

# Object features used by humans and monkeys to identify rotated shapes

**Kristina J. Nielsen**

Max Planck Institute for Biological Cybernetics,  
Tübingen, Germany



**Nikos K. Logothetis**

Max Planck Institute for Biological Cybernetics,  
Tübingen, Germany



**Gregor Rainer**

Max Planck Institute for Biological Cybernetics,  
Tübingen, Germany



Humans and rhesus monkeys can identify shapes that have been rotated in the picture plane. Recognition of rotated shapes can be as efficient as recognition of upright shapes. Here we investigate whether subjects showing view-invariant performance use the same object features to identify upright and rotated versions of a shape. We find marked differences between humans and monkeys. While humans tend to use the same features independent of shape orientation, monkeys use unique features for each orientation. Humans are able to generalize to a greater degree across orientation changes than rhesus monkey observers, who tend to relearn separate problems at each orientation rather than flexibly apply previously learned knowledge to novel problems.

Keywords: object recognition, view-invariance, *Macaca mulatta*

Citation: Nielsen, K. J., Logothetis, N. K., & Rainer, G. (2008). Object features used by humans and monkeys to identify rotated shapes. *Journal of Vision*, 8(2):9, 1–15, <http://journalofvision.org/8/2/9/>, doi:10.1167/8.2.9.

## Introduction

Object constancy – the ability to recognize an object despite the variations in appearance introduced by viewing the object from different angles – is considered to be a hallmark of object recognition. Clearly, under most circumstances, we have no problem recognizing a familiar object in a new orientation. Nonetheless, having to recognize an object from a novel viewpoint in general comes at a cost, as it requires longer processing times and may lead to higher error rates. Under certain conditions, the effects of viewpoint on performance disappear, and objects are recognized in a view-invariant manner. This seems to be the case when objects can be identified by few, very distinctive features, which remain diagnostic object characteristics despite changes in object rotation. For example, view-invariant performance has been observed for small stimulus sets, as well as stimuli that can be identified by small numbers of features (Lawson & Jolicoeur, 2003; Newell, 1998; Tarr, Bülthoff, Zabinski, & Blanz, 1997; Wilson & Farah, 2003). It has furthermore been demonstrated that drawing an observer's attention to the unique rotation-invariant features of an object can turn a view-dependent performance into a view-invariant performance (Liter, 1998; Wilson & Farah, 2003). Finally, repeated exposure to a stimulus set allows subjects to adopt a view-invariant strategy. Training effects are

limited to stimuli seen during the training sessions, and are thought to be generated because subjects become aware of the unique, rotation-invariant features of an object (Jolicoeur, 1985; Jolicoeur & Milliken, 1989). These findings pertain to rotations of objects in the picture plane, as well as object rotations in depth. The results are consistent with the idea that subjects show view-invariant behavior if they can use distinctive, rotation-invariant features to recognize objects. However, it remains to be shown that subjects indeed use the same feature to identify an object despite changes in view. Here we use “Bubbles” (Gosselin & Schyns, 2001) to directly determine whether observers showing view-invariant performance use the same feature to identify an object irrespective of its orientation.

Rhesus monkeys are the major animal model for human perception. When monkeys have to identify an object presented from multiple viewpoints after rotation in depth, they usually show a view-dependent behavior. However, training with multiple views of an object can generate view-invariant behavior, as shown in experiments by Logothetis et al. and Wang et al. (Logothetis, Pauls, Bülthoff, & Poggio, 1994; Logothetis, Pauls, & Poggio, 1995; Wang, Obama, Yamashita, Sugihara, & Tanaka, 2005). These studies indicate that monkeys – similar to humans – show view-invariant behavior if objects can be distinguished by few distinctive features. Both groups trained monkeys to discriminate target

objects from distractor objects. View-dependency of performance was assessed by presenting the target objects in different orientations, and observing the monkeys' discrimination performance. In the study by Logothetis et al. (1994 and 1995), the monkeys immediately performed view-invariant when targets and distractors were very differently shaped everyday objects. In contrast, identification of artificial target objects, which were much more similar to the artificial distractor objects, became view-invariant only after training. More recently, Wang et al. (2005) directly showed that the similarity between targets and distractors determined whether the monkeys could generalize across changes in viewpoint. Thus, these studies suggest that the availability of diagnostic features is an important determinant of how monkeys identify rotated objects. In this study, we perform the same experiment as with the human observers to identify the features of rotated shapes used by monkeys.

We tested humans and monkeys under very similar conditions, so the results allow a comparison of behavioral strategies across species. Recently, we showed that even if tested under the same conditions, monkeys and humans may use different strategies to discriminate between natural scenes (Nielsen, Logothetis, & Rainer, 2006). Furthermore, by looking at eye movement patterns, it has been suggested that monkeys predominantly direct their attention to low-level image features, whereas the allocation of attention is driven more strongly by high-level scene interpretations in humans (Einhäuser, Kruse, Hoffmann, & König, 2006). This study reports an additional interspecies comparison of cognitive strategy and performance, namely that of the strategy used for solving a complex generalization task such as the recognition of partially occluded objects presented from different views.

## Methods

### Subjects

Five human observers (2 males and 3 females) were tested. All subjects were naïve as to the purpose of the experiments. Informed consent was obtained from all subjects. Subjects had normal or corrected-to-normal vision. Testing sessions usually lasted around 1 hour, with subjects completing about 1,000 trials in this time. Subjects returned to the lab for additional sessions, with one session run per day. We collected 3,000 trials for all observers with the exception of AK, who performed 6,000 trials.

Two adult male Rhesus monkeys (*Macaca mulatta*) participated in the experiments. Before the experiments, a metal head post and a scleral search coil (Judge, Richmond, & Chu, 1980) were implanted under aseptic conditions (Lee, Simpson, Logothetis, & Rainer, 2005).

Monkeys had restricted access to water, but received their daily amount of liquid during the experimental sessions, and were provided with dry food ad libitum. All studies involving the monkeys were approved by the local authorities (Regierungspräsidium Tübingen), and were in full compliance with the guidelines of the European Community (EUVD, European Union directive 86/609/EEC) and the National Institutes of Health for the care and use of laboratory animals. The monkeys were tested daily, and performed between 500 and 1,000 trials per day. A total of 28,000 trials was analyzed for monkey G00, and 12,300 trials for monkey B98.

### Task and stimuli

Three geometric shapes were shown as black surfaces, centered on a gray background (see Figure 2). Each shape could be shown upright or rotated around the center of the shape (rotation in the picture plane). All stimuli were presented centrally. Their long axis subtended between 4.4 and 4.9 deg of visual angle.

For human observers, trials began with the presentation of a yellow fixation spot for 500 ms, followed by one of the stimuli for 500 ms (see Figure 1). Each stimulus was surrounded by a thin white frame of 12 by 12 deg of visual angle. Observers responded after the presentation of the stimulus by pressing designated keys on the numerical keypad of a standard computer keyboard. Each of the shapes in a stimulus set was associated with a specific response key, independent of the shape's orientation. The subjects were informed about the response mapping before the start of the first session by providing them with a printout of the upright stimuli and their assigned response keys. Subjects then performed 20 training trials with the upright stimuli only, followed by trials with upright and rotated shapes. All subjects could immediately perform the task at ceiling with these shapes. No constraints were imposed on reaction time, and eye movements were not monitored. We have shown as part of a previous experiment that imposing fixation constraints has no influence on the results of the Bubbles experiments (Nielsen et al., 2006). With the exception of observer AK, no feedback was given about the correctness of an answer. Observer AK received feedback in form of a + or – shown after the response for 300 ms. Because the results of this observer were not different from the results of the other observers, we conclude that feedback does not affect our results.

For the monkey observers, trials began with the presentation of a yellow fixation spot and the sounding of a tone (see Figure 1). The monkeys were required to fixate the fixation spot for 100 ms, after which time the spot was replaced by one of the three stimuli for 300 ms. During stimulus presentation, the monkeys had to keep their gaze within 3 deg from the center of the screen. At stimulus offset, three white squares (the targets) were

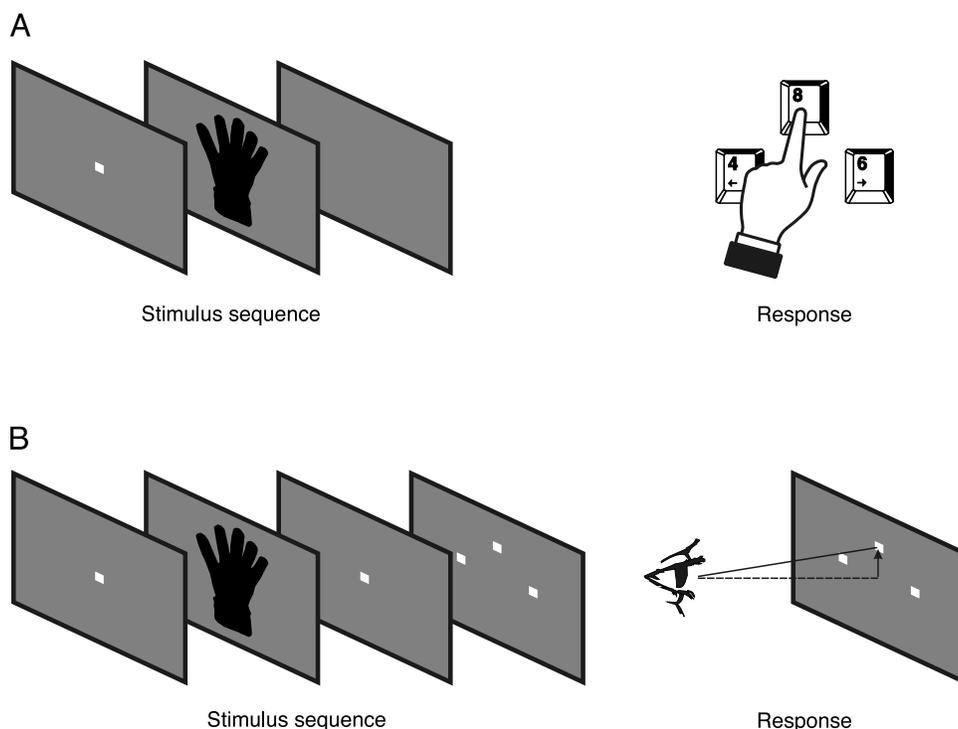


Figure 1. Experimental paradigm. (A) Paradigm for human observers. (B) Paradigm for monkey observers. In both panels, the left side shows the sequence of stimuli as they appear on the screen. The right side indicates the response modality (button press for human observers; saccade for monkey observers).

presented at 6 deg eccentricity. Each of these squares was associated with one of the shapes. A saccade to the correct target was rewarded by a drop of juice. Rotated versions of a shape were associated with the same target as the upright shape. Monkeys were taught the association between the upright shapes and the saccade targets by introducing a brightness cue between the targets, with the correct target being brighter than the other targets. This brightness cue was gradually removed as the monkeys' performance improved. A brightness cue was used only for the upright shapes; the monkeys learned the association between rotated shapes and response targets through trial-and-error.

The diagnostic regions of the upright shapes and their rotated versions were determined using Bubbles (Gosselin & Schyns, 2001). During the Bubbles sessions, the differently oriented versions of each shape, as well as

the different shapes, were presented in a pseudorandom order. During Bubbles trials, shapes are shown behind trial-unique occluders (for examples, see Figure 2). Occluders were generated as described previously (Nielsen et al., 2006). Briefly, occluders consisted of non-transparent surfaces, punctured by round windows (bubbles). Each occluder had a size of 6 by 6 deg of visual angle, corresponding to 256 by 256 pixels. When shapes were presented behind the occluders, parts of the shapes were visible through the bubbles. Bubbles had a profile of a 2D Gaussian with a standard deviation of 14 pixels. Bubbles were randomly positioned in the occluders, with the restriction that the center of each bubble fell within 3 deg of visual angle from the center of the screen to remain within the boundaries of the occluder, and the centers of two bubbles were not identical. The number of bubbles per occluder was adapted to each subject's



Figure 2. Appearance of the stimuli during the Bubbles session. This figure shows the same stimulus behind four different occluders. For this example, each occluder was generated by randomly placing three windows (bubbles) in an otherwise non-transparent surface.

performance. For the human observers, we used a staircase protocol for this purpose. Staircases were run independently for each image in a stimulus set, and converged to a performance of 75% correct. After every fourth trial of an image, the numbers of bubbles were updated. The number was decreased by three if the image had been identified correctly in the last four trials, and increased by two if less than three trials had been correct. For monkeys, we used a modified staircase procedure. During a session, the numbers of bubbles were identical for each image so as not to serve as a potential cue. Initially, bubbles numbers were set to a value at which the monkeys could perform the task at ceiling performance. After 15 trials, the number of bubbles was successively decreased by a fixed amount until the monkey's performance dropped below 70% correct. At this point, the number of bubbles was reset to the initial value, and the cycle was restarted. We maintained presentation of unoccluded stimuli throughout the Bubbles sessions (10% of trials for human observers and 40% for monkeys) as a baseline control of performance.

## Setup

Monkeys performed experiments in acoustically shielded chambers. Eye movements were monitored using the scleral search coil technique (Robinson, 1963) and digitized at 200 Hz. Stimuli were presented on a 21" monitor (Intergraph 21sd115, Intergraph Systems, Huntsville, AL, USA) with a resolution of 1024 by 768 pixels, and a refresh rate of 75 Hz. Background luminance of the monitor was set to 41 cd/m<sup>2</sup>, and the monitor was gamma corrected. The monitor was placed at a distance of 95 cm from the monkey. Stimuli were generated in an OpenGL-based stimulation program under Windows NT. Similar equipment was used for human observers, who were seated 85 cm from the monitor (background luminance of 27 cd/m<sup>2</sup>).

## Data analysis

Analyses were carried out in Matlab (The Mathworks, Natick, MA, USA). To analyze the Bubbles data, we compared the occluders from trials in which a stimulus was correctly identified with the occluders from incorrect trials. More specifically, we compared for each pixel in the occluder the distribution of occluder values in the correct trials against the distribution of occluder values in the incorrect trials using a Kolmogorov-Smirnov test. Occluder values ranged from 0 (occluder pixel non-transparent) to 1 (occluder pixel transparent). The resulting *p* values were Bonferroni corrected for the number of occluder pixels to account for the multiple comparisons. Diagnostic regions were defined to include all pixels with *p* values below the 5th percentile for a particular stimulus.

The amount of overlap expected by chance for a pair of diagnostic regions was estimated using a Monte-Carlo simulation. In each run of the Monte-Carlo simulation, the diagnostic region of the rotated shape was randomly repositioned. For this purpose, the largest continuous subregion of each diagnostic region was approximated as a polygon using Matlab's "regionprops" command. Repositioning of the approximated diagnostic region was achieved by adding a random offset in horizontal and vertical direction to the polygon's center. Since the diagnostic regions were often positioned close to the center of the stimuli, offsets were limited to the range of -128 to +128 pixels, so that most of the diagnostic region remained within the boundaries of the stimulus. Since only the largest subregion of the diagnostic region of the rotated shape was used in this computation, the Monte-Carlo results underestimate the overlap expected by chance. However, the polygon sizes accounted for between 88% and 96% of the sizes of the original diagnostic regions. One thousand repetitions were run for each pair of diagnostic regions, and the critical amount of overlap was estimated as the 95th percentile of the generated distribution.

## Results

### Experiments with human observers

In previous studies, viewpoint-independent behavior was usually observed for small stimulus sets consisting of very different stimuli (Jolicoeur, 1985; Takano, 1989). Because we were interested in the usage of shape features during view-invariant performance, we used a stimulus set consisting of three common objects with very different shapes (a bottle, a hand, and a drain pipe; see Figure 3). These shapes were shown as black surfaces on a gray background. Rotation of an object in depth can lead to the disappearance of one feature, and the appearance of new features. In this case, observers have to use different features to recognize shapes, depending on the viewing angle. However, when objects are rotated in the picture plane, the same features are visible at each orientation. Since the usage of features can thus more easily be compared across orientations for rotations in the picture plane, we restricted the experiment to these rotations.

The subjects always had to discriminate between the three aforementioned shapes. Shapes were shown for 500 ms on a computer screen. After the presentation, the subjects had to press one of three keys on the computer keyboard, with each shape being assigned to a particular key. Observers learned the discrimination task with upright stimuli only. Thereafter, they were tested with rotated shapes. All subjects could immediately generalize the task from upright shapes to other orientations. We

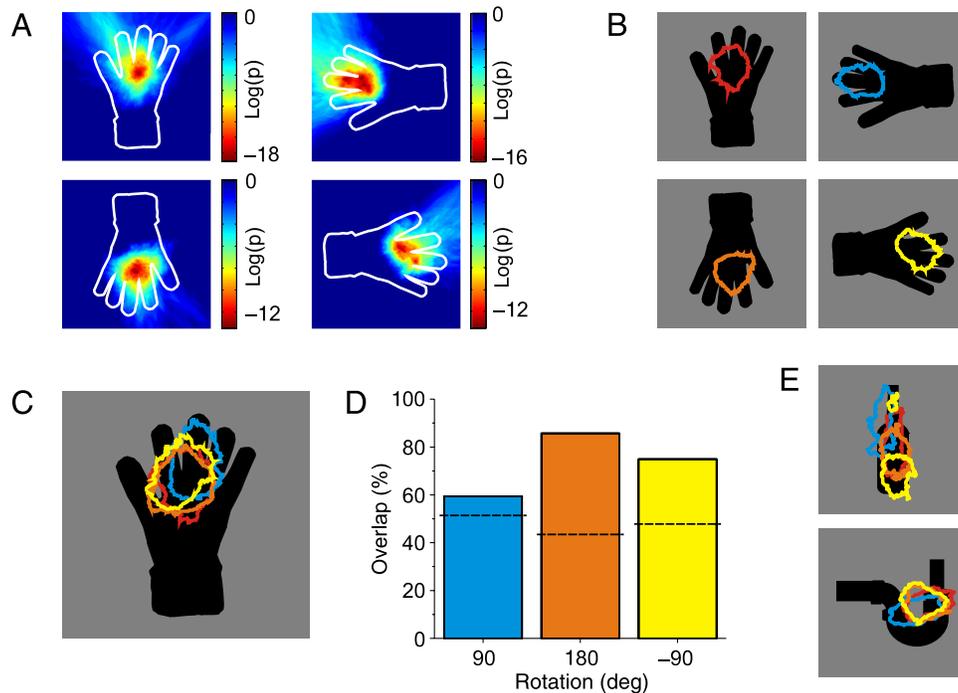


Figure 3. Data of an exemplar subject (VB). (A) Results of the Kolmogorov-Smirnov test performed on the Bubbles data. The color indicates the significance of differences between occluders from correct and incorrect trials. The outline of the shape is superimposed on each plot as a reference. (B) Diagnostic regions for the differently oriented versions of a shape. (C) Diagnostic regions superimposed on the upright shape. For this plot, the diagnostic regions of rotated shapes were rotated to the upright. The same colors are used as in (B) to indicate the different stimulus orientations. (D) Amount of overlap computed from (C). Dashed lines indicate the level of overlap that has to be exceeded to reach a level of  $p < .05$ , as determined by the Monte-Carlo simulations. (E) Diagnostic regions for the other two shapes in the set. Again, the diagnostic regions of the rotated shape versions were first brought to the upright before plotting. The same color scheme as in (B) and (C) was used.

proceeded to use Bubbles to determine which shape regions were used by the observers to identify each shape. During the Bubbles sessions, shapes could appear upright and rotated. The subjects continued to identify each presented shape irrespective of its orientation by pressing one of the designated keys on the computer keyboard. On every trial, a shape was seen behind a randomly generated occluder, which consisted of an occluding surface punctured by randomly placed round windows (see [Methods](#)). Parts of the shape were visible through these windows. Depending on which parts were visible, observers could or could not identify the shape. Hence, the stimulus features supporting behavior – the “diagnostic regions” – were identified by comparing the occluders from trials in which a shape was correctly identified to the occluders from trials with incorrect responses. We used Kolmogorov-Smirnov tests to compare occluders from correct and incorrect trials (see [Methods](#)). This method has previously been successfully applied to detect the features used by observers to discriminate between sets of natural scenes (Nielsen et al., 2006). As an example, the results of the Kolmogorov-Smirnov test are shown for one subject and a selected shape (the hand) in [Figure 3A](#). This subject was tested with shapes presented at four different orientations, each separated by 90 deg. The plots indicate whether

occlusion reliably influenced the subject’s performance for a particular part of the hand shape. A region adjacent to the middle finger consistently showed the strongest influences of occlusion, independent of the orientation at which the shape was presented.

To quantify the degree to which observers used the same shape features to identify upright and rotated versions of a shape, we first computed the diagnostic region for each stimulus, which consisted of the 5% stimulus pixels with the lowest  $p$  values. Thus, diagnostic regions contained the stimulus regions with the strongest influence of occlusion, and always had the same size. The resulting diagnostic regions for the exemplar case are shown in [Figure 3B](#). For the rest of the analysis, we used the diagnostic regions determined for the upright shapes as a reference. To quantify similarity in diagnostic regions across stimulus rotations, each rotated shape, and with it its diagnostic region, was first rotated to the upright. We then computed the overlap between the reoriented diagnostic region of a rotated shape’s versions, and the diagnostic region of the upright version of the shape. Overlap was measured in percentage of the size of the upright diagnostic region. [Figure 3C](#) plots the actual upright diagnostic region for the exemplar case, as well as the diagnostic regions of the other orientations after

rotation to the upright. The resulting overlap is plotted for the tested orientations in Figure 3D. The results for the other two shapes, obtained for the same subject, are plotted in Figure 3E.

Five human observers were tested in the experiments. Two observers were tested with shapes presented at 0 deg, 90 deg, 180 deg, and  $-90$  deg (where positive angles indicate counterclockwise rotations). Three additional observers were presented with shapes rotated by smaller angles. In these experiments, shapes were shown at 0 deg,  $\pm 30$  deg, and  $\pm 90$  deg. The results for the 90 deg orientations were the same across both groups of observers: no significant difference in overlap between the two groups,  $t$  test, 90 deg:  $t(13) = -0.6$ ,  $p = .5$ ,  $-90$  deg:  $t(13) = 1.8$ ,  $p = .09$ . Furthermore, the general pattern of results was the same in both groups, allowing us to pool the results of all subjects. We verified that all subjects maintained view-invariant behavior throughout the Bubbles sessions. For this purpose, control trials were

introduced in the Bubbles sessions. In control trials, shapes were shown without occluders, and therefore could be used to monitor a subject's behavior on the basic discrimination task. In the control trials, the average performance yielded a level of  $96.0\% \pm 1.4\%$  correct responses. There was no difference in the performance rate between different orientations, repeated measures ANOVA,  $F(4) = .2$ ,  $p = 1.0$ .

Figure 4A (black circles) plots the average overlap between the diagnostic regions of the rotated shapes (aligned to the upright orientation) and the corresponding upright diagnostic regions. This plot indicates that the average overlap was not strongly influenced by the orientation of a shape. Indeed, an ANOVA on the overlap resulted in a nonsignificant influence of stimulus orientation,  $F(4) = .95$ ,  $p = .4$ . For all tested stimulus orientations, the diagnostic regions for the rotated shapes overlapped about 60% of the diagnostic regions of the upright shapes.

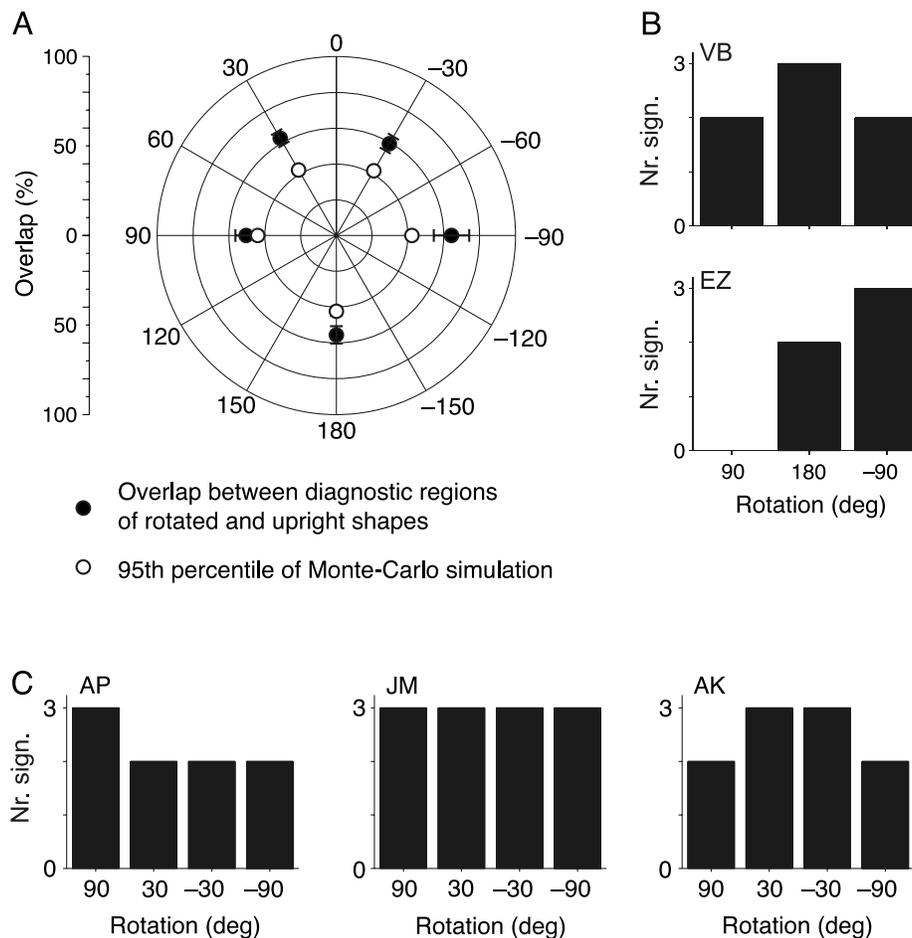


Figure 4. Consistency in the diagnostic regions for upright and rotated shape versions (human observers). (A) Black circles: Average overlap of rotated and upright diagnostic regions for different stimulus orientations, computed across observers and shapes. White circles: Average overlap necessary to reach  $p < .05$ , as determined by the Monte-Carlo simulations. Error bars denote the SEM. (B and C) Data of individual subjects. The plots indicate for each subject how many shapes had a significant overlap at a particular orientation. (B) Subjects tested with shapes presented at 0 deg, 90 deg, 180 deg, and  $-90$  deg. (C) Subjects tested with stimuli rotated by 0 deg,  $\pm 30$  deg, and  $\pm 90$  deg.

The observed overlap between diagnostic regions suggests that subjects had a tendency to use the same features to identify upright and rotated versions of a shape. To further test this conclusion, we computed the overlap expected under the assumption that subjects randomly select shape features to identify rotated shape versions. Because of their computation, all diagnostic regions covered 5% of the full image. Thus, some overlap between any two diagnostic regions is expected even for random placement of the regions, with the amount of overlap expected by chance depending on the shape of the involved diagnostic regions. To compute whether the overlap between any rotated diagnostic region and its corresponding upright diagnostic region exceeded the level expected by chance, we used the following Monte-Carlo simulation to estimate the chance level. The upright diagnostic region was kept fixed, and the diagnostic region for the rotated shape of interest was rotated to the upright as before. In every repetition of the simulation, the latter region was then randomly repositioned, and the overlap between the repositioned region and the upright region was computed. By these means, we could determine the level of significance of the observed overlap for each pair of rotated and upright diagnostic region. We could also compute the overlap that needed to be exceeded to reach a level of  $p = .05$  for each pair of diagnostic regions. This critical overlap was averaged across shapes and subjects for each rotation, and is plotted as a reference in [Figure 4A](#) (white circles). [Figures 4B](#) and [4C](#) plot for each subject and orientation the number of shapes at which the overlap exceeded the level of  $p < .05$ .

For most of the subjects, the observed overlap was significantly larger than chance in almost all cases. Notably, there was no consistent influence of stimulus orientation on the number of shapes for which a significant overlap was obtained. The only exception was observer EZ, for which the overlap of none of the shapes rotated by 90 deg reached significance. However, the  $p$  values for two of the shapes at this orientation were .06 and .07; for the third shape, the  $p$  value was .2. Thus, the overlaps for two of the shapes are almost significant, making the results of this observer more similar to the others. Nonetheless, it is also possible that observer EZ – in contrast to all other observers – used unique features to identify the shapes rotated by 90 deg.

In conclusion, our data suggest that human observers showing view-invariant recognition in most cases use similar features to identify upright and rotated versions of a shape.

## Experiments with monkey observers

The monkeys were initially trained to discriminate between the three shapes when shown upright. During each trial, one of the shapes was shown for a short time, and then replaced by three squares (the targets). Each

target was associated with one of the shapes, and a saccade to the correct target was rewarded by a drop of juice. We then tested their capability to generalize the task to rotated versions of the shapes. For this test, the shapes were rotated in steps of 30 deg in the picture plane. The differently oriented versions of all shapes were presented in random order to the monkeys. Rotated versions of a shape were associated with the same target as the upright shape, and a saccade to this target was rewarded with a drop of juice.

[Figure 5A](#) plots the performance of the two monkeys in the generalization test. These plots show that while the performance of the monkeys for the upright stimuli was around 90% correct, their performance deteriorated rapidly as the shapes were rotated away from this orientation. Performance levels at any stimulus orientation were compared against the chance level of 33% correct using a  $\chi^2$ -test, applying a Bonferroni correction to adjust for the 12 comparisons. Monkey B98 seemed to transfer to stimuli rotated counterclockwise by 30 deg; however, the performance failed to reach the level of  $p < .05$  when the Bonferroni correction was applied ( $\chi^2 = 5.7$ ,  $p = .02$  uncorrected). For any of the larger rotations, the monkey's performance was not significantly better than chance ( $p > .05$ ). Results for monkey G00 were similar. The monkey could generalize from the upright shapes to shapes rotated by 30 deg in any direction (30 deg:  $\chi^2 = 30.0$ ,  $p < .001$ ; -30 deg:  $\chi^2 = 10.0$ ,  $p = .02$ ), and to shapes rotated counterclockwise by 60 deg ( $\chi^2 = 15.0$ ,  $p = .001$ ). Performance was not better than chance for the rest of the rotations ( $p > .05$ ).

Since both monkeys could not generalize the task to stimuli rotated by more than 30 deg away from upright, they had to be trained on the rotated stimuli before the rest of the testing could be carried out. In these training sessions, the monkeys performed the shape discrimination task on stimulus sets consisting of the upright version of each shape, and shapes with increasingly larger rotations. Thus, in the first training sessions, monkeys were trained with the upright stimuli and stimuli rotated by  $\pm 30$  deg. The next stimulus set consisted of the upright stimuli and stimuli rotated by  $\pm 60$  deg, and so forth. Each stimulus set was kept unchanged until the monkeys could identify all stimuli in the set correctly at least in 75% of the trials on two consecutive days. In each trial, the monkeys received feedback about the correctness of their response. No additional cue was given to the monkeys about the association between stimuli and saccade targets. Thus, the appropriate responses had to be learned by trial-and-error. The number of trials necessary to reach the criterion level of 75% correct for a specific stimulus orientation can thus be taken as a measure for the difficulty to acquire the task for this orientation. Monkey B89 required 1,860 trials on stimuli rotated by 30 deg to reach the criterion. Additional 1,666 trials were required until the criterion was met for the 60 deg rotations, and further 1,278 for the 90 deg rotations. Thus, this monkey required about the

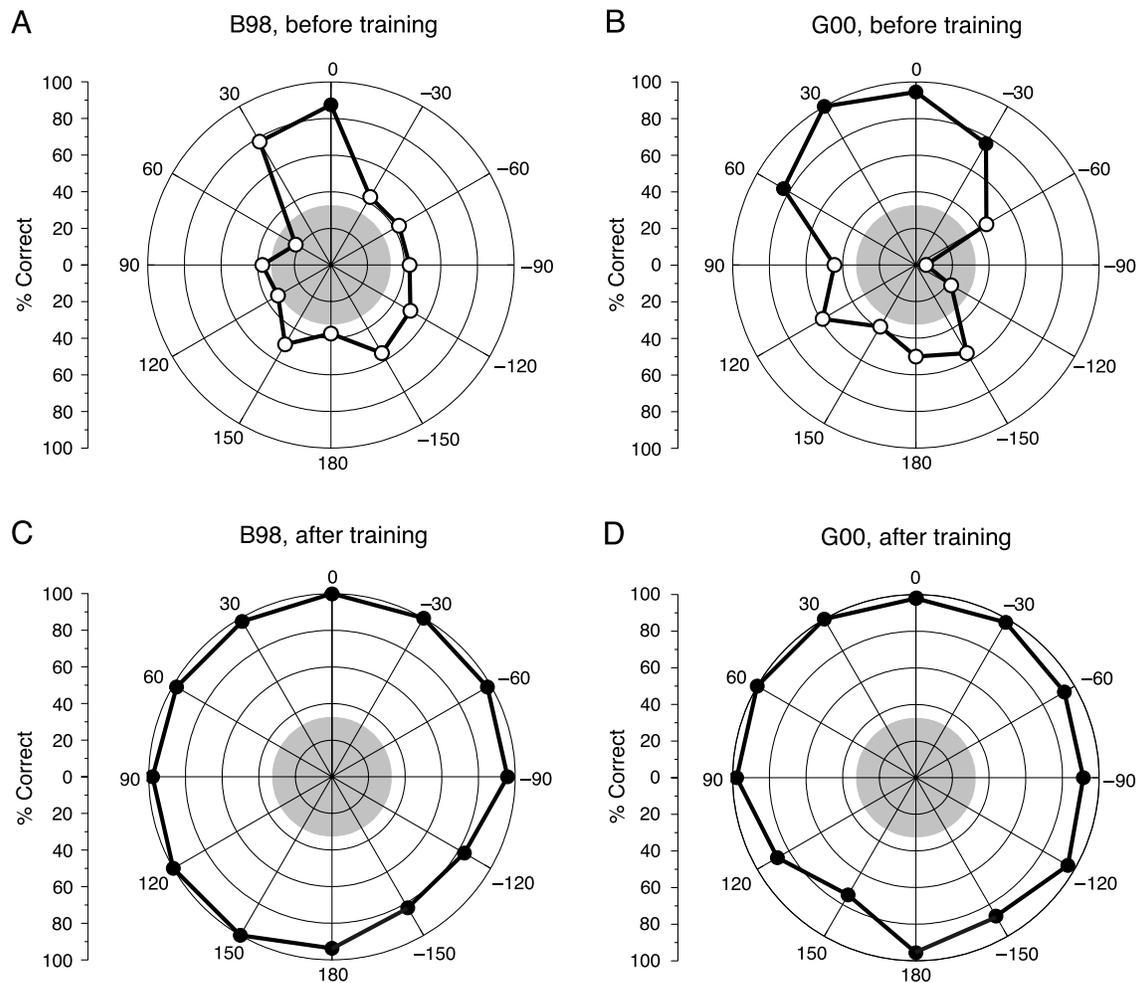


Figure 5. Performance of monkeys for rotated shapes. In the polar plots, each symbol represents the performance of a monkey with shapes presented at a specific orientation (computed over 10 to 20 repetitions per orientation for A and C, and about 50 repetitions for B and D). Closed symbols indicate performance levels significantly different from chance ( $\chi^2$ -test,  $p < .05$  after Bonferroni correction for multiple tests). Open symbols indicate performances not different from chance. The gray circle corresponds to the chance level.

same amount of training for each of the possible orientations up to 90 deg. As shown, monkey G00 could immediately generalize to stimuli rotated by 30 deg. It took this monkey 5,969 trials to meet the criterion level for stimuli rotated by 60 deg, followed by further 2,862 trials for the 90 deg rotations. Thus, it seems for this monkey that there was some transfer of training on rotated shapes to novel rotations.

The monkeys were trained on rotated shapes until they were correctly discriminating stimuli rotated up to 90 deg away from the upright. At this point, we again tested their generalization performance for all rotations up to 180 deg from the upright. Note that the monkeys had not been exposed to stimuli rotated by angles larger than 90 deg since the initial generalization test. The monkeys' performance in this second generalization test is plotted in Figure 5B. The additional training on stimuli rotated up to

90 deg was a sufficient experience to allow both monkeys to generalize to any stimulus orientation in the picture plane (performance at all stimulus orientations better than chance,  $\chi^2$ -test, Bonferroni corrected  $p < .001$  for all tests).

Since with sufficient training the monkeys seemed to be able to perform the discrimination task independent of stimulus orientation, we proceeded to use Bubbles to determine whether this generalization performance was based on using a particular feature to identify a shape, independent of the shape's orientation. In the Bubbles sessions, the stimulus set consisted of the upright versions of the three shapes, and shapes rotated by  $\pm 30$  deg and  $\pm 90$  deg. Thus, the stimulus set consisted of rotated shapes to which at least one of the monkeys spontaneously generalized, and of rotated versions that required additional training. The monkeys identified all stimulus

orientations equally well during the Bubbles sessions. Across all stimuli and rotations, they maintained a performance level of  $98.5\% \pm 0.4\%$  correct on the unoccluded shapes, which were presented throughout the Bubbles sessions (see [Methods](#)). There was a small effect of stimulus orientation on the performance level, repeated measures ANOVA,  $F(4) = 3.1$ ,  $p = .04$ , which was due to a difference in performance on shapes rotated by 30 deg and 90 deg counterclockwise (paired  $t$  test,  $t(5) = 2.6$ ,  $p = .05$ ). None of the other comparisons yielded significant results ( $p > .05$  for all comparisons). Most critically, this means that the performance levels for the upright stimuli were not different from the performance levels for any of the rotated stimuli.

The same analysis as previously applied to the human data was carried out on the Bubbles data collected for the monkeys. For monkey B98, the first stimulus (the bottle) had to be excluded from further analysis, because we could not determine a diagnostic region for the upright stimulus. As an example for the behavior of the monkeys, the diagnostic regions computed for one of the monkeys

are shown in [Figure 6A](#) and [6B](#). For comparison, the results of one of the human observers tested with the same stimulus orientations are plotted in [Figure 6C](#). These figures show a marked difference between the behavior of monkey and human observer: The diagnostic regions of the human observer were largely centered on the same shape features, irrespective of the shape's orientation. In contrast, the monkey observer relied on different features to identify differently oriented versions of the same shape.

The average results obtained from both monkeys confirmed this observation. The average overlap is plotted in [Figure 6D](#). Black regions in this figure indicate the average overlap,  $\pm 1$  *SE*, necessary to reach a level of  $p < .05$ . [Figure 6E](#) shows the individual data of both monkeys by plotting the number of shapes per orientation for which the overlap was significantly larger than expected by chance ( $p < .05$ ). In contrast to the human data, the overlap – the degree to which the same features were used to identify upright and rotated shape versions – depended on the stimulus orientation for the monkeys. A repeated measures ANOVA performed on the overlap

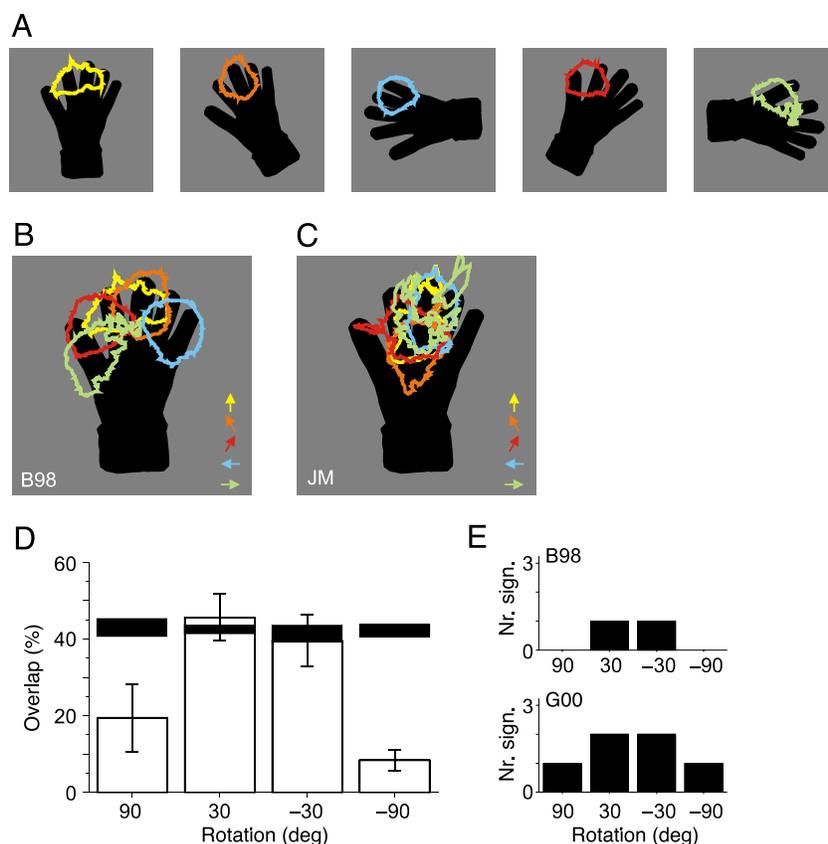


Figure 6. Consistency in the diagnostic regions for upright and rotated shape versions (monkey observers). (A) Exemplar data for monkey B98, showing the diagnostic regions for each rotated version of the hand. (B) The same data as in (A), but after rotating all diagnostic regions to the upright. The color of the arrows plotted on the right indicate the orientation for each diagnostic region. (C) Exemplar data for a human observer (JM) tested with the same stimulus orientations. Data are plotted as in (B). (D) Average overlap observed across both monkeys and all shapes. Error bars denote the *SEM*. For each orientation, the black region indicates which overlap on average needed to be exceeded to reach  $p < .05$ . The upper edge of the black region is placed at the average critical overlap  $+1$  *SEM*, the lower edge at the average  $-1$  *SEM*. (E) Number of significant overlaps per rotation.

yielded a significant influence of orientation,  $F(3) = 6.6$ ,  $p = .007$ . The most pronounced difference was observed when comparing the overlap for 30 deg and 90 deg rotations. While the overlap was on average 43% for both 30 deg rotations, it only reached 14% for the 90 deg rotations, a difference that was highly significant, paired  $t$  test,  $t(9) = 4.3$ ,  $p = .002$ . A similar trend was seen for both monkeys individually. Overall, the overlap for only a few pairs of upright and rotated diagnostic regions reached the level of  $p < .05$ . However, most of these cases were observed for the 30 deg rotations. Thus, the monkeys showed a strong tendency to use different shape features the further the shapes were rotated away from the original. This pattern of behavior is different from the one observed for human observers. A direct comparison between the levels of overlap reached at each stimulus rotation furthermore revealed that in general the monkeys were using less similar shape features than humans. With the exception of shapes rotated by  $-30$  deg, the average overlap was significantly lower for the monkeys than for the humans,  $t$  test, 30 deg:  $t(12) = 2.4$ ,  $p = .03$ ; 90 deg:  $t(18) = 2.6$ ,  $p = .02$ ;  $-30$  deg:  $t(12) = 1.9$ ,  $p = .08$ ;  $-90$  deg:  $t(18) = 6.2$ ,  $p < .001$ . Interestingly, there was a significant correlation between the performance in the initial generalization test (see Figures 5A and 5B) and the degree to which the monkeys used orientation-independent shape features after training with rotated shapes. The better the initial performance for a specific orientation, the

more similar were the features used to identify the upright and the rotated shape (Pearson correlation coefficient between overlap and performance in the initial generalization test:  $r = .60$ ,  $p = .005$ ). This suggests that for rotation angles for which the monkeys could not immediately identify the shapes, they had a larger tendency to use novel shape features to solve the task.

When analyzing the results of the monkeys, we noticed that it seemed as if the monkeys were using a fixed spatial region to identify each shape, without adapting the position of this region to the rotation of the shape (see Figure 6A). To test for this possibility, we again computed the overlap between the diagnostic regions of upright and rotated versions of a shape. However, since we were interested in whether diagnostic regions remained at a fixed spatial location despite stimulus rotation, we directly computed the overlap between the diagnostic region of a rotated shape and the diagnostic region of the corresponding upright shape, without bringing the rotated diagnostic region to the upright. As before, a Monte-Carlo simulation was carried out to compute the amount of overlap expected by chance. Again, the simulation consisted of randomly repositioning the diagnostic regions for the rotated shape, and computing the overlap between the upright and the repositioned diagnostic region.

The results of this analysis are plotted in Figure 7. They show a similar pattern as Figure 6. The amount of overlap strongly depended on the orientation of a stimulus,

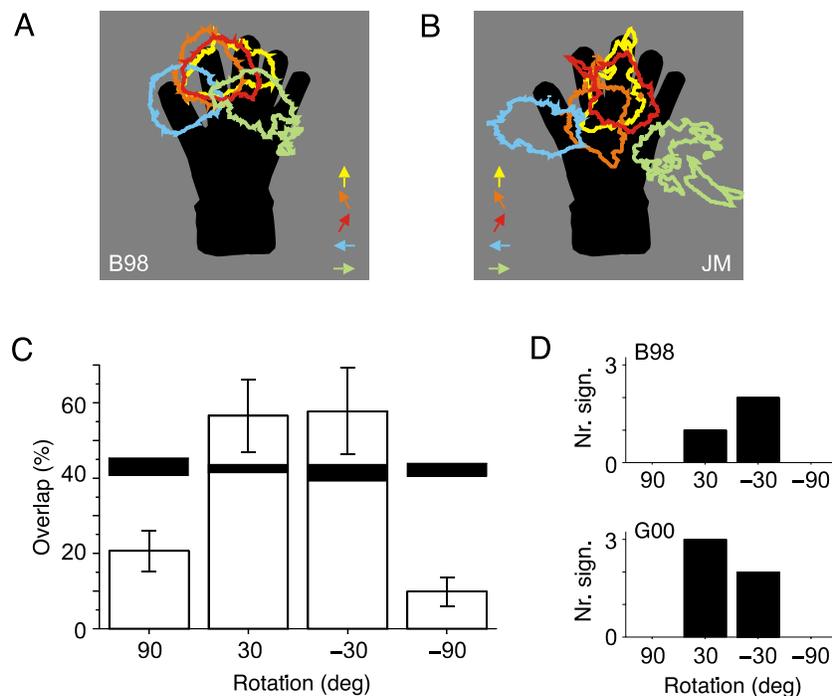


Figure 7. Overlap between diagnostic regions for rotated and upright shapes, without taking the orientation of shapes into account. (A and B) The same diagnostic regions are plotted as in Figures 6B and 6C but without rotation of the diagnostic regions to the upright. (C) Average overlap between the diagnostic regions of rotated and upright shapes, without taking the orientation of a rotated shape into account. (D) Number of significant overlaps per orientation. Layout of (C) and (D) as in Figure 6.

repeated measures ANOVA,  $F(3) = 15.3$ ,  $p < .001$ , with larger overlaps for the 30 deg than the 90 deg rotations, paired  $t$  tests,  $t(9) = 5.9$ ,  $p < .001$ . With the exception of the  $-30$  deg rotation, the overlaps were similar to the ones computed before, paired  $t$  tests, 30 deg:  $t(4) = -1.4$ ,  $p = .3$ ; 90 deg:  $t(4) = -0.1$ ,  $p = .9$ ;  $-30$  deg:  $t(4) = -3.1$ ,  $p = .04$ ;  $-90$  deg:  $t(4) = -.4$ ,  $p = .7$ . As a control, we performed the same analysis on the data for human observers. For the human data, the amount of overlap was significantly reduced if diagnostic regions of rotated shapes were not first rotated to match the upright orientation. This was the case for all orientations, paired  $t$  tests, 30 deg:  $t(8) = 2.4$ ,  $p = .04$ ; 90 deg:  $t(14) = 5.9$ ,  $p < .001$ ;  $-30$  deg:  $t(8) = 3.5$ ,  $p = .009$ ;  $-90$  deg:  $t(14) = 9.4$ ,  $p < .001$ ; 180 deg:  $t(5) = 4.4$ ,  $p = .007$ .

Thus, it seems that while the human data can better be explained by assuming that fixed shape features, not a fixed spatial region, are used to identify rotated shape versions, the same is not the case for the monkey data. For the 90 deg rotations, only a few shapes had overlaps which were significantly larger than chance, independent of how the overlap was computed. These results suggest that the monkeys were using different sets of features to identify the upright versions of a shape, and the versions rotated by  $\pm 90$  deg. For the 30 deg rotations, larger overlaps were obtained. However, the same amount of overlap was reached with and without taking the orientation of a shape into account before computing the overlap. Thus, the data can equally well be explained by assuming that the monkeys were using the same features to identify different versions of a shape, or by assuming that they used a fixed region in space irrespective of shape orientation.

## Discussion

We tested whether view-invariant object recognition is based on particular shape features. View invariance was examined for rotations in the picture plane. Our results show that irrespective of the in-plane orientation at which a shape was presented, human observers relied on the same shape features for identification. In contrast, monkeys identified each rotated version of a shape using a unique set of features. This was the case even though the testing was similar for both species.

Our results for human observers are consistent with previous findings. Both training and feature-based attention can lead to view-invariant performance, which has been explained by assuming that both processes enhance an observer's knowledge about which shape features are informative (Jolicoeur, 1985; Liter, 1998). The consistency with which human observers rely on the same shape features irrespective of a shape's orientation provides a basis for these findings. It remains to be shown that these results can not only be obtained for the small, limited

stimulus set employed here, but also for more complex categorization tasks closer to object recognition problems encountered in the real world. Certainly, our results will not hold if object rotations in depth make previously diagnostic features disappear, and new diagnostic features appear. However, our data show that as long as object modifications allow it, human observers have a bias to use the same diagnostic features despite object modifications. It could be argued that a small stimulus set like ours artificially introduces invariance in behavior. The results of the monkey observers with their lower degree of invariance demonstrate that the stimulus set, although very limited, could nonetheless evoke very different behaviors. We have chosen this stimulus set for two reasons: First, one of the main interests of this study was the comparison between human and monkey behavior. The chosen task – admittedly very simple for humans – was already rather complicated for monkeys. We therefore did not make the task more complicated. Second, we chose a limited stimulus set for which human observers immediately showed view-invariance to limit instabilities during testing. The Bubbles paradigm requires an extensive amount of trials for each shape in the stimulus set. Thus, throughout testing the subjects receive sufficient training on each shape to eventually generate view-invariant performance, even if they initially show a dependency on the view. These changes in behavior during testing were avoided by using a stimulus set that allowed view-invariant behavior from the beginning.

In this study, we tested the strategies of Rhesus monkeys that have to distinguish between shapes presented at different rotation angles. Other studies have previously assessed the performance of other species on the same task. For example, Hollard and Delius (1982) showed that pigeons match rotated shapes to their upright counterparts in a delayed match-to-sample task. Recently, it has been shown using Bubbles that pigeons use non-accidental shape properties when discriminating between 2D projections of 3D objects (Gibson, Lazareva, Gosselin, Schyns, & Wasserman, 2007). These shape properties support recognition of a shape from different views, as they remain distinctive despite rotations of the object (Biederman, 1987). Gibson et al. (2007), however, did not test whether pigeons use the same shape features to identify a shape across all orientations. The results obtained in our study make this an interesting question to be pursued in the future.

A number of plausible reasons exist for differences in behavior between monkeys and humans, which can roughly be split into two classes. On the one hand, different behaviors might simply reflect the fact that humans understand the task different than monkeys, and therefore use different strategies to solve it. On the other hand, it is possible that objects are encoded differently in the human and monkey brain, leading to different capabilities to solve a task. Obviously, both explanations are not mutually exclusive and might well be linked.

Before discussing possible differences in the neural encoding of objects further, we will discuss reasons biasing monkeys to solve the task using different behavioral strategies than humans. First, it seems possible that familiarity with the stimulus set in the initial testing sessions could influence an observer's strategy. For human observers, all three shapes represented familiar, everyday objects, which they could clearly identify from their silhouettes. Assuming that monkeys can as well correctly interpret 2D pictures, at least two of the three shapes (the hand and the bottle) represent objects commonly encountered by our monkeys. Since the monkeys exhibited the same behavior for all three shapes, it seems unlikely that shape familiarity prior to testing is a reason underlying species-specific task strategies. In addition, Jolicoeur has demonstrated that stimulus orientation influences naming times similarly for common and uncommon objects (Jolicoeur, 1985).

Task difficulty could be a second reason leading to different strategies in humans and monkeys. Our stimulus set was simple enough that humans immediately generalized and showed view-invariant performance, but the task was obviously more difficult for the monkeys as they required a lot of training to reach view-invariant performance. It therefore remains possible that using a stimulus set for which the monkeys immediately show view-invariance could lead to a more similar behavior in both species. Yet, it needs to be kept in mind that our stimulus set, which consisted of only three, very distinctive shapes, was already very limited.

A third reason for strategic differences in humans and monkeys may lie in the general understanding of the task. We have no possibility to address this issue further, and it remains a possible explanation for our results. However, we can at least rule out that monkeys solved the task differently because they never perceived the equivalence of upright and rotated shapes, instead treating each rotated shape as a new, individual stimulus. After training the monkeys with stimuli rotated by maximally 90 deg, we tested their performance on shapes rotated by much larger angles (up to 180 deg). The monkeys had only once, and then only very briefly, been exposed to shapes rotated by such large angles. Nonetheless, they were able to identify shapes at all rotation angles correctly. This strongly suggests that after sufficient training the monkeys were generalizing from the learned to unlearned stimulus orientations.

Fourth, differences in strategies might be generated by the chosen testing paradigm. This point is valid in general, but is additionally motivated by concerns that have been raised regarding the Bubbles paradigm (Murray & Gold, 2004a). Bubbles involves the repeated presentation of stimuli occluded in such a way that only isolated stimulus fragments remain visible. It has been argued that this way of stimulus presentation biases subjects in favor of adopting a strategy based on local object features, and against one involving more global or holistic object

properties. This and other concerns regarding Bubbles have been more thoroughly discussed in a recent series of papers (Gosselin & Schyns, 2004; Murray & Gold, 2004a, 2004b). Obviously, a change in strategy is in itself a concern. Changes in strategy introduced by the testing paradigm have an additional relevance here if the chosen paradigm influences monkeys differently than humans. The influences of a method such as Bubbles on behavior can only be assessed by comparison with another method that (a) determines diagnostic object features in the same quantitative manner, and (b) has been shown not to influence behavior. As there is a lack of quantitative methods for this purpose, it remains difficult – at least in our opinion – to establish under which circumstances the “true” behavior of an observer can be determined. Currently, reverse correlation (Ahumada & Lovell, 1971) would be most suited for a comparison. Reverse correlation determines diagnostic object features by the influences of additive noise on behavior. Future studies should address the question of whether Bubbles and reverse correlation give comparable results, both for monkeys and humans. In our view, Bubbles has the advantage that it simulates a condition that often occurs naturally, namely the partial occlusion of objects, with the added benefit that partial occlusion influences the perception of humans and monkeys alike (Fujita, 2001; Kovács, Vogels, & Orban, 1995; Osada & Schiller, 1994; Sugita, 1999). Therefore, we think it is likely that if Bubbles influences observers' strategies, it at least does so similarly for humans and monkeys.

Finally, one more possible reason for behavioral differences between humans and monkeys should be mentioned. Recent experiments have suggested that monkeys put more emphasis on local than on global stimulus properties (Anderson, Peissig, Singer, & Sheinberg, 2006; Einhäuser et al., 2006). While a global interpretation of the shape should be independent of stimulus orientation, the most distinctive local features may well change with different stimulus orientations.

Differences in behavior might be generated because of purely behavioral biases that differ between species. In contrast, they might also be a reflection of different brain capacities for different tasks. Here, differences in the encoding of objects in the human versus the monkey brain could be the reason for species-specific solutions for the rotation task.

For rotations in depth and in the picture plane, two classes of models are commonly used to explain the mechanisms supporting object recognition across changes in viewpoint: view-invariant models and multiple-views models. These models can be distinguished by whether they assume that an object is encoded in memory using a single, view-invariant instance, or whether multiple, view-specific templates are stored for each object (e.g., see Lawson, 1999; Tarr & Bülthoff, 1998; Tarr & Pinker, 1989). The behavior of the human subjects on our small stimulus set is consistent with the view-invariant models,

since subjects always used the same feature to identify an object. The only possible exception is observer EZ, whose performance suggested a mix of behaviors explained both by view-invariant and multiple-views models. The monkeys' strategies can be better described by the multiple-views model, as it seems that the monkeys use different templates to identify the same shape at different rotation angles.

View-dependent or -independent behavior might not necessarily be generated because of differences in the memory encoding of objects. Instead the type of information an observer extracts from an object – and whether or not it supports view-invariant behavior – might determine performance on rotated objects (Schyns, 1998). In this sense, our data demonstrate that very different types of information can be used to support view-invariant performance. Despite the fact that humans and monkeys used different shape features, both species performed view-invariant throughout the testing. Intuitively, it seems that the monkeys' strategies should not easily support view-invariant performance, as they identified each rotated version of a shape using a unique set of features. However, view-invariance can be achieved by basing object identification on view-dependent image features. For example, a recent model by Ullman generates view-invariance from view-dependent features by linking these features to form so-called “abstract” features (Ullman, 2007). The abstract features then allow viewpoint-independent object recognition. Possible mechanisms for the formation of abstract features include the direct observation of an object undergoing transformation, as well as establishing the interchangeability of object features in a common context. It is an interesting possibility that the monkeys performed view-invariant after training because they had learned the equivalence of the unique shape features used to identify each shape version. The long training necessary for them to reach view-invariant performance might then reflect the time necessary to link these different features.

Finally, the initial generalization capabilities of monkeys versus humans may hint at another interesting possibility regarding the neural encoding of objects. Monkeys initially could generalize to shapes rotated by about 30 to 60 deg, which is in agreement with other studies testing the generalization capabilities of monkeys for rotations in depth (Logothetis et al., 1994; Wang et al., 2005). Interestingly, the tuning width of neurons in the inferotemporal cortex of monkeys, a brain region strongly implicated in object recognition processes (Logothetis & Sheinberg, 1996; Tanaka, 1996), was reported to be on the order of 60 deg for rotations of objects in the picture plane (Logothetis et al., 1995). Behavioral and neural tuning widths therefore show some agreement for the monkeys. Mixed results exist for the influence of object rotation on neural responses in the human brain, suggesting both view-dependent and view-independent object representations in the higher cortical visual areas (Gauthier et al.,

2002; Grill-Spector et al., 1999; James, Humphrey, Gati, Menon, & Goodale, 2002; Vuilleumier, Henson, Driver, & Dolan, 2002). The widths of neuronal tuning curves for object rotations have so far not been studied systematically in the human brain. If the monkeys initial generalization performance indeed is a function of the tuning widths of inferotemporal neurons, then our data for the human observers suggest that tuning curves of neurons representing shapes in the human brain should be broader than the ones found in the monkey.

## Acknowledgments

This work was supported by the Max Planck Society. G.R. is a DFG Heisenberg investigator (RA 1025/1-1). We thank B. Dillenburger, C. Wehrhahn, and M. Wilke for comments on the manuscript.

Commercial relationships: none.

Corresponding author: Kristina J. Nielsen.

Email: nielsen@salk.edu.

Address: The Salk Institute for Biological Studies, 10010 North Torrey Pines Rd, La Jolla, CA 92037, USA.

## References

- Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.
- Anderson, B., Peissig, J. J., Singer, J., & Sheinberg, D. L. (2006). XOR style tasks for testing visual object processing in monkeys. *Vision Research*, *46*, 1804–1815. [[PubMed](#)]
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147. [[PubMed](#)]
- Einhäuser, W., Kruse, W., Hoffmann, K. P., & König, P. (2006). Differences of monkey and human overt attention under natural conditions. *Vision Research*, *46*, 1194–1209. [[PubMed](#)]
- Fujita, K. (2001). Perceptual completion in rhesus monkeys (*Macaca mulatta*) and pigeons (*Columbia livia*). *Perception & Psychophysics*, *63*, 115–125. [[PubMed](#)]
- Gauthier, I., Hayward, W. G., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2002). BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron*, *34*, 161–171. [[PubMed](#)] [[Article](#)]
- Gibson, B. M., Lazareva, O. F., Gosselin, F., Schyns, P. G., & Wasserman, E. A. (2007). Nonaccidental properties underlie shape recognition in mammalian

- and nonmammalian vision. *Current Biology*, *17*, 336–340. [PubMed]
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271. [PubMed]
- Gosselin, F., & Schyns, P. G. (2004). No troubles with bubbles: A reply to Murray and Gold. *Vision Research*, *44*, 471–477. [PubMed]
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203. [PubMed] [Article]
- Hollard, V. D., & Delius, J. D. (1982). Rotational invariance to visual pattern recognition by pigeons and humans. *Science*, *218*, 804–806. [PubMed]
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2002). Differential effects of viewpoint on object-driven activation in dorsal and ventral streams. *Neuron*, *35*, 793–801. [PubMed] [Article]
- Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition*, *13*, 289–303. [PubMed]
- Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 200–210. [PubMed]
- Judge, S. J., Richmond, B. J., & Chu, F. C. (1980). Implantation of magnetic search coils for measurement of eye position: An improved method. *Vision Research*, *20*, 535–538. [PubMed]
- Kovács, G., Vogels, R., & Orban, G. A. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *Journal of Neuroscience*, *15*, 1984–1997. [PubMed] [Article]
- Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica*, *102*, 221–245. [PubMed]
- Lawson, R., & Jolicoeur, P. (2003). Recognition thresholds for plane-rotated pictures of familiar objects. *Acta Psychologica*, *112*, 17–41. [PubMed]
- Lee, H., Simpson, G. V., Logothetis, N. K., & Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron*, *45*, 147–156. [PubMed] [Article]
- Liter, J. C. (1998). The contribution of qualitative and quantitative shape features to object recognition across changes of view. *Memory & Cognition*, *26*, 1056–1067. [PubMed]
- Logothetis, N. K., Pauls, J., Bülthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology*, *4*, 401–414. [PubMed]
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563. [PubMed]
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, *19*, 577–621. [PubMed]
- Murray, R. F., & Gold, J. M. (2004a). Troubles with bubbles. *Vision Research*, *44*, 461–470. [PubMed]
- Murray, R. F., & Gold, J. M. (2004b). Reply to Gosselin and Schyns. *Vision Research*, *44*, 479–482.
- Newell, F. N. (1998). Stimulus context and view dependence in object recognition. *Perception*, *27*, 47–68. [PubMed]
- Nielsen, K. J., Logothetis, N. K., & Rainer, G. (2006). Discrimination strategies of humans and rhesus monkeys for complex visual displays. *Current Biology*, *16*, 814–820. [PubMed] [Article]
- Osada, Y., & Schiller, P. H. (1994). Can monkeys see objects under conditions of transparency and occlusion. *Investigative Ophthalmology & Visual Science*, *35*, 1664.
- Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering*, *10*, 137–145. [PubMed]
- Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. 147–179. [PubMed]
- Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, *401*, 269–272. [PubMed]
- Takano, Y. (1989). Perception of rotated forms: A theory of information types. *Cognitive Psychology*, *21*, 1–59. [PubMed]
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139. [PubMed]
- Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, *67*, 1–20. [PubMed]
- Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science*, *8*, 282–289.
- Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, *21*, 233–282. [PubMed]

- Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Sciences, 11*, 58–64. [[PubMed](#)]
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience, 5*, 491–499. [[PubMed](#)]
- Wang, G., Obama, S., Yamashita, W., Sugihara, T., & Tanaka, K. (2005). Prior experience of rotation is not required for recognizing objects seen from different angles. *Nature Neuroscience, 8*, 1568–1575. [[PubMed](#)]
- Wilson, K. D., & Farah, M. J. (2003). When does the visual system use viewpoint-invariant representations during recognition? *Cognitive Brain Research, 16*, 399–415. [[PubMed](#)]