PERGAMON

Neuropsychologia 41 (2003) 421–441

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Less impairment in face imagery than face perception in early prosopagnosia

Pascale Michelon [a,*], Irving Biederman [b]

[a] Department of Psychology, Washington University, Campus Box 1125, One Brookings Drive, St. Louis, MO 63130-4899, USA
[b] University of Southern California, Los Angeles, CA, USA

Received 6 November 2001; received in revised form 20 August 2002; accepted 20 August 2002

Abstract

There have been a number of reports of preserved face imagery in prosopagnosia. We put this issue to experimental test by comparing the performance of MJH, a 34-year-old prosopagnosic since the age of 5, to controls on tasks where the participants had to judge faces of current celebrities, either in terms of overall similarity (Of Bette Midler, Hillary Clinton, and Diane Sawyer, whose face looks least like the other two?) or on individual features (Is Ronald Reagan's nose pointy?). For each task, a performance measure reflecting the degree of agreement of each participant with the average of the others (not including MJH) was calculated. On the imagery versions of these tasks, MJH was within the lower range of the controls for the agreement measure (though significantly below the mean of the controls). When the same tasks were performed from pictures, agreement among the controls markedly increased whereas MJH's performance was virtually unaffected, placing him well below the range of the controls. This pattern was also apparent with a test of facial features of emotion (Are the eyes wrinkled when someone is surprised?). On three non-face imagery tasks assessing color (What color is a football?), relative lengths of animal's tails (Is a bear's tail long in proportion to its body?), and mental size comparisons (What is bigger, a camel or a zebra?), MJH was within or close to the lower end of the normal range. As most of the celebrities became famous after the onset of MJH's prosopagnosia, our confirmation of the reports of less impaired face imagery in some prosopagnosics cannot be attributed to pre-lesion storage. We speculate that face recognition, in contrast to object recognition, relies more heavily on a representation that descibes the initial spatial filter values so the metrics of the facial surface can be specified. If prosopagnosia is regarded as a form of simultanagnosia in which some of these filter values cannot be registered on any one encounter with a face, then multiple opportunities for repeated storage may partially compensate for the degraded representation on that single encounter. Imagery may allow access to this more complete representation.

Keywords: Apperceptive prosopagnosia; Visual mental imagery; Visual perception

1. Introduction

Evidence from different domains of investigation suggests that visual imagery and visual perception share representations and processes (see [26] for a review). In the neuropsychological domain, cases of patients showing parallel impairments in imagery and perception have been considered as evidence that these activities share neural structures. However, dissociated impairments have also been reported, with individuals manifesting impaired perception with intact imagery [4,5] or vice versa [29,38]. Such observations have led some authors [5,18,26,27] to propose that although imagery and perception activate the same store of representations, they can differ in terms of processes. Parallel deficits in imagery and perception would result from impaired shared processes, such as those involved in accessing previously stored material. In contrast, isolated imagery deficits would result from impairment of processes specific to the imagery activity (such as the image generation processes) and isolated perceptual deficits would result from impairment of processes specific to the perceptual activity (such as edge detection and feature grouping). Imagery in individuals with a perceptual deficit could still be manifested by using visual representations processed and stored before the lesion occurred. In the face recognition domain, a few cases of preserved face imagery in prosopagnosia have been reported [3,15]. As expected from the standard account, these individuals could only imagine faces of persons learned before the lesion.

Two broad types of prosopagnosia are usually distinguished: apperceptive (or perceptual) and associative (or mnesic) prosopagnosia [16]. Apperceptive prosopagnosia is

* Corresponding author. Tel.: +1-314-935-4138; fax: +1-314-935-7588.
E-mail addresses: pmichelo@artsci.wustl.edu (P. Michelon), bieder@usc.edu (I. Biederman).

0028-3932/02/$ – see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0028-3932(02)00176-8

defined by a perceptual deficit affecting the visual processing of faces. In associative prosopagnosia the deficit is at the memory level: faces can be correctly processed but associated information (personality, name, etc.) can no longer be retrieved. As Young et al. [43] suggest, only patients with apperceptive agnosia may show preserved face imagery. Patients with associative prosopagnosia could not access visual representations of faces from names.

In this study, we report the case of an apperceptive prosopagnosic patient, MJH, whose pattern of performance—less impaired imagery, relative to his perceptual recognition, of faces acquired after the onset of his prosopagnosia—is not readily, prima facia, reconciled with the standard accounts of the relation between perception and imagery. We do, however, offer a speculation as to how these results might be accommodated by a general model of face representation.

2. Case description

At the age of 5, MJH fell off an 8 ft high platform and sustained lesions in his left visual cortex (Area 18) and right fusiform gyrus, as confirmed by an MRI scan. At the time of testing, MJH was 34. MJH underwent several neuropsychological evaluations (in 1979 at local hospitals, in 1986 at Harvard Medical School, and in 1996 at the University of Southern California (Biederman and coworkers, unpublished data). Results of these evaluations are presented in Table 1. They suggest that MJH is now alert, attentive, socially functional, and has no general intellectual deficit. He has a degree from a college oriented to individuals with various kinds of disabilities. He is quite outgoing and has a good sense of humor. Previously, he was employed as a writer for television sitcoms and he currently performs as a motivational speaker for various corporations. However, MJH suffers from several problems such as tunnel vision beyond approximately 10° in the right visual field (stemming from his lesion in Area 18), particularly in the lower portion of that field. He has difficulty comprehending abstract drawings, a mild learning disability involving arithmetic, spelling, and handwriting, slowed reading speed, poor sense of direction, impaired motor speed and coordination and a history of tics.

MJH is above average at naming object pictures, missing only three in the Boston Naming Test. He manifests no problem in recognizing people from the sound of their voice but he is extremely poor at recognizing familiar faces and matching or distinguishing unfamiliar faces. Given a forced choice test in which one had to select a celebrity from a pair of pictures, the other being a noncelebrity, MJH was correct on only 22 out of 40, or 55%. The celebrities were limited to those that were highly familiar according to MJH, according to his self report [30]. Normal controls averaged 39 out of 40, or 97.5%. Whereas normal subjects were almost always able to respond immediately, MJH often required approximately 10 s to make his choice. In attempting to identify individuals from pictures of their faces, MJH relies on local cues. In Mangini and Biederman's test, he was able to name five of the celebrities based on such cues, e.g. Woody Allen: hair, hairline, and big glasses; Goldie Hawn: mole on lip and long blonde hair; Pope John Paul: no one wears a "cap" like that. (If we eliminate these five instances, then he was 17 out of 35, or 48.6%.) In another test, MJH required several minutes to identify a picture of Albert Einstein from a set of 10 distractor pictures [2]. Normal individuals were able to perform the task at RSVP rates (one picture every 200 ms). He succeeded in this task solely on the basis of Einstein's long hair.

In a forced choice match-to-sample task in which one of two pictures is identical to the sample and the distractor is on a morphed dimension of being a different individual of the same sex and same expression, MJH requires far more stimulus "energy" (i.e. image differences) than normal subjects to reach an accuracy level by staircase threshold method of 75% [30]. Insofar as the three pictures were presented simultaneously with no pressure to respond quickly, neither memory nor familiarity nor retrieval of a person identity node [10] are required to perform this task. Indeed, one does not even have to employ any aspect of face processing to perform this task insofar as the task only requires matching two

Table 1
MJH's performance on 12 classical neuropsychological tests

Test	MJH's score
WAIS	Average range, except difficulty with complex constructions
WISC	Average range
Boston Naming Test	Average range
Benton Judgment of Line Orientation Test	Average range
WRAT	11.8 (reading), 9.5 (spelling), 7.8 (arithmetic)
Rey Auditory Verbal Learning Test	11/15 words encoded after five presentations, nine words produced on immediate recall and seven words after a delay
Gray Oral Reading test	46th percentile; 73rd when given more time to read the text in the comprehension test
Rey–Osterrieth Complex Figure	Disorganized and distorted copy, delayed and immediate recalls about the same
Hooper Visual Organization test	Nine failed out of first 21. MJH identified the isolates of images, but could not integrate them
Groove Peg Board	25 pegs placed in 79 s with his right hand but only seven with his left hand (same time span)
Random Letter Cancellation Test	One target omitted in the lower right and two targets in the lower left (which he later self corrected)

stimuli for physical identity. MJH's deficit is thus of a fundamental inability to employ visual information to *individuate* faces. When the difference between matching and distractor faces are on a dimension of expression (happy–unhappy) or gender, then MJH is far less deviant compared to normal subjects. This accords well with MJH's more normal capacity to judge gender, attractiveness, age, expression, and direction of gaze from pictures of faces (Kalocsai and Biederman, unpublished data [30,32]).

In the discussion we will propose a neural computational account of face recognition (that proposed by Biederman and Kalocsai [7]) to provide a framework for understanding MJH's deficit.

3. Experiment 1: judgment of whole faces

To assess the functioning of configural aspects of MJH's face perception and imagery, we first studied MJH's and controls' ability to perceive and imagine highly similar faces. The Imagery task we used was modeled after a task previously used to test configural imagery [3,15,43]. In this task participants were presented with triads of names of celebrities familiar to those tuned to US popular culture in the year 2000. For each triad, they decided which of these people looks least like the other two. In the triads we created, all faces were fairly similar, in that the individuals were of the same gender, race, hair coloring, and without any obviously distinguishing facial feature (e.g. Bette Midler, Hillary Clinton, Diane Sawyer). The assumption is that such highly similar triads induce reliance on holistic processing of the faces, rather than a search for isolated facial features.

Two perceptual tasks were also created in which photographs of the faces of the same celebrities were used. The design of the perceptual tasks was identical to that of the Imagery task except that the participants were also asked to name the faces. In a first perceptual task, the No Hair task (NH), the participants were presented with only the oval shape of the faces from forehead to chin, without hair. In a second perceptual task, the Hair task (H), we used the full photographs of the faces with hair. MJH, as many other prosopagnosic patients, reports that hair is one of the more important cues he uses to identify a person, so it was possible that MJH would rely on the configuration of the hair rather than the face. Using the same faces in these two perceptual tasks allowed us to isolate performance with or without hair as a cue.

3.1. Method

3.1.1. Participants

We tested MJH and a group of 10 control participants (four females; 20–34 years) who participated voluntarily in the experiment. In all the experiments, the control subjects were undergraduate and graduate students from the University of Southern California.

3.1.2. Materials and apparatus

In the Imagery task, participants saw 60 names of American celebrities, grouped into triads (Appendix A). We selected only celebrities for whom MJH reported that he "knew what they look like." He also said that he could form a mental image of the faces of these celebrities even though he knew he most likely would not be able to recognize them from a picture. The presentation of these celebrities on television and film is often accompanied by introductions, voice, and other identifying cues which would allow MJH to code specific identifying features for those celebrities, e.g. a mole above the left side of the jaw for Dolly Parton. We assumed that control participants would know the celebrities we selected, as these were very famous. This was confirmed by their results to a rating task where the mean percentage of "not known" celebrities was 2.5%, S.D. = 2.6. In the perceptual tasks, 60 color photographs of the faces of the celebrities used in the Imagery task were used, as illustrated in Fig. 1. These photographs were scanned from magazines. Each face image subtended a visual angle of 6° horizontally and 7° vertically. Photographs in the H task represented faces with hair. Photographs in the NH task were presented within ovals that occluded the hair, showing only the faces. Names or faces (depending on the task) were organized by triad. In each triad, faces were of the same sex, approximate age (except for one triad with Leonardo DiCaprio), and similar hair color and hair cut. A pre-test with six judges was run to check that visual mental imagery was necessary to perform the task. Judges were presented with triads of names of celebrities and for each triad, decided which face was the least similar to the other two. All the judges said they

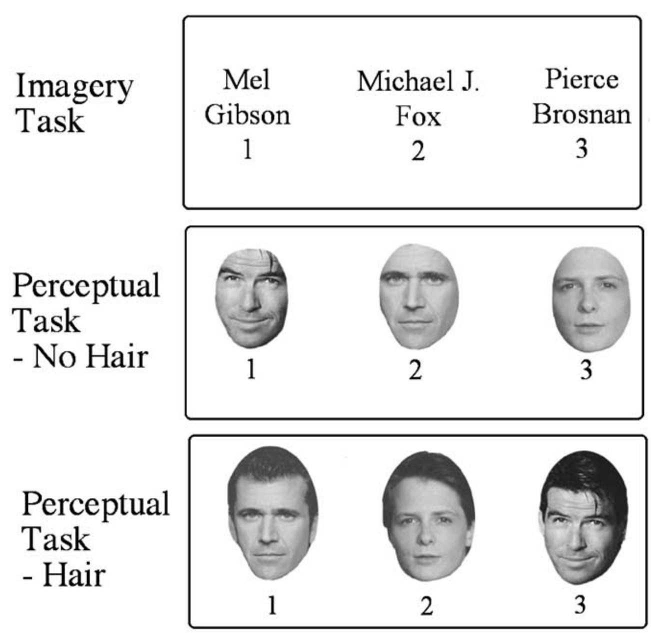

Fig. 1. Example of stimuli used in the three Imagery and Perceptual tasks in Experiment 1. Top row: Name task; middle row: No Hair task; bottom row: Hair task.

imagined the three faces to compare them in every triad. Although a majority of judges chose the same deviant face per triad, some variation appeared which was taken as evidence that faces were fairly similar within triads.

In a triad, names or faces were positioned horizontally, as illustrated in Fig. 1. Below each name or face was a number: 1, 2 or 3. The order of presentation of the 20 triads was randomly varied between the three tasks. The spatial arrangement of the names or faces in each triad was balanced in such a way that every name or face was presented once in the left most position, once in the center and once in the right most position.

3.1.3. Design and procedure

All the participants performed the Imagery task first, then the NH task, and finally the Hair task. Instructions were presented on the screen before each task.

In the Imagery task, each trial started with a fixation point presented for 1000 ms. Then, after a blank screen of 500 ms, the stimulus (three names) was presented and remained on until participants responded. With the stimulus, at the bottom of the screen, a question ("Which person looks least like the other two?") was presented. Participants were asked to imagine the three faces and to answer the question by pressing the numerical key (1, 2 or 3) corresponding to their choice. Response times were recorded. The instructions stressed the importance of using only the face to answer the question and no other attributes such as hair, profession, etc. (even though an attempt was made to keep these factors relatively homogeneous within each triad). After participants had responded, a blank screen of 500 ms ended the trial.

Trials in both perceptual tasks were composed of the same events as in the Imagery task except that photographs of faces were presented instead of words. A second event was also added 500 ms after participants had decided which face was least like the other two: The stimulus appeared again with a different question ("Can you name these celebrities?"). Responses were recorded by the experimenter who then initiated the next trial.

3.1.4. Measure of agreement

We did not consider responses given by the judges in the pre-test or even our own judgments as correct responses, as had been done in previous studies [3,15,43]. We wanted a measure of "agreement" that would reflect variation between individuals. For example, for the triad "Bette Midler, Hillary Clinton, Diane Sawyer," four out of the six judges chose Bette Midler as the least similar person, two chose Hillary Clinton and no one selected Diane Sawyer. If the majority was the criterion, then Bette Midler would be considered as the correct (i.e. least deviant) response. However, if MJH had chosen Hillary Clinton, it would be difficult to consider his response as incorrect as other normal people would have made the same choice.

To derive a more general measure of performance, we computed for MJH, as well as for each control participant, a mean percentage of deviation from the consensus. The deviation score was computed according to the proportion for each choice made by the 10 normal participants (not counting MJH). If six participants had chosen Bette Midler, three Hillary Clinton and one Diane Sawyer, the proportion of participants choosing Bette Midler was 60%, Hillary Clinton 30%, and Diane Sawyer 10%. If MJH had chosen Hillary Clinton then the percentage of his deviation from the consensus would be $100 - 30$, or 70%. In the same triad, the percentage of deviation of a participant choosing Diane Sawyer would be 100%, as he would have been the only one to make this choice.

4. Results

Table 2 shows the mean percentage of deviation and the mean response time (RT) for MJH and for the controls in each task. First, to assess whether MJH was in the controls' range, for each task, we compared MJH's results to those of the controls. Then, to test whether MJH benefited from seeing faces, which was expected for controls, the difference between imagery and perceptual tasks were measured for both MJH and controls.

4.1. Imagery task

In the Imagery task, MJH's RTs were about 9.3 s longer that those for the controls, $t(9) = 6.44$, $P < 0.001$. MJH's percent deviation score, 62.00%, was within the range of the controls (range of 50.55–68.43%) and within 1 S.D. from the controls' average score. It was however significantly greater than the mean deviation score of the controls, $t(9) = 2.38$, $P < 0.05$. Two control participants showed more deviation from the consensus than that shown by MJH. A possible limitation of the inference that MJH's percent deviation was in

Table 2
Mean percentage of deviation from the consensus and RT (s) in the Imagery task and the two perceptual tasks (No Hair and Hair tasks) of Experiment 1

	Percent deviation from consensus			RT (s)		
	Imagery task	No Hair task	Hair task	Imagery task	No Hair task	Hair task
MJH	62.00	58.50	60.50	22.172	24.083	9.243
Control participants	57.59 (5.85)	44.85 (4.33)	46.15 (4.18)	12.885 (4.557)	13.185 (6.210)	10.063 (5.375)

Standard deviations are given in parentheses.

the controls' range may be that both the controls and MJH performed only slightly better than chance, which would have been 66%. However, a *t*-test showed that controls performed significantly better than chance, $t(9) = 4.31$, $P < 0.01$. Nonetheless, we cannot rule out the possibility that the relatively small difference between MJH and the controls can, in part, be attributable to a ceiling effect on the deviation scores.

An item analysis was performed in the Imagery task to assess whether items that were of low agreement for the controls were also of low agreement for MJH. This yielded a positive, but non-significant correlation, $r = 0.332$, $z(18) = 1.3$, n.s.

4.2. Perceptual No Hair task

In the perceptual No Hair task, MJH's mean response time was much longer, by about 10.9 s, than that of the controls, $t(9) = 5.55$, $P < 0.01$. His percent deviation score, 58.5% was outside the controls' range (38.39–52.82%), more than 3 S.D. from the controls' mean score and was significantly higher than the mean deviation of the controls, $t(9) = 9.96$, $P < 0.01$. In addition, MJH correctly named only 6.66% of the celebrities, compared to 96.16% correct for the control group. The misidentifications made by MJH were mostly "I do not know" responses rather than naming the wrong person.

4.3. Perceptual Hair task

In the perceptual Hair task, MJH's mean response time was not different from that of the controls, $t(9) < 1$. However, his percent deviation score, 60.5%, was again outside the controls' range (38.69–50.94%), more than 3 S.D. from the controls' average score, and significantly higher than the mean deviation score of the controls, $t(9) = 10.85$, $P < 0.01$. Moreover, MJH correctly identified only 20% of the celebrities, compared to 97.5% correctly identified by the control participants. Note that misidentifications made by MJH were mostly, as in the No Hair task, "I do not know" responses. However, MJH also responded with an incorrect name on a larger proportion of the Hair trials (26%) than in the No Hair trials (15%).

4.4. Imagery versus Perceptual tasks

In comparison to their performance in the Imagery task, controls reduced their proportion of disagreement both in the No Hair task, $t(9) = 7.68$, $P < 0.01$, and in the Hair task, $t(9) = 4.35$, $P < 0.01$. Controls' RTs did not differ between the three tasks.

An item analysis over the 20 triads showed that, in contrast to controls, the availability of the pictures for MJH produced only a negligible lowering of his deviation scores compared to that which he obtained on the Imagery task. MJH, however, responded faster in the Hair task in comparison to the No Hair task, $t(19) = 4.76$, $P < 0.05$, as well as in comparison to the Imagery task, $t(17) = 2.70$, $P < 0.05$. The d.f. for this test is 17 instead of 19 because in the Imagery task MJH was unable to imagine one of the faces for two of the 20 triads.

Although there was no correlation in the deviation scores over the individual items between MJH and the controls in the Imagery task, there was (surprisingly) a positive correlation, $r = 0.593$, between MJH and controls in the No Hair task, $z(20) = 2.81$, $P < 0.01$, but no correlation between MJH and the controls' performance in the Hair task ($r = -0.04$).

We also performed an item analysis across tasks for MJH and the controls, separately. For MJH, there was a negative correlation, $r = -0.69$, between the Imagery and the No Hair tasks, $z(18) = 3.34$, $P < 0.01$. There was no correlation between the Imagery and the Hair tasks ($r = 0.12$) or between the two perceptual tasks ($r = 0.07$). In contrast, for the controls, there was a positive correlation, $r = 0.78$, between the two perceptual tasks, Hair and No Hair, $z(20) = 4.39$, $P < 0.01$, and no correlation between the Imagery task and the two perceptual tasks (No Hair task: $r = 0.14$; Hair task: $r = 0.11$).

5. Discussion

Our results confirmed MJH's impairment in face recognition in that he was only able to name 4 out of the 60 celebrities in the No Hair task whereas the controls were able to name, on average, 58 out of the 60 pictures. Naming a face requires an extra stage beyond that required for just recognizing the face as familiar or not. When he could not name the face (incorrect answer) MJH never indicated even a faint feeling of familiarity for that face. This observation is consistent with the results of Mangini and Biederman [30] in the celebrity–non celebrity forced-choice task, in which feelings of familiarity could have guided a correct choice. A subsequent experiment [34] also indicates that MJH manifests no covert recognition.

Perhaps the most remarkable result of this investigation is that MJH claims to have imagery of faces at all. But is there any information in his imagery, at least with respect to performance on this task? Whereas his deviation scores were markedly greater than those of the controls on the perceptual tasks, with no control showing greater deviance than him, his deviation score on the Imagery task was only moderately, though significantly, greater than that of the controls, with two of the control participants having deviation scores greater than his. However, to the extent that the percent agreement scores provide evidence of informative imagery for the perceptual conditions, the data may indicate that there is virtually no information in the face imagery of the *normal participants* in that there is no correlation over items for the normal participants between the agreement scores of the Imagery and Perceptual tasks. It is possible that the low correlation between the Imagery and Perceptual tasks for the normal participants represents high variability on the imagery tasks, but this would confirm that the Imagery task, even for the normal participants, may contain little useful information for this task. A possible reason for the low correlations is that the face images generated in imagery did not match the particular pictures of that celebrity selected in the present experiment. It is also possible that the above chance deviation scores for the normal participants and MJH, indicating that the agreement was greater than what would be expected from chance, did not arise from imagery about the face but knowledge about the personality, profession, etc. of the celebrities. However, reliance on such non-face information would have been expected to produce stronger correlations between the Imagery and Perceptual tasks.

These results thus leave some ambiguity as to whether the smaller difference in the agreement scores for MJH and the controls on the Imagery task relative to the perceptual tasks should be interpreted as: (a) indicating some preservation of information in MJH's facial imagery; or (b) merely a consequence of the imagery of normal participants containing little of the information that that is obtained when actually looking at a face. It is possible that the consensus for the controls increased in both perceptual tasks compared to the Imagery task, because judgments were based on the same pictures rather than on somewhat different memory and/or imagery skills for the imagery tasks. But this possibility would be consistent with the proposition that the relatively low discrepancy between MJH and the controls reflects not MJH's preserved imagery but the absence of informative imagery (or at least, more variability in imagery skills) in the controls. (We note that the evidence for preserved imagery is stronger in Experiment 2.)

So whether this first experiment should be interpreted as indicating preserved face imagery is, admittedly, arguable. We present these data because they are consistent with previous observations [4,5] using a method that has been employed by other investigators [3,15,43]. Moscovitch and Moscovitch [35] adopted a criterion of impairment performance that is beyond 2 S.D. from the mean of the controls. Because MJH's performance on the Imagery task was within 1 S.D. of the controls, it would thus not be classified as impaired (MJH's performance on the perceptual tasks was greater than 2 S.D. from the mean and thus would be regarded as impaired). However, because of the high variability of the controls on that task, the near floor level of their agreement scores and because MJH's agreement score, given the variability of the controls, was still significantly lower than the mean of the controls, it may not be reasonable to accept the 2 S.D. criterion for classification of MJH as unimpaired on the Imagery task.

As noted previously, for the controls, items of low agreement in the Imagery task were not necessarily of low agreement in the perceptual tasks. There was a correlation between the agreement in the two perceptual tasks (Hair and No Hair) which may show that controls were using the same type of facial cues to make their judgement in both tasks. MJH did not show the same pattern of results. For him, there was a negative correlation between the Imagery and the No Hair perceptual tasks and no correlation between the two perceptual tasks. The absence of a correlation between the two perceptual tasks suggests that MJH may not have been using the same cues in the two tasks, consistent with our conjecture of greater variability in face encoding for MJH. He may have been unable to refrain from using hair in the Hair task but was forced to use facial cues in the No Hair task. That MJH was trying to use facial cues in the No Hair task may be evidence by the finding that items of low agreement in this task tended to be the same for both MJH and controls. Insofar as there were few local features that could distinguish the faces in the triad, the perceptual impairment of MJH is consistent with the hypothesis that his prosopagnosia results from a deficit in configural processing.

MJH's deviation from the consensus was not different between the two perceptual tasks. This result may suggest that, in the Hair task, MJH could focus on the faces and avoid using the hair to perform the task. However, MJH answered more quickly in the Hair task than in the No Hair task. He seemed more confident although his percent deviation score did not improve. Also, although very impaired, MJH's naming performance was better in the Hair task (20%) than in the No Hair task (6.66%), consistent with the classical observation that prosopagnosics use hair as a major cue to individuation. The improvement of performance could also be attributed to the repetition of the faces. However, MJH made more wrong guesses in the Hair task than in the No Hair task. When guessing MJH used names he had seen in the Imagery task and tried to match them with facial cues. For instance, when presented with Mel Gibson's face in the Hair task MJH's answered "Bruce Willis, because of the hair line." It thus seems more likely that MJH did not benefit from the repetition of the faces but tried to use the hair cut to identify faces, which helped him but was not an overall successful strategy.

If it is accepted that MJH showed relatively less impairment in face imagery than face perception, our results replicate previous reports of preserved configural imagery in prosopagnosia [3,15] with one major difference. In these previous studies, the authors proposed that configural imagery was preserved because mental images of faces were generated from representations of faces stored before the lesion. However, in contrast to these other cases, MJH's prosopagnosia began so early in life that his knowledge of most of the celebrity faces used in our experiment had to be acquired after the lesion occurred. Only about one-third of our presented celebrities (13 of 39) might have been famous before 1970 and those individuals would not have been well known to a 5-year-old. In Section 13 we proposed an explanation as to how MJH might have acquired facial representations after his lesion that could be accessed through imagery despite his impaired perceptual processes. This explanation holds that percepts of a face that are degraded or limited to only a few features could be integrated over repeated exposures to partially correct for degraded or missing information in any one percept. If imagery could access such representations then performance on an Imagery task might rival that of a perceptual task for a prosopagnosic.

6. Experiment 2: judgment of facial features

In this experiment we tested MJH's ability to judge facial features both from perception and from memory. We assessed whether MJH would be impaired in a task requiring perceptual judgments about individual facial features but might be in the normal range when judging the same features from imagery. We used a task modeled after Bartolomeo et al. [4] and Young et al. [43]. In the imagery version of the task, MJH and controls were instructed to imagine faces of American celebrities. Each participant was then asked six questions about the features of each face (e.g. Ronald Reagan, "Is his nose pointy?"). There were six questions specific to each face although some questions were posed for several of the faces (Appendix A). In the perceptual version of the task, the participants were presented with photographs of faces and were asked the same six questions as the ones asked in the Imagery task. In this perceptual version of the task, MJH and the control participants were also asked to name the celebrities.

6.1. Method

6.1.1. Participants

MJH was compared to a group of 10 control participants (six females; age range: 19–25 years) who participated voluntarily in the experiment. None of the controls participated in Experiment 1.

6.1.2. Materials and apparatus

In the Imagery task, 12 names of American celebrities were used. These celebrities constituted a subset of those used in Experiment 1. In the Perceptual task, 12 color photographs of the faces of the celebrities used in the Imagery task were used. For each face, six questions about individual features of the face were designed. The same questions were used in the imagery and perceptual versions of the task. Half of the questions required a "no" answer and the other half a "yes" answer. Each question required a "yes" and a "no" answer an equivalent number of times. For example, the question "Is his–her nose pointy?" required a "no" answer when asked about Ronald Reagan, and Sylvester Stallone but a "yes" answer when asked about Clint Eastwood and Julia Roberts. Questions were designed and then used in a pre-test with five judges to check that visual mental imagery was needed to answer them. All judges confirmed that imagery was needed. The order of presentation of the faces/names and the order of the questions for each face was reversed for half the control participants.

6.1.3. Design and procedure

MJH as well as the 10 control participants performed the Imagery task first, and after a break of 15 min, the Perceptual task.

In the Imagery task each trial began with the name of a celebrity presented for 1000 ms. Looking at a blank screen, participants were asked to imagine the face of the celebrity. Once they had as good an image as possible of the face they pressed a key. Response time was registered. After 500 ms the first question appeared and stayed on the screen until participants responded. To respond, participants pressed either the key marked 'yes' or the key marked 'no'. After 500 ms the second question appeared. Participants responded using the same procedure. After the six questions were presented using this procedure, the message "next person" appeared for 1000 ms. The next trial began after a 1000 ms interval.

Trials in the Perceptual task were composed of exactly the same events as in the Imagery task except that photographs of faces were presented instead of names. For each face, the photograph first stayed on the screen until the participant recognized the celebrity. Participants pressed a key as soon as they recognized the face and then said aloud the name of the celebrity, which was recorded by the experimenter. The six questions were then presented one after the other on the screen using the same procedure as in the Imagery task.

7. Results

In this experiment, as in Experiment 1, responses could not be considered as absolutely correct or incorrect. So, again, we computed, for MJH as well as for each control participant, a mean percentage of deviation from the consensus and mean RTs for each task. Results are summarized in Table 3. To assess whether MJH was in the controls' range, we compared mean percentages of deviation from the consensus and mean RTs for both MJH and controls in each task. Finally we tested: (a) whether MJH and controls' performances were different depending on the nature of the questions (questions about facial features or non-facial features) (Table 4); and (b) whether MJH's imagery performance was

Table 3
Mean percentage of deviation from the consensus and RTs (s) for answering questions about facial features of famous people (Experiment 2)

	Deviation from consensus		RT (s)	
	Imagery	Perception	Imagery	Perception
MJH	31	26.5	3.257	3.173
Control participants	27.8 (2.5)	18 (2.3)	1.452 (0.264)	1.770 (0.412)

Standard deviations are given in parentheses.

Table 4
Mean percentage of deviation from the consensus for answering questions about facial and non-facial features of famous faces (Experiment 2)

	Face questions		Non-Face questions	
	Imagery	Perception	Imagery	Perception
MJH	37.08	31.46	18.75	16.66
Control participants	31.09 (5.27)	20.22 (2.68)	16.9 (3.37)	10.5 (1.86)

Standard deviations are given in parentheses.

different depending on whether he recognized the faces in the perceptual task (Table 5).

7.1. Imagery task

In the Imagery task, MJH's mean latency for generating the image of a celebrity's face was 2 s faster ($M = 1.273$ s) than that of the control participants ($M = 3.627$ s, S.D. = 3.324 s), although the difference, because of high variability among the controls, fell just short of significance, $t(9) = 2.23$, $P = 0.05$. However, MJH needed more time than the controls to answer questions about individual features of the imagined faces, $t(9) = 21.62$, $P < 0.001$. MJH's percentage of deviation in the Imagery task, 31%, was within the control range (23.4–33%) and 1.5 S.D. from the controls' average deviation. It was however significantly greater than the mean deviation of the controls, $t(9) = 4.24$, $P < 0.01$. One control participant showed more deviation from the consensus than did MJH.

7.2. Perceptual task

When shown the pictures, MJH identified only 50% of the faces compared to 100% for the controls and his naming RTs, when correct, were considerably slower ($M = 5.774$ s) than those of the controls ($M = 1.193$ s, S.D. = 0.460 s), $t(9) = 29.87$, $P < 0.01$. Very likely MJH was benefiting from the exposure to the faces and names in the prior experiment. In the perceptual task, MJH's percentage of deviation in answering questions about individual facial features was outside the controls' range (15–22%), almost 4 S.D. above the controls' average score and significantly higher than the mean deviation of the controls, $t(9) = 11.08$, $P < 0.01$. Moreover, these judgments were made more slowly than that of the controls, $t(9) = 10.21$, $P < 0.01$.

Table 5
Mean percentage of deviation from the consensus for answering questions about facial and non-facial features of recognized famous faces (Experiment 2)

	Face questions		Non-face questions	
	Imagery	Perception	Imagery	Perception
MJH	27.3	27.3	17.1	21.4
Control participants	31.1 (5.3)	20.22 (2.7)	16.9 (3.4)	10.5 (1.9)

Standard deviations are given in parentheses.

7.3. Imagery versus perceptual tasks

Control participants were in markedly greater agreement in the perceptual task compared to the Imagery task, $t(9) = 33.49$, $P < 0.01$. However, they responded faster in the Imagery task than in the perceptual task, $t(9) = 2.46$, $P < 0.05$. An item analysis over the 72 trials showed that, in contrast, the reduction in MJH's deviation score on the perceptual task relative to the Imagery task was not significant, $t(71) < 1$. His speed of responding did not differ reliably on the two tasks, $t(52) = 1.1$.

We performed an item analysis to see whether items of low agreement in the Imagery task were also of low agreement in the perceptual task. Results showed a correlation between the imagery and perceptual tasks for both MJH, $r = 0.27$, $z(72) = 2.31$, $P < 0.03$, and the control participants, $r = 0.56$, $z(72) = 5.25$, $P < 0.01$. Additional correlation tests were computed to see whether MJH and controls' judgements were correlated within the same task. Results showed a positive correlation both in the Imagery task, $r = 0.56$, $z(72) = 5.27$, $P < 0.01$, and in the perceptual task, $r = 0.41$, $z(72) = 3.66$, $P < .01$.

7.4. Face versus non-face features

The 72 questions consisted of 48 that concerned facial features, and 24 that concerned non-face features namely hair (22 questions) and glasses (2 questions). Table 4 shows mean performance as a function of feature type. The control participants were in substantially greater agreement for the non-face questions than the face questions both on the Imagery task, $t(9) = 7.55$, $P < 0.01$, and in the Perceptual task, $t(9) = 8.24$, $P < 0.01$. The greater agreement overall for the Perceptual task compared to the Imagery task was evident for both face, $t(9) = 7.02$, $P < 0.01$, and non-face questions, $t(9) = 4.72$, $P < 0.01$. For MJH, despite the 2:1 ratio in agreement scores between face and non-face items, this difference was not significant, neither in the Imagery task, $t(23) = 1.54$, nor in the Perceptual task, $t(23) = 1.68$, suggesting low power in this comparison. The moderately greater agreement shown by MJH in the Perceptual task compared to the Imagery task was not significant either for face questions, $t(47) < 1$, or for non-face questions, $t(23) < 1$.

On the Imagery task, although in the controls' range and within 2 S.D. from the controls' average score, MJH's scores were significantly more deviant than the controls for face questions only (range = 20.02–39.41%, with one participant more deviant than MJH), $t(9) = 3.41$, $P < 0.01$, but not for non-face questions (range = 13.41–24.31%, with three participants more deviant than MJH), $t(9) = 1.64$.

As in Experiment 1, the discrepancy between MJH and the controls was more marked on the Perception task. MJH's score was more than 4 S.D. above the controls' average score and was more deviant than the controls for both the face questions (range = 17.37–25.82%), $t(9) = 12.58$, $P < 0.01$, and non-face questions (range = 7.86–14.8%), $t(9) = 9.93$, $P < 0.01$, with no control participants showing more deviance than MJH on either task.

7.5. Recognized versus non-recognized faces

Insofar as MJH recognized only half of the celebrities' faces (and the control subjects recognized all of them) we compared MJH and controls only for those faces which MJH did recognize, separately for face and non-face questions. Results are shown in Table 5. On the Imagery task, for both face and non-face questions, MJH's mean deviation score was in the normal range—in fact his score was *below* the mean of the controls for the face questions—and less than 1 S.D. from the average controls' score but he was well above the normal range and more than 2 S.D. from the controls' average score in the perceptual tasks for both types of questions. MJH was in substantially greater agreement with the controls on the non-face questions than the face questions but, unlike the controls, he showed no reduction (instead, a slight increase) in his agreement scores in the perceptual tasks.

8. Discussion

MJH had longer latencies and was half as accurate as the control participants for naming celebrities, despite his prior experience with these photos in Experiment 1. This result confirms his difficulties in identifying people from a picture of their faces. That MJH's naming performance in this experiment (50% correct) was much better than in Experiment 1 (20%) is understandable in light of his prior experience. In both the experiments, MJH had seen the names of all the celebrities in the Imagery task before he performed the perceptual naming task. Another important factor in his greater naming accuracy in Experiment 2 was that fewer celebrities (12) were used in Experiment 2 than the 60 used in Experiment 1. Consequently, MJH may have been able to remember more names from the Imagery task in Experiment 2 than in Experiment 1. As MJH can imagine faces, he may be able, when a limited pool of names is available, to use his visual knowledge of the faces as an aid in a naming task.

MJH's performance in judging facial features from memory was less impaired, relative to the controls, than when he was judging the pictures themselves. As in Experiment 1, MJH's mean deviation score was significantly greater than that of the controls and his RTs were considerably and significantly greater than those of the control participants. By the Moscovitch and Moscovitch [35] criterion, MJH's imagery score was within the normal range in that it was less than 2 S.D. away from the controls' mean score (whereas in the Perceptual task it was more than 2 S.D. away from the controls' mean). This pattern of results suggests less impaired imagery than perception. In this Imagery task, performance was much better than chance (unlike Experiment 1), so there is little uncertainty that MJH (and controls) could perform these judgments of facial features from imagery. Moreover, unlike Experiment 1, there was information in the imagery of the control participants in that their imagery responses were significantly positively correlated with their responses in the Perceptual tasks. In the Perceptual task, MJH was not only slower than controls but his performance was out of the normal range. Relative to controls, he performed better in the Imagery task than in the Perceptual task, a pattern that was particularly clear when the data were restricted to those faces that MJH could recognize. The controls had better performance in the Perceptual task but responded more slowly than they did in the Imagery task. This result suggests that it was difficult to make the required judgments based on the photographs even for normal participants.

For both MJH and the controls, items of low agreement in the Imagery task were also of low agreement in the Perceptual task, although the correlation was weaker for MJH than for controls. In the same way, items of low agreement for MJH were also of low agreement for controls in both the Imagery and Perceptual tasks. These results suggest that, in both tasks, MJH used similar facial information as that used by controls. Although some of this information might have been so strongly associated with the celebrity (e.g. Jay Leno's protruding chin, Marilyn Monroe's blond hair) as to be accessible without imagery, much of it was not.

Both in the Imagery and Perceptual tasks, controls and MJH were markedly better in the non-face questions about hair and glasses than in the questions about face features. MJH's imagery for these non-face features was approximately equivalent to those of the controls. These results allow the possibility that the reduced impairment of MJH's imagery was a consequence of equivalent performance on the non-face questions. But, once non-face questions were removed from the analysis, MJH's performance in imagery was still less impaired than his performance in perception compared to controls, which suggest relatively preserved Imagery for both types of features.

Relative to controls, MJH's judgments about facial features are impaired in perception but somewhat preserved in imagery. As is proposed in Section 13, these results can be described in terms of the model proposed by Biederman and Kalocsai [7]. To our knowledge, such a dissociation between intact imagery and damaged perception for facial features has been reported only once previously [3]. In Bartolomeo et al.'s study, the Perceptual task required matching a facial feature to the corresponding item out of three distractors, which were features either isolated or in the context of a face. This perceptual task thus differed from the imagery task, which consisted in answering questions about facial features from memory. This task difference makes it difficult to draw inferences as to the level of performance shown on the two tasks. Other than that study, no previous studies have systematically compared perception and imagery of facial features in prosopagnosia. Thus, it is unclear whether the dissociation observed could also happen in other patients. Indeed, in two other studies of prosopagnosia, preserved facial feature-based imagery was observed but no Perceptual task testing feature-based processing was conducted [42,43]. Conversely, a report of impaired perception of facial features did not assess imagery of facial features [39].

As only one previous study reported a dissociation between imagery and perception for facial features we designed a third experiment to be able to replicate our findings.

9. Experiment 3: judgment of features of emotions

In this experiment we evaluated MJH's ability to imagine and perceive features of faces expressing emotions. All faces were unfamiliar to the participants. We used such stimuli to test whether MJH's perceptual deficit would extend to unknown faces. The task was modeled after Bowers et al. [9]. In the imagery version of the task, participants were asked to imagine faces expressing an emotion and for each face, six questions about the features of the face were asked (e.g.: angry face, "Are the eyes squinted?"). In the perceptual version of the task, photographs of faces expressing emotions were used. Participants were first asked to name the emotion expressed by each face and then to answer the same six questions as the ones used in the Imagery task. To each face corresponded a specific set of six questions, although some questions were repeated from one face to another (Appendix A).

9.1. Method

9.1.1. Participants

MJH was compared to the same group of 10 control participants who participated in Experiment 2.

9.1.2. Materials and apparatus

In the Imagery task, six-word prompts were used: sad face, happy face, surprised face, fearful face, angry face, disgusted face. In the Perceptual task, six black and white digitized photographs from Ekman and Friesen [17] were used. Each face subtended a visual angle of 6° horizontally and 7° vertically. Each of these faces expressed one of the six emotions used in the Imagery task. To characterize each emotion we used the classification scheme proposed by Ekman and Friesen [17]. For each face, six questions about individual features of the face were designed. The same questions were asked in the Imagery and Perceptual tasks. Half of the questions required a "no" answer and the other half a "yes" answer. Each question required a "yes" and a "no" answer an equivalent number of time. For instance, the question "Are the eyes squinted?" required a "no" answer when asked about a fearful face but a "yes" answer when asked about a disgusted face. A pre-test with five judges assessed whether visual imagery was needed to answer them. All judges confirmed that imagery was needed.

9.1.3. Design and procedure

MJH as well as the 10 control participants performed the Imagery task first, and after a break of 15 min, the Perceptual task. The procedure was identical to that used in Experiment 2 except that words referring to facial expression (Imagery task) or faces expressing emotions (Perceptual task) were used as stimuli. In the Perceptual task, participants were asked to name the expressions before they answered the questions. The order of presentation of the emotions and the order of the questions for each emotion was reversed for half the control participants.

10. Results

Results are summarized in Table 6.

10.1. Imagery task

In the Imagery task, MJH's mean latency for generating images of faces expressing an emotion was faster ($M = 2.833$ s) than that of the control participants ($M = 6.120$ s, S.D. $= 4.411$ s), $t(9) = 2.35$, $P < 0.05$. However, MJH was slower than controls to answer questions about individual features of the imagined expressions, $t(9) = 14.57$, $P < 0.01$. MJH's score of deviation in the Imagery task, 31.6%, was within the lower range of the control participants' score (24–33.3%) and within 1 S.D. from the controls' average score. It was however significantly different from the mean controls' score, $t(9) = 3.16$, $P < 0.02$. Three participants showed more deviation from the consensus than did MJH.

10.2. Perceptual task

In the Perceptual task, MJH accurately identified less emotions (50%) than controls (90%, S.D. = 11.65, range: 66.67–100, with half the participants at ceiling) according to the classification provided by Ekman and Friesen [17], $t(9) = 10.85$, $P < 0.01$. However, MJH, as the controls, were 100% accurate in identifying a happy face. Most of

Table 6
Mean percentage of deviation from the consensus and RTs (s) for answering questions about facial features of faces expressing emotions (Experiment 3)

	Deviation from consensus		RT (s)	
	Imagery	Perception	Imagery	Perception
MJH	31.7	27.8	5.022	4.521
Control participants	28.5 (3.2)	12.8 (3.4)	2.330 (0.584)	2.216 (0.391)

Standard deviations are given in parentheses.

the errors that MJH made were confined to the same emotions that some of the normal participants confused. MJH was also considerably slower ($M = 6.016$ s) than control participants ($M = 1.998$ s, S.D. $= 1.217$ s) to name emotions, $t(9) = 10.44$, $p < 0.01$. When answering questions about individual features of the perceived facial expressions, MJH was slower than controls, $t(9) = 18.64$, $P < 0.01$. Moreover, MJH's percentage of deviation in the Perceptual task was outside the controls' range (8.6–19%), more than 4 S.D. above the controls' average score and significantly higher compared to the mean of the controls, $t(9) = 13.95$, $P < 0.01$.

10.3. Imagery versus Perceptual tasks

Control participants had greater consensus (lower deviation scores) in the Perceptual task than in the Imagery task, $t(9) = 45.46$, $P < 0.01$, but were not faster on the Perceptual task, $t(9) < 1$. An item analysis over the 36 trials showed that, in contrast, MJH did not enjoy greater consensus on the Perceptual task, $t(35) < 1$, nor was he faster on that task, $t(27) = 1.3$.

We performed an item analysis to see whether, for both MJH and the control participants, there was a correlation between the Imagery and Perceptual tasks in term of percentage of agreement per item. There was no correlation between the two tasks either for MJH ($r = -0.06$), or for the control participants ($r = 0.03$). Additional correlations were computed to assess whether MJH and controls' performance over items correlated within a task. Results showed a positive correlation in the Imagery task, $r = 0.46$, $z(36) = 2.86$, $P < 0.01$, but not in the Perceptual task, $r = 0.08$.

10.4. Recognized versus non-recognized facial expressions

MJH recognized only 50% of the facial expressions. Half of the controls also made recognition errors. This allowed us to compare MJH and five of the controls for both recognized and non-recognized facial expressions. Results are illustrated in Table 7.

Results did not differ between recognized and non-recognized facial expressions. For both, in the Imagery task, MJH's deviation score was within the normal range (23–32 and 15–50% for recognized and non-recognized expressions, respectively) and within 1 S.D. from the controls' average score. In contrast, in the Perceptual task, MJH's performance was outside the normal range (8–18.5 and 2–22% for recognized and non-recognized expressions, respectively) and more than 2 S.D. from the controls' average score for both recognized and non-recognized expressions.

Table 7
Mean percentage of deviation from the consensus for answering questions about features of faces expressing emotions for recognized and non-recognized expressions (Experiment 3)

	Recognized expressions		Non-Recognized expressions	
	Imagery	Perception	Imagery	Perception
MJH	25.5	24.4	37.7	31.1
Controls	27.5 (3.8)	12.7 (4.2)	27 (14.3)	12.5 (9.2)

Standard deviations are given in parentheses.

11. Discussion

Results again revealed a dissociation between relatively preserved imagery and impaired perception. In the Imagery task, the same pattern of result as the one observed in Experiments 1 and 2 indeed suggests relatively preserved imagery. MJH's mean deviation score was different from that of the controls and his responses were slower. However, MJH's score was within the control range. Three control participants had greater deviation scores than MJH. In the Perceptual task, MJH was not only slower but his judgments were not consistent with those of the control participants and considerably outside the range of the controls. MJH, in contrast to control participants, was better in the Imagery task than in the Perceptual task.

In contrast to what was observed in Experiment 2, results showed no correlation between agreement for a specific item in the Imagery task and agreement for the same item in the Perceptual task, neither for the controls nor for MJH. This result may suggest that representations used in the Imagery task were different, both for MJH and controls, from images presented in the Perceptual task. In contrast to Experiment 2, items of low agreement for MJH were the same as for controls in the Imagery task only, which stress the dissociation between imagery and perception for MJH.

Imagery for single features of faces expressing emotions was relatively preserved in MJH, though slower, but perception of these features was impaired. This was true independently of whether the facial expression was recognized or not. This result shows that for MJH, perception of facial features is impaired even when unknown faces are perceived and the task is not that of individuation. So, the pattern of results observed in Experiment 2 with famous faces was replicated here and extended to unknown faces. It may be surprising that a prosopagnosic patient manifests difficulties in perceiving facial features because: (a) prosopagnosia is often thought of as a configural deficit; and (b) many prosopagnosics use such features as an identification aid. However, the local features employed by MJH and other prosopagnosics tend to be of high contrast and qualitative, i.e. a beauty mark, a distinctive hairline. In our experiment, judgments on facial features required processing fine metric aspects. It may be possible that such fine metric properties are provided or make sense only in the context of the whole face, which cannot be represented by MJH.

In this experiment, MJH's accuracy in naming emotions was lower than that of controls.[1] However, his impairment seems less severe than for naming celebrities. When asked to describe how he identified facial expressions, MJH reported a strategy that suggests the use of facial features rather than facial configurations: "I only need one or two aspects to get an expression." As identifying emotions seems to require configural processing [11], MJH's strategy may need more time to be successful. This would explain why, when given as much time as needed, MJH seems to perform almost as well as controls in an emotion naming task (Kalocsai and Biederman, unpublished data; [30]). However, when time is limited, as it is the case in this experiment and in real life in general, MJH makes errors and although above chance, cannot reach the level of control participants.

Bruce and Young's [10] model of face perception posits independent routes for the processing of facial expressions and face identity. Single unit recordings in the macaque also suggest that individuation of faces and perception of emotions are subserved by separate cortical loci [24]. Some prosopagnosic patients show preserved abilities at identifying facial expressions [33,40], whereas others, like MJH, show some deficit [33]. Associated deficits do not necessarily imply that facial expression processing and identity processing are computed by the same system. It is possible that a lesion damages two independent systems at the same time if they are anatomically related. Alternatively, the lesion could damage an earlier stage that provides perceptual input to both systems.

12. Experiments 4a, 4b and 4c: imagery for non-face objects

In the following three experiments we tested MJH's ability to imagine non-face visual stimuli (i.e. animals, artifacts). Our aim was: (a) to test whether MJH could imagine non-face visual objects in the same way he could imagine faces; and (b) if he could, to test whether he was slower and somewhat less accurate from controls, as was the case when he was imagining faces.

12.1. Experiment 4a

In this experiment, we tested MJH's visual imagery for color. We used a classical neuropsychological task that asks participants to retrieve the color of common items from memory [21].

12.1.1. Method

In all three experiments, MJH's performance was compared to the performance of the same 12 control participants (11 females; age range: 17–21 years) who participated voluntarily.

We used 21 common items that had characteristic colors but were not verbally associated with their colors (e.g. football, lettuce). Items were read by the experimenter one after the other. MJH and control participants were asked to name the characteristic color of each object. The experimenter recorded the responses.

12.1.2. Results and discussion

MJH gave the same response as the majority of the control participants for 80.95% of the items. For each control participant we calculated how many responses were identical to the responses given by the majority of the 11 other control participants. Controls gave the same response as the majority for 90.5% of the stimuli (S.D. = 7.6) (range: 76.19–100%, four participants were at ceiling, i.e. >95%). As in face imagery tasks, MJH's performance was within the controls' performance range and less than 2 S.D. from the controls' average score. Moreover, two participants gave less congruent responses that did MJH. However, MJH's score was significantly worse than the controls' mean score, $t(11) = 4.35$, $P < 0.01$.

12.2. Experiment 4b

In this experiment we tested MJH's visual imagery for metrics of non-face visual objects, using questions about the length of animals' tails relative to the lengths of their bodies (long versus short), a classical test in neuropsychology [19]. Coarse metrics would be distinguished from fine metrics of facial features assessed in Experiment 2. To be able to record response times we designed a computerized version of the test. We used 20 animals' names not verbally associated with tails, 10 with a long tail (e.g. lizard) and 10 with a short tail (e.g. bear). A pre-test done with five judges helped us to decide which animals belonged to which category. For each animal, the participants' task was to decide whether the animal's tail was long in proportion to its body size. To make sure participants understood the task, an example

[1] This result was surprising as in previous testing MJH seemed good at identifying emotions (Kalocsai and Biederman, unpublished data). In our experiment, an emotion was expressed by one particular face only. To give MJH the possibility to identify the same emotion on several faces we designed a control experiment in which each emotion was expressed by five different faces. We used 60 faces from Ekman and Friesen [17]. Facial expressions were surprise, happiness, fear, disgust, anger, and sadness. Faces were presented for 3 s each. Participants (MJH and a group of 14 controls) were asked to name the emotion expressed by each face.

Results showed that the naming accuracy of controls was high: 83.4% (S.D. = 7.96). However, some emotions were more difficult to identify than others were: happiness (100% correct identifications, S.D. = 7.96), sadness (95.7%, S.D. = 6.46), surprise (91.4%, S.D. = 11), anger (77.14%, S.D. = 10.7), disgust (76.48%, S.D. = 18.23), and fear (63.57%, S.D. = 22.05). MJH correctly identified only 50% of the emotions. However, he was above chance level (16.66%) for all the facial expressions except anger: happiness (100%), sadness (40%), surprise (80%), anger (10%), disgust (50%), and fear (20%). MJH was significantly less accurate ($P < 0.01$) than controls for five emotions out of six: sadness, $t(13) = 31.26$, anger, $t(13) = 23.47$, fear, $t(13) = 7.39$, disgust, $t(13) = 5.43$, and surprise, $t(13) = 3.87$. These results confirmed that MJH was impaired at naming facial expressions.

(peacock) was given with the instructions. It is usually assumed that participants only have visual knowledge to help them perform such a task. As a consequence, imagery is required.

12.2.1. Method

Animals' names were presented one by one on the screen until the participant answered by pressing a key marked Yes or a key marked 'no'. Response times were registered.

12.2.2. Results and discussion

MJH's mean percent correct of 85% was within the controls' range, less than 1 S.D. from the controls' mean score and not different from the controls' mean of 89.58% (S.D. = 10.76) (range: 65–100%), $t(11) = 1.47$. Note that two participants gave less correct responses than MJH did and that three were at ceiling, i.e. >95%. However, MJH's mean RTs for correct responses (3.775 s) was longer than the mean response time ($M = 1.529$ s, S.D. = 0.563 s) of the control participants, $t(11) = 13.82$, $P < 0.01$. As the same controls participated in Experiments 4a and 4b we tested whether participants who were less accurate in Experiment 4a were also less accurate in Experiment 4b. Results showed no correlation between the two experiments.

So although MJH's performance accuracy was not different from that of controls, his use of object imagery, as his use of face imagery, was slower than that the control participants.

12.3. Experiment 4c

In this experiment we tested MJH's visual imagery for non-face visual objects using a size comparison task. Pairs of names of objects were presented one by one on the screen. MJH and control participants were to decide which object of each pair was bigger. Previous results showed that in such a task, the closer the objects' size, the longer it takes to compare them [27,36]. Since this function is similar to the one obtained when participants make direct perceptual comparisons, the task is assumed to be mediated by the use of internal analog representations that contain relative size information; that is, visual mental images [36].

12.3.1. Method

We used Paivio's [36] normative data. Paivio [36] collected mean size ratings for 176 items. Based on these ratings, he created different types of pairs of objects that varied depending on the size ratio between the objects (high, medium, and low). Results in a size comparison task revealed a linear increase of response times as a function of the size ratio, which was taken as an evidence for the use of visual images representing relative size of objects.

Twelve pairs of each type, high, medium and low size ratios, were selected. Each pair was presented on the screen until the participants answered. A number, 1 or 2, corresponded to each name in a pair. Participants indicated their response (i.e. which one is bigger) by pressing either the key marked 1 or the key marked 2. For half of the trials the biggest object of the pair corresponded to the number 1 and for the other half to the number 2. Response times were registered.

12.3.2. Results and discussion

Mean RTs and percent errors are summarized in Table 8. Given that the function linking RT and size ratio has been well established, for both MJH and controls we performed planned comparisons (one-tailed t-tests with alpha corrected according to Rom's procedure).

For the control group, there was a monotonic increase in RTs with a reduction in the size ratio replicating Paivio's [36] results, $F(2, 22) = 15.12$, $P < 0.001$. RTs were longer to medium size ratio trials than to high size ratio trials, $t(11) = 3.52$, $P < 0.01$ and were longer to low size ratio trials than to medium size ratio trials, $t(11) = 3.1$, $P < 0.01$. Although accuracy was high overall ($M = 90.97\%$ correct) for the controls, there was a significant effect of size ratio, $F(2, 22) = 19.13$, $P < 0.01$. Planned comparisons showed that error rates on medium size ratio trials were lower than those for low size ratio trials, $t(11) = 5.23$, $P < 0.01$, however, medium and high size ratio trials did not differ, $t(11) < 1$.

As the same controls participated in Experiments 4a–c, we tested whether performance and RTs were correlated between experiments. The only significant correlation appeared between responses times in the Tail task (Experiment 4b) and the Size task (Experiment 4c). Controls who were faster in the Tail task were also faster in the Size task, $z = 2.6$, $P < 0.01$. However, there was no correlation between the two tasks in terms of correct responses.

Overall, MJH performed the comparisons more slowly ($M = 3.984$ s) than controls ($M = 1.529$ s, S.D. = 0.563 s), $t(11) = 15$, $P < 0.01$. MJH's longer RTs were true of all three levels of size similarity, high size ratio trials, $t(11) = 23.11$, $P < 0.01$, medium size ratio trials, $t(11) = 7.18$, $P < 0.01$, and low size ratio trials, $t(11) = 12.37$, $P < 0.01$.

Table 8
Mean response time (s) and mean percentage errors in the size comparison task (Experiment 4c)

	RTs (s)			Error rates (%)		
	High	Medium	Low	High	Medium	Low
MJH	5.683	4.050	6.337	8.34	25.00	16.67
Control participants	1.793 (0.583)	2.156 (0.874)	2.559 (1.058)	4.17 (5.62)	2.78 (5.43)	20.14 (10.33)

Standard deviations are given in parentheses.

MJH's overall error rate was within the controls' performance range ($M = 9.03$, S.D. $= 4$, range: 16.67–2.78%, one participant was at ceiling (>95%) and four were close to ceiling with 94.44%). One participant had a higher overall error rate than MJH. MJH's error rates were, however, significantly higher than those of the controls, $t(11) = 5.17$, $P < 0.01$. MJH's error rates exceeded those for the controls for high, $t(11) = 2.57$, $P < 0.05$, and medium size ratios, $t(11) = 14.17$, $P < 0.01$, but not for low size ratio, $t(11) = 1.16$. However, MJH's error rates were within less than 1 S.D. from the controls' average error rates for high and low size ratios and more than 4 S.D. above the controls' average for medium size ratios.

MJH's RTs and error rates did not increase monotonically with a decreasing ratio in the sizes of the objects. An item analysis over the 36 trials (12 per size ratio) showed that MJH's mean RTs to low size ratio trials tended to be longer than to medium size ratio trials, $t(8) = 1.56$, $P = 0.07$. However, in contrast to the expected relation between size ratio and performance, medium size ratio trials were actually shorter than those to high size ratio trials, although this effect fell short of significance, $t(8) = 2.26$, $P = 0.97$. This departure from monotonicity may be explained by a speed for accuracy tradeoff in that the error rates for the medium size ratio items were markedly higher than those for the low size ratio objects.

The pattern of results shown by MJH in this non-face imagery experiment is similar to that observed in face imagery experiments. MJH's performance was slower and less accurate (or in lower agreement) than that of the controls. However, MJH's percent correct score was in the range of the control participants and only 1 S.D. away from the controls' mean score. Moreover, results suggest that the function linking RTs and size ratio may be preserved for MJH.

13. General discussion

13.1. Summary of the results

The main results of this case study can be summarized as follows. MJH, a prosopagnosic since the age of 5, showed: (a) severely impaired perception but possibly less impaired imagery of facial configurations (Experiment 1); (b) severely impaired perception but somewhat preserved imagery of single facial features (Experiments 2 and 3); (c) severe impairment in naming familiar faces but less severe impairment in naming facial expressions (Experiments 1–3); (d) somewhat impaired imagery for color and non-facial visual objects (Experiments 4a–c); and (e) a general reduced speed in making judgments from imagery for both faces and non-face objects.

Although in imagery tasks MJH's mean accuracy/agreement scores were always lower than those of the controls and his responses always slower, it must be emphasized that the criterion for designating that his imagery was only "moderately" impaired was: (a) that MJH was within the control range; (b) that at least one of the control participants had an agreement/accuracy score that was worse than MJH's; and (c) that MJH was within 2 S.D. from the controls' mean score. In every experiment MJH's accuracy/agreement score was within the controls' range in imagery but that was never true in the perceptual tasks. Moreover, in every experiment, in imagery tasks, MJH's score was within 1 S.D. away from the controls' mean score whereas in perceptual tasks, his score was more than 2 S.D. away from that of the mean of the controls.

Somewhat complicating this interpretation of MJH's preserved imagery (compared to perception) relative to the controls, is the variability indicated by the poor agreement scores for the controls on the face imagery task in Experiment 1. But the same pattern was found in Experiment 2 where the controls had much greater agreement when performing the face imagery tasks.

Note that RT pattern showed that MJH's imagery was slower than normal. It is unlikely that such a deficit can be completely attributed to MJH's motor problems. Instead it may reflect a degraded memory representation.

An arguable result from the first experiment is a dissociation between impaired perception and possibly preserved imagery for facial configurations. Some previously reported case studies observed a similar dissociation [3,15]. A second, more consistent, result from Experiments 2 and 3 is a dissociation between impaired perception and relatively preserved imagery (or less impaired imagery) for facial features. Such a dissociation has been reported once in the literature by Bartolomeo et al. [3]. These authors proposed that imagery for facial configurations and/or features was preserved because: (a) face representations stored before the lesion were still available; and (b) imagery and perception share representations.

Such reasoning cannot hold in MJH's case as his prosopagnosia occurred so early in life that he could not have a pre-lesion intact store of face representations. Yet MJH's performance was relatively preserved in facial imagery tasks. Below we offer a speculation as to how these results might be accommodated by a general model of face recognition. We first discuss why faces may be processed in a manner that differs from most object processing, even when the to-be-discriminated objects are as similar as the to-be-discriminated faces.

13.2. Why a deficit specific to faces?

The processing of faces is often described as configural, i.e. as dealing with the interrelation between features in contrast to feature- or part-based processing [12,19,37]. Part-based processing is thought to be important for the shape-based identification of basic-level and most instances of subordinate-level classification of non-face objects [6,8] but not for the identification of faces, although it does play a role in face processing as well [13].

But how could a "configural representation" be conceptualized in neurocomputational terms? And why would a different form of representation be employed for faces and objects? Biederman and Kalocsai [7] have argued that the requirements for representing the fine metrics of the sculpting of a complex surface such as a face can be met by a representation initially computed by a V1 hypercolumn-like array that records and stores the multiscale, multioriented values of spatial (Gabor) filters in a two-dimensional coordinate space, as proposed in the *t*-jet model of von der Malsburg and his colleagues [28,41]. The activation values of medium- and larger-scale kernels whose receptive fields are centered on one part of a face are affected by variations in contrast occurring in distant regions of the face, leading to the configural phenomena evident in face recognition.

It is important to note that the memorial representation of a face, while preserving aspects of the original Gabor activation values computed in V1 itself (and some neurons in V2 and V4) would be in a non retinotopically mapped area, such as IT in the macaque [45] or the fusiform face area (FFA) in humans [25]. Unlike neurons at the earlier retinotopically mapped stages (V1–V4), these neurons would manifest a high degree of translation and size invariancethus the activation values would have to be relative in scale and position, defined in terms of cycles per object (face) rather than cycles per degree, and defined in object-centered, rather than retinotopic-centered, coordinates. The lesion that produces prosopagnosia then, would not be at the earlier stages where the initial filtering was performed, which would result in deficits in both face and non-face processing, but in the mapping of these activation values to the later stage or at the later stage itself (or both). The apparent preservation of MJH's and other prosopagnosics' facial imagery, both by subjective report and as evidenced here and in prior studies, suggests that for such individuals there is a deficit in the mapping, with some preservation of storage. Such a deficit would also accommodate Mangini and Biederman's [30] results in which MJH revealed sizable deficits relative to controls in a simultaneous match-to-sample task with face stimuli in which the matching and distractor choices were of different individuals of the same sex and same neutral expression. When the matching and distractor faces differed in gender or expression, MJH's performance was toward the lower end of those of the normal controls. Adding noise to the images resulted in much larger costs for MJH, relative to the controls, for all three tasks (gender, expression, and individuation), but markedly so for the individuation task.

Mangini and Biederman [31] have recently extended Biederman and Kalocsai's [7] proposal to provide a more explicit account of later-stage face processing, particularly in prosopagnosia. They argue that the deficit can be modeled in terms of a noisy covariance matrix by which the early stage representations, i.e. Gabor kernel activation values, are combined to form faces. If faces are reconstructed from the covariance matrix, adding noise does not make the reconstructed faces look noisy, only more similar to each other. This accords with the prosopagnosic's complaint that "all faces look the same." Greater exposure to instances of a particular gender and expression allow more accurate reconstruction, as all faces are but one of two genders and can be classified in terms of a dozen or so expressions.

Object recognition would employ the same Gabor jet representation in its initial stages (V1–V4). However, the output of the jets would be mapped onto units modeled as a structural description. Such a description would be based on nonaccidental (qualitative) differences in parts and relations, themselves based on contours at orientation and depth discontinuities. In the case of fine subordinate-level discriminations, this process would be achieved at a small scale [8]. A structural description describes the parts and relations among these parts and these elements can be mapped onto language and employed for object reasoning [8]. A structural description would also be activated when looking at a face, although it would typically be insufficient for distinguishing among similar faces. It would, nonetheless, specify aspects of the face area such as whether the person was wearing glasses or had a texture region (mustache or beard) or a mole or scar in a particular location or hair of a particular color and style. In general, both types of representation—spatial and structural description—would be computed for both objects and faces but the structural descriptions would typically be insufficient to distinguish similar faces from a large and unspecified set of faces.

The gain in the ability to represent the metric variations of a complex surface, enabled by a spatial representation such as that described in the present account, comes with a number of costs [7]. These costs constitute the signature phenomena of face recognition: In contrast to object recognition, the recognition of faces greatly suffers from: (1) planar rotation (assuming that humans cannot mentally recompute the kernel values produced by a two-dimensional rotation); (2) changes in contrast polarity; and (3) moderate changes in three-dimensional orientation and lighting direction. These three types of changes all affect the filter values but not the orientation and depth discontinuities that allow an object to be readily recognized from its line drawing [7]. The inability to describe verbally differences in similar faces derives from the incapacity to map thousands (if not millions) of Gabor coefficients and their positions directly onto language. Biederman and Kalocsai [7] tested the dependence of face recognition on the spatial composition of an image and found that whereas the matching of highly similar faces was very much affected by a change in the precise spatial composition, the matching of equally similar chairs was not.

13.3. Less impaired face imagery than face perception in prosopagnosia: a speculation

From the foregoing analysis, the difference between object and face representations is that the latter places greater reliance on the spatial representation to capture the metrics of a complex surface, viz. the face. If the full surface is to be represented in terms of the Gabor jet model, then jets whose receptive fields are centered throughout the surface have to be registered and stored. We speculate that MJH's prosopagnosia may be a condition in which he cannot simultaneously map the full lattice of jets to a representation that is sufficient for individuating faces, so only a subset of the jets is represented. This limitation could properly be regarded as a form of simultanagnosia. Indeed, MJH often complains that cluttered scenes, in general, cause him great difficulty. With faces, MJH searches for local distinctive features, such as the hairstyle or a beauty mark, which would be specified by a structural description. Note that this explanation of prosopagnosia may be limited to cases of apperceptive prosopagnosia, where there is a presentation of perceptual deficits, in contrast to cases of associative prosopagnosia, where there is a presentation of more conceptual deficits [16]. We do not regard the phenomenon of preserved face imagery in prosopagnosia as a test of Biederman and Kalocsai's [7] theory. The theory merely offers a framework within which one might understand the phenomenon.

Standard accounts of the imagery and perception interaction predict that face imagery and perception share the same representations. So how could face imagery, as assessed by tasks that require configural and metric processing, be preserved in prosopagnosia, particularly for faces that were learned after the onset of the malady? We speculate that on any one encounter, only a subregion of the face may be represented in high resolution so that a single feature (or a few features) might be extracted from the attended regions. However, with repeated encounters with a face, noisy or missing kernel values in one image may not be the same values that are missing on subsequent encounters. Consequently, the prosopagnosic individual has an opportunity to improve his representation in memory through integration of these partial percepts, although on any one encounter his perception is impaired. Although the memory representation may be improved with repeated encounters, as it is with all of us, it would by no means provide a high fidelity representation of the face for the prosopagnosic. Often the context or voice in which a famous or familiar face is experienced provides strong cues to the identity of the person so there is little ambiguity as to which person identification node (to use [10], terminology) the facial features are to be associated. Indeed, MJH's prosopagnosia is most strongly manifested in situations that exclusively provide face information (without other cues such as voice and gait) as to the particular individuals that he might encounter. Once he knows the composition of a small group of familiar people (as when he visits the laboratory), he is successful in distinguishing them by using such cues as hair color, gender, and facial features such as a mustache or location of a mole. However, as he notes, this process is effortful, "like doing difficult mental arithmetic." The use of contextual cues to access distinguishing features can be readily assessed by his chance performance in a forced-choice celebrity task ("Which of these two individuals is famous?") but his moderately high accuracy (72%, although with very long response times) in a name–picture verification task (Is this Bill Clinton?) (Mangini, unpublished data).

Another possible interpretation of the results comes from Moscovitch and Moscovitch's [35] proposition that in normal subjects the object and face systems interact to achieve recognition of inverted faces. The object system would build a facsimile of a face including parts of the face and their relations and transfer it to the face system, where it is used to achieve recognition of the inverted face. Such a collaboration between the two systems may explain the performance of MJH. Indeed, he may be using face representations computed initially in the object system rather than in the face system. In the terms of Biederman and Kalocsai's [7] model this would amount to using a structural description of a face rather than a spatial representation. MJH's better performance in imagery than in perception would then be explained by the fact that such facsimile representations (i.e. structural descriptions) may be more efficient in imagery tasks than in perceptual tasks because no identification is needed in imagery. That MJH's imagery is in the range of controls but not completely normal would be explained by the fact that, even in imagery tasks, facsimile representations are less adequate than configural ones.

13.4. Could such a dissociation be observed in other individuals?

Our results are supported by previous results from studies of covert face recognition in prosopagnosic patients showing covert recognition of faces encountered after the onset of the illness [14]. Such a result suggests that perceptual representations can be created for new faces that can be used covertly but not overtly for face recognition. However, MJH's performance may reflect a different type of mechanism for two reasons. First, in contrast to MJH, the patients reported in these studies were people who had had prior normal experience with faces. Second, we observed no covert face recognition in MJH [34] which suggests that representations build by MJH can not be used for covert recognition although they seem to be usable for imagery.

It is unclear whether the dissociation we observed between imagery and perception could be found again with other individuals who have no pre-lesion storage.

Additional evidence to MJH's case would come from cases of developmental (by contrast to acquired) prosopagnosia with intact face imagery. Unfortunately, there are only a few studies of developmental prosopagnosia and as far as we know none have tested for face imagery. Ariel and Sadeh [1] reported a case of congenital visual agnosia and prosopagnosia in an 8-year-old child. They tested visual imagery for non-face objects only. Their results are consistent with ours. Indeed, this child who was agnosic since his birth could draw and recall shape and color of objects from memory as well as other children. Ariel and Sadeh [1] did not propose an interpretation for this observation which they described as puzzling. However, an explanation can be proposed using the framework described earlier. In this framework, both face and object perception start with a Gabor jet representation from which a structural description is derived. Face perception would rely more on the stored Gabor representation and object perception on the structural description [7]. As the child described by Ariel and Sadeh [1] is prosopagnosic we can infer that he may present a perceptual deficit at the level of the activation of a (complete) Gabor representation in the latter face representation areas, such as the FAA. So, object perception may also be impaired at the same level (whether or not agnosia can be characterized as a deficit in simultaneously registering Gabor jets is beyond this discussion). Object imagery, presumed to share the same representations as object perception, would thus be impaired. However, an opportunity to improve the representation in memory may exist with repeated encounters with the same object, which would explain why object imagery may be relatively well preserved in that child. If Ariel and Sadeh [1] had recorded response times they may, however, have observed longer imagery latencies for that child compared to controls.

That face and object perception, and thus imagery, may use the same type of Gabor representations would also explain why MJH showed a mild impairment (longer latencies and higher error rates in some tasks than controls) in object imagery tasks, although he is not object agnosic. Indeed, damaged Gabor representations used in object perception, although not central to object perception, may lead to a slower object imagery. Note that whether faces are "special" and processed by a specific system or by a more general visual system dealing with all types of visual objects (face and non-face) is still under debate [20,22]. In the framework developed here, we assume that both faces and objects share the same first stage of the recognition system.

13.5. Recognizing face identity versus recognizing facial expressions

Also of importance was the observation that MJH has a deficit both in face identity processing and in facial expression processing. Models of face perception propose two independent routes for the processing of facial expression and face identity [10]. This hypothesis is supported by several experimental results [23,24,44]. Some prosopagnosic show preserved facial expressions processing [33,40], whereas others do not [33]. MJH shows no impairment in discriminating happy from unhappy [30]. In the current experiment he was perfect in judging when a face was happy. His deficit only appeared when discriminating among more negative emotions, including surprise. It is possible that happy faces, with their upturned mouths, provide the kind of qualitative cue that MJH uses with high accuracy in object recognition, leading, perhaps, to his normal performance on the Boston Naming Test.

The case of MJH adds evidence supporting associated deficits. However, mere association of deficits is not strong neuropsychological evidence. Indeed, two different systems can be damaged at the same time if they are anatomically related. So associated deficits in identity and facial expression processing do not necessarily mean that the two are processed by the same system. Moreover, in the case of MJH, the deficit for identifying facial expressions seemed not as great as the deficit for identifying faces. There are several possible explanations for this dissociation. One is that individuation requires classification into one of thousands, if not tens of thousands, of classes whereas expression discrimination requires less than 15. As a consequence, MJH may have had an opportunity to store facial expression representations but only a few familiar face representations before the lesion occurred. Another explanation is that there may be more stimulus energy supporting expression differences than individuation of faces. However, with strict controls of these variables, Mangini and Biederman [30] showed that MJH still shows a markedly greater deficit for individuation than expression or sex discrimination. Finally, the dissociation may be explained by the existence of two independent systems, one being less impaired than the other.

14. Conclusions

To conclude, we reported the case of a prosopagnosic who has some imagery for faces that he has difficulty perceiving. These are faces that he would not have learned prior to his lesion. This a priori paradoxical pattern of results is explained in the context of a new account of prosopagnosia in which only subsets of Gabor jets can be represented during perception. We propose that these partial percepts can be improved in memory through the error correction afforded through repeated exposures to the face. This improved memorial representation would allow a prosopagnosic patient to visualize faces encountered after the onset of the lesion.

Acknowledgements

Support was provided by grants from the Human Frontiers Science Program and the James S. McDonnell Foundation. We thank MJH for cooperating in our investigation. His good will and need for understanding prosopagnosia have been extraordinarily helpful in this enterprise. We also thank Bradley Duchaine for providing control participants' data in the control emotion naming experiment and Morris Moscovitch and two anonymous reviewers for thoughtful comments on a previous draft of this paper.

Appendix A

List of triads used in Experiment 1 with the proportion of each name selected by the control participants in the

Imagery task (I), No Hair task (NH) and Hair task (H). Choices made by MJH are indicated by an asterisk (∗).

Stimulus 1	Stimulus 2	Stimulus 3
Whoopie Goldberg	Whitney Houston	Oprah Winfrey
60 (I)	40 (I)	0* (I)
30 (NH)	70* (NH)	0 (NH)
30* (H)	70 (H)	0 (H)
Mel Gibson	Michael J. Fox	Pierce Brosnan
40 (I)	40* (I)	20 (I)
20 (NH)	70 (NH)	10* (NH)
10 (H)	80 (H)	10* (H)
Bette Midler	Hillary Clinton	Diane Sawyer
86* (I)	14 (I)	0 (I)
30 (NH)	40 (NH)	30* (NH)
20 (H)	60 (H)	20* (H)
Goldie Hawn	Dolly Parton	Joan Rivers
20 (I)	20* (I)	60 (I)
10 (NH)	0 (NH)	90* (NH)
10 (H)	0 (H)	90* (H)
Jim Carrey	Elvis Presley	Tim Allen
30 (I)	30* (I)	40 (I)
20* (NH)	0 (NH)	80 (NH)
20* (H)	10 (H)	70 (H)
Sharon Stone	Madonna	Marilyn Monroe
80* (I)	0 (I)	20 (I)
70 (NH)	10* (NH)	20 (NH)
60* (H)	10 (H)	30 (H)
Magic Johnson	Bryan Gumbel	O.J. Simpson
44 (I)	28* (I)	28 (I)
80* (NH)	20 (NH)	0 (NH)
70* (H)	30 (H)	0 (H)
Sean Connery	Billy Cristal	Burt Reynolds
10* (I)	80 (I)	10 (I)
10 (NH)	90* (NH)	0 (NH)
10* (H)	90 (H)	0 (H)
Roseanne Barr	Jane Pauley	Monica Lewinsky
0 (I)	50* (I)	50 (I)
40 (NH)	50* (NH)	10 (NH)
20 (H)	80* (H)	0 (H)
Joe Pesci	John Goodman	Richard Nixon
11* (I)	67 (I)	22 (I)
0 (NH)	90* (NH)	10 (NH)
0 (H)	100 (H)	0* (H)
John Travolta	Bruce Willis	Harrison Ford
40* (I)	20 (I)	40 (I)
70 (NH)	0* (NH)	30 (NH)
50* (H)	0 (H)	50 (H)
Heather Locklear	Jennifer Aniston	Julia Roberts
11* (I)	11 (I)	78 (I)
0 (NH)	10 (NH)	90* (NH)
0 (H)	30 (H)	70* (H)
Tom Cruise	Kelsey Grammer	Jack Nicholson
50 (I)	50* (I)	0 (I)

Appendix A. (*Continued*)

Stimulus 1	Stimulus 2	Stimulus 3
60 (NH)	30 (NH)	10* (NH)
67 (H)	33 (H)	0 (H)
Ronald Reagan	Danny Devito	Sylvester Stallone
10 (I)	80* (I)	10 (I)
0* (NH)	20 (NH)	80 (NH)
10 (H)	10 (H)	80* (H)
Sandra Bullock	Cher	Demi Moore
30* (I)	60 (I)	10 (I)
50* (NH)	30 (NH)	20 (NH)
20* (H)	40 (H)	40 (H)
Donal Trump	David Letterman	Robin Williams
20 (I)	50 (I)	30* (I)
50 (NH)	30* (NH)	20 (NH)
50* (H)	30 (H)	20 (H)
Leonardo DiCaprio	Robert Redford	Rod Stewart
50 (I)	10 (I)	40 (I)
70 (NH)	0* (NH)	30 (NH)
50 (H)	0 (H)	50* (H)
Kim Bassinger	Barbara Streisand	Pamela Anderson
10 (I)	70 (I)	20 (I)
0 (NH)	80 (NH)	20* (NH)
0 (H)	80 (H)	20* (H)
Bill Clinton	Clint Eastwood	Tom Brokaw
38* (I)	62 (I)	0 (I)
50 (NH)	50 (NH)	0 (NH)
60* (H)	40 (H)	0 (H)
George W. Bush	Richard Gere	Jay Leno
20 (I)	30 (I)	50* (I)
0 (NH)	50 (NH)	50* (NH)
0* (H)	50 (H)	50 (H)

List of questions used in Experiment 2 with the proportion of each answer made by the control participants in the Imagery task (I) and the Perceptual task (P). Choices made by MJH are indicated by an asterisk (∗).

	Imagery		Perception	
	Y	N	Y	N
Jay Leno				
Is his hair gray?	100*	0	100*	0
Are his eyes big?	40*	60	10*	90
Is his chin protruding?	100*	0	100*	0
Does he have sideburns?	60*	40	10*	90
Is his nose short and straight?	20	80*	30*	70
Does he have thick eyebrows?	80*	20	90	10*
Marilyn Monroe				
Is her hair dark?	0	100*	0	100*
Does she have full lips?	90*	10	100*	0
Does she have a mole?	100*	0	100*	0
Are her eyebrows thick?	20	80*	0*	100

Appendix A. (*Continued*)

	Imagery		Perception	
	Y	N	Y	N
Is her chin square?	20	80*	20	80*
Is her hair blond?	100*	0	100*	0
Clint Eastwood				
Are his eyes big?	20*	80	0	100*
Is his nose pointy?	90	10*	100*	0
Is his hair blond?	0	100*	0	100*
Does he have thin lips?	90	10*	100*	0
Does he have a pointy chin?	80*	20	30	70*
Does he have a wide mouth?	50*	50	20*	80
Woody Allen				
Is his nose big?	90*	10	100*	0
Is his hair blond?	10*	90	0*	100
Are his eyebrows bushy?	50*	50	20*	80
Is there a part in his hair?	70*	30	40*	60
Is there a dimple in his nose?	20	80*	20	80*
Does he have glasses?	90*	10	100*	0
Goldie Hawn				
Is her hair blond?	100*	0	100*	0
Does she have a wide mouth?	90	10*	100*	0
Does she have glasses?	0	100*	0	100*
Is her chin square?	0	100*	40	60*
Does she have big eyes?	90*	10	70*	30
Is her nose big?	0	100*	0	100*
Richard Nixon				
Is his chin protruding?	60	40*	20	80*
Is there a dimple in his nose?	40	60*	50	50*
Does he have a large forehead?	80*	20	90*	10
Is his hair blond?	0	100*	0	100*
Do his ears stick out?	60*	40	90	10*
Does he have sideburns?	40*	60	10	90*
Sylvester Stallone				
Does he have droopy eyes?	90*	10	100	0*
Is his hair dark?	100*	0	90*	10
Is his nose pointy?	30	70*	40	60*
Does he have a wide mouth?	30*	70	10*	90
Is his chin square?	80*	20	100*	0
Is there a part in his hair?	10*	90	60*	40
Julia Roberts				
Is her nose pointy?	90*	10	100	0*
Does she have a mole?	10	90*	0	100*
Are her eyebrows bushy?	0	100*	0	100*
Does she have big eyes?	90*	10	100*	0
Does she have a wide mouth?	90	10*	100*	0
Is her face round?	10	90*	10	90*
Ronald Reagan				
Does he have full lips?	30	70*	0	100*

Appendix A. (*Continued*)

	Imagery		Perception	
	Y	N	Y	N
Is his hair dark?	100*	0	100*	0
Does he have a pointy chin?	30	70*	0	100*
Are his eyebrows bushy?	40*	60	20	80*
Is his nose pointy?	60	40*	30	70*
Is there a part in his hair?	90*	10	100*	0
Robert Redford				
Is his hair blond?	80*	20	100*	0
Is his chin square?	80*	20	100*	0
Do his ears stick out?	20	80*	10	90*
Are his eyebrows bushy?	50	50*	0	100*
Does he have sideburns?	40	60*	50	50*
Is his hair dark?	0	100*	0	100*
J.F. Kennedy				
Does he have a large forehead?	80	20*	20	80*
Is his chin protruding?	20*	80	40	60*
Is there a part in his hair?	90*	10	100*	0
Is his face round?	70	30*	80	20*
Is his hair dark?	50*	50	60*	40
Is his nose short and straight?	60*	40	60	40*
Elvis Presley				
Is his hair dark?	100*	0	100*	0
Does he have thin lips?	20	80*	60*	40
Is his chin protruding?	80	20*	100*	0
Does he have sideburns?	100*	0	100*	0
Does he have droopy eyes?	60	40*	40	60*
Is his hair gray?	0	100*	0	100*

List of questions used in Experiment 3 with the proportion of each answer made by the control participants in the Imagery task (I) and the Perceptual task (P). Choices made by MJH are indicated by an asterisk (∗).

	Imagery		Perception	
	Y	N	Y	N
Sadness				
Are the corners of the lips turned down?	90*	10	90*	10
Are the lips opened?	30	70*	10	90*
Is the nose wrinkled?	10	90*	30	70*
Are the corners of the eyes turned down?	100*	0	100	0*
Are the eyebrows drawn together?	20*	80	100	0*
Are the eyes squinted?	10	90*	60	40*
Disgust				
Is the nose wrinkled?	80*	20	100*	0
Are the eyes opened wide?	20*	80	10	90*

Appendix A. (*Continued*)

	Imagery		Perception	
	Y	N	Y	N
Are the corners of the lips turned down?	80*	20	90*	10
Are the eyes squinted?	90	10*	90*	10
Are the corners of the lips turned up?	30	70*	20*	80
Crows feet wrinkles at corner of the eyes?	50	50*	100	0*
Fear				
Are the eyes squinted?	0*	100	0	100*
Are the teeth clinched?	30	70*	100*	0
Are the corners of the eyes turned down?	50	50*	30	70*
Are the lips opened?	70*	30	90*	10
Are the eyes opened wide?	90*	10	100*	0
Are the lips sealed tight?	30	70*	0	100*
Surprise				
Are the eyebrows up?	100*	0	100*	0
Are the corners of the lips turned down?	0	100*	0	100*
Is the forehead free of wrinkles?	20*	80	10	90*
Is the mouth opened?	100*	0	90*	10
Are the eyes opened wide?	100*	0	100*	0
Are the eyebrows drawn together?	0	100*	10	90*
Happiness				
Are the corners of the lips turned up?	90*	10	100*	0
Crows feet wrinkles at corner of the eyes?	80	20*	100*	0
Are the eyebrows drawn together?	0	100*	0*	100
Are the teeth clinched?	0	100*	0*	100
Is the forehead free of wrinkles?	60*	40	80*	20
Are the corners of the lips turned down?	0	100*	10	90*
Anger				
Are the eyes opened wide?	30*	70	0*	100
Are the eyebrows drawn together?	100*	0	80	20*
Is the mouth opened?	40	60*	0	100*
Are the eyebrows up?	30*	70	0	100*
Are the eyes squinted?	60*	40	100*	0
Are the lips sealed tight?	70*	30	100*	0

References

[1] Ariel R, Sadeh M. Congenital visual agnosia and prosopagnosia in a child: a case report. Cortex 1996;32:221–40.

[2] Arnaud M (Producer), Morel D. (Director). La Fin des Secrets. Program aired on ARTE television channel (France and Germany), November, 1999.

[3] Bartolomeo P, Bachoud-Levi A, De Gelder B, Denes G, Barba GD, Brugieres P, et al. Multiple-domain dissociation between impaired visual perception and preserved mental imagery in a patient with bilateral extrastriate lesions. Neuropsychologia 1998;36:239–49.

[4] Bartolomeo P, Bachoud-Levi A-C, Denes G. Preserved imagery for colours in a patient with cerebral achromatopsia. Cortex 1997;33:369–78.

[5] Behrmann M, Moscovitch M, Winocur G. Intact visual imagery and impaired visual perception in a patient with visual agnosia. Journal of Experimental Psychology: Human Perception and Performance 1994;20:1068–87.

[6] Biederman I. Recognition by components: a theory for human image understanding. Psychological Review 1987;94:115–47.

[7] Biederman I, Kalocsai P. Neurocomputational bases of object and face recognition. Philosophical Transactions of the Royal Society London: Biological Sciences 1997;352:1203–19.

[8] Biederman I, Subramaniam S, Bar M, Fiser J, Kalocsai P. Subordinate-level object classification reexamined. Psychological Research 2000;62:131–53.

[9] Bowers D, Blonder LX, Feinberg T, Heilman KM. Differential impact of right and left hemisphere lesions on facial emotion and object imagery. Brain 1991;114:2593–609.

[10] Bruce V, Young A. Understanding face recognition. British Journal of Psychology 1986;77:305–27.

[11] Calder AJ, Young AW, Keane J, Dean M. Configural information in facial expression perception. Journal of Experimental Psychology: Human Perception and Performance 2000;26:527–51.

[12] Carey S, Diamond R. From piecemeal to configurational representation of faces. Science 1977;195:312–4.

[13] Collishaw SM, Hole GJ. Featural and configurational processes in the recognition of faces of different familiarity. Perception 2000;29:893–909.

[14] De Haan EHF, Young AW, Newcombe F. Face recognition without awareness. Cognitive Neuropsychology 1987;4:385–415.

[15] De Renzi E, Di Pellegrino G. Prosopagnosia and alexia without object agnosia. Cortex 1998;34:403–15.

[16] De Renzi E, Faglioni P, Grossi D, Nichelli P. Apperceptive and associative forms of prosopagnosia. Cortex 1991;27:213–21.

[17] Ekman P, Friesen WV. Pictures of facial affect. Palo Alto (CA): Consulting Psychologists Press; 1976.

[18] Farah MJ. The neural basis of mental imagery. In: Gazzaniga MS, editor. The cognitive neurosciences. Cambridge (MA): MIT Press; 1995. p. 963–75.

[19] Farah MJ, Hammond KM, Levine DN, Calvanio R. Visual and spatial mental imagery: dissociable systems of representation. Cognitive Psychology 1988;20:439–62.

[20] Farah MJ, Rabinowitz C, Quinn GE, Liu GT. Early commitment of neural substrates for face recognition. Cognitive Neuropsychology 2000;17:117–23.

[21] Farah MJ, Wilson KD, Drain M, Tanaka JN. What is “special” about face perception? Psychological Review 1998;105:482–98.

[22] Gauthier I, Logothetis NK. Is face recognition not so unique after all? Cognitive Neuropsychology 2000;17:125–42.

[23] George MS, Ketter TA, Gill DS, Haxby JV. Brain regions involved in recognizing facial emotion or identity: an oxygen-15 PET study. Journal of Neuropsychiatry and Clinical Neurosciences 1993;5:384–94.

[24] Hasselomo ME, Rolls ET, Baylis GC. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Investigative Ophthalmology and Visual Science 1989;41:225.

[25] Kanwisher N. Domain specificity in face perception. Nature Neuroscience 2000;3:759–63.

[26] Kosslyn SM. Image and brain. The resolution of the imagery debate. Cambridge (MA): MIT Press; 1994.

[27] Kosslyn SM, Murphy GL, Bemesderfer ME, Feinstein KJ. Category and continuum in mental comparisons. Journal of Experimental Psychology: General 1977;106:341–75.

[28] Lades M, Vortbruggen JC, Buhmann J, Lange J, von der Malsburg C, Wurtz RP, et al. Distortion invariant object recognition in the dynamic link architecture. IEEE Transactions and Computers 1993;42:300–11.

[29] Luzzatti C, Davidoff J. Impaired retrieval of object–colour knowledge with preserved colour naming. Neuropsychologia 1994;32:933–50.

[30] Mangini MC, Biederman I. Differentiating expression, gender and identity in faces: comparing normals, the ideal observer, and a prosopagnosic. In: Proceedings of the Paper Presentation at the Meeting of the Vision Sciences Society, Sarasota, FL, May 2001.

[31] Mangini MC, Biederman I. Prosopagnosics have lower internal noise? In: Proceedings of the Poster Presentation at the 2nd Annual Meeting of the Vision Sciences Society, Sarasota, FL, May 2002.

[32] Mangini MC, Biederman I, Kosta A. Is greater accuracy for gender than person discrimination of faces a consequence of class uncertainty? Evidence from normals and a prosopagnosic. In: Proceedings of the Paper Presentation at the Meetings of the Association for Research in Vision and Ophthalmology, Ft. Lauderdale, FL, May 2000.

[33] McNeil JE, Warrington EK. Prosopagnosia: a reclassification. Quarterly Journal of Experimental Psychology: Human Experimental Psychology 1991;43:267–87.

[34] Michelon P, Biederman I. Automatic face processing for gender but not identity in prosopagnosia, submitted for publication.

[35] Moscovitch M, Moscovitch D. Super face-inversion effects for isolated internal or external features, and for fractured faces. Cognitive Neuropsychology 2000;17:201–19.

[36] Paivio A. Perceptual comparisons through the mind's eye. Memory and Cognition 1975;3:635–47.

[37] Rhodes G. Looking at faces: first-order and second-order features as determinates of facial appearance. Perception 1988;17:43–63.

[38] Riddoch MJ. Loss of visual mental imagery: a generation deficit. Cognitive Neuropsychology 1990;7:249–73.

[39] Schweich M, Bruyer R. Heterogeneity in the cognitive manifestations of prosopagnosia: the study of a group of single cases. Cognitive Neuropsychology 1993;10:529–47.

[40] Tranel D, Damasio AR, Damasio H. Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. Neurology 1988;38:690–6.

[41] Wiscott L, Fellous J-M, Kruger N, von der Malsburg C. Face recognition by elastic bunch graph matching. IEEE Pattern Recognition and Machine Intelligence 1997;19:775–9.

[42] Young AW, Hellawell DJ, Van De Wal C, Johnson M. Facial expression processing after amygdalotomy. Neuropsychologia 1996;34:31–9.

[43] Young AW, Humphreys GW, Riddoch MJ, Hellawell DJ, De Haan EHF. Recognition impairments and face imagery. Neuropsychologia 1994;32:693–702.

[44] Young AW, McWeeny KH, Hay DC, Ellis AW. Matching familiar and unfamiliar faces on identity and expression. Psychological Research 1986;48:63–8.

[45] Young MP, Yamane S. Sparse population coding of faces in the inferotemporal cortex. Science 1992;256:1327–31.