ELSEVIER

COGNITION

Cognition 57 (1995) 217-239

Categorical effects in the perception of faces

James M. Beale*, Frank C. Keil*

Department of Psychology, Cornell University, Ithaca, NY 14853, USA

Received April 25, 1994, final version accepted February 9, 1995

Abstract

These studies suggest categorical perception effects may be much more general than has commonly been believed and can occur in apparently similar ways at dramatically different levels of processing. To test the nature of individual face representations, a linear continuum of "morphed" faces was generated between individual exemplars of familiar faces. In separate categorization, discrimination and "better-likeness" tasks, subjects viewed pairs of faces from these continua. Subjects discriminate most accurately when face-pairs straddle apparent category boundaries; thus individual faces are perceived categorically. A high correlation is found between the familiarity of a face-pair and the magnitude of the categorization effect. Categorical perception therefore is not limited to low-level perceptual continua, but can occur at higher levels and may be acquired through experience as well.

1. Introduction

The categorization of natural objects in the world, whether they be faces, vehicles or vegetables, is normally thought to involve many levels of processing ranging from general beliefs about the category as a whole to psychophysical invariants that occur across instances. It is normally assumed that categorical perception effects are due to perceptual processes at the psychophysical level. Yet, these low-level categorical perception effects may be indicative of general cognitive processes (Harnad, 1987). This raises the question of whether the techniques for studying low-level perceptual categories might be effectively applied to the study of higher-level categories as well.

* Correspondence to either author. J.M. Beale, E-mail: James_Beale@Cornell.edu. F.C. Keil, E-mail: FCK1@Cornell.edu

SSDI 0010-0277(95)00669-9

The colors of a rainbow illustrate categorization along a natural sensory continuum. A smooth range of light frequencies are present, yet we perceive bands of color rather than a gradual continuum of color change. We perceive discrete shifts in hue. It is easier to discriminate two colors of differing shades when they cross color boundaries (green–yellow) than when they are within the same category (green–green), even though the differences in wavelength are identical for the two pairs (Bornstein & Korda, 1984). Similar results have been found through studies of adults in other cultures, on infants during early development, and of other trichromatic primate species – all of whom conform to approximately the same set of color categories as found in adult English speakers (for review see Bornstein, 1987). In addition to converging behavioral evidence, neurophysiological studies of the lateral geniculate nucleus have found color-sensitive cells responding within four color ranges, roughly corresponding to blue, green, yellow, and red (DeValois & DeValois, 1975). The existence of these color-sensitive cells has generally been interpreted as evidence that color categorization stems from innate mechanisms in the early stages of visual processing.

Another area in which well-established categorical perception effects have been found is that of speech perception. For example, the phonemes /be/, /de/ and /ge/ are stop-consonants which differ in their place of articulation; these three phonemes differ from one another along a continuum of starting frequencies of the transition of the second formant. When subjects are presented with equally spaced stimulus pairs along this continuum, pairs which cross phoneme boundaries are easiest to discriminate (Liberman, Harris, Hoffman, & Griffith, 1957). Similar categorical discontinuities in speech perception have been found in studies of phonemic variation along a continuum of voice onset time (VOT), such as between the consonants /ba/ and /pa/ (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

A number of different theories have been proposed to explain these speech categorization effects. According to early motor theory explanations, speech categories depend on physical differences during articulation (Liberman et al., 1967). It has also been suggested that categories result from natural discontinuities in the auditory continua (Pastore et al., 1977). Interestingly, chinchillas display the same perceptual discontinuities for VOT as do humans (Kuhl & Miller, 1975), which seems to support this latter view. However, it has also been found that the categorical boundaries for VOT continua differ between native, monolingual speakers of Spanish and English (Williams, 1977). Given the contrasting evidence, it has more recently been concluded that speech categories are both constrained by general acoustic properties and subject to tuning through experience (Rosen & Howell, 1987).

According to the standard explanation of the phenomena, categorical perception effects occur when there is a natural physical continuum along which real stimuli normally vary. Harnad (1987) has suggested that the creation of these low-level perceptual discontinuities not only furnish the building blocks for higher level categories, but may provide "a representative model for the categorization process in general". If so, categorical perception effects might not all result from innate processing constraints but might also be expected for more artificial continua, at higher perceptual levels and where no single dimension of variation is obvious.

Faces comprise a class of objects for which human beings attain a high level of expertise at an early age (Diamond & Carey, 1977). As with other objects, faces naturally undergo a number of transformations that change the relations among their features (e.g., changes in expression). They may also be viewed under a variety of conditions which change the prominence of these features (e.g., changes in lighting). Differences between individual faces are even more striking, varying along a multitude of dimensions, yet we are able to recognize familiar faces despite such wide variations and transformations. Infants are capable of discriminating facial expression by the age of 3 months (Barrera & Maurer, 1981; Kuchuk, Vibbert & Bornstein, 1986) and perhaps even at birth (Field et al., 1983; Field, Woodson, Greenberg, & Cohen, 1982).

Facial expressions appear to be perceived categorically (Etcoff & Magee, 1992). Facial expressions naturally vary along a number of continua (e.g., from happy to sad, angry to afraid). As indicated by categorization and discrimination measures for stimuli varying along these continua, pairs of faces which straddled category boundaries are discriminated more easily than are within-category pairs. Although these stimuli were distributed along full continua between the expressions, they were perceived as belonging to discrete categories. Since faces normally exhibit a wide range of expressive states, the categorization of facial expressions serves to disambiguate natural continua. Cells have been identified in the temporal visual cortex that respond selectively to both facial expressions and face identify (Hasselmo, Rolls, & Baylis, 1989; Perrett et al., 1984). As in color and speech perception, it could be argued that the categorization of facial expressions occurs at a low level of perception and is innately specified.

The following studies address the question of whether individual faces are perceived categorically. Categorical effects were found for facial expressions (Etcoff & Magee, 1992), but continua are more plausible between expressions than between individual faces. It has been suggested that the identification of facial expression and face recognition are accomplished by independent perceptual mechanisms (Bruce and Young, 1986; Young & Bruce, 1991). A built-in mechanism for emotion recognition might be feasible given that the meanings attributed to facial expressions are generally constant between individuals and across cultures (Ekman, Freisen, & Ellsworth, 1982). However, the ability to recognize specific individuals must be learned and the continua between individual faces are not naturally occurring. It thus seems less reasonable to expect that individual faces would be perceived categorically or that face recognition is innately determined.

It might be, however, that there are general constraints on category formation that apply to individual face recognition. Given the multidimensional nature of faces, gradual and continuous transformations between faces cutting across all these dimensions might well be seen in a gradual manner. Thus, with non-naturally occurring continua, there may be no perceptual distortion and thus no category boundary effects. Alternately a broader class of "categorical perception" effects may be based less on simple dimensions or bottom-up factors and more on general principles governing category construction. These two competing accounts are the focus of the studies reported here.

EXPERIMENT 1 (DISCRIMINATION AND CATEGORIZATION TASKS)

2. Method

Paired categorization and discrimination tasks were used in this study to test the nature of face perception and representation. Subjects were presented with non-naturally occurring continua between individual familiar faces.

2.1. Subjects

Thirty-two Cornell University undergraduates participated, 16 in each of two experimental conditions.

2.2. Materials

Photographic quality images of famous faces were chosen as base images. All faces were fully frontal, with both ears visible, and a neutral facial expression. Grey-scale images were scanned into a Macintosh™ computer where they were edited in Adobe Photoshop™ to remove backgrounds and everything below the chin. Backgrounds were then filled with a neutral grey. All images were scaled to 243×328 pixels at 72 dpi. Stimuli were generated along linear continua using a "morph" program (Morph™) which, given any two images as endpoints, can produce a linear continuum of images between the two end images.

The morph algorithm has two primary components: warp and fade. "Warping" between two images is accomplished by Delaunay tesselation, a type of finite element analysis which uses linear triangulation. When the algorithm is applied, neighboring control points are connected into optimal triangular regions with non-crossing line segments, resulting in a planar graph. This partitions the image such that all pixels within a particular triangular region are closer to the control points at the triangle's vertices than to any other control points. Warping from one image to another shifts the control points from their initial positions (in one image) to their final positions (in the other) along linear trajectories. All control points are shifted by an equal percentage of the total distance between their initial and final positions; for example, a 30% morph warps all control points 30% of the distance along the linear path between their initial and final positions. The locations of intervening pixels in the images are linearly interpolated across the planar surface based on the positions of their nearest control points, which define the local triangular region. (See Wolberg, 1990, for a more thorough description of Delaunay tesselation.) The warping of the image during morphing can be visualized as the stretching of a rubber sheet; as anchor points on the corners are shifted, all intervening locations also shift in position while maintaining their relations with neighboring locations. The second component of the morph algorithm is a gradual "fade" between the values of corresponding pixels in the two end images. Thus for the 30% morph image, the values of all corresponding pixels are set to a weighted average with the values in the final image contributing 30% while those of the initial image contribute 70%.

Control points were chosen in the following manner. Initially, a single point was placed in the center of each pupil in the two end images. Next, points were placed along the outermost edges of the faces. These early points brought the images roughly into register. The 50% morph image was then consulted to determine what additional points must be added. Any discontinuities or blurring in the morph image were corrected by placing additional points in the two end images. The process of point placement was repeated until all oddities, blurrings, or discontinuities in the 50% morph image were eliminated. For all face-pairs, approximately 300 reference points were entered to "map" the points between the two images of each face-pair (see Figs. 1 and 2). A series of nine images were then produced at 10% increments as a continuum between each of two face-pairs: John F. Kennedy to Bill Clinton and Pete Townsend to Sylvester Stallone. The resultant continua consisted of 11 images with equal steps between them (see Fig. 3a,b).

2.3. Procedure

Stimuli were presented to subjects using an experiment-running program (SuperLab™) on a Macintosh color monitor in grey-scale mode. Each subject viewed stimuli from either the Kennedy/Clinton or Townsend/Stallone continuum. All subjects were presented with a discrimination task followed by a categorization task; these tasks are analogous to those used by Etcoff and Magee (1992). The discrimination task followed an "ABX", matching to sample paradigm. On each trial, three images were displayed consecutively: the first two for 750 ms each and the third for 1 s. A 1 s ISI

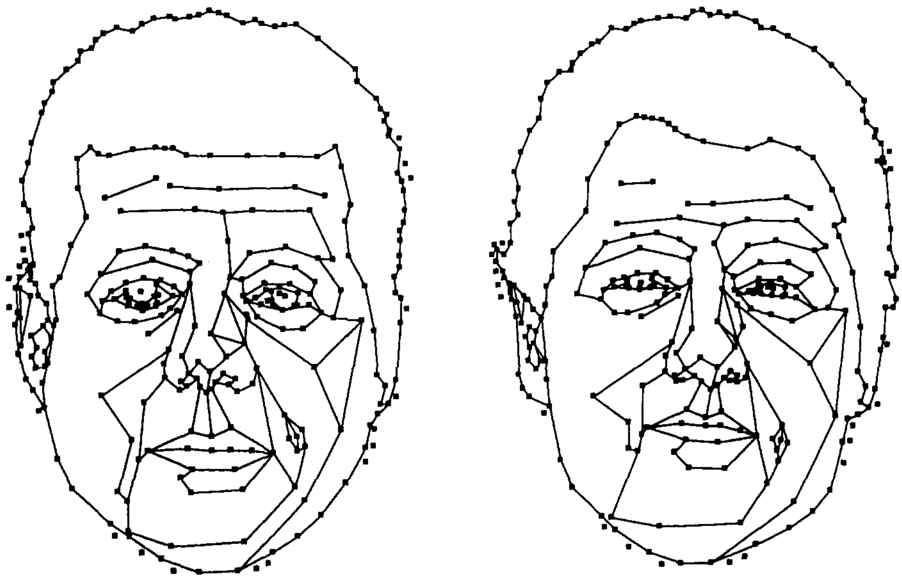

Fig. 1. Morph points for Kennedy and Clinton. For this pair, approximately 300 points are used to "map" locations on one face to locations on the other in a one-to-one correspondence.

consisting of a blank white screen separated consecutive stimuli. The first and second images (A and B) always differed by 20%, or two steps along the linear continuum. The third image (X) was always the same as one of the two previous images. Subjects pressed a key on the keyboard to indicate whether the third image was the same as the first or second image. All nine two-step pairings of the 11 images were presented in each of four orders (ABA, ABB, BAA, BAB) resulting in 36 combinations. Each combination was presented twice to each subject and the resulting 72 trials were fully randomized.

The categorization task consisted of a forced-choice categorization. Stimuli along a face-pair continuum were presented one at a time in random order for 750 ms followed by a blank white screen. After each image was presented, subjects pressed a key corresponding to either Kennedy or Clinton (or Townsend or Stallone). Each image was presented eight times resulting in 88 randomly ordered trials.

3. Results

In both the Kennedy/Clinton and Townsend/Stallone conditions, subjects judged stimuli as belonging to distinct categories with a sharp boundary

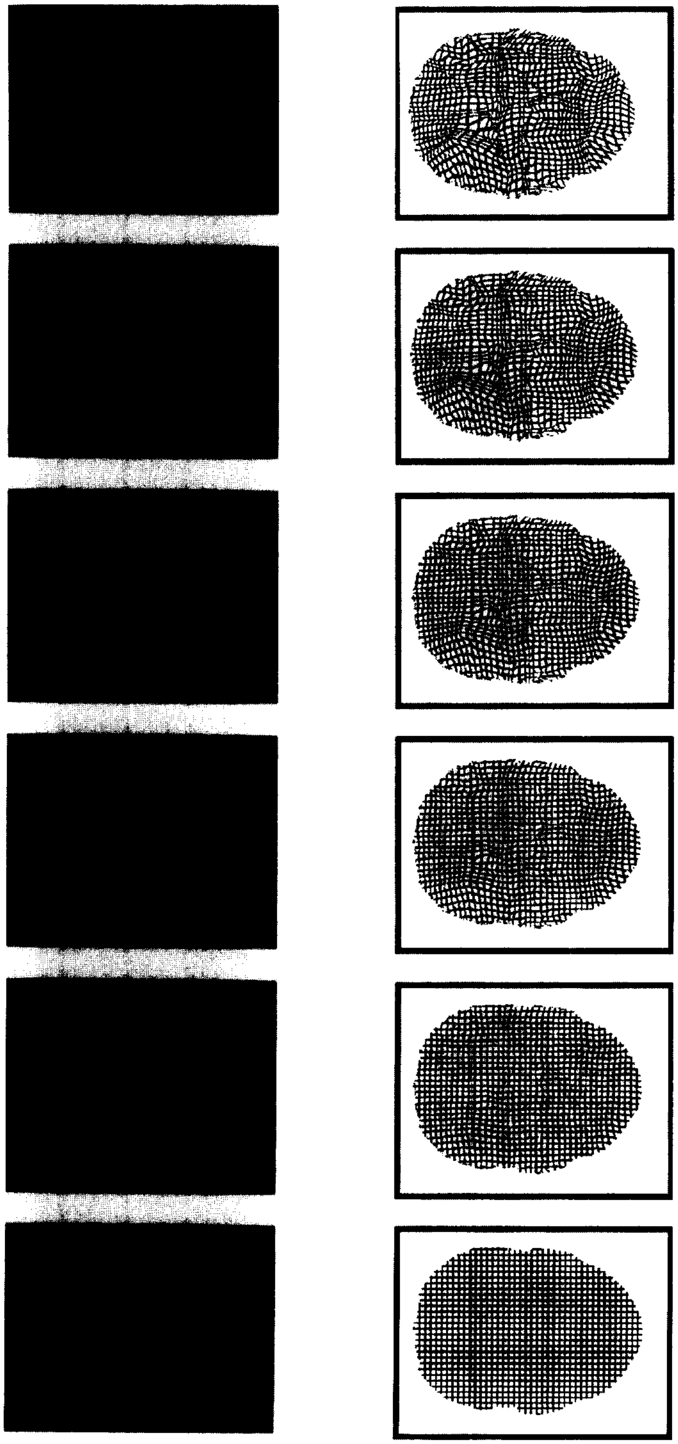

Fig. 2. When the morph algorithm is applied, each point is moved along a linear path from its start position to its final position. The program also gradually shifts the pixel intensity from its initial to final value for corresponding pixels. The bottom row displays the linear shift from Kennedy to Clinton. The top row shows examples of the morphed images.

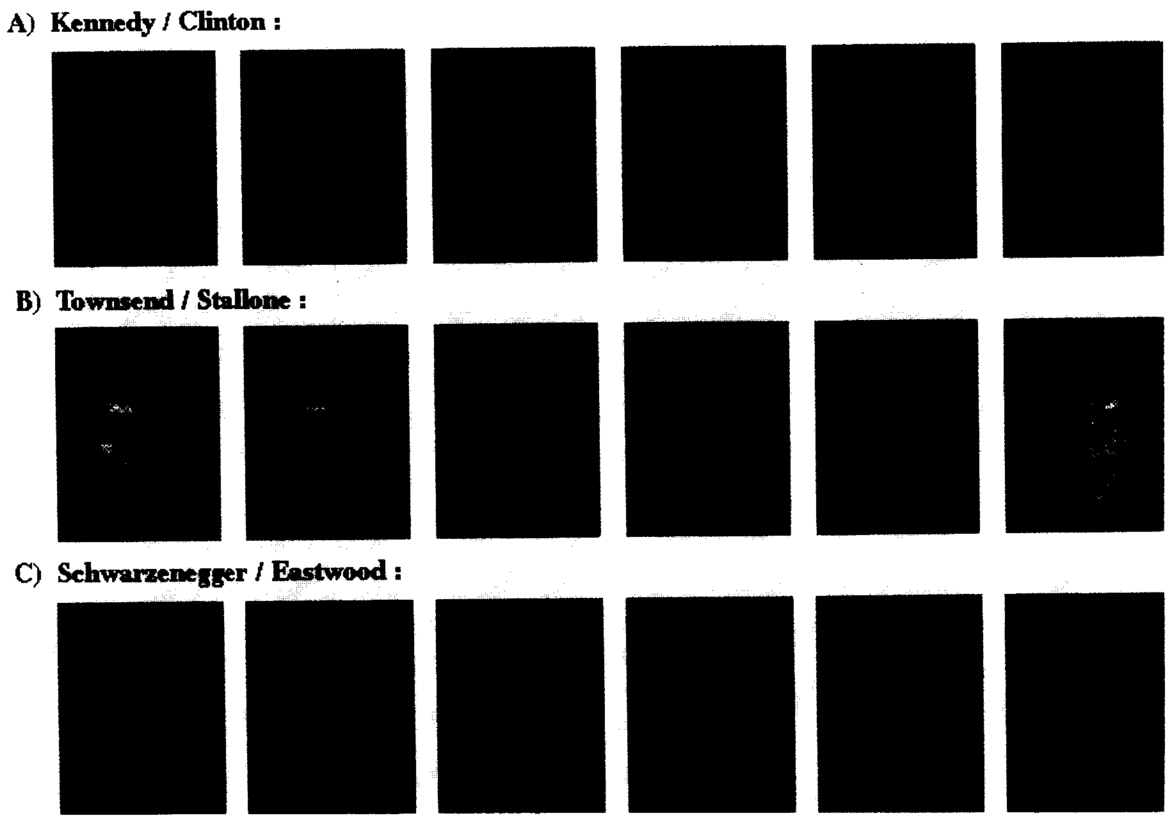

Fig. 3. Stimuli used in Experiment 1 (a and b) and Experiment 2 (a, b, and c). Six of the 11 stimuli in each of the two morphed continua are shown. Consecutive images differ by 20%. In order, they represent the 1%, 20%, 40%, 60%, 80%, and 99% images. Images at 1% and 99% were used (not 0% and 100%) to ensure that all images had undergone the morph process.

between them. From the eight presentations per subject, a between-subject average was computed for each image in the categorization task. The categorization task data in both conditions showed a clear shift in identity judgements (see Fig. 4). However, this shift is not definitive since it may be an artifact of the two-choice forced-choice judgements by subjects. By defining a 33% and 66% cut-off for the category boundary, the categorization task data were used to predict performance in the discrimination task. If stimuli along a continuum are categorically perceived, a peak in accuracy would be expected in the discrimination task for the two-step pair that straddles the boundary. From the eight presentations per subject of each two-step pair, percent accuracy scores were computed for each subject. On an informal subject-by-subject analysis, for over half of the subjects a peak in discrimination was correctly predicted from individual categorization task data. However, percent accuracy scores averaged between subjects provided a more direct measure of the overall effect (see Fig. 4); all further studies report only between-subject data. Planned comparisons were performed on the accuracy scores at the predicted peaks; in each condition, accuracy for the pair that straddled the boundary was contrasted with the mean accuracy on all the other pairs combined. Pairs at the predicted peaks were found to be significantly higher in both conditions: Kennedy/Clinton $F(1, 15) = 12.626$, $p < .003$; Townsend/Stallone $F(1, 15) = 18.178$, $p \leq .0007$. Addi-

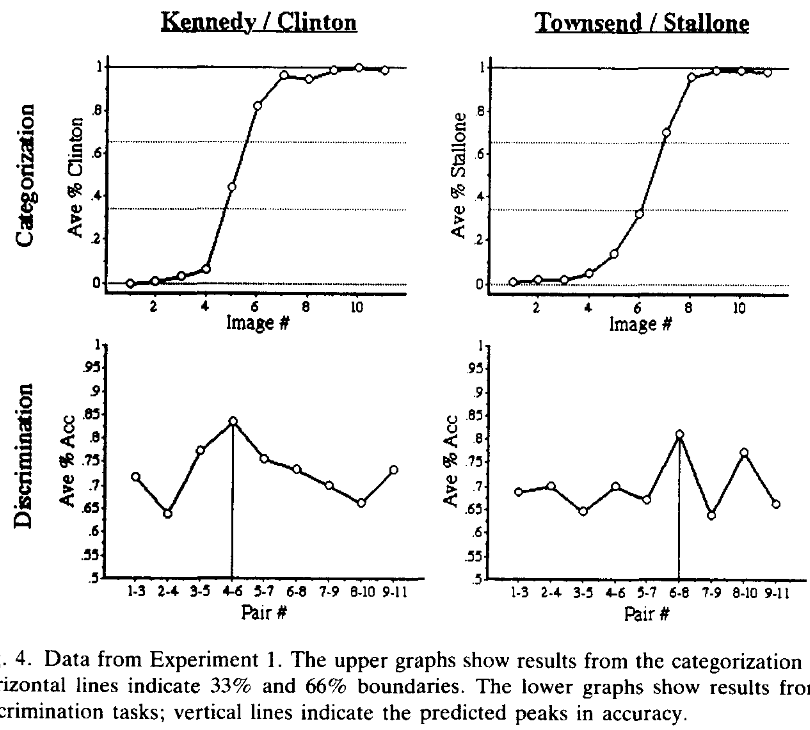

Fig. 4. Data from Experiment 1. The upper graphs show results from the categorization tasks; horizontal lines indicate 33% and 66% boundaries. The lower graphs show results from the discrimination tasks; vertical lines indicate the predicted peaks in accuracy.

tionally, face-pairs to either side of the predicted peaks partially straddle the category boundary. Since crossing the category boundary should cause an increase in accuracy, we would expect these pairs to also be more easily discriminable than are fully within-category pairs. This expectation seems to hold for the Kennedy/Clinton but not the Townsend/Stallone condition.

4. Discussion

Both continua of face stimuli are perceived categorically. While the results do suggest categorical perception, categorization might not be occurring at the level of face perception. The discrimination task allows subjects to use any available information within the images to make their judgements. Instead of perceiving the faces holistically, subjects could be focusing on local feature differences between the images. Thus, the results may not display face perception as such but rather categorical discrimination of lower-level features. For example, if subjects were able to focus on small changes in pixel intensities or contrast levels, focusing on these changes might account for the pattern of the data. Similarly, the data might indicate a categorical shift in the shape of the nose or hair rather than of the overall face. However, it is unclear why the midpoint of the continuum would be where such changes in pixel intensities would be most pronounced.

EXPERIMENT 2 (BETTER-LIKENESS AND CATEGORIZATION TASKS)

5. Method

Since local differences within the images may influence performance, a new task was developed to bias subjects towards more holistic processing. In this new task, rather than merely discriminating between the images, subjects must judge which of two images is the better likeness of a particular person. When presented with two images taken from a single continuum, each image is closer to a particular end of the continuum. On each trial of the "better-likeness" (B-L) task, subjects judge which of two stimuli was more like a particular person. Accurate discrimination is thus correctly judging which of the two images is closer to a particular end of the continuum (e.g., "more like Kennedy"). To discriminate, subjects must rely on their own mental representation of these individual faces. As in Experiment 1, the endpoints of all face-pairs in this study are the faces of highly familiar individuals; thus, subjects already possessed visual representations of these individual faces.

5.1. Subjects

Sixteen Cornell University undergraduates participated.

5.2. Materials

The Kennedy/Clinton and Townsend/Stallone continua from the previous study were used as stimuli. An additional face-pair continuum from Arnold Schwarzenegger to Clint Eastwood was also included. These new images conformed to all of the criterion as described in the previous study, except that both faces were smiling (see Fig. 3a–c).

5.3. Procedure

Each subject was presented with the better-likeness (B-L) task followed by three categorization tasks. On each trial in the B-L task subjects determined which of two images, presented simultaneously, was closer to a particular end of one of the continua. Prior to the presentation of stimuli on each trial, the following question appeared on the computer screen: "Which one of the following two faces looks more like ____?" where the blank contained the name of one of the six individuals used as stimuli. The name was always appropriate to the stimulus continuum presented, but was counterbalanced between the two appropriate names. Subjects then saw both images of a two-step pair (A and B) simultaneously for 1 s followed by a white screen with the numbers "1" and "2" below the area where each image had appeared. Subjects would respond by pressing a key (1 or 2) to indicate whether the left or right image was more like the person who had been named. Two-step pairs from all three continua were randomly intermixed. The 27 two-step pairs were presented in each of four orders (name A–AB, name A–BA, name B–AB, name B–BA) resulting in 104 combinations. Each combination was presented only once.

Stimuli for each categorization task consisted of one of the three morphed continua. The method of these tasks was identical to that described in Experiment 1, except that each image was presented only twice.

6. Results

Subjects judged stimuli along all continua as belonging to distinct categories with a sharp border between them. Between-subject averages were computed for each image in the categorization task. In all three conditions a clear shift was found (see Fig. 5). The 33% and 66% cut-offs for the categorical boundary were used to predict performance in the B-L

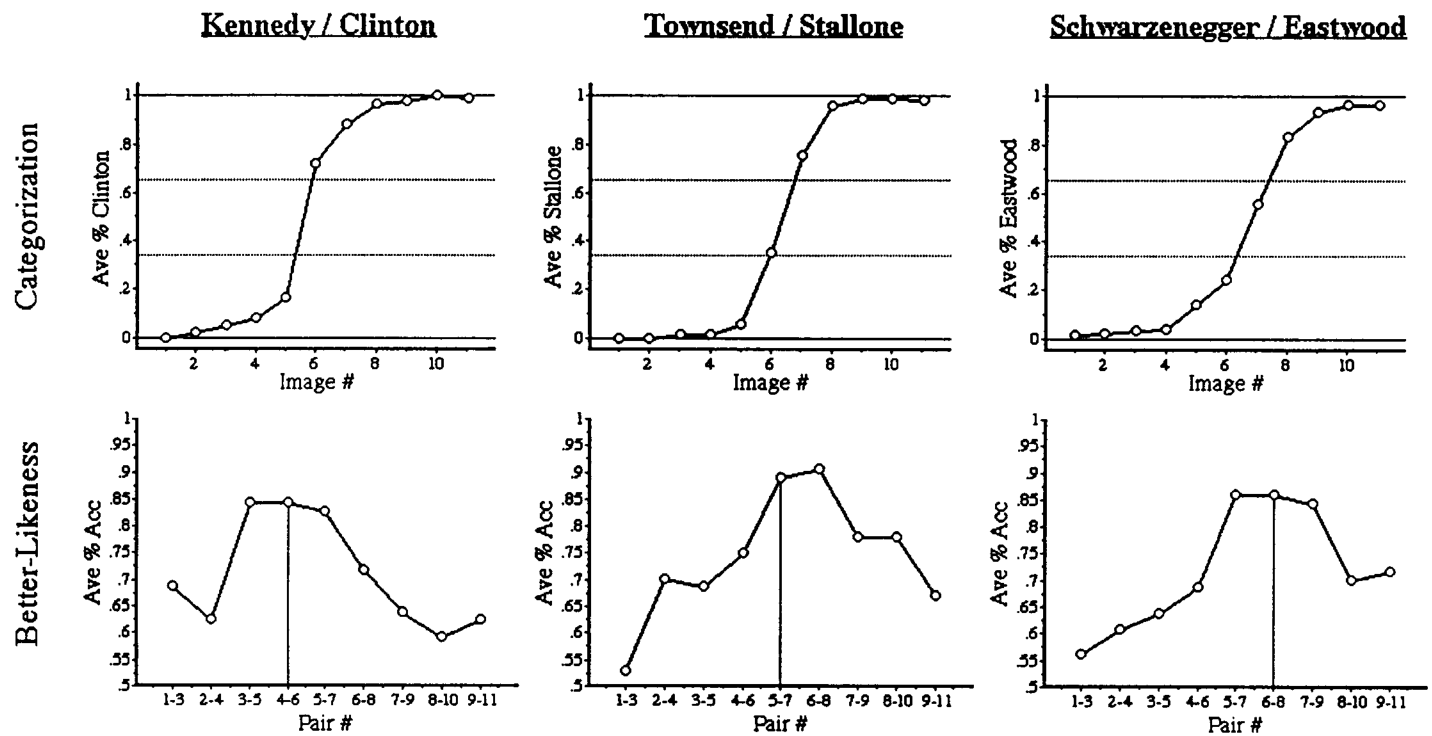

Fig. 5. Data from Experiment 2. The upper graphs show results from the categorization tasks; horizontal lines indicate 33% and 66% boundaries. The lower graphs show results from the discrimination tasks; vertical lines indicate the predicted peaks in accuracy.

task. Peaks in accuracy were predicted in the B-L task for the two-step pairs that straddled the boundary in the categorization task. Percent accuracy scores were computed and averaged between subjects (see Fig. 5). Planned comparisons were performed on the accuracy scores of the two-step pairs which crossed the boundary to determine if they were higher than the mean accuracy of all other pairs combined. All three comparisons were significant: Kennedy/Clinton $F(1, 15) = 7.06$, $p < .0342$; Townsend/Stallone $F(1, 15) = 38.901$, $p < .0001$; Schwarzenegger/Eastwood $F(1, 15) = 8.696$, $p \leq .0050$. In all three conditions, face-pairs which only partially straddle the category boundary were also easier to discriminate than within-category pairs.

7. Discussion

Individual face stimuli are perceived categorically. Peaks in discrimination accuracy were found for those face-pairs which fully straddle the category boundaries. All B-L task data have single-peaked distributions centered on the predicted peaks. In the discrimination task data for Townsend/Stallone in Experiment 1, the 8–10 pair was a fully within-category pair which would have been significant if it had been the predicted peak (see Fig. 4); between it and the actual predicted peak was the lowest point of discriminability in that condition (the 7–9 pair). Unlike the discrimination task data of Experiment 1, no other peaks in accuracy are found in the B-L task data aside from pairs which straddle the category boundaries (either fully or partially). In addition, cross-category pairs show higher accuracies while within-category pairs show lower accuracies than comparable pairs in the discrimination tasks of Experiment 1. In these respects, the B-L task data are much more stable, with less noise and cleaner peaks than were seen in the discrimination task data for the same face pairs. To the extent that the B-L task constrains subjects' discriminations to judgements of faces as a whole, categorization is occurring at the level of face perception and not local features. Although face representations are flexible enough to deal with "unnatural" variations in features, their bounds seem to be sharply delimited.

Categorical effects have thus far been shown for three pairs of highly familiar faces. Before considering further implications, the generality of this effect needs to be explored. Are all faces perceived in a categorical manner, and if not where do the effects occur? Two contrasting predictions are possible: (1) perhaps it is intrinsic to how we code faces, even after the briefest single exposure, to build them into representations with sharp categorical boundaries; alternatively, (2) categorical effects in recognition may result from expertise with specific faces. If so, the effect should vary with the familiarity of the individual face; highly familiar faces should show the strongest categorical effect, while unfamiliar faces show the lowest.

EXPERIMENT 3 (BETTER-LIKENESS AND CATEGORIZATION TASKS)

8. Method

8.1. Subjects

Sixteen Cornell University undergraduates participated.

8.2. Materials

This study used pairs of faces at differing levels of familiarity, including completely unfamiliar faces. For the familiar faces, familiarity was acquired through subjects' exposure in the media (i.e., television, film, etc.). While the exact degree of exposure to these faces was not experimentally administered, levels of familiarity were taken into account. Preceding this study, a database of 38 faces was collected. Each face within this database conformed to the criterion described in Experiment 1, namely: photographic-quality grey-scale images, fully frontal, with both ears visible, neutral facial expression, cropped below the chin, neutral grey backgrounds, and scaled to 243×328 pixels at 72 dpi.

In an independent rating task, 16 subjects rated all 38 faces on 9-point scales for familiarity, distinctiveness, apparent age, and photographic quality. Using the sum of the squared difference scores, four new face-pairs were selected such that within-pair differences were minimized on all four scales. The four face-pairs differed between-pairs in familiarity as follows: Dustin Hoffman (7.47)/Michael Douglas (7.73), high familiarity; Carey Grant (5.56)/Jack Lemmon (5.05), moderate familiarity; Barry Tubb (2.69)/Christopher Atkins (2.81), low familiarity; Kevin Burns (1.88)/Jason Harris (2.56), low familiarity; this final face-pair was comprised of previously novel, unfamiliar faces. Morphed continua with 10% increments were generated for each of these face-pairs (see Fig. 6).

8.3. Procedure

Since some faces were known to be unfamiliar, subjects were presented with an initial training screen containing the original faces of each of the four face-pairs in two rows, four to a row, in alphabetical order. The name of the individual was given below each face. Subjects were allowed to study these faces until they felt they could match the correct name to each face. Subjects were then presented with a B-L task followed by four categorization tasks. Two-step pairs from all four continua were randomly intermixed in the B-L task. Stimuli for each categorization task consisted of one of the

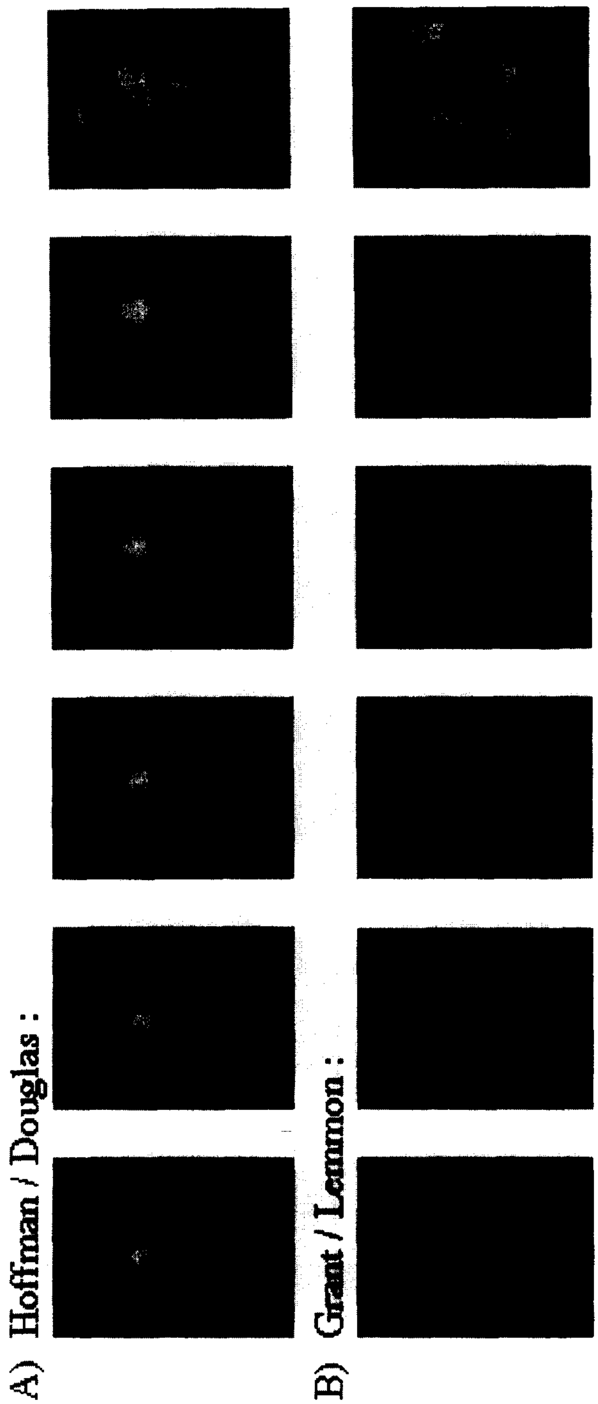

Fig. 6. Stimuli used in Experiment 3. Six of the 11 stimuli in each of the four morphed continua are shown. Consecutive images differ by 20%. In order, they represent the 1%, 20%, 40%, 60%, 80%, and 99% images.

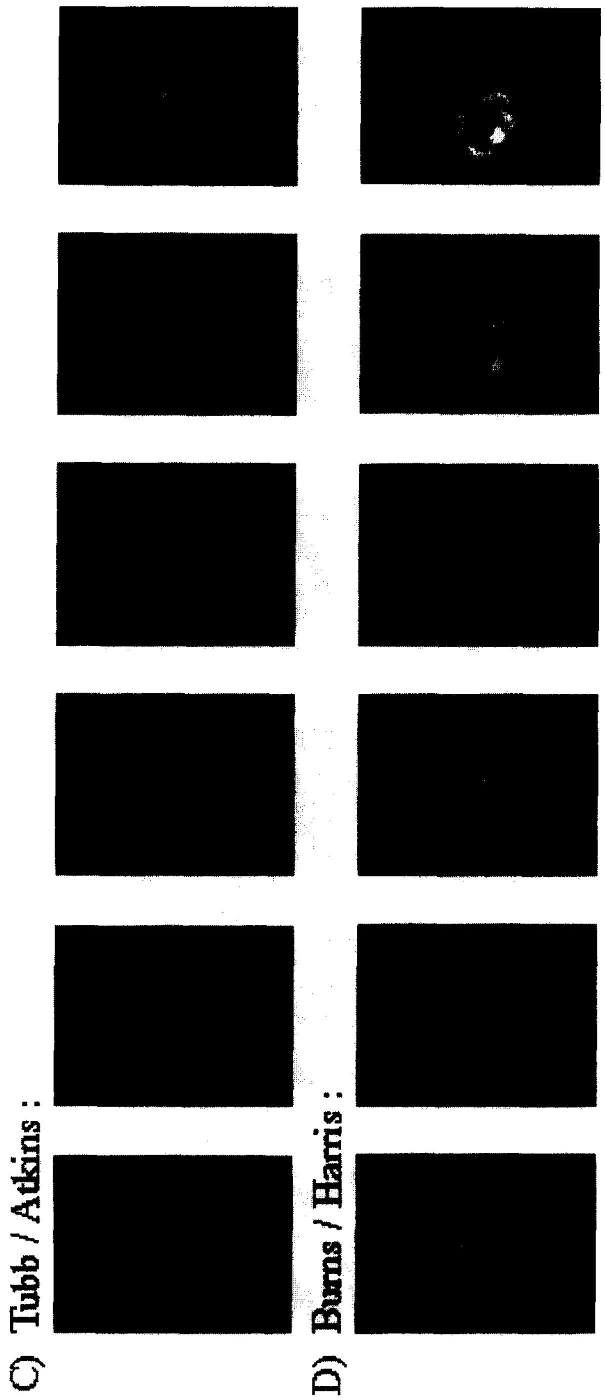

Fig. 6. *(Continued)*

four morphed continua. Other aspects of these tasks were identical to those in Experiment 2.

9. Results

Between-subject averages were computed for each image in the categorization task. In all four conditions a clear shift was found (see Fig. 7). The 33% and 66% cut-offs for the categorical boundary were used to predict performance in the B-L task. Peaks in accuracy were predicted in the B-L task for the two-step pairs that straddled the boundary in the categorization task. Percent accuracy scores were computed and averaged between subjects (see Fig. 7). Planned comparisons were performed on the accuracy scores of the two-step pairs which crossed the boundary to determine if they were different from the within category pairs. Both the Hoffman/Douglas and the Grant/Lemmon comparisons were significant: $F(1, 15) = 7.105$, $p < .0078$ and $F(1, 15) = 3.286$, $p \leq .0450$, respectively. The remaining comparisons

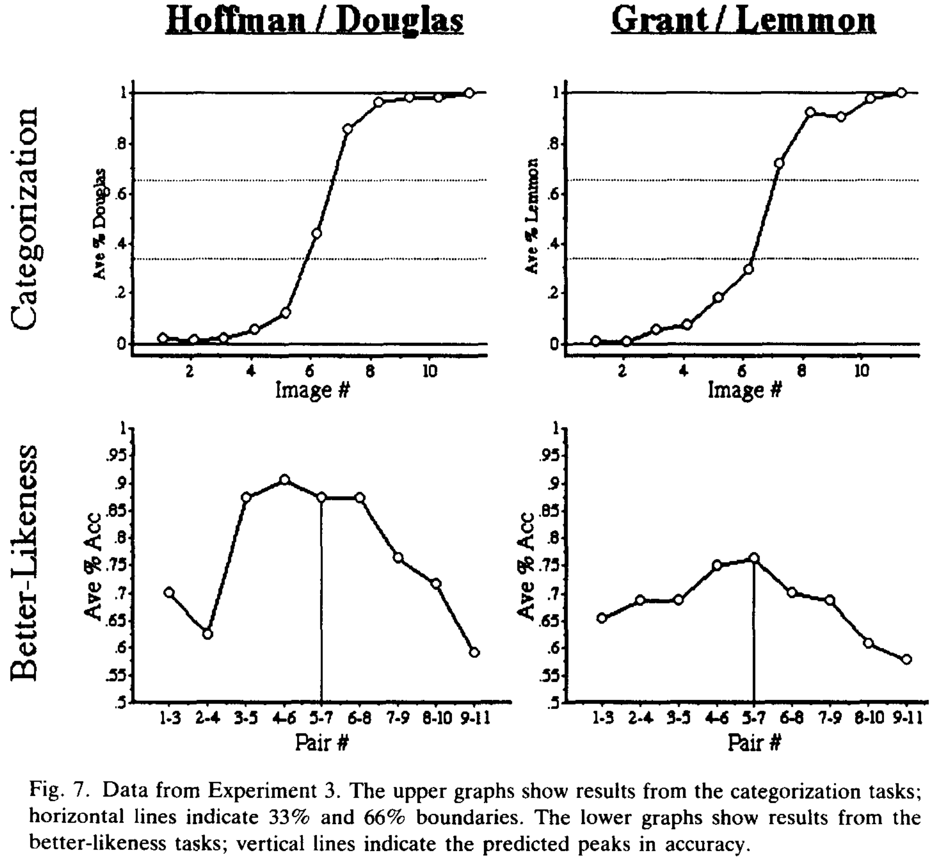

Fig. 7. Data from Experiment 3. The upper graphs show results from the categorization tasks; horizontal lines indicate 33% and 66% boundaries. The lower graphs show results from the better-likeness tasks; vertical lines indicate the predicted peaks in accuracy.

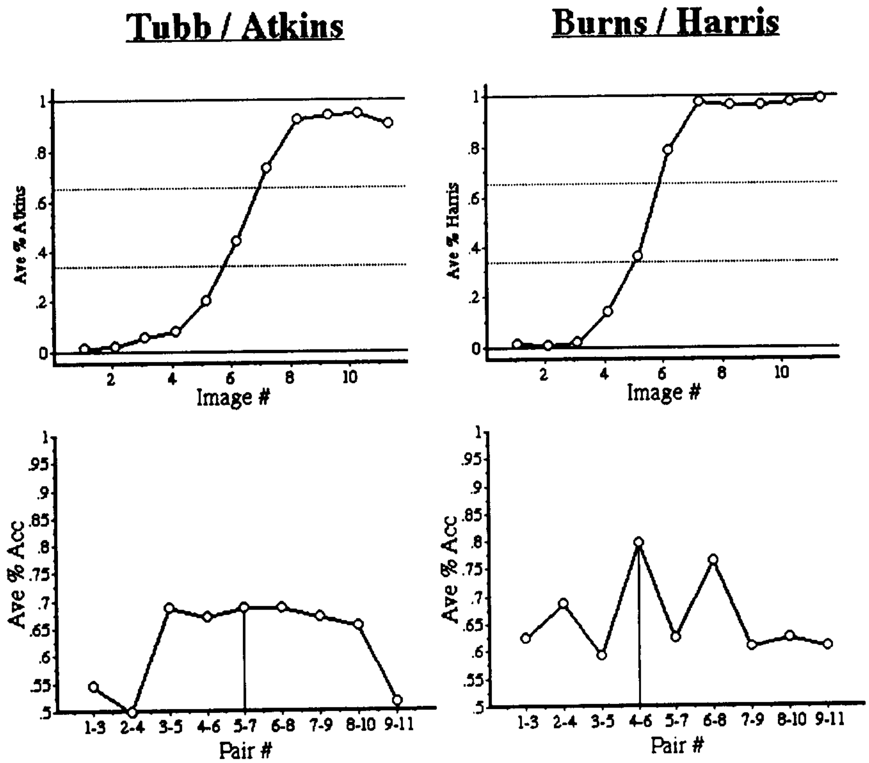

Fig. 7. (*Continued*)

were not significant: Tubb/Atkins $F(1, 15) = 1.462$, $p \leq .1227$; Burns/Harris $F(1, 15) = .48$, $p \leq .2495$.

10. Discussion

In this experiment familiarity was systematically varied across individual face-pairs to study its effect on categorization. Two low-familiarity conditions were included to offset the use of only highly and moderately familiar face-pairs in Experiment 2: Kennedy (8.44)/Clinton (8.94); Schwarzenegger (8.94)/Eastwood (8.63); Townsend (4.00)/Stallone (8.94). Taken together with the results from Experiment 2, it is clear that subjects' familiarity with particular face stimuli is highly predictive of the degree of the categorization effect. There is a high correlation between the average familiarity rating of a face-pair and the magnitude of the categorization effect, $r = -.848$ (see Table 1).

In this final study, the level of familiarity of the individual faces was systematically studied based on previous levels of exposure; this allowed us

Table 1
Correlation of familiarity with magnitude of effect

Pair	Average familiarity	Magnitude (p-value)
Schwarzenegger/Eastwood	8.78	.0050
Kennedy/Clinton	8.69	.0342
Hoffman/Douglas	7.60	.0078
Townsend/Stallone	6.47	<.0001*
Grant/Lemmon	5.31	.0450
Tubb/Atkins	2.75	.1227 n.s.
Burns/Harris	2.22	.2495 n.s.

Note: Familiarity ratings are the averages of the individual ratings for the members of each face-pair; p-values indicate the magnitude of the categorization effect as exhibited by the planned contrasts of the predicted peaks in accuracy. All values are from the better-likeness task data of Experiments 2 and 3. The familiarity of the face-pairs is correlated with the magnitude of the categorization effect ($r = -.848$).

* The low p-value for the Townsend/Stallone pair may result from a relatively large mismatch in individual familiarity scores, 4 and 8.94 respectively.

to establish the existence of the phenomenon, while eliminating the need for time-consuming laboratory training of subjects. Studies of trained familiarity for faces have been planned and should provide additional information about the nature of face representations. A variety of techniques are currently being developed to determine what type and how much exposure is necessary for categorical perception effects to appear.

GENERAL DISCUSSION

Familiar faces are perceived categorically at the level of individual face representation. Thus, categorical perception is not limited to low-level perceptual continua, but can occur at higher levels as well. Since morphed continua between individual faces are not naturally occurring, categorical boundaries for specific continua need not be innately constrained; categorical perception effects can be acquired through experience.

The better-likeness task of Experiments 2 and 3 used modified psychophysical techniques to demonstrate "higher-level" categorical perception. As such the results reported in this series of studies challenge previous conceptions of categorical perception, indicating that they are much more general than has commonly been believed. However, the full scope of this phenomenon is yet to be determined and many additional questions remain to be answered.

What information is used to code individual faces? What is the nature of the representation? One possibility is that the underlying structure of faces is represented in terms of acceptable (and unacceptable) deformations of component elements. While facial variations certainly add to the complexity of face processing, it has recently been suggested that "non-rigid variations created by expressive movements of the face may not actually make face recognition a more difficult problem . . . but may actually facilitate discrimination within a class of objects that all share the same overall structure" (Bruce, 1994).

Any face can be described in terms of its constituents that remain constant due to underlying bone structure and those that vary due to deformations of the soft tissues, the muscles and skin (as well as motions of the jawbone). As the structural invariants are learned for any given face, the acceptable range of deformations can be computed more accurately. Once the bone structures of a face are represented, the boundaries for the representation could then be determined based on general principles of the deformability of substances.[1] Since the bone structure invariants for novel faces are unknown, no categorical bounds should exist between individual unfamiliar faces.

In addition to explaining the lack of category boundaries for unfamiliar faces, the "structural invariants" hypothesis predicts that the effect should vanish for any face-like stimuli which do not deform. One example of famous non-deforming "faces" is the muppets of Jim Henson. Individual muppets are easily recognized by many individuals yet their faces are semi-rigid, the motion of their jaws being their primary deformation. Preliminary results of pilot studies using morphed continua between muppet "faces" indicate no categorical effects, thus providing tentative support for the structural invariants hypothesis. However, whether or not the structural invariants hypothesis is correct, it makes no mention of underlying mechanisms.

By another strategy, rather than focusing on facial structures that remain constant, representations may take advantage of natural facial variations. One way to represent these variations is as vectors within a multidimensional feature space, or face space. Variations in any given feature can be thought of as a single dimension, or vector. Each instance of a face can thus be defined as a conjunction of vectors which specify a point in feature space (see Valentine, 1991, for a more complete description). Computational models using multidimensional vector space representations have been used successfully to categorize faces on the basis of sex, to discriminate familiar from unfamiliar faces (O'Toole, Abdi, Deffenbacher, & Valentin, 1993), and to recognize individual faces, even when partially occluded (Turk & Pentland, 1991).

If faces are represented in terms of component vectors, several predictions can be made. Since individual faces vary, the representation of any single individual must define a region, rather than a point, in face space. Different individuals would occupy different regions of face space by virtue of their dissimilar features. The results of Experiments 1, 2, and 3 indicate

[1] Thanks to Ulric Neisser for raising this possibility in discussion.

that higher familiarity is associated with greater distortion of face information; as subjects become familiar with any given face, the nature of their internal representations change. According to this model, changes due to experience should serve to separate various instances from one another in feature space. However, the mechanism by which these changes occur are not well explained and could take many different forms.

Evidence suggests that faces are represented in terms of their similarity to and difference from previously experienced exemplars. Researchers have found advantages in subjects' ability to recognize highly distinctive faces (e.g., Light, Kayra-Stuart, & Hollander, 1979; Rhodes, Brennan, & Carey, 1987). Other studies have suggested that people make use of averaged face information when making familiarity judgements; when shown face stimuli that are the average of a set of unfamiliar faces, subject judge them to be highly familiar (Bruce, Doyle, Dench, & Burton, 1991; Valentine & Bruce, 1986). In light of these findings, a number of prototype and exemplar-based models have been proposed to account for face recognition phenomena (see Valentine, 1991).

Standard prototype models are similar to those that describe faces in terms of multidimensional vectors; the primary difference is that, in prototype models, each instance is represented in terms of its similarity to some standard, or prototypical exemplar. Individual faces might then be recognized by matching to the nearest stored exemplar. Although prototype theories describe how faces might be represented and recognized, they provide little explanation of how category boundaries are formed or how they change with experience. The mechanisms by which faces are perceived categorically have yet to be adequately accounted for by any theoretical approach.

There are many different phenomena in numerous domains that fall under the term "categorical perception" and these effects vary in the degree to which they are affected by experience. Although color categorization seems to depend on innate mechanisms, phoneme categorization most likely has both innate and acquired components, and face categorization appears to be primarily an acquired phenomenon. This may mean that categorical perception does not result from a single, specific mechanism, but rather is a general information processing strategy taking many different forms which serve to simplify sensory input. Rather than asking "Is categorical perception innate or acquired?" we would do better to ask "What are the relative contributions of innate mechanisms and of experience for particular categorical perception effects?"

Although a face must be familiar before category boundaries are found, how much and what kind of experience is necessary? Further studies in which subjects are trained to become familiar with individual faces will help in answering this question. These techniques could also be used to test for categorical effects in the perception of objects in other domains and to help us understand why such effects occur. As such, these additional studies should deepen our understanding of the nature and formation of object categories. If categorical perception is a basic perceptual mechanism, as has been suggested, the results of these studies may also help us to develop a more general theory of knowledge acquisition and structure as well.

Acknowledgements

This work was supported by NIH grant #5-R01-HD23922. We would like to thank Andrew F. Hayes and Daniel J. Simons for statistical advice, Daniel T. Levin for help with software and images, and Barbara L. Finlay, David J. Field, James E. Cutting and Grant Gutheil for comments on drafts. Portions of this work were presented at the Psychonomic Society 34th annual meeting in Washington, DC on November 7, 1993 and at the 3rd Neural Computation and Psychology Workshop at the University of Stirling, Scotland, on August 31, 1994.

References

Barrera, M., & Maurer, D. (1981). The perception of facial expression by 3-month-olds. *Child Development, 52*, 203–206.

Bornstein, M.H. (1987). Perceptual categories in vision and audition. In S. Harnad (Ed..), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Bornstein, M.H., & Korda, N.O. (1984). Discrimination and matching within and between hues measured by reaction times: some implications for categorical perception and levels of information processing. *Psychological Research, 46*, 207–222.

Bruce, V. (1994). Stability from variation: the case of face recognition. *Quarterly Journal of Experimental Psychology, 47*, 5–28.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition, 38*, 109–144.

Bruce, V., & Young, A.W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–327.

DeValois, R.L., & DeValois, K.K. (1975). Neural coding of color. In E.C. Carterette & M.P. Friedman (Eds.), *Handbook of perception* (Vol. 5). New York: Academic Press.

Diamond, R., & Carey, S. (1977). Development changes in the representation of faces. *Journal of Experimental Child Psychology, 23*, 1–22.

Ekman, P., Freisen, W.V., & Ellsworth, P. (1982). What are the similarities and differences in facial behavior across cultures? In P. Ekman (Ed.), *Emotions in the human face*. New York: Cambridge University Press.

Etcoff, N.L., & Magee, J.J. (1992). Categorical perception of facial expressions. *Cognition, 44*, 227–240.

Field, T., Woodson, R., Cohen, D., Greenberg, R., Garcia, R., & Collins, K. (1983). Discrimination and imitation of facial expressions by term and preterm neonates. *Infant Behavior and Development, 6*, 485–489.

Field, T., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science, 218*, 179–181.

Harnad, S. (1987). Introduction: psychological and cognitive aspects of categorical perception: a critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Hasselmo, M.E., Rolls, E.E., & Baylis, G.C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research*, *32*, 203–218.

Kuchuk, A., Vibbert, M., & Bornstein, M.H. (1986). The perception of smiling and its experiential correlates in 3-month-old infants. *Child Development*, *57*, 1054–1061.

Kuhl, P.K., & Miller, J.D. (1975). Speech perception by the chinchilla: voiced–voiceless distinction in the alveolar plosive consonants. *Science*, *190*, 69–72.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.

Liberman, A.M., Harris, K.S., Hoffman, H.S., & Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358–368.

Light, L.L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 212–228.

O'Toole, A.J., Abdi, H., Deffenbacher, K.A., & Valentin, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America A*, *10*, 1–7.

Pastore, R.E., Ahroon, W.A., Baffuto, K.J., Friedman, C., Puleo, J.S., & Fink, E.A. (1977). Common-factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 686–696.

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves, M.A. (1984). Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, *3*, 197–208.

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

Rosen, S., & Howell, P. (1987). Auditory, articulatory, and learning explanatory explanations of categorical perception in speech. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition*. New York: Cambridge University Press.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*, 71–86.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race on face recognition. *Quarterly Journal of Experimental Psychology*, *43A*, 161–204.

Valentine, T., & Bruce, V. (1986). Recognizing familiar faces: the role of distinctiveness and Familiarity. *Canadian Journal of Psychology*, *40*, 300–305.

Williams, L. (1977). The perception of stop consonant voicing by Spanish–English bilinguals. *Perception and Psychophysics*, *21*, 289–297.

Wolberg, G. (1990). *Digital image warping*. Los Alamitos, CA: IEEE Computer Science Society Press.

Young, A.W., & Bruce, V. (1991). Perceptual categories and the computation of "grandmother". *European Journal of Cognitive Psychology*, *3*, 5–49.