REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL 15 Jan 92 TO 14 Jan 96

4. TITLE AND SUBTITLE

REPRESENTATIONS OF SHAPE IN OBJECT RECOGNITION AND LONG-TERM VISUAL MEMORY

5. FUNDING NUMBERS

F49620-92-J-0169

61102F
2313/AS
2313/BS

6. AUTHOR(S)

Michael J. Tarr

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Yale University
Dept of Psychology
P.O. Box 208205
New Haven CT 06520-8205

8. PERFORMING ORGANIZATION REPORT NUMBER

AFOSR-TR-96

0303

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NL
110 Duncan Ave Suite B115
Bolling AFB DC 20332-8080

Dr John F. Tangney

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

19960624 319

13. ABSTRACT (Maximum 200 words)

Our research has focused on the mechanisms used in visual object recognition. Recent psychophysical results suggest that human perceivers often rely on viewpoint-specific (view-based) representations in conjunction with normalization procedures. Over the past year we have explored the degree to which view-based representations are also appearance based. Specifically we have found that visual recognition across many tasks is sensitive to changes in image properties such as illumination, color, and material (texture). These results indicate that object representations are information rich and that abstract part-based structural-descriptions will not account for much of human recognition performance. Other work has focused on how view-based representations are organized and function across changes in viewpoint. Using 3D stimuli rotated in depth we have investigated the role of task, ranging from basic-level to subordinate-level discriminations, how view-based representations generalize from known members of a class to unfamiliar members of that class, and how perceptual expertise is acquired and influences recognition strategies. We have also been investigating the mechanisms used for discriminating between highly similar objects, e.g., faces. In particular, we are interested in how such exemplar-specific

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(U)

processes relate to/interact with more generic recognition mechanisms, as well as how such processes are affected by distortions or noise. Finally, we have been using these and other results to develop a new exemplar-based approach to object recognition/representation. This model assumes three types of associations between input: 1) Across-exemplar associations at the level of coarse features of bounding contours; 2) Across-view associations between qualitatively distinct viewpoints for a specific exemplar; 3) Within-exemplar associations at the level of fine features throughout the image. Crucially, associations are relatively weak at initially encoding, but become reinforced through experience. We are currently investigating the degree to which such a model may account for a range of behavioral phenomena, including different categorical levels of recognition and variation in image feature sensitivity with different tasks.

FINAL TECHNICAL REPORT.

AFOSR/NL Grant No. F49620-92-J-0169

REPRESENTATIONS OF SHAPE IN OBJECT RECOGNITION AND LONG-TERM VISUAL MEMORY

Michael J. Tarr
Brown University
Department of Cognitive and Linguistic Sciences
Box 1978
Providence, RI 02912

June 5, 1996

Final Technical Report for Period 15 January 1992 - 14 January 1996

Distribution Statement

Prepared for

Dr. John F. Tangney (Program Manager)
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH/NL
110 Duncan Avenue Suite B115
Bolling AFB, DC 20332-0001

2. Objectives

Objectives remain the same as originally stated. We are investigating the representations and mechanisms underlying human object recognition using both psychophysical and computational methods.

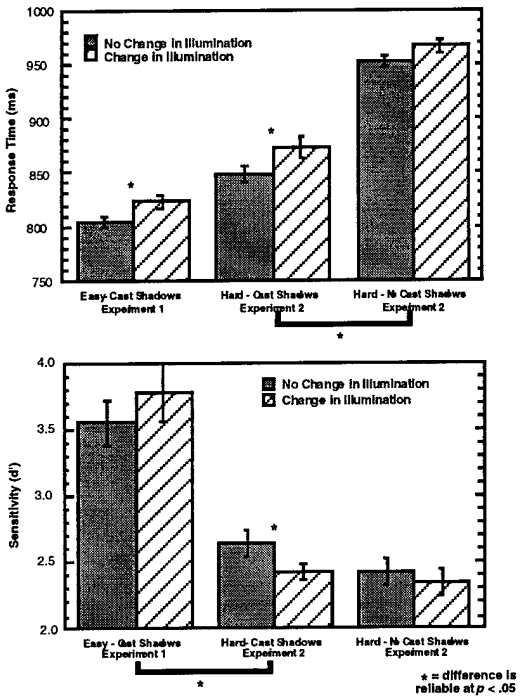
3. Status of Effort

Our research has focused on the mechanisms used in visual object recognition. Recent psychophysical results suggest that human perceivers often rely on viewpoint-specific (view-based) representations in conjunction with normalization procedures. Over the past year we have explored the degree to which view-based representations are also appearance based. Specifically we have found that visual recognition across many tasks is sensitive to changes in image properties such as illumination, color, and material (texture). These results indicate that object representations are information rich and that abstract part-based structural-descriptions will not account for much of human recognition performance. Other work has focused on how view-based representations are organized and function across changes in viewpoint. Using 3D stimuli rotated in depth we have investigated the role of task, ranging from basic-level to subordinate-level discriminations, how view-based representations generalize from known members of a class to unfamiliar members of that class, and how perceptual expertise is acquired and influences recognition strategies. We have also been investigating the mechanisms used for discriminating between highly similar objects, e.g., faces. In particular, we are interested in how such exemplar-specific processes relate to/interact with more generic recognition mechanisms, as well as how such processes are affected by distortions or noise. Finally, we have been using these and other results to develop a new exemplar-based approach to object recognition/representation. This model assumes three types of associations between input: 1) Across-exemplar associations at the level of coarse features of bounding contours; 2) Across-view associations between qualitatively distinct viewpoints for a specific exemplar; 3) Within-exemplar associations at the level of fine features throughout the image. Crucially, associations are relatively weak at initially encoding, but become reinforced through experience. We are currently investigating the degree to which such a model may account for a range of behavioral

phenomena, including different categorical levels of recognition and variation in image feature sensitivity with different tasks.

4. Accomplishments/New Findings

• *The Role of Illumination and Texture in Object Representations*

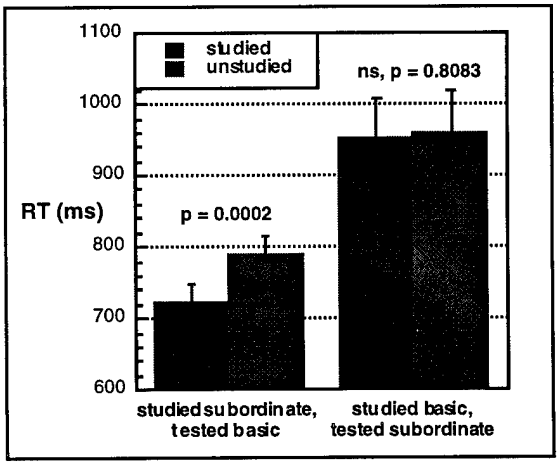


In collaboration with Daniel Kersten (University of Minnesota), we have been investigating the role of illumination and texture in object recognition. Using a variety of tasks (e.g., sequential matching and naming) we have found that visual recognition performance is sensitive to changes in image properties such as illumination and material. Specifically, the effects of illumination are often considered to be discounted early in visual processing — to test this we manipulated whether the same-shaped objects appeared in the same or in a different illumination in a matching task. Results, as shown in the adjacent figure, show that observers are sensitive to changes in illumination in judging two images to be from the same object (in terms of a 20 ms cost in response time) and that this sensitivity increases as the discrimination task becomes more difficult (in terms of a decrease in sensitivity as measured by d'). However, we find this cost for a change in illumination only when cast shadows are present on each object -- without cast shadows there is no reliable cost for a change in illumination. Interestingly, the absence of cast shadows produces a second effect -- there is a large increase in overall response times to make the shape discrimination. Together, these results suggest that a) object representations are image based, encoding the effects of illumination; b) image-based mechanisms become increasingly prominent as shape discriminations become more subtle (e.g., between different models of planes); c) one reason why the effects of illumination are encoded may have to do with the benefit they provide in terms of disambiguating novel shapes. That is, while there is a small cost in terms of illumination invariance, this is far outweighed by the potential benefit in terms of

information about the shape of the object. Further experiments have been exploring these effects for familiar objects, specifically human faces. Here it appears that there is a larger cost for a change in illumination (50 ms), but that the absence of shadows does not lead to slower response times. Thus, shadows may provide maximal information for learning about unfamiliar shapes, but less of a benefit one shapes are learned and therefore consistent with a known shape model.

In related work, we have been assessing the degree to which object representations are material sensitive — the assumption being that along with shape, material is a fundamental property of object representations. In our initial experiments we have manipulated material primarily through surface texture creating both material-consistent and material-inconsistent objects (e.g., a furry bunny versus a steel bunny). First results are quite promising — in a naming priming task, we have found large and significant priming for naming an object in the same material versus changing the material. These effects are comparable to those found for changes in viewpoint and much larger than those found for a range of other stimulus transformations. Such findings are being explored in further experiments that manipulate the diagnosticity of the material, the category of the object, and the task. Overall, these results provide some indication that shape-based approaches to recognition may be significantly enhanced by the inclusion of surface properties such as texture or material.

• *Task-Dependent Recognition Strategies*

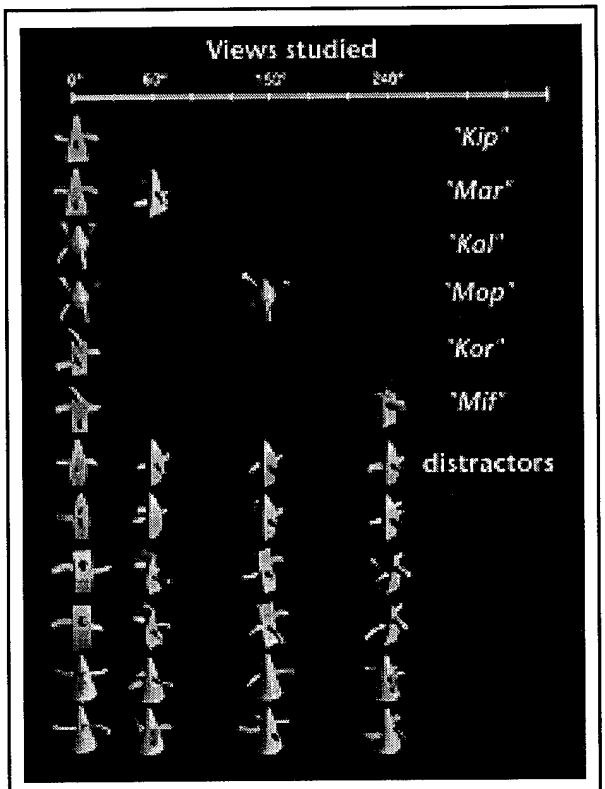


We have been exploring the impact of different recognition tasks, subordinate-level to basic-level, on strategies used for successful recognition. In our first experiments we are investigating whether the representations of familiar objects (planes, cars, animals, etc.) activated during one type of task generalize to another type of task. Using a name-matching paradigm, we are manipulating whether the match occurs at the basic or subordinate level. Although there is some reason to suppose that

representations activated during subordinate-level tasks will prime basic-level tasks, many theories of recognition assume that basic-level tasks access only abstract/coarse representations that are insufficient for subordinate-level recognition. However, we have been developing a framework for recognition in which the same neural units support both levels of recognition — in such an approach, regardless of the task at hand, encountering a given exemplar of a class will prime both the exemplar itself and the class (the primary reason for this is that basic-level representations are not considered abstractions, but rather the pooled output of the exemplars within that basic-level class). Initial results suggest that subordinate-level access does prime basic-level access, but not vice-versa (see figure). This finding indicates that the representations used for subordinate-level recognition are not separable from those used for basic-level categorization, and, in particular, object recognition may be thought of as activating a pool of exemplars which reach a recognition threshold more rapidly for the more generic basic level than for the specific subordinate level.

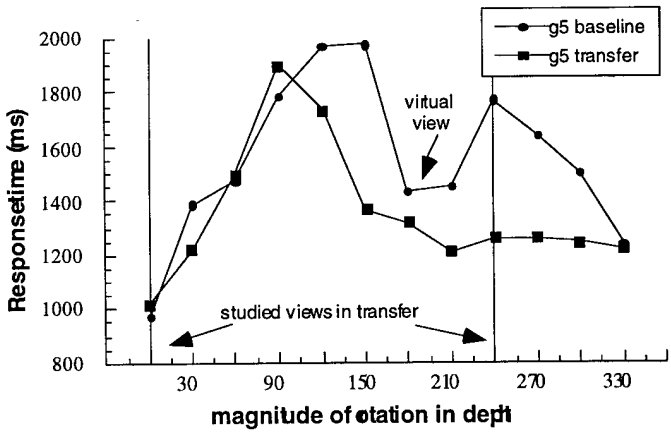
• *View-based Class Generalization*

Image-based approaches to object representation have been criticized as "template-like" and inadequate for generalizing across different exemplars of a familiar class. In particular, it has been suggested that they rely on rigid views that are specific to image features and attributes. We investigated this issue in the recognition of exemplars from novel 3D object categories. The experiment was designed to test whether, in the context of viewpoint-specific recognition, view-specific information for some exemplars would generalize to similar views of other members of the same basic-level class. Such generalization would be useful when new exemplars of a familiar category are seen for the first time. For instance, it may be possible to infer what a new model of car will look like from an unseen viewpoint based on the experienced view of the new exemplar and the knowledge of cars encountered in the past.



Objects used in the transfer Experiment. Subjects in the Transfer group practiced recognizing all shown viewpoints, while subjects in the Baseline group only practiced the 0° view.

Methodology. Six targets and six distractors were used, organized in six pairs of objects sharing the same central part, but with a different arrangement of the smaller parts (see figure). Subjects in the *Transfer* group practiced recognizing the objects from a small number of viewpoints generated by rotations in depth around the vertical axis. Half of the targets (one of each pair) were presented at 0° and at another orientation (60°, 150° or 240°, clockwise). The other half of the targets appeared as frequently but only at 0°. Distractors appeared in all mentioned orientations. For the *Baseline* group, the procedure was identical except that the targets were only practiced at 0°. The practice sessions were followed by a "surprise" block which consisted for both groups in the randomized presentation of all targets and distractors at 12 possible orientations.



Response times for correct responses at recognition of object g5 in the surprise block, for the Transfer and Baseline groups.

Results. For all targets showing an orientation effect (4 of 6), there was evidence for "virtual views" -- that is, generalization from the studied 0° view to the 180° view. In other words, targets that were practiced only at 0° or at 0° and another orientation (60° or 240°) were recognized faster than would be predicted from the studied views only. Furthermore, for the two pairs of objects for which orientation effects were found, comparison

between the Transfer and the Baseline groups provided strong evidence for generalization of studied views to other exemplars. An example is shown in the figure of response times, for recognition of object g5 ("Kip") during the surprise block. Again, the only difference between the two groups is that subjects in the Transfer group studied another pairwise similar object, gp5 ("Mar"), at 240° in the practice sessions. This experience generalized to object g5, even though subjects had to distinguish between the two objects.

Such results indicate that viewpoint-specific representations not only support discrimination between visually similar exemplars, but also allow generalization across exemplars, in particular, between visually similar views of the same objects and visually similar exemplars.

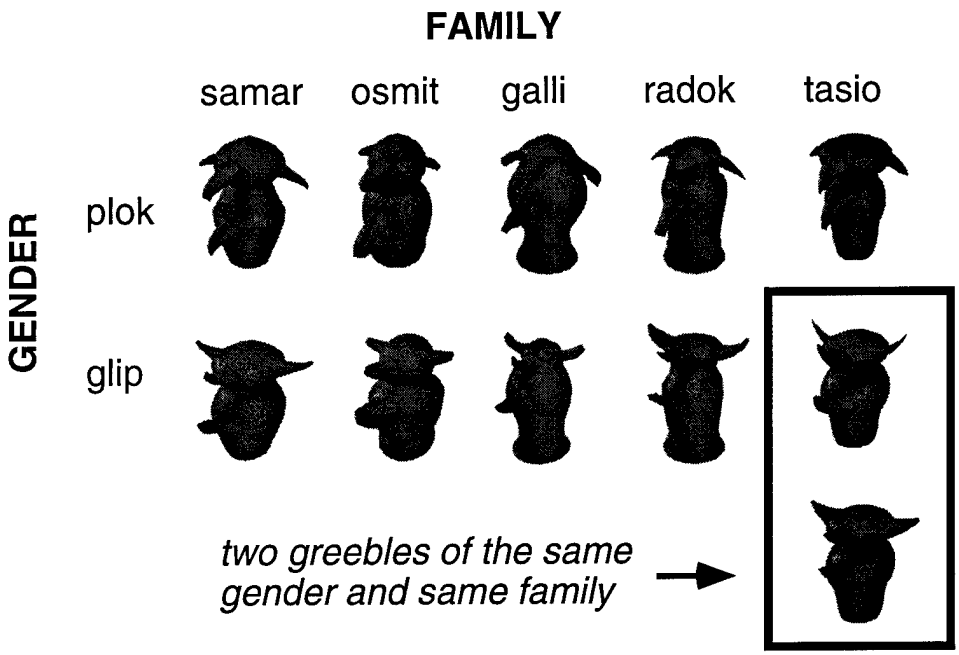
• *Face recognition*

Past research has demonstrated that inverted faces appear to be more difficult to recognize than a variety of other inverted objects, such as houses, airplanes, or landscapes. These results have led some to claim that there is a dedicated mechanism for processing faces that is disproportionately affected by inversion. We have run several experiments that test the claim that large inversion effects for recognition are specific to faces. The experiments here involve systematically rotating faces within the picture plane while measuring recognition performance. One study found that performance at correct face recognition is linearly related to the angular disparity of the face from the canonical upright orientation. By directly comparing faces to appropriate control stimuli, a second experiment revealed that recognition performance is the same for misoriented faces and misoriented control objects. Finally, we found that the difficulty recognizing inverted faces can be explained by a generic viewed-based model of representation which is not face-specific. These results suggest that faces, as a stimulus class, show a strong inversion effect because they share a single, vertical, highly canonical orientation, as well as a highly homogeneous configuration.

• *Perceptual expertise*

Until recently, the easiest way to study the role of expertise in object recognition was to do correlational studies involving novices and extant experts for a given class, for instance

dogs or birds. Other techniques included correlating age with face recognition abilities in developmental studies. The expertise training studies that we conducted demonstrates that experts may be created in a relatively short time span (10 hours), in particular, demonstrating some behavioral effects only previously obtained with faces.



In our first experiment, we demonstrated the possibility of creating "experts" for the recognition of novel non-face stimuli ("Greebles", see figure) in a relatively brief period of extensive training. We tested whether experts for this class of non-face stimuli were sensitive to configural changes much as subjects have been found to be with upright human faces but not control stimuli. Configural sensitivity was tested in a paradigm used with faces by Tanaka & Sengco (1996). They found that subjects were better at forced-choice recognition of face features in the context of the intact face than when they were presented in isolation or in the context of the original face but with other features displaced (e.g., eyes moved apart). Tanaka and Sengco also tested recognition of parts of houses, inverted or scrambled faces and did not find a similar sensitivity to changes in configuration. Interestingly, we found that experts (but not novices) at Greeble recognition demonstrated such a sensitivity to configural changes. That is, Greeble parts were better recognized in the context of intact Greebles relative to the recognition of the same parts in isolation. Moreover, experts recognized Greeble parts better in a studied configuration as compared to Greeble parts in a transformed configuration (in which the top parts were rotated 15° each towards the front).

In this first study, subjects 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) in a label-verification paradigm. To be considered experts, participants had to become as fast on the individual level as they were on the more categorical levels. 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.

In a second expertise study conducted at Oberlin College in collaboration with Prof. James Tanaka, we attempted to improve the expertise training procedure. The new training required the subjects to learn more individuals and used a naming task in addition to the label-verification task, in order to increase the difficulty level and diversify the experts' experience with the Greebles. Also, all experts were trained for a fixed amount of time (about 9 1/2 hours). Experts' performance reached an asymptote after about 6 hours of training. This indicates that a shorter training procedure, less costly and less laborious for the subjects, could be used in further experiments.

Study 2 also included other measures previously used with face stimuli. In most cases, effects found with faces (but not other categories) were replicated with Greeble experts. In particular, Greeble experts' part recognition is sensitive to contrast inversion just as it appears to be to sensitive to configural changes. Moreover, Greeble experts recognized composites made of parts of different Greebles more slowly when the separable parts are arranged in a valid Greeble configuration. Finally, expertise appears to be orientation specific in that Greeble experts are not sensitive to configural changes when they are tested with inverted Greebles (Study 1) and produce orientation-sensitive patterns for naming that increase with increasing misorientation. Thus, experts are only experts at the orientations for which they are trained -- for unfamiliar orientations, they effectively perform as novices.

• *Orientation Priming*

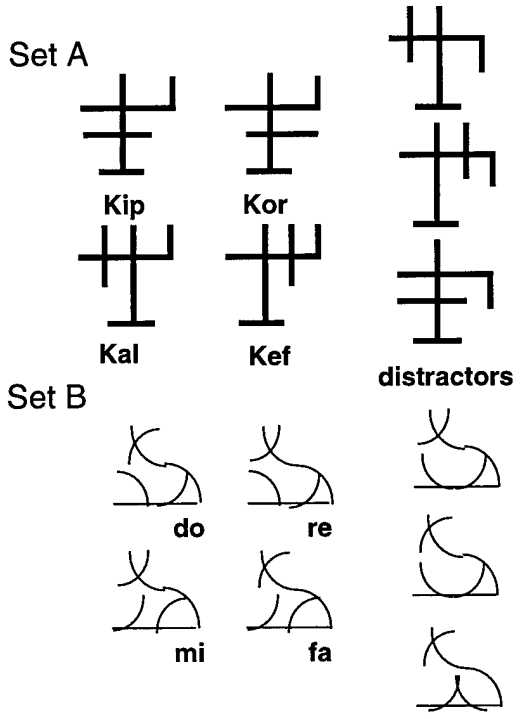
Orientation-priming across objects has often been taken as evidence for orientation-invariant representations because it may imply that orientation and shape information can be encoded separately. In this context, the explanatory power of a view-based framework would be increased by demonstrating orientation priming across objects in

the context of orientation-dependent recognition. We conducted a series of experiments exploring the possibility of orientation priming in identification judgments of novel stimuli. In all experiments, subjects were first trained to identify homogeneous 2D novel stimuli and later tested for recognition of those targets among distractors in different picture-plane orientations. Of interest was whether orientation cues could reduce the robust orientation generally found with such stimuli.

Cooper and Shepard (1973) found that prior orientation information, in the absence of shape information, does not reduce the orientation effect for a standard/mirror version judgment on letters. The results of Experiment 1 replicated that finding for an identification judgment of 2D novel objects stimuli (first set in figure), when the orientation cue was an arrow preceding each object.

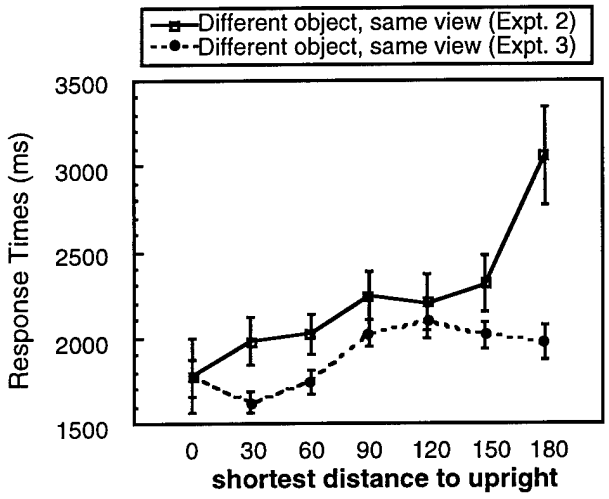
Experiment 2 tested whether the orientation information would be more useful if given in the form of a similar target in the same orientation, when both targets have to be identified. We examined sequential effects, dividing the trials according to the characteristics of the preceding trial: a target could be preceded by the same object in the same view (SoSv), a different object in the same view (DoSv), the same object in a different view (SoDv) or a different object in a different view (DoDv). Experiment 2 revealed that targets are not efficient orientation primes when there is no overall contingency between their orientation information and the orientation of the next stimulus.

Experiment 3 tested an "image-based" account for our earlier results -- that orientation information in the form of a similar target reduces the orientation effect in a context where the orientation information is maximally predictive of the incoming stimulus orientation (because of the activation of a class-general representation during earlier recognition). Such a situation was achieved by



Two sets of homogeneous objects used in the orientation priming experiments.

a *blocking* manipulation, in which trials were organized in 12 blocks of 15 trials, each block at one of twelve orientations in the picture plane (0°, 30°, 60°..., 330°). The orientation effect for DoSv trials in this blocked procedure led to a significant reduction of the orientation effect, especially at large orientations, compared to homologous trials from Experiment 2 (see figure).



Response times for correct responses in Experiments 2 (orientation random) and 3, (orientation blocked) ,for trials preceded by a different object in the same viewpoint.

Experiment 4 tested the hypothesis that the orientation priming obtained by blocking trials by orientation does not rely on subjects being aware of the blocking. The orientation priming by blocking was replicated in a paradigm using short blocks of trials (6 trials) intermixed with random trials, leaving the subjects unaware of the blocking. This result indicates that an automatic mechanism is most likely responsible for the orientation priming. Moreover, orientation priming was found to be linearly related to the serial position within the block, providing further support for an image-based account.

The results from Experiments 1 to 4 do not address whether the prime is necessarily a visually similar object or simply another object which is recognized at the same orientation. To test this possibility, we chose a design in which targets from two basic-level categories are alternated so that the prime is always *from a different category than the target.* An image-based account predicts no orientation-priming in this case, while if normalization of a frame of reference is mediating the orientation-priming, priming should also occur in those conditions. Subjects learned the names of eight objects organized in two homogeneous categories (both sets in the earlier figure) and were tested for their identification in trials for which orientation was either blocked or randomized. In both conditions targets of the two classes were constrained to alternate (A-B-A-B...). The results did not produce any evidence for a reduction of the orientation

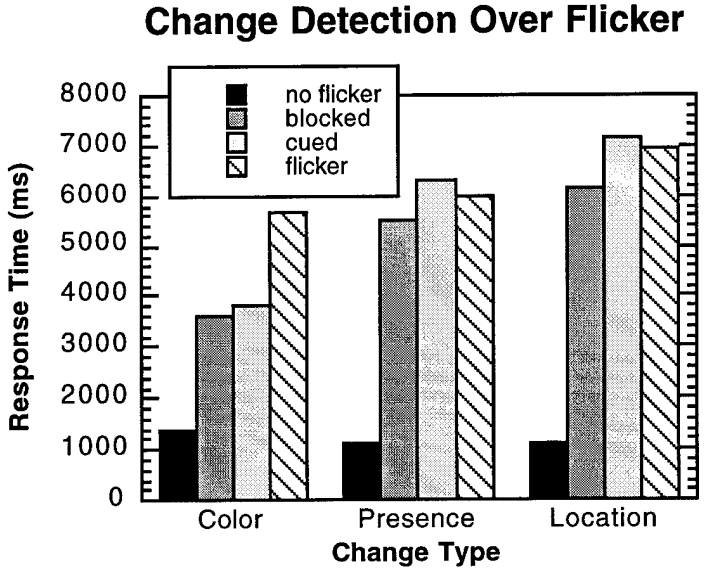
effect in the blocking manipulation when targets alternated between two different categories, despite the fact that subjects were explicitly told of the blocking manipulation. Contrasting this to the orientation priming found in Experiment 4 when subjects were unaware of the orientation blocking and the blocks only 6 trials long, there seems to be strong evidence for orientation priming being an automatic by-product of the activation of orientation-specific representations accessed during recognition.

• *Scene Perception*

A series of scene perception experiments were conducted using realistic photographs of natural scenes. Our purpose was to determine which object features and relations are most important for scene perception and how scene representations differ from those of individual objects. We used the "flicker" paradigm introduced by Rensink, O'Regan, and Clark (1995) in which one element of a scene continuously alternates between its original appearance and a noticeably changed appearance. A brief blank field disrupts the scene at the moment that the change occurs and presumably delocalizes transients produced by the change which may otherwise draw the perceiver's attention. The crucial assumption of this technique is that visual properties that are more salient within the representation of the scene are preserved across such transients and, therefore, are easier to detect when changed. Within this paradigm we also examined the degree to which cueing particular sets of features influenced detection of changes within a scene. Our assumption was that cueing would differentially enhance detection of changes in features that were typically less salient in visual memory.

So far we have completed one project consisting of four different experiments. The first was a replication of the original flicker study which showed that changes were detected much faster when they occurred in foreground objects than when they involved the background of a scene. The second experiment was a control in which we measured the detection of changes without the presence of flicker. In the last two experiments we used two different cueing manipulations to investigate the role of attention and expectations in scene perception. In one experiment scenes were blocked according to the type of change (color, location, or presence) that occurred. In the other,

scenes containing each type of change were randomly intermixed, but a cue informing participants as to the type of change was provided prior to each trial.

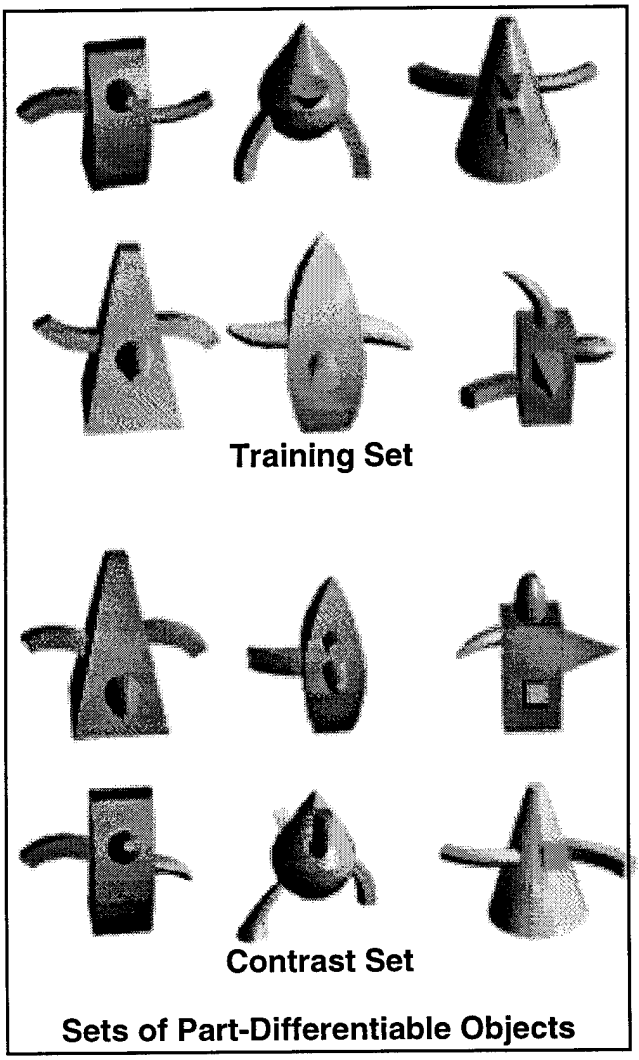


In addition to confirming the foreground/background effect, several other results stand out. First, regardless of the type of change, changes were detected significantly faster when there was no flicker. Second, cueing made no difference in the detection of changes in foreground objects. Third, cueing facilitated the detection of color changes in background information, but not location or presence changes. Finally, there was no difference between the degree of facilitation obtained for blocked vs. trial-by-trial cueing. These results suggest that attention is initially directed towards elements of the scene that perceivers consider to be informative or interesting, e.g., foreground objects. Moreover, it is likely that such elements are represented with the greatest salience in visual memory. There is apparently less salience in the representation of background information, and based on the cueing advantage, even less salience in the representation of color information in the background. Planned follow-up experiments will use computer graphics psychophysics to create both synthetic familiar and nonsense scenes, thereby allowing us to manipulate properties such as context, familiarity, viewpoint, texture, or 3D location.

• Concurrent encoding of viewpoint-dependent and viewpoint-invariant object representations

Current theories of object recognition have posited both viewpoint-dependent and viewpoint-invariant modes of object representation. However, it is still unclear as to what conditions determine how perceptual mechanisms apply such representations under different contexts in learning and recognition. We have completed a project in which we

have demonstrated that regardless of the role of viewpoint during initial encoding, subjects apparently encode *both* types of representations. Specifically, subjects were initially taught a set of objects, the training set, that could be immediately recognized equally well at all viewpoints: in one case 2D line drawings similar to those used in Tarr and Pinker (1990) and in the other case 3D part-differentiable objects (where a small number of qualitatively different parts is sufficient to discriminate one object from all others in the set).

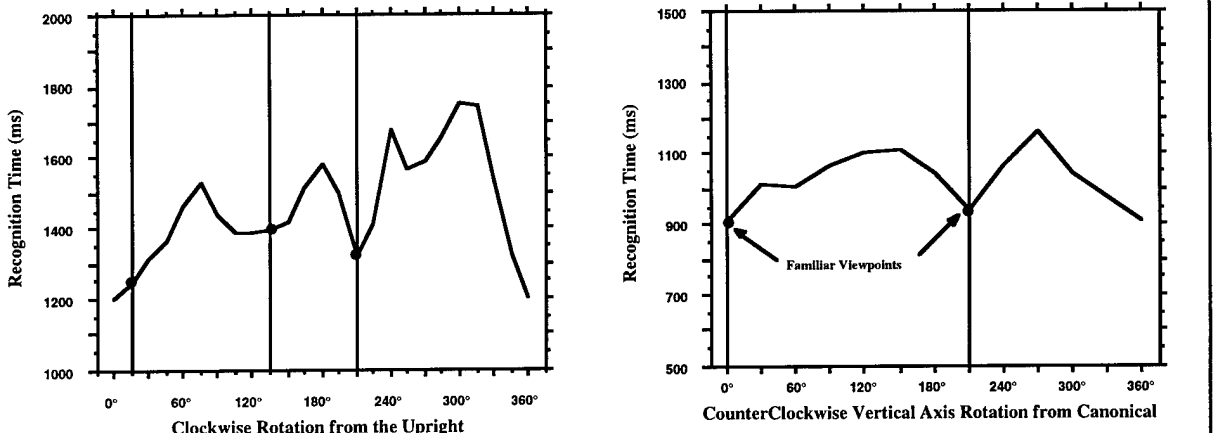


Sets of Part-Differentiable Objects

After familiarization, subjects were given extensive practice recognizing the objects from a select set of viewpoints generated by rotations in the image-plane or in depth (depending on the stimulus set). As predicted, in both instances, recognition performance was immediately equivalent at all tested viewpoints, indicating that viewpoint-invariant mechanisms and representations were employed during this phase. Following practice at recognition across several days, subjects were taught an equivalent number of new objects, referred to as the contrast set. The critical manipulation is that combined with the objects in the training set, no single object could be differentiated by a qualitative description of parts (as in Biederman's, recognition-by-components theory, Biederman & Gerhardstein, 1993) or by simple one-dimensional ordering of parts (see Tarr & Pinker, 1990). To assess the impact of including these new objects, additional unfamiliar viewpoints were also added during this phase. Two crucial predictions were made: (1) introducing the contrast set would result in a shift to viewpoint-dependent recognition

mechanisms; (2) viewpoint-dependent effects would be systematically related to the nearest previously seen viewpoint despite the previous lack of effects of viewpoint.

As shown in the two graphs below, in the final phase of each experiment, both predictions were obtained. For both 2D rotated in the image-plane and 3D objects rotated in depth, there is now a significant effect of viewpoint on naming time. Crucially, this pattern is systematic to the nearest familiar viewpoint, indicating that subjects did encode a viewpoint-specific object representation at each observed viewpoint.



Recognition of Familiar Objects Following Training in Select Viewpoints

A control experiment verified that these viewpoint-dependent effects are not simply due to the addition of viewpoints and objects. This study employed the identical 2D training set used in the previous experiment, but employed a contrast set that did not require subjects to rely on complex part relations across more than a single dimension. Under such conditions it was predicted that, despite the introduction of new viewpoints and objects, viewpoint-dependent patterns would not be obtained. Results for the familiar training objects confirmed this: no systematic pattern of response times across orientation was observed. Overall these results indicate that there is no "default" recognition mechanism. Rather the visual system apparently encodes at least two distinct types of object representations, one viewpoint invariant and one viewpoint

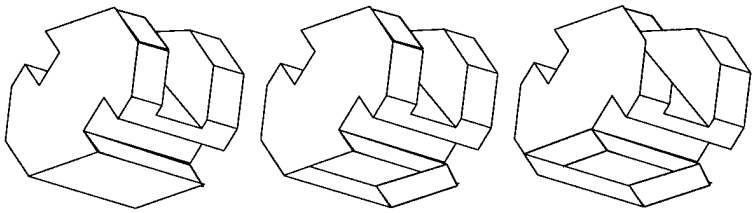
dependent, and utilizes each along with appropriate recognition mechanisms in accordance with the perceptual information necessary for accomplishing a given task.

• *Memory for Impossible Objects*

The study of human object recognition and human memory have been considered separately for many years. Recently, however, this gap has been bridged by the advent of several implicit memory paradigms for studying memory for objects. One such paradigm, introduced by Schacter, Cooper, and Delaney (1990), tests subjects' memory for the 3-D structure of objects. In an initial study phase, subjects view line drawings of novel objects while performing a study task, such as deciding what direction each object faces. Following this study task, subjects perform an object possibility test, in which they decide whether line drawings represent possible objects, that could potentially exist as real, 3-D objects, or impossible objects, that have "structural ambiguities" rendering them unable to actually exist in 3-D (see figure below). Some test objects were presented during the study task whereas others are completely new; as with other implicit memory tasks, memory for studied items can be inferred if subjects perform differently on studied than on unstudied test items. This advantage for studied compared to unstudied items is known as a priming effect. One important finding coming out of studies employing this paradigm has been that subjects often demonstrate priming on the object possibility test for possible, but not for impossible, objects (Schacter et al., 1990). Cooper and Schacter (1992) considered this to be evidence that possibility priming is based on encoded information about the 3-D structure of objects. Since impossible objects do not have definable 3-D structures, Cooper and Schacter argued, they cannot be primed.

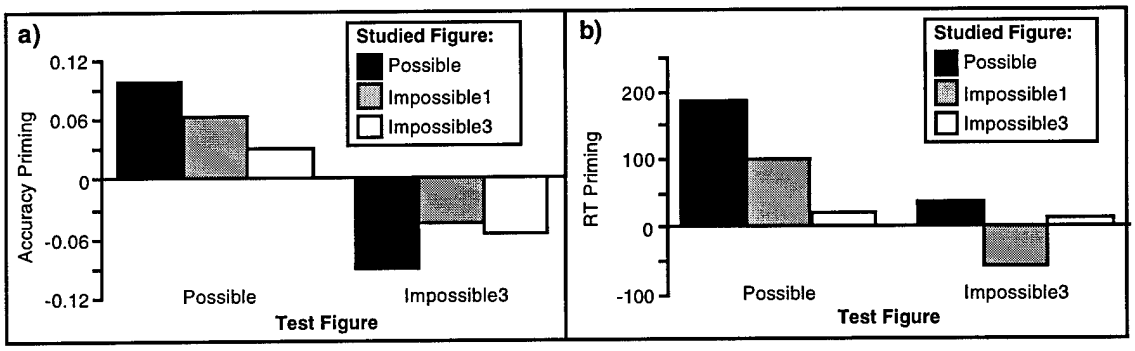
While a logical analysis of the object possibility task strongly implies that possibility decisions should, in fact, be based on memory for object structure, we reasoned that priming might also be found for the possible portions of impossible objects. That is, while some lines and surfaces in impossible drawings cannot exist in 3-D, other surface combinations are perfectly plausible. Furthermore, some objects can be "more impossible" than others, leading to the inference that the more possible structural information is available in a line drawing, the more priming should be evidenced for that drawing. To test these ideas, we developed a new stimulus set in which line drawings of

possible objects were matched with line drawings of two types of impossible objects containing one and three impossible portions (see figure). We call these objects possible, impossible1, and impossible3 objects, respectively. In the study phase of these experiments, we had subjects study some of each type of object. Then in the test phase, we had subjects perform the object possibility task on both the possible and the impossible3 versions of the studied objects, as well as some unstudied objects.



Three versions of an object: possible; impossible1, and impossible3

Part (a) of the figure below shows results from an experiment in which subjects had only 45 ms to view test figures. Under these demanding stimulus conditions, accuracy (proportion correct responses) serves as the dependent measure. Results for possible test objects are shown on the left side of the figure. Here, we can see that our hypothesis was confirmed: possible studied objects, which contained the most amount of 3-D structural information, primed possibility decisions the most, while impossible1 and impossible3 objects, which contained progressively less valid structural information, primed possibility decisions less and less. Results for impossible3 test objects are shown on the right side of the figure. Here, we found that having studied an object made subjects less likely to respond correctly that objects were impossible. Importantly, this negative priming effect was greatest if subjects had studied the object in its possible version. This result is again consistent with our hypothesis that subjects utilize encoded information about the 3-D structure of possible portions of objects in the process of making possibility decisions. Subjects were not always able to extract enough information from test flashes to be absolutely confident about their possibility decisions. Therefore, subjects' memory for the possible portions of studied objects led them to believe that those objects were possible, regardless of whether the flashed test objects were possible or impossible. However, since subjects remembered more structural information about possible than about impossible studied objects, larger priming effects, both positive and negative, were found when studied objects had been possible.



Accuracy and response times from object possibility experiments. a) When test objects were flashed for 45 ms; b) When test objects were shown for up to 5 sec.

Part (b) of the above figure shows results from a second experiment in which we allowed subjects enough time (up to 5 seconds) to view test objects so that they could almost always make correct possibility decisions. Since accuracy rates were at ceiling, response time served as the primary dependent measure in this experiment. Results for possible test objects, on the left of the figure, confirm our findings from the previous experiment: the most amount of priming was demonstrated when objects were studied in their possible versions, less priming was demonstrated for impossible1 studied objects, and the least amount of priming was demonstrated for impossible3 studied objects. Impossible3 test objects were not significantly primed in any of the conditions in this experiment (right side of the graph). This result, which we have subsequently replicated many times, indicates that encoded knowledge about the 3-D structure of objects neither helps nor hinders the perception of impossible portions in line drawings of the objects.

In total, these results provide strong support that object possibility priming is based on encoded information about 3-D object structure. However, our findings also indicate that this structural information need not be encoded in an "all-or-none" fashion, as was originally claimed by Cooper and Schacter (1992). The graded effects of object impossibility found in the present experiments are also inconsistent with the kind of abstract structural representations proposed by theorists such as Marr (1982) and Biederman(1987), which would be expected to be either computable (for possible studied objects) or not (for impossible studied objects).

• *Viewpoint-specific implicit memory effects*

In a second line of research utilizing the object possibility paradigm, we tested priming for objects that had been rotated in the picture plane between study and test. Here, we were interested solely in performance on possible test objects primed by possible studied objects. As in the latter experiment described above, we allowed subjects up to 5 seconds to view test objects, and utilized response time as the dependent measure. Object possibility results from two experiments indicate that priming declines gradually from a robust 250-ms for unrotated objects to essentially 0-ms for objects rotated 60° or more in the picture plane. These results were contrasted with those from an old/new recognition tasks, in which exactly the same objects were shown but subjects were required to decide whether objects had or had not been seen during the earlier study task. Here, performance declined consistently throughout the entire range of 0° - 180° rotations employed in the experiments.

To account for these patterns of effects, we have posited an object recognition model utilizing viewpoint-specific representations of object structure. When making old/new recognition decisions, the perceived structure of test objects must be explicitly matched with these encoded structural representations. The further test objects are rotated from studied viewpoints, the more difficult this matching process becomes. In the process of making possibility decisions, on the other hand, structural representations are only accessed incidentally, since these decisions do not require explicit identification of particular objects. Our results indicate that such incidental contact with encoded structural representations is only helpful for making possibility decisions when objects are rotated 60° or less in the picture plane.

Explicit identification of objects, as is required in traditional object recognition tasks such as naming and old/new recognition, is relatively uncommon in everyday experience. For example, we do not exclaim "This is a mug" every time we take a sip of coffee. Implicit memory tasks such as object possibility, in which access to object representations is incidental rather than explicit, can potentially provide important converging evidence with naming and old/new recognition on the nature of encoded object representations used in object recognition processes. Results from the experiments outlined above support the viewpoint-dependent representations inherent in models such as those offered by Edelman and his colleagues (Bülthoff, Edelman, & Tarr, 1995). Another experiment from our lab indicates that information about object

size is not included in the representations utilized during the object possibility task. Somewhat surprisingly, however, yet another experiment suggests that information about object color may be important to the structural representations accessed while making possibility decisions. We are currently following up on all three of these lines of research.

• *Lexical and Perceptual Encoding of Spatial Relations*

William Hayward and I completed a series of experiments investigating the nature of *qualitative* spatial relations encoded between objects in a scene (or between parts of an object). Such relations are an essential element of many structural-description theories of object representation (i.e., Hummel & Biederman, 1992). Specifically, we have examined the possibility that the restricted meanings of spatial prepositions used in language reflect a similar qualitative encoding of spatial relations in the visual representation system. As detailed in the attached paper (accepted pending revision to *Cognition*), a series of four experiments indicate that linguistic descriptions and the visual encoding of space share common structures for the relations "above" "below" "left" and "right". Across four experiments objects were presented in a scene where one, the reference object, always appeared in the center, and the other, the figural object, appeared in one of many positions on a 7x7 grid surrounding the reference object. Results from the first two experiments indicate that perceivers have a preference to apply spatial terms in a qualitative manner — for example, applying "above" when the figural object is directly vertical relative to the referent. Secondly, while the same spatial terms certainly apply to other relations between objects, they do so in a gradient that decreases in both frequency of application and assessed appropriateness with distance from the preferred axis.

A similar pattern was obtained in two experiments that employed perceptual judgments with scenes configured as in the first two studies. One study required subjects to use spatial memory to recall the position of the figural object relative to the reference object. A second study required subjects to judge whether the figural object was in the same location relative to the reference object in two sequential frames (which shifted randomly in screen position so that subjects could not simply note the absolute position of the figural object between frames). In both studies performance was highest at spatial positions where the figural object was axially aligned with the reference object.

Such results suggest that there is a correspondence between qualitative spatial representations found in the visual system and the categorical form referred to in language (i.e., we refer to objects being simply *above* rather than precisely how far above). Given this correspondence, we may begin to explore the specifics of spatial relations within objects using both linguistic and non-linguistic tasks. For example, one paradigm may employ sequentially presented images of similar objects where the relations between parts vary. While the magnitude of quantitative changes in spatial relations are expected to influence performance, qualitative changes are predicted to have a far greater impact on performance. This and related paradigms may be used to assess the qualitative boundaries of part relations within objects, as well as possible similarities to linguistic descriptions of such relations.

5. Personnel Supported

Michael Tarr	Faculty
Vlada Aginsky	Graduate Student (1st year; Brown University)
Isabel Gauthier	Graduate Student (3rd year)
Pepper Williams	Graduate Student (3rd year)
Alan Ashworth	Graduate Student (Ph.D., May 1995)
William Hayward	Graduate Student (Ph.D., May 1995)

Undergraduates: Steven Messé, Yale University, 1992-1993; Douglas Bitting, Yale University, 1993; Scott Yu, Yale University, 1993-1994; KaRin Turner, Yale University, 1993-1994; Joan Weisman, Yale University, 1995; Marion Zabinski, Yale University, 1993-1995; James Servidea, Yale University, 1994-1995; Jame Rosoff, Yale University, 1994-1995; Tami Edwards, Yale University, 1995; Scott Klemmer, Brown University, 1995

6. Publications

Tarr, M. J., Kersten, D., & Bülthoff, H. H. Why the visual system might encode the effects of illumination. *In preparation.*

Ashworth, A. R. S. III, & Tarr, M. J. Mental representations of faces and their transformations. *In preparation.*

Tarr, M. J. The concurrent encoding of viewpoint-dependent and viewpoint-invariant information during object recognition. *In preparation.*

Tarr, M. J., & Gauthier, I. Geometric and class constraints in visual object recognition: The representation of objects in unfamiliar views. *In preparation.*

Gauthier, I., & Tarr, M. J. Becoming a "greeble" expert: Exploring mechanisms for face recognition. Submitted to *Vision Research.*

Hayward, W. G., & Tarr, M. J. Testing conditions for viewpoint invariance. Submitted to the *Journal of Experimental Psychology: Human Perception and Performance.*

Tarr, M. J., & Bülthoff, H. H. (Eds.) Image-based object recognition in man, monkey, and machine. To appear as a special issue of *Cognition.* Contributors: P. Schyns & H. Ando, T. Poggio & T. Vetter, N. Logothetis, D. Perrett, M. Goodale, S. Ullman, H. H. Bülthoff, M. J. Tarr.

Tarr, M. J., & Kriegman, D. J. Toward understanding human object recognition: Aspect graphs and view-based representations. *Psychological Review.* [Accepted pending revisions.]

Williams, P., & Tarr, M. J. Object decision priming and recognition sensitivity for different types of impossible figures. Submitted to *JEP:LMC.*

Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. To what extent do unique parts influence recognition across changes in viewpoint? Max-Planck Technical Report #22, Max-Planck Institut für biologische Kybernetik, Tübingen, Germany. Submitted to *Psychological Science.*

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? *Journal of Experimental Psychology: Human Perception and Performance, 21(6),* 1494-1505.

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

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 5*(3), 247-260.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition, 55,* 39-84.

Tarr, M. J., & Black, M. J. (1994). A computational and evolutionary perspective on the role of representation in vision. *Computer Vision, Graphics, and Image Processing: Image Understanding, 60*(1), 65-73.

Tarr, M. J., & Black, M. J. (1994). Reconstruction and Purpose. *Computer Vision, Graphics, and Image Processing: Image Understanding, 60*(1), 113-118.

Tarr, M. J. (1994). Visual representation. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior* Vol. 4 (pp. 503-512). San Diego: Academic Press.

Tarr, M. J. (1993). Is a picture really worth a thousand words? *Computational Intelligence, 9* (4), 356-359.

Tarr, M. J. (1993). From perception to cognition. *Behavioral and Brain Sciences, 16* (2), 251-252. [Commentary on target article: Landau & Jackendoff, "What" and "where" in spatial language and spatial cognition].

7. Interactions/Transitions (a. Participation/Presentations)

Braje, W. L., Kersten, D. J., Tarr, M. J., & Troje, N. F. Shadows and illumination influence face recognition. *The Annual Meeting of The Association for Research in Vision and Ophthalmology (ARVO),* Ft. Lauderdale, FL, April, 1996.

Tarr, M. J. Features of Recognition. *Workshop on Object Recognition*, Smith-Kettlewell Eye Research Institute, San Francisco, CA, January 2-6, 1996.

Gauthier, I., & Tarr, M. J. Becoming a Greeble expert: Exploring the face recognition mechanism. *36th Annual Meeting of the Psychonomic Society*, Los Angeles, CA, November 10-12, 1995.

Hayward, W. G., & Tarr, M. J. When does human object recognition use outline shape? *36th Annual Meeting of the Psychonomic Society*, Los Angeles, CA, November 10-12, 1995.

Ashworth, A., & Tarr, M. J. Recognizing rotated faces. *36th Annual Meeting of the Psychonomic Society*, Los Angeles, CA, November 10-12, 1995.

Bülthoff, H. H., Tarr, M. J., Blanz, V., & Zabinski, M. To what extent do unique parts influence recognition across changes in viewpoint? *European Conference on Visual Perception*, Tübingen, Germany, August, 1995.

Tarr, M. J., Hayward, W. G., Gauthier, I., & Williams, P. Is object recognition mediated by viewpoint-invariant parts or viewpoint-dependent features? *European Conference on Visual Perception*, Tübingen, Germany, August, 1995.

Blanz, V., Vetter, T., Bülthoff, H. H., & Tarr, M. J. What object attributes determine canonical views? *European Conference on Visual Perception*, Tübingen, Germany, August, 1995.

Gauthier, I., & Tarr, M. J. Generalizations of viewpoint-specific representations to new exemplars of a category. *66th Annual Meeting of the Eastern Psychological Association,* Boston, MA, April, 1995.

Ashworth, A., & Tarr, M. J. Rotating faces. *66th Annual Meeting of the Eastern Psychological Association,* Boston, MA, April, 1995.

Tarr, M. J. Multiple views: Behavioral evidence for a theory of human object recognition. *NECI Vision Workshop*, NEC Research Institute, Princeton, NJ, March, 1995.

Kersten, D., Tarr, M. J., & Bülthoff, H. H. Object recognition depends on illumination. *Annual Meeting of The Association for Research in Vision and Ophthalmology (ARVO)*, Ft. Lauderdale, FL, May, 1995.

Tarr, M. J. Common mechanisms for the recognition of faces and objects. *ATR Symposium on Face and Object Recognition - 95*, Kyoto, Japan, January 17-20, 1995.

Tarr, M. J., Hayward, W. G., Gauthier, I., & Williams, P. Geon recognition is viewpoint dependent. *35th Annual Meeting of the Psychonomic Society*, St. Louis, MO, November 11-13, 1994.

Williams, P., Crowder, R. G., & Tarr, M. J. The basis of the "bias" effect in object decision priming. *35th Annual Meeting of the Psychonomic Society*, St. Louis, MO, November 11-13, 1994.

Tarr, M. J. Allocation of views: Behavioural evidence for a theory of human object recognition. *17th Annual Meeting of the European Neuroscience Association*, Vienna, Austria, September 5-8, 1994.

Kersten, D., Tarr, M. J., & Bülthoff, H. H. Illumination dependency in human object recognition. *European Conference on Visual Perception*, Eindhoven, Netherlands, September 5-8, 1994.

Williams, P., & Tarr, M. J. 3D possibility of both studied and tested objects affects object decision performance. *Annual Meeting of the American Psychological Society*, Washington, DC, June 30-July 3, 1994.

Hayward, W. G., & Tarr, M. J. Viewpoint effects in the recognition of natural stimuli. *Annual Meeting of the American Psychological Society*, Washington, DC, June 30-July 3, 1994.

Hayward, W. G., & Tarr, M. J. Spatial language and spatial representation. *65th Annual Meeting of the Eastern Psychological Association*, Providence, RI, April 15-17, 1994.

Tarr, M. J. Conditions for viewpoint dependence and viewpoint invariance in human object recognition. *Lake Ontario Visionary Establishment XXIII Conference on Perception and Cognition*, Niagara Falls, Canada, February 10-11, 1994.

Tarr, M. J., & Chawarski, M. C. The concurrent encoding of object-based and view-based object representations. *34th Annual Meeting of the Psychonomic Society*, Washington, DC, November 5-7, 1993.

Tarr, M. J. Invited panel member, special session on purposive vision. *International Joint Conference on Artificial Intelligence*, Chambery, France, August, 1993.

Tarr, M. J., & Kriegman, D. J. A formal basis for understanding view-based representations in humans. *Workshop on Visual Perception: Computation and Psychophysics*, Cape Cod, MA, January, 1993.

Colloquia: Department of Cognitive and Neural Systems, Boston University, February, 1993; Department of Psychology, University of Toronto, March, 1993; Department of Cognitive and Linguistic Sciences, Brown University, April 1993; ONR Workshop on Cognitive Neuroscience, Pittsburgh, PA, October 1993; Department of Psychology, Columbia University, November 1993; Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany, December 1993; Center for Cognitive Science, Rutgers University, April 1994; ONR Workshop on Image Understanding, Washington, DC, April 1994; Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany, August 1994; Wesleyan University, November 1994; ATR Human Information Processing Research Laboratories, Kyoto, Japan, January 1995; Boston College, March 1995; Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany, August 1995; University of Leuven, Belgium, August 1995; Cambridge Basic Research Institute, October 1995; University of California, Santa Barbara, November 1995; NEC Research Institute, March 1996; Memory Disorders Research Unit, Boston University Medical School, March 1996; Department of Psychology, University of Minnesota, June 1996.

8. Inventions and Patents

None.

9. Honors/Awards

- Nominated for an American Psychological Association Early Career Research Award (1996).
- Memberships: ARVO, American Psychological Society, Psychonomic Society, *Behavioral and Brain Sciences* Associate, American Psychological Association, Eastern Psychological Association.
- Organizer of a workshop on scene perception at the Max-Planck Institute in Tübingen, GERMANY in the summer of 1996.

- Co-instructor (with Stephen Palmer) for the Psychology of Perception course at the First International Summer Institute in Cognitive Science (FISI-CS) at the State University of New York, Buffalo, NY, July 5-29, 1994.

- Founder and organizer of the Pre-Psychonomics meeting on research in object perception and memory (OPAM), 1993, 1994.

- Consulting Editor, *Journal of Experimental Psychology: Human Perception and Performance*; *Psychological Bulletin; Psychological Science*