Normal recognition of emotion in a prosopagnosic

Bradley C Duchaine, Holly Parker, Ken Nakayama
Vision Sciences Laboratory, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA;
e-mail: brad@wjh.harvard.edu
Received 19 November 2002, in revised form 22 February 2003

Abstract. In the leading model of face perception, facial identity and facial expressions of emotion are recognized by separate mechanisms. In this report, we provide evidence supporting the independence of these processes by documenting an individual with severely impaired recognition of facial identity yet normal recognition of facial expressions of emotion. NM, a 40-year-old prosopagnosic, showed severely impaired performance on five of six tests of facial identity recognition. In contrast, she performed in the normal range on four different tests of emotion recognition. Because the tests of identity recognition and emotion recognition assessed her abilities in a variety of ways, these results provide solid support for models in which identity recognition and emotion recognition are performed by separate processes.

1 Introduction
Bruce and Young’s (1986) influential model of face recognition posits that there are separate mechanisms for the recognition of facial identity and the recognition of facial expressions of emotion, and this distinction is supported by studies from neuropsychology, neurophysiology, and neuroimaging. There are a number of reports of individuals who have lost the ability to recognize facial identity yet retain the ability to recognize facial expressions of emotion (Bowers et al 1985; Etcoff 1984; Evans et al 1995; Jones and Tranel 2001; Nunn et al 2001; Tranel et al 1988). Conversely, other neuropsychological patients have shown the opposite pattern (Humphreys et al 1993; Kurucz and Feldmar 1979; Young et al 1993). Similarly, single-cell recording in monkeys (Hasselmo et al 1989; Perrett et al 1984), recording and stimulation in humans (Fried et al 1982; Kreiman et al 2000), and neuroimaging in humans (Haxby et al 2002; Puce et al 1998) have found separate brain regions that respond preferentially to facial identity and facial expressions. The independence of these abilities is further supported by dissociations within these abilities. It appears that the recognition of facial identity is performed by at least two dissociable mechanisms (Le Grand et al 2001; Moscovitch et al 1997), and selective impairments in the recognition of particular facial expressions suggest that multiple mechanisms are responsible for recognizing facial expressions (Adolphs 2002; Gray et al 1997; Harmer et al 2001; Lawrence et al 2002; Sprengelmeyer et al 1996).

Although neuropsychological dissociations have provided some of the strongest evidence for the independence of mechanisms for identity and emotion recognition, many of the cases documenting impaired identity recognition and normal emotion recognition were assessed with a small number of tests of identity and emotion recognition because this issue was not the focus of the case studies. For example, in four of the five cases cited above only one type of emotion recognition test was used (Etcoff 1984; Jones and Tranel 2001; Nunn et al 2001; Tranel et al 1988), and in only one case were two tests of emotion recognition used (Evans et al 1995). Bowers et al (1985) used a variety of emotion tests, but their analysis dealt only with differences between groups based on lesion location and not dissociations within individuals. In addition, support for the dissociation in their patients was apparent on the emotion tasks only when performance on one of the identity tasks was equated between groups, because
the group showing differentially impaired emotion recognition was also impaired with identity recognition.

As a result, there are no comprehensive demonstrations of normal emotion recognition with impaired identity recognition and so the double dissociation between these abilities is open to question. To address this issue, we have tested a prosopagnosic with six tests of identity recognition and four tests of emotion recognition. By using many tests examining different aspects of each competence and using a variety of designs, we provide more conclusive evidence that the abilities can dissociate.

2 Case history

NM is a 40-year-old left-handed teacher who reports great trouble recognizing facial identity. She is a quite social person who enjoys interacting with others, and so her face recognition impairments are sometimes very distressing for her. Because NM is a teacher, person recognition is critical, and she uses a number of non-facial routes to recognition including hair, body shape, context, voice, and characteristic face and body movements. These strategies help NM recognize some people, but she still has trouble recognizing many individuals.

NM became aware of her face recognition problems during her teenage years, but it was not until she read a newspaper advice column that mentioned prosopagnosia that she realized, as an adult, that it was a recognized condition. NM knows of no events that are likely to have caused brain damage, and she reports that her mother also has difficulty recognizing faces but we have not confirmed this with testing yet. In cases in which there is no reason to suspect brain damage, such prosopagnosics are usually called developmental prosopagnosics.

Like some prosopagnosics, NM has great difficulties with navigation (Jones and Tranel 2001; McConachie 1976). She describes her traveling as “bumbling” from place to place, and reports that she has trouble understanding how the neighborhoods that she has lived in are laid out. She reports no difficulties with everyday object recognition, but tests of individual-item object recognition (sometimes called subordinate level) with categories such as natural landscapes, horses, and cars have shown that she does have some problems when forced to recognize particular objects. She does not use glasses or contact lenses.

3 Tests of low-level vision and basic-level object recognition

In order to determine if lower-level visual problems contribute to NM’s recognition difficulties, we tested her low-level vision. We also tested her performance on basic-level object recognition.

3.1 Visual acuity

NM was tested on Snellen acuity from 20 feet. Her binocular acuity was 20/20, her right eye was 20/25, and her left eye was 20/40.

3.2 Contrast sensitivity: Pelli–Robson

The Pelli–Robson contrast sensitivity test is a commercially available test of contrast sensitivity (Pelli et al 1988). Participants are asked to identify letters on the chart, and the contrast of the letters decreases as observers proceed through the test. With the standard scoring procedure, NM achieved scores of 1.80 using her right eye, 1.65 using her left eye, and 1.95 using both eyes. These scores are in the normal range (Mantyjarvi and Laitinen 2001).

3.3 Birmingham Object Recognition Battery

The Birmingham Object Recognition Battery (Riddock and Humphreys 1993) consists of tests designed to assess functioning at number of levels in the visual system. In the
copying test, NM had no trouble copying geometric figures or simple objects. The next four tests all require participants to judge whether a particular aspect of two stimuli presented side-by-side is the same or different. NM performed normally on all of these tests: 27/30 on Length Match, 24/30 on Size Match, 26/30 on Orientation Match, and 32/40 on Position of Gap Match. These results along with her Pelli–Robson score indicate that her elementary perception is normal.

On the Overlapping Figures test, NM had no difficulty. She also performed normally on the Minimal Feature Match, Foreshortened Match, and Object Decision. NM enjoys sketching, but she found it challenging to draw the animals in the Drawing from Memory test. Interestingly, NM typically draws people from behind rather than from the front, because she often uses hair to recognize others.

3.4 Basic-level object naming

We presented NM with 100 figures from the Snodgrass and Vanderwart (1980) set of line drawings, and asked her to name them at the basic level (e.g., car, scissors, helmet) (Rosch et al. 1976). The first 50 items were presented for 2 s per item and the second 50 were presented for 1 s per item. NM found this task very easy and named the objects without any problem.

3.5 Summary

The results of these tests indicate that NM has no trouble with elementary perception and basic-level object recognition. This indicates that her face recognition difficulties do not stem from problems with low-level perception or basic-level object recognition but from impairments to higher-level mechanisms.

4 Does NM have problems recognizing facial identity?

NM reports that she has great difficulty recognizing facial identity, and we tested her with six tests of face recognition. Together, these tests probe face recognition in a number of different ways. As we show below, they confirmed that NM is, in fact, severely impaired with the recognition of facial identity from static images.

4.1 Face One in Ten

The Face One in Ten (OIT) tests participants’ ability to discriminate between different faces despite extreme changes in illumination (Duchaine 2000). It simulates everyday recognition better than many tests in two respects. First, it requires recognition of test images that the participant has not studied, so face recognition is required rather than simply stimulus or image recognition (Hay and Young 1982). Second, subjects repeatedly view the different faces used in the test, so target discrimination requires that participants recognize the particular target face rather than simply discriminating between faces that are familiar and those that are not. Finally, as in everyday recognition, most faces presented are not the face for which the participant is searching.

Participants were asked to recognize 15 photos of a target face, which differed greatly in illumination, out of 150 photos presented one at a time (see figure 1). Three different target faces were used so there were 450 trials (3 × 150) with a total of 45 (3 × 15) target presentations. The faces in these black-and-white photographs were cropped so that only the internal facial features were visible. In the study phase, participants were asked to memorize the target face from three different photos that were cycled through three times for 3 s per photograph. Following this, participants were told that they would be presented with test faces, one at a time, and they were asked to respond as quickly as possible with a mouse click whether or not the photo displayed the target face. Participants were also informed that the target faces would appear on approximately 10% of the trials. None of the study faces was used as a test face, and the 135 distractor faces consisted of 15 different images of nine individuals. The 150 test
photographs were broken into three groups of 50, and the target study faces were presented prior to each set. Participants were not provided with any feedback about their performance.

The first 50 trials (one set) are used as a practice set so 400 trials were included in the data analysis. A signal detection analysis showed that NM’s discrimination was far worse than that of the controls (see figure 7). The measure of discrimination, $d'$, for thirteen undergraduate controls was 3.61 with a standard deviation (SD) of 0.49. NM’s $d'$ of 1.66 places her 4 SDs below the mean. Her response times were also more than twice as long as those of the controls. Controls averaged 774 ms (SD = 121 ms) when responding “target present” and 530 ms (SD = 87 ms) when responding “target absent”. NM’s “target present” mean was 2016 ms, and her “target absent” mean was 1240 ms.

4.2 Old/new discriminations

In these two tests, participants were presented with ten black-and-white photographs of women and were later required to discriminate these ‘old’ faces from ‘new’ faces that they had not seen before. Unlike the Face OIT, this test requires participants to memorize and recognize the faces of ten individuals. Identical images of each face are used in this test, but the number of images used makes strategies relying on unique features of each target image (size of smile, skin color, etc) unlikely to result in normal performance.

High-school yearbook photos of Caucasian women were cropped so that very little hair was visible (see figure 2). In the study phase, each of the 10 target faces was shown for 3 s, and the target faces were cycled through twice. In the test phase, 50 images were presented one at a time, and participants mouse clicked to indicate whether the image was ‘old’ or ‘new’ as quickly as possible without sacrificing accuracy. The 50 images consisted of each of the 10 target faces presented twice ($10 \times 2$) and 30 different new faces.

Two separate tests using this paradigm were conducted. Each used a different set of faces, and the two tests were run on different days. On face set 1, the $d'$ for the fifteen graduate student controls was 3.06 (SD = 0.55). NM’s $d'$ was 1.65; this is more than 2.5 standard deviations below the mean. Her response times were approximately 2 SDs slower than the mean. The control means for hits and correct rejections were
962 ms (SD = 227 ms) and 948 ms (SD = 214 ms), respectively. NM’s means were 1553 ms and 1554 ms.

The results for face set 2 were similar to those for face set 1. The $d'$ of the same fifteen graduate student controls was 2.91 (SD = 0.40) whereas NM’s $d'$ was 1.29, which is 4 SDs below the mean. Controls averaged 977 ms (SD = 201 ms) on hit trials, and 975 ms (SD = 197 ms) for correct rejections. NM’s hit mean of 1074 ms was comparable to that of the controls, and her correct rejection mean of 1415 ms was 2 SDs slower than that of the controls (see figure 7 for the results from these tests).

4.3 Warrington Recognition Memory for Faces
The Warrington Recognition Memory for Faces (RMF—Warrington 1984) is similar to the Old/New discriminations discussed in the previous section, but it uses 50 target faces rather than 10 faces so it requires memorization and recognition of an even larger set of target faces. Participants viewed 50 faces for 3 s a piece in the study phase, and responded whether they found each face “pleasant” or “unpleasant”. In the test phase, pairs made up of an ‘old’ face (an image identical to that used in the study phase) and a ‘new’ face were presented simultaneously, and participants were asked to choose the previously viewed face. There were 50 items, and the published control mean for individuals in the 40–44-year-old age group is 44.8 (SD = 3.3). NM’s score of 26 was at chance, and so she was extremely impaired on this test (see figure 7).

It is worth noting that problems with the RMF do not imperil the interpretation of NM’s score. The problems exist because the stimuli in the RMF include a considerable amount of non-facial information so it is possible for participants to perform normally on the test without using the facial information (Duchaine and Weidenfeld 2003; Nunn et al 2001). However, an impaired score does demonstrate impaired face recognition, because the participant was unable to score normally despite the availability of both facial and non-facial information.

4.4 Famous Faces
Normal performance on ‘famous face’ tests requires that participants access previously existing face memories in order to recognize novel views of the celebrities’ faces. Participants were presented with the images of twenty-five celebrities cropped so that minimal hair and clothing were visible (see figure 3). Each image was presented for 10 s and participants had 5 s between images to respond with the celebrity’s name or some other piece of unique identifying information (political office held, movie role, etc). Unlike some prosopagnosics, NM was able to identify many of the faces in this test (Duchaine 2000), but her errors revealed her difficulties with face recognition and the alternative routes to recognition that she used. She was able to correctly identify fifteen
celebrities, and afterward indicated that she was familiar with and had significant exposure to images of twenty-three of the twenty-five individuals. Sixteen undergraduate control subjects averaged 23.6/25 with an SD of 1.3, and so NM’s performance was 6 SDs below the mean of the control group.

Some of her errors demonstrated her impairment quite clearly. When presented with Arnold Schwarzenegger in a color photograph, she identified him as Ronald Reagan. A bit of Schwarzenegger’s hair was visible, and it looked similar to Reagan’s pompadour. NM also hesitantly identified a photograph of a young Elvis Presley as Bill Clinton. Later, when I showed her the image again, she said that Elvis’s grin led her to identify it as Clinton. Above we mentioned that hair and characteristic emotional expressions are routes that NM reported relying on, and these remarkable confusions make her use of them quite clear.

4.5 Different Views test
In the face recognition tests described above, frontal shots of faces were always used. In the ‘different views’ test, participants were asked to recognize individuals across changes in viewpoint. Participants were presented with a frontal view of a face for 3 s and then required to decide which one of 3 three-quarter profile photos showed the individual in the frontal shot. The three-quarters profile photos appeared immediately after the frontal shot was removed from the screen. Sixteen undergraduate control participants averaged 27.1/30 (SD = 1.3); NM’s score of 25/30 placed her score on the low end of the normal range (see figure 7). NM reported that she performed well on the test, because she was able to use a feature-matching strategy (primarily relying on the eyebrows) despite the changes in the view of the faces.

4.6 Face recognition summary
NM’s face recognition is clearly impaired. Performance on the unfamiliar face tests can be fairly compared, because all participants have equivalent exposure to the faces. On these tests, her accuracy scores were 4.0, 2.6, 4.1, 5.7, and 1.6 SDs below the control averages. In addition, her response times tended to be much longer than the control averages (see figure 7 for the results from all these tests).

5 Is NM’s recognition of facial expressions impaired?
NM believes that she can discriminate facial expressions of emotion normally, and reports that she sometimes uses characteristic facial expressions to determine identity. Next we present four experiments assessing NM’s recognition of facial expressions of emotion. To increase the chances of finding an impaired score on an emotion test, we have tested her using widely differing methods which tap into different aspects of emotion perception.
5.1 Emotion Hexagon

In the Emotion Hexagon, subjects must identify the emotion depicted by a model. Following the method used in Young et al (1996), stimuli for the Emotion Hexagon were created by morphing between the 6 different facial expressions of emotion drawn from the Ekman and Friesen series (1976). Two male and two female models were used in our experiment, and 5 morphed images were created between each of the following emotion pairs for each model: happiness–surprise, surprise–fear, fear–sadness, sadness–disgust, disgust–anger, and anger–happiness (see figure 4). The morphs were created from blends of 90%–10%, 70%–30%, and 50%–50%. Each image was displayed for 1000 ms, and participants were instructed to respond with a key-press to indicate the predominant emotion. Our set consisted of 120 images, and each image was shown twice during the experiment. Because there were 6 response options, response times were not analyzed.

Responses to the 50%–50% morphs were not analyzed, and responses were considered correct only if the participant indicated the predominant emotion. On the 92 90%–10% blends, twenty-five undergraduate control subjects averaged 79.0 correct responses (SD = 7.8). NM responded correctly to 89 of the 92 faces so she was well above the mean and nearly perfect with these faces. The controls averaged 72.6 correct on the 70%–30% blends (SD = 7.5), and NM’s score of 77 again placed her above the mean. In addition, her distribution of errors showed that she did not have a selective deficit for a particular emotion. NM’s results for the 90% and 70% morphs are shown combined in figure 7.

5.2 Reading the Mind in the Eyes test

Baron-Cohen and colleagues (1997, 2001) developed the ‘Reading the Mind in the Eyes’ test in order to assess the attribution of emotions in adults. Unlike the previous test, response options consisted of many different, often subtly differing emotional states. In this test, 36 photographs of faces that show only the eye region are presented to participants one at a time. Four terms for emotional states are presented along with each eye region, and participants must decide which adjective best describes the emotional state of the model (see figure 5). We modified the test slightly by scanning the images and presenting them on a computer screen, but the scanned images lost little detail and our control participants scored very similarly to controls tested with the paper version. Participants were given an unlimited amount of time to provide their answers, which were recorded by an experimenter.

NM’s score of 26/36 places her in the normal range when compared with both our controls and the test developers’ controls. The same twenty-five undergraduate controls used for our version of the Emotion Hexagon averaged 28.6/36 (SD = 2.6) (see figure 7), and the one hundred and three undergraduates tested by Baron-Cohen et al (2001) averaged 28.0 (SD = 3.5). Baron-Cohen et al (2001) also tested one hundred and twenty-two controls from the general population, and their mean was 26.2 (SD = 3.6). Although this is not a timed test, NM had typical response latencies.
5.3 Emotion Matching

The Emotion Matching test assesses the ability to categorize emotional expressions as the same despite changes in the models portraying the emotions. Subjects were presented with a photograph of an individual drawn from Ekman and Friesen's set (1976) depicting surprise, disgust, happiness, or neutrality. This photograph was removed from the screen and immediately afterward photographs of three new individuals, each depicting a different emotion, were presented. Participants were asked to indicate with a key-press which of these three faces depicted the same emotional expression as the individual presented initially.

There were 8 items per emotional expression so there were a total of 32 items. The same twenty-five undergraduate control participants averaged 29.9 (SD = 1.9). NM's score of 30 places her in the normal range (see figure 7). Her response times were also within the normal range.

5.4 Emotional Intensity

This test assesses a participant’s ability to judge the relative intensity of differing emotional expressions. The stimuli used were morphs between a neutral emotional expression and faces portraying anger, disgust, and fear (see figure 6). Five different models were used for each emotion series. The difference between any two faces in a
morphed continuum was always centered on the 50% morph and the difference was determined by comparing their positions. For example, the difference between a morph that is 42% neutral/58% anger is 16%. On each trial, pairs of faces from a continuum were presented sequentially for 300 ms per face, and participants had to decide which face depicted a more intense emotional expression. The pairs differed by 8%, 16%, 24%, 32%, or 64%, and there were 10 pairs presented for each difference. Because three emotions were tested, there were 150 trials total. Table 1 presents NM’s results, and it shows that she performed similarly to the undergraduate control participants for all three emotions.

### Table 1. Comparison of the control means and NM’s performance on the Emotional Intensity task.

<table>
<thead>
<tr>
<th>Difference</th>
<th>Anger Control %</th>
<th>Anger NM’s %</th>
<th>Disgust Control %</th>
<th>Disgust NM’s %</th>
<th>Fear Control %</th>
<th>Fear NM’s %</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>96.3</td>
<td>100</td>
<td>92.7</td>
<td>100</td>
<td>64</td>
<td>100</td>
</tr>
<tr>
<td>32</td>
<td>93.3</td>
<td>100</td>
<td>91.3</td>
<td>100</td>
<td>32</td>
<td>100</td>
</tr>
<tr>
<td>24</td>
<td>90.0</td>
<td>100</td>
<td>88.7</td>
<td>90</td>
<td>24</td>
<td>100</td>
</tr>
<tr>
<td>16</td>
<td>80.0</td>
<td>60</td>
<td>75.3</td>
<td>70</td>
<td>16</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>68.3</td>
<td>70</td>
<td>64.3</td>
<td>70</td>
<td>8</td>
<td>100</td>
</tr>
</tbody>
</table>

#### 5.5 Emotion recognition summary

Our four tests probed NM’s emotion recognition abilities in a variety of ways. Her scores ranged from 1 SD below the mean to 1 SD above the mean, so she demonstrated normal performance on all of the tests. These tests indicate that NM does not have any problems recognizing facial expressions of emotion.

#### 6 Discussion

We have summarized all of NM’s test results in figure 7, and it shows that she has a deficit for facial identity recognition but not for facial emotion recognition. We tested her

![Figure 7. NM’s performance in z values for the identity recognition tests and the emotion recognition tests [z = (control mean – NM’s score)/control SD]. Negative z values for accuracy represent performance worse than the control mean and negative z values for response times represent durations longer than the control mean. The vertical dashed line at z = −2.0 indicates the border between impaired scores and normal scores.](image-url)
with six tests of facial identity recognition, and her face recognition problems were clearly manifest on these tests. In contrast, she performed normally on four tests requiring the recognition of facial expressions of emotion. These results accord with her belief that she has great difficulty recognizing facial identity while having no trouble recognizing facial expressions of emotion.

We used a variety of identity and expression tests in order to more firmly ground our conclusions about her abilities. The tests of identity recognition varied in the number of faces involved in the task (one/many), whether recognition of novel views was required, and the time duration between memorization of a face and recognition of the face (immediate/intermediate/weeks). Similarly, our tests of expression recognition differed in terms of the information presented (whole face/eye region), the nature of the discrimination (between different emotions/within one emotion), the intensity of the expression, and the expressions used. Given her consistent outcome on such a wide variety of tests, we can confidently conclude that her recognition of facial identity is impaired while her recognition of facial expressions of emotion is normal.

Neuropsychological dissociations between these abilities have provided some of the most powerful evidence in favor of the separate modules proposed by Bruce and Young (1986), but the past cases demonstrating impaired identity-recognition with normal expression-recognition were not based on varied and systematic tests. Our case provides one-half of the double dissociation between these abilities, and three reports have shown normal identity-recognition with impaired expression-recognition (Humphreys et al 1993; Kurucz and Feldmar 1979; Young et al 1993). This double dissociation along with the other lines of evidence discussed in section I make a strong case for the computational independence of identity and expression recognition.

Finally, neuroimaging (Hoffman and Haxby 2000; Kanwisher et al 1997; Puce et al 1998), neurophysiological studies (Desimone 1991; Hasselmo et al 1989; Perrett et al 1984), and lesion studies (Barton et al 2001; Landis et al 1986) indicate that the recognition of facial identity depends largely on fusiform structures, whereas perception of changeable aspects of the face such as emotional expression is performed in the superior temporal sulcus. This suggests that NM’s impairments may lie in the fusiform gyrus while her superior temporal sulcus mechanisms operate normally. It also suggests that NM is more likely to perceive lip movements, eye gaze, and other changeable aspects of the face than prosopagnosics with impaired recognition of facial expressions of emotion, and future experiments will address this question.

Acknowledgments. We are grateful to NM for her time, interest, and enthusiasm. Daphne Maurer, Rick Le Grand, and Phil Cooper kindly provided us with NM’s acuity results. This research was supported by grants from NIH (F32 MH64246-02 & R01 EY13602).

References
Adolphs R, 2002 “Neural systems for recognizing emotion” Current Opinion in Neurobiology 12 169 – 177
Barton J, Cherkasova M, O’Connor M, 2001 “Covert recognition in acquired and developmental prosopagnosia” Neurology 57 1161 – 1168
Bruce V, Young A, 1986 “Understanding face recognition” British Journal of Psychology 77 305 – 327
Duchaine B, 2000 “Normal configural processing in developmental prosopagnosia” NeuroReport 11 79–83
Duchaine B, Weidenfeld A, 2003 “An evaluation of two commonly used tests of unfamiliar face recognition” Neuropsychologia 41 713–720
Ettcoff N, 1984 “Selective attention to facial identity and facial emotion” Neuropsychologia 22 281–295
Humphreys G W, Donnelly N, Ridoch M J, 1993 “Expression is computed separately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence” Neuropsychologia 31 173–181
Jones R D, Tranel D, 2001 “Severe developmental prosopagnosia in a child with superior intellect” Journal of Clinical and Experimental Neuropsychology 23 265–273
Le Grand R, Mondloch C J, Maurer D, Brent H F, 2001 “Early visual experience and face processing” Nature 410 890
Moscovitch M, Winocur G, Behrmann M, 1997 “What is special about face recognition? Nineteen experiments on a person with visual object agnosia, and dyslexia but normal face recognition” Journal of Cognitive Neuroscience 9 555–604