VISUAL COGNITION, 1996, *3* (2), 81–117

Categorical Perception of Morphed Facial Expressions

Andrew J. Calder
University of Durham, Durham, UK
MRC Applied Psychology Unit, Cambridge, UK

Andrew W. Young
MRC Applied Psychology Unit, Cambridge, UK

David I. Perrett
University of St Andrews, St Andrews, Fife, UK

Nancy L. Etcoff
Harvard Medical School. Massachusetts General Hospital, Boston, Massachusetts, USA

Duncan Rowland
University of St Andrews, St Andrews, Fife, UK

Using computer-generated line-drawings, Etcoff and Magee (1992) found evidence of categorical perception of facial expressions. We report four experiments that replicated and extended Etcoff and Magee's findings with photographic-quality stimuli. Experiments 1 and 2 measured *identification* of the individual stimuli falling along particular expression continua (e.g. from happiness to sadness) and *discrimination* of these stimuli with an ABX task in which stimuli A, B, and X were presented sequentially; subjects had to decide whether X was the same as A or B. Our identification data showed that each expression continuum was perceived as two distinct sections separated by a category boundary. From these identification data we were able to predict subjects' performance in the ABX discrimination task and to demonstrate better discrimination of cross-boundary than within-category pairs; that is, two faces identified as different expressions (e.g. happy and sad) were easier to discriminate than two faces of equal physical difference identified as the same expression (e.g. both happy).

Requests for reprints should be sent to A. J. Calder, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge, CB2 2EF, UK. Email: andy.calder@mrc-apu.cam.ac.uk
This research was supported by ESRC Grant R000234003. We are grateful to Professor P. Ekman for giving us permission to use and reproduce photographs from the Ekman and Friesen (1976) Pictures of Facial Affect series.

Experiments 3 and 4 addressed two new issues arising from Etcoff and Magee's (1992) data and the results of our own Experiments 1 and 2: (1) that they might reflect artefacts inherent in the use of single continua ranging between two prototypes – for example, a range effect or an anchor effect, (2) given that the ABX procedure incorporates a short-term memory load, discrimination data obtained with this task might reflect a short-term memory rather than a perceptual phenomenon. We found no support for either of these reinterpretations and further evidence of categorical perception.

A common requirement in perception is to assign the things we see or hear to discrete categories. Assignment to categories simplifies the task of interpreting events in the external world and selecting appropriate actions. However, it can carry a hidden cost; we can be relatively insensitive to physical changes in a stimulus if they occur within a perceptual category, and more sensitive to changes of the same physical magnitude occurring across a boundary between two perceptual categories. For example, there are abrupt discontinuities in the perception of initial consonants as voice onset time is increased, and perceived colour can change suddenly across certain wavelengths. Many examples of such phenomena have been described (Harnad, 1987). Collectively, they are known as categorical perception.

The hallmark of categorical perception, then, is greater sensitivity to a physical change when it crosses the boundary between two perceptual categories than to the same change occurring within a particular category. In the case of colour, for example, psychophysical studies have confirmed that we are relatively insensitive to changes in wavelength if they occur within a region belonging to a colour category, and more sensitive to changes of the same physical magnitude occurring across the boundary between two colours (Bornstein & Korda, 1984).

Most demonstrations of categorical perception have used unidimensional stimulus changes (voice onset time in phoneme perception, wavelength in colour perception, etc.), and it is only recently that developments in computer graphics have made it possible to explore categorical perception effects with multidimensional stimuli, such as faces.

Facial expressions of emotion form an excellent candidate for the exploration of categorical perception with multi-dimensional stimuli. People are very skilled at understanding each other's facial expressions. We know that babies are very interested in faces (Johnson, Dziurawiec, Ellis, & Morton, 1991) and that they show precocious ability to respond to different facial expressions (Field, Woodson, Greenberg, & Cohen, 1982). We also know that, for tests using a fixed range of alternative choices, certain configurations of facial features resulting from specific patterns of facial muscle movements are perceived throughout the world as corresponding to particular basic emotions (Ekman, 1992, 1994). Moreover, selective changes in ability to recognize emotion from the face have been reported after brain injury; sometimes patients may remain able to recognize other social cues, such as identity, from the face even though they have problems in understanding facial emotion (Etcoff, 1984; Young, Newcombe, de Haan, Small, & Hay, 1993). PET studies have also shown differences between brain regions involved in the analysis of facial identity and expression (Sergent, Ohta, MacDonald, & Zuck, 1994).

These facts are consistent with the long evolutionary history of facial expressions of emotion (Darwin, 1872; Ekman, 1973), but we know little about the perceptual basis of how emotions are recognized. One of the fundamental issues that is still disputed concerns whether facial expressions are perceived as varying continuously along certain underlying dimensions or as belonging to qualitatively discrete categories (Ekman, 1982; Ekman, Friesen, & Ellsworth, 1972). Findings of categorical perception of facial expressions are thus of particular interest, as one might use existing theories to argue either way. For example, the idea of basic universally recognized emotions would suggest categorical perception, whereas dimensional accounts would not. Similarly, whereas the small number of universal facial expressions of emotion might lead one to expect categorical perception, the ease with which we can interpret differences in the intensity of each emotion would not.

The first investigation of categorical perception of expressions was made by Etcoff and Magee (1992), who used an algorithm devised by Brennan (1985) to create morphed images lying along the continuum between any two face drawings by adjusting the differences between the locations of corresponding points in one drawing and the other. To this end, Etcoff and Magee (1992) converted photographs from the Ekman and Friesen (1976) series of pictures of facial affect into line-drawings based on the locations of 169 landmark points, joined into 37 separate lines. The computer was then used to create several series of drawings representing equal interpolated steps between two different facial expressions posed by the same individual.

With these stimuli, Etcoff and Magee (1992) found striking evidence of categorical perception of facial expressions. They measured *identification* of the individual stimuli falling along a particular expression continuum (e.g. from happiness to sadness) and *discrimination* between pairs of these stimuli with an ABX task in which stimuli A, B, and X were presented sequentially; subjects had to decide whether X was the same as A or B. In this ABX discrimination task, Etcoff and Magee (1992) found that people were more accurate at detecting the differences between pairs of drawings that crossed a subjective category boundary (such as between a drawing seen as happy in the identification task and a drawing seen as sad) than they were at detecting equal physical differences that lay within a category (i.e. between two drawings that would be identified as happy or two drawings identified as sad).

Although clearly important, Etcoff and Magee's (1992) findings were made with line-drawings, which are inevitably low in realism because they lack pigmentation and other surface-based cues. However, the principles of Brennan's (1985) technique have been extended by Benson and Perrett (1991a, 1991b) to allow manipulation of photographic-quality images. This is achieved by specifying a large number of facial feature points, using these to divide face photographs into a mesh of triangular sections, and then deforming each of these sections in line with changes in the locations of the specified points. To create interpolated ("morphed") images depicting the continuum between two faces (say, pictures of a person with happy and sad expressions) in photographic quality, the positions of the features in one photograph are thus moved towards their positions in the other photograph, as if the image were imprinted on a rubber sheet. This procedure allows us to explore categorical perception of faces using morphed images depicting facial continua in photographic quality.

For our experiments, then, morphs of facial expressions from the Ekman and Friesen (1976) series were made using the Benson and Perrett (1991a, 1991b) algorithms; by varying the relative weights given to each of the two prototypical expressions in these blends, new images lying at equidistant points along the continuum from one of the contributory prototypes to the other were created.

In Experiments 1 and 2 three expression continua (happiness → sadness, sadness → anger, and anger → fear) were used to replicate Etcoff and Magee's (1992) findings with photographic-quality images. Etcoff and Magee (1992) had investigated the seven facial expressions from the Ekman and Friesen (1976) series (happiness, sadness, fear, anger, surprise, disgust, neutral) across eight continua (happy–sad, angry–sad, angry–afraid, angry–disgusted, happy–neutral, neutral–sad, happy–surprised, sur-prised–afraid). Their findings showed categorical perception for the six continua that did not involve a facial expression of surprise.

The continua chosen for the present Experiments 1 and 2 thus formed a subset of the continua investigated by Etcoff and Magee (1992), drawn from those they found to give clear categorical perception effects, and involving four of the seven facial expressions they used. Our results were consistent with the suggestion that the continua were perceived categorically. Experiments 3 and 4 then addressed two important issues arising from such data, which have seldom been explored in the literature on categorical perception: (1) that they might be a consequence of using single continua ranging between two prototypes—for example, a range effect or an anchor effect. Were this so, phenomena usually taken to indicate categorical perception might be attributed to the structure of the experiments themselves, rather than reflecting something that subjects brought with them to the experiment. Clearly, categorical perception is much more interesting if it reflects what is already in our heads, not what is created for experiments. (2) Given that our discrimination task incorporated a short-term memory load, the data might reflect a short-term memory rather than a perceptual phenomenon. We were able to rule out both of these alternative explanations, providing further evidence that categorical perception is inherent in the functioning of the system used to perceive facial expressions.

EXPERIMENT 1

Method

Materials

Three photographic-quality expression continua were prepared using Benson and Perrett's (1991b) algorithms, a happiness → sadness continuum with face PF from the Ekman and Friesen (1976) series, a sadness → anger continuum with face WF, and an anger → fear continuum with face EM. These continua are shown in Figure 1.

Our reasons for choosing these faces and continua for an initial experiment were that, given the substantial resources needed to render and transform images in photographic quality, it was considered prudent to draw on a subset of the wide range of continua investigated by Etcoff and Magee (1992). In addition, we selected images where the two photographs of the poser were as similar as possible in general quality (same head positioning, same lighting, etc.). Furthermore, we opted for pairs of photographs that would not show gross changes in physical features, and especially from open to closed mouths; such changes can introduce easily verbalised differences between images, which might over-ride effects due to the perception of expressions per se. Mean recognizabilities of the emotions for each face by Ekman and Friesen's (1976) normal subjects were as follows: PF happiness 100%, PF sadness 100%, WF sadness 79%, WF anger 100%, EM anger 83%, EM fear 92%. However, these recognizabilities relate to a much more difficult six-way choice between happiness, sadness, fear, anger, surprise, and disgust, not to the simpler two-way choices used in our identification task for Experiment 1.

Each continuum was prepared by blending the two prototype expressions posed by the same person (e.g. happiness and sadness) in 10 equal steps. This gave 11 face images (the 2 prototypes and 9 morphs), numbered 0–10, with 10% increments between each expression. For the happiness–sadness continuum, for example, face 0 was the happy prototype (by definition, 0% sad, 100% happy); face 1 (10% sad, 90% happy) was shifted one-tenth towards the sad prototype; face 2 (20% sad, 80% happy) was shifted two-tenths towards the sad prototype, and so on, until the sad prototype was reached at face 10 (100% sad, 0% happy).

The preparation of each continuum was done in the same way. Preparation of the happiness → sadness continuum is now described in detail, to illustrate the process. The procedure involved three stages.

Stage 1: Delineation. Each photograph was frame-grabbed at a resolution of 512 × 720 pixels with 256 grey levels, and 186 points were positioned manually onto the photograph of PF's face with a happy expression from the Ekman and Friesen (1976) series. The locations of these points were specified in terms of anatomical landmarks, with each facial feature represented by a set number of points (e.g. the mouth was represented by 22 points and each eyebrow by 8

FIG. 1. The three continua of morphed face images used in Experiment 1: PF: happiness → sadness, WF: sadness → anger, and EM: anger → fear.

points). The points were then joined to produce a delineated representation comprising 50 feature contours. Exactly the same method was applied to the photograph of the same model posing sadness. Hence, across the two prototype expressions (happiness and sadness) there was conformity with respect to the anatomical positioning of the 186 points, but the exact spatial position of points might vary; for example, the corners of the mouth were turned up in the happy face and down in the sad face, but the shape of the hairline was the same in both.

Stage 2: Shape Interpolation. A continuum of face shapes was generated between the two delineated prototype face shapes (in our example, PF happy and PF sad). This was achieved by taking the delineation data for the two prototype images and calculating the vector difference for each landmark. For example, consider the point describing the tip of the nose; this has a location on the PF happy face of (x1,y1) and a location on the PF sad face of (x2,y2). Equations describing the vector from (x1,y1) to (x2,y2) were used to obtain positions for the point at the tip of the nose, which moved in a straight line from the location of that point in the PF happy face (x1,y1) to the location of that point in PF's sad face (x2,y2). This process was repeated for each of the 186 feature points, to generate the 9 face shapes that would interpolate linearly between the two original face shapes. Hence, to produce a face halfway between happy and sad (face 5 in our 0–10 continuum), the points in the happy face were moved 50% closer to the same points in the sad face. Likewise, face images 1, 2, 3, etc. in the continuum were created by moving the points in the happy face the corresponding distance (10%, 20%, and 30% closer to sad, respectively).

Stage 3: Producing a Continuous-tone Image. The final stage created a continuous-tone (photographic-quality) image for each of these interpolated face shapes. This was achieved by taking both of the prototype faces and "warping" or "stretching" them (as if they were printed on a rubber sheet) to the new shape, so that all points representing the same feature were aligned across images. The two faces, now with the same intermediary face shape, were then blended with the appropriate weight. For example, in the 90% happy, 10% sad morph (face 1 in the happiness → sadness continuum), the pixel intensities in each section of the image were arrived at by deforming the happy face 10% towards the sad prototype, and the sad face 90% towards the happy prototype, and then blending the grey levels in these two contributory images in the ratio of nine parts from the happy prototype to one part from the sad prototype.

Subjects

Thirty six members of the MRC Applied Psychology Unit subject panel participated in the experiment. All were between the ages of 27 and 39 years and had normal or corrected-to-normal vision. They were paid for participating.

Design and Procedure

The subjects were divided into three groups of 12. Each group completed an ABX discrimination task and a forced-choice identification task with the same continuum (one of PF: happiness → sadness, WF: sadness → anger, and EM: anger → fear). The ABX discrimination task assessed the subjects' ability to discriminate between faces in an expression continuum, and the identification task determined how they labelled the same faces' expressions.

There was a small variation in picture quality between some of the endpoint faces (0 and 10) and the morphed faces (1 → 9) in the three continua. We were concerned that subjects might be able to use this difference to discriminate between pairs of faces that included either of the endpoints. Therefore, in the ABX discrimination task the stimulus set comprised faces 1 → 9 only. All subjects performed the discrimination task first, followed by the identification task.

ABX Discrimination Task. On each trial a central fixation cross was presented for 250 msec, followed by a blank interval of 250 msec, and then three successive face images from one of the continua. The first and second faces (A and B) were displayed for 750 msec each, and the third face (X) for 1 sec; they were separated by blank intervals of 1 sec. These times were chosen to emulate those used by Etcoff and Magee (1992). Faces A and B were always different, and on half of the trials face X was identical to face A and on the other half to face B. Subjects were asked to make a button press response to indicate whether X matched A or B.

Faces A and B differed by either 2 steps (e.g. morphs 1 and 3), 3 steps (e.g. morphs 1 and 4), or 4 steps (e.g. morphs 1 and 5) along the appropriate continuum. For each continuum, all 18 possible face pairs drawn from morphed images 1 → 9 (7 two-step pairs, 6 three-step, and 5 four-step) were presented eight times (twice in each of the four orders: ABA, ABB, BAA, BAB), giving a total of 144 trials per continuum.

To acquaint subjects with the ABX procedure, the experiment began with 24 practice trials, selected at random. So that subjects could assess their performance in this practice phase, an error was signalled by a bleep from the computer; there was no feedback in the experimental trials.

Identification Task. After completing the ABX discrimination task, subjects were asked to identify the expression on each face from the same continuum. Each identification trial began with a 250-msec presentation of a central fixation cross, followed by a blank interval of 250 msec, and then a face from the continuum seen in the ABX discrimination task. Subjects were asked to identify the face's expression by making a two-choice button-press response. The two choices offered corresponded to the expressions posed in the prototype faces; for example, for the happiness → sadness continuum the subjects had to identify the expressions as either "happy" or "sad". Subjects were shown the 9 morphed face images (faces 1 → 9) from the ABX discrimination task and the two prototypes (faces 0 and 10). Each face was presented 8 times in random order, giving a total of 88 trials per continuum.

The stimuli were presented on a 256 grey-scale 16-in Macintosh colour monitor using the Psyscope™ 1.04 (ABX discrimination task) and Superlab™ 1.68 (identification task) presentation packages. In both the discrimination and identification tasks each face subtended a horizontal visual angle of 3.6° and a vertical visual angle of 5.2°.

Results

Results are summarized in Figure 2. The upper graph in Figure 2 shows a summary of the subjects' identification data, and the bottom graph is a summary of their discrimination data in the two-step ABX discrimination trials. Data from the three-step and four-step trials in the ABX discrimination task are not reported because the subjects' performance in these two conditions approached ceiling.

The top left-hand section of Figure 2 shows the mean percentage of times subjects identified the images from the happiness → sadness continuum as happy (filled squares) or sad (open squares). Face images 0, 1, 2, 3, and 4 were mainly identified as happy. In contrast, faces 6, 7, 8, 9, and 10 were consistently categorised as sad. Face 5 was identified as happy and sad with approximately equal frequency. Hence, the identification data from the happiness → sadness continuum showed an abrupt category shift, with the boundary lying between faces 4 and 6. The sadness → anger and anger → fear continua (top middle and top right-hand sections of Figure 2, respectively) showed similar patterns, with the category boundaries lying between images 3 and 5 in each.

To analyze the data from the discrimination and identification tasks, we predicted subjects' performance in the ABX discrimination task from their identification data and compared these predictions to the observed ABX discrimination results (Liberman, Harris, Hoffman, & Griffith, 1957; Pisoni & Lazarus, 1974).

The technique of deriving a predicted discrimination curve from identification data and relating it to observed discrimination performance is widely used in the categorical perception literature, but the formula we adopted was novel. To calculate the subjects' predicted performance, we based our formula on the view that this would depend on two contributory factors: (1) their ability to make use of the constant physical differences between each pair of images, regardless of their expressions, and (2) any additional contribution from categorical perception of the facial expressions. As an estimate of the first factor (ability to make use of the physical differences between each pair), we used the mean of the discriminabilities for the ABX pairs at the ends of each continuum (i.e. faces 1,

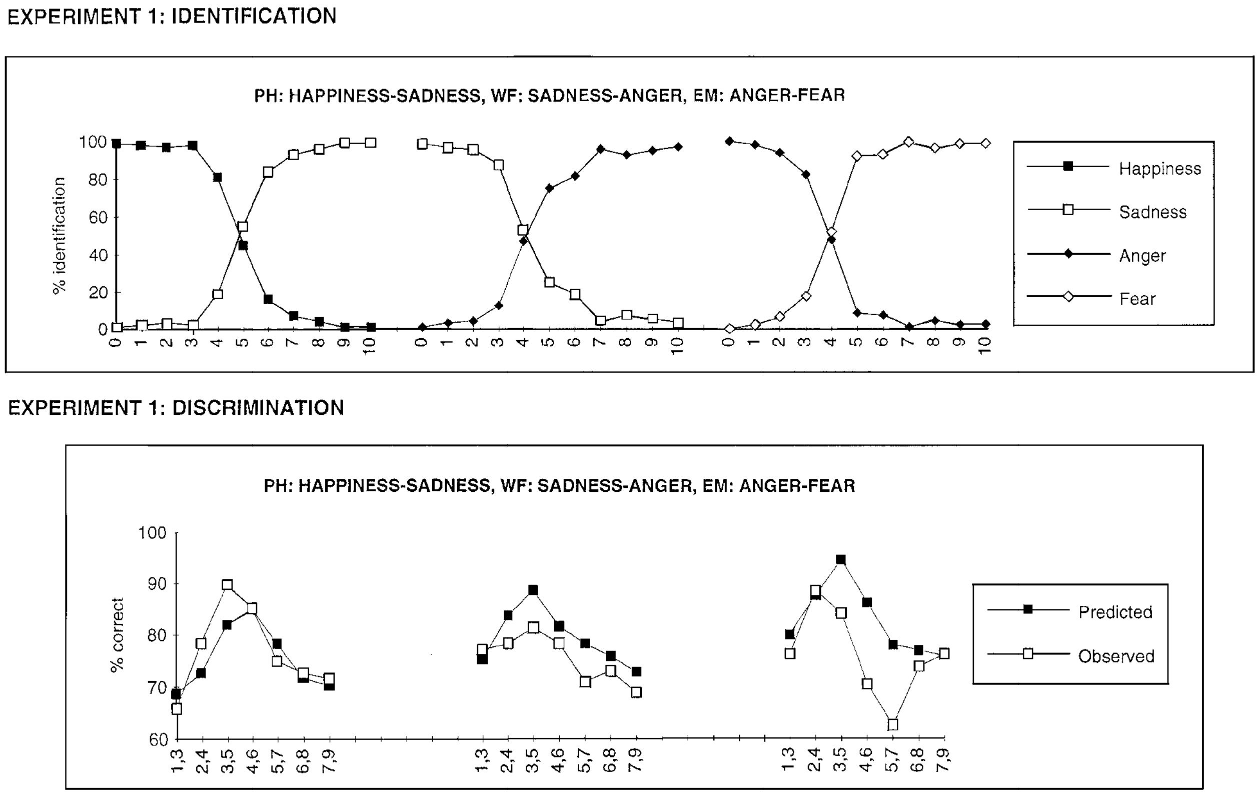


FIG. 2. Identification data (top) and ABX discrimination data (bottom) for the three expression continua used in Experiment 1: happiness → sadness (left), sadness → anger (middle), and anger → fear (right). The identification data show the frequency (mean percentage) with which each face in a continuum was identified as a given expression. The discrimination data show subjects' predicted and observed performance with the two-step face pairs. The measure of performance was the mean percentage of times each face pair was discriminated correctly.

3 and faces 7, 9), where any contribution of categorical perception of expressions would be minimal as these stimuli were assigned to the appropriate category nearly 100% of the time in the identification task. To estimate the second factor (contribution from categorical perception of the facial expressions), we used the differences between the rates at which the members of each pair were assigned to a particular expression category in the identification task. These two factors were then combined by adding 0.25 of the identification difference for the relevant pair to the estimate of the overall physical difference, to yield the predicted performance curve for each continuum. The intention of multiplying the identification difference by this constant was simply to bring the range of variability of predicted and observed curves into line with each other. This makes it easier to see the fit between the curves on a graph but does not affect the correlation. Of course, this particular way of combining the estimates of the two factors is arbitrary, because we have no *a priori* way of knowing which will have the greater influence. However, the 1:4 ratio, which was derived by trial and error for Experiment 1, works reasonably well not only for Experiment 1, but for all of the experiments reported here.

Figure 2 shows subjects' mean percent predicted performance (filled squares) and their observed performance (open squares) with the PF: happiness → sadness, WF: sadness → anger, and EM: anger → fear morphs. The fit between the predicted and observed performance was assessed by looking at the correlation of these for each continuum. The significance of these correlations was then determined by converting each to a t-value, using the procedure described by McNemar (1962). This showed significant correlations of predicted and observed performance for the happiness → sadness, $r = 0.88$, $t(5) = 4.06$, $p < 0.01$, and sadness → anger, $r = 0.82$, $t(5) = 3.23$, $p < 0.05$ continua, but the correlation for the anger → fear continuum did not reach statistical significance, $r = 0.60$, $t(5) = 1.66$, $p > 0.1$. When data for the predicted and observed performances across all three continua were considered together (i.e. as if they had come from a single continuum across Figure 2, running happiness → sadness → anger → fear), a significant correlation was again observed, $r = 0.62$, $t(19) = 3.47$, $p < 0.01$.

Because the formula used to derive predicted ABX performance above was novel and inevitably contained an arbitrary element, as we had no independent way of estimating the relative influence of its two components, we sought to confirm whether these findings would also hold with the widely used formula of Liberman et al. (1957). The main difference between these formulas is that Liberman et al. (1957) predicted discrimination on the assumption that categorical perception will be the only contributory factor, and they did not include any estimate of differences due to other factors.

In practice, using the Liberman et al. (1957) formula to predict discrimination from our identification data produced exactly the same pattern of results, with significant correlations of predicted and observed discrimination performance for the happiness → sadness, $r = 0.82$, $t(5) = 3.20$, $p < 0.05$, and sadness → anger, $r = 0.77$, $t(5) = 2.71$, $p < 0.05$ continua, whereas the correlation for the anger → fear continuum did not reach statistical significance, $r = 0.55$, $t(5) = 1.47$, $p > 0.1$. When data for the predicted and observed performances across all three continua were considered together (as if they had come from a single continuum running happiness → sadness → anger → fear), a significant correlation was again observed, $r = 0.67$, $t(19) = 3.97$, $p < 0.001$.

These correlational analyses examined the relation between predicted and observed performance for discrimination of each pair of images. However, there are other possible criteria for categorical perception. As we have already indicated, the most widely used criterion is that discrimination should show greater sensitivity to a physical change when it crosses the boundary between two perceptual categories than to the same change occurring within a particular category. We therefore compared observed ABX discriminability for the pairs closest to each boundary (faces 4,6 for the happiness → sadness continuum, and faces 3,5 for the sadness → anger and anger → fear) to the average observed discriminability for pairs lying well within each category (pairs 1,3 and 7,9 in each continuum). A two factor ANOVA examined the effects of type of pair (within-category or cross-boundary; repeated measure) and stimulus continuum (happiness → sadness, sadness → anger, or anger → fear). This showed that discrimination was better for cross-boundary pairs: type of pair, $F(1, 33) = 18.09$, $p < 0.001$. There was no main effect of stimulus continuum, $F<1$, and no type of pair × continuum interaction, $F(2, 33) = 1.06$, $p > 0.1$, indicating that this enhanced cross-boundary discrimination held across all three continua.

Exactly the same pattern was revealed by an analysis that compared observed ABX discriminability for the pairs closest to each boundary (faces 4, 6 for the happiness → sadness continuum, and faces 3, 5 for sadness → anger and anger → fear) to the average observed discriminability for all other pairs in the relevant continuum. As before, discrimination was better for cross-boundary pairs: type of pair, $F(1, 33) = 16.95$, $p < 0.001$. There was no main effect of stimulus continuum $F<1$, and no type of pair × continuum interaction, $F<1$, showing again that this enhanced cross-boundary discrimination held across all three continua.

Discussion

The results of Experiment 1 were consistent with Etcoff and Magee's (1992) finding that facial expressions are perceived categorically. The identification data from all three continua showed the characteristic step function associated with categorical perception, with reliable assignment of the morphs at each end of the continuum to the nearest prototype category and a rapid shift in the middle of the continuum. From these identification data we were able to predict subjects' ABX discrimination performance across all three expression continua taken together, and the correlations were significant for the individual happiness →

sadness and sadness → anger continua. However, the correlation of observed and predicted discrimination performance for the anger → fear continuum did not reach statistical significance. This was perhaps surprising because Etcoff and Magee (1992) found categorical perception of all three types of continua used in Experiment 1. Further analyses of ABX discriminability for within-category and cross-boundary pairs showed that discrimination was better for cross-boundary pairs in all continua.

Overall, then, the results of Experiment 1 were encouraging. Note, however, that although the key categorical perception effect of enhanced discrimination of pairs of stimuli falling across rather than well within category boundaries was found, inspection of Figure 2 shows that peak discriminabilities did not always align perfectly with category boundaries, and that within-category discriminabilities were not constant. Such observations held for most of the experiments reported here; we therefore discuss their implications in the General Discussion.

Having established initial findings consistent with the categorical perception hypothesis, it was considered important to demonstrate that these results could be generalized across stimulus materials—that is, that categorical perception of facial expressions is not confined to the particular continua used in Experiment 1. In Experiment 2, we therefore used a repeated-measures design to examine subjects' performance with four examples of the same three types of expression continua.

EXPERIMENT 2

Method

Materials

Four happiness → sadness, four sadness → anger and four anger → fear continua were prepared from pictures of four female models (MF, MO, NR, SW; Ekman & Friesen, 1976). Each continuum was prepared by blending two pictures of the same person posing with two different expressions, using the method outlined for Experiment 1; hence, there were a total of 12 continua, each containing 11 equally spaced face images.

Faces were chosen so that the four photographs of each poser (happy, sad, angry, and afraid) were as similar as possible in general quality (same head positioning, same lighting, etc.). However, in order to get four models who would meet these criteria, we removed the restriction used in Experiment 1 of avoiding changes from open to closed mouths; this restriction had to be removed for the happy expressions, all of which showed open mouths with teeth visible. Mean recognizabilities of the emotions across the four posers for Ekman and Friesen's (1976) normal subjects were as follows; happiness 98%, sadness 91%, anger 100%, fear 85%. Note again, though, that these relate to a six-way choice between happiness, sadness, fear, anger, surprise and disgust, not the simpler two-way choices used in our identification task for Experiment 2.

The stimuli were arranged by type of continuum into three sets; thus, one set contained the four happiness → sadness continua, another set the four sadness → anger continua, and the third set the four anger → fear continua.

Subjects

Twelve post-graduate and undergraduate students at the University of Durham took part in the experiment. All were between the ages of 23 and 37 years and had normal or corrected-to-normal vision. They were paid for participating.

Design and Procedure

Each subject performed an ABX discrimination task and an identification task with all 12 continua. Trials were blocked by continua; hence, there were 12 blocks of discrimination and 12 blocks of identification trials.

Individual subjects were tested in three sessions separated by approximately 24 hours. In one session a subject completed four blocks of ABX discrimination trials and four blocks of identification trials with one set of expression continua—for example, the four happiness → sadness continua. Subjects always performed the four blocks of discrimination trials first in each session. The order of viewing the continua in the discrimination and identification tasks was counterbalanced across subjects.

ABX Discrimination Task. In Experiment 1, subjects performed the ABX discrimination task with face images 1 → 9 because of concern that some of the endpoint faces (0 and 10) were of slightly different quality (the morphing process can make some outlines a little less sharp). This difference was negligible in the Experiment 2 stimuli, so the 11 faces (0 → 10) from each of the 12 continua were presented in both the discrimination and identification tasks.

The discrimination task used the ABX procedure as described for Experiment 1. Each trial began with a central fixation cross, followed by three face images (face A then face B and then face X); subjects were asked to decide whether face X matched face A or face B. Presentation times and stimulus sizes were as for Experiment 1.

As in Experiment 1, faces A and B differed by either 2, 3, or 4 steps along the 11-level continuum. For each continuum, (e.g. MF:happiness → sadness, NR:anger → fear, etc.) all 24 possible pairs (9 two-step, 8 three-step, and 7 four-step) were presented twice in random order, giving a total of 48 trials per continuum. Trials were counterbalanced so that all four possible orders of ABX stimuli (ABA, ABB, BAA, BAB) occurred with equal frequency in the two-step, three-step, and four-step trials.

Each session began with a block of 24 ABX practice trials containing faces from a continuum that was not seen in the experimental trials. In addition, each block of 48 experimental trials began with 4 dummy trials selected at random.

Identification Task. The general procedure for the identification task was identical to that used in Experiment 1. However, whereas in Experiment 1 each face in a continuum was presented 8 times (88 trials per continuum), 4 presentations were used in Experiment 2 (44 trials per continuum). The order of presentation of each face in a given continuum was random.

Results

Data from the three-step and four-step discrimination trials are not reported because subjects' performance in these two conditions approached ceiling.

The upper graph in Figure 3 shows data from the identification task. Each section of the graph shows data pooled across four continua—for example, the happiness → sadness graph (Figure 3, top left) shows the mean percentage of times the faces in the four happiness → sadness continua were identified as "happy" or "sad".

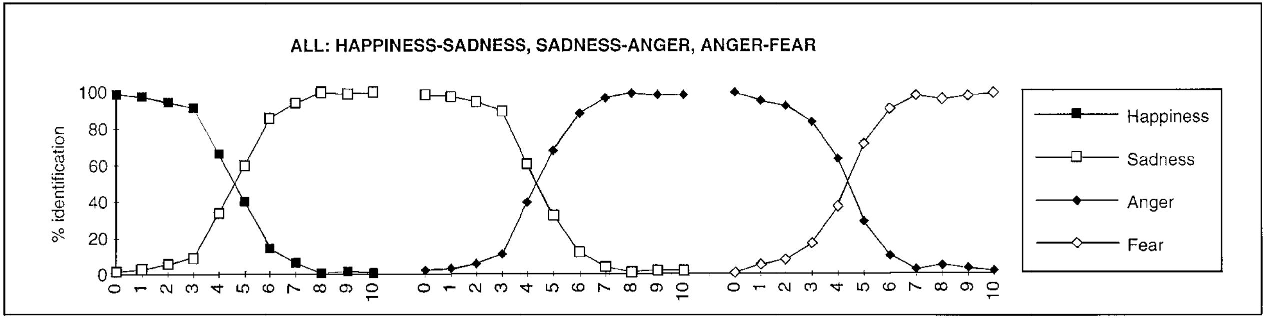
The three sets of continua again showed category boundaries that occurred in the same places as for the expression continua used in Experiment 1: happiness → sadness 4,6, sadness → anger 3,5, and anger → fear 3,5. Again we calculated subjects' predicted discrimination performance with each continuum using the method described for Experiment 1. The lower graph in Figure 3 shows subjects' predicted (filled squares) and mean percentage correct observed performance (open squares) for the ABX discrimination task.

As before, the relationship between the predicted and observed performances was assessed by correlation. Statistically significant correlations were found for the continua placed end to end—happiness → sadness → anger → fear, $r = 0.75$, $t(25) = 5.61$, $p < 0.001$—and individually for the sadness → anger, $r = 0.89$, $t(7) = 5.06$, $p < 0.01$, and anger → fear, $r = 0.76$, $t(7) = 3.12$, $p < 0.02$ continua. However, there was no correlation of predicted and observed discrimination performance across the happiness → sadness continuum, $r = 0.01$, $t(7) = 0.04$, $p > 0.1$.

An additional correlational analysis using the Liberman et al. (1957) formula to predict discrimination from our identification data produced the same pattern of results, with significant correlations of predicted and observed discrimination performance across the continua placed end to end—happiness → sadness → anger → fear, $r = 0.45$, $t(25) = 2.53$, $p < 0.05$—for the sadness → anger continuum, $r = 0.83$, $t(7) = 3.96$, $p < 0.01$, and the anger → fear continuum, $r = 0.72$, $t(7) = 2.73$, $p < 0.05$, but not for the happiness → sadness continuum $r = 0.01$, $t(7) = 0.04$, $p > 0.1$.

As for Experiment 1, we also compared observed ABX discriminability for the pairs closest to each boundary (faces 4,6 for the happiness → sadness continuum, and faces 3,5 for sadness → anger and anger → fear) to the average observed discriminability for pairs lying well within each category (pairs 1,3 and 7,9 in each continuum). A two-factor ANOVA examined the effects of type of pair (within-category or cross-boundary; repeated measure) and stimulus

EXPERIMENT 2: IDENTIFICATION



EXPERIMENT 2: DISCRIMINATION

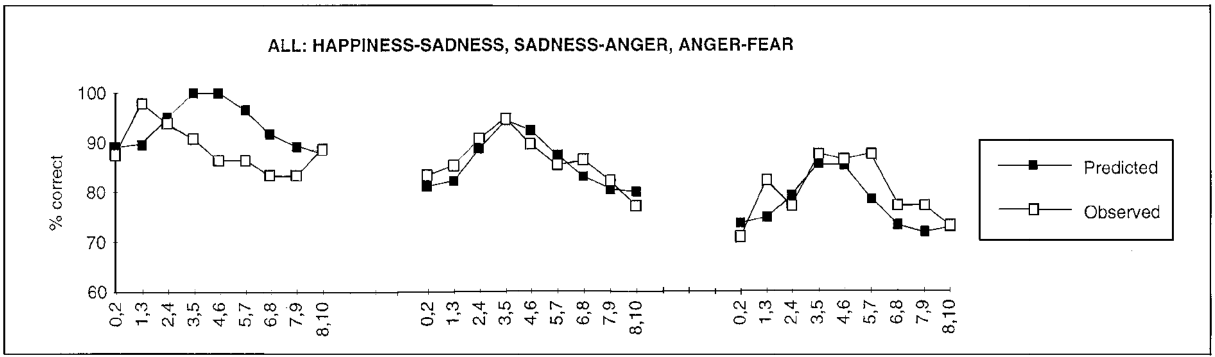


FIG. 3. Identification data (top) and ABX discrimination data (bottom) for the three types of expression continua used in Experiment 2: happiness → sadness (left), sadness → anger (middle), and anger → fear (right). The identification data show the frequency (mean percentage) with which faces in the same type of continua were identified as a given expression. The discrimination data show subjects' predicted and observed performance with the two-step face pairs. The measure of performance was the mean percentage of times each face pair was discriminated correctly.

continuum (happiness → sadness, sadness → anger, or anger → fear; repeated measure). In this analysis, the main effect of type of pair did not reach statistical significance, $F(1, 11) = 3.22$, $0.1 > p > 0.05$, and there was no main effect of stimulus continuum, $F < 1$. However, a significant type of pair × continuum interaction was found, $F(2, 22) = 4.34$, $p < 0.05$. Post hoc Tukey tests ($\alpha = 0.05$) showed that enhanced cross-boundary over within-boundary discrimination held for the sadness → anger and anger → fear continua, but not for happiness → sadness. This was exactly the pattern found by correlating observed and predicted performance.

The same pattern was also revealed by an analysis that compared observed ABX discriminability for the pairs closest to each boundary (faces 4,6 for the happiness → sadness continuum, and faces 3,5 for sadness → anger and anger → fear) to the average observed discriminability for all other pairs lying within the relevant continuum. This analysis showed a significant main effect of type of pair, $F(1, 11) = 5.81$, $p < 0.05$, no main effect of stimulus continuum, $F(2, 22) = 2.29$, $p > 0.1$, and a type of pair × continuum interaction, $F(2, 22) = 4.10$, $p = 0.05$, in which enhanced cross-boundary discrimination held for the sadness → anger and anger → fear continua, but not for the happiness → sadness.

Discussion

There was a good fit between the predicted and observed discrimination values across all three types of expression continuum considered together, and individually for the sadness → anger and anger → fear continua, but not for the happiness → sadness continua. Further analyses of observed ABX discriminability for within-category and cross-boundary pairs showed the same pattern of enhanced cross-boundary discrimination for the sadness → anger and anger → fear continua, but not for happiness → sadness. Hence in Experiment 2 we have consistent evidence of categorical perception for sadness → anger and anger → fear continua, but not for happiness → sadness continua.

The failure to find evidence of categorical perception for the happiness → sadness continua used in Experiment 2, either in the form of a good fit between predicted and observed discrimination performance or of enhanced cross-boundary discriminability, probably reflects a problem with the stimuli. In all four happy faces the mouths were open, revealing the teeth, whereas in the four sad faces they were closed. Consequently, when these two expressions were blended, the teeth in the happy face changed from white to grey in distinct stages before blending into the sad face. From the discrimination data it appears that subjects used these gradations in teeth colour to help differentiate between the face images. This is reflected in the sudden increase in performance with face pair 1,3, followed by a gradual decline as the teeth cues become less visible (Figure 3, bottom left). Recall that Experiment 1 found evidence of categorical perception of a happiness → sadness continuum, so in principle it is not impossible to obtain that result. However, the mouth was closed in both the happy and sad expressions used in Experiment 1. Without this precaution, the contribution to ABX discrimination from categorical perception of the morphed facial expressions is swamped by the more easily used non-expressional cues resulting from the readily perceived physical change in the colour of the teeth.

With this caveat concerning the results for the happiness → sadness continua in Experiment 2, Experiments 1 and 2 taken together provide evidence of categorical perception for all three types of expression continuum used (happiness → sadness, sadness → anger, and anger → fear). These results are consistent with Etcoff and Magee's (1992) findings, replicating their results with more naturalistic images.

EXPERIMENT 3

In each block of trials, Experiments 1 and 2 used single continua ranging between two prototype expressions. A problem with this sort of design is that during the course of the experiment subjects can readily learn the prototypes and the range of stimuli in-between. In addition, the faces at the ends of each continuum occur less frequently in the ABX discrimination task; for example, in the two-step trials face 0 is seen in the 0,2 condition alone, whereas face 2 occurs in both the 0,2 and 2,4 conditions. The effect of such factors might be to bias subjects' correct discrimination responses towards the middle of each continuum (Poulton, 1975).

These potential artefacts apply as much to Etcoff and Magee's (1992) findings as to our own. Thus, there is at least a small possibility that categorical perception data for facial expressions can be accounted for by these sorts of effects. In order to exclude such explanations, Experiment 3 used morphs of a single face with three different expressions to create a continuum that ranged from fear → happiness → anger → fear. Hence, the continuum had no fixed endpoints, and all morphs could be used equally often in the ABX discrimination task.

Method

Materials

All possible continua were prepared using the method described for Experiment 1 from three pictures of the same male face (JJ; Ekman & Friesen, 1976) posing the expressions fear, happiness, and anger. This gave a total of three continua (fear → happiness, happiness → anger, and anger → fear), each containing 11 faces; note that there are only three continua, because happiness → fear is the same as fear → happiness, etc. We then removed the prototype expressions (fear, happiness, and anger—faces 0 and 10 on the 0–10 numbering

FIG. 4. Morphed face images used in Experiments 3 and 4: fear → happiness, happiness → anger, anger → fear.

scheme used in Experiments 1 and 2) and every second morphed face (2, 4, 6, and 8) from the three continua. This left five morphed faces (1, 3, 5, 7, and 9) in each continuum, with a total of 15 morphed faces in all. These stimuli are shown in Figure 4.

When these 15 morphed faces were arranged end to end, the effect was to create a single continuum ranging from fear through to happiness through to anger and back to fear. All 15 morphed faces were seen in the ABX discrimination and identification tasks.

JJ's face was chosen to create this continuum because the images were of consistent general quality (same head positioning, same lighting, etc.), and did not include any changes from open to closed mouth. Mean recognizabilities of each emotion by Ekman and Friesen's (1976) normal subjects were as follows: happiness 100%, anger 76%, fear 96%. However, these relate to a six-way choice between happiness, sadness, fear, anger, surprise, and disgust, not to the simpler three-way choice used in our identification task for Experiment 3.

Subjects

Twelve members of the MRC Applied Psychology Unit subject panel participated in the experiment. All were between the ages of 23 and 46 years and had normal or corrected-to-normal vision. They were paid for participating.

Design and Procedure

Subjects performed the ABX discrimination task first; the general procedure for this task was the same as that described in Experiments 1 and 2. Faces A and B were always 1 step apart in the 15 morphs used (e.g. faces 1 and 3; this is equivalent to what would have been a two-step difference in the numbering scheme used in Experiments 1 and 2). All 15 of the one-step pairs lying along the entire continuum were presented 8 times (twice in each of the four possible orders: ABA, ABB, BAA, and BAB), giving a total of 120 trials, which were arranged into a random order. Because the continuum had no fixed endpoints, the 15 faces occurred with equal frequency.

The experiment began with a set of 24 practice trials selected at random from the block of 120 ABX discrimination trials. Errors in the practice trials were signalled with a bleep from the computer; there was no feedback in the experimental trials.

Having completed the discrimination task, subjects performed an identification task with the same continuum. The presentation procedure was almost identical to that described for Experiments 1 and 2. However, because the continuum was prepared from three expressions (fear, happiness, and anger) rather than two, the subjects were asked to identify the facial expressions with one of three labels: "afraid", "happy", and "angry". Each of the 15 morphed faces was presented 8 times in random order.

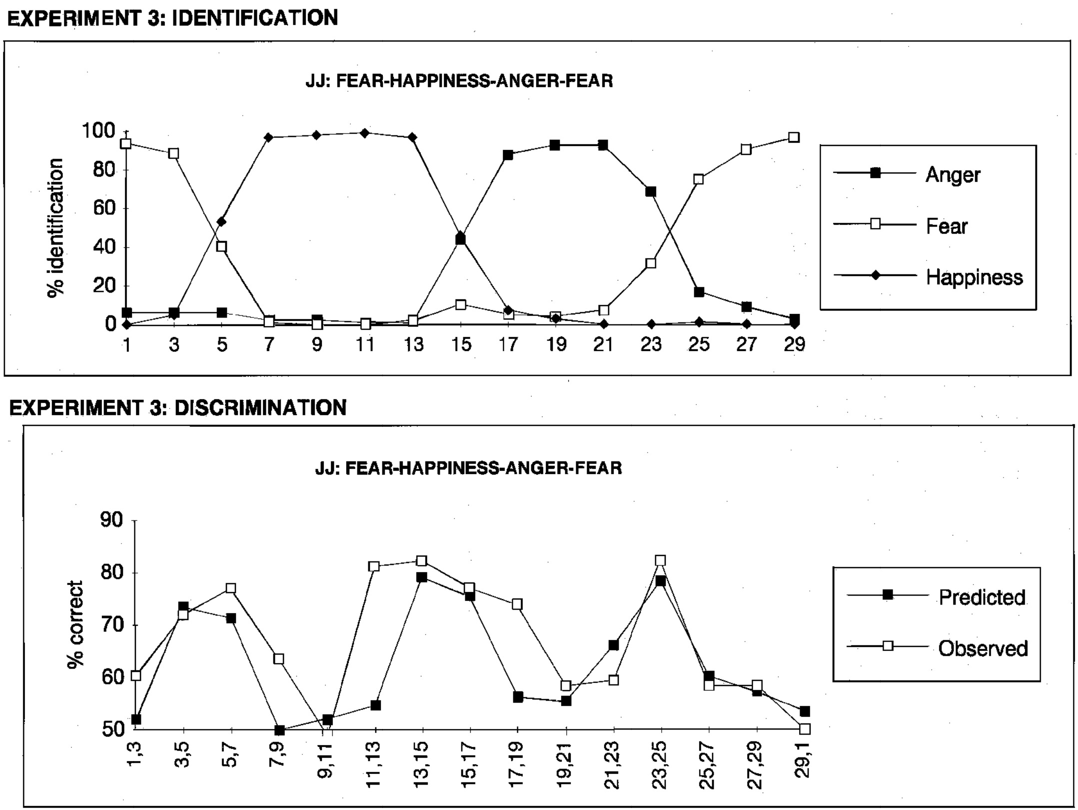


FIG. 5. Identification data (top) and ABX discrimination data (bottom) for the continuum used in Experiment 3: fear → happiness → anger → fear. The identification data show the frequency (mean percentage) with which each face in the continuum was identified as fear, happiness, or anger. The discrimination data show subjects' predicted and observed performance given as the mean percentage of times each face pair was discriminated correctly.

Results

The upper graph in Figure 5 shows the mean percentage of identification responses for the fear → happiness → anger → fear continuum. For convenience, in Figure 5 the morphed faces are numbered as if they formed an ascending series running fear → happiness → anger → fear (1, 3, 5, 7, 9 for fear → happiness; 11, 13, 15, 17, 19 for happiness → anger; 21, 23, 25, 27, 29 for anger → fear). In reality, the start of this endless continuum is arbitrarily determined.

The identification data indicated that subjects perceived the continuum as falling into three distinct sections separated by category boundaries: a fear section (morphs 25, 27, 29, 1, 3), a happiness section (morphs 7, 9, 11, 13), and an anger section (morphs 17, 19, 21, 23). The remaining two morphs (5 and 15) fell on the fear → happiness and happiness → anger boundaries respectively; there is not an equivalent face in the anger → fear continuum because the category shift from anger to fear fell between morphs 23 and 25. Note that the ambiguous morphs 5 and 15 were principally identified as the expressions in the immediately preceding and following sections of the continuum, and *not* as the third expression, which did not contribute to that particular range. For example, face 5 from the fear → happiness range was principally identified as “afraid” or “happy”, and not as “angry”. This implies that each boundary was perceived as part of a continuum ranging between two distinct expressions, rather than as a region of general uncertainty where subjects were guessing.

From the identification data we calculated subjects’ predicted performance in the ABX discrimination task using the basic method outlined for Experiment 1. Recall that the formula used to calculate subjects’ predicted performance in Experiments 1 and 2 incorporated two components: (1) an estimate of subjects’ ability to make use of the constant physical differences between each pair of images, regardless of their expressions, and (2) an estimate of any additional contribution from categorical perception of the facial expressions. For the first factor (ability to make use of the physical differences between each pair) we used the mean of the discriminabilities for the ABX pairs at the ends of each section of the continuum (i.e. morphs 29,1 and 9,11 for the fear → happiness section. morphs 9,11, and 19,21 for the happiness → anger section; morphs 19,21 and 29,1 for the anger → fear section). To estimate the second factor (contribution from categorical perception of the facial expressions), we used the sum of the differences between the rates at which the members of each pair were assigned to each expression category in the identification task. As in Experiments 1 and 2, these two factors were then combined by adding 0.25 of the identification difference for the relevant pair to the estimate of the overall physical difference, to yield the predicted performance curve for each section of the continuum. The only differences between the formula used to derive the predicted discrimination curve in Experiment 3 and the formula used for

Experiments 1 and 2 are thus those necessarily introduced by the change from using two to three possible expression categories.

The predicted ABX discrimination values are plotted together with subjects' mean percent correct observed performance in the lower graph in Figure 5. There was a significant positive correlation between the predicted and observed values, $r = 0.68$, $t(13) = 3.36$, $p < 0.01$.

An additional correlational analysis using the Liberman et al. (1957) formula to predict discrimination from our identification data also showed a significant positive correlation between the predicted and observed values, $r = 0.69$, $t(13) = 3.41$ $p < 0.01$.

Again, we also tested whether observed discrimination performance showed greater sensitivity to a physical change when it fell in the boundary region between two perceptual categories than to the same change occurring within a particular category. In Experiment 3, the endless continuum can be thought of as being divided into three principal sections, ranging from fear → happiness, happiness → anger, and anger → fear. In each section, two of the ABX discrimination pairs we tested lay close to one of the prototype images (pairs 1,3 and 7,9 for fear → happiness, pairs 11,13 and 17,19 for happiness → anger, and pairs 21,23 and 27,29 for anger → fear); we will call these within-category pairs. In contrast, two other pairs in each section of the continuum were distant from the prototype and fell in the boundary region between two categories (pairs 3,5 and 5,7 for fear → happiness, pairs 13,15 and 15,17 for happiness → anger, and pairs 23,25 and 25,27 for anger → fear); we will call these cross-boundary pairs. In addition, three stimulus pairs (faces 9,11, faces 19,21, and faces 29,1) involve parts of two different sections of the continuum; these were not used in the analysis.

We therefore compared mean ABX discriminability for within-category and cross-boundary pairs in each section of the continuum. A two-factor ANOVA examined the effects of type of pair (within-category or cross-boundary; repeated measure) and section of continuum (fear → happiness, happiness → anger, or anger → fear; repeated measure). This showed that discrimination was better for cross-boundary than for within-category pairs: type of pair, $F(1, 11) = 14.91$, $p < 0.01$. There was also a main effect of section of continuum, $F(2, 22) = 6.84$, $p < 0.01$, showing that discrimination was poorer in the anger → fear region. However, there was no significant type of pair × section of continuum interaction, $F(2, 22) = 1.72$, $p > 0.1$, indicating that enhanced cross-boundary discrimination held across all three sections of the continuum.

Discussion

Experiment 3 found evidence of categorical perception of images from a continuum that was composed only of morphed faces, with no fixed beginning or end. Observed ABX discrimination performance correlated significantly with the performance curve predicted from identification data, and the relative discriminabilities of within-category and cross-boundary pairs in each region of the continuum showed enhanced performance for cross-boundary pairs.

Given the structure of the continuum, these results cannot reflect range effects, anchor effects, or learning of prototypes during the experiment (in fact, prototypes were never shown). They thus provide further support for the view that categorical perception of facial expressions reflects what is in subjects' heads at the start of the experiment, rather than artefacts created by reliance on single continua.

EXPERIMENT 4

One final issue that we wanted to address was the possible role of short-term memory. The ABX discrimination procedure incorporates a visual memory load; in order to decide whether face X matches face A or face B, one must hold representations of A and B in memory. Hence, although the ABX discrimination task is usually interpreted as if it were a test of perceptual processing, it also taps short-term memory function. It is possible, then, that the pattern of discrimination data (heightened discrimination at the category boundary) found by Etcoff and Magee (1992) and ourselves constitutes a short-term memory phenomenon rather than a perceptual effect. Given that we have argued as if our data support the latter interpretation, it was important to provide evidence to support this claim. Hence, in Experiment 4 we replicated the results of Experiment 3 with a discrimination task that does not incorporate a memory load as an essential component; for this purpose we used a same–different matching paradigm. Subjects were asked to decide whether a pair of simultaneously presented face morphs from the continuum used in Experiment 3 were identical or different. They also completed an identification task with morphs from the same continuum.

Method

Subjects

Twelve members of the MRC Applied Psychology Unit subject panel participated in the experiment. They were between the ages of 25 and 42 years and had normal or corrected-to-normal vision. They were paid for participating. None of the subjects had taken part in Experiments 1 to 3.

Design, Materials and Procedure

Subjects saw the continuum (fear → happiness → anger → fear) used in Experiment 3 in a same–different matching task and an identification task.

Same–Different Matching Task. Each trial began with a 250-msec presentation of a central fixation cross followed by two simultaneously presented face morphs, positioned horizontally to the right and left of centre. The morphs were either identical or one step apart on the continuum (e.g. morphs 1 and 3). Subjects were instructed to make a button-press response to indicate whether they thought the faces were the same or different as quickly and accurately as possible. Reaction times and accuracy were recorded. All 15 possible same pairs and 15 one-step different pairs (i.e., 1,3; 3,5; 5,7; etc.) were presented eight times in random order, giving a total of 240 trials. Each morphed face occurred with equal frequency on the left and right of the screen.

Identification Task. Having completed the same–different matching task, the subjects performed an identification task. The design of the identification task was identical to that used in Experiment 3.

Results

The upper graph in Figure 6 shows the mean percentages of times each face in the continuum was categorised as "afraid", "happy", or "angry".

As in Experiment 3, the morphs were perceived as lying along three distinct sections of the continuum, with the boundaries lying in the same places as found in Experiment 3. Again, faces lying in the boundary regions between two contributory expressions (e.g. fear and happiness) were rarely identified as the third, non-contributory expression (e.g. anger for the fear → happiness continuum).

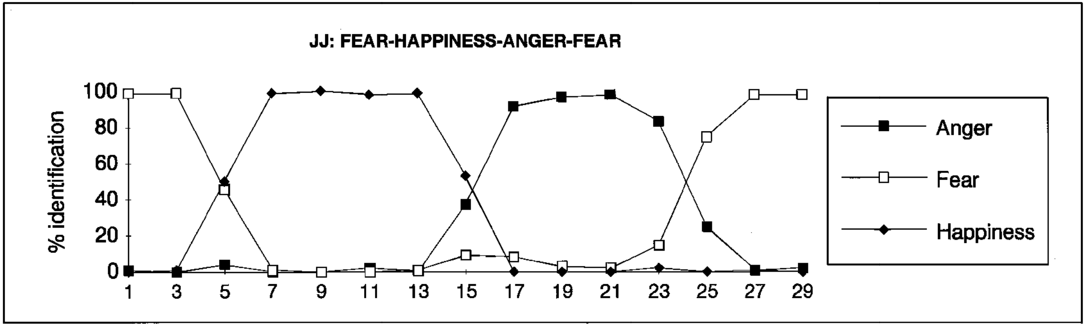
Unlike the forced-choice ABX discrimination procedure used in Experiments 1–3, same–different matching is susceptible to response biases. Inspection of the data from Experiment 4 showed that, as might be expected given the small differences between the pairs of morphs used, subjects tended to respond "same" unless they could detect a difference. Hence, they made few errors to "same" pairs, but many errors involving the misclassification of "different" pairs as "same".

Two different measures of discrimination performance were therefore used– the percentages of different pairs successfully detected, and the signal detection statistic d', which is uncontaminated by response bias.

First, for the percentages of different pairs successfully detected, the observed performance is shown in the lower graph in Figure 6. From the identification data we calculated subjects' predicted performance for the "different" response trials in the same–different matching task. This was done using the same formula as for Experiment 3. The predicted (closed squares) and observed (open squares) discrimination values are plotted in the lower graph in Figure 6. There was a significant relationship between the predicted and observed values for percentage correct different responses, $r = 0.54$, $t(13) = 2.31$, $p < 0.05$.

An additional correlational analysis using the Liberman et al. (1957) formula to predict discrimination from our identification data showed a positive correlation between the predicted and observed values, which did not reach statistical significance at the 0.05 level, $r = 0.45$, $t(13) = 1.79$, $0.1 > p > 0.05$.

EXPERIMENT 4: IDENTIFICATION



EXPERIMENT 4: DISCRIMINATION: BASED ON 'DIFFERENT' RESPONSES

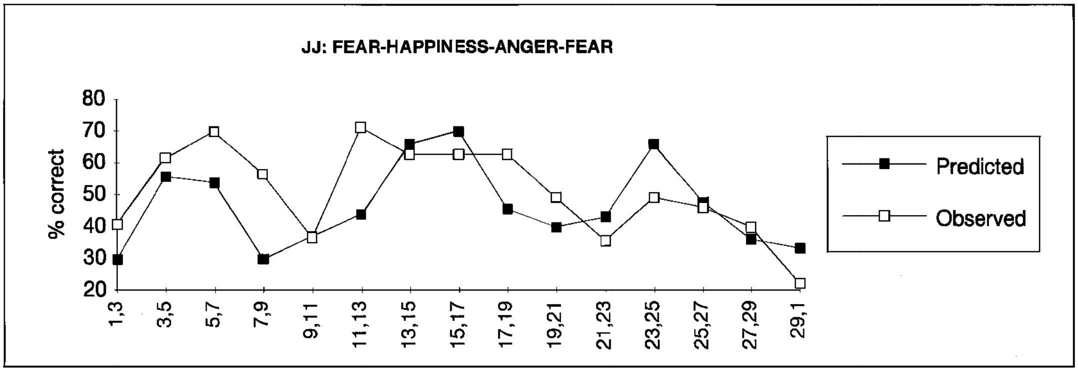


FIG. 6. Identification data (top) and same–different matching data (bottom) from Experiment 4. The identification data show the frequency (mean percentage) with which each face in the continuum was identified as fear, happiness, or anger. The discrimination data show subjects' predicted and observed performance given as the mean percentage of times each face pair was discriminated correctly.

We also tested whether the observed discrimination performance showed greater sensitivity to a physical change when it crossed the boundary between two perceptual categories than to the same change occurring within a particular category. To do this, we again divided the continuum into three principal sections ranging from fear → happiness, happiness → anger, and anger → fear, as for the analysis of Experiment 3. In each section, two of the pairs we tested were within-category pairs lying close to one of the prototype images (pairs 1,3 and 7,9 for fear → happiness, pairs 11,13 and 17,19 for happiness → anger, and pairs 21,23 and 27,29 for anger → fear), and two other pairs in each section were cross-boundary pairs falling in the region between two categories (pairs 3,5 and 5,7 for fear → happiness, pairs 13,15 and 15,17 for happiness → anger, and pairs 23,25 and 25,27 for anger → fear).

We therefore compared mean same–different discriminability for within-category and cross-boundary pairs in each section of the continuum. A two-factor ANOVA examined the effects of type of pair (within-category or cross-boundary; repeated measure) and section of continuum (fear → happiness, happiness → anger, or anger → fear; repeated measure). This showed that discrimination was better for cross-boundary than for within-category pairs: type of pair, $F(1, 11) = 8.86$, $p < 0.05$. There was also a main effect of section of continuum, $F(2, 22) = 5.93$, $p < 0.01$, showing that discrimination was again poorer in the anger → fear region. However, both of these main effects were qualified by a significant type of pair × section of continuum interaction $F(2, 22) = 5.18$, $p < 0.05$, which indicated that enhanced cross-boundary discrimination was not consistently found across all three sections of the continuum. Post hoc Tukey tests ($\alpha = 0.05$) showed that differences between within-category and cross-boundary pairs were only significant in the fear → happiness and anger → fear sections of the continuum, not in the happiness → anger section.

As a second measure of discrimination performance, we calculated observed sensitivity values for the signal detection statistic d'; these are shown in Figure 7. To derive a predicted curve for d' performance, we again used the formula set out for Experiment 3, but we weighted the second factor (sum of identification differences) at 0.01 instead of 0.25 to allow for the different numerical scale of d' as compared to percentage correct. Again, there was a significant relationship between the predicted and observed discrimination values, d', $r = 0.53$, $t(13) = 2.23$, $p < 0.05$, based on our own formula. An additional correlation analysis using Liberman et al. (1957) formula showed a positive correlation between predicted and observed values that did not reach statistical significance at the 0.05 level, $r = 0.44$, $t(13) = 1.76$, $0.1 > p > 0.05$.

As Experiment 4 used the same continuum as Experiment 3 and only differed in the discrimination measure used, we were interested to compare the results from the two groups of subjects with different discrimination tasks: ABX and same–different matching. There was a significant correlation between the

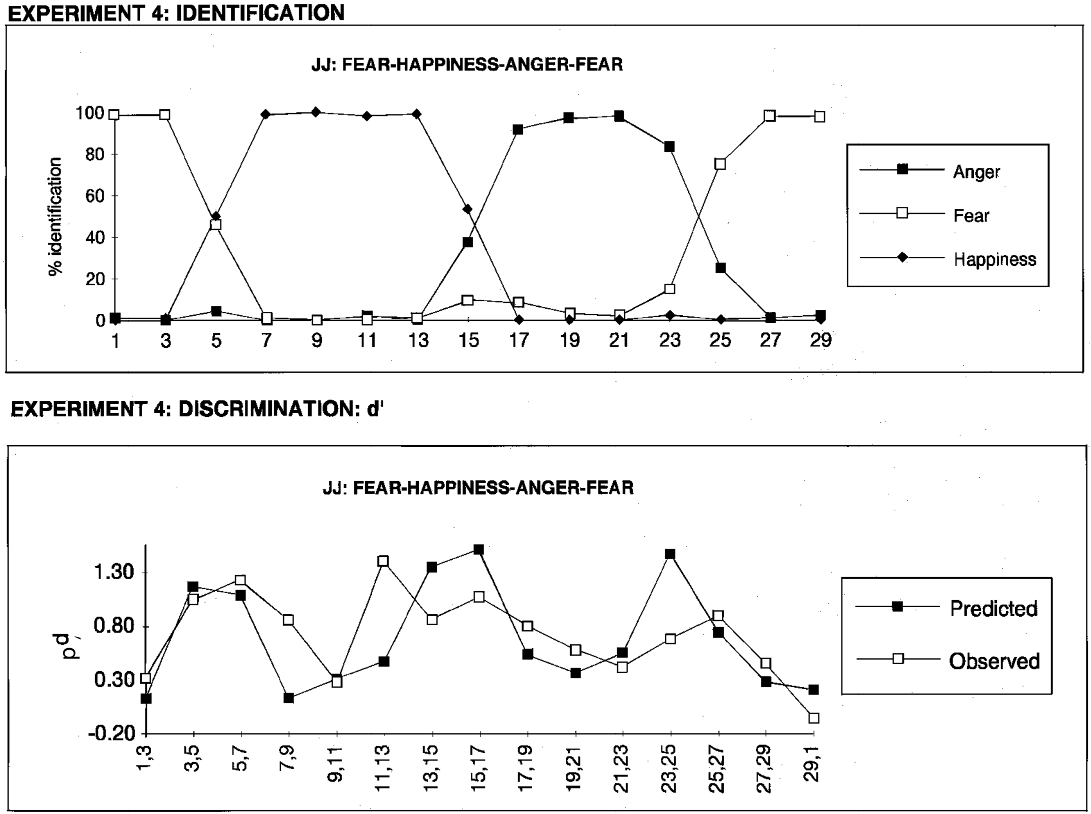


FIG. 7. Predicted and observed sensitivity (d') values in the Experiment 4 same–different matching task.

observed mean percentage of correct "different" responses recorded in Experiment 4 and subjects' observed mean percentage of correct responses in the ABX task from Experiment 3, $r = 0.84$, $t(13) = 5.58$, $p < 0.001$. This suggests that ABX discrimination and same–different matching tasks were tapping the same abilities.

In Experiment 4, we were able to record subjects' reaction times in the discrimination task. Reaction times had not been used as a dependent variable in the ABX discrimination task chosen for Experiments 1–3 because with ABX discrimination it is not clear whether response times might be affected by the order of the stimuli; for example, subjects may be faster to respond when face B = X than when face A = X, and it is quite likely that they are doing different things in making these comparisons to the immediate or one-before-last items.

Mean reaction times to make a correct "different" response in Experiment 4 are plotted in Figure 8. As for our other correlational analyses, we calculated subjects' predicted performance for this measure using a formula incorporating two factors: (1) an estimate of subjects' ability to make use of the constant physical differences between each pair of images; for this, we used the mean reaction times for the ABX pairs at the ends of each section of the continuum (i.e. morphs 29,1, and 9,11 for the fear → happiness section; morphs 9,11, and 19,21 for the happiness → anger section; morphs 19,21, and 29,1 for the anger → fear section); (2) an estimate of any additional contribution from categorical perception of the facial expressions, based on the sum of the differences between the rates at which the members of each pair were assigned to each expression category in the identification task. These two factors were then combined by subtracting twice the identification difference for the relevant pair from the estimate of the overall physical difference, to yield the predicted performance curve for each continuum.

Subjects' predicted and observed mean reaction times to make a correct "different" response are plotted in Figure 8. The predicted and observed values were significantly correlated, $r = 0.52$, $t(13) = 2.17$, $p < 0.05$. An additional correlational analysis using the Liberman et al. (1957) formula also showed a significant correlation between predicted and observed reaction times, $r = 0.53$, $t(13) = 2.27$, $p < 0.05$.

Again, we also compared mean same–different discriminability for within-category and cross-boundary pairs in each section of the continuum. A two-factor ANOVA examined the effects of type of pair (within-category or cross-boundary; repeated measure) and section of continuum (fear → happiness, happiness → anger, or anger → fear; repeated measure). This showed that discrimination was better for cross-boundary than for within-category pairs: type of pair, $F(1, 11) = 6.40$, $p < 0.05$. There was no main effect of section of continuum, $F < 1$, and no type of pair × section of continuum interaction, $F < 1$, which indicated that enhanced cross-boundary discrimination (as reflected in faster reaction times for correct responses) held across all three sections of the continuum.

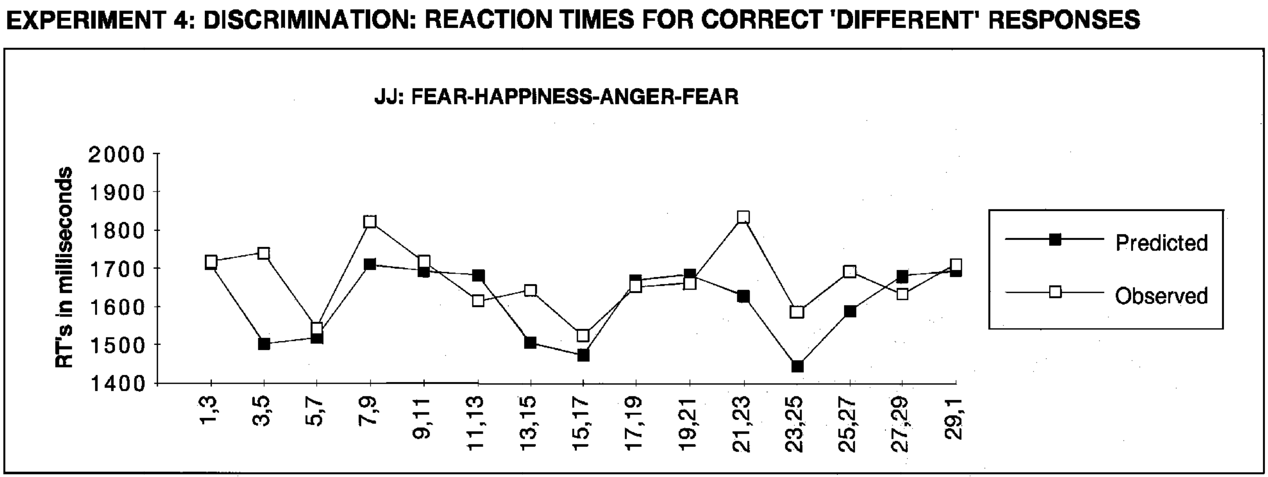


FIG. 8. Predicted and observed mean reaction times to make a correct “different” response in the Experiment 4 same–different matching task.

Discussion

Experiment 4 showed evidence of categorical perception with a same–different matching paradigm, using three dependent measures: correct responses to different pairs, sensitivity values (d'), and reaction times. For each measure, observed discrimination performance correlated significantly with performance curves predicted from identification data using our own formula. With the Liberman et al. (1957) formula, correlations of predicted and observed discrimination performance approached significance for correct responses to different pairs and for sensitivity values (d'), and there was a significant correlation between predicted and observed reaction times. Additional analyses of the relative discriminabilities of within-category and cross-boundary pairs showed enhanced performance for cross-boundary pairs in the fear → happiness and anger → fear sections of the continuum in terms of correct responses to different pairs, and for all sections of the continuum in terms of reaction time.

The pattern of discrimination performance shown with the same–different matching task in Experiment 4 was also found to be highly correlated with the observed performance of the ABX discrimination task used in Experiment 3. As we discussed earlier, the same–different matching task can be considered a test of perceptual processing, whereas the ABX discrimination task taps both perceptual and short-term memory function. Of course, it is *possible* that people still make use of some form of short-term storage even in the same–different matching task; however, the important point is that same–different matching with simultaneously presented pairs does not in itself *demand* that they do so, and it thus gives a rather better estimate of perceptual discrimination than does the sequential ABX task. Given the high degree of concordance between the data from these two tasks, it seems reasonable to conclude that categorical perception of facial expressions is more likely to be a perceptual than a short-term memory phenomenon.

GENERAL DISCUSSION

Etcoff and Magee (1992) found evidence of categorical perception of morphed line drawings of facial expressions. In this paper we have replicated and extended their principal findings with more naturalistic stimuli: continuous-tone (photographic-quality) images. It is true that, because of the substantial work involved in rendering images to these high standards, we have only investigated a subset of the emotions and continua explored by Etcoff and Magee (1992). For pragmatic reasons, it seemed to us best to begin by demonstrating clear and consistent effects with this limited set of four emotions from the six (excluding neutral) in the Ekman and Friesen (1976) series. In addition, we wanted to address two important questions raised by Etcoff and Magee's (1992) study and our own Experiments 1 and 2. (1) Could these data be accounted for by artefacts inherent in the use of single continua? (2) Given that the ABX discrimination task incorporates a short-term memory load, could the results reflect a short-term memory phenomenon rather than a perceptual effect?

In order to replicate Etcoff and Magee's results with photographic-quality morphs, Experiments 1 and 2 examined three types of single continua (happiness → sadness, sadness → anger, and anger → fear). Each continuum consisted of evenly spaced faces moving from one prototype expression to another (e.g. happiness → sadness). Between Experiments 1 and 2, we found evidence of categorical perception of all three types of single continua.

Having replicated the basic phenomenon, we then addressed two key issues that arise in this type of work and have seldom been explored in other investigations of categorical perception effects. Several studies have shown that there are problems with using single continua in psychophysical experiments because they can distort the subjects' performance towards the middle of the range (Poulton, 1975). This effect can occur for a number of reasons: learning the stimulus range, anchor effects, or learning prototypes. Hence, it was not clear whether the patterns of discrimination data found in Experiments 1 and 2 reflected something that subjects brought with them to the experiment or phenomena that might be attributed to the structure of the experiments themselves.

To address this issue, Experiment 3 used a continuum prepared from three prototype expressions (fear → happiness → anger → fear). This continuum had no physical endpoints. In addition, the continuum only contained morphed faces, because the three prototype expressions were removed. Despite these precautions, subjects' identification and discrimination data showed the characteristic patterns associated with categorical perception. Each of the three sections (fear → happiness; happiness → anger; anger → fear) embedded in the continuum showed a clear category boundary, with no intrusions from the third possible expression. Additionally, an analysis of the ABX discrimination data showed that subjects' predicted performance, which we derived from their identification data, gave a good fit to their observed performance.

In Experiment 4, we addressed the role of short-term memory in the ABX discrimination task. The ABX procedure incorporates a memory load—that is, matching X with A or B necessitates holding representations of A and B in memory. Hence, it was possible that the discrimination data from Etcoff and Magee (1992) and our own Experiments 1 to 3 reflected a short-term memory phenomenon rather than a perceptual effect. To distinguish between these two explanations, Experiment 4 used a discrimination task that did not incorporate a memory component: a same–different matching paradigm. The subjects were presented with the continuum seen in Experiment 3, and the data from the two Experiments were compared. The results of Experiment 4 still showed evidence of categorical perception of facial expressions, and subjects' performance in the two discrimination tasks (Experiment 3 using ABX, and Experiment 4 using same–different matching) was highly correlated.

The results of Experiments 3 and 4, then, provide further support for the idea that facial expressions are perceived categorically, because they exclude alternative explanations such as range effects, anchor effects, or short-term memory function.

Two aspects of these data require further discussion: (1) the relative merits of line drawing and photographic-quality stimuli in this type of research, and (2) the nature and implications of the categorical perception effects we observed.

Line Drawings and Photographic-quality Stimuli

We have replicated and verified Etcoff and Magee's (1992) principal findings with continuous-tone images. Having established this, it is interesting to compare Etcoff and Magee's results with our own.

Experiment 2 examined subjects' performance with four examples of happiness → sadness, sadness → anger, and anger → fear expression continua. Etcoff and Magee used a very similar design to examine the same three continua in their study, so it is particularly relevant to compare these two sets of data. The first point of interest is the marked resemblance between the identification data from the two studies; the happiness → sadness and anger → fear continua are particularly similar. The second point to note is that line-drawing and continuous-tone images produce rather different patterns of discrimination data. In the present study we found that the discriminability of two face morphs increased in line with their proximity to the category boundary but was not zero within each category; this pattern is consistent with studies showing categorical perception of phonemes and colours, where within-category discrimination remains above-chance (Bornstein & Korda, 1984; Liberman et al., 1957). Etcoff and Magee (1992) also noted that within-category discrimination was above chance for line drawings of facial expressions, but they found a slightly different pattern of relatively constant within-category discrimination with a sudden increase in performance for the two morphs that straddled the category boundary.

We think that there is a relatively simple explanation for these differences between our discrimination data and those of Etcoff and Magee (1992). Line-drawings are minimalist representations of faces that leave out much of the information present in photographs. In particular, line drawings preserve two-dimensional shape information but eliminate much of the face's pigmentation and cues to its three-dimensional structure derived from shading.

What are the effects of these differences? There is a striking resemblance between the identification data obtained with line-drawings and photographic-quality morphs. Hence, it is clear that line-drawings contain sufficient information to *identify* facial expressions. In contrast, photographic-quality images contain not only information needed to identify facial expressions but also much information that is irrelevant to the perception of expression (e.g. hair texture, skin blemishes). When subjects are asked to *discriminate* between two continuous-tone morphs, there are therefore expression-relevant and expression independent cues that they can use. In comparison to photographic-quality morphs, line-drawing morphs contain a predominance of the expression-relevant shape cues, and this would account for the different patterns of discrimination data.

One conclusion that could be drawn from our results, therefore, is that line-drawing morphs may offer a relatively pure medium within which to examine categorical perception of facial expressions relatively uncontaminated by other perceptual differences. This observation does not detract from the significance of this present study. Line drawings clearly lack ecological validity. Hence, it was important to establish that Etcoff and Magee’s (1992) results could be replicated with more natural-looking stimuli. It is useful to know that this can be done, but it is equally important to recognize the limitation that photographic-quality morphs will contain many changes that are not relevant to the perception of expression, and hence add noise to discrimination performance if this is intended as a measure of categorical perception of expression.

Nature and Implications of the Categorical Perception Effects

We turn now to consider the nature and implications of the categorical perception effects we observed.

The basic claim of categorical perception is that a *physical* continuum need not be *perceptually* linear. Although this seems a straightforward idea, non-linearity can be indexed in a number of different ways. We have noted five different criteria for non-linearities that might be associated with categorical perception that have been used in the literature:

1. There should be abrupt shifts in identification along the continuum.
2. Observed discriminability of stimulus pairs should correlate with discrimination performance predicted from identification.
3. Discrimination should be better for cross-category than within-category stimulus pairs.
4. Peak discriminability should align precisely with category boundaries in identification.
5. Within-category discriminability should be constant.

In terms of these criteria, our results usually met (1), (2), and (3), but not (4) and (5). Although the key categorical perception effect of enhanced discrimination of pairs of stimuli falling across rather than well within category boundaries (criterion 3) was found for all of our experiments, we noted that peak discriminabilities did not always align perfectly with category boundaries, and that within-category discriminabilities were not constant.

We do not think this is particularly surprising. Criteria 4 and 5 will really only apply in circumstances where categorical perception is the overwhelmingly predominant influence. But a stimulus like the face has perceptual and social significance extending well beyond its expression, important though that may be. Hence, we would not want to claim that categorical perception of expressions is the *only* influence on subjects' discrimination performance in our experiments. As we pointed out in discussing the differences between results with line-drawings (Etcoff & Magee, 1992) and our continuous-tone images, photographic-quality morphs will contain many changes that are not relevant to the perception of expression but can be used to assist discrimination. Consistent with this point, Etcoff and Magee's (1992) findings with line-drawings, in which expression-relevant cues are indeed relatively dominant, did often meet criteria 4 and 5.

The present experiments have therefore provided evidence of categorical perception of morphed facial expressions, but we do not seek to claim that this is the sole contributory factor in discrimination tasks.

A particularly striking feature of the identification data in Experiments 3 and 4 was that morphs lying in the regions where identification changed from one emotion to another were still identified as one of the two nearest prototype expressions and hardly ever as the third, which did not contribute to that particular range. For example, face 5 from the fear → happiness range was principally identified as "afraid" or "happy", and not as "angry". As we noted, this implies that each boundary was still perceived as part of a continuum ranging between two expressions, rather than as a region of general uncertainty. This is as would be expected on the hypothesis that basic emotions each have their own distinct prototype, but it is less easy to reconcile with a dimensional account. If emotions are analyzed in terms of underlying dimensions instead of as discrete perceptual categories, we might expect that morphed continua between images that lie across these dimensions will at some point enter regions where the expression becomes indeterminate. For the limited set of emotions used in Experiment 3 and 4, this did not happen.

The results of Experiments 3 and 4, therefore, suggest that we might in future wish to add a sixth potential criterion for categorical perception: When moving along a continuum from one category prototype to another, there should be no intrusions from irrelevant prototypes in the identification data. We plan to explore this possibility in further work using a wider range of facial expression continua.

Because of the strong biological background evident in facial expressions of emotion (Darwin, 1872; Ekman, 1973), it might be thought that the evidence of categorical perception we have found reflects a relatively fixed influence from a dedicated underlying biology. Although this assumption is often made, in part because other categorical perception effects (e.g. for speech or colour perception) also seem to be linked to underlying biological predispositions, we see no compelling reason to accept it. Instead, it seems to us more likely that categorical perception will prove to be an emergent property of population coding in the nervous system; in other words, we suspect that categorical perception effects arise whenever populations of cells become tuned to distinct categories. The crucial test will therefore be to determine whether categorical perception effects can be created by learning, and some of the other aspects of face perception (such as recognition of identity) offer interesting possibilities for further research directed at this issue.

To a great extent, categorical perception research has been dictated by technology. Consequently, the majority of studies in this area have explored the perception of primary perceptual categories, such as colours and phonemes. Modern imaging techniques allow us to extend the domain of study to the perception of complex multidimensional stimuli—in this case, facial expressions. The finding of enhanced discrimination at category boundaries for these multidimensional stimuli is consistent with the idea that facial expressions are perceptually coded in terms of their conformity to prototype configurations corresponding to basic emotion categories, but also with the more general proposition that categorical perception is a fundamental property of the categorization process.

REFERENCES

Benson, P.J., & Perrett, D.I. (1991a). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology, 3,* 105–135.

Benson, P.J., & Perrett, D.I. (1991b). Synthesising continuous-tone caricatures. *Image and Vision Computing, 9,* 123–129.

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.

Brennan, S.E. (1985). Caricature generator: Dynamic exaggeration of faces by computer. *Leonardo, 18,* 170–178.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Ekman, P. (Ed.). (1973). *Darwin and facial expression: A century of research in review.* New York: Academic Press.

Ekman, P. (Ed.). (1982). *Emotion in the human face.* Cambridge: Cambridge University Press.

Ekman, P. (1992). Facial expressions of emotion: An old controversy and new findings. *Philosophical Transactions of the Royal Society, London, B335,* 63–69.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268–287.

Ekman, P., & Friesen, W.V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W.V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon.

Etcoff, N.L. (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia, 22,* 281–295.

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

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

Harnad, S. (Ed.). (1987). *Categorical perception: The groundwork of cognition.* Cambridge: Cambridge University Press.

Johnson, M.H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19.

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.

McNemar, Q. (1962). *Psychological statistics* (3rd edition), New York and London: Wiley.

Pisoni, D.B., & Lazarus, J.H. (1974). Categorical and noncategorical modes of speech perception along with voicing continuum. *Journal of the Acoustical Society of America, 55,* 328–333.

Poulton, E.C. (1975). Range effects in experiments on people. *American Journal of Psychology, 88,* 3–32.

Sergent, J., Ohta, S., MacDonald, B, & Zuck, E. (1974). Segregated processing of facial identity and emotion in the human brain: A PET study. *Visual Cognition, 1,* 349–369.

Young, A.W., Newcombe, F., De Haan, E.H.F., Small M., & Hay, D.C. (1993). Face perception after brain injury: Selective impairments affecting identity and expression. *Brain, 116,* 941–959.

Manuscript received 1 June 1995

Revised manuscript received 15 September 1995