of expertise.

One of the most salient characteristics of face recognition is that faces have similar features organized in similar configurations. Therefore, an adequate set of

---

Correspondence to: Isabel Gauthier, Department of Psychology, Yale University, PO Box 208205, New Haven, CT 06520-8205, tel: (203) 432-4567, fax: (203) 432-7172, e-mail: zaza@minerva.cis.yale.edu, or Michael J. Tarr, Department of Cognitive & Linguistic Sciences, Brown University, Box 1978, Providence, RI 02912, tel: (401) 863-1148, fax: (401) 863-2255, e-mail: Michael_Tarr@brown.edu. This work was supported by grants to MJT from the Air Force Office of Scientific Research, contract number F49620-92-J-0169, and from the Office of Naval Research, contract number N00014-93-1-0305. IG was supported by a FCAR post-graduate scholarship. We thank Scott Yu for creating the Greebles, Jay Servidea for creating the experts, Pepper Williams and James Tanaka for the comments on earlier versions of this paper, and Bob Abelson for naming the Greebles.

control stimuli should share this constraint. For this reason, sets of exemplars from a single visually homogeneous category such as species of birds or breeds of dogs have been used as control stimuli. However, it is not only the homogeneity of the subset of objects actually used in the experiment that matters -- for familiar classes of objects, the space of *all known* exemplars is also crucial. Thus, the apparent homogeneity of a control set may be insufficient if the larger class is not homogeneous (as in the case of houses or landscapes, Diamond & Carey, 1986).

A second characteristic of face recognition is that faces are typically recognized at the exemplar-specific level. Thus, while we often recognize most objects at the *basic level* (e.g., “chair or dog,” see Rosch et al., 1976), faces are generally recognized at the most extreme *subordinate level* (e.g., “Jim or Max”). Consequently, it is important that control tasks addressing face-specific effects require the recognition of control stimuli at the subordinate level (e.g., distinguishing between several dogs of the same breed).

A third characteristic of face recognition is that humans are highly expert at the very difficult task of discriminating between individual faces. Although expertise is difficult to define, it seems clear that it should be more than simply a practice effect in which performance improves with experience. One empirical definition that has been used and which we will adopt here is a *qualitative* shift in processing. Tanaka and Taylor (1991) found such a shift for bird experts who were as fast to recognize objects at the subordinate level (“robin”) as they were at the basic level (“bird”). In contrast, non-experts are consistently faster on basic-level discriminations as compared to subordinate-level discriminations. Similarly, because humans are face experts, judgments of face identity (subordinate level) are as fast as judgments that are more categorical, for instance gender (Tanaka, personal communication). Therefore, because expertise interacts with the level of categorization, it is important that control tasks addressing face-specific effects use stimuli for which the participants are experts.

Based on such criteria, studies that have used bird or dog recognition by experts appear to have adequately matched control tasks to face recognition (Diamond & Carey, 1986; Rhodes & McLean, 1990). Indeed, these studies have found evidence for nominally face-specific effects with non-face stimuli. However, there are three limitations to using such controls. First, from a practical standpoint, experts within a given domain may be difficult to recruit. Second, from a theoretical standpoint, extant experts are already trained and, as such, do not provide the experimenter with any opportunity to manipulate the origin or the level of expertise. Third, from an empirical standpoint, several researchers (Carey & Diamond, 1994; Diamond & Carey, 1986; Johnson & Morton, 1991) have emphasized that the inversion effect in dog judges requires 10 years of experience with a specific breed, which is also the time it takes for children to perform in the normal adult range on face encoding tasks (Carey & Diamond, 1994). This long onset to attain expertise suggests that a comparable level of competence may not be obtainable in the time-course of an experiment (or at least one we would wish to run). The study presented here addresses these limitations by attempting to create experts for the subordinate-level recognition of a homogeneous set of non-face stimuli (“Greebles”; Figure 1). In particular, we examine whether, given extensive experience with some Greebles, participants exhibit sensitivity to configural information with unfamiliar Greebles. If indeed experts can be created in the laboratory, this would provide a tool for the investigation of face recognition and, more generally, visual expertise.

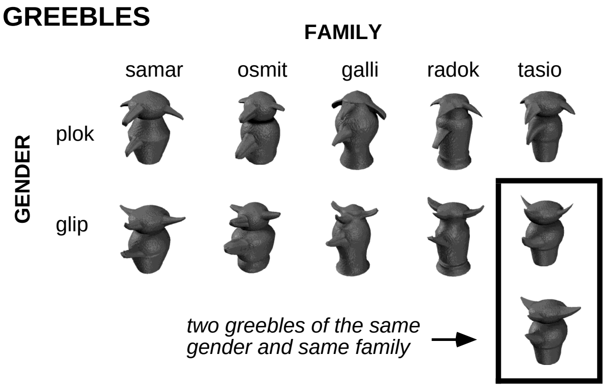


Figure 1. Meet the Greebles. Sample objects chosen from a set of 60 control stimuli for faces. Each object can be categorized at the Greeble, family, gender, and individual levels. The Greebles were created by Scott Yu using Alias Sketch! 3D modeling software.

In the present study we chose to investigate the nominally face-specific sensitivity to changes in configuration (Tanaka & Farah, 1993; Tanaka & Sengco, 1996). In prior studies, control stimuli for faces were houses, inverted faces, or scrambled faces. Given possible non-equivalence between these sets and normal faces, we used stimuli specifically constrained to be similar to faces along several dimensions, Greebles, as our control set. Moreover, we manipulated the level of expertise, so that this variable was not confounded with stimulus class. As discussed earlier, the stimulus transformations used in Tanaka and colleagues’ experiments were independent of the information required to perform the forced-choice recognition judgment. This same manipulation was used here to assess sensitivity to configural transformations. Therefore, if the parts of each Greeble are encoded independently, then the patterns of performance observed for the Isolated parts, the Transformed configuration, and the Studied configuration conditions are predicted to be equivalent. On the other hand, if the parts of each Greeble are encoded in a configural manner, that is, the positions of individual parts are dependent on one another, then performance is predicted to be best in the Studied configuration condition, poorer in the Transformed configuration and the Isolated parts conditions. Crucially, this pattern is expected to be more pronounced for experts than novices. Moreover, an interaction of expertise with orientation is expected, that is, for experts, the recognition of parts in upright Greebles should be more sensitive to configural transformations than the recognition of parts in inverted Greebles.

Method

Participants. Thirty-two undergraduates at Yale University participated in the experiment in return for course credit and/or payment.

Design and Materials. Sixty photorealistically rendered 3D objects (Greebles) were generated with Alias Sketch! (Alias Research Inc., Toronto) on an Apple Macintosh. All Greebles have four protruding parts organized in approximately the same spatial configuration on a vertically-oriented central part. The set is organized orthogonally along two categorical dimensions, such that each Greeble is a member of one of two “genders” and one of five “families” (Figure 1). There are five central part shapes each defining one of the five families. The gender difference is defined by the orientation of the parts relative to the central part, either all pointing upward or downward. Although some of the parts are very similar to each other, every individual part is unique within the set.

From this set, 30 Greebles (3 individuals from each gender x family combination) were used during expertise training, while 24 unfamiliar Greebles (12 of each gender) were used in the novice-level and the expertise-level test phases. Nonsense words were used as names to designate the three kinds of parts, the two genders, the five families, and each individual. For purposes of expertise training, 10 Greebles (5 of each gender) were given individual names. For the novice-level and the expertise-level test phases, four sets (“Plok1, Plok2, Glip1, Glip2”) of six Greebles within the same gender category were crossed with four sets (“A, B, C, D”) of six novel names to produce four testing conditions: Plok1A-Glip1B, Plok2C-Glip2D, Glip1D-Plok1C, Glip2B-Plok2A. There were four experts and four novices tested for recognition of parts of Greebles in each of these four possible orders (for a total of 16 experts and 16 novices). For each of the 24 unfamiliar Greebles used in the test phases, two versions were generated, one with the top pair of parts in its original position and one with each of these two parts moved 15° around the vertical axis towards the front. Three distractors were created for each of these two versions, with one of the three kinds of parts replaced in each distractor by a foil part (drawn from within the same subset of 6 objects). Finally, three images were also created for each target, showing each target and foil part in isolation.

The experiment was performed on an Apple Macintosh LC 475 equipped with a Sony Trinitron 13” color monitor with a resolution of 640 x 480 pixels (72 pixels per inch). The Greebles were all the same purple shade, with an overhead light and a stippled texture. Images were about 6.5 x 6.5 cm and presented in the middle of the screen on a white background. Participants sat about 60 cm from the screen, yielding a display area subtending approximately 6.2° x 6.2° of visual angle.

Procedure. The experiment consisted of three phases: 1) testing of sensitivity to configural changes in novices, 2) expertise training, and, 3) testing of sensitivity to configural changes in experts. See Table 1 for a detailed description of the training and testing procedure. We now review the procedures used for novices and experts.

Participants who served as novices first learned the names of the 3 kinds of Greeble parts (from the top to the bottom of an object, *boges, quiff, dunth*, Figure 2a). No further training was given. Participants were then tested for forced-choice recognition of parts with upright and inverted Greebles (Figure 2b). For each of the two orientations, the names of six different Greebles were learned. Each name was shown for 1 sec in the middle of the screen followed by a Greeble that the participant could look at for as long as desired. Six Greebles were studied in this way six times each, in a random order for a total of 36 learning trials. Following this, forced-choice recognition of the parts was assessed. On each trial, a prompt was shown on the screen specifying one part of a particular target (e.g., “PIMO’S BOGES”) followed by two pictures side-by-side on the screen. Participants selected whether the right or left image contained the designated part by pressing one of two keys. There were three conditions randomized together: 1) Studied configuration: the two choices were the specified part and a foil part, both in the context of the Greeble specified in the prompt; 2) Transformed configuration: the two choices were of the specified part and a foil part, both in the context of the Greeble specified in the prompt but with the top parts

	Number of Trials	**Novices**	**Experts**
Learn generic names of Greeble parts and specific names for 6 upright Greebles	36	⊗	
Recognition of parts at upright in the Studied, Transformed, and Isolated conditions	54	⊗	
Learn to associate specific names with 6 *new* inverted Greebles	36	⊗	
Recognition of parts at inverted in the Studied, Transformed, and Isolated conditions	54	⊗	
Examples of the three levels of categorization	75	⊗	
Learn the names of first 5 individuals and blocks of 60 trials of a yes/no categorization paradigm for each level of categorization (gender, family, individual)	720		⊗
Learn the names of 5 more individuals and blocks of 60 trials of a yes/no categorization paradigm for each level of categorization (gender, family, individual)	360		⊗
Training Cycle			
Blocks of trials at the individual level	180		⊗
Blocks with the three levels randomized, (yes/no categorization task)	360		⊗
Until performance on the individual level is indistinguishable from one of the other two levels			
Learn generic names of Greeble parts and specific names for 6 *new* upright Greebles	36		⊗
Recognition of parts at upright in the Studied, Transformed, and Isolated conditions	54		⊗
Learn to associate specific names with 6 *new* inverted Greebles	36		⊗
Recognition of parts at inverted in the Studied, Transformed, and Isolated conditions	54		⊗

Table 1. Testing and training procedure for novices and experts at Greeble recognition.

moved 15° towards the front; 3) Isolated part: the two choices were of the specified part and a foil part, both in isolation on the screen. Following this testing with upright Greebles, six different Greebles were learned in an inverted orientation and the recognition of their parts was assessed with the Studied configuration, Transformed configuration, and Isolated part conditions using inverted Greebles.

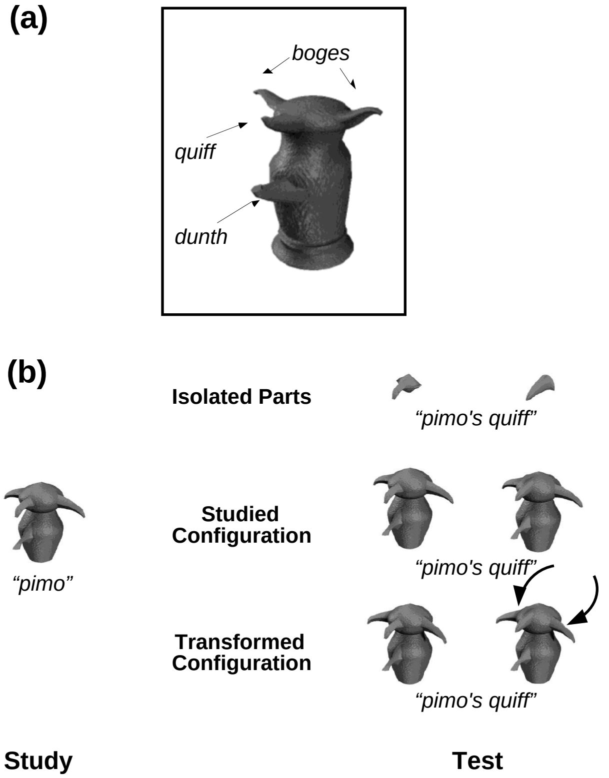


Figure 2. a) Novel names assigned to the Greeble parts. b) Example of the forced-choice recognition paradigm used to test novices and experts. Participants were shown a single Greeble at study and then were tested with pairs of images showing a part of the studied Greeble and a distractor part. Parts appeared in isolation, in the Studied configuration, or a Transformed configuration and participants judged whether the left or right image contained the specified part from the studied Greeble. Arrows indicate the stimulus changes in the Transformed configuration. Note that while the 15° rotation of the top parts is quite subtle, experts (but not novices) report noticing this change.

Participants who served as experts first went through extensive training to make them “experts” at Greeble recognition. They practiced recognizing 30 Greebles at three levels of categorization: the gender, family, and individual levels. Each of the 30 Greebles had a visually defined gender and family category while only ten of the objects were given individual names (the others were part of a “none-of-the-above” category at the individual level). Each category was taught to participants by showing a series of examples from that category followed by repeated blocks of 60 trials of a label-verification paradigm for each level of categorization. Each label-verification trial was initiated with a fixation cross in the middle of the screen for 500 ms, followed by a label shown for 1,000 ms designating a gender, family, or individual. After 250 ms, a Greeble replaced the label and it remained on the screen until the participant responded as to whether the Greeble matched the label. After an average of 6 runs at each level (60 trials per run), there was a cycle of two types of tasks: the first included 180 trials of practice at the individual level and the second included 360 trials divided into two blocks of 180 randomized trials, with 60 trials for each of the three levels of categorization. The large number of individual level trials in the first task provided more experience on the most difficult level and the second task allowed a comparison between the three levels when participants could not predict the level from one trial to the next.

To be considered experts, participants had to reach a pre-specified criterion during the mixed blocks. Comparisons were made on the three levels of categorization for the ten objects for which individual names were assigned. To reach the criterion, the average response time for individual-level recognition had to be statistically equivalent to the response time for at least one of the two other levels (measured by pairwise t-tests with individual alpha levels of .05). Experts reached the criterion after an average of 3,240 trials (ranging from 2,700 to 5,400) spread across a total of 7 to 10 one-hour sessions (Figure 3)[1]. After reaching the criterion, experts were tested for the recognition of parts of 12 new Greebles (6 upright, 6 inverted) in the identical procedure in which novices were tested (Studied, Transformed, Isolated).

Proportion correct and response times were analyzed with three-way ANOVAs including two within-subject and one between-subject factors: Orientation (Upright/Inverted) x Presentation Condition (Studied, Transformed, Isolated) x Expertise (Novice/Expert). Only response times for correct trials were analyzed and they were submitted to a log transformation before analysis (to normalize the typically skewed RT distribution). Mean RTs for all 12 cells of the design are shown in Table 2.

---

[1] Note that generic experience with the stimuli was found to be insufficient to develop expertise. The artist who created the Greeble set took just as long to reach the criterion of expertise as complete novices. This is not to say that the large amount of experience this person had with the Greebles did not result in some type of expert processing of this category, only that this knowledge did not transfer to the part recognition task. Similarly, expertise with faces is thought not to transfer to inverted faces (Yin, 1969; Moses, Ullman, & Edelman, 1996).

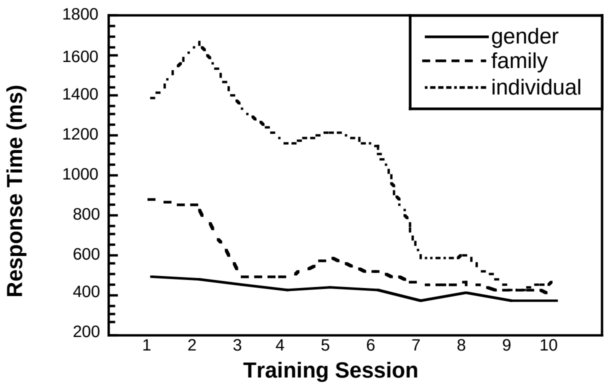


Figure 3. Expertise training. Example of the progression of response times for recognizing Greebles at the gender, family, and individual levels with increasing expertise. Data is shown for one participant because participants reached the criterion after different numbers of training sessions (see text for details regarding the criterion).

The ANOVA revealed that experts were reliably faster, $F(1,30) = 8.21$, $p < .01$, and marginally more accurate, $F(1,30) = 3.65$, $p = .06$, than novices; inverted Greebles were responded to reliably faster, $F(1,30) = 18.42$, $p < 0.001$, but were not more accurately recognized, $F < 1$, than upright Greebles; presentation conditions varied reliably from each other for both response time, $F(2,60) = 38.84$, $p < .001$, and accuracy, $F(2,60) = 9.07$, $p < .001$. The main effect of orientation on response time may be attributed to the fact that participants were always tested first with upright Greebles, and, thus may have the advantage of having practiced the forced-choice recognition task when they encountered inverted Greebles. Note, however, that these main effects do not address the crucial predictions of this study. Rather, these focus on the interaction analyses specifically comparing the two changed conditions, Isolated parts and Transformed configuration, to the Studied configuration condition, crossed with the level of expertise and the orientation of the stimuli. These comparisons, all significant according to Scheffé's post-hoc tests ($p < .05$), are presented next.

UPRIGHT GREEBLES

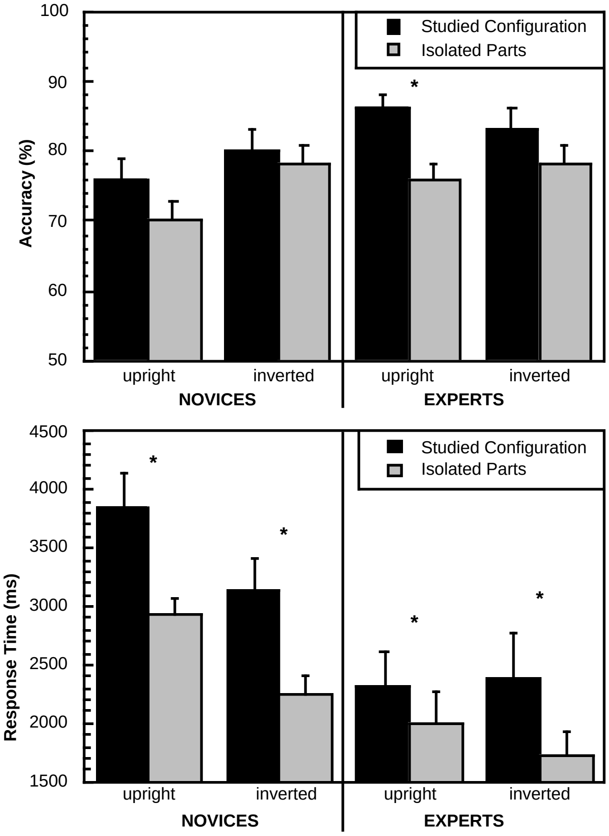
	top parts	middle part	bottom part	mean
NOVICES				
Transformed	2845 / 89	4255 / 79	3581 / 71	3560/80
Studied	3341 / 86	4354 / 71	3863 / 68	3853/76
Isolated parts	2835 / 78	3671 / 61	2262 / 72	2923/70
EXPERTS				
Transformed	2382 / 88	2855 / 85	2609 / 80	2695/86
Studied	2257 / 93	2472 / 90	2038 / 82	2306/87
Isolated parts	1670 / 87	2319 / 73	2125 / 73	1991/76

INVERTED GREEBLES

	top parts	middle part	bottom part	mean
NOVICES				
Transformed	2278 / 77	3331 / 75	3148 / 77	2919/76
Studied	2632 / 83	4024 / 77	2733 / 80	3129/80
Isolated parts	2270 / 82	2145 / 71	2286 / 80	2234/78
EXPERTS				
Transformed	1572 / 93	2394 / 90	1896 / 83	2204/85
Studied	1969 / 82	2829 / 85	1172 / 83	2382/83
Isolated parts	1443 / 80	1974 / 77	1422 / 77	1717/79

Table 2. Response times (ms) and percent correct for the recognition of the three types of parts for upright and inverted Greebles by novices and experts.

Isolated parts vs. Studied configuration. As shown in Figure 4, for novices, the Isolated parts and the Studied configuration conditions were not reliably different in terms of accuracy, but response times were reliably faster for the Isolated parts condition relative to the Studied configuration condition, presumably because there is considerably less information to process when the parts are presented in isolation. This response time advantage for the Isolated parts condition relative to the Studied configuration also holds for novices with inverted Greebles and for experts with both upright and inverted Greebles. Although response times were not reported in their paper, a similar pattern was also observed by Tanaka and Sengco (1996) for the recognition of parts of faces (J. Tanaka, personal communication).



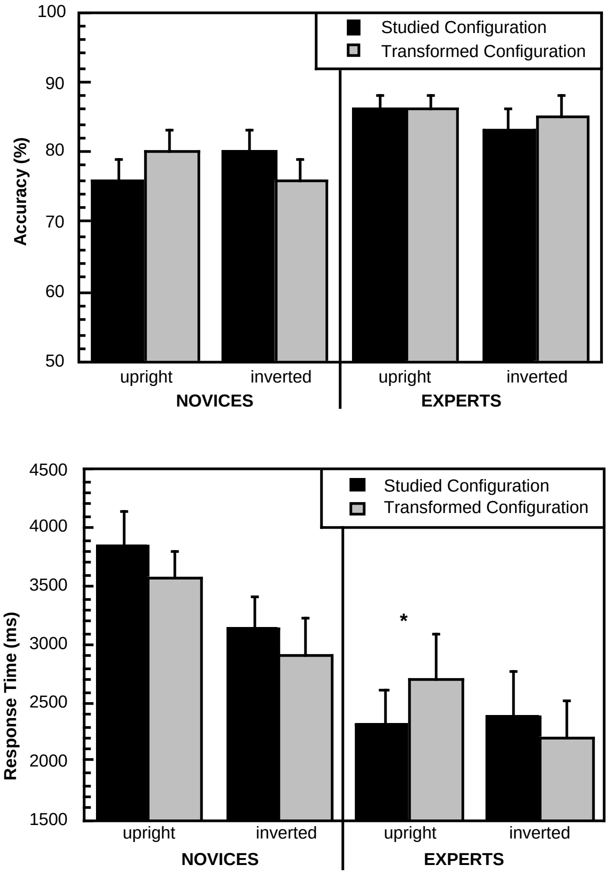
* p < 0.05, Scheffé's test

Figure 4. Accuracy and response times for correct trials in the part recognition test, for the Studied configuration and Isolated parts conditions. Results are reported for novices and experts with both upright and inverted Greebles. Error bars reflect the standard error between subjects, while the Scheffé tests are repeated measures.

Across both expertise level and stimulus orientation, the response time advantage for isolated parts manifests itself as a speed-accuracy tradeoff as participants were always faster and less accurate in the Isolated parts condition relative to the Studied configuration condition. However, the cost for experts with upright Greebles *cannot* be explained by this speed-accuracy tradeoff because the experts showed at least as large a response time difference between the Isolated parts and Studied configuration conditions with inverted Greebles as they showed with upright Greebles, yet the effect in accuracy was obtained only for upright Greebles. Moreover, there is no reliable increase in the Studied-Isolated difference between novices and experts. Finally, there is some hint that the Isolated-Studied difference may be in part due to the homogeneity of the Greeble set and the subtle part discrimination task, rather than to the level of expertise. In particular, although not reliable, the direction of the Isolated-Studied difference for accuracy is the same as for the other three groups (novices with both upright and inverted Greebles and experts with inverted Greebles) and this difference was consistent across the three types of Greeble parts (Table 2). Interestingly, this effect could be akin to the object-superiority effect obtained by Gyoba, Arimura, and Maruyama (1980) in which a learned perceptual schema can generate contextual expectations facilitating recognition. Supporting this argument, Tanaka et al. (1996) have recently reported that children as young as 6 years of age remember individual parts of faces better in the context of complete faces as compared to the same parts in isolation. This suggests that the object advantage may occur earlier than configural sensitivity during the process of acquiring perceptual expertise. In this context, the fact that experts did not show a reliable difference from novices is less surprising, since the Isolated-Studied contrast may test a different process than the Transformed-Studied contrast.

Transformed configuration vs. Studied configuration. As shown in Figure 5, for novices, the Transformed configuration and the Studied configuration conditions were not reliably different in terms of either accuracy or response times. For experts, however, response times to upright Greebles were reliably slower in the Transformed configuration condition relative to the Studied configuration condition. Crucially, this difference represents a qualitative change in the recognition behavior of experts -- in contrast, the accuracy difference obtained in the Isolated-Studied comparison for experts was only a change in magnitude -- thus, the preferred explanation here is that the expertise manipulation produced the speed advantage for the Studied configuration condition over the Transformed configuration condition. Supporting this interpretation, a two-factor ANOVA on log(RT) revealed a main effect for Expertise, $F(1,30) = 10.8$, $p < .005$, and a near-reliable interaction between the Expertise (novice/expert) and Condition (transformed/studied), $F(1,30) = 3.85$, $p = .059$. Also significant was the fact that the Transformed-Studied difference was consistent across the three types of Greeble parts (Table 2).

Based on informal debriefings following testing, none of the novices reported noticing the moved parts in the Transformed configuration condition. In contrast, some of the experts spontaneously reported that the top parts of some Greebles had been moved and all of the experts responded affirmatively when asked if they had noticed the transformation.



* $p < 0.05$, Scheffé's test

Figure 5. Accuracy and response times for correct trials in the part recognition test, for the Studied configuration and Transformed configuration conditions. Results are reported for novices and experts with both upright and inverted Greebles. Error bars reflect the standard error between subjects, while the Scheffé tests are repeated measures.

Discussion

Face processing shows disproportionate costs for configural changes (Tanaka & Farah, 1993). Although this "face-specific" effect has been interpreted as evidence for a face-exclusive mechanism, we wondered whether this pattern could be explained by a more general recognition mechanism fine-tuned by experience with homogeneous stimuli. We investigated this possibility by testing sensitivity to configural transformations for novices and experts with homogeneous non-face stimuli -- Greebles. Several findings stand out as relevant to the question of face-specific recognition mechanisms. First, our results suggest that the previously obtained object-superiority effect for faces holds for the recognition of parts taken from members of a visually homogeneous non-face object class. Greeble parts, in particular, were better recognized in the context of intact Greebles relative to the recognition of the same parts in isolation. This advantage was no different for experts as compared to novices and both groups showed a similar pattern of behavior with inverted Greebles. Thus, it seems that the visual properties of the objects and/or the task, rather than the level of expertise, were responsible for the difference. We also found a general response time advantage for isolated parts over the Studied configuration -- while this finding does not account for the accuracy difference displayed by experts with upright Greebles, it does suggest caution in interpreting the results of the part recognition paradigm in that response times are not typically reported (Tanaka & Sengco, 1991; Tanaka & Farah, 1993; Tanaka et al., 1996). In contrast, Tanaka and Farah (1993) did not find an object-superiority effect with either inverted or scrambled faces or houses, all sets of homogeneous objects. Our belief is that this discrepancy indicates an important advantage to using novel objects as control stimuli: inverted and scrambled faces are "wrong" versions of an overlearned stimuli, and the entire category of houses contains much more variation in the configuration of their features than do faces. Thus, prior experience of participants with the more typical instances of faces and houses could prevail over the experimentally-created proximal qualities of the stimuli, especially if the participants are not extensively trained on the modified versions of the stimuli.

Second, our results suggest that the training procedure rendered the experts more sensitive to a subtle change in the configuration of the parts, even when this change was performed on a part that they were instructed to ignore. In particular, experts recognized Greeble parts better in the Studied configuration as compared to Greeble parts in the Transformed configuration. What is not entirely clear is why our participants showed this sensitivity in response time while Tanaka's participants showed it in accuracy. Of course, psychophysical models rarely allow one to predict *a priori* whether a difference between conditions will manifest itself in one dependent measure or the other. Supporting our interpretation of this effect, however, is that the expert recognition of all three types of Greeble parts was sensitive to this transformation, in accordance to the findings with faces. This effect of configural information was not present in the novices' data, nor was it found for experts with inverted Greebles. Thus, it appears to represent a qualitative shift in recognition behavior produced by the expertise training.

These results offer some insights into the recognition patterns found for faces by Tanaka and his colleagues. In particular, they obtained an advantage for the Studied configuration of a face over both isolated parts of the face and a Transformed configuration of the face. Here, we dissociated these conditions with regard to their dependence on experience and found that sensitivity to these transformations was not specific to faces. It should be noted that the question of whether Greeble experts' sensitivity to configural changes is specific to the training orientation should be addressed more specifically in a design in which testing is counterbalanced across the upright and inverted conditions.

Conclusions

The present study shows how extensive practice with previously-novel non-face objects can lead to some of the recognition effects typically associated with faces. We found that expertise training changed novices, who were presumably processing Greebles with their “default” object recognition system, into experts, who were not only faster and more accurate but displayed a greater sensitivity to configural changes. This effect of expertise acquisition on the part recognition paradigm can be compared to Stroop interference (Stroop, 1935). Robust interference is found in the Stroop task when subjects have to name the color of incongruently colored color terms. This interference is due to the automaticity of reading that has been acquired over years of practice. In a similar fashion, the acquisition of Greeble expertise leads to interference from information that experts have learned to process automatically. This is demonstrated by the fact that our experts cannot ignore this more global information, even when it would be more efficient to do so (e.g., in the Transformed condition). In contrast to the Stroop effect, not much is known about the learning process that leads to face or Greeble expertise, nor can our experiment illuminate the particular features that are used by experts. The only evidence regarding this issue stems out of studies on the features used for face recognition, for instance Rhodes (1988) reported that both first-order (e.g., the appearance of the parts) and second-order features (e.g., the spatial relations between the parts), as well as global inferred features such as age and weight, appear to be encoded in face representations. While novices may rely on first-order features, expertise acquisition may lead them to use second-order features and even perhaps higher-order features.[2] The similarities of the pattern obtained here for Greeble part recognition to that obtained for recognition of face parts suggests that Greeble experts employed mechanisms similar to those implicated in face recognition. Assuming this to be the case, an important question is: Did training lead novices to abruptly switch from one type of processing to another, or did a more continuous shift of the type of processing occur?

Consideration of single-cell recording work with monkeys suggests a speculative but intriguing possibility. First, Perrett and Oram (1993) suggested that the configural sensitivity found for some “face cells” -- temporal lobe neurons selectively activated by faces -- could be produced by a combination of inputs most selective for complex assembled features. For example, cells responsive to two eyes side-by-side or a nose above a mouth could be combined to produce a sensitivity to the overall face configuration. Second, K. Tanaka et al. (1996), working with anesthetized monkeys, have recently investigated the minimal stimulus features necessary and sufficient to activate individual neurons in infero-temporal (IT) cortex. They have found that the critical features of these cells are moderately complex (e.g., an eight point star shaped pattern or a green square above a red circle) and may be thought of as an “alphabet” of features that could be combined to code complex objects. It is possible that the complex features for which IT cells appear to be selective are not fixed but can be modified as the result of structured experience such as expertise at subtler levels of discrimination. Indeed, Logothetis and Pauls (1995) have demonstrated that IT neurons can become highly selective for previously novel stimuli. In our experiment, expertise training may have led to the assembly of complex feature-detectors, extracted from the statistical properties of the Greeble set that proved useful for performing the training discriminations (for example, the orientation of the Boges is diagnostic for distinguishing between the two genders). Such a system could presumably make use of the recurrent spatial configuration across the set and of the probabilities of co-occurrence for parts and contours of different Greebles (e.g., for a similar statistical approach to object representation, see Edelman, 1995). For instance, there would be no need to represent the Boges of a Greeble separately since they always occurred in redundant pairs (much as eyes or halves of a face). If expertise is a result of a large proportion of cells becoming selectively tuned to multiple parts that frequently co-occur, then experts would be expected to show a cost for the recognition of parts in isolation or in a Transformed configuration. In accordance with this idea, there is some evidence that categorization tasks with novel objects can lead to the creation of new perceptual features, that is, assemblies of parts that were diagnostic for the required categorization judgment (Schyns & Murphy, 1991; 1994). Moreover, in the case of both faces and other objects, these temporal lobe visual “feature detectors” have been found to be viewpoint-dependent (Logothetis & Pauls, 1995; Miyashita & Chang, 1988; Perrett & Oram,1993). If the configural cues acquired during expertise are indeed mediated by associations between and tuning of these cells, degradation of expert performance with orientation changes should be expected, as was found here.

In summary, we hypothesized that the putatively face-specific sensitivity to configural changes might be explained by a more general recognition mechanism fined-tuned by experience with homogeneous stimuli. The present results with Greebles provide some evidence that this is indeed the case -- experts showed greater sensitivity to a change in a studied Greeble

---

[2] Interestingly, several recent models suggest that the perceptual system may be tuned in a similar manner based on experience -- in particular, in terms of the self-organization that may occur in early vision (Field, 1994; Weiss & Edelman, 1995).

configuration than did novices. These results suggest that expertise at discriminating between visually similar objects, such as Greebles *or* faces, produces the obtained sensitivity to configural transformations. More generally, we believe that such results illuminate the point that visual representations and mechanisms are not steady states and, as such, it is essential to consider how they change with experience. As Johnson and Morton (1991) have argued in their work on infants' face recognition, only a combination of both the cognitive and the biological perspectives can provide an answer to this fascinating question.

REFERENCES

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

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. (1995). Representation, similarity, and the chorus of prototypes. *Minds and Machines*, *5*, 45-68.

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

Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation, 6*, 559-601.

Gyoba, J., Arimura, M., & Maruyama, K. (1980). Visual identification of line segments embedded in human face patterns. *Tohoku Psychologica Folia*, *39*, 113-120.

Johnson, M. H., & Morton, J. (1991). *Biology and Cognitive Development: The Case of Face Recognition*. Oxford, UK: Blackwell.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representation in the primate. *Cerebral Cortex*, *3*, 270-288.

Miyashita, Y., & Chang, H. S. (1988). Neural correlate of pictorial short-term memory in the primate temporal cortex, *Nature*, 331; 68-70.

Perrett, D. I., & Oram, M. W. (1993). Neurophysiology of shape processing. *Image and Vision Computing*, *11*, 317-333.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, *17*, 43-63.

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

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.

Schyns, P. G., & Murphy, G. L. (1991). The ontongeny of units in object categories. In *The XIII Meeting of the Cognitive Science Society* (pp. 197-202). Hillsdale, NJ: Lawrence Erlbaum.

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. (1988). Face perception and the right hemisphere. In L. Weiskrantz (Eds.), *Thought without language* (pp. 108-313). New York: Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *12*, 242-248.

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

Tanaka, J. W., & Sengco, J. A. (1996). Features and their configuration in face recognition. *Unpublished manuscript*, Oberlin College.

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, J. W., Kay, J. B., Grinnell, E., & Stansfield, B. (1996). Face recognition in young children: When the whole is greater than the sum of its parts. *Unpublished manuscript*, Oberlin College.

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

Weiss, Y., & Edelman, S. (1995). Representation of similarity as a goal of early visual processing, *Network, 6*, 19-41.

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