Research Article

Do Humans and Baboons Use the Same Information When Categorizing Human and Baboon Faces?

Julie Martin-Malivel,[1] Michael C. Mangini,[1] Joël Fagot,[2] and Irving Biederman[1]

[1]Department of Psychology, Neuroscience Program, University of Southern California, and
[2]CNRS, Mediterranean Institute for Cognitive Neuroscience, Marseille, France

ABSTRACT—*What information is used for sorting pictures of complex stimuli into categories? We applied a reverse correlation method to reveal the visual features mediating categorization in humans and baboons. Two baboons and 6 humans were trained to sort, by species, pictures of human and baboon faces on which random visual noise was superimposed. On ambiguous probe trials, a human-baboon morph was presented, eliciting "human" responses on some trials and "baboon" responses on others. The difference between the noise patterns that induced the two responses made explicit the information mediating the classification. Unlike the humans, the baboons based their categorization on information that closely matched that used by a theoretical observer responding solely on the basis of the pixel similarities between the probe and training images. We show that the classification-image technique and principal components analysis provide a method to make explicit the differences in the information mediating categorization in humans and animals.*

Pigeons and monkeys can be trained to classify pictures of natural or artificial objects into distinct categories (Delorme, Richard, & Fabre-Thorpe, 2000; Herrnstein, 1990; Herrnstein & Loveland, 1964; Roberts & Mazmanian, 1988; Schrier & Brady, 1987; Vogels, 1999a), such as "trees" or "humans," and they can generalize their behavior to previously unseen exemplars of those categories. However, it is not clear what information the animals use to achieve their classifications. Some researchers have argued that performance reflects the animals' "natural concepts" (Herrnstein & Loveland, 1964), implying that the animals recognize the objects represented in the pictures and use the information at a conceptual level. Others have proposed that accurate categorization of pictures can be mediated by nonconceptual processes (Fagot, Martin-Malivel, & Dépy, 1999), in which animals focus on a subset of perceptual features associated with each category, such as the presence of a distinctive color patch (D'Amato & van Sant, 1988). In this case, animals could categorize the stimuli using image characteristics without necessarily recognizing the objects represented in the pictures (Fagot, 2000). Identifying the types of information[1] that control categorization is crucial for evaluating conceptual abilities of animals and for interpreting data obtained in studies of object recognition and the neural basis of visual recognition.

To infer which aspects of a visual stimulus an animal utilizes in a categorization task, researchers have traditionally manipulated the physical properties of pictures, for instance, by scrambling images or deleting object parts (reduction procedure; Cerella, 1979; Martin-Malivel & Fagot, 2001b; Vogels, 1999a). The inherent limitation of this procedure is that it is dependent on the experimenter's a priori judgment about which aspects of the picture are relevant to the animal's classification. Because pictures are complex and polymorphous stimuli,

Address correspondence to Julie Martin-Malivel, Psychology Department, Emory University, 532 N. Kilgo Circle N.E., Atlanta, GA 30322, e-mail: jmalive@emory.edu, or to Michael C. Mangini, Brain and Cognitive Sciences Department, MIT, NE20-443, 77 Massachusetts Ave., Cambridge, MA 02139, e-mail: mangini@mit.edu.

---

[1]We use the term *information* as shorthand for any pictorial features that may subserve categorization, not as a reference to Shannon's entropic measure.

untested perceptual features may contribute to the discrimination (Fagot, 2000).

The present study was aimed at identifying the information used by baboons and humans to categorize human and baboon faces. We used an adaptation of the classification-image (CI) procedure (Ahumada, 1996, 2002; Ahumada & Lovell, 1971). The CI method (also termed reverse correlation) has been employed with complex images such as faces only recently (Gold, Murray, Bennett, & Sekuler, 2000; Gosselin & Schyns, 2003; Mangini & Biederman, 2001, 2004), and only in human subjects. Another method, Bubbles (Gosselin & Schyns, 2001; Schyns, Bonnar, & Gosselin, 2002), estimates which regions of a picture a subject finds informative, but unlike the CI technique, does not estimate the specific information used by the observer. This method has recently been applied to pigeons (Gibson, Wasserman, Gosselin, & Schyns, 2005). The differences between the CI technique and Bubbles have been discussed previously (Gosselin & Schyns, 2004; Mangini & Biederman, 2004; Murray & Gold, 2004). Because a CI analysis provides a rich set of data from simple classification responses of any observer (animal, human, or machine), it is well suited to the study of comparative cognition.

In the current study, adult humans and baboons were tested using identical human and baboon pictures. The data were collected from test trials using a perceptually ambiguous probe stimulus created to resemble human and baboon faces equally. When presented in high levels of visual noise, repetitions of the identical underlying ambiguous stimulus elicited "human" responses on some trials and "baboon" responses on others. Because the only aspect of the stimulus that differed across presentations was the randomly generated visual noise that perturbed the underlying ambiguous image, one can assume that the response on a given trial depended on which of the observer's categories some feature of the noise pattern matched more closely. By contrasting the average of the visual noise patterns that elicited "human" responses with the average of the patterns that elicited "baboon" responses, we were able to discover the visual information that the subjects used to determine their responses.

This method compares linear estimates of the information used by subjects. We do not assume that all of human and baboon vision—especially higher-level vision—is itself linear (see Mangini & Biederman, 2004, for a discussion of the implications of modeling a classification as a linear process). There is evidence from single-unit primate studies in V4 and the inferotemporal cortex (Kobatake & Tanaka, 1994; Pasupathy & Connor, 1999) that neural responses are not linear. Nonetheless, one can ask whether humans and baboons differ in this linear component. We should also point out that the vast majority of the statistics employed in behavioral science and neuroscience involve linear comparisons, as when reaction time distributions, spike rates, or blood-oxygen-level-dependent responses are compared, and whenever a correlation coefficient is calculated.

METHOD

Participants and Apparatus

The participants were 2 adult Guinea baboons (*Papio papio*; B06 and B07), 1 male and 1 female from the animal facility of the CNRS (Marseille, France), and 6 humans (H1–H6), 2 male and 4 female undergraduates from the University of Southern California. They were trained to discriminate pictures of human and baboon faces in a computer-controlled go/no-go task. The experiments were approved by the Comité Régional d'Éthique pour l'Expérimentation Animale de la Région Provence and by the Institutional Review Board of the University of Southern California (USC). The experimental stimuli were created with Matlab, using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The humans were tested at USC. An in-house Matlab program controlled stimulus creation, stimulus presentation, and response recording. The humans responded by pressing the space bar on the keyboard. The baboons were tested in Marseille. An in-house program written in Turbo Pascal 5.0 controlled stimulus presentation and response recording. The baboons responded by manipulating a joystick. To ensure that image brightness and contrast would be controlled across presentation systems, we adapted the images to the monitor used for each species. Photometric measurements were recorded from the presentation monitor in Marseille and fitted with a gamma function describing the relation between pixel values and monitor brightness. For the monitor used with humans, the gamma value was calculated using a matching procedure and corrected with color lookup-table adjustments available in the Psychophysics Toolbox. The stimuli were presented at 128×128 pixels, subtending a visual angle of $4.5°$.

Training

Half the subjects had to give a "go" response (baboons moved the joystick; humans pressed the space bar) when a human face was shown and a "no-go" response (no movement of the joystick or space bar for 2 s) when a baboon face was shown. The other half had to apply the reverse rule. An auditory tone followed each correct response; in addition, baboons received food pellets, and humans received a credit of 1¢. The baboons were already familiar with the setup and had been trained previously on go/no-go tasks requiring joystick manipulation and pattern discrimination (Martin-Malivel & Fagot, 2001a, 2001b). The human subjects were instructed that they would be making go/no-go responses, but were given no verbal instructions indicating what stimuli they would be seeing or which response to make for which stimuli.

Training for both species involved five phases, as illustrated in Figure 1. In Phase 1, the subjects categorized 60 gray-scale pictures of nonfrontal views of unknown humans and baboons. After performance exceeded 90% correct for both categories, generalization was demonstrated in Phase 2 with 60 previously unseen exemplars of the two categories (half nonfrontal and half frontal views). All subjects performed above 90% correct in the
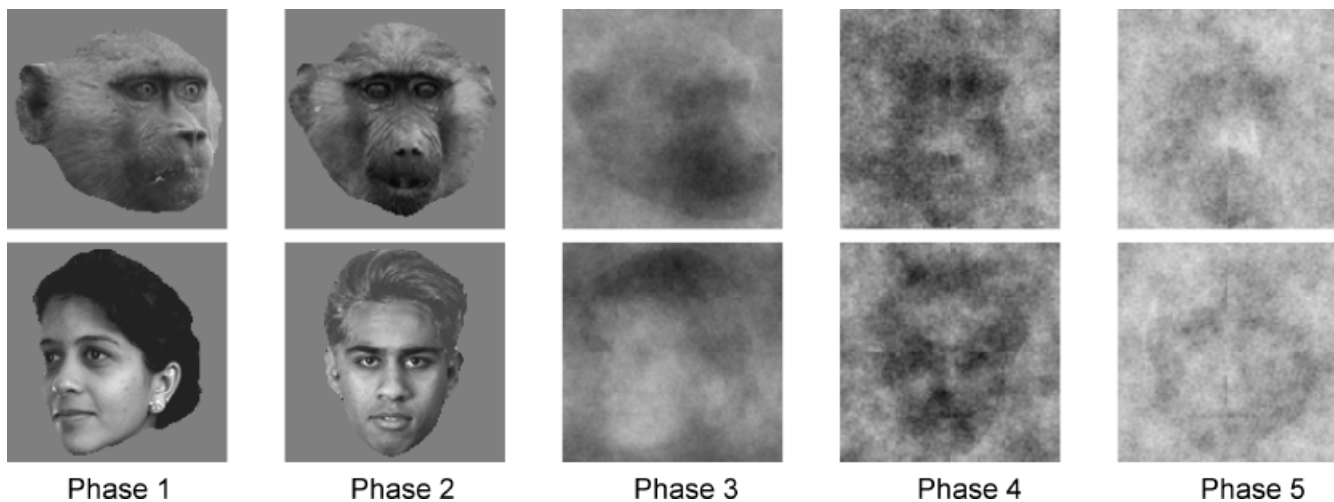
Fig. 1. Training stimuli and phases. In Phase 1, the subjects learned to categorize 60 gray-scale pictures (nonfrontal views) as "human" or "baboon." Phase 2 was a generalization phase in which subjects categorized 60 new exemplars, including frontal and nonfrontal views, of the two categories. In Phase 3, subjects categorized the original nonfrontal views, on which sinusoidal random noise had been superimposed. Phase 4 was a generalization test in which frontal views were shown in high levels of noise, and Phase 5 was a generalization test in which frontal views of warped faces were shown in high levels of noise.

first session. In Phase 3, they categorized the original nonfrontal views, on which sinusoidal random noise had been superimposed. The underlying images were equated in spatial frequency content, contrast, and average luminance. We have previously described in detail the method used to create the sinusoidal noise (Mangini & Biederman, 2004). The signal-to-noise ratio was progressively decreased over the training sessions, thus rendering the discrimination more difficult. Training continued until subjects achieved a performance level of 80% correct with a signal-to-noise ratio of 0.8. Phase 4 consisted of a generalization test in which high-level noise was superimposed on the frontal views used in Phase 2 and was completed when subjects were able to achieve accuracy of 80% or greater. Phase 5 consisted of a generalization test in which high-level noise was superimposed on 30 frontal views of warped faces, in which the positions of the mouth, nostrils, and eyes and the external shape were altered to be identical in all images (see the top two rows in Fig. 2). Subjects performed above chance level. Overall, the humans required, on average, 1,610 training trials (range: 1,470–2,130) to achieve criterion in the training phases; the baboons required 2,040 and 2,280 trials.

We have previously discussed the advantages of using sinusoidal noise instead of Gaussian white pixel noise (Mangini & Biederman, 2004; see also Biederman & Kalocsai, 1997). Computational simulations demonstrated that with moderate numbers of trials (several hundred to tens of thousands), sinusoidal noise produced a better estimate of the template than did white noise.

CI Computation

After training, the CI method was used to infer the information employed by the humans and baboons in categorizing the pic-

tures. Subjects completed a categorization test in which high levels of noise were superimposed on the images (Fig. 2). In this test, 75% of the trials were baseline trials. The baseline stimuli consisted of the 30 warped human and baboon faces used in Phase 5, with different superimposed random visual noise for each trial (Fig. 2, two top rows). Intermixed with baseline trials were probe trials (25% of trials), on which an ambiguous face, a morph between baboon and human faces, was presented (ambiguous trials). Subjects were not informed about these ambiguous trials, for which there was no correct response. Each subject received 200 sessions of 40 trials (30 baseline trials, 10 ambiguous trials). The ambiguous trials were rewarded randomly at the rate at which baseline trials were responded to correctly in the previous testing session. The perceptual appearance of the ambiguous stimulus varied substantially with the noise pattern, so that it was more humanlike in some trials and more baboonlike in other trials (Fig. 2, bottom row). Responses on these ambiguous trials provided data useful for determining which changes to an image induced a "human" or "baboon" response.

For each trial, the noise pattern was assigned to one of the two categories on the basis of the observer's response. After 8,000 trials, we acquired approximately 2,000 categorized noise patterns for a given observer (ambiguous trials). The average intensity for "baboon" classifications was computed, as was the average intensity for "human" classifications. The CI was then computed for each subject by subtracting the average "baboon" pattern from the average "human" pattern (a detailed description of this procedure is given in Mangini & Biederman, 2004). These CIs provide empirically derived linear estimations of the information utilized by each individual subject (Ahumada, 2002).
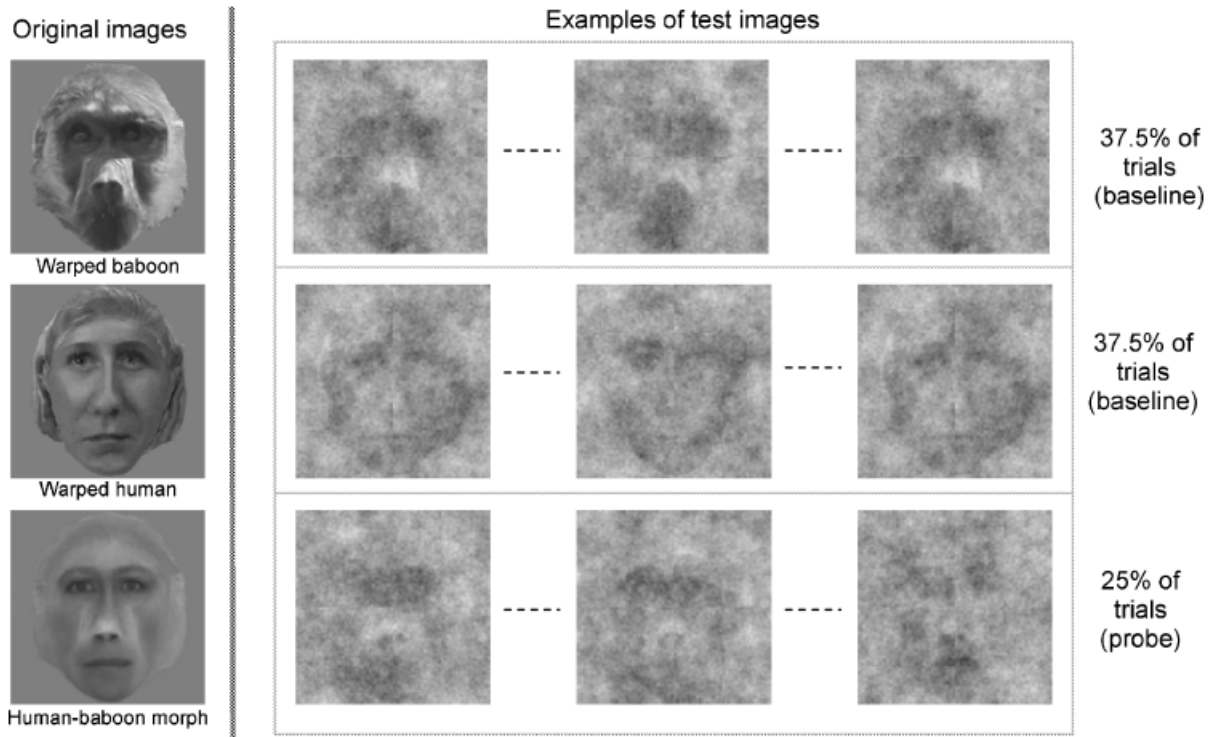
Fig. 2. Illustration of the test stimuli with superimposed random noise. Examples of the original clear images used to create the test stimuli are presented on the left. The top two rows present examples of images presented on baseline trials (75% of the trials): warped frontal views of either baboon (15 individuals) or human (15 individuals) faces. The bottom row presents the unique ambiguous human-baboon morph (left) and examples of this morph with superimposed random visual noise, as presented on probe trials (ambiguous trials; 25% of the trials). This morph was created from 10 previously unseen warped images of baboon and human faces, similar to those shown on the left in the top two rows. Note that the superimposed noise on the morph stimulus rendered the morph more humanlike on some trials, and more baboonlike on others.
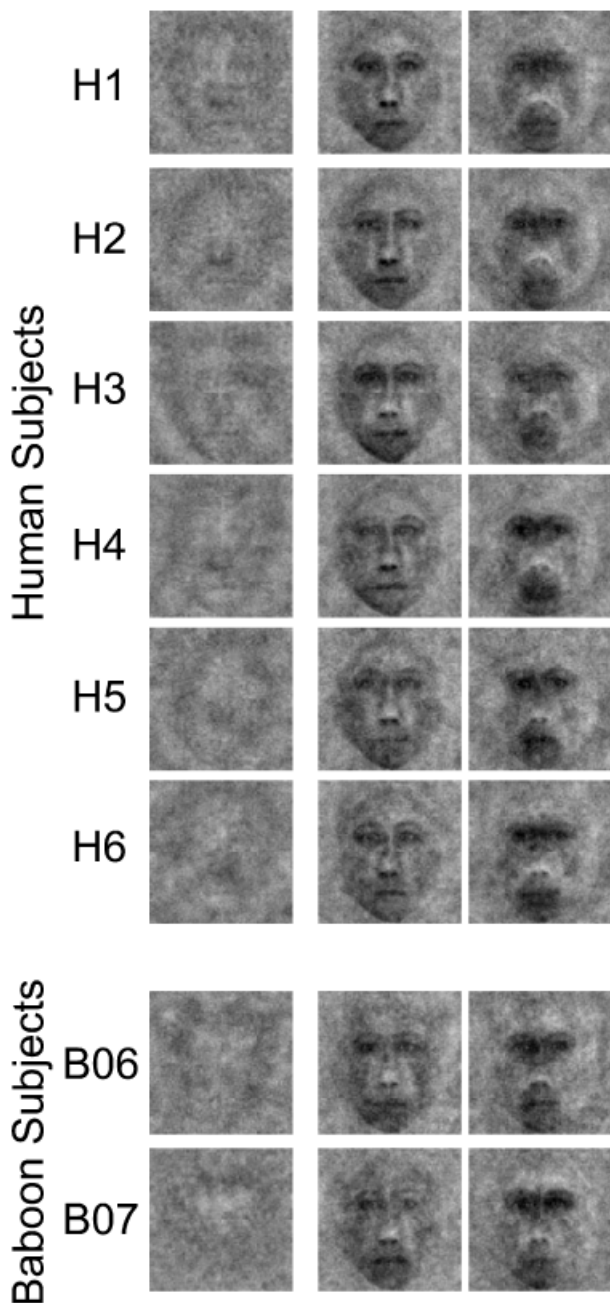
RESULTS

During testing, the signal-to-noise ratio was varied so that each subject's performance was approximately 80% correct. In the final six testing sessions, subjects achieved stable levels of performance at average signal-to-noise ratios of 0.5 for the baboons and 0.4 for the humans. A t test showed no threshold difference between species, $t(6) = 1.225$, $p_{rep} = .67$. Because the two species performed at approximately the same level of accuracy in a perceptually challenging task, it would be tempting to assume that they used the same visual information. However, the CI method allowed us to directly compare the information used by different subjects.

The first column of Figure 3 presents the CIs, which illustrate the information used by each subject to categorize the pictures as "human." The middle column re-creates the typical image that elicited a "human" response from each subject (i.e., the combination of the observer's CI and the original ambiguous human-baboon morph). The third column shows the typical image that elicited a "baboon" response (i.e., the result of subtracting the CI from the original ambiguous morph). These reconstructions make explicit how the dark and light patches of each CI affect the ambiguous image.

Inspection of the CIs suggests that the human and baboon subjects utilized different sources of information to perform the categorization. A coarse contrast between the eyes and the surrounding head region appeared to be the baboons' primary feature for determining species assignment. The raw CIs for human subjects show that, although they also utilized information around the eyes and the top of the head, they additionally used some detailed facial features in the lower half of the face. For example, contrast energy around the nose and chin line appeared in the CIs for all human subjects, indicating that they utilized these areas to make their classification. Neither of the 2 baboon subjects used these cues.

Bootstrap Analysis of the CIs

A bootstrap analysis provided a test of whether the baboons' and humans' CIs differed significantly. First, to characterize the variability within the human population, we compared 1,000 bootstrap sample CIs from a given human subject with 1,000 bootstrap sample CIs from the remaining 5 human subjects. For each of the 1,000 pairs of samples, similarity was measured as the correlation between the pixels of the two CIs. This leave-one-out procedure was repeated for every human, yielding a distribution of 6,000 similarity measures.

Fig. 3. Individual results. The column on the left shows the classification image (CI) for each human (H1–H6) or baboon (B06, B07) subject. Each CI was obtained by subtracting the average of the noise patterns that induced the subject to respond ''baboon'' from the average of the noise patterns that induced the subject to respond ''human.'' The CIs presented here show the information that each subject used to categorize the pictures as ''human.'' The middle column shows the result of adding the CI to the original ambiguous morph, thus re-creating the typical image eliciting a ''human'' response. The column on the right shows the result of subtracting the CI from the original ambiguous morph, thus re-creating the typical image eliciting a ''baboon'' response.

Next, to determine the similarities between the baboons' and humans' use of information, we compared 1,000 bootstrap samples pulled from each baboon's data with bootstrap data from each of the six leave-one-out human populations. Averaging across the 6,000 similarity measures gave the average similarity between a baboon's information use and the humans' information use. The percentile rank of the average similarity value was less than 0.5% (for B06) and 0.8% (for B07). The across-species differences were significantly greater than the within-humans differences (B06: $p_{\text{rep}} = .966$; B07: $p_{\text{rep}} = .956$).

Principal Component Analysis (PCA)

Although the bootstrap analysis provided a test for the planned comparison of baboons' and humans' information use, we also subjected the CIs (unlabeled as to which species they came from) to a PCA. The PCA extracted the dimensions accounting for the greatest variance among the individuals' CIs. The first principal component answered the question: What is the greatest difference among the eight CIs? The first step in the PCA was to subtract the mean of all the CIs from all the images. This was done so that the first principal component would represent the dimension that accounted for the most variance among the images (rather than the mean of the images themselves). The principal components were computed as the eigenvectors of the inner product of the matrix containing all of the mean-subtracted CIs as vectors. The eigenvalues of these components represent the amount of variance accounted for by each principal component.

Results showed that the first principal component (after the mean of all subjects had been subtracted) accounted for 20% of the variance (chance = 14%). Because there is no standard method for testing the statistical significance of a principal component, we performed a Monte Carlo simulation. The null hypothesis tested was that the eight human and baboon CIs were no more redundant than a set of eight "random observer" CIs. These random CIs were generated by taking 1,000 randomly generated noise patterns and randomly assigning them to the "human" and "baboon" categories without any consideration of their visual structure. The eigen-decomposition of the eight random CIs was performed as described for the human and baboon data. The eigenvalues for the random observers were recorded, and the process was repeated 10,000 times to obtain a null distribution. The eigenvalue of the first principal component of our human-baboon PCA was greater than the eigenvalues of all 10,000 Monte Carlo simulations ($p_{\text{rep}} = .996$), showing that there was more redundancy in the eight human and baboon CIs than could be expected by chance. In other words, the 20% of variance accounted for by the first principal component represents a significantly reliable difference among the subjects.

Figure 4a shows the correlation between each subject's CI and the first principal component. This component clearly splits the subjects into two distinct groups: humans and baboons. The images to the left of the graph illustrate the information subjects used to classify the images as human (left) and baboon (right). The top pictures (row 1) illustrate the information that humans tended to use, and the bottom pictures illustrate the information
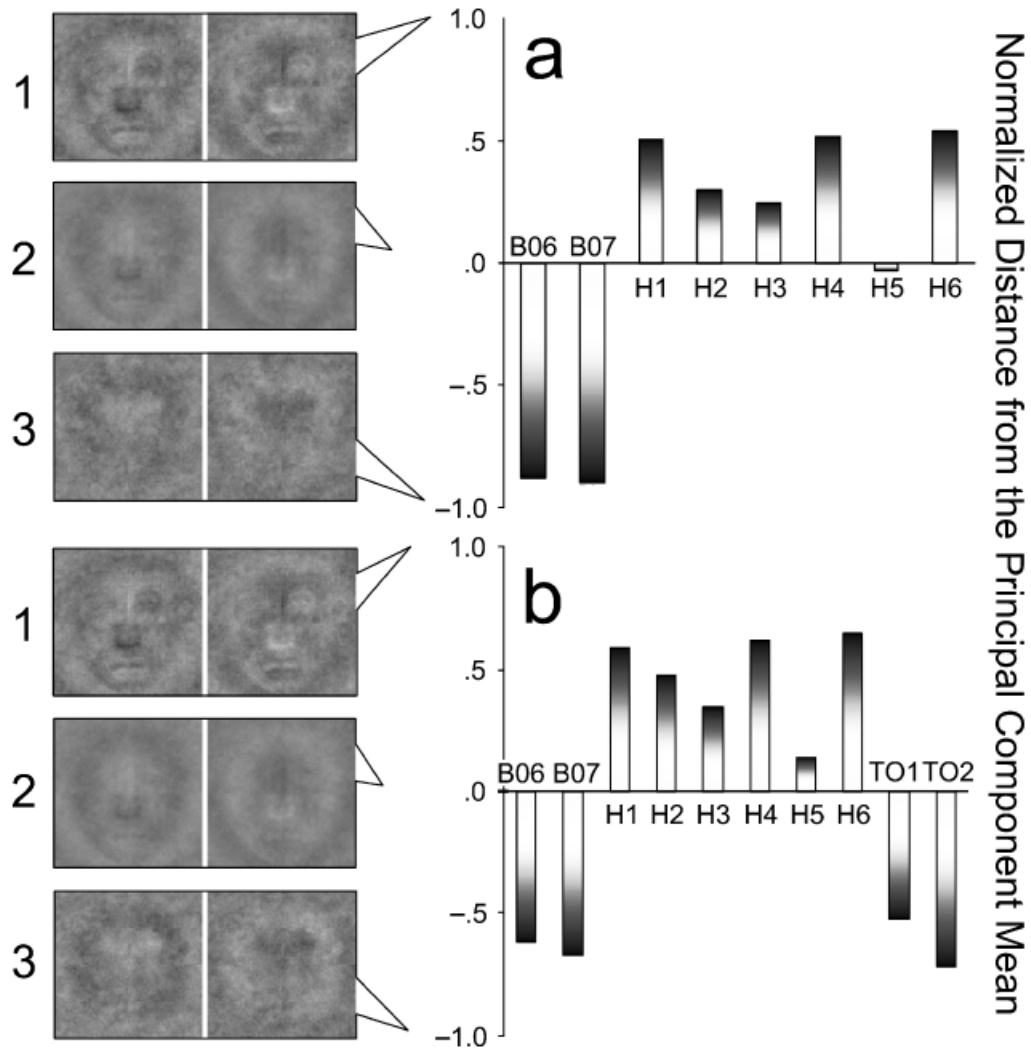
Fig. 4. Results of the principal component analyses (a) for human and baboon subjects and (b) for human and baboon subjects and two runs of the theoretical observer (TO). The individual bars represent the correlations between each subject's classification image and the first eigenvector. The images along the ordinate of each graph illustrate the dimension accounting for the greatest variance among subjects. The three rows depict, respectively, the information used by subjects with high positive projections (+1) on the first principal component, by subjects at the midpoint (0 projection), and by subjects with high negative projections (−1) to categorize faces as human (left) or baboon (right). The subjects with high positive projections tended to use many of the finer details across the entire face, including the eyes, nose, lips, and jawline, whereas those with high negative projections tended to use a coarse contrast between the eyes and the surrounding face to make their discriminations.

that the baboons used (row 3). Inspection of these images reveals that the baboons categorized the faces mainly on the basis of contrast between the eyes and the surrounding head region, responding "human" in the case of a lighter eye area and "baboon" in the case of a darker eye area, whereas the humans paid attention to a broader range of facial features, such as the shape of the nose and mouth and facial contour.

DISCUSSION

The results demonstrate that different species given identical training regimens can utilize substantially different information

to achieve comparable performance. This difference could not be discovered from experiments in which only the performance over all trials was computed, without considering the responses to individual trials. By correlating observers' responses to random noise patterns trial by trial, we were able to determine that the two species used different information to achieve successful categorization. The PCA shows that the largest difference among the CIs was the difference between the CIs of humans and baboons. Both baboons relied heavily on the coarse contrast between the eyes and the surrounding face, unlike humans, who relied on both coarse and detailed information across the entire face.

These results suggest that the eye region was the primary focus of attention for the baboons. This finding is reminiscent of previous findings showing that monkeys freely viewing images of conspecifics tend to look longer at the eye region than at other parts of the face (*Papio hamadryas*: Kyes & Candland, 1987; *Macaca mulatta*: Gothard, Erickson, & Amaral, 2004). Because humans have also been found to focus primarily on the eye region (Henderson, Williams, & Falk, 2005), this is unlikely to explain our species differences. Before one can speculate on the cognitive, social, or evolutionary causes that may underlie the difference in the CIs of our humans and baboons, one should consider the inherent information constraints the task and the stimuli impose on any observer. Any system with the capability of encoding the training images (simply as the pixels that appear on the computer display) and storing them in memory should be able to perform above chance on this classification task. Such a theoretical observer (TO) needs no prior knowledge of humans or baboons, no social interests, and no cognitive representations other than the training images themselves. How similar were our human and baboon subjects to such a TO? That is, to what extent did the subjects utilize information that could be learned from the training images, rather than their knowledge of the appearance of humans and baboons based on their extensive exposure with members and images of the two species?

To address this question, we created a computational simulation mimicking the performance of a naive TO. The Phase 5 warped training stimuli, uncontaminated by noise, were used to compute a discrimination template. This was done by computing the pseudoinverse of a matrix containing all the training images as vectors of pixel values and multiplying the pseudoinverse by a vector containing species indicators for every image (1 for human images, and -1 for baboon images). The result was a linear least-squares template (Bishop, 1995).

In the simulation, the TO referred to this template to classify the image in each test trial as belonging to the human or the baboon category. Each noisy ambiguous image was correlated with the template; if the result was positive, the face was classified as "human," and if the result was negative, the face was classified as "baboon." The simulation was run twice, using the same number of trials as for the human and baboon subjects. CIs for each run of the TO were computed and directly compared with the human and baboon CIs. A PCA was run to determine the dimension that accounted for the greatest variance among the 10 CIs (6 human CIs, 2 baboon CIs, and 2 TO CIs). The first principal component accounted for 18% of the variance (chance = 11%, Monte Carlo calculation: $p_{rep} = .996$) and revealed that the CIs for the baboons and the TO were highly similar and markedly different from those of the humans (Fig. 4b).

The computational simulation reveals what is the simplest and nearly optimal strategy to solve this two-alternative picture categorization task, which involves a limited set of images. Is it possible that the TO reflects the optimal strategy to categorize real-life humans and baboons in natural settings? Because the TO was trained on a limited number of 2-D images (15 pictures per category), photographs taken under specific lighting conditions, we do not believe that it fully reflects the optimal way to sort human and baboon 3-D faces in natural settings. The finding that the baboons' and TO's results were very similar demonstrates that the baboons were highly efficient at ascertaining the relevant information for this specific visual task. It suggests that they relied on the perceptual appearance of the warped training and test stimuli so that their CIs resembled that of the TO. This result is consistent with findings indicating that animals do not always process pictorial stimuli the way humans do, and do not readily recognize objects from 2-D representations (Fagot, 2000). We believe that even in situations in which monkeys recognize represented objects in laboratory tasks, they tend to use low-level stimulus characteristics, when available, to solve the tasks. For instance, we (Martin-Malivel & Fagot, 2001b) showed that baboons considered only facial contours to discriminate clear pictures of known caregivers.

Though our baboons regularly encounter both other baboons and humans in their daily lives, their reliance solely on the training images means that they did not refer to this experience when categorizing our stimuli. Why they chose not to use their everyday experience (or could not do so) is beyond the scope of the present investigation. In contrast, information not specified by the training stimuli seemed to play a key role for the human subjects. We assume that for the humans, conceptual representations of real-life human and baboon faces formed prior to the experiment were critical.

Our findings should encourage detailed investigations of how animals perceive and process 2-D images in laboratory settings. Animals are broadly used as models to investigate the nature and neural bases of various cognitive phenomena, such as categorization (Delorme et al., 2000; Herrnstein & Loveland, 1964; Vogels, 1999a, 1999b), face perception (Bruce, 1982; Parr & de Waal, 1999; Perrett et al., 1986), object recognition (Vogels, 1999b), imitation (mirror neurons; Rizzolatti & Arbib, 1998), and social behavior (Brothers, Ring, & Kling, 1990). These studies often rest on the assumption that the animals process pictorial information in the same manner as humans do. This is not necessarily the case, however, even when their performance is strikingly similar to that of humans, as demonstrated in the present study.

The CI method used here can characterize the information employed by different species and allows explicit comparison of subjects' performance with theoretical models to provide a greater understanding of interspecies differences at an information-processing level. Given the need to use animal data to understand the neural bases of human perception and cognition, the development of methods that make explicit the representations underlying human and animal perception is essential.

REFERENCES

Ahumada, A.J., Jr. (1996). Perceptual classification images from vernier acuity masked by noise [Abstract]. *Perception*, *26*, 18.

Ahumada, A.J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, *2*(1), Article 8. Retrieved April 2002 from http://journalofvision.org/2/1/8/

Ahumada, A.J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *352*, 1203–1219.

Bishop, C.M. (1995). *Neural networks for pattern recognition*. Oxford, England: Oxford University Press.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Brothers, L., Ring, B., & Kling, A. (1990). Response of neurons in the macaque amygdala to complex social stimuli. *Behavioural Brain Research*, *41*, 199–213.

Bruce, C. (1982). Face recognition by monkeys: Absence of an inversion effect. *Neuropsychologia*, *20*, 515–521.

Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 68–77.

D'Amato, M.R., & van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 43–55.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorization of natural scenes does not rely on colour clues: A study in monkeys and humans. *Vision Research*, *40*, 2187–2200.

Fagot, J. (Ed.). (2000). *Picture perception in animals*. Hove, England: Psychology Press.

Fagot, J., Martin-Malivel, J., & Dépy, D. (1999). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? *Current Psychology of Cognition/Cahiers de Psychologie Cognitive*, *18*, 923–949.

Gibson, B.M., Wasserman, E.A., Gosselin, F., & Schyns, P.G. (2005). Applying Bubbles to localize features that control pigeons' visual discrimination behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 376–382.

Gold, J.M., Murray, R.F., Bennett, P.J., & Sekuler, A.B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, *10*, 663–666.

Gosselin, F., & Schyns, P.G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271.

Gosselin, F., & Schyns, P.G. (2003). Superstitious perceptions reveal properties of memory representations. *Psychological Science*, *14*, 505–509.

Gosselin, F., & Schyns, P.G. (2004). No troubles with bubbles: A reply to Murray and Gold. *Vision Research*, *44*, 471–477.

Gothard, K.M., Erickson, C.A., & Amaral, D.G. (2004). How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? *Animal Cognition*, *7*, 25–31.

Henderson, J.M., Williams, C.C., & Falk, R.J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*, 98–106.

Herrnstein, R.J. (1990). Level of stimulus control: A functional approach. *Cognition*, *37*, 133–166.

Herrnstein, R.J., & Loveland, D.H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549–551.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*, 856–867.

Kyes, R.C., & Candland, D.K. (1987). Baboon visual preferences for region of the face. *Journal of Comparative Psychology*, *101*, 345–348.

Mangini, M.C., & Biederman, I. (2001). Differentiating expression, gender, and identity in faces: Comparing normals, the ideal observer, and a prosopagnosic. *Journal of Vision*, *1*(3), Abstract 331. Retrieved December 2001 from http://journalofvision.org/1/3/331

Mangini, M.C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science*, *28*, 209–226.

Martin-Malivel, J., & Fagot, J. (2001a). Cross-modal integration and conceptual categorization in baboons. *Behavioural Brain Research*, *122*, 209–213.

Martin-Malivel, J., & Fagot, J. (2001b). Perception of pictorial human faces by baboons: Effects of stimulus orientation on discrimination performance. *Animal Learning & Behavior*, *29*, 10–20.

Martin-Malivel, J., Mangini, M., Fagot, J., & Biederman, I. (2002, March). *Using reverse correlation to infer the representation of human and baboon faces by human and baboon observers*. Paper presented at the annual meeting of the Vision Sciences Society, Sarasota, FL.

Murray, R.F., & Gold, J.M. (2004). Troubles with bubbles. *Vision Research*, *44*, 461–470.

Parr, L.A., & de Waal, F.B.M. (1999). Visual kin recognition in chimpanzees. *Nature*, *399*, 647.

Pasupathy, A., & Connor, C.E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*, 2490–2502.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Perrett, D.I., Mistlin, A.J., Potter, D.D., Smith, P.A.J., Head, A.S., Chitty, A.J., Broenimann, R., Milner, A.D., & Jeeves, M.A. (1986). Functional organization of visual neurones processing face identity. In H.D. Ellis, M.A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 187–198). Dordrecht, the Netherlands: Martinus Nijhoff.

Rizzolatti, G., & Arbib, M.A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*, 188–194.

Roberts, W.A., & Mazmanian, D.S. (1988). Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal*

of Experimental Psychology: Animal Behavior Processes, *14*, 247–260.

Schrier, A.M., & Brady, P.M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): Effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 136–143.

Schyns, P.G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402–409.

Vogels, R. (1999a). Categorization of complex visual images by rhesus monkeys: Part 1. Behavioural study. *The European Journal in Neuroscience*, *11*, 1223–1238.

Vogels, R. (1999b). Categorization of complex visual images by rhesus monkeys: Part 2. Single-cell study. *The European Journal in Neuroscience*, *11*, 1239–1255.