Super-recognizers: People with extraordinary face recognition ability

Richard Russell[1], **Brad Duchaine**[2], and **Ken Nakayama**[1]

[1]Harvard University, Cambridge MA, USA

[2]Institute of Cognitive Neuroscience, University College London, London, UK

Abstract

We tested four people who claimed to have significantly better than ordinary face recognition ability. Exceptional ability was confirmed in each case. On two very different tests of face recognition, all four experimental subjects performed beyond the range of control subject performance. They also scored significantly better than average on a perceptual discrimination test with faces. This effect was larger with upright than inverted faces, and the four subjects showed a larger 'inversion effect' than control subjects, who in turn showed a larger inversion effect than developmental prosopagnosics. This indicates an association between face recognition ability and the magnitude of the inversion effect. Overall, these 'super-recognizers' are about as good at face recognition and perception as developmental prosopagnosics are bad. Our findings demonstrate the existence of people with exceptionally good face recognition ability, and show that the range of face recognition and face perception ability is wider than previously acknowledged.

Developmental prosopagnosia is a condition marked by exceptionally poor face recognition ability despite normal vision and absence of brain damage or other cognitive deficits. The first case of developmental prosopagnosia (alternately called congenital or hereditary prosopagnosia) was only described in 1976 (McConachie, 1976), and the condition was thought to be very rare (Kress & Daum, 2003). However, recent years have witnessed estimates suggesting that 2 – 2.5% of the general population has the condition (Kennerknecht et al., 2006; Nakayama et al., 2006) and rapid growth in the number of cases reported (Barton, Cherkasova, Press, Intrilligator, & O'Connor, 2003; Behrmann & Avidan, 2005; Duchaine & Nakayama, 2006b; Le Grand et al., 2006).

The widespread use of terms such as 'condition', 'disorder', and 'impaired' to describe developmental prosopagnosia indicates a prevailing notion of face recognition ability as being either normal (i.e. roughly average) or pathological. The prevalence of this notion may be due in part to the apparent lack of people who are as far above average at face recognition as developmental prosopagnosics are below average. Finding such people would support an alternate notion of a broad distribution of face recognition ability, with (at least some cases of) developmental prosopagnosia representing the lower tail of the distribution.

We have been contacted by several people who with telling anecdotes have made persuasive claims to having superior face recognition abilities. The goal of the present study was to evaluate these claims through objective testing. Specifically, we sought to determine whether these people are indeed much better than normal at recognizing faces. A secondary goal was to determine whether they are also better than average at perceiving subtle

Correspondence to: Richard Russell, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge MA 02138, USA, Telephone: +1 (617)495-3884 x5, Fax: +1 (617) 495-3764, rrussell@wjh.harvard.edu.

differences between simultaneously presented faces and to measure their face inversion effect.

SUBJECTS

Following recent media coverage of our work on developmental prosopagnosia, several people contacted us because they believe that they are the opposite of prosopagnosic—that they are exceptionally *good* at recognizing faces. We have tested three of these people; CS, a 26 year old female PhD student, CL, a 40 year old female homemaker, and JJ, a 36 year old female municipal employee. We also tested MR, a 31 year old male computer programmer. MR was referred by his roommate (a developmental prosopagnosic who we tested in our laboratory) because MR had told his roommate that he thought himself to be much better than normal at recognizing faces.

They describe their face recognition abilities in strong terms. "It doesn't matter how many years pass, if I've seen your face before I will be able to recall it. It happens only with faces." (CS) "My boyfriend at the time used to call me a 'freak of nature' when it came to recognizing faces." (CL). All describe situations in which they correctly recognized near strangers who they had not seen for years and who had since undergone major changes of appearance (e.g. growing from a child into an adult or adopting a radically different hairstyle). All report being able to recognize actors playing minor characters or 'extras' in movies, television, and advertisements from other roles they have played: "I always knew I was very good at recognizing people from television and movies when they appeared in other things, no matter how small the role" (JJ). In comparison, many developmental prosopagnosics report not having much interest in film or television, because their inability to recognize the different characters makes it difficult to follow the story.

These subjects report altering their social behavior to accommodate the fact that they frequently recognize other people who do not recognize them. "I've learned to stop surprising people with bizarre comments like, 'Hey, weren't you at that so-and-so concert last fall… I recognize you'. Before that I'd occasionally make people uncomfortable with my recognition" (MR). "I do have to pretend that I don't remember [people], however, because it seems like I stalk them, or that they mean more to me than they do when I recall that we saw each other once walking on campus four years ago in front of the quad!" (CS).

FACE RECOGNITION

The purpose of this study was to determine objectively whether these people are much better than average at recognizing faces. Existing clinical tests of face recognition ability were designed to distinguish normal from sub-normal ability and are not sufficiently difficult to distinguish normal from super-normal ability. To address this issue, we designed new, difficult tests of face recognition ability.

Method

Tests

Before They Were Famous (BTWF): This is a famous face recognition test, in which subjects view 56 photographs of famous individuals and attempt to name them or provide uniquely identifying descriptions. To make the test very difficult, we selected photos taken of the people before they became famous. Many of the photographs were taken when the famous people were still children, and so correct identification requires generalization across substantial change in the appearance of the face. Figure 1 shows example stimuli similar to those used in the BTWF test. Images are presented for three seconds, but the subject has unlimited time to respond. Responses are scored as correct only if the subject correctly

names the person or provides a uniquely identifying description. Non-uniquely identifying descriptions (e.g. "he's a politician" or "she is an actress") are scored as incorrect. A weakness of famous face tests (including the BTWF test) is that previous exposure to the stimulus identities as well as face recognition ability determines performance. Thus the BTWF test alone is insufficient to assess face recognition ability. To address this limitation we also used a face recognition test in which subjects are trained to recognize new faces, so that exposure to the target faces is equated across subjects.

Cambridge Face Memory Test (CFMT) long form: The second test of face recognition is an adapted version of the Cambridge Face Memory Test (CFMT) (Duchaine & Nakayama, 2006a). The CFMT involves learning to recognize six unfamiliar male faces from three different views and then testing recognition of these faces in a three-alternative forced-choice task. The test starts easy, by testing recognition with the same images used during training. The second section is more difficult, using novel images that show the target faces from un-trained views and lighting conditions. The third section is more difficult still, consisting of novel images with visual noise added. The CFMT, with 72 trials, is insufficiently difficult to reliably distinguish normal from super-normal ability, so we made a long form of the test. The CFMT long form has a fourth section with 30 very difficult trials that come at the end of the original CFMT. Figure 2 shows example stimuli from the CFMT and the CFMT long form. The CFMT long form and the BTWF test, like the CFMT, will be available free of charge when used for research purposes.

Our goal was to make the fourth section much more difficult than the earlier sections. We selected target and distractor images that differed greatly from the training images and the previous test images. Six trials each use profile images (shown in Figure 2), images cropped to include only the internal features (brows, eyes, noses, mouths) of the faces, completely *uncropped* images (so that hair, ears, and necks are visible—as shown in Figure 2), images of smiling faces, and images of other emotional expressions.

Also to make the fourth section more difficult, the distractor images repeat much more frequently. This reduces the difference in familiarity between target and distractor faces. In the first three sections of the CFMT, there are 46 distinct distractor identities, each appearing on approximately 1/23 of the trials (while each target appears on 1/6 of the trials). In the fourth section, only the 14 distractor identities that appear most frequently in the first three sections (appearing on approximately 1/16 of the trials) are used. These 14 distractor identities appear on approximately 1 / 6.5 of the trials (as in earlier sections, the targets appear on 1/6 of the trials). Finally, the images in the fourth section contained even more visual noise than the images in the third section.

In 21 control subjects, the fourth section was indeed more difficult (M = 45%, SD = 12%) than the previous sections (third section M = 66%, SD = 19%, second section M = 75%, SD = 16%, third section M = 98%, SD = 5%). Performance on the five trial types of the fourth section did not differ significantly (ranging from 40–49%).

Subjects

CL, CS, and JJ were tested with the BTWF test and the CFMT long form. MR, who was tested before the CFMT long form was created, was tested with the BTWF test and the CFMT short form. Because the CFMT is a memory test, he could not be given the long form after having already taken the short form. The CFMT long form and BTWF tests were administered to 25 control subjects (11 female, mean age 30 years, S.D. 14), whose age did not significantly differ from the experimental subjects.

Results

Performance on the CFMT short form and BTWF face recognition tests is presented in Fig 3a, and performance on the CFMT long form and BTWF face recognition tests is presented in Figure 3b. Note that the y-axis—performance on the BTWF test—is the same in both parts of Figure 3. Each point represents a single subject. Because the CFMT short form is a subset of the items in the CFMT long form, the set of subjects is the same in both parts of Figure 3, with the exception of subject MR. Performance on the BTWF test was correlated $r = 0.70$, $p < 0.001$, N = 29, with the CFMT short form and $r = 0.72$, $p < 0.001$, N = 28, with the CFMT long form. The high correlation between the two measures supports the idea that they are both testing the same underlying processes.

On the short form of the CFMT, all four experimental subjects performed at ceiling—CS, CL, and JJ received the maximum score of 72; MR made one error. On both forms of the CFMT, all experimental subjects performed better than any control subject. On the BTWF test, the experimental subjects performed better than all but one control subject, who received the same score as CS. The experimental subjects performed significantly better than the control subjects on both the CFMT long form, $t_{26} = 3.60$, $p < 0.001$ and the BTWF test, $t_{27} = 6.20$, $p < 0.001$. Effect sizes for comparisons between control and experimental subjects were very large for both tests—CFMT long form, $d = 2.94$; BTWF test, $d = 3.57$.

On both tests of face recognition, performance by the experimental subjects was much higher than performance by the control subjects. This provides objective evidence in support of their subjective experience of having significantly better than normal face recognition ability. Because their recognition ability greatly exceeds that of normal people, we call them 'super-recognizers'.

FACE PERCEPTION

Having identified a group of individuals who are exceptionally good at face recognition, we can investigate their performance at other tasks, to see whether exceptional face recognition associates with other exceptional abilities. One ability that is closely related to face recognition is face perception, which we operationalize here as the ability to discriminate between simultaneously presented faces. In this experiment we examined whether the super-recognizers ability comes only from superior memory, or also from superior perception. We also tested developmental prosopagnosics, who represent the very low end of face recognition ability.

People are impaired at perceiving and recognizing images when they are inverted (turned upside-down), and this impairment is stronger for images of faces than of other objects or scenes. This 'face inversion effect' is often diminished or non-existent in prosopagnosics, and another goal of this experiment was to measure the magnitude of the inversion effect in the super-recognizers.

Methods

Participants—In addition to the four super-recognizers, we tested 26 new control subjects (15 female, mean age 33 years, S.D. 13), and 26 developmental prosopagnosic subjects (14 female, mean age 42 years, S.D. 14 years), neither group being significantly different in age than the super-recognizers. The developmental prosopagnosic subjects came from our study pool, having contacted us through our website (www.faceblind.org), and were tested in our laboratory. All reported having very poor face recognition ability and scored very low on a test of famous face recognition, an old/new face recognition test, and the Cambridge Face Memory Test (CFMT).

Test—To test the ability to perceive subtle differences between faces in the absence of memory, we administered the Cambridge Face Perception Test (CFPT) (Duchaine, Germine, & Nakayama, 2007). This test consists of trials in which subjects must sort a set of six frontal views of faces by similarity to a target face shown from a 3/4 viewpoint. The six faces were created by morphing the target face with six other faces, in varying proportion. In this way the six faces vary systematically in similarity to the target face. Because each of the six faces was produced by combining the target face with a different face, their appearances vary from the target face in different ways, which should make the task require face-specific processing. The test includes eight upright and eight inverted trials that alternate in a fixed, pseudo-random order, allowing measurement of the inversion effect for face perception. Figure 4 shows an example of a correctly sorted trial from the CFPT.

Results

Performance on the CFPT is shown in Figure 5. Repeated measures ANOVA with orientation as a within-subjects effect found main effects of orientation, $F_{1,53} = 114.31$, $p < 0.001$, and subject group, $F_{2,53} = 15.34$, $p < 0.001$, and a significant interaction between orientation and subject group, $F_{2,53} = 9.29$, $p < 0.001$. Overall, the super-recognizers performed the task better than the control subjects, who in turn were better than the prosopagnosics. As indicated by the interaction, the differences between subject groups were much greater for upright faces than for inverted faces. This can be appreciated by noting that the differences between the groups are much larger on the x-axis than on the y-axis (which cannot be attributed to floor performance in the inverted condition; chance performance is 94 errors). Consistent with recent findings (Megreya & Burton, 2006), performance on the upright and inverted conditions was significantly correlated ($r = 0.626$, $p < 0.001$, N = 56).

Because the range of performance was so large, we measured the size of the inversion effect as the difference between upright and inverted performance, normalized to upright performance ((upright – inverted)/upright). Super-recognizers had a larger inversion effect than controls, who in turn had a larger effect than developmental prosopagnosics. The mean inversion effect for super-recognizers was 2.3 (SD 0.2), while it was 1.1 (SD 0.7) for controls, and 0.3 (SD 0.3) for developmental prosopagnosics. These group differences were significant, $F_{2,55} = 27.38$, $p < 0.001$, indicating that subject groups with better face recognition performance also have a larger face inversion effect.

With upright faces, effect size for the comparison between control subjects and super-recognizers, $d = 1.48$, was very similar to the effect size for the comparison between control subjects and developmental prosopagnosics, $d = 1.59$. Compared with the face recognition tasks, the upright face perception task yielded a smaller effect size for the comparison between control subjects and super-recognizers. For the inversion effect with the CFPT, the effect size comparing controls and super-recognizers was $d = 2.39$; comparing controls and developmental prosopagnosics the effect size was $d = 1.35$. Overall, the results from the face perception task support the notion that the super-recognizers' abilities are not limited to face memory—they extend to face perception, albeit to a weaker degree. However, the difference does not much extend to generic differences in perceptual discrimination—the effect size between the super-recognizers and the control group at discriminating inverted faces was only $d = 0.29$, and between the developmental prosopagnosics and the control group was $d = 0.73$.

These results suggest that face recognition and face perception ability are associated. To further test this idea, we ran 23 new control subjects on *both* the CFMT long form and the CFPT. Performance on the CFMT long form was correlated $r = 0.667$, $p < 0.001$, N = 23, with the CFPT upright, but only $r = 0.042$, $p = 0.851$, N = 23, with the CFPT inverted. This

provides additional support for the notion that (upright) face recognition ability is associated with upright face perception ability, but not with inverted face perception ability.

GENERAL DISCUSSION

All four experimental subjects performed far above average on two different tests of face recognition. On a test of familiar face recognition ability (BTWF) and a test that requires learning unfamiliar faces (CFMT), their scores were better than any of the 25 control subjects. The results support their claim of having extraordinary face recognition ability. This provides initial evidence for the existence of people with face recognition ability that is far above average: face 'super-recognizers'.

Identifying people as 'super-recognizers' because they are much better than average at face recognition raises several questions. Do super-recognizers form a distinct group? We have no evidence that their face recognition processes are qualitatively different than normal. It seems more likely that super-recognizers are simply the high end of a broad distribution of face recognition ability. Though the cut-off is likely arbitrary, we believe it is relevant that the people we tested arrived independently at the conclusion that they are much better than average. How much better than normal are the super-recognizers described here, and is it a meaningful difference? One approach to addressing this question is to compare the difference between control and super-recognizer performance with the difference between control and developmental prosopagnosic performance. Most developmental prosopagnosics we have tested in our laboratories score around 2–3 standard deviations below normal on the CFMT short form. In comparison, three super-recognizers scored around 2 standard deviations above the mean on the CFMT long form. Similarly, on the CFPT, effect sizes were very similar for the comparisons between prosopagnosics and controls, and between super-recognizers and controls. In both face recognition and face perception, *the super-recognizers are about as good as many developmental prosopagnosics are bad*. This suggests the possibility that many cases of developmental prosopagnosia represent the low end of normal face recognition ability rather than a qualitatively different kind of face processing.

Whether one or more distributions underlie the range of face recognition ability, the existence of super-recognizers indicates that this range is larger than has been appreciated. On one end of this range lie developmental prosopagnosics, some of whom even have difficulty recognizing members of their nuclear family. On the other end of this range lie super-recognizers, who frequently recognize complete strangers out of context after many years. The discovery that the range of face recognition ability is wide rather than narrow is relevant in applied and theoretical contexts.

Various social institutions are premised upon the false assumption that all people have similar face recognition ability. Many occupations rely critically on face recognition. Given the estimated prevalence of prosopagnosia, it is likely that thousands of prosopagnosic security personnel are checking ID photographs daily. It would be preferable for people in these occupations to have average or even superior face recognition ability. Face recognition ability could be assessed as part of the selection process for such jobs. Similarly, the criminal justice system relies heavily on eyewitness identification for investigating and prosecuting crimes (Wells & Olson, 2003). Eyewitness identification accuracy is significantly correlated with performance on standardized face recognition tests (Morgan et al., 2007); witness face recognition ability could be assessed after a crime as a factor to be considered in postdicting eyewitness identification accuracy.

Our third experiment, on face perception and the face inversion effect, illustrates how comparing the super-recognizers and developmental prosopagnosics can be used to investigate basic mechanisms of face processing. The super-recognizers were better at perceiving differences between faces than control subjects, who in turn were better than developmental prosopagnosics. Extending recent work with normal (Wilhelm, Herzmann, Kunina, & Sommer, 2007) and prosopagnosic (Duchaine, Germine, & Nakayama, 2007) populations, this indicates an association between face recognition ability and face perception ability across the *entire* range of face recognition ability. This suggests that variation in face recognition ability is not solely a function of memory, but also of perception.

Acquired and developmental prosopagnosics often have small or non-existent face inversion effects (Behrmann & Avidan, 2005; de Gelder & Rouw, 2000; Duchaine, Germine, & Nakayama, 2007; Duchaine, Yovel, Butterworth, & Nakayama, 2006). The current work replicates this finding and extends it by showing that super-recognizers, with exceptionally good face recognition, have *larger* face inversion effects than normal controls. This indicates that the lack of inversion effects in prosopagnosics may not indicate qualitatively different processing, as argued by de Gelder & Rouw (2000) and Duchaine et al. (2006), because the size of the face inversion effect is associated with face recognition ability across the entire range of face recognition ability. The continuous distribution of inversion decrements (Figure 5) further supports the notion of quantitative rather than qualitative differences between prosopagnosic and normal face processing.

Our discovery of super-recognizers demonstrates that people can not only be much worse than average at face recognition (as in developmental prosopagnosia), but also much better than average. This provides support for the possibility that developmental prosopagnosia—in some cases—represents a low-functioning version of normal face processing rather than an impaired version.

Acknowledgments

REFERENCES

Barton JJ, Cherkasova MV, Press DZ, Intrilligator JM, O'Connor M. Developmental Prosopagnosia: A study of three patients. Brain and Cognition. 2003; 51:12–30. [PubMed: 12633587]

Behrmann M, Avidan G. Congenital prosopagnosia: face-blind from birth. TRENDS in Cognitive Sciences. 2005; 9(4):180–187. [PubMed: 15808500]

de Gelder B, Rouw R. Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? NeuroReport. 2000; 11:3145–3150. [PubMed: 11043539]

Duchaine B, Germine L, Nakayama K. Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. Cognitive Neuropsychology. 2007; 24(4):419–430. [PubMed: 18416499]

Duchaine B, Nakayama K. The Cambridge Face Memory Test: Normal performance and an investigation of its validity using inverted performance and prosopagnosic subjects. Neuropsychologia. 2006a; 44(4):576–585. [PubMed: 16169565]

Duchaine B, Nakayama K. Developmental prosopagnosia: a window to content-specific face processing. Current Opinion in Neurobiology. 2006b; 16:166–173. [PubMed: 16563738]

Duchaine B, Yovel G, Butterworth EJ, Nakayama K. Prosopagnosia as an impairment to face-specific recognition mechanisms: Elimination of the alternative hypotheses in a developmental case. Cognitive Neuropsychology. 2006; 23:714–747. [PubMed: 21049351]

Kennerknecht I, Grueter T, Welling B, Wentzek S, Horst J, Edwards S, et al. First Report of Prevalence of Non-Syndromic Hereditary Prosopagnosia (HPA). American Journal of Medical Genetics Part A. 2006; 140A:1617–1622. [PubMed: 16817175]

Kress T, Daum I. Event-related potentials reflect impaired face recognition in patients with congenital prosopagnosia. Neuroscience Letters. 2003; 352:133–136. [PubMed: 14625041]

Le Grand R, Cooper PA, Mondloch CJ, Lewis TL, Sagiv N, de Gelder B, et al. What aspects of face processing are impaired in developmental prosopagnosia? Brain and Cognition. 2006; 61:139–158. [PubMed: 16466839]

McConachie HR. Developmental Prosopagnosia: A single case report. Cortex. 1976; 12:76–82. [PubMed: 1261287]

Megreya AM, Burton AM. Unfamiliar faces aren't faces: Evidence from a matching task. Memory & Cognition. 2006; 34:865–876. [PubMed: 17063917]

Morgan CAI, Hazlett G, Baranoski M, Doran A, Southwick S, Loftus E. Accuracy of Eyewitness Identification is significantly associated with performance on a standardized test of face recognition. International Journal of Law and Psychiatry. 2007; 30:213–223. [PubMed: 17449097]

Nakayama K, Garrido L, Russell R, Chabris CF, Gerbasi M, Duchaine BC. Developmental prosopagnosia: phenotypes and estimated prevalence. Soc. Neurosci. Abstracts. 2006 2006-A-7594-SfN.

Wells GL, Olson EA. Eyewitness Testimony. Annual Review of Psychology. 2003; 54:277–295.

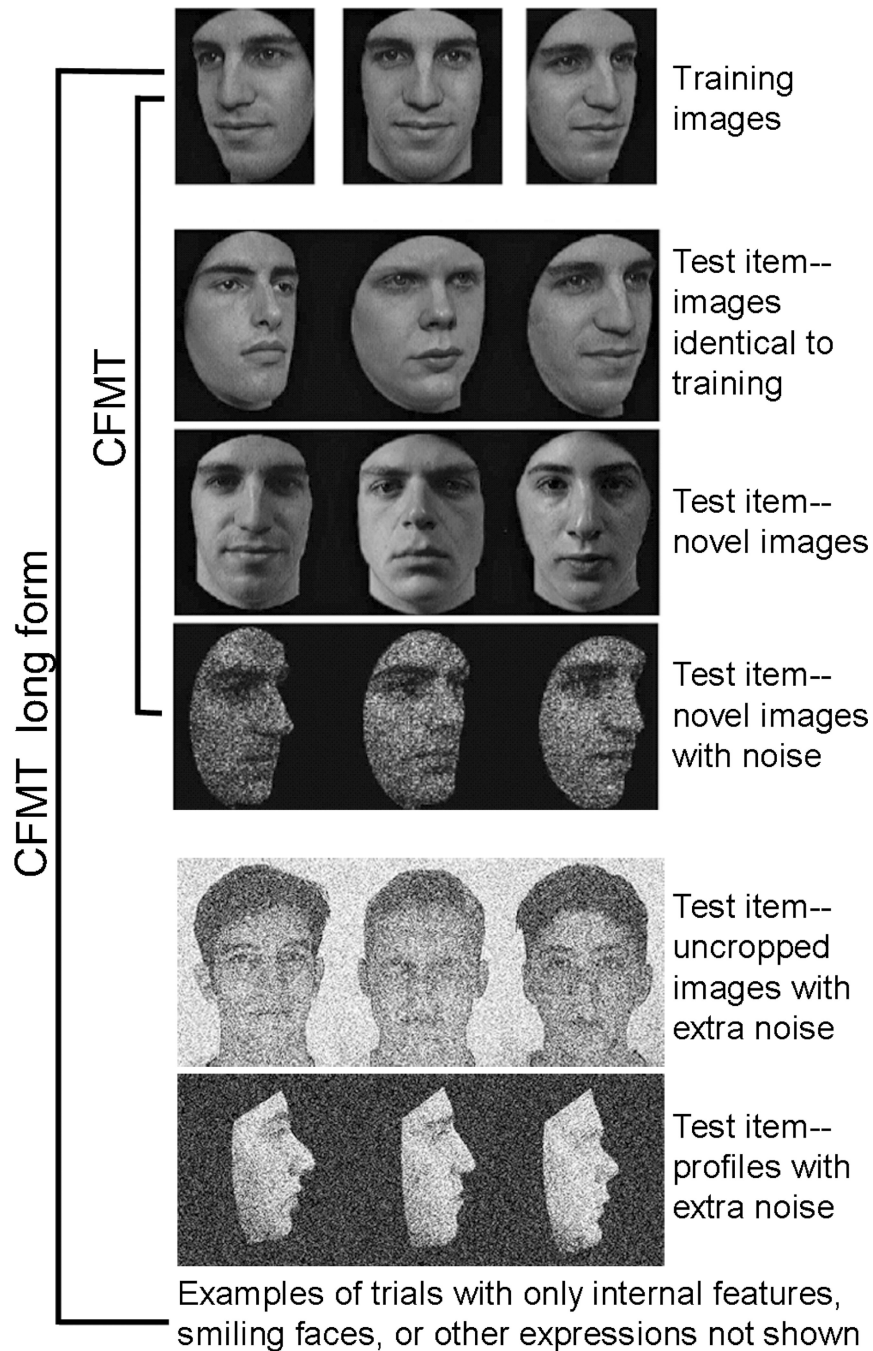Wilhelm O, Herzmann G, Kunina O, Sommer W. Face Cognition: A set of Distinct Mental Abilities. 2007 Available from Nature Precedings <http://hdl.handle.net/10101/npre.2007.1385.1>.

Figure 1.
Examples of faces similar to those used in the BTWF test[1].

Figure 2.
Examples of faces similar to training and test images in the CFMT. None of these particular images was used in the test. The top row shows training views of a target face. The following three rows show test items from the CFMT. The CFMT long form includes all the CFMT items, plus an additional 30 items. The bottom two rows show test items from the CFMT long form.
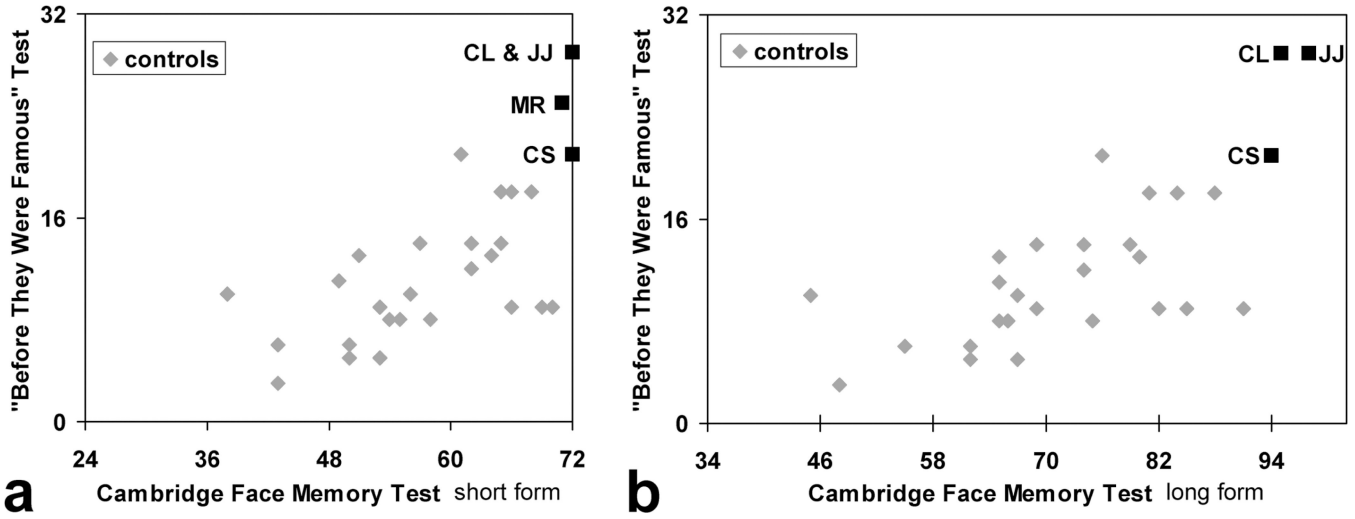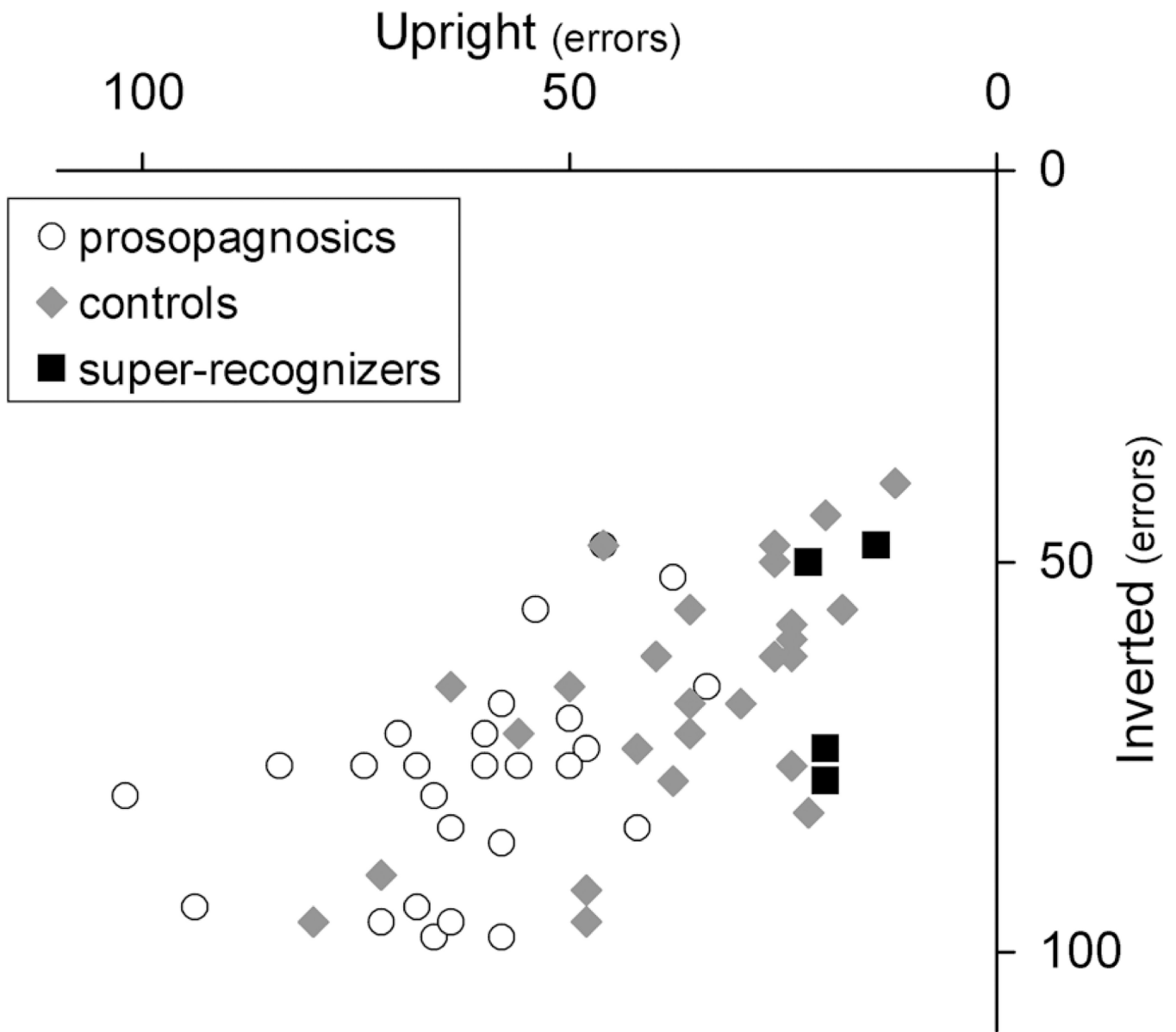
Figure 3.
a) Performance on BTWF test plotted against CFMT short form. b) Performance on BTWF test plotted against CFMT long form. The dependent variable in both tests is the number correct. Minimum values on each axis represent chance performance on the given task; maximum performance is 56 on the BTWF test, 72 on the CFMT short form, 102 on the CFMT long form.
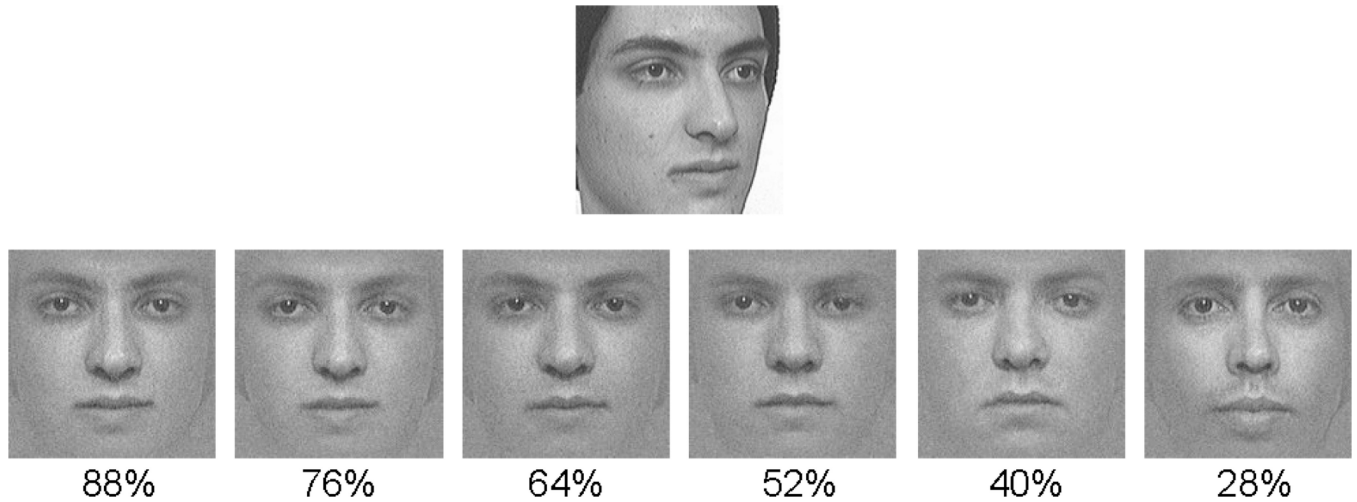
Figure 4.
Images from an item in the CFPT. Numbers under each image indicate the percentage of the target face in the image. In a test item, the six frontal shots are presented in a random order, and participants sort them based on similarity to the target image (the ¾ view at top). Turn the page upside down to experience the effect of inversion.

Figure 5.
Performance of individual subjects on the upright (x-axis) and inverted (y-axis) portions of the CFPT test of face discrimination. Performance is measured in errors (each space that a face must be moved be in the correct location is one error); smaller numbers indicate better performance.