The Contribution of Symmetry and Motion to the Recognition of Faces at Novel Orientations

THOMAS A. BUSEY*
Indiana University, Bloomington, Indiana

SAFA R. ZAKI

Williams College, Williamstown, Massachusetts

In Press, Memory and Cognition

* Please send correspondence to Thomas A. Busey, Department of Psychology, Indiana University, Bloomington, IN 47405, email: busey@indiana.edu

Abstract

In three experiments we examine the recognition of faces at novel orientations. While performance tends to decay as difference between the study and test angles increases, an orientation that is symmetric with respect to the study orientation shows strong performance, in many cases better than the frontal view. We investigate the properties of this particular facility in perception and memory tasks. Symmetrized faces show surprisingly different patterns of behavior than unsymmetrized faces, despite the fact that many faces are already fairly symmetric. In memory experiments, subjects show robust symmetric orientation effects, and can differentiate between the original study views and the symmetric orientation. In a third experiment we demonstrate that smooth motion improves performance at the symmetric orientation, while two control motions do not. Together the three experiments support the view that multiple representations are at work during the recognition of faces at the symmetric orientation, and that during memory tasks subjects tend to rely on representations that are more robust against texture asymmetries and that may include limited depth information.

When recognizing faces at different orientations, subjects' recognition performance often degrades as the angle increases between the studied and test views. Thus while we can recognize faces from a variety of viewpoints, even under conditions in which we have been exposed to only a single viewpoint of a novel face (e.g., Rock & DiVita, 1987; Tarr & Pinker, 1989; 1990; Bülthoff & Edelman, 1992), performance is best when the study and test views match (Patterson & Baddeley, 1977). However, when generalizing to novel viewpoints, one particular orientation stands out as having special properties. Several authors, (see Vetter, Poggio & Bülthoff, 1994; Vetter, Hurlbert & Poggio, 1995; Liu, Knill & Kersten, 1995; Troje & Bülthoff, 1996) have pointed out that the *symmetric* orientation could be used by the observer in conjunction with properties of the human head to enable good generalization and therefore accurate recognition. This explanation takes advantage of the approximate bilateral symmetry of the human head. For example, a head viewed at a 45° angle will appear as an approximate mirror image when viewed at a -45° angle, despite the fact that the other side of the head is now visible (Troje & Bülthoff, 1996). Human observers perform quite well at the symmetric orientation, and in fact perform better than at the frontal view. This strong recognition of this particular novel orientation is known as the *symmetric orientation effect*.

A variety of theoretical approaches have been put forth to account for object and face recognition data, and several have been applied to account for recognition at the symmetric orientation. While a vigorous debate has emerged in the object recognition literature over the adequacy of different models (see Hummel, 2000), the cumulative data now suggests that multiple representations may be at work. The present work considers the nature of the representations that support strong recognition of symmetric orientations in perception and memory experiments.

The groundbreaking psychological research that addressed face recognition at different viewpoints relied primarily on paradigms that tap visual working memory, such as same-different tasks (Troje & Bülthoff, 1996; 1998; Troje, 1998). While this paradigm is appropriate for addressing the perceptual variables of interest in these studies, in many cases face recognition requires longer memory retention intervals in order to support social interactions. In the present work we generalize existing work on face recognition at different viewpoints to memory paradigms that address how observers perform when asked "Have you seen this person before?" rather than "Is this the same face?". Although both are face-recognition tasks, memory tasks are require longer retention intervals, and introduce the necessity to consider multiple faces in memory prior to making a decision. Thus the two paradigms may stress different sources of information or representations.

Symmetric orientation effects have the potential to help us understand the nature of information that enables generalization across orientations. To organize the possible representations that may be at work, below we briefly summarize the extant models of face and object recognition as they apply to our tasks. These provide a basis for the interpretation of our results.

Models of Symmetric Orientation Recognition

Vetter, Poggio & Bülthoff (1994) and Vetter & Poggio (1994) have pointed out that symmetric-orientation recognition of bilaterally-symmetric objects such as faces could be accomplished with transformations applied to image-based representations that work in a 2-dimensional coordinate system. The salient landmark features are reflected across the vertical meridian and then matched to the face shown at the symmetric orientation. No knowledge of the 3-dimensional structure of the head is required, and the technique is specific to bilaterally-symmetric objects. Models that rely on this transformation accurately capture many of the aspects of human recognition, including the fact that the symmetric orientation is better recognized than the frontal view. In addition, the finding that changes in lighting prove devastating to performance provides support for these models because it suggests that humans do not extract enough 3-dimensional structure from the head to compensate for lighting changes (Troje & Bülthoff, 1996).

This image-based transformation is an extension of a class of theories known as view-based models, which seek to explain as much of object recognition as possible without resorting to representations other than transformations applied to 2-dimensional views. These 2-dimensional views are seen as holistic representations that enable matching whole images to images stored in memory. The images in memory are not exact copies of the raw image, since preprocessing allows normalization for location and size on the retina. In an extended version, a hierarchy of views is possible, such that parts of objects are stored as separate views. However, this moves the model toward a second class of models as described next.

A second class of models has been termed *structural description* models, because the models process relevant features into parts (e.g., geons or other feature primitives) and then represent the relations between these features independently of the features themselves (e.g. Biederman, 1987; Hummel & Stankiewicz, 1996a, 1996b, 1998; Stankiewicz & Hummel, 1996). These models stress additional processing of the image, including abstractions that go beyond image-based information.

Thus one dimension along which to view the two classes of models is the degree of abstraction that is assumed: view-based models stress the use of relatively few abstractions and rely primarily on fairly veridical copies of image information. Faces share similar geon compositions, and the emphasis of this type of description may be on individual features or relational information as augmented by a holistic module. Structural description models emphasize more abstractions that explicitly encode relational information beyond that contained in a view-based representation. However, recent versions of structural description models have also moved closer to the view-based models. For instance, the Jim.2 model (Hummel & Stankiewicz, 1996b) includes a holistic representation that retains elements of a view-based model.

Both image-based and structural description models are consistent with the symmetric orientation effect. The reflection about the vertical axis initially proposed by Vetter, Poggio and Bülthoff (1994) allows an image-based transformation to accurately capture this effect. Likewise, in one instantiation of a structural description model by Hummel & Biederman (1992), the horizontal relations between relevant features are not coded as 'left-of' or 'right-of', but simply as 'beside'. Thus to the structural description model, the original view at 45° would be viewed as nearly identical to the -45° view, thereby providing strong symmetric orientation recognition. The structural description model would parse a frontal view with a different set of relations, and thus would have more difficulty matching the frontal view to the original view, thereby accurately accounting for the human difficulties with the frontal view. The structural description models might have difficulty accounting for the performance degradation seen with changes in illumination, since the structural description models are designed to be robust against these shading changes. However, if illumination changes influenced the nature of the structural description construction, then these data are also consistent with structural description models. Despite the fact that both types of models account for basic effects, it is possible to tease apart the different representations with additional manipulations. The goal of the present work is to address when these different representations, as characterized by the two classes of models, may be at work in both recognition and memory experiments.

Our first experiment was designed to replicate existing effects in the literature with fully-textured faces in order to address the role of texture in the symmetric orientation process. We find that a fairly minor symmetrization transformation applied to the faces produces a dramatically different pattern of results across different orientations. We then consider data from two memory experiments that extend the symmetric orientation effects to longer-term retention intervals, which may require more durable representations that abstract more information away from the raw image. Together all three experiments suggest that symmetric orientation effects depend on several different sources of information that makes symmetric orientation recognition surprisingly robust in memory paradigms and more dependent on exact image matches in perception experiments. This suggests a greater role for view-based models in perceptual tasks, but abstractions similar to those proposed in structural-description models may contribute more in memory experiments.

Experiment 1

The optimal stimulus for an image-based system with a horizontal reflection process is the mirror reversed image, which was explored by Troje & Bülthoff (1998). In an interesting manipulation, they introduced asymmetric lighting which produced a drastic drop in performance at the symmetric orientation, in part because the pattern of shadows changes when the face rotates but the asymmetric lighting does not. This is consistent with an image-based system that encodes illumination along with the structure of the head instead of abstracting higher-level features that are robust against lighting changes. In Experiment 1 we extend this work to address conditions such as those introduced by their asymmetric lighting, in which observers might rely on representations that are relatively robust to texture asymmetries. However, rather that rely on asymmetric lighting, we chose instead to use the natural texture asymmetries inherent in textured faces.

The initial experiments that explored symmetric orientation effects used 3-dimensional headscan models that lacked texture (Troje & Bülthoff, 1996; 1998; Troje, 1998). In many cases this was done for good reason, because texture elements may interfere with experiments designed to address the abstraction of 3-dimensional structure. However, texture tends to survive the kinds of head-turn motions that faces typically undergo, which may affect the relation between the original view, the frontal view and the symmetric orientation view. Texture elements tend to be more amenable to abstractions such as "mole on left cheek", and thus untextured faces may emphasize a view-based system that tends to rely on 2-dimensional shape information. Note that here we are using the term "texture" somewhat loosely, since it is difficult to separate texture from shape information. A headscan reports both 3d structure information as well as a color map overlaid on the shape information. When working with laser scans it is possible to remove texture from the face by setting each voxel to a uniform color, and we operationally define texture as the color of the voxel. However, this is not pure texture since this does not remove the dependencies on shape once it is added back in. Thus the distinction between texture and shape information may lie more in the scale of the information associated with each source of information: shape information tends to be derived from large portions of the head, while texture information can be obtained mainly (but not entirely) from small regions of the head. In addition, a rotating 3-dimensional object will change its 2-d projection as it rotates, but texture, especially if it is relatively uniform, may survive these changes. Subjects may adopt a strategy that relies on texture because of this survival, but this may lead to potential problems if the observer becomes overly reliant on texture and then the texture changes when the symmetric orientation is tested. Thus texture asymmetries have the potential to reveal the strategies that subjects adopt when generalizing to novel orientations.

To address the role of texture, one approach might be to compare textured and untextured faces directly. However, textured faces contain much more information and may produce better performance, making direct performance comparisons subject to scale dependencies. In addition, performance on the untextured faces is likely to rely heavily on the particular technique used to remove texture. We chose an alternative that preserves texture but moves the faces closer to the original untextured stimuli. We used an image processing technique that symmetrized the faces. In a symmetrized face the texture and 3D structure matches exactly in both the original and symmetric view, and this affords an image-based system an ideal opportunity to match the original and symmetric views. The unsymmetrized faces will have somewhat different textures on the two sides, thus potentially reducing performance if an image-based system is used, because salient feature that are reflected across the midline will not find a match. However, if performance remains relatively unchanged then observers are relying on representations that are not harmed by texture mismatches. The results will suggest when different representations (as characterized by view-based and structural-description models) are at work, will aid in the interpretation of prior results, and will also allow comparisons with the memory data from Experiment 2, which differ in surprising ways.

We included one other manipulation that we thought might affect the nature of the information used by observers when recognizing faces at the symmetric view. The horizontal reflection process suggested by Vetter, Poggio and Bülthoff (1994) takes advantage of the fact that faces are bilaterally symmetric. However, this is true only for faces that are upright or fully inverted. To put pressure on this reflection process, we introduced a 30° tilt into some of the test images. This tilt was applied in the image plane after the head was rotated around the vertical axis. When faced with a tilted test image, a reflection process will not find a close match to the untilted study image, and other information may be require to solve the recognition problem. Thus introducing a tilt may reveal the existence of representations that work in addition to the image-based reflection process such as an abstraction process as described by structural-description models. While a view-based model might assume a rectification process prior to reflection, the results of this manipulation bear on whether human observers do in fact perform this operation or instead suffer from reduced performance as a result of the tilt.

The goal of this experiment is to address the nature of the information at use when generalizing faces to new orientations. While not a pure test between the two classes of models (which have tended to become more similar over time), the results suggest when a more template-based match might be at work, and when observers may tend to rely on more local features or texture elements that are more robust to manipulations that change the surface appearance of the face. The tilt manipulations may induce other representations that might not rely on exact texture matches that are the hallmark of view-based models as they were originally conceived. These other representations may result in an interaction between tilt and the symmetrization procedures, since an unsymmetrized face may not harm a more abstract representation if it becomes active for a tilted face.

Method

Observers

The observers were Indiana University undergraduate students. There were 20 observers in Experiment 1. The observers received course credit for their participation.

Stimuli

The stimuli were 16 laser scan models of human heads obtained from a Cyberware(TM) laser scanner. During scanning, the heads of the scanned faces were covered with a beige cap that hid most of the hair. The laser scanner produces a depth map of the head, mapping the deformations of the face from a cylinder at 512 (horizontal) by 256 (vertical) positions. Red, green, and blue color values are also measured at each point.

During scanning the faces were asymmetrically lit, which produces a small luminance difference across the head (see Figure 1). We obtained the head scans from other researchers and therefore had no control over this aspect of scanning. This texture asymmetry is embedded in the RGB values, and thus will translate as the face rotates. Note that this is unlike the asymmetric illumination of Troje & Bülthoff (1998), which did not translate as the face rotates. This asymmetry is in addition to any naturally occurring textural or structural asymmetry. The symmetrization procedure described below will remove this asymmetry.

The heads were symmetrized as follows. First, the depth map (which provides the 3-dimensional structure of the head) and the corresponding luminance map were treated as 512 pixel by 256 pixel images. In the depth map, the value at each location provides the deformation from a cylinder that gives the face its structure. The corresponding pixel in the luminance map provides the color at that location once the luminance map has been molded to fit the depth map. We then used graphics programs to flip both texture and depth maps horizontally which creates the mirror face. The next step was to place control points on the important features of both of the texture maps. Then, using morphing procedures (Beier &

Neely, 1992) we combined the original face with the mirror face by warping the locations of the features onto an average set of locations (computed individually for each face) and then performing a cross-fade blend that combined the two images. This results in a completely symmetric face at both the level of the structure and the texture that looks quite realistic. We noted a slight blurring effect in the texture as a result of the blending process, so we introduced a slight sharpening effect in the texture maps prior to 3D rendering of the head models. Figure 1 shows an example head and its symmetrized version.

The models were then converted into files that could be photo-realistically rendered in the POV-Ray rendering package. These were rendered using a light source positioned above and behind the location of the virtual camera. The heads were rotated around the vertical axis, which extends through the middle of each head.

Each face had an unsymmetrized and a symmetrized version. As noted above, the symmetrization procedure tended to slightly blur aspects of the texture, and of course the symmetrized faces looked more symmetric. In order to prevent subjects from using these artifacts as cues to the correct answer in the same/different task, we chose distractors from the same class of faces. Thus if an unsymmetrized face was used as the study face, the distractor test face would also be an unsymmetrized face.

We took special care to insure that the center of rotation of the unsymmetrized faces was placed at the center of the head as determined by the distance between the ears. We also used the tip of the nose and the maximum points on the ears to align the rotation of the head such that at 0° the head was pointing straight ahead as much as possible, although perfect alignment is difficult with asymmetric heads. These checks are important so that additional asymmetries are not introduced by imperfect rotations or alignments.

The tilt manipulation was performed in the image plane using Photoshop after rendering in POVRay. This makes shadows rotate with the face rather than changing with the tilt.

The stimuli were presented in full color on two 21" monitors that mirrored the monitor of a Macintosh 7200 PowerMac computer. Timing and stimulus control was provided by the VideoToolbox library of c routines (Pelli & Zhang, 1991). Data collection was provided by a two-button keypad with millisecond resolution.

All within-subject manipulations involved random assignments across trials, and were not blocked. All experiments used within-subject designs, with the exception of Experiment 2 which used one between-subject manipulation as discussed in a later section.

Design and Procedure

The stimuli were presented in a sequential same-different task. A study face appeared for 1440 ms and was replaced by a pattern mask composed of parts of various faces. This mask remained on the screen for 1000 ms, and then was replaced by a test face. The study face was positioned slightly to the left of the center of the monitor, and the test face was positioned slightly to the right of center. The test face remained on the screen until all subjects had responded.

The study image was always an upright face presented at either 40° or -40°. The test face was either at the same angle, the opposite angle, or the frontal view. In addition, the test face could have been tilted 30° to the left or right, which was applied in the image plane. All conditions were counterbalanced.

Observers were given instructions that they should respond 'same' if the same person appeared at test even if the person appeared at a new orientation. They were given practice trials and feedback to insure that they understood the instructions. Observers were also asked to respond as quickly as possible, maintaining a high degree of accuracy. We gave them feedback via a flashing LED if their response exceeded 800 ms on any particular trial.

Subjects completed 480 trials which included 10 repetitions of each of the 48 conditions (3 angles (same, frontal, opposite) x symmetrization (symmetrized vs. unsymmetrized) x 3 levels of tilt (upright, tilted left, tilted right) x 2 levels of facing (positive angels vs. negative angles) x target present/absent). Testing lasted approximately 90 minutes.

Results

Figure 2 shows the sensitivity (d') data for Experiment 1, which were submitted to a 3-way repeated measures analysis of variance, with symmetrization, tilt, and test angle as factors. We found no differences between faces studied at 40° and -40°, and thus we averaged over these conditions. The left side of Figure 2 presents data from the unsymmetrized faces, which show a very different pattern from the symmetrized faces (right side of Figure 2).

First, consider the unsymmetrized faces (left side of Figure 2). For the upright unsymmetrized faces, the drop in performance between the original angle and the symmetric angle is larger for upright faces than for the tilted faces, as noted in the left side of Figure 2. These effects lead to a significant interaction between tilt (upright vs. inverted) and test angle (original vs. symmetric orientation) for the unsymmetrized faces ($F(1,19)=4.634$, $p < 0.05$).

In addition to these effects, the unsymmetrized faces show no evidence of a symmetric orientation effect: Performance for both tilted and untilted faces at the symmetric orientation view is below that of the frontal view.

For the symmetrized faces, a different pattern of results occurs, as shown in the right side of Figure 2.

Performance at the symmetric orientation is high in both untilted and tilted conditions, with tilt extracting a small performance decrement. This represents a different ordering of the symmetric orientation views than with the unsymmetrized faces, with additive effects for test angle and tilt. The interaction between tilt and test angle is not significant for the symmetrized faces ($F(1,19) < 1$).

Unlike the unsymmetrized faces, the symmetrized faces show strong symmetric orientation effects: Performance at the symmetric orientation is above the frontal view for the untilted faces, and slightly above the frontal view for the tilted faces.

The overall three-way interaction between test angle (original, frontal or symmetric angle), tilt, and symmetrization is marginally significant ($F(2,38)=3.06$, $p=0.059$).

Discussion

The results of the d' (sensitivity) analyses indicate that symmetrized faces show strong symmetric orientation effects, while unsymmetrized faces do not. In addition, for unsymmetrized faces, performance drops a great deal at the symmetric orientation relative to the original orientation, and this performance decrease is mitigated once a tilt is introduced into the test images. Symmetrized faces show only additive effects of tilt and test angle.

Why should the pattern of results in Figure 2 be so different for unsymmetrized and symmetrized faces? One possibility is that observers first try to rely on an image-based representation such as a view-based reflection and match, perhaps because of its computational ease. This produces a poor match at the symmetric orientation for unsymmetrized faces due to texture and structural asymmetries and results in a large performance drop at the symmetric orientation. However, the tilt manipulation may disrupt the image-based system, and bias the observer to rely on a representation that includes other sources of information such as individual features, their relations, or additional higher-level abstractions like 3D structure as suggested by structural description models. This results in a smaller performance decline between the original and symmetric orientation, perhaps because this alternative representation is better able to overcome texture and structural asymmetries.

The results from the upright unsymmetrized faces and the symmetrized faces are consistent with prior work by Troje & Bülthoff (1998). Based on their results with asymmetric lighting and untextured faces, we would expect texture asymmetries to produce steep performance declines, and they do. In addition, we would expect strong symmetric orientation effects for the symmetrized faces, which also can be seen in the right side of Figure 2. However, an image-based model, even one with a 2D rotation built in to overcome the tilt manipulation, may have difficulty with the interaction between tilt and test angle that was observed only for unsymmetrized faces. This results from the fact that although the tilt can be removed, perhaps with some performance cost associated with the tilt-rectification process, this rectification process would not interact with the texture asymmetries. Thus the model would predict similar performance declines between original and symmetric orientations for upright and tilted test images. Instead we see a smaller decline for tilted images than for upright images.

One implication from the present results is that the use of the untextured faces in prior speeded same-different tasks may have emphasized the role of image-based information. In our data, we see strong symmetric orientation effects only for symmetrized faces, and this effect seems to disappear for unsymmetrized images, at least those with our texture asymmetries. When variables such as tilt are introduced to the test image, the observer may switch to a more abstract representation that is not tied so directly to the original image, and this produces a less of a decline at the symmetric orientation, perhaps because texture asymmetries no longer play as strong a role.

One benefit of these possible abstractions is that they may be easier to store for longer retention intervals, and thus may support longer-term recognition of faces, which is the topic of Experiments 2 and 3.

Symmetric Orientation Recognition in Memory

Our next step is to generalize the findings from perceptual experiments to the domains of intermediate and long-term memory. We first address whether symmetric orientation effects are found in memory paradigms, and then ask whether observers have access to the original and symmetric orientation when making old/new memory judgments. Finally, we will address the role of motion in the abstraction of limited 3-dimensional depth information to facilitate symmetric orientation recognition. We chose to use the unsymmetrized faces in Experiments 2 and 3 for their ecological value rather than comparing symmetrized and unsymmetrized faces in each experiment, which would have doubled the number of conditions. Although we did not find symmetric orientation effects with the unsymmetrized faces in Experiment 1, we hypothesized that we might in a memory task because different representations could be used that might be more robust against texture asymmetries. This would provide strong evidence that different processes were at work in the two domains.

Experiment 2

There were two goals in Experiment 2. First, to our knowledge the symmetric orientation effects have not been established in recognition memory, despite the general agreement that the results of perceptual matching tasks should be extended to memory (see Troje & Bülthoff, 1998). Second, if symmetric orientation effects are found in a memory task, can the observer discriminate between the original view and the symmetric orientation view? Intuitively, one may think this trivial. However, in longer-term memory the directionality of the face may be poorly encoded as a result of the fact that orientation is usually irrelevant for identification. For example, in a classic memory demonstration, Nickerson and Adams (1979) produced drawings of pennies that included versions with Lincoln's face reversed. Fewer than half of the observers could pick the correct penny out of the 15 examples. In the neuropsychological literature, Turnbull and McCarthy (1996) describe a patient who could not discriminate between an object and its mirror image *while both were clearly visible*. The results suggest that orientation information is fragile or possibly stored in a location that can be damaged, while leaving object recognition intact. Finally, Price and Gilden (2000) demonstrate that when observers are asked to remember a rotating, translating object, the observers decouple the two events and retained only the translation direction. The translation direction tends to be more informative because it indicates an object that may require interaction or uniquely identifies the object because two objects typically cannot occupy the same location simultaneously. Rotation direction may not require an adaptive behavior because any particular state of the object can be achieved by either clockwise or counter-clockwise rotation.

Thus previous research suggests that rotation direction is relatively uninformative and observers apparently do not encode it. Similarly in the present experiments, the orientation (left- or right-facing) of the face is usually irrelevant for identification purposes, since faces can be observed from either orientation. Thus we may observe a similar dissociation between identity and orientation in face recognition. Note that some structural description models (e.g. Hummel & Biederman, 1992) replace 'left-of' and 'right-of' relations with a 'beside' relation, and thus a structural description model may not be able to discriminate between an object and its mirror image unless additional model components explicitly encode these relations.

To assess whether observers could distinguish between representations of the original and symmetric orientation view, we conducted Experiment 2 as a between-subjects design. Half of the observers were asked to say 'old' if they recognized the face, even if it was at a new orientation. We term these the Inclusion instructions. The other observers were asked to say 'old' *only* if the face appeared at its original orientation. We term these instructions the Exclusion instructions.

We use an old/new recognition paradigm in which observers studied 12 faces. Each of these is studied in one of 5 possible study orientations, which were 70°, 40°, 10°, -20° and -50°. They were then tested with all 5 orientations, as well as with 12 distractor faces also shown at all 5 orientations.

We are interested in whether symmetric orientations are better recognized than other novel orientations, but we chose study and test angles that were asymmetric around the frontal view. Our rational for this choice was as follows. The symmetrized faces in Experiment 1 showed strong symmetric orientation effects, while the unsymmetrized faces did not, contrary to evidence provided by untextured faces in the literature. This suggests that the choice of stimuli and test angles may have in part dictated the information that observers use to make recognition judgments at symmetric orientations. In the case of the unsymmetrized faces in Experiment 1, observers may have chosen to use relatively raw image-based information that did not provide a good match to the unsymmetrized image at the symmetric orientation. This may have led to the poor performance observed for these stimuli. To de-emphasize reliance on raw image-based information, in Experiment 2 we deliberately chose test angles that were at least 10° away from the actual symmetric orientation, which tends to strongly affect the profile of the face. We believe that this is also more naturalistic: rarely do we view a face at exactly the original study view or exactly the symmetric view. We also chose to use the original (unsymmetrized) faces rather than the symmetrized versions for the same naturalistic reasons. Finding strong symmetric orientation effects even under these conditions would demonstrate that the representation stored in memory included enough information to overcome texture asymmetries and changes in orientation that affect the profile view of the face.

Method

Observers

There were 167 Indiana University students in the Inclusion task and 140 students in the Exclusion task. They received course credit for their participation.

Stimuli

The stimuli were 24 head models, which included the 16 unsymmetrized heads from Experiment 1 plus an additional 8 head models that shared similar characteristics.

As with the previous studies, the models were converted into files that could be photo-realistically rendered in the POV-Ray rendering package. These were rendered using a light source positioned above and behind the location of the virtual camera. The heads were rotated around the vertical axis, which extends through the middle of each head. The study orientations were 70°, 40°, 10°, -20° and -50°. The same orientations were used at test.

The stimuli were presented in full color on two 21" monitors that mirrored the monitor of a Macintosh 7200 PowerMac computer. Timing and stimulus control was provided by the VideoToolbox library of c routines (Pelli & Zhang, 1991). Data collection was provided by a Macintosh Centris 610 computer with 6 external keypads. Up to 6 observers participated at one time.

Design and Procedure

Twelve faces were chosen as target faces for the entire experiment, and 12 were reserved as distractors. Each of the 12 faces was randomly assigned to one of 5 study orientations, which was rotated across groups so that each face appeared at each study orientation an approximately equal number of times. The faces were shown for 2 seconds, and were separated by 2 seconds. Following the presentation of the study faces, the observers were shown 120 test trials, which consisted of 60 target faces (12 targets at all 5 orientations) and 60 distractor faces. Those observers in the Inclusion condition were told to respond 'old' if a studied face appeared *at any orientation*. Those observers in the Exclusion condition were told to respond 'old' if a studied face appears *only at the identical orientation*. In order to prevent observers in the Exclusion condition from explicitly encoding the orientation at study, the response instructions were provided only after the conclusion of the study portion of the experiment for the observers in the Exclusion condition.

To orient the observers to the task and introduce our stimuli, prior to the study session observers were given a brief practice session in which they viewed 3 practice faces at various orientations. These faces were not included in the later study or test session. Prior to the test session, the observers were given 6 practice test trials. In particular, observers were shown two test trials early in the practice test session, one that had an old face at its original orientation, and one in which an old face appeared at a new orientation. Observers in the Inclusion condition were told to respond 'old' to both images of this person, since a view of this person was shown in the practice study session. Observers in the Exclusion condition were told to respond 'old' only to the face at the identical orientation and the practice test session was given at the end of the study session. Feedback was given during the practice test session to ensure that observers understood that they should respond 'old' to a studied face only to orientations that conform to their instructions.

Results and Discussion

We discuss the data from the Inclusion and Exclusion instruction observers separately. All analysis of variance models were computed as 2-way repeated measures ANOVAs with the interaction with subjects as the relevant error term for main effects and interactions.

Inclusion Instructions

The overall data for the Inclusion instructions are found in the left panels of Figure 3. In general we find higher performance when the study and test angles correspond (circled points) and performance decreases as difference between the study and test angles increases. The exception is found for the 70° and -50° study conditions at and near the symmetric orientations, as shown in the lower left panel of Figure 3. For those faces studied at a -50° angle, performance at the 40° test angle shows improvement over the 10° test angle (which is near the frontal view). This improvement is significant ($F(1,166)=5.9$, $p<0.05$). Likewise for those faces studied at a 70° angle, performance at the -50° test view is vastly improved over performance at the -20° test view (also near the frontal view). This improvement is also significant ($F(1,166)=19.4$, $p<0.01$).

These comparisons illustrate that symmetric orientation effects do occur in recognition memory, despite the fact that the test angles did not correspond exactly with the symmetric orientations (for example, observers studied a face at 70°, but -50° was the closest test angle to the symmetric orientation of -70°) and despite the fact that the faces were not symmetrized.

Beta values were computed from hit and false alarm rates. The values were quite close to optimal (~1.1) and show no effects of study angle, test angle, nor an interaction between the two (all p-values > 0.05).

Exclusion Instructions

The overall data for the Exclusion instructions are found in the right panels of Figure 3. Again we find strong performance when the study and test angles correspond (circled points). However, the d' data fall off more quickly as the angular difference increases.

If observers can accurately distinguish between the original study view and the symmetric orientation view, then we may not expect the symmetric orientation effects described for the Inclusion data above. Note that the Exclusion instructions were only given after the study session to avoid problems with the subject explicitly encoding the orientation of the face. The lower-right panel of Figure 3 shows the d' results for the Exclusion task for the same study angles that demonstrate symmetric orientation effects in the Inclusion task. Unlike the Inclusion instruction data, we find a complete absence of symmetric orientation effects for the Exclusion task. Given these combined results, it appears that observers can distinguish between the original and symmetric orientations. That is, whatever representation is used to recognize faces at the symmetric orientation in memory tasks, it appears to be distinguishable (and rejected when necessary) by the subjects from the representation used to recognize the original orientation.

Beta values were computed from the hit and false alarm rates for the Exclusion task. The beta values were larger (~1.4) than in the Inclusion task, but below the optimal beta value of 9. Study angle did not produce systematic differences in beta, but test angle reached significance ($F(4,556) = 3.6$, $p<0.05$) as did the interaction between study and test angle ($F(16, 2224) = 2.5$, $p<0.05$). Inspection of the 25 beta values suggests that this interaction results from higher beta values when study and test angles correspond. Both the main effect of test angle and the interaction may have resulted from the conservative criterion required by the Exclusion instructions which may have made it difficult for observers to adjust their criterion for different test views. Alternatively, this may reflect a failure of the normality assumptions that underlie the beta calculation, which could result from behavior at the tails of distributions. This explanation might also account for the fact that d' values were actually slightly higher in the Exclusion condition. Both beta and d' are somewhat sensitive to changes in the false alarm rate, and when the false alarm rates are already small due to conservative responding, small differences in the false alarm rates might produce unexpected d' and beta values. Note that while these differences limit comparisons across Experiments 1 and 2 in terms of absolute d' values. However, comparisons within experiments are still valid. In addition, we find very similar patterns of data when analyzing hit rates alone.

On possible complication with the present design is that all 5 orientations were tested for each study face. In addition, each distractor face was shown in all 5 orientations. This was done to increase the power of the data, but there is a concern that this repeated testing may have artificially raised performance on the symmetric orientation, or otherwise interacted with the experimental factors. To address this, we analyzed the data separately for each time a face was tested. The data is noisier since there is only 1/6th of the data in each analysis. However, the same pattern of results is evident: the symmetric orientation effects are present even in the first testing of each face, and there is no evidence that the magnitude changes across testings. The Exclusion task shows no evidence of symmetric orientation effects at any of the 6 testings. Thus repeated testing cannot account for our effects. The only effect that changed systematically with testing was the distractor false alarm rate, which went up fairly dramatically, from values initially around .34 to final values of .44 in the Inclusion task, and from .22 to about .27 in the Exclusion task. This change did not interact with test angle in either experiment, but the high values may explain why our d' values are fairly small in both instruction conditions.

In summary, we find evidence for large symmetric orientation effects in recognition memory. That we find such effects even when using unsymmetrized faces and test angles that are 10-20° away from the symmetric study angle suggests that these effects are fairly robust and that observers have abstracted enough information about the faces to overcome texture asymmetries and orientation changes.

The data from the Exclusion instruction task suggests that observers seem to be able to distinguish between representations that support recognition of the original view and the symmetric orientation view. Structural description models would require explicit storage of left-of and right-of relations in order to discriminate between the original and mirror-image views.

Comparison with Experiment 1

The strong symmetric orientation effects found in Experiment 2 contrast sharply with the absence of such effects in Experiment 1. There are a number of methodological differences between the same-different paradigm used in Experiment 1 and the memory paradigm of Experiment 2. However, the unsymmetrized faces used in Experiment 1 were also used in Experiment 2, and the range of angles is also approximately the same. Thus it is of interest to compare the *qualitative* effects observed in the two experiments. We saw no symmetric orientation effects for the unsymmetrized faces in Experiment 1 (see Figure 2, left panel), using performance at the frontal view as a comparison. However, we saw strong symmetric orientation effects for the Inclusion instructions in Experiment 2 (see Figure 3). The short memory delay and speeded nature of the same/different paradigm in Experiment 1 may have emphasized the use of relatively raw image-based information that performs poorly when applied to the unsymmetrized faces. This is consistent with a view-based template match model. However, this view-based information may be difficult to store over longer periods of time, forcing observers in the memory experiment to resort to more abstract representations such as those suggested by structural description models that are more robust in memory and across different orientations.

Regardless of the interpretation, this comparison shows that symmetric orientation effects can be quite strong in memory paradigms, even when unsymmetrized faces are used and tested at angles that do not exactly match the symmetric orientation.

Experiment 3

The robust symmetric orientation effects observed in Experiment 2 suggest that in recognition memory observers rely on a representation that is robust against texture asymmetries, perhaps by abstracting higher-level features and/or spatial relations from the raw image-based code as suggested by structural description models. This supports strong symmetric orientation recognition in memory over longer retention intervals. In Experiment 3 we introduced an additional manipulation at study that we hoped would further encourage the use of representations that involve similar kinds of abstractions that create a more robust representation. During the study portion we rotated the heads around the vertical axis ±15° away from the study orientation in a smooth motion sequence. This gives the appearance of a head shaking back and forth and provides limited depth information around the study angle. We hoped that such a manipulation might provide enough structure-from-motion to allow observers to encode some depth information about the shape of the head. This abstracted information could potentially improve recognition at the symmetric view. Note also that the static faces used in many recognition and memory studies are somewhat different from the real world, where motion, rather than static images, is the norm. Thus our motion manipulation provides an important generalization to real-world face recognition.

Two lines of evidence suggest that motion can affect the nature of the information stored in memory. McNamara, Diwadkar and Blevins (1998) addressed the viewpoint-dependent nature of scene perception using displays that provided (or inhibited) apparent motion. They showed observers a perspective display of 5 colored dots on a computer monitor. The dots could be rotated to simulate a new viewpoint, and when alternated with the appropriate ISI, strong apparent motion was reported by observers. In a control condition, the dots jumped from one side of the screen to the other as well as rotated, which prevents a percept of apparent motion. They then tested the observer's memory for the configuration of dots at novel orientations using an old/new recognition paradigm. Distractor stimuli were composed of the same colored dots in new spatial configurations. They found that recognition for views that fell in between the two views (i.e. the interpolated views) was as good as recognition for the actual study views. These did not hold for the interpolated views in the condition in which apparent motion was prevented. McNamara et al. (1998) concluded that the motion created 'virtual views' in memory in between the endpoint views that facilitated recognition for these interpolated views. Their data suggest that the apparent motion allowed the observers to extract enough structural information about the locations of the dots to infer their relative placement for the interpolated views.

Our face stimuli are more complex than the 5 dots used by McNamara et al.(1998). This rich structure adds additional information that could be used by observers, especially information defined by constraints in which information from one source guides the processing of information from another source. A large literature suggests that depth relations (e.g., Hildreth, Grzywacz, Adelson & Inada, 1990) and possibly the 3-dimensional structure of objects can at least partially be recovered from an object by placing it in motion (e.g., Wallach & O'Connell, 1953; Todd & Bressan, 1990). However, the recovery of depth from motion depends in part on constraints provided by the stimulus. One source of information that may provide these constraints is a line of *coplanar points*. The structure of faces is for the most part bilaterally symmetric, and as a result faces have a series of features that run down the center of the forehead, the nose and the chin that make up a particularly salient line of coplanar points. Pizlo and Stevenson (1999) have demonstrated that coplanar lines, like the one running down the middle of faces, can provide rich structural information when placed into motion. This requires that the observer identify the points as coplanar, which should not be a issue given the fact that humans have vast experience with faces. Pizlo and Stevenson (1999) find that shape constancy is best achieved by a stimulus containing planar contours, and that other relations such as symmetry and topological stability also contribute. Human faces contain both bilateral symmetry and a salient line of coplanar points, suggesting that motion may provide strong structural information through the use of these invariants. Thus motion information may provide at least limited structural information about the face that may assist the recognition of faces at symmetric orientations. Note that merely placing an object into motion may not automatically provide additional structural information that would allow recognition of novel viewpoints: Bülthoff and Edelman (1992) compared motion versus static study conditions for wire-frame and amoeba objects, and found no benefit for the motion conditions. These objects were without constraints such as symmetry and readily-identifiable coplanar points, and as a result the motion information did not prove particularly useful when generalizing to novel viewpoints. This reinforces the conclusion by Pizlo and Stevenson (1999) that it is the interaction between regularities in the stimuli and motion that provides structural information.

Several authors have provided recent investigations of how motion might enhance recognition of human faces, using a variety of motion manipulations. A recent review by O'Toole, Roark and Abdi (2002) differentiates between rigid and non-rigid motion. Rigid motion, such as that produced by head turns and nods may contribute to structure-from-motion as suggested by Pizlo and Stevenson (1999) for simpler stimuli. This motion can contribute to identity judgments (Hill & Johnston, 2001) and improve recognition of the faces at similar orientations (Lander, Christie & Bruce, 1999). Non-rigid motion such as expression changes have also been investigated, and several authors have suggested that this enhances priming (Thornton & Kourtzi, 2002) and recognition (Knappmeyer, Thorton & Bülthoff, 2003). Pike, Kemp, Towell, and Phillips (1997) compared rigid motion that encompassed the entire 360° rotation of the head with two static conditions and found better recognition for static test views with the full motion. However, Christie and Bruce (1998) found no benefit for moving stimuli at either study or test. These differences may result from whether familiar or unfamiliar faces are used (O'Toole et. al, 2002).

The work most similar to the present manipulations is that of Hill, Schyns and Akamatsu (1997), who used motion sequences with untextured 3D laserscan models of faces. Their results suggest that motion enhances face recognition, although they did not address symmetric orientation effects. They rotated faces from one profile to another and back through a sequence of 5 frames, presenting each frame for 100 ms. In a control condition the ordering of the frames was randomized. Their motion experiment did not allow an investigation of symmetric orientation effects (or performance at any novel orientations), since all 5 test views spanning both sides of the frontal view were shown in the study sequence. The authors report only an overall increase in performance in the motion condition relative to the random motion control, which could imply only that motion stimuli are easier to look at or more informative than random motion displays. This leaves open the question of how motion enhances the recognition of faces: does it provide more overall information by virtue of the smooth nature of the motion, or does motion enable the acquisition of specific types of information (such as limited 3-dimensional structure) that can assist recognition at particular novel viewpoints such as the symmetric orientation? If it does, this would suggest that motion allows observers to extract additional information than that provided by a single image-based representation.

We constructed smooth motion sequences by rotating the face around the vertical axis and compared this condition to two control conditions designed to reduce or eliminate the motion percept while holding constant the total information available in the sequence. For each of the two study views, five frames were generated for each face, which includes two views that are ± 7.5° from the study view, and two views that were ±15° away from the study view. Rather than use all 5 study views from Experiment 3, we limited our study views to 70° and 35°, and generated frames around these views to create motion sequences. The top images in Figure 4 show the 5 views centered around the 35° study angle, and the bottom row shows the 5 test orientations.

We used three motion sequences. The first provides a smooth motion percept by sequentially showing the 5 frames for 180 ms each for a total of 5760 ms. This results in a rich sensation of the head smoothly rotating back and forth for 4 complete cycles. A random motion control condition was generated by taking the frames from the smooth motion condition and randomizing the sequence such that no two frames repeated. While some accidental sequences will produce a slight perception of motion, overall the sequence shows a face randomly jerking from one view to the next, preventing acquisition of a smooth motion percept. The second control condition reorders the frames so that a slow motion sequence is shown; each view is visible for 720 ms, allowing the face to move through only one complete cycle in 5760 ms. It should be stressed that only the ordering of the frames changed across the motion conditions; the total display time is constant, as is the duration of each frame (although the slow motion case has frames that are ordered such that they repeat in order to produce slow motion). In addition, the static information available to the observer is the same across the three conditions. Thus the only differences between the three conditions was the nature of the motion itself.

We anticipated that the differences between the control and experimental conditions might be somewhat small, and to ensure that we measured any effects of motion on the perception of the symmetric view, we tested the exact symmetric orientation. Thus the five test orientations were 70°, 35°, 0°, -35° and -70°. A face was shown *either* at 70° or 35° in one of the 3 motion conditions, and tested using static images at all 5 orientations.

Method

Observers

Observers were 149 Indiana University students who received course credit for their participation.

Stimuli

The head models were identical to those used in Experiment 2. Movie sequences were created by rendering 4 views around the 70° and 35° study views that were ±7.5° and ±15° away from the study views. These were presented in full color on the PowerMac 7200, which is able to write the individual images in less than a screen refresh, making the effective ISI between frames 0 ms. Each frame of the motion sequence was on display for 180 ms, for a total of 5760 ms per sequence.

Design and Procedure

Twelve faces were used in the study session and shown in one of the 6 conditions (2 study orientations x 3 motion conditions) according to a randomized counterbalancing schedule. The instructions were similar to the Inclusion instructions in Experiment 2, with the exception that the practice-study period contained faces that were placed into motion. Observers were told that they would have to recognize the faces at novel orientations, and the practice test session verified that they understood that a previously-seen face at test at a novel orientation was still an old face.

Following the practice study and practice test session, the observers viewed the 12 study faces and then made old/new judgments on all 120 test faces. These faces were the 12 target faces shown at 5 test orientations (70°, 35°, 0°, -35° and 70°), as well as the 12 distractor faces also shown at the 5 test orientations.

Results and Discussion

Observers studied only left-facing faces in Experiment 3, and this may introduce small criterion shifts for right-facing faces. Memory performance is expected to be lower for right-facing test angles, and observers might shift their decision criterion to compensate. As anticipated, there were small but significant differences between the false alarm rates, which were .42, .45, .43, .40 and .39 with a common SEM = 0.010. As a result, we report only sensitivity (d') data for Experiment 3.

The data for Experiment 3 are shown in Figure 5 for the two study orientations, and demonstrate similar effects. The abscissa shows the 5 test orientations, and the three motion conditions produce the three curves. The error bars represent one standard error of the mean.

Data from the 70° Study View

The left panel of Figure 5 shows the data from faces that rotated around the 70° study view. Performance is highest for angles tested at the study view of 70°, and falls for the frontal view, and then shows a recovery for the smooth motion condition. A repeated-measures ANOVA revealed an interaction between motion condition (Smooth, Random or Slow) and test angle ($F(8, 1312)=2.4$, $p<0.05$). The critical comparison between smooth and random motion sequences reveals a significant interaction ($F(4, 656)=3.8$; $p<0.01$); in addition, the smooth and slow motion condition also interact with test angle ($F(4, 656)=2.66$; $p<0.05$). Inspection of Figure 5 reveals that these interactions do not result from better performance for the smooth motion condition at the study orientation of 70°, where there were no differences between the three conditions ($F(2, 328) = 2.08$, $p = .13$). In addition, there was no main effect of motion condition ($F(2,328)<1$), indicating that the smooth motion condition did not provide better performance overall.

Instead, the interactions between motion condition and test angle derive from differences between the three conditions at the -35° and -70° test angle, which are near the symmetric orientation. Restricting the analyses to the -70° test angle reveals a significant difference between the Smooth and Random conditions ($F(1, 164)=4.54$, $p<0.05$), although not between the Smooth and Slow conditions ($F(1,164)=1.55$, $p>0.05$). When the nearby test angle of -35° is included in the analysis, the smooth motion condition dominates both the Random condition ($F(1,164)=4.81$, $p<0.05$) and the Slow condition ($F(1,164)=6.85$, $p<0.01$).

All of these findings are consistent with a process that uses aspects of the motion to derive a representation that assists recognition of the symmetric orientation view. Interestingly, the beneficial effects of smooth motion appear to be restricted to angles at or near the symmetric orientation, and do not facilitate the original view, which might instead still rely on relatively raw image-based information such as a template match. Thus observers make use an image-based match where possible, and rely on other representations where the match fails.

Data from the 35° Study View

The right panel of Figure 5 shows the data from the 35° study view, and show effects similar to those from the 70° study view in the left panel of Figure 5. As with the 70° data, the interaction between motion and test angle is significant ($F(8, 1312)= 3.16$, $p<0.01$). The critical comparison between smooth and random motion sequences again reveals a significant interaction ($F(4, 656)=4.3$; $p<0.01$); in addition, the smooth and slow motion condition also interact with test angle ($F(4, 656)=4.45$; $p<0.01$).

As with the 70° study view data, there is no main effect of motion condition, indicating that the smooth motion condition does not provide more information overall ($F(2,328)=1.29$, $p<0.05$). Nor are there significant differences between the three motion conditions at the 35° study view ($F(2,328)<1.0$). Instead, the interactions result in part from superior performance of the smooth motion condition at the symmetric test angle of -35°, although this effect was only marginally significant ($F(1,164)=2.93$, $p = .089$).

For the 35° study condition, the beta values show a small effect of test orientation ($F(1,328) = 2.5$, $p<0.05$) but no effect of motion and no interaction between motion and test angle. This main effect stems from slightly lower beta values at the 30° test view. There were no effects of beta for the 70° test orientation.

Motion Assists the Recognition of the Symmetric Orientation View

The recovery of performance at the symmetric orientation view only for the smooth motion and the interaction between motion and test angle are both consistent with a process that uses motion to derive a representation that assists in the recognition of the symmetric orientation view. How might this be accomplished, and why should it depend on speed? We offer one possible mechanism that is based on the recovery of local 3-dimensional information in the form of the contour of the face. There are likely other possible explanations based on image-based transformations such as flow-fields or improved image interpolation. Further work may be needed to distinguish between alternative representations.

Extraction of Local 3D Structure

As discussed above, Pizlo and Stevenson (1999) have suggested that motion may enhance the perception of lines of *coplanar points*, and this might provide good recognition at the symmetric orientation. Faces are (mostly) bilaterally symmetric, and as a result have a series of features run down the center of the forehead and down the nose and chin that make up a particularly salient line of coplanar points. Observers must identify the points as coplanar, which should not be an issue given the fact that humans have vast experience with faces. Pizlo and Stevenson (1999) find that shape constancy is best achieved by a stimulus containing planar contours, and that other relations such as symmetry and topological stability also contribute. Our faces contain both bilateral symmetry and a salient line of coplanar points, suggesting that motion may provide strong structural information through the use of these invariants. Thus motion information may provide at least limited structural information about the face that may assist the recognition of faces at symmetric orientations. Note that merely placing an object into motion may not automatically provide additional structural information that would allow recognition of novel viewpoints: Bülthoff and Edelman (1992) compared motion versus static study conditions for wire-frame and amoeba objects, and found no benefit for the motion conditions. These objects were without constraints such as symmetry and readily-identifiable coplanar points, and as a result the motion information did not prove particularly useful when generalizing to novel viewpoints. This reinforces Pizlo and Stevenson's (1999) conclusion that it is the interaction between regularities in the stimuli and motion that provides structural information.

Our test stimuli were static, thus requiring observers to extract information about the contour of the face from a single image. The frontal view provides essentially no information about the contour of the face, but the symmetric orientation view does allow some contour information to be inferred, thus supporting good recognition at the symmetric orientation view, but only for the smooth motion condition.

Performance at the study view

For both the 70° and 35° study views, performance at the study view did not differ for the three motion conditions. Why is there no benefit overall for the smooth motion condition? One explanation is that all three motion conditions provide enough image-based information that allows an image-based representation such as a template match to recognize the original view. However, this representation does not work well for the symmetric orientation due to texture or structural asymmetries, and thus the observer relies on additional information, possibly including motion-assisted contour shape information. The data from Experiment 3 are consistent with a model in which observers rely on information derived from smooth motion to enhance the recognition at the symmetric orientation view, which augments the image-based representation. This extraction of depth information requires combining information across orientations, and there may be specific timing constraints on this process that make the smooth motion condition generalize better to the symmetric orientation, but prevents the slow motion condition from similarly generalizing. Note that the total static information was held constant in all three motion conditions, and the different motion conditions were generated by re-ordering the frames in the movies to produce different sequences.

In summary, the results of Experiment 3 suggest that observers not only extract information from a raw image-based representation and form abstractions that overcome the texture asymmetries (as was seen in Experiments 1 and 2) but also use information derived from smooth motion to enhance performance at the symmetric test angle. One possible explanation for this selective improvement is that observers use regularities in coplanar points to estimate depth parameters. This depth information could be used at test, but only for test angles that provide similar estimates of depth. The frontal view, for example, is very difficult to extract depth estimates from, and therefore we see no improvement at this test angle. The original study angle can still rely on the raw image-based representation for all three conditions, which leads to no improvement at the original angle. In combination these two effects provide no overall benefit of smooth motion, but improved performance at the symmetric test angle.

Accounts of Additional Effects

In addition to the differences between motion conditions described above, Figure 5, right panel, shows that performance at the 70° test view for items studied at 35° also shows large differences between the three conditions ($F(2,328)=8.25$, $p<0.01$). Surprisingly, the smooth motion condition is much worse than either the random or slow motion conditions. We did not anticipate this result, and we offer only this admittedly post-hoc explanation.

For our stimuli, information in the 70° view about the ridge line of the eyes or the protrusion of the mouth (as estimated by the corner of the mouth) suggests more depth in the face than is actually present, because the 70° view may be misidentified as a pure profile view (see Figure 6). If observers try to obtain contour information from the static 70° view they will not match the contour provided by smooth motion rotations around 35°. That is, the smooth motion condition places a representation into memory that includes shape information that is difficult to match to the static 70° test angle, and this produces decrements in performance at the 70° test angles for the smooth motion condition.

Although it may be difficult to go from a 35° study stimulus to a 70° test stimulus, the reverse may be possible. The rotating 70° study stimulus may provide enough contour information to enable good recognition at the -35° test stimulus, which would account for the slightly better recognition performance at -35° than -70° in Figure 5. All of these differences support the idea that different representations are used at the original and symmetric orientations. The test views that correspond to the original orientation can be done with a template match since they don't have texture mismatches with the study images. However, the template match mechanism does not generalize well to nearby orientations. This may be why the 35° test angle does not benefit from smooth motion while -35° benefits due to its proximity to the symmetric orientation, which does benefit from the smooth motion at study.

General Discussion

The results of these three experiments demonstrate surprising differences between perception and memory that suggest that different representations are used to support recognition of the symmetric viewpoint. Although the symmetric orientation represents a special case, it is unlikely that the visual system would develop a representation specific to a single orientation, and thus it is likely that the conclusions obtained from the symmetric orientation data apply more generally to recognition at viewpoints that differ from the original orientation. Thus we are able to make more general conclusions about the nature of the representations at work in perception and memory tasks.

The results of Experiment 1 suggest that strong performance at the symmetric orientation depends on a close match between the study and test view images once the symmetric orientation view has been horizontally flipped. Since performance is disrupted by texture asymmetries, but this disruption is ameliorated by introducing a tilt to the test image, it seems likely that observers are relying on a template-match representation as suggested by view-based models for the upright images, but a more flexible representation able to overcome texture asymmetries in the tilted case. These more flexible representations may have properties similar to those described by structural description models such as local features, their relations, and possible 3D shape information.

In memory, the strong symmetric orientation performance and generalization to this orientation supported best by smooth motion all suggests the reliance on a representation that involves abstractions derived from the raw image information. The ability to overcome texture asymmetries and generalize to views that are not exactly the symmetric orientation are properties that are more consistent with structural description models, which is further supported by the abstraction of limited 3D shape information in the smooth motion condition of Experiment 3. Two other kinds of motion preserved the same static information across conditions, but apparently did not lead to the kind of recovery of depth information that enhances symmetric orientation recognition.

Why would perception and memory rely on different representations? One possibility is that view-based representations are easy to store and use, especially when conditions favor them. Comparing a test face to a single study face over short test intervals seems to provide a favorable (change) environment as long as texture asymmetries are not present. This representation may not be durable enough to survive in memory experiments, or perhaps the process of trying to compare a test face to multiple items in memory disrupts the template-match process. Instead, observers may view test faces as individuals and consider non-visual information such as that described by structural description models or even more abstract information such as similarity to known individuals or whether the person looks happy. Since this more abstract information is more likely to survive changes in orientation (for instance, expression can be determined from all orientations used), this switch in the nature of the representations would support better generalization, especially to the symmetric orientation where the abstraction of information from the test image would more closely match that from the study image. A related explanation suggests that differences across domains results from the use of information at different spatial scales. If a memory representation were coarser and lost specific texture information that could be used to make a template match, performance at the symmetric orientation might not suffer because texture asymmetries are no longer noticed.

Whatever the nature of these representations, the processes that enable recognition of the symmetric orientation are under cognitive control, since observers can and do differentiate between real and symmetric orientations if necessary, as in Experiment 2. This occurs even when orientation is not explicitly encoded (or at least observers were not instructed to do so).

An open question is how recognition of the symmetric orientation view maps on to related fields such as the identification of symmetry within an object (such as a face seen in a frontal view). Despite a fairly well-developed field (See Tyler, 1994, 1995, Wagemans, 1995, for reviews), no consensus has been reached about the underlying neurophysiological underpinnings of symmetry detection. However, such a model, when developed, may also be able to account for the image-based representation proposed by Troje and Bülthoff (1998). Whether it can be extended to include elements of motion and categorical relations between individual features, or whether an additional representation is required, must then be addressed. This remains a rather active area of research in our laboratory.

Author Notes

The authors gratefully acknowledge Alice O'Toole, who provided the 3D laserscan models. Portions of this research were supported by an NIH grant to Thomas Busey.

References

Beier, T., & Neely, S. (1992). Feature-based image metamorphosis. *Computer Graphics, 26,* 35-42.

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

Bülthoff, H.H. & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Science USA, 89,* 60-64.

Christie,F & Bruce,V. (1998) The role of dynamic information in the recognition of unfamiliar faces. *Memory and Cognition, 26,*780-790.

Hildreth, E., Grzywacz, N., Adelson, E., & Inada, V. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics, 48,* 19-36.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology, 11*, 880-885.

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

Hummel, J. E. & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480-517.

Hummel, J. E. (2000). Where view-based theories break down: The role of structure in shape perception and object recognition. In E. Dietrich and A. Markman (Eds.). *Cognitive Dynamics: Conceptual Change in Humans and Machines* (pp. 157 - 185). Hillsdale, NJ: Erlbaum.

Hummel, J. E. & Stankiewicz, B. J. (1996a). Categorical relations in shape perception. *Spatial Vision, 10,* 201-236.

Hummel, J. E., & Stankiewicz, B. J. (1996b). An architecture for rapid, hierarchical structural description. In T. Invi & J. McClelland (Eds.), *Attention and performance*

XVI: Information integration in perception and communication (pp. 93-121). Cambridge, MA: MIT Press.

Hummel, J. E., & Stankiewicz, B. J. (1998). Two roles for attention in shape perception: A structural description model of visual scrutiny. *Visual Cognition, 5,* 49-79.

Knappmeyer, B., Thornton, I.M., & Bülthoff, H. H. (2003). Facial motion can bias the perception of facial identity. *Vision Research 43*, 1921 -1936.

Lander, K., & Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory and Cognition, 27*, 974-985.

Liu, Z., Knill, D.C. & Kersten, D. (1995). Object classification for human and ideal observers. *Vision Research, 35,* 549-568.

McNamara, T. P., Diwadkar, V. A., & Blevins, W. A. (1998). Representations of apparent rotation. *Abstracts of the Psychonomic Society, 3,* p. 35.

Nickerson, R. S., and Adams, M. J. (1979). Long term memory for a common object. *Cognitive Psychology, 11*, 287-307.

O'Toole, A. J., Roark, D. & Abdi, H. (2002). Recognizing moving faces: A Psychological and Neural Synthesis, *Trends in Cognitive Sciences*,6, 261-266.

Patterson, K. and Baddeley, A. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3*:406-417.

Pelli, D. G. & Zhang, L (1991). Accurate control of contrast on microcomputer displays. Vision Research, 30, 1033-1048.

Pike, G., Kemp, R., Towell, N. and Phillips, K. (1997). Recognising moving faces: the relative contribution of motion and perspective view information. *Visual Cognition, 4,* 409-39.

Pizlo, Z., & Stevenson, A. (1999). Shape constancy from novel views. *Perception & Psychophysics, 61*, 1299-1307.

Price, C. M. & Gilden, D. L. (2000). Representations of motion and direction. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 18-30.

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

Stankiewicz, B. J., & Hummel, J. E. (1996). MetriCat: A representation for subordinate and basic level classification. *Proceedings of the 18th Annual Meeting of the Cognitive Science Society* (pp. 254-2259). Hillsdale, NJ: Erlbaum.

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

Tarr, M., & Pinker, S. (1990). When does human object recognition use a viewer-centered reference frame? *Psychological Science*, *1*, 253-256.

Thornton, I.M., & Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception*, *31* (1), 113 - 132.

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, *48*, 419-430.

Troje, N. (1998). Generalization to novel views of faces: Psychophysics and models concerning the role of bilateral symmetry. In: *Downward processing in the perception representation mechanism*, C. Taddei-Ferretti and C. Musio (eds), World Scientific, Singapore, New Jersey, London, Hong Kong, pp. 171-188.

Troje, N., & Bülthoff, H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research*, *36*(12), 1761-1771.

Troje, N., & Bülthoff, H. (1998). How is bilateral symmetry of human faces used for recognition of novel views? *Vision Research*, *38*, 79-89.

Turnbull, O., & McCarthy, R. (1996). Failure to discriminate between mirror-image objects: A case of viewpoint independent object recognition? *Neurocase*, *2*, 63-72.

Tyler, C. W. (1994). The perception of symmetry. Part I: Theoretical aspects [Special Issue]. *Spatial Vision*, *8*, 4.

Tyler, C. W. (1995). Empirical aspects of symmetry perception. *Spatial Vision*, *9*, 1-7.

Ullman, S. (1996). High-level Vision: Object Recognition and Visual Cognition. Cambridge MA: MIT Press.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32,* 193-254.

Vetter, T. & Poggio, T. (1994). Symmetric 3D objects are an easy case for 2D object recognition. *Spatial Vision*, *8*, 443-453.

Vetter, T., Hurlbert, A. & Poggio, T. (1995). View-based models of 3D object recognition: Invariance to imaging transformations. *Cerebral Cortex, 5,* 261-269.

Vetter, T., Poggio, T., & Bülthoff, H.H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. *Current Biology, 4,* 18-23.

Wagemans, J. (1995). Detection of visual symmetries. *Spatial Vision*, *9*, 9-32.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45,* 205-217.

Figure Headings

Figure 1. Example stimuli from Experiment 1, shown at -40°, 0°and 40° rotations.

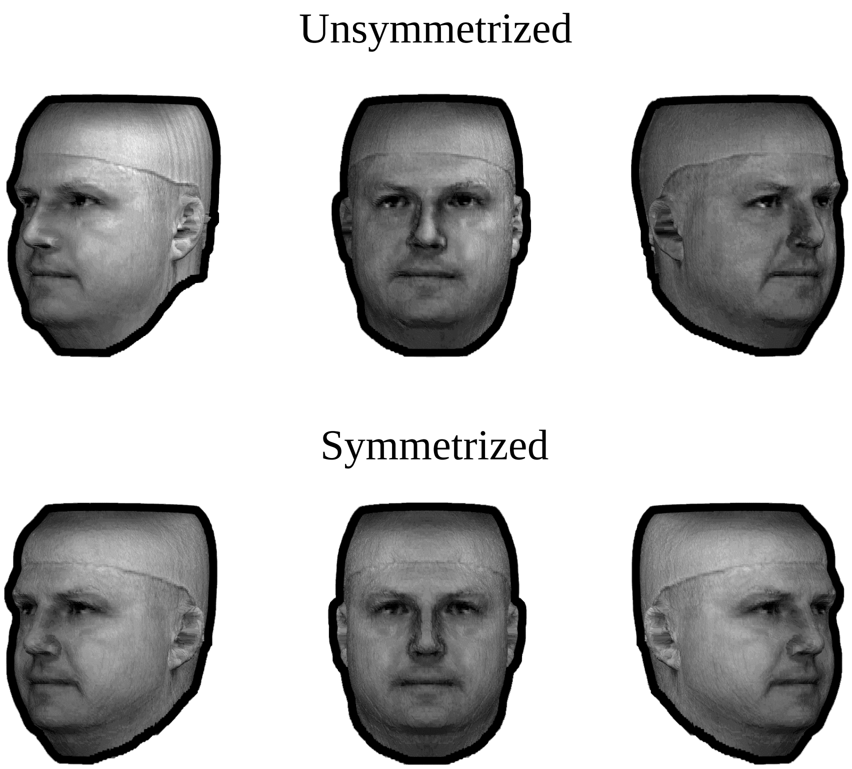
Figure 2. Sensitivity (d') results from Experiment 1. Data from unsymmetrized faces are plotted on the left side, while data from symmetrized faces are plotted on the right side. Stippled bars show data from the frontal views for each condition. Error bars represent 1 standard error of the mean (SEM).

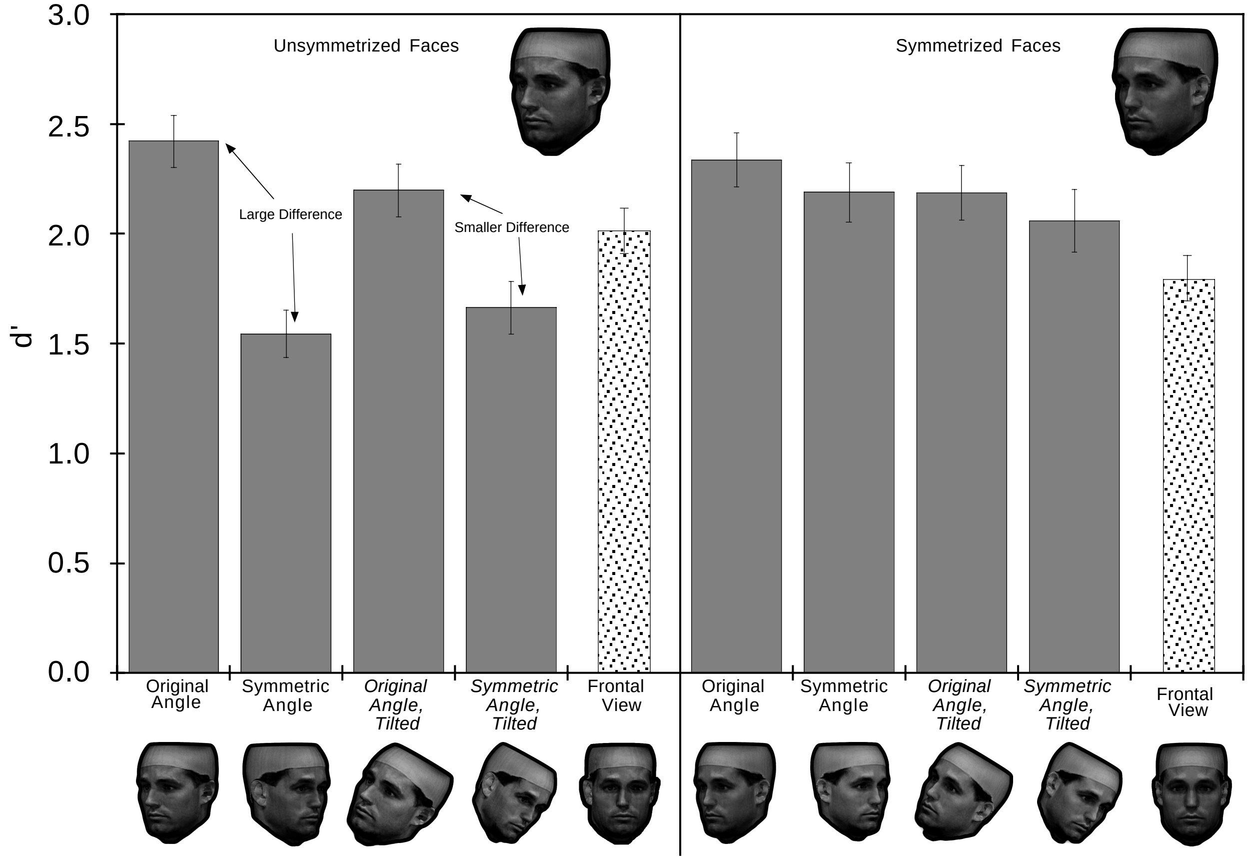
Figure 3. Sensitivity (d') data from Experiment 2 for all Study Angles. **Left Columns**: Results from the Inclusion instructions show large symmetric orientation effects at the relevant test angles for -50° and 70° study conditions (top left panel). **Right Columns:** Results from the Exclusion instructions show a complete absence of symmetric orientation effects. Error bars represent ± 1 SEM. Circled points represent those conditions where study and test angles correspond.

Figure 4. Experiment 3 stimuli. Top row: 5 frames of a motion sequence centered around the 35° viewpoint. These were shown for 5 complete cycles at study. Bottom row: Five test viewpoints used in Experiment 3, which were static.

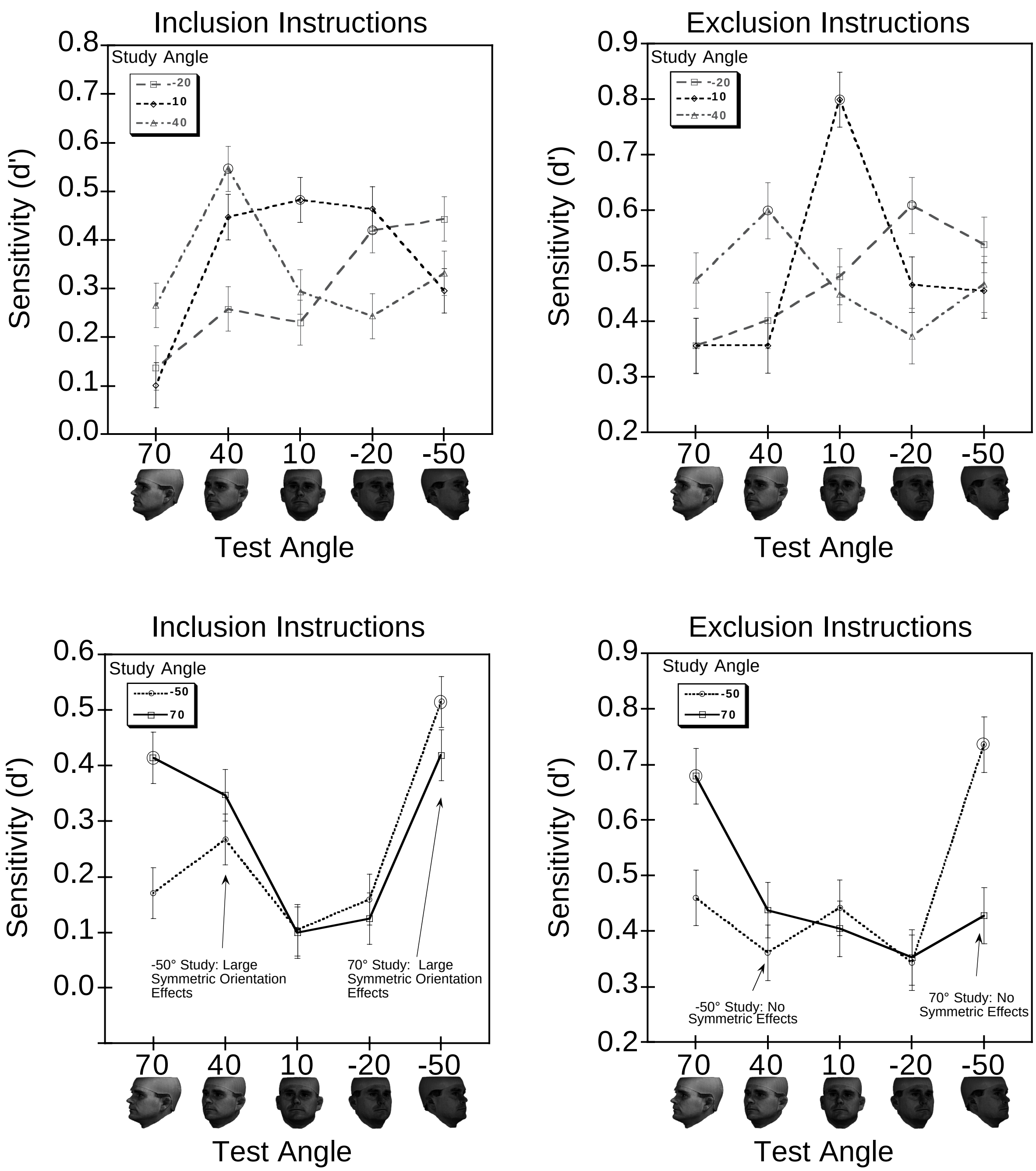
Figure 5. Data from Experiment 3. **Left Panel**. Sensitivity (d') values for faces studied at -70° and tested at all 5 test angles for the three motion conditions. **Right Panel**. Sensitivity (d') values for faces studied at -35°. Error bars represent 1 SEM. Note the selective improvement at the symmetric test angle for the smooth motion conditions.

Figure 6. Faces rotated to 70° and 90°. The 70° head may appear as a clean profile until the 90° is seen for comparison. This may disrupt the estimates of depth obtained from the profile views and lead to reduced performance for the smooth motion condition at the 70° test orientation.

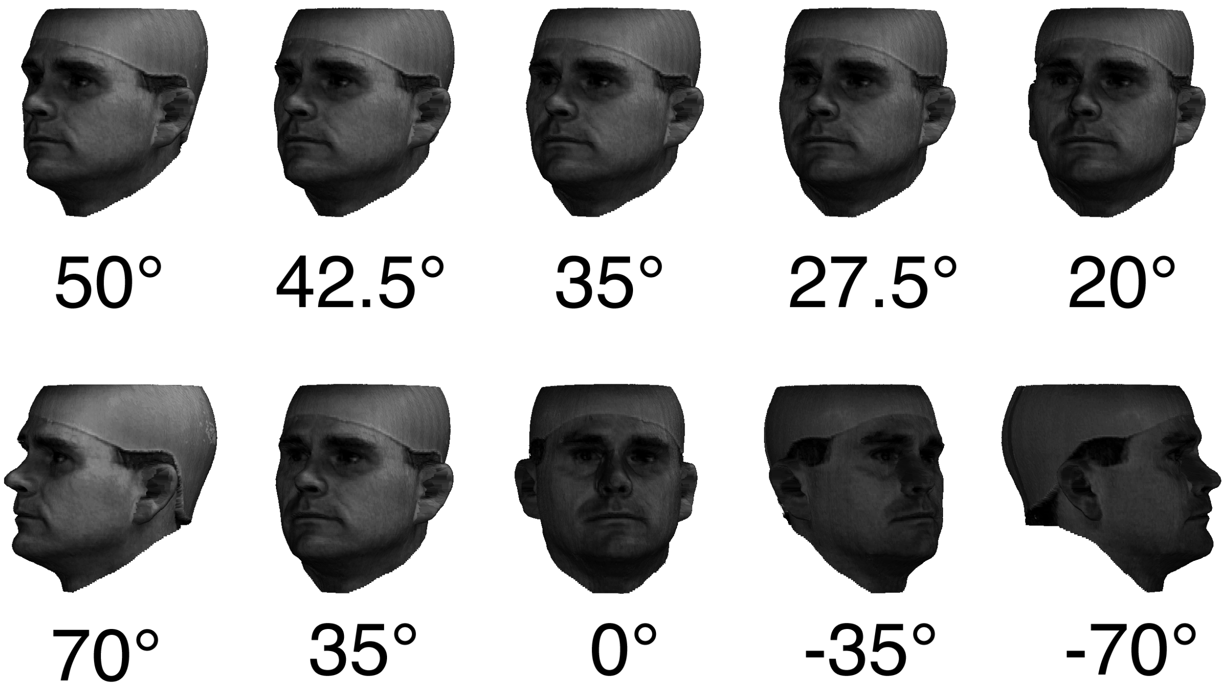

Unsymmetrized
Symmetrized

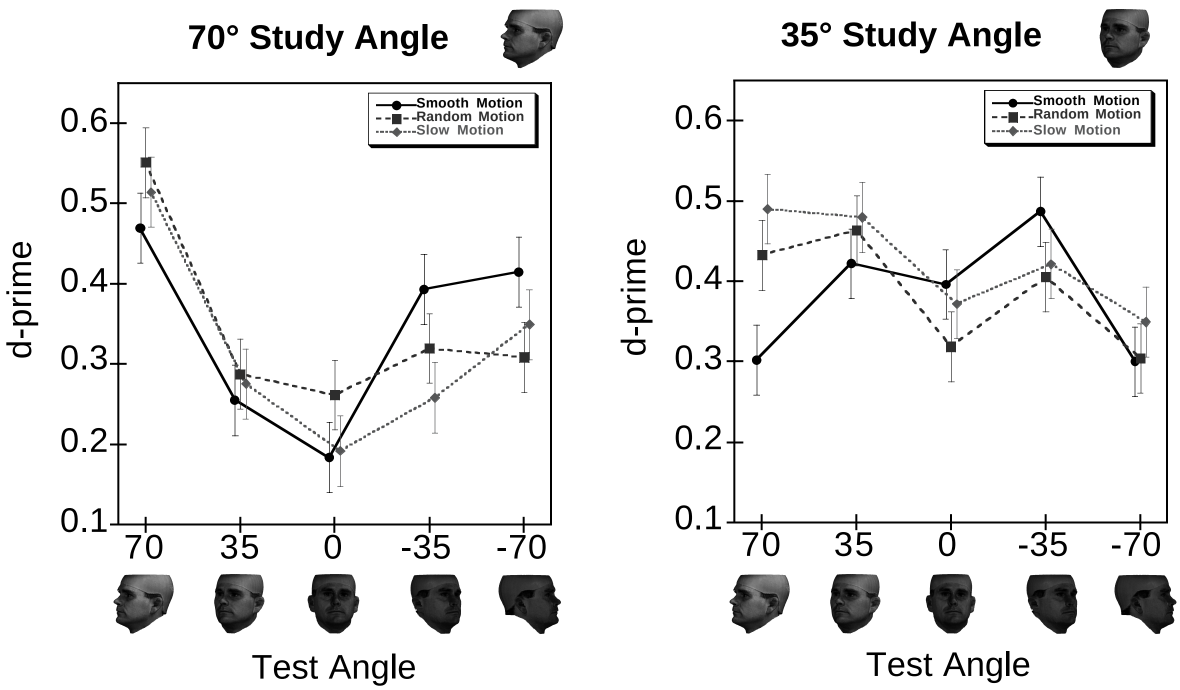


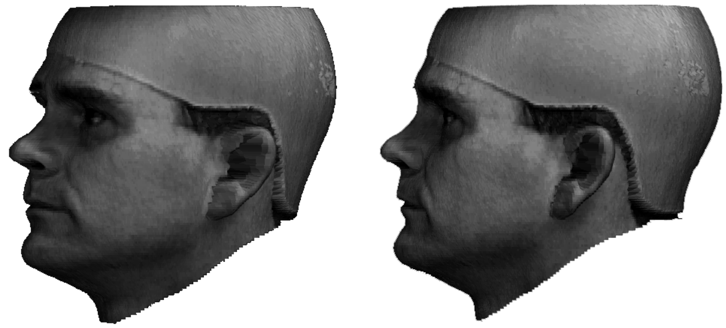
3.0
2.5
2.0
1.5
1.0
0.5
0.0
d'
Unsymmetrized Faces
Symmetrized Faces
Large Difference
Smaller Difference
Original Angle
Symmetric Angle
Original Angle, Tilted
Symmetric Angle, Tilted
Frontal View
Original Angle
Symmetric Angle
Original Angle, Tilted
Symmetric Angle, Tilted
Frontal View



Inclusion Instructions
Study Angle
-20
10
-40
Sensitivity (d')
Test Angle
70
40
10
-20
-50
Exclusion Instructions
Study Angle
-20
10
-40
Sensitivity (d')
Test Angle
Inclusion Instructions
Study Angle
-50
70
-50° Study: Large Symmetric Orientation Effects
70° Study: Large Symmetric Orientation Effects
Exclusion Instructions
Study Angle
-50
70
-50° Study: No Symmetric Effects
70° Study: No Symmetric Effects


50°
42.5°
35°
27.5°
20°
70°
35°
0°
-35°
-70°


70° Study Angle
Smooth Motion
Random Motion
Slow Motion
d-prime
0.6
0.5
0.4
0.3
0.2
0.1
70
35
0
-35
-70
Test Angle
35° Study Angle
Smooth Motion
Random Motion
Slow Motion
d-prime
0.6
0.5
0.4
0.3
0.2
0.1
70
35
0
-35
-70
Test Angle



70° 90°