What's in the name? Categorical perception for unfamiliar faces can occur through labeling

M. KIKUTANI, D. ROBERSON, AND J. R. HANLEY
University of Essex, Colchester, England

The conditions under which categorical perception (CP) occurs for unfamiliar faces are unclear. Although CP is generally found only for familiar faces, it has been reported for unfamiliar faces after brief training (Levin & Beale, 2000) or even without any learning of the original faces (Campanella, Hanoteau, Seron, Joassin, & Bruyer, 2003). Three experiments investigated whether CP can be observed for an unfamiliar morphed face continuum without preexposure to the endpoints of the continuum (Experiment 1); with brief exposure to the endpoints (Experiment 2); or with exposure to named endpoints (Experiment 3). CP was always observed for matched pairs of famous faces. However, CP for unfamiliar faces was induced only when participants observed names paired with the endpoint faces before the start of the experiment. The results suggest that CP effects for unfamiliar faces can be observed extremely rapidly when clear category labels are presented.

Categorical perception (CP) occurs when a continuous variation along a perceptual dimension comes to be perceived as a series of discrete regions separated by sharp boundaries between labeled categories (Harnad, 1987). Discrimination of two very similar stimuli becomes more accurate when they straddle a category boundary than when they fall toward the center of either category (Goldstone, 1994). Researchers investigating face processing have found CP for facial expressions of emotion (Calder, Young, Perrett, Etcoff, & Rowland, 1996), race (Levin & Angelone, 2002), and gender (Bülthoff & Newell, 2004), and for the identity of well-known faces (Beale & Keil, 1995; Stevenage, 1998; Levin & Beale, 2000). In the present study, we investigate the extent to which CP can be observed for the identification of faces that were unfamiliar before the experiment.

Beale and Keil (1995) first reported CP between two familiar faces in two slightly different tasks. They used an AB–X discrimination task in which A and B were morphs differing by 20% from each other, presented sequentially, followed by a target X that was identical to either A or B. Participants selected which morph (A or B) was identical to X. They found superior discrimination accuracy when A and B straddled the category boundary (e.g., 60% John F. Kennedy/40% Bill Clinton, paired with 40% Kennedy/60% Clinton) than when they were both from the same identity category (e.g., 80% Kennedy/20% Clinton vs. 60% Kennedy/40% Clinton). They found similar results when using a “better-likeness” task in which paired morphs were presented simultaneously followed by a question: “Which looks more like Kennedy?” They also compared the results for familiar faces with morphed continua generated from unknown faces. Only familiar

continua showed CP, suggesting that, for unfamiliar faces, no category representation had been established, so all pairs of morphs were equally difficult to discriminate.

Other researchers, however, have found evidence of CP with preexperimentally unfamiliar faces (McKone, Martini, & Nakayama, 2001; Stevenage, 1998; Viviani, Binda, & Borsato, 2007). Stevenage used photographs of two identical twins that participants were initially unable to discriminate. Participants then rated photographs of the same twin and photographs of the different twins for similarity. Next, participants were trained to associate names to each of the faces, and then the rating task was repeated. Photographs of the same twin were rated as more similar than they had been before training, whereas photographs of the different twins were rated as more different. Thus, appropriate training can induce CP even for two almost identical, unfamiliar faces. However, 60–70 training trials lasting around 30 min were required to learn the names of the faces. Viviani et al. reported the development of CP for unfamiliar faces after around 300 trials of an AB–X recognition task, even for continua where both the identity and the gender varied. McKone et al.'s CP effect occurred after participants had been trained for between 1,000 and 10,000 trials.

Levin and Beale (2000) reported CP for unfamiliar faces after much shorter preexposure to the faces before the experimental task began. During preexposure, two unfamiliar faces were presented with arbitrarily allocated names (e.g., “Mike” or “Paul”) and participants were asked to study them until they could confidently associate the names with the faces. On average, an observation of the faces and names for only 40 sec successfully induced a discrimination advantage for between-category over within-category pairs in a better-

D. Roberson, robedd@essex.ac.uk

likeness task. This finding suggests that a long familiarization period (Stevenage, 1998) is not necessary to induce CP for pictures of unfamiliar faces.

Only two studies have reported CP for unfamiliar faces without prolonged prior exposure to the faces. Levin and Angelone (2002) found an effect, but only for a continuum between black and white faces, not for same-race continua. Participants might have shown CP for a cross-race continuum because of shading or shape differences, rather than identity per se. Campanella, Hanoteau, Seron, Joassin, and Bruyer (2003) did find an effect without manipulating race, although even here the magnitude of CP effect was larger for familiar faces. In a delayed same-different matching task, participants showed significantly better between-category than within-category discrimination for unfamiliar face pairs. However, Campanella et al. did not control the distinctiveness of the faces. Angeli, Davidoff, and Valentine (2008) suggested that Campanella et al.'s CP effect may have arisen as an artifact because, when the distinctiveness of the faces varies, the perceptual space is larger between morphs that are close to the distinctive face. Thus, the physical difference may have been greater in the between-category discrimination trials than in the within-category trials.

In the present study, we investigated whether CP for unfamiliar face identification is linked to the learning of verbal labels for each face. An explanation of CP that requires activation of verbal labels has been proposed in other domains (Pilling, Wiggett, Özgen, & Davies, 2003; Roberson & Davidoff, 2000). This account suggests that between-category pairs of stimuli are discriminated more easily because they activate different verbal labels, whereas the members of a within-category pair activate the same label even if they are physically different. Consistent with this account, CP is eliminated if a secondary verbal interference task is undertaken that prevents participants from naming each stimulus (Roberson & Davidoff, 2000; Winawer et al., 2007). In both Stevenage (1998) and Levin and Beale (2000), names were allocated to the faces during familiarization. In a more recent study by Viviani et al. (2007), endpoint faces were labeled as "Face A" and "Face B" in a study phase that preceded a categorization task. Thus, it remains unclear whether CP resulted from the establishment of a facial representation that is activated independently of access to a person's name (or category label), or whether CP requires a learned association between the facial representation and the label.

Three experiments used continua of facial images that were well controlled for distinctiveness. An X-AB two-alternative forced choice paradigm was used for all experiments. Participants saw a target stimulus, followed by two side-by-side test stimuli, and had to report which stimulus matched the target. The same two sets of morphed face continua (familiar and unfamiliar) were used in all experiments. Experiment 1 was a baseline study to investigate whether CP occurs for unfamiliar as well as for familiar faces without prior exposure to the endpoint faces. We then investigated whether verbal labeling is essential for CP. In Experiment 2, the unlabeled endpoint faces of both continua were presented for participants to study for 1 min before the discrimination task. If new categories can be

generated quickly from the endpoints and CP can occur from perceptual learning, the unfamiliar face continuum should elicit CP. In Experiment 3, the endpoint faces were presented with associated names ("George," "Luke") for participants to study for 1 min before the discrimination task. If CP for faces requires long-term perceptual learning (Beale & Keil, 1995; Stevenage, 1998; Viviani et al., 2007), better discrimination of between-category than of within-category pairs should occur only for familiar face continua in all three experiments. If CP for faces can be established with no preexposure (Campanella et al., 2003), then it should be observed for unfamiliar faces in all three experiments. If face categories can be established from rapid perceptual learning in less than a 1-min exposure (Levin & Beale, 2000), then it should be observed for both familiar and unfamiliar face continua in Experiments 2 and 3. If, however, it is critically dependent on the association of names to faces, it should be found for only familiar faces in Experiments 1 and 2, but for both familiar and unfamiliar continua in Experiment 3.

PRETESTS OF MATERIALS

Distinctiveness Ratings

Thirty-five famous faces and 34 unfamiliar faces were rated for distinctiveness on a 7-point scale (1 = *least distinctive*, 7 = *most distinctive*) by 120 undergraduates from the University of Essex. All faces were of white males. All images were cut below the chin, resized to 250 × 250 pixels, and printed in grayscale on white paper. The two faces rated least distinctive were chosen from both groups. We also calculated, for each face, a *z* score based on the overall distribution of distinctiveness judgments for the whole set of faces. The famous face pair consisted of Matt Damon (mean distinctiveness = 3.7, *z* = -1.76) and Antonio Banderas (mean distinctiveness = 3.8, *z* = -1.55). The unfamiliar face pair consisted of face L (mean distinctiveness = 2.8, *z* = -1.45) and face G (mean distinctiveness = 3.1, *z* = -1.05). Thus, the difference in *z* scores was 0.21 for famous faces and 0.40 for unfamiliar faces.

Materials

The four images were frontal views edited to remove ears and most of the hair. All faces had neutral expressions. They were blended (using Morphue version 1.85) to create a familiar and an unfamiliar continuum. Hue, brightness, and contrast levels of the morphs were adjusted to minimize the difference within each continuum. Intervals of 10% between the endpoint images were applied to create nine intermediate morph images for each continuum (10/90, 20/80, 30/70, 40/60, 50/50, 60/40, 70/30, 80/20, and 90/10).

Classification. A classification task (Levin & Beale, 2000) was carried out to identify the category boundary for each of the two continua. Thirty-two undergraduates from the University of Essex (26 female, 6 male, mean age = 22.4) completed this task. The two endpoint faces were presented side by side at the top of the computer screen, and a morphed image was presented centrally below them. Participants indicated by a keypress which endpoint (left or right) the morph most resembled. Each of the nine morphed images was presented three times, giving 27 trials per block. The endpoint faces remained on the screen throughout the classification task, so there was no memory component to this task. Trials within the block were randomized across participants and block order (familiar/unfamiliar) and the positions of the endpoint faces were counterbalanced. Figures 1 and 2 show the mean classification of each morph from Banderas to Damon and from G to L.

Pairs of morphs from each continuum were selected with a 20% interval between members of the pair. Following Beale and Keil (1995),

Figure 1. Classification data for famous face continuum (percentage that each morph was classified as Banderas).

between-category pairs were chosen so that one member of the pair was classified with one endpoint on fewer than 33% of trials, whereas the other morph was classified with the same endpoint on more than 66% of trials. Only the 60%–40% pair met the criterion, so additional between-category pairs were created with morphs 35%–55% and 45%–65%. The 10%–30%, 20%–40%, 60%–80%, and 70%–90% pairs were defined as within-category pairs. The Appendix shows examples of the within- and between-category morphs.

EXPERIMENT 1

This experiment investigated whether CP could occur for a morphed face continuum generated from two unfamiliar faces without any prior exposure to the faces, when the distinctiveness of the faces for both the familiar and unfamiliar continua was matched.

Method

Participants. Twenty undergraduates participated (3 male, 17 female; mean age = 23.2 years).

Procedure. An X–AB match-to-sample design was used (Bülthoff & Newell, 2004). Morphs A and B always differed from each other by 20%, and X was always identical to either A or B. Following a 750-msec fixation cross, the target morph (X) was shown for 1,000 msec on the center of the screen, followed by a blank screen for 1,000 msec. The paired morphs (A and B) were then shown for 2,000 msec or until a response was made, followed by a 500-msec intertrial interval. Participants were asked to indicate by keypress as quickly and accurately as possible which test morph (A or B) was identical to the target. After five practice trials, each pair was presented four times so that X was identical to A and B twice. The positions of Morphs A and B were counterbalanced. Nine pairs were presented, giving 36 trials in a block for each continuum. Trials were randomized within blocks, and the order of presentation of blocks was randomized.

Results and Discussion

An alpha level of .05 was adopted for all experiments. Figure 3 illustrates the mean percentage accuracy for each pair of familiar and unfamiliar faces. Mean accuracy for familiar faces was 67.5% for within-category pairs and 82.1% for between-category pairs. Mean accuracy for un-

Figure 2. Classification data for unfamiliar face continuum (percentage that each morph was classified as G).

Figure 3. Mean percentage of correct responses for within-category pairs (W) and between-category pairs (B) of familiar and unfamiliar faces in Experiment 1 with standard error bars.

familiar faces was 70.6% for within-category pairs and 68.3% for between-category pairs.

A 2 (face continuum: familiar vs. unfamiliar) $\times$ 2 (pair type: between vs. within) within-subjects ANOVA on mean correct responses revealed a significant main effect of familiarity [$F(1,19) = 5.12$, $MS_e = 110.226$] and of pair type [$F(1,19) = 6.926$, $MS_e = 109.062$], as well as a significant interaction [$F(1,19) = 10.044$, $MS_e = 141.736$]. Examination of the simple main effects showed that between-category pairs were discriminated significantly better than within-category pairs for familiar faces [$F(1,19) = 19.498$, $MS_e = 109.062$] but not for unfamiliar faces [$F(1,19) = 0.481$, $MS_e = 109.062$]. There was also better discrimination of familiar than of unfamiliar face pairs for between-category pairs [$F(1,19) = 17.150$, $MS_e = 110.226$] but not for within-category pairs [$F(1,19) = 0.886$, $MS_e = 110.226$].

Between-category pairs were discriminated significantly better than the within-category pairs only in the familiar face continuum. The results were consistent with the pattern reported by Beale and Keil (1995), but differed from those reported by Campanella et al. (2003). Without preexposure to the endpoint faces, CP was not observed for unfamiliar faces when the distinctiveness of the endpoint faces was matched.

It might have been the case that CP was observed only for familiar faces because, even though the distinctiveness of the chosen pairs of faces was closely matched, the unfamiliar face pair was the most typical of a large set and might have been particularly difficult to distinguish to the degree necessary for CP. To control for this possibility, we repeated the AB-X task using two highly distinctive unfamiliar faces, with a group of 20 participants who had not taken part in Experiment 1. Mean distinctiveness ratings of the two faces were 4.72 ($z = 1.48$) and 5.28 ($z = 2.31$), respectively. Mean accuracy was 74.06 for within-category pairs, and 69.17 for between-category pairs. This

difference was not significant [$t(19) = 1.36$, $p > .1$]. Even with much more distinctive faces, CP was not found for unfamiliar faces.

EXPERIMENT 2

In Experiment 2, the endpoint faces used in Experiment 1 were presented for 60 sec, so participants had the opportunity to become familiar with the endpoint faces before the discrimination task (Levin & Beale, 2000). If CP can result from a minimal period of perceptual learning without verbal coding, it should be observed for both continua in this experiment.

Method

Participants. Twenty-two undergraduates (8 male, 14 female; mean age = 19.4 years) participated.

Materials and Procedure. The materials and procedure were identical to those of Experiment 1 with one addition: At the beginning of each block of trials, the two appropriate endpoint faces were presented side by side for 1 min. The positions of the images were counterbalanced across participants. Participants were instructed to learn the two faces, which, they were told, would be blended into a series of images for the discrimination task. The X-AB discrimination task was identical to that used in Experiment 1.

Results and Discussion

Figure 4 illustrates the mean percentage accuracy for each pair of familiar and unfamiliar faces. Mean accuracy for familiar faces was 68.5% for within-category pairs and 79% for between-category pairs. Mean accuracy for unfamiliar faces was 66.7% for within-category pairs and 66.6% for between-category pairs.

A 2 $\times$ 2 ANOVA showed a significant main effect of familiarity [$F(1,20) = 9.175$, $MS_e = 113.374$] and a significant interaction [$F(1,20) = 4.929$, $MS_e = 118.052$]. The effect of pair type was not significant [$F(1,20) = 3.773$, $MS_e = 153.508$]. Tests of simple main effect showed

Figure 4. Mean percentage of correct responses for within-category pairs and between-category pairs of familiar and unfamiliar faces in Experiment 2 with standard error bars.

that between-category pairs were discriminated significantly better than within-category pairs for familiar faces [$F(1,20) = 7.564$, $MS_e = 153.508$], but not for unfamiliar faces [$F(1,20) < 1$, $MS_e = 153.508$]. There was also better discrimination of famous faces than of unfamiliar faces for between-category pairs [$F(1,20) = 14.016$, $MS_e = 113.374$] but not for within-category pairs [$F(1,20) = 0.291$, $MS_e = 113.374$].

The results showed a pattern identical to that found in Experiment 1. Exposure to the endpoint faces from which the unfamiliar face continuum was created did not produce a performance on the between-category task that was superior to the within-category task. Therefore, merely presenting the unfamiliar faces without category labels for 1 min does not produce CP. In Experiment 3, we investigated whether CP would be induced if the endpoint faces were accompanied by their names at preexposure.

EXPERIMENT 3

Method

Participants. Twenty-two undergraduates participated (5 male, 16 female; mean age = 19.8 years).

Procedure. The experimental procedure was identical to that of Experiment 2, except that names were presented underneath each endpoint face at study. Familiar faces were always presented with their first names (Matt or Antonio); the unfamiliar faces were presented with the names Luke and George and were matched for frequency and for number of syllables. The allocation of names to the unfamiliar faces was counterbalanced. The X-AB task was identical to that used in Experiments 1 and 2.

Results and Discussion

Mean percentage accuracy for each pair of familiar and unfamiliar faces is shown in Figure 5. Mean accuracy for familiar faces was 67.5% for within-category pairs and

84.5% for between-category pairs. Mean accuracy for unfamiliar faces was 62.3% for within-category pairs and 70.8% for between-category pairs.

A 2×2 ANOVA showed significant main effects of familiarity [$F(1,21) = 18.35$, $MS_e = 105.585$] and pair type [$F(1,21) = 23.145$, $MS_e = 154.193$]. The interaction was not significant [$F(1,21) = 2.975$, $MS_e = 133.753$]. Examination of the simple main effects revealed superior cross-category discrimination for both familiar [$F(1,21) = 20.59$, $MS_e = 1,154.193$] and unfamiliar [$F(1,21) = 5.13$, $MS_e = 154.193$] faces. Unlike in Experiments 1 and 2, here the unfamiliar face continuum was also perceived categorically after the endpoint faces had been presented with names. A 60-sec exposure was enough for participants to learn the new face-name pairings. This experiment replicated the study of Levin and Beale (2000) and suggests that online learning of new faces is possible, but CP occurs only if they are associated with verbal labels.

GENERAL DISCUSSION

In this study, we found no evidence of CP for completely unfamiliar faces when the distinctiveness of the endpoint faces was strictly controlled (Experiment 1). We found this result both with two highly typical and with two highly distinctive faces and in spite of a larger z -score difference in distinctiveness between the pair of faces than for the familiar pair (for which CP was found). This result supports the conclusion of Angeli et al. (2008) that previous reports of CP for unfamiliar faces without preexposure (Campanella et al., 2003) may have resulted from unequal distinctiveness of the endpoint faces.

In Experiment 2, a 60-sec observation of the endpoint faces did not induce CP between two previously unfamiliar

Figure 5. Mean percentage of correct responses for within-category pairs and between-category pairs of familiar and unfamiliar faces in Experiment 3 with standard error bars.

faces. Might it be reasonably argued that 60 sec is insufficient for perceptual learning to take place? It is well known that matching internal features is faster for familiar faces, whereas external feature matching is faster for unfamiliar faces (Ellis, Shepherd, & Davies, 1979). Bonner, Burton, and Bruce (2003) have shown that after a 30-sec exposure, internal face matching becomes significantly faster and external matching becomes slower than at pretest. So a 60-sec observation of the two endpoint faces should have induced some perceptual learning of the two unfamiliar faces (Bonner et al., 2003; Clutterbuck & Johnston, 2005) in Experiment 2. Our results therefore suggest that perceptual familiarity alone may be insufficient for CP and argue against the hypothesis that CP is generally a low-level perceptual phenomenon, at least in the domain of face identity.

In Experiment 3, the pairing of to-be-learned faces with arbitrarily assigned labels successfully induced CP with 60 sec of observation. The results of Experiments 2 and 3 suggest strongly that between-category pairs of unfamiliar face morphs are easier to distinguish than within-category pairs only when the category labels are learned with the endpoint images. The association of names to the endpoint images may help to delineate the boundary between the two categories.

It is possible that providing names in the inspection phase of Experiment 3 induced participants to pay more attention to the labeled faces than the participants in Experiment 2 did (even though both sets of participants received the same instructions to study the faces), or simply motivated the participants more. This seems unlikely, however, because within-category discrimination was superior in Experiment 2.

In the AB-X paradigm, when a participant compares the test stimuli and a stored representation of the target,

if one test item is both physically different from the target and has a different label, it can be swiftly rejected at both the exemplar and the category levels. However, where both test stimuli have the same label as the target, the correct selection can be made only at the exemplar level. In this case, however, exemplar-level comparison (different) and category-level comparison (same) conflict, making performance more error prone. If separate subsystems process exemplar-level information and category-level information in parallel (Marsolek, 2004), then performance should be better for cross-category judgments than for within-category judgments as soon as a clear category boundary is marked. Labeling the endpoint faces appears to facilitate this, even within the time frame examined here.

However, Angeli et al. (2008) did not find CP for unfamiliar faces in a better-likeness task, despite the fact that participants studied the endpoint faces with names before the experiment. This difference might be the result of differential task demands, because the better-likeness task requires participants to explicitly relate each presented pair of stimuli back to the studied endpoints ("Which one looks most like Jamie?"), whereas the AB-X task used here requires participants only to match each target morph to one of two test stimuli. However, Angeli et al.'s study used a between-subjects design, testing participants in different countries on a single continuum of faces, so it is also possible that the different populations employed different response criteria. Furthermore, their design contained only one pair that straddled the category boundary. Thus, it was possible to make a within- versus between-category comparison only for this single pair against all within-category pairs. In our design, two additional between-category pairs were created (see above). This,

combined with the within-subjects design, may have increased the experimental power of our paradigm.

Viviani et al. (2007) suggested that the lack of a CP effect for unfamiliar faces in previous experiments arose because mixing discrimination trials from different continua "might have been detrimental for establishing a robust memory trace" (p. 457). In the present study, Experiments 1 and 2 did not find CP for unfamiliar faces, even with only one set of familiar faces and one set of unfamiliar faces. Nor is it likely that the failure to find CP resulted because the unfamiliar faces tested were either more similar to each other or more typical than the familiar faces, because similar results were found for a pair of highly distinctive faces, both in this experiment and by Angeli et al. (2008).

Experiments that have previously reported CP for unfamiliar faces have found it to be the result of a gradual learning process, across the course of hundreds or thousands of trials (Stevenage, 1998; Viviani et al., 2007). Moreover, in the experiments of Viviani et al., the same participants also carried out the classification task that established the boundary before participating in experimental tasks, so additional learning time was available, although performance on the classification task did not predict the locus of CP effects. Viviani et al. therefore argued against the view that perceptual learning modifies the representational space in any immutable fashion, since they found the location of the category boundary to be modified by the range of stimuli presented in the course of their experiments. They suggested, instead, that the range of the distribution of facial features in any particular experimental session determines where the boundary will be marked. The present data do not address the possibility of longer term perceptual learning. They do, however, demonstrate that CP effects can be induced very rapidly if the endpoints of two categories are labeled, even if neither face is particularly distinctive.

In conclusion, these experiments showed that, when the distinctiveness of the endpoint faces is matched, CP for face identity can be induced rapidly by verbal coding. However, the experiments reported here, like those of other researchers (Angeli et al., 2008; Beale & Keil, 1995; Levin & Beale, 2000) used only one image of each face during training. It would be interesting to investigate whether names would be necessary for CP if faces were learned in a more realistic way, such as by observing images from different views or moving images of a face.

AUTHOR NOTE

We are grateful to Dan Levin and two anonymous reviewers for their comments on an earlier draft of the manuscript. Correspondence should be addressed to D. Roberson, Department of Psychology, University of

Essex, Wivenhoe Park, Colchester CO4 3SQ, England (e-mail: robedd@essex.ac.uk).

REFERENCES

- ANGELI, A., DAVIDOFF, J., & VALENTINE, T. (2008). Face familiarity, distinctiveness, and categorical perception. *Quarterly Journal of Experimental Psychology*, **61**, 690-707.
- BEALE, J. M., & KEIL, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, **57**, 217-239.
- BONNER, L., BURTON, A. M., & BRUCE, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, **10**, 527-536.
- BÜLTHOFF, I., & NEWELL, F. N. (2004). Categorical perception of sex occurs in familiar but not unfamiliar faces. *Visual Cognition*, **11**, 823-855.
- CALDER, A. J., YOUNG, A. W., PERRETT, D. I., ETTOFF, N. L., & ROWLAND, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, **3**, 81-117.
- CAMPANELLA, S., HANOTEAU, C., SERON, X., JOASSIN, F., & BRUYER, R. (2003). Categorical perception of unfamiliar facial identities, the face-space metaphor, and the morphing technique. *Visual Cognition*, **10**, 129-156.
- CLUTTERBUCK, R., & JOHNSTON, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology*, **17**, 97-116.
- ELLIS, H. D., SHEPHERD, J. W., & DAVIES, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, **8**, 431-439.
- GOLDSTONE, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, **123**, 178-200.
- HARNAD, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 1-25). Cambridge: Cambridge University Press.
- LEVIN, D. T., & ANGELONE, B. L. (2002). Categorical perception of race. *Perception*, **31**, 567-578.
- LEVIN, D. T., & BEALE, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception & Psychophysics*, **62**, 386-401.
- MARSOLEK, C. [J.] (2004). Abstractionist versus exemplar-based theories of visual word priming: A subsystems resolution. *Quarterly Journal of Experimental Psychology*, **57A**, 1233-1259.
- MCKONE, E., MARTINI, P., & NAKAYAMA, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 573-599.
- PILLING, M., WIGGETT, A., ÖZGEN, E., & DAVIES, I. R. L. (2003). Is color "categorical perception" really perceptual? *Memory & Cognition*, **31**, 538-551.
- ROBERSON, D., & DAVIDOFF, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition*, **28**, 977-986.
- STEVENAGE, S. V. (1998). Which twin are you? A demonstration of induced categorical perception of identical twin faces. *British Journal of Psychology*, **89**, 39-57.
- VIVIANI, P., BINDA, P., & BORSATO, T. (2007). Categorical perception of newly learned faces. *Visual Cognition*, **15**, 420-467.
- WINAWER, J., WITTHOFT, N., FRANK, M. C., WU, L., WADE, A. R., & BORODITSKY, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, **104**, 7780-7785.

APPENDIX
Examples of Within- and Between-Category Familiar
and Unfamiliar Faces Used in All Three Experiments

Familiar Within-Category Pair

Banderas 90%

Banderas 70%

Familiar Between-Category Pair

Banderas 60%

Damon 60%

Unfamiliar Within-Category Pair

G90%

G70%

Unfamiliar Between-Category Pair

L60%

G60%

(Manuscript received August 1, 2007;
revision accepted for publication February 25, 2008.)