

The prototype effect in face recognition: Extension and limits

ROBERTO CABEZA

University of Alberta, Edmonton, Alberta, Canada

VICKI BRUCE

University of Stirling, Stirling, Scotland

and

TAKASHI KATO and MASAOMI ODA

ATR Human Information Processing Research Laboratories, Kyoto, Japan

The prototype effect in face recognition refers to a tendency to recognize the face corresponding to the central value of a series of seen faces, even when this central value or prototype has not been seen. Five experiments investigated the extension and limits of this phenomenon. In all the experiments, participants saw a series of faces, each one in two or more different versions or exemplars, and then performed a recognition test, including seen and unseen exemplars and the unseen prototype face. In Experiment 1, a strong prototype effect for variations in feature location was demonstrated in oldness ratings and in a standard old/new recognition test. Experiments 2A and 2B compared the prototype effect for variations in feature location and variations in head angle and showed that, for the latter, the prototype effect was weaker and more dependent on similarity than for the former. These results suggest that recognition across feature variations is based on an averaging mechanism, whereas recognition across viewpoint variations is based on an approximation mechanism. Experiments 3A and 3B examined the limits of the prototype effect using a face morphing technique that allows a systematic manipulation of face similarity. The results indicated that, as the similarity between face exemplars decreases to the level of similarity between the faces of different individuals, the prototype effect starts to disappear. At the same time, the prototype effect may originate false memories of faces that were never seen.

How do we build up stable representations of new faces from variations in viewpoint, expressions, and lighting? This question has been somewhat neglected despite the enormous amount of recent interest in face recognition (e.g., see Bruce, Cowey, Ellis, & Perrett, 1992). The prototype effect in face recognition can provide useful information about this issue.

The *prototype effect* refers to a tendency to respond to the central value of a series of varying exemplars, even when this central value or prototype has not been experienced. The prototype effect has been observed mainly on classification tests (Homa, Cross, Cornell, Goldman, & Schwartz, 1973; Homa, Dunbar, & Nohre, 1991; Homa, Goldhardt, Burrue-Homa, & Carson Smith, 1993; Homa & Vosburgh, 1976; Posner & Keele, 1968) and recognition

tests (Bruce, Doyle, Dench, & Burton, 1991; Franks & Bransford, 1971; Homa et al., 1993). Posner and Keele (1968), for example, trained participants to classify into different categories dot patterns that were variations of some underlying prototype patterns. In a subsequent classification test, the nonstudied prototype patterns were classified as fast and as accurately as studied exemplars and better than nonstudied exemplars. In recognition tests, the prototype effect is usually revealed by false alarms to a nonstudied prototype pattern, and it has been demonstrated with several different materials, including geometric forms (Franks & Bransford, 1971; Homa et al., 1993), kinesthetic movements (Solso & Raynis, 1979), numbers and words (Solso, Heck, & Mearns, 1993), and human faces (Bruce et al., 1991; Solso & McCarthy, 1981).

The *prototype effect in face recognition* has typically been investigated using Identikit faces (Inn, Walden, & Solso, 1993; Malpass & Hughes, 1986; Solso & McCarthy, 1981). Solso & McCarthy, for example, created a prototype Identikit face consisting of a particular combination of hair, eyes, nose + chin, and mouth and then produced exemplars of this prototype by varying one, two, three, or four of its components (75%, 50%, 25%, or 0% exemplars, respectively). Participants tried to memorize

The experiments described in this paper were conducted while the first author held a postdoctoral post at ATR labs, and the second author was in receipt of a collaborative research award from ATR. The authors thank Shigeru Akamatsu and Yoh'ichi Tohkura for the provision of research facilities, and Morris Moscovitch, Rob Althoff, Robert Solso, and Pepper Williams for comments on a previous version of this paper. Correspondence should be addressed to R. Cabeza, Department of Psychology, University of Alberta, P220 Biological Sciences Bldg., Edmonton, AB T6G 2E9, Canada (e-mail: cabeza@psych.ualberta.ca).

three 75%, four 50%, and three 25% exemplars and later performed a recognition test with confidence ratings. Seen exemplars were correctly recognized, but the unseen prototype was recognized with greater confidence. Bruce et al. (1991) investigated the prototype effect in face recognition with a different technique. Using computerized face reconstruction software (Mac-a-Mug Pro), which allows high-quality line-drawn faces to be produced, they created exemplars from prototype faces by displacing internal features (eyebrows, eyes, nose, and mouth) up and down by a certain number of pixels (e.g., 3, 6, 9, 12, and 15 pixels). Although this form of variation does not occur naturally, it corresponds roughly to natural forms of face variations, such as the ones associated with aging, expression movements, and so on. At study, participants rated the apparent age and masculinity–femininity of face exemplars (incidental learning) and then received a recognition test. In a forced-choice recognition test (Experiments 1–6), participants preferred the unseen prototype to other unseen exemplars, and they could not distinguish between the unseen prototype and seen exemplars. In a test in which they had to rate their confidence of having seen exactly the same picture (Experiments 7 and 8), participants rated the prototype as high as seen exemplars and higher than unseen exemplars.

These studies showing prototype effects in the recognition of previously studied facial exemplars are potentially illuminating about the mechanisms by which representations of novel faces become built up from varying exemplars. The experiments described above strongly suggest that, at least for variations of features or their placements in full-face images, successive exemplars of the same face may be stored or retrieved in a way that favors the nonstudied average of the exemplar set. (We return later to the controversy over the precise mechanism that might yield such effects.) However, these previous experiments are limited both in the nature of the materials used and in the tasks that have been used to demonstrate prototype effects. To be an ecologically valid mean of assessing how representations of faces are established, the task at test should be to distinguish old faces from new ones, rather than to recognize exactly the same pictures studied, and the faces shown should be realistic rather than line-drawn.

In the present experiments, we adopted these more ecologically valid methods to investigate the limits of the prototype effect in face recognition. Previous studies have shown the prototype effect when participants rated whether each test exemplar was old or new using *Identikit* (e.g., Solso & McCarthy, 1981) or line-drawn (e.g., Bruce et al., 1991) faces in the front view. In Experiment 1, we studied the prototype effect not only with oldness ratings but also using a standard old/new recognition test, in which participants were asked to rate whether the face, rather than the exemplar, was old or new. We used high-quality color photographs of faces.

Second, we extended the investigation of the prototype effect in face recognition to variations between and

within different viewpoints. A strong prototype effect has typically been observed for face changes within the same view, such as the exchange (e.g., Solso & McCarthy, 1981) or displacement (e.g., Bruce et al., 1991) of face features. In contrast, Bruce (1994) reported some preliminary data suggesting that the prototype effect may be difficult to obtain when the exemplars show variations in head angle. This finding was interpreted as indicating that recognition across feature variations reflects an averaging or superimposing mechanism that does not operate across different views. The reason why averaging or superposition of exemplars can operate within but not between head angles is because viewpoint changes affect the position of facial features within the visual image, rendering any superposition or averaging invalid. Recognition across different views is more likely to involve an approximation or interpolation mechanism (Bülhoff & Edelman, 1992; Poggio & Edelman, 1990). If exemplars are stored separately, their prototype may still be advantaged in recognition because it is more similar to more stored exemplars than is any novel exemplar. However, prototype recognition based on similarity to stored exemplars should be more dependent on probe–exemplar similarity than prototype recognition based on averaging or superposition. This reasoning suggests that the prototype effect for variations in viewpoint should be more dependent on probe–exemplar similarity than the prototype effect for feature variations. This prediction was tested in Experiments 2A and 2B.

Finally, building on the findings of Experiments 1 and 2, Experiments 3A and 3B examined the limits of the prototype effect using a novel technique that allows a systematic manipulation of face similarity: morphing of different faces. We investigated whether the prototype effect would disappear when face exemplars became different enough to be attributed to different individuals.

Each experiment had two phases. In the first phase, participants rated faces according to some incidental dimension (e.g., masculinity). Each individual's face appeared in the sequence of images to be rated in two or more different versions, or exemplars (e.g., with different internal feature locations or different head angles). In the second phase, participants performed a face-exemplar recognition test (as in Bruce et al.'s, 1991, Experiments 7 and 8) or a face-identity recognition test. In the face-exemplar recognition test, all faces were old, but they are presented in different versions (exemplars), some of which were old and some of which were new. For each exemplar, participants had to indicate their confidence of having seen exactly the same version of the face at study. In the face-identity recognition test, in contrast, old faces were mixed with completely new faces, and participants had to indicate as fast as possible for each face whether it was old or new. Both tests included three versions of each critical face: one of the seen exemplars (a studied version of a studied face), one unseen exemplar (a nonstudied version of a studied face), and the prototype (a nonstudied version of a studied face, which falls in the center of the

variation range of the studied versions). In the present experiments, the prototype effect was said to occur when recognition of the unseen prototype (as old) was (1) significantly higher than recognition of the noncentral unseen exemplars and (2) not significantly lower than recognition of seen exemplars.

EXPERIMENT 1

Experiment 1 investigated the prototype effect for feature shifts, and it had three objectives. The first was to determine whether the strong prototype effect for feature displacements found by Bruce et al. (1991) with schematic faces (Mac-a-Mug Pro) could be replicated with high-quality color photographs (see Figure 1). The second objective was to examine whether the results obtained with the front-view prototype can be generalized to other prototype views. Experiment 1 investigated the prototype effect for two prototype views: the front view (see Figure 1, right column) and the 45°-left view (see Figure 1, left column). The third aim was to determine whether the results obtained with the face-exemplar recognition test can be generalized to a standard face-identity recognition test. It is assumed (see Bruce, 1994) that the results of these tests are equivalent, but no empirical data has been provided to support this assumption.

Method

Participants. Thirty-two students participated in the experiment. In all the experiments and pilot studies reported in this article, the participants were undergraduate Japanese students who were paid for their help.

Materials. The materials were constructed from the photographs of the faces of 24 Japanese men in their 20s (see example in Figure 1). The photographs were taken in a professional photo studio, but the models were not professional models. The background (blue) and lighting (frontal) was the same in all pictures. The digitized images had a size of 256 × 256 pixels and a definition of 72 pixels per inch and 32 bits per pixel. The size of the heads (from the top of the hair to bottom of the chin) was about 250 pixels. Part of the neck was visible, but all clothes were eliminated by covering them with the background color. Figure 1 shows examples of the prototype and exemplar faces. The prototype was either the natural front view or the natural 45°-left view. Exemplars were created by displacing the internal features of the face (eyebrows, eyes, nose, and mouth) up or down, by either 8 or 12 pixels. Features were displaced using the software Morph, which does not produce any visible "joint."

The faces of 16 men were divided into eight sets of 2 men each, which were counterbalanced across conditions. In the face-identity recognition test, four sets were presented both at study and at test, and four sets were presented only at test as the distractors. Half of the sets (studied and nonstudied) were assigned to the front-view prototype condition, and half were assigned to the left view prototype. Half of the sets (front and left) were assigned to a feature shift condition, and half were assigned to a head-angle variation condition. The head-angle manipulation was investigated in a more controlled fashion in Experiments 2A and 2B; hence, its results in Experiment 1 are not reported. For each face in the feature-location condition, the 8-pixel-up and the 8-pixel-down exemplars were included in the study list. At test, the seen exemplar was one of these pictures, and the unseen exemplar was the 12-pixel exemplar in the opposite direction.

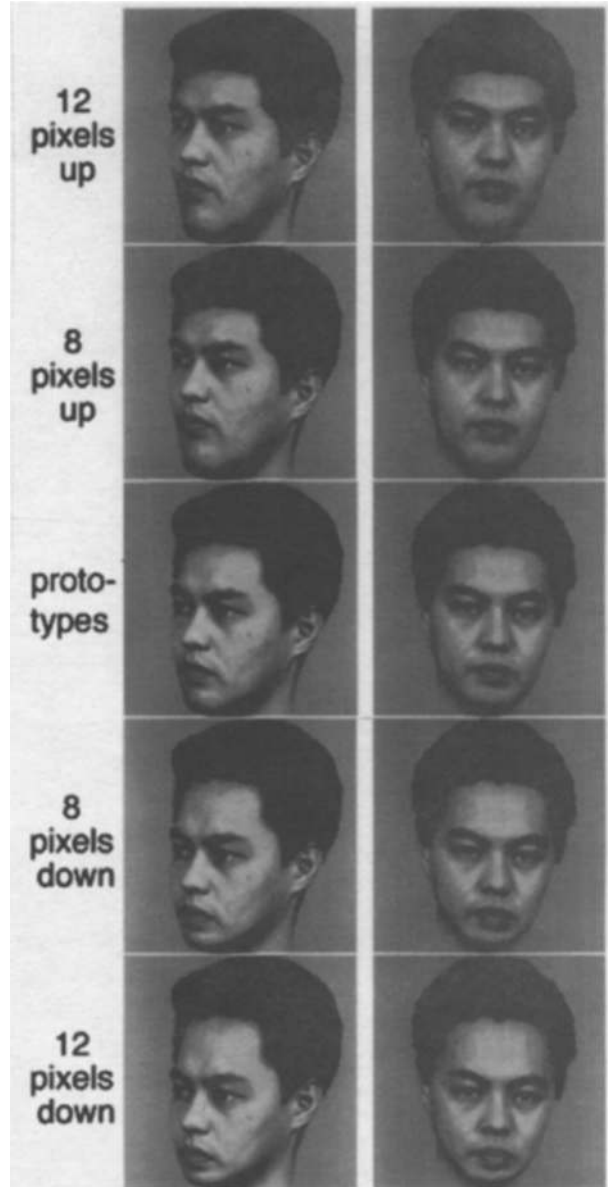


Figure 1. Examples of the materials of Experiments 1: 8- and 12-pixel feature-location exemplars of the front and 45°-left prototypes.

Procedure. The experiment was conducted on a Macintosh Quadra 950 with a 24-bit video card and a color monitor with a resolution of 72 dpi. The software was SuperLab Version 1.6 (Cedrus, Wheaton, MD). In the study phase, each face was displayed for 12 sec and was followed by the masculinity (1 = "woman-like" male face; 8 = "man-like" male face) and age (1 = 18–21 years; 8 = 46–49 years) scales, which remained on the screen until the experimenter entered participant's verbal response. The presentation sequence seen by each participant comprised a total of 16 exemplars, the 8-pixel-up and 8-pixel-down versions of two heads in front view and of two heads in 45°-left view, plus viewpoint variations of four more heads. Each picture was presented twice, and rated on both scales in each presentation. Presentation order was randomized for each participant. The participants were told that, even though similar faces would appear, they should rate the unique appearance of each face with-

out considering their previous responses. No test was mentioned during the study phase, and, hence, learning was incidental.

After the study phase, half of the participants performed a face-exemplar recognition test, and half performed a face-identity recognition test. The face-exemplar recognition test included only the studied faces. For each studied face, there were three pictures: the prototype, one of the two seen exemplars (a studied version of a studied face), and one unseen exemplar (a nonstudied version of a studied face). In the instructions, the participants were told that all the faces that would appear on the screen had been presented previously but that some of the pictures were old and some were new. They were instructed to indicate using a 10-point scale (0 = *definitely not seen picture*; 9 = *definitely seen picture*) how confident they were that exactly the same picture had appeared in the previous phase. The instructions warned the participants that they had to recognize pictures, not faces or people, and, consequently, if they thought that a face had been seen before but the picture was not identical, they should answer negatively (0–4). The participants told their ratings to the experimenter, who typed them on the keyboard. The face-identity recognition test included the seen, unseen, and prototype pictures of the studied faces mixed with an equal number of equivalent pictures of nonstudied faces. The participants were told that seen and unseen faces would appear on the screen, and their task was to answer as fast as possible for each face whether it had been seen on the previous phase (press “Z” key) or not (press “/” key). Unlike the face-exemplar recognition test, if the participants thought that a different picture of the same face had been seen in the study phase, they had to answer “yes.”

Results and Discussion

The results of the face-exemplar and face-identity recognition tests in Experiment 1 are shown in Table 1. In the face-identity recognition test, rejection of nonstudied faces was almost perfect (less than 3% of false alarms), and differences in reaction times (RTs) were generally nonsignificant; accordingly, only recognition responses (seen or unseen) to studied faces are reported and discussed here. The alpha level for all statistical tests in this article was set at .05.

Observation of the face-exemplar recognition ratings in Table 1 suggests a strong prototype effect in both prototype views. Confirming this impression, a 3 (test cue) \times 2 (prototype view) analysis of variance (ANOVA) yielded a significant main effect of test cue [$F(2,30) = 12.90$, $MS_e = 30.17$], and planned contrasts indicated that pro-

totypes [$F(1,30) = 25.27$, $MS_e = 59.10$] and seen exemplars [$F(1,30) = 9.90$, $MS_e = 23.16$] were both rated reliably higher than unseen exemplars and that prototypes were rated almost significantly higher than seen exemplars [$F(1,30) = 3.53$, $MS_e = 8.27$, $p < .07$]. Thus, replicating the results of Bruce et al. (1991), there was a strong prototype effect for feature shifts. The main effect of prototype view and the test cue \times prototype view interaction were both nonsignificant, indicating that the prototype effect for the feature-location manipulation is not dependent on prototype view.

Despite differences in task demands and participants, the results of the face-identity recognition test replicated the main results of the face-exemplar recognition test. The ANOVA yielded a significant main effect of test cue [$F(2,30) = 6.18$, $MS_e = 0.07$], a nonsignificant effect of prototype view, and a nonsignificant interaction. The bottom row of Table 1 suggests that the participants had difficulty rejecting unseen exemplars in the front-view condition, but not in the 45°-left condition. We do not have a good explanation for this nonsignificant tendency. Planned contrasts indicated that the participants produced “seen face” responses more often to unseen prototypes than to unseen exemplars [$F(1,30) = 11.91$, $MS_e = 0.14$], and even than to seen exemplars [$F(1,30) = 5.30$, $MS_e = 0.06$]. Thus, there was a strong prototype effect. Since two cells in the feature-location ANOVA do not show variance at all ($M = 1.00$), nonparametric tests were also performed. These analyses confirmed the significant main effect of test cue and the nonsignificant effect of prototype view.

To summarize, Experiment 1 provided three main findings. First, the results (acquired using realistic high-quality color photographs) replicated the strong prototype effect for feature displacements found by Bruce et al. (1991) with schematic faces. Second, the prototype effect was not affected by prototype view (front vs. 45°-left), suggesting that findings obtained with a front-view prototype can be generalized to other prototype views. Third, the face-identity recognition test showed essentially the same results as the face-exemplar recognition test. This finding is consequential because the face-exemplar recognition test has been used in several prototype studies (e.g., Bruce et al., 1991; Malpass & Hughes, 1986), and there was no evidence about whether its results could be generalized to a task much more similar to that of everyday face recognition.

EXPERIMENTS 2A AND 2B

Experiments 2A and 2B investigated the hypothesis that recognition across variations occurring within the same view is based on an averaging or superposition mechanism, whereas recognition across different views is based on an approximation to stored exemplars. One prediction of this hypothesis is that the prototype effect should be more sensitive to probe-exemplar similarity in the case of head-angle variations than in the case of

Table 1
Mean Recognition Ratings
(0 = *Definitely Unseen Picture*; 9 = *Definitely Seen Picture*)
in the Face-Exemplar Recognition Test, and
Mean Recognition Responses (0 = *Unseen Face*; 1 = *Seen Face*)
in the Face-Identity Recognition Test in Experiment 1

Test Cues	Prototype View	
	45°-left	Front
Face-Exemplar Recognition		
Prototype	8.4	8.5
Seen	7.4	8.0
Unseen	6.7	6.4
Face-Identity Recognition		
Prototype	1.00	1.00
Seen	.97	.91
Unseen	.84	.97

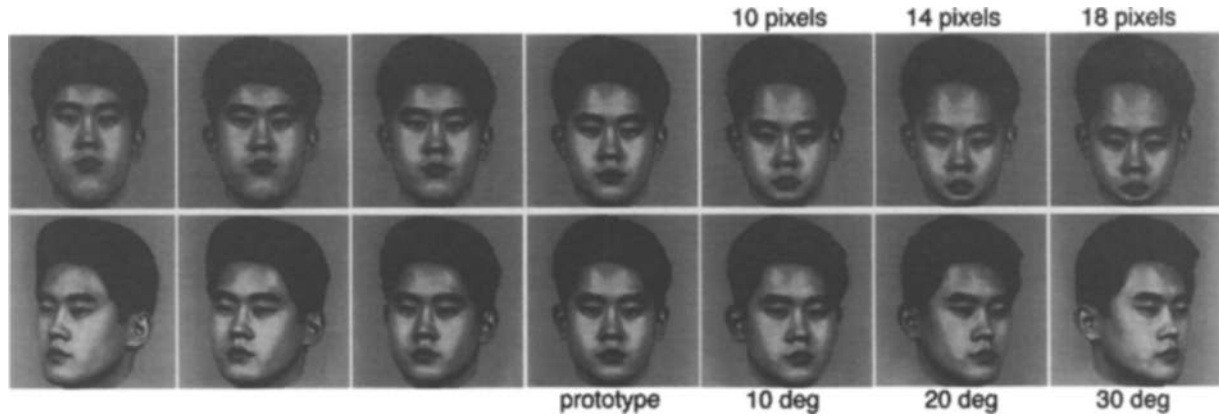


Figure 2. Examples of the materials of Experiments 2A and 2B: A prototype face (central face in both rows) and its corresponding feature-location (top row) and head-angle (bottom row) face exemplars.

feature-location variations. Experiments 2A and 2B compared the prototype effect for head-angle variations of 10°, 20°, and 30° with the prototype effect for feature-location variations of 10, 14, and 18 pixels (see Figure 2). The only difference between Experiments 2A and 2B was on the rating scale used during study. In Experiment 2A, an age-rating scale was used. Since age rating might encourage attention to feature-shift variations rather than to head-angle variations, in Experiment 2B we employed a rating task likely to encourage attention to head-angle variations rather than to feature shifts: This was a rating of gaze direction of the studied heads. The test phase used in Experiments 2A and 2B was the exemplar recognition test. Experiment 1 had shown that the exemplar recognition and identity recognition tests yielded similar results, and the former proved a more sensitive tool that required fewer additional faces for the test sequence than identity recognition.

In order to make meaningful comparisons between the storage and retrieval of feature versus viewpoint exemplar variations, we needed first to examine the similarity properties of the two scales of variations employed. Our prediction was that viewpoint variations would show more sensitivity to probe-exemplar similarity than would feature variations. Therefore, it was important to establish that values on the range of viewpoint variations we selected were not intrinsically more different from one another than values within the range of feature variations we chose. Before running Experiments 2A and 2B, two pilot studies were conducted in order to determine which sizes of feature displacement involve a similar degree of probe-exemplar similarity as head-angle variations of 10°, 20°, and 30°. In the first pilot study, participants were presented with pairs of photographs, the prototype and one exemplar (feature-location or head-angle variation), and were asked directly to rate the similarity between the two pictures. In the second pilot study, a method similar to that of the main experiments was used, with the difference that only one exemplar per face was presented at study, rather than two. With only one exemplar at study,

no prototype effect should occur, and recognition of the unseen prototype pattern should be a function of its similarity to the seen exemplar. In other words, when only one exemplar is presented, recognition becomes a measure of remembered similarity.

The two pilot studies provided converging evidence supporting the decision of matching head-angle variations of 10°, 20°, and 30° with feature-location variations of 10, 14, and 18 pixels. The first pilot study showed that 10/14/18 pixel manipulation produced a stronger effect on perceived similarity than did 10°/20°/30° manipulation and that variations of 10, 14, and 18 pixels were more noticeable than variations of 10°, 20°, and 30°, respectively. Likewise, the second pilot study showed a significant difference on remembered similarity between the 10- and 18-pixel exemplars, but not between 10° and 30° exemplars. Thus, if the effect of probe-exemplar similarity on the prototype effect were basically the same for feature-location variations and for head-angle variations, an experiment comparing 10/14/18 pixel variations with 10°/20°/30° variations should find a stronger effect of similarity for feature-location variations than for head-angle variations. Since the prediction to be tested was exactly the opposite, matching 10°, 20°, and 30° with 10, 14, and 18 pixels represented an extremely conservative choice.

Method

Participants. Thirty-six participants participated in Experiment 2A, and 72 participants participated in Experiment 2B.

Procedure. Variation type (feature-location, head-angle) was manipulated within participants, and variation size (small, medium, large) was varied between participants. The six critical faces were divided into two sets (two males and one female; one male and two females), which were counterbalanced across feature-location and head-angle conditions. The seen exemplars in small, medium, and large conditions were the 10°, 20°, and 30° exemplars, respectively, in the head-angle condition, and the 10-, 14-, and 18-pixel exemplars, respectively, in the feature-location condition (see Figure 2). The unseen exemplars were, respectively, the 20°, 30°, and 40° exemplars and the 14-, 18-, and 20-pixel exemplars. Faces were presented for 4 sec and rated on one scale. Each picture was presented (and rated) twice. The only difference between Experiment 2A and

Experiment 2B was in the rating scale used at study. In Experiment 2A, it was an age scale (1 = 18–21 years; 8 = 46–49 years), whereas in Experiment 2B, it was a gaze-direction scale (1 = “the person in the picture is looking straight at you”; 8 = “the person in the picture is looking away from you”).

In order to avoid ceiling effects, 40 study filler were also included. To prevent the participants from easily rejecting prototypes at test because no natural (i.e., unmanipulated) front views had been presented in the rating phase (though all feature-location variants were based on the front view), the natural front-view prototypes of 4 of the 20 filler faces were also incorporated in the study list. Twelve studied filler pictures (1 feature-location exemplar of each of 4 faces; 1 head-angle exemplar of each of 4 faces; 4 prototypes) were also tested. Four nonstudied filler prototypes were tested as well, so that the number of seen and unseen filler prototypes in the test was identical. The filler exemplars and prototypes at study and test were always the same in the three between-participants conditions (small, medium, large). In sum, the study list included the 12 critical pictures (6 critical faces \times 2 exemplars) plus the 44 filler pictures. The face-exemplar recognition test consisted of 18 critical pictures (prototype, seen, and unseen exemplars of the 6 critical faces) plus 12 studied filler pictures (exemplars and prototypes) and 4 nonstudied filler prototypes.

Results and Discussion

The results of Experiments 2A and 2B are shown in Table 2. Since the results of these experiments are almost identical, they are described together. Wherever two F values are mentioned, the first (F_A) corresponds to Experiment 2A and the second (F_B) corresponds to Experiment 2B.

First, a 3 (variation size) \times 2 (variation type) \times 3 (test cue) ANOVA was conducted. The main effect of variation type was reliable [$F_A(1,33) = 29.50$, $MS_e = 81.48$; $F_B(1,69) = 52.64$, $MS_e = 229.72$], reflecting higher overall ratings in the feature-location condition than in the head-angle condition. The main effect of test cue was significant [$F_A(2,66) = 13.35$, $MS_e = 23.40$; $F_B(2,138) = 18.88$, $MS_e = 53.58$], and planned contrasts indicated that seen exemplars were rated significantly higher than unseen exemplars [$F_A(1,66) = 3.81$, $MS_e = 6.67$; $p < .055$, $F_B(1,138) = 11.67$, $MS_e = 33.12$], indicating that the participants could in general distinguish between seen and unseen exemplars. The variation size \times variation type interaction was nonsignificant. The variation size \times test cue interaction was reliable [$F_A(4,66) = 11.42$, $MS_e = 20.01$; $F_B(4,138) = 7.40$, $MS_e = 21.01$], reflecting a stronger effect of variation size on the prototype than on the other test cues. The variation type \times test cue interaction was also reliable [$F_A(2,66) = 11.14$, $MS_e = 24.20$; $F_B(2,138) = 15.69$, $MS_e = 37.61$], since the difference between feature-location and head-angle conditions occurred particularly for the prototype. Finally, the critical three-way interaction between variation size, variation type, and test cue was significant [$F_A(4,66) = 4.11$, $MS_e = 8.92$; $F_B(4,138) = 2.51$, $MS_e = 6.02$].

This significant interaction between variation size, variation type, and test cue is consistent with the prediction that the prototype effect should be more affected by

Table 2
Mean Ratings
(0 = Definitely Unseen Picture; 9 = Definitely Seen Picture)
in the Face-Exemplar Recognition Test as a Function of
Variation Size (Large, Medium, and Small) and Variation Type
(Head Angle, Feature Location) in Experiments 2A and 2B

Test Cue	Head Angle			Feature Location		
	Large	Medium	Small	Large	Medium	Small
Experiment 2A						
Prototype	2.8	4.6	7.0	7.1	7.2	7.8
Seen	7.8	6.4	6.6	7.7	7.0	7.6
Unseen	6.9	5.3	7.1	7.8	6.6	6.8
Experiment 2B						
Prototype	3.3	4.3	6.2	6.4	7.1	7.9
Seen	6.9	7.1	6.4	7.0	7.4	7.5
Unseen	5.7	6.2	5.3	6.6	7.4	7.2

variation size for head-angle variations than the one for feature-location variations. However, since this interaction can also reflect variations in seen and unseen exemplars, it is necessary to confirm an interaction involving the differences between prototype and seen ratings. With this objective, a 3 (variation size) \times 2 (variation type) ANOVA was performed on the seen – prototype differences. Figure 3 shows seen – prototype differences for Experiment 2A. The main effect of variation size was significant [$F_A(2,33) = 19.30$; $F_B(2,69) = 9.32$], indicating an overall influence of variation size on the prototype effect. The main effect of variation type was also reliable [$F_A(1,33) = 17.77$; $F_B(1,69) = 28.53$], indicating a stronger prototype effect for feature-location variations than for head-angle variations. More importantly, as suggested by Figure 3, the critical interaction between variation type and variation size was significant [$F_A(2,33) = 7.50$; $F_B(2,69) = 3.4$]. This interaction corroborates the prediction that the prototype effect would be more affected by variation size for head-angle variations than for feature-location variations.

In sum, the results of Experiment 2A and 2B confirmed the prediction of a stronger effect of similarity on the prototype effect for head-angle variations than for feature-location variations. This result is consistent with the idea that stable recognition across variations occurring within the same view is based on averaging stored exemplars, whereas stability across viewpoint variations is based on an approximation to stored exemplars.

However, there are two points of concern about the results of Experiments 2A and 2B. First, it is possible to argue that the levels of head-angle variation were not matched in terms of similarity to the levels of feature-location variation. We selected these levels to work against the prediction that the prototype effect would be more dependent on similarity for head-angle variations than for feature-location variations, which was nevertheless confirmed. Yet, it could be argued that the matching is problematic. Second, it is possible to argue that recogni-

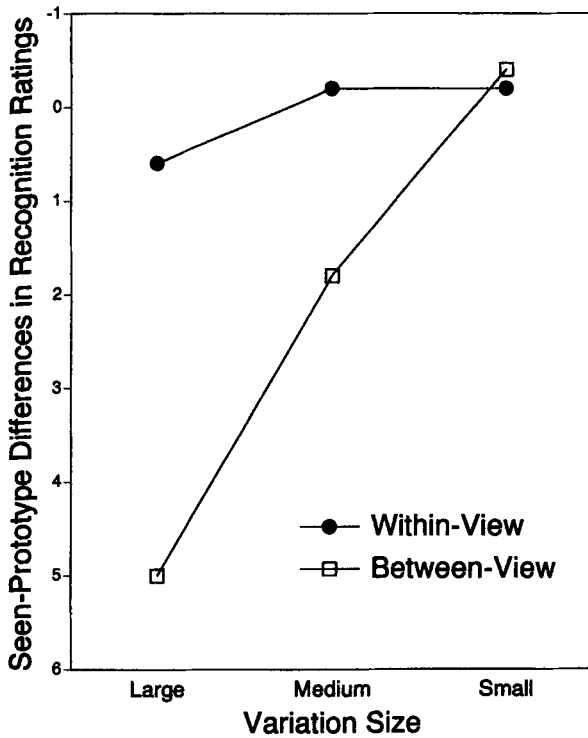


Figure 3. Experiment 2A: Seen — prototype differences in recognition ratings in the face-exemplar recognition test as a function of type (within-view = feature location; between-view = head angle) and size of face variation.

tion ratings for unseen exemplars in the feature-location condition were too high. This could be a problem, because it is conceivable that, if recognition ratings for unseen exemplars had been lower, recognition ratings for the prototype would have also been lower. It should be noted, however, that, in Experiments 2A and 2B, it was not possible to compare directly the results for the prototype and for the unseen exemplars, because their distance in pixels from the seen exemplars was quite different. In the large condition, for example, the difference between the prototype and the seen exemplars was 18 pixels, whereas the distance between seen and unseen exemplars was only 2 pixels. A direct comparison between prototype and unseen exemplars ratings would be fair only if their distance from seen exemplar was the same (e.g., a 36-pixel unseen exemplar in the large condition). However, this is difficult to achieve with the feature displacement technique because internal features cannot be displaced more than 25 pixels without approaching the borders of the face. The morphing technique employed in Experiments 3A and 3B did not have this limitation, allowing a more radical manipulation of face similarity. Given these two problems, the results of the Experiment 2A and 2B should be interpreted with caution.

A further possible point of concern about Experiment 2 is that the viewpoint variations only assessed variation

around the full-face viewpoint, and it is possible that the full-face might be a canonical or privileged viewpoint with special properties. However, we have conducted further unpublished experiments (Bruce, 1995) with variations around angled-view prototypes with similar results (see also data reported by Bruce, 1994), and, therefore, we do not think the results of this experiment are artificial in this respect.

EXPERIMENTS 3A AND 3B

Experiments 2A and 2B showed that in the case of within-view face variations, such as the feature-location manipulation, a strong tendency to recognize the unseen prototype can occur even for large changes in similarity (e.g., between 18-pixel-up and 18-pixel-down exemplars). What is the limit of this prototype effect? Does it disappear when face exemplars are so different that they can be attributed to different individuals?

In order to investigate this issue, we developed a new technique for manipulating exemplar similarity: blending, or morphing, the faces of different people (see Figure 4). In sets of six different faces (e.g., Faces A–F), one face was randomly chosen as “vertex” (Face A) and was morphed with each of the other five faces (Faces B–F), creating five composite faces (Faces 1–5), which were used as seen exemplars (studied versions of studied faces) or unseen exemplars (nonstudied versions of studied faces). The prototypes were produced by averaging groups of four exemplars. Exemplar similarity was manipulated by varying the proportion of the vertex face (e.g., Face A) and another face (e.g., Face B) in the composite face—for example, 70% and 30%, 60%, and 40%, and so forth. Due to the larger proportion of a common component (Face A), 70% composites are more similar to each other than are 60% composites, and so on (see Figure 5). A similarity level of 100% would correspond to perfect identity, and a similarity level of 0% would correspond to the average similarity between the faces of different individuals in the sample population. An identity-rating pilot study showed that the morphing manipulation is an effective method for manipulating exemplar similarity: As the proportion of the vertex face decreased, the tendency for attributing the composite faces to different individuals increased.

Experiment 3B investigated the prototype effect from a level of similarity of 60% to a level of similarity of 10%. The problem investigated was whether the prototype effect would disappear when the level of exemplar similarity approaches the one between different faces in the population. This is important to establish if we want to argue that what is being tapped in the face prototype effect is the means by which stable representations of individual faces are established. Whatever the mechanism is, it needs to lump together exemplars that are likely to belong to the same person’s face but keep distinct representations for different individuals’ faces.

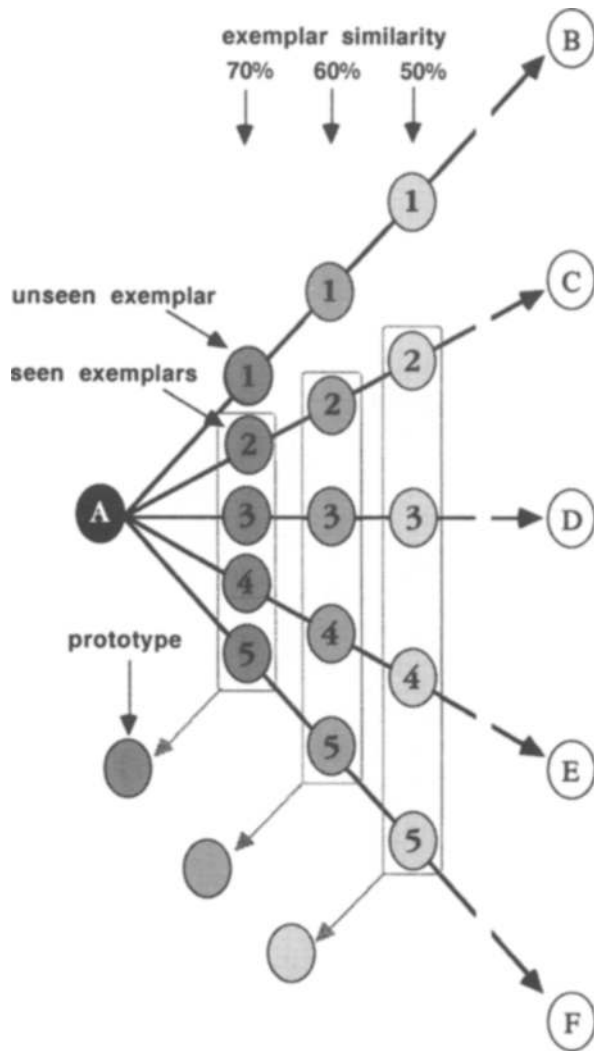


Figure 4. Schematic representation of a method used in Experiments 3A and 3B to manipulate exemplar similarity by blending a vertex face (A) with other faces (B-F) in different proportions (e.g., 70%-50%).

Before Experiment 3B, a preliminary experiment (Experiment 3A) was conducted in order to determine whether a strong prototype effect, such as that obtained with the feature-location manipulation, could be found at all using the morphing technique and whether the face-exemplar and face-identity recognition test (cf. Experiment 1) would produce similar results in the case of the morphing manipulation. Experiment 3A, the identity-rating pilot study, and Experiment 3B are reported in separate sections below.

Method of Experiment 3A

Participants and Design. Fifty students participated in the experiment. The design had two within-participants factors (exemplar similarity [70%, 60%, 50%] and test item [prototype, seen, unseen]) and one between-participants factor (test [face-exemplar recognition, 36 participants; face-identity recognition, 24 participants]).

Materials. The critical materials were constructed from the front-view pictures of 54 women and 54 men between 20 and 40 years of age ($M = 29$ years, $SD = 7$ years). The pictures were similar to the ones used in the previous experiments. The 108 pictures were divided into 18 sets (9 female, 9 male) of six faces, on the basis of two characteristics important for blending: visibility of nostrils and presence of hair on the forehead. Blending was conducted as described before (see Figure 4). First, using a computerized face processing system (Craw, Kato, Costen, & Robertson, 1994), 35 landmarks were manually placed on equivalent locations in each face (e.g., right corner of the left eye). Second, in each set, one face was randomly chosen as "vertex" and was blended with each of the other five faces at three different levels: 70%, 60%, and 50% of the vertex face. Blending (morphing) was performed by distorting the shape of both faces to the weighted average shape and obtaining the weighted average of each pixel. As illustrated in Figure 5, 70% composites were more similar to each other than 60% composites. This assumption was later supported by the results of the identity-rating pilot study in Experiment 3B. Third, the size of each composite face was reduced until the distance between the top of the head (Landmark 22) and the bottom of the chin (Landmark 29) was 200 pixels. Finally, an oval line (closed B-spline curve) was automatically fitted around each face using some of the peripheral landmarks, and everything outside the oval was covered with black.

For each of the three levels of exemplar similarity, one prototype face was created by averaging Exemplars 2-5 (see examples in Figures 4 and 5) or Exemplars 1-4. The prototypes were created using the same computer system used for creating the exemplars and were almost identical in picture quality. The exemplars included in the prototype (Exemplars 2-5 or 1-4) were used as seen exemplars, and the one not included (Exemplar 1 or Exemplar 5) was used as unseen exemplar. Thus, in each of the 18 sets of faces, there were five 70% exemplars, five 60% exemplars, five 50% exemplars, two prototypes of 70% exemplars, two prototypes of 60% exemplars, and two prototypes of 50% exemplars (total = 368 pictures). The 18 sets (9 male and 9 female) were combined into six groups of three sets each (2 female and 1 male, or 2 male and 1 female), which were counterbalanced, one group per condition, across the three conditions in the face-exemplar recognition test (70%, 60%, and 50%) or across the six conditions in the face-identity recognition test (studied, 70%, 60%, and 50%; nonstudied, 70%, 60%, and 50%).

Procedure. The procedure was very similar to that of the previous experiments. The study list included four exemplars for each of three sets in the 70%, three sets in the 60%, and three sets in the 50% condition (total = 36 pictures). For half of the participants, Exemplars 1-4 were presented at study, Exemplar 1 was the seen exemplar in the test, and Exemplar 5 was the unseen exemplar in the test. For the other half, Exemplars 2-5 were studied, Exemplar 5 was the seen exemplar, and Exemplar 1 was the unseen exemplar. Thus, the same exemplar that was tested as seen was also tested as unseen, and vice versa. Each exemplar was presented for 6 sec and was then rated on an age scale. Immediately after the end of the study phase, the participants started the face-exemplar recognition test or the face-identity recognition test. The face-exemplar recognition test included one prototype, one seen exemplar, and one unseen exemplar for each set of the nine studied sets (total = 27 pictures). The face-identity recognition test included the same pictures plus pictures equivalent to prototype, seen, and unseen exemplars for each of the nine nonstudied sets (total = 54 pictures).

Results and Discussion of Experiment 3A

The results of face-exemplar and face-identity recognition tests are shown in Table 3. As in Experiment 1, differences in RTs in the face-identity recognition test were generally nonsignificant, and, hence, only recognition



Figure 5. Examples of the materials of Experiment 3A (70%–50%) and Experiment 3B (60%–10%). Exemplar similarity decreases from left (70%) to right (10%). The prototypes in the bottom row were obtained by averaging Exemplars 1–4.

responses are reported and discussed here. Despite the differences in participants and task demands, the results of the face-exemplar recognition test and face-identity recognition test were almost identical, and, hence, they are described together. Wherever two F values are reported, the first (FE) corresponds to the face-exemplar recognition test, and the second (FI) corresponds to the face-identity recognition test.

A 3 (exemplar similarity: 70%, 60%, 50%) × 3 (test cue: prototype, seen, unseen) ANOVA yielded significant main effects of exemplar similarity [$F_E(2,70) = 26.09$, $MS_e = 69.63$; $F_1(2,46) = 9.79$, $MS_e = 0.62$] and test cue [$F_E(2,70) = 50.25$, $MS_e = 109.01$; $F_1(2,46) = 28.74$, $MS_e = 0.82$] and a reliable interaction between them [$F_E(4,140) = 2.93$, $MS_e = 6.33$; $F_1(4,92) = 5.85$, $MS_e = 0.18$]. Pairwise comparisons indicated that recognition of prototypes [$F_E(1,70) = 88.24$, $MS_e = 191.41$; $F_1(1,46) = 49.82$, $MS_e = 1.43$] and seen exemplars [$F_E(1,70) = 59.76$, $MS_e = 129.63$; $F_1(1,46) = 35.12$, $MS_e = 1.01$], was higher than recognition of unseen exemplars, with no reliable

difference between recognition of prototypes and seen exemplars. Thus, a strong prototype effect for the morphing manipulation was observed in both the face-exemplar recognition test and the face-identity recognition test. The results in Table 3 suggest that as exemplar-similarity decreased from 70% to 50%, recognition of unseen exemplars dropped more rapidly than recognition of the prototype and seen exemplars, which decreased at a similar rate. This impression was confirmed by separate ANOVAs showing that the cue type × similarity interaction was significant between unseen exemplars and seen exemplars [$F_E(2,70) = 3.15$, $MS_e = 9.16$; $F_1(2,46) = 8.80$, $MS_e = 0.27$] and between unseen exemplars and the prototype [$F_E(2,70) = 4.18$, $MS_e = 9.04$; $F_1(2,46) = 6.49$, $MS_e = 0.23$], but not between the prototype and seen exemplars. Thus, recognition of the prototype behaved similarly to recognition of seen exemplar, and both behaved differently than recognition of unseen exemplars.

In summary, Experiment 3A provided two main findings. First, the results extended previous research (Bruce

et al., 1991; Solso & McCarthy, 1981) by demonstrating a strong prototype effect for the novel face-morphing manipulation. This prototype effect provides a more direct support for the idea of a superimposing or averaging mechanism than the prototype effects obtained with other techniques. In the Identikit manipulation (Inn et al., 1993; Malpass & Hughes, 1986; Solso & McCarthy, 1981), the prototype is a combination of the most common features of seen exemplars, and, in the feature-location manipulation (Bruce et al., 1991; present Experiments 1, 2A, and 2B), the prototype is the central value of seen exemplars, whereas in the morphing manipulation, the prototype is the actual average of seen exemplars. Second, the results confirmed the conclusion of Experiment 1 that the results of the face-identity recognition test can be generalized to standard yes/no recognition tests, such as the face-exemplar recognition test.

Identity-Rating Pilot Study

Twelve participants participated in the identity rating pilot study. For the 18 sets of faces used in Experiment 3A, 40%, 30%, 20%, and 10% exemplars and prototypes were constructed with the same method employed in that experiment. The 18 sets were counterbalanced across six exemplar similarity conditions (60%, 50%, 40%, 30%, 20%, and 10%). Examples of these conditions are shown in Figure 5. In the identity-rating pilot study, two exemplars were presented one above the other and remained on the screen until the participant entered an identity rating (0 = *definitely different people*; 9 = *definitely the same person*). All 10 possible combinations of two of the five exemplars in each of the 18 sets were presented (total = 180 trials).

The average identity ratings for exemplars with a similarity between 60% and 10% are shown in Figure 6. A one-way repeated measures ANOVA showed a highly significant effect of exemplar similarity [$F(5,55) = 55.54$, $MS_e = 42.76$], indicating that the blending technique was effective in producing changes on perceived identity. All pairwise comparisons, except the one between 20% and

10%, were significant. The correlation between the morphing percentages (60%, 50%, 40%, etc.) and the participants' identity ratings was highly significant ($r = .81$, $p < .0001$). This indicates that the face-morphing manipulation is an effective method for varying the degree of similarity between face exemplars.

Method of Experiment 3B

Twenty-four students participated in Experiment 3B. The procedure of Experiment 3B was identical to the face-exemplar recognition condition of Experiment 3A, except that there were six within-participants conditions (60%, 50%, 40%, 30%, 20%, and 10%) instead of three (70%, 60%, and 50%). The same 18 sets used in the identity-rating pilot study were counterbalanced across the six exemplar-similarity conditions (60%, 50%, 40%, 30%, 20%, and 10%).

Results and Discussion of Experiment 3B

The mean recognition ratings in the face-exemplar recognition test in Experiment 3B are shown in Figure 6. An inspection of Figure 6 suggests that (1) overall, recognition ratings declined as exemplar similarity decreased, (2) there was a strong prototype effect, (3) recognition of unseen exemplars declined faster than recognition of prototypes and seen exemplars, and (4) recognition ratings for the prototype were higher or equal than those for seen exemplars (prototype effect) until a level of exemplar similarity of 20%, whereas, below this level, prototypes were rated lower than seen exemplars. Statistical analyses confirmed these four impressions. First, a 6 (exemplar similarity: 60%, 50%, 40%, 30%, 20%, 10%) \times 3 (test cue: prototype, seen, unseen) ANOVA yielded a reliable main effect of exemplar similarity [$F(5,115) = 20.38$, $MS_e = 77.44$]. Second, pairwise comparisons showed that ratings for prototypes [$F(1,46) = 132.44$, $MS_e = 409.24$] and seen exemplars [$F(1,46) = 110.51$, $MS_e = 341.48$] were both reliably higher than ratings for unseen exemplars, but not reliably different between each other. Third, the main effect of test cue was significant [$F(2,46) = 81.31$, $MS_e = 251.26$], and there was a reliable interaction between test cue and exemplar similarity [$F(10,230) = 6.60$, $MS_e = 12.83$]. Separate ANOVAs showed significant similarity \times cue interactions between unseen exemplars and prototypes [$F(5,115) = 8.13$, $MS_e = 15.20$] and between unseen exemplars and seen exemplars [$F(5,115) = 9.02$, $MS_e = 16.82$]. Finally, there was a significant similarity \times cue interaction between prototypes and seen exemplars [$F(5,115) = 3.10$, $MS_e = 6.48$], and pairwise contrasts indicated that prototypes were rated as high as seen exemplars at 60%, 50%, 30%, and 20%, higher than seen exemplars at 40% [$F(1,115) = 10.43$, $MS_e = 21.82$], and lower than seen exemplars at 10% [$F(1,115) = 4.00$, $MS_e = 8.33$].

In sum, the results Experiment 3B showed that, as exemplar similarity decreased, recognition ratings for unseen exemplars declined faster than recognition ratings for prototypes and seen exemplars and that there is a clear prototype effect up to a similarity level of 20%. At a similarity level of 10%, the prototype effect began to weaken. These results suggest that the mechanism support-

Table 3
Mean Recognition Ratings
(0 = Definitely Unseen Picture; 9 = Definitely Seen Picture)
in the Face-Exemplar Recognition Test,
and Mean Recognition Responses
(0 = Unseen Face; 1 = Seen Face)
in the Face-Identity Recognition Test

Test Cues	Exemplar Similarity		
	70%	60%	50%
Face-Exemplar Recognition			
Prototype	7.4	6.9	6.3
Seen	7.2	6.3	6.1
Unseen	6.3	4.8	3.8
Face-Identity Recognition			
Prototype	.92	.90	.82
Seen	.90	.82	.82
Unseen	.86	.69	.49

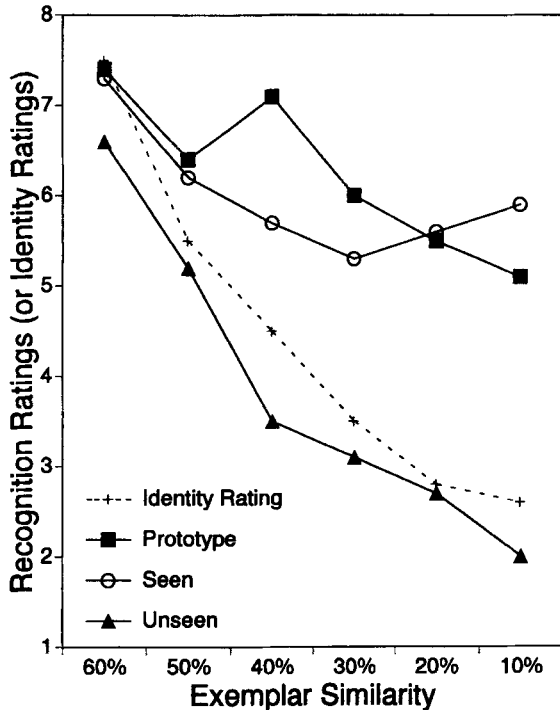


Figure 6. Experiment 3B: Ratings in the identity-rating pilot study (0 = definitely different people; 9 = definitely the same person) and in the face-exemplar recognition test (0 = definitely unseen picture; 9 = definitely seen picture).

ing the prototype effect is robust enough to resist a considerable degree of variability between seen exemplars, and, at the same time, it does not operate across faces belonging to different individuals.

GENERAL DISCUSSION

The present experiments provided four main outcomes. First, the experiments provided evidence supporting the generalizability of the prototype effect in face recognition. Second, they supplied data suggesting that stable face recognition across variations depends on different mechanisms for face variations within the same view and for face variations across viewpoints. Third, they yielded information concerning the limits of the prototype effect in face recognition. These three results are discussed in order in the following sections.

Generalizability of the Prototype Effect in Face Recognition

The present results supported the generalizability of the prototype effect on face recognition. First, in all the reported experiments, strong prototype effects were obtained with high-quality color photographs of faces. This indicates that the results obtained with Identikit (Inn et al., 1993; Malpass & Hughes, 1986; Solso & McCarthy, 1981) or line-drawn faces (Bruce et al., 1991) can be gen-

eralized to realistic face recognition conditions. Second, a similar prototype effect was obtained not only for a prototype face in the front view but also for a prototype face in a 45° view. This finding suggests that the prototype effect is not limited to the front view, for which it has been typically investigated (e.g., Bruce et al., 1991; Malpass & Hughes, 1986; Solso & McCarthy, 1981). Third, a strong prototype effect was found with a new face manipulation technique: morphing of different faces. Thus, the prototype effect occurs not only when face features are exchanged (e.g., Solso & McCarthy, 1981) or displaced (Bruce et al., 1991) but also when they are elastically altered in shape and color. Finally, a strong prototype effect was obtained not only with the face-exemplar recognition test used in previous studies (e.g., Bruce et al., 1991; Malpass & Hughes, 1986; Solso & McCarthy, 1981) but also with the face-identity recognition test. The face-identity recognition test is much more representative of real face recognition, where the task is to recognize an individual having earlier been exposed to some set of variations of that face, rather than to recognize the variations (pictures) themselves. It is important to note that the face-exemplar recognition test and the face-identity recognition test involve very different task demands: In the first task, the correct response to the prototype is negative (the prototype is an unseen picture), whereas, in the second task, it is positive (the prototype is a seen face). Thus, in the face-exemplar recognition test, the prototype effect is observed as a failure to reject an unseen picture, whereas, in the face-identity recognition test, it is detected as an improved accuracy to correctly recognize a seen face. This result confirms what has been an assumption in much of the earlier work on prototype effects in face recognition (e.g., Bruce, 1994)—that equivalent prototype effects should be found in terms of difficulty to reject the prototypical face as an unseen picture or as a facilitation to recognize it as a seen face.

Face Recognition Across Feature Variations and Across Head-Angle Variations

In Experiments 2A and 2B, feature-location exemplars tended to be fused together, even when quite different from each other, whereas head-angle exemplars tended to be treated apart, even if they were quite similar to each other. These results are consistent with the hypothesis that stable recognition across variations within the same view involves an averaging or superimposition mechanism (Bruce, 1994), whereas stable recognition across variations in head-angle involves an interpolation mechanism.

The marked effect of similarity on face recognition across views is consistent with exemplar-based recognition theories (e.g., Hintzman, 1986; Nosofsky, 1988, 1991). This model assumes that a recognition probe activates stored exemplars in parallel as a function of their similarity to the probe, and, hence, they predict that recognition of unseen views should be a function of head angle. The pronounced effect of similarity for head-angle vari-

ations is also consistent with evidence that recognition of objects in unseen views does not necessarily involve a three-dimensional model of the object, and it can be based on an approximation to a limited number of stored two-dimensional views of the object (Bülthoff and Edelman, 1992; Poggio & Edelman, 1990). Finally, the present results are consistent with evidence that face recognition is view dependent. Animal research has shown that there are cells in the superior temporal sulcus of the macaque brain that respond to specific views of the head (e.g., Perrett, Mistlin, & Chitty, 1989), and cognitive studies have demonstrated that matching of faces is impaired when there are changes in viewpoint (e.g. Bruce, Valentine, & Baddeley, 1987). Thus, people may store only some views in memory and recognize other views by means of an interpolation or approximation mechanism.

In contrast, the finding that feature-location exemplars tended to be fused together even when quite different from each other is more consistent with the notion of an averaging mechanism. An averaging mechanism would take face exemplars as input and would output a pattern formed by the mean value of each feature dimension (Posner, Goldsmith, & Welton, 1967). An alternative mechanism would be one that outputs a pattern incorporating the most frequently experienced features (Neumann, 1977). Malpass and Hughes (1986) contrasted these two possible mechanisms by varying the eyes, noses, mouths, and chins of Identikit face exemplars from lighter/finer (coded "1") to darker/coarser (coded "5"). In an asymmetric feature distribution (e.g., nine "1," one "2," five "3," one "4," and one "5"), the averaging model predicts that the output pattern will be composed of mean feature values (e.g., "3"), whereas the attribute frequency model predicts that it will be composed of modal feature values (e.g., "1"). Malpass and Hughes tended to support the attribute frequency model. However, one limitation of Malpass and Hughes's study is that an averaging mechanism is inherently unlikely when changes are only qualitative, such as shifts between different types of Identikit features (e.g., two types of mouths). The feature displacement technique, on the other hand, involves a quantitative change and has produced results (Bruce et al., 1991, and present experiments) that can be easily explained by an averaging model. Moreover, the strong prototype effect obtained with the morphing manipulation in Experiments 3A and 3B provides a direct support for the averaging model, because the pictures of the prototypes were actually produced by averaging seen exemplars.

Thus, the present results suggest that stable face recognition across variations is achieved by different mechanisms for variations within the same view and for variations in head angle. In the case of feature variations, stability seems to involve an averaging mechanism that yields a prototype effect that is strong and resistant to exemplar similarity; in the case of viewpoint variations, stability seems to involve an interpolation mechanism that yields a prototype effect that is weak and dependent on exemplar similarity.

Nevertheless, given the difficulties involved in equating feature and view variations in perceived and remembered similarity, the present results should be interpreted with caution. Moreover, it is a goal for future research to determine whether the assumption of different mechanisms is strictly necessary or whether recognition across feature and head-angle variations can be accounted for by a single memory mechanism interacting with possible differences in the manner in which different types of image are encoded. For example, a common exemplar-similarity mechanism could account for the different patterns of findings obtained across feature variations compared with viewpoint variations if more attention was paid to some kinds of variation than to others. This was one aim of the comparison between Experiments 2A and 2B, which examined how the prototype effects were influenced by instructions that drew attention to either source of variation. We found that this instructional manipulation did not alter the basic pattern of our findings, but there may be more intrinsic differences between the coding of head angle (which provides an important cue to direction of attention) and encoding of feature variation (which does not provide such social cues) that have still to be explored.

The Limits of the Prototype Effect

The results of Experiment 3B suggest that when exemplar similarity decreases to the level of similarity between different faces in the population, the prototype effect starts to disappear. If we assume that the function of the prototype effect is to facilitate personal identification across variations, then this result makes good sense, because this function would be partially lost if the prototype effect occurred not only across different versions of the same face but also across the faces of different individuals.

At the same time, the results of Experiment 3B suggest that a face recognition mechanism based on the prototype effect can fail under certain conditions. Even though the prototype effect formally disappeared at the 10% level (recognition ratings for prototype were significantly lower than those for seen exemplars), at this level, the tendency to recognize the unseen prototype as a seen picture was still quite strong. Since face exemplars at the 10% level were generally attributed to different individuals in the identity-rating pilot study, this result indicates that, under certain circumstances, we may falsely recognize a face we had never seen just because it is close to the average of a subset of faces we had seen in the past. Indeed, this "false memory" for faces that resemble many others in the population appears to provide one of the strong contributions to the effects of facial distinctiveness in recognition memory (e.g., see Bartlett, Hurry, & Thorley, 1984; Bruce, Burton, & Dench, 1994; Light, Kayra-Stuart, & Hollander, 1979; Vokey & Read, 1992).

This false memory illusion resembles one that is now under intense scrutiny in the domain of verbal memory: memory for words not presented in lists. In the standard paradigm (Deese, 1959; Roediger & McDermott, 1995),

participants hear a list of words (e.g., *thread, pin, eye, sewing, sharp, point, haystack, pain, injection*, etc.) that are all highly associated with a word that is not presented (e.g., *needle*). In subsequent tests, they display a strong tendency to falsely recall or recognize the nonpresented common associate (e.g., Roediger and McDermott, 1995). This memory illusion is very similar to the prototype effect investigated here, with the difference that it occurs at a semantic level rather than at a perceptual level. Yet, the phenomenon of memory for nonstudied words has been recently demonstrated for lists of phonologically similar words (Schacter, Verfaellie, & Anes, 1997), and, hence, the semantic/perceptual dimension is not enough to differentiate the two phenomena.

Summary

In summary, the present results pointed out extensions of and the limits to the phenomenon of the prototype effect in face recognition. It occurs for realistic photographs of faces in different views, both as difficulty to reject the prototype as an unseen picture and as facilitation to recognize it as a seen face. It occurs not only for feature exchanges and displacements but also for plastic morphing transformations. On the other hand, in the case of view variations, it is weak and similarity-dependent, probably reflecting a different recognition mechanism. Finally, it tends to not operate across the faces of different people; however, under certain conditions, it may originate memory for faces that were never seen.

REFERENCES

BARTLETT, J. C., HURRY, S., & THORLEY, W. (1984). Typicality and familiarity of faces. *Memory & Cognition*, *12*, 219-228.
 BRUCE, V. (1994). Stability from variation: The case of face recognition. The M. D. Vernon Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *47A*, 5-28.
 BRUCE, V. (1995). [Prototype formation across head-angle variations]. Unpublished raw data.
 BRUCE, V., BURTON, A. M., & DENCH, N. (1994). What's distinctive about a distinctive face? *Quarterly Journal of Experimental Psychology*, *47A*, 119-141.
 BRUCE, V., COWEY, A., ELLIS, A. W., & PERRETT, D. I. (1992). *Processing the facial image*. Oxford: Oxford University Press, Clarendon Press.
 BRUCE, V., DOYLE, T., DENCH, N., & BURTON, M. (1991). Remembering facial configurations. *Cognition*, *38*, 109-144.
 BRUCE, V., VALENTINE, T., & BADDELEY, A. (1987). The basis for the 3/4 view advantage in face recognition. *Applied Cognitive Psychology*, *1*, 109-120.
 BÜLTHOFF, H., & EDELMAN, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, *89*, 60-64.
 CRAW, I., KATO, T., COSTEN, N., & ROBERTSON, G. (1994). *Methods for improving principal component analysis coding of faces for recognition: A testbed* (TR-H No. 104). Kyoto: ATR Human Information Processing Research Laboratories.
 DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17-22.

FRANKS, J. J., & BRANSFORD, J. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology*, *90*, 65-74.
 HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.
 HOMA, D., CROSS, J., CORNELL, D., GOLDMAN, D., & SCHWARTZ, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology*, *101*, 116-122.
 HOMA, D., DUNBAR, S., & NOHRE, L. (1991). Instance frequency, categorization, and the modulating effect of experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 444-458.
 HOMA, D., GOLDFHARDT, B., BURRUEL-HOMA, L., & CARSON SMITH, J. (1993). Influence of manipulated category knowledge on prototype classification and recognition. *Memory & Cognition*, *21*, 529-538.
 HOMA, D., & VOSBURGH, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 322-330.
 INN, D., WALDEN, K. J., & SOLSO, R. L. (1993). Facial prototype formation in children. *Bulletin of the Psychonomic Society*, *31*, 197-200.
 LIGHT, L. L., KAYRA-STUART, F. & HOLLENDER, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning & Memory*, *5*, 212-228.
 MALPASS, R. S., & HUGHES, K. D. (1986). Formation of facial prototypes. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. Young (Eds.), *Aspects of face processing* (pp. 154-162). Dordrecht: Martinus Nijhoff.
 NEUMANN, P. G. (1977). Visual prototype formation with discontinuous representation of dimensions of variability. *Memory & Cognition*, *5*, 187-197.
 NOSOFSKY, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 700-708.
 NOSOFSKY, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 3-27.
 PERRETT, D. I., MISTLIN, A. J., & CHITTY, A. J. (1989). Visual neurones responsive to faces. *Trends in Neurosciences*, *10*, 358-384.
 POGGIO, T., & EDELMAN, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*, 263-266.
 POSNER, M. I., GOLDSMITH, R., & WELTON, K. E., JR. (1967). Perceived distance and classification of distorted patterns. *Journal of Experimental Psychology*, *77*, 353-363.
 POSNER, M. I., & KEELE, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *83*, 304-308.
 ROEDIGER, H. L., III, & MCDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 803-814.
 SCHACTER, D. L., VERFAELLIE, M., & ANES, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology*, *11*, 331-342.
 SOLSO, R. L., HECK, M., & MEARNES, C. (1993). Prototype formation in very short-term memory. *Bulletin of the Psychonomic Society*, *31*, 185-188.
 SOLSO, R. L., & MCCARTHY, J. E. (1981). Prototype formation of faces: A case study of pseudo-memory. *British Journal of Psychology*, *72*, 499-503.
 SOLSO, R. L., & RAYNIS, S. A. (1979). Prototype formation of imaged, kinesthetically and visually presented geometric forms. *Journal of Experimental Psychology: Human Perception & Performance*, *5*, 701-712.
 VOKEY, J. R., & READ, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition*, *20*, 291-302.

(Manuscript received May 28, 1997;
 revision accepted for publication October 23, 1997.)