PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Becoming a Face Expert [and Discussion]

Susan Carey, S. De Schonen and H. D. Ellis

Phil. Trans. R. Soc. Lond. B 1992 **335**, 95-103
doi: 10.1098/rstb.1992.0012

Becoming a face expert

SUSAN CAREY

Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, U.S.A.

SUMMARY

Young children do not form representations of newly encountered faces as efficiently as do adults. A first step in explaining this difference, like any age-related change, is locating its source. A major source of the improvement is acquisition of knowledge of faces *per se*, as opposed to age-related changes in general pattern encoding or memorial skills.

Two consequences of expertise at individualizing members of classes that share a basic configuration are known: a large inversion effect and a caricature advantage. It is possible that both of these effects reflect increased reliance, with expertise, on configuration distinguishing features. Several phenomena that indicate that inversion interferes with the encoding of configural aspects of faces are reviewed. Finally, developmental data are presented that confirm the suspicion that there are at least two distinct sources of the vulnerability of face encoding to inversion, perhaps reflecting two distinct senses of 'configural encoding' of faces, only one of which is implicated in adult expertise at face encoding.

1. INTRODUCTION

We read momentary expression, character, age, sex, and, of course, personal identity from faces. Here I am concerned with the mental processes underlying adult expertise at recognition of individual faces. A familiar face is identified in about 0.5 s, in spite of the very many faces stored in memory, and in spite of the high degree of similarity among faces. Adults can successfully encode large numbers of new faces from photographs inspected only briefly (e.g. 50 seen for 5 s each) and subsequently pick these from distractors at recognition rates of over 90%. Furthermore, once well encoded, representations of faces are not interfered with by newly encoded representations. One demonstration study found above 90% recognition of yearbook photos of schoolmates, independent of class size between 90 and 900, and independent of elapsed time from graduation between 3 months and 35 years (Bahrick *et al.* 1975)!

2. WHY STUDY THE DEVELOPMENT OF FACE PERCEPTION?

The development of face recognition stands as a paradigm case of perceptual development; thus, explaining it is an important goal for the field of developmental psychology. Moreover, details of how this capacity is put together during development bear on controversies concerning adult face encoding. In this review, I develop points within each theoretical context: understanding perceptual development in general, and understanding face perception in adults. With respect to general issues of development, I discuss the problem of locating the source of age-related changes. The question is whether children are terrible at face recognition for reasons having anything to do with their being children, as opposed to their being novices. With respect to face recognition in general, I discuss how data from development bear on explaining why faces are so hard to encode when they have been turned upside down.

3. THE COURSE OF DEVELOPMENT OF FACE PERCEPTION

Evolution has provided the baby a running start at face recognition. Neonates preferentially track moving schematic faces, in contrast to other patterns of comparable complexity, including upside-down schematic faces (Goren *et al.* 1975; Johnson *et al.* 1991). And within days, babies have formed representations that support discrimination of their mother's face from a stranger's face (Bushnell *et al.* 1989; Walton & Bower 1991). During the first 6 months, the baby comes to discriminate young from old faces, male from female faces (see, for example, Fagan (1979)). By 5–7 months, babies succeed at encoding new faces from minimal exposure, subsequently discriminating these from faces they have not seen before.

In spite of this impressive beginning, face recognition undergoes protracted development. Compared with normal adult levels of skill, young children are profoundly deficient at face encoding. On some clinical tasks, children under 10 years perform at a level diagnostic of right hemisphere brain damage, whereas 10-year-old children, although worse that adults, perform in the normal adult range (Benton & Van Allen 1973; Carey *et al.* 1980). Figure 1 shows the typical developmental function on recognition memory tasks.

Phil. Trans. R. Soc. Lond. B (1992) **335,** 95–103
Printed in Great Britain

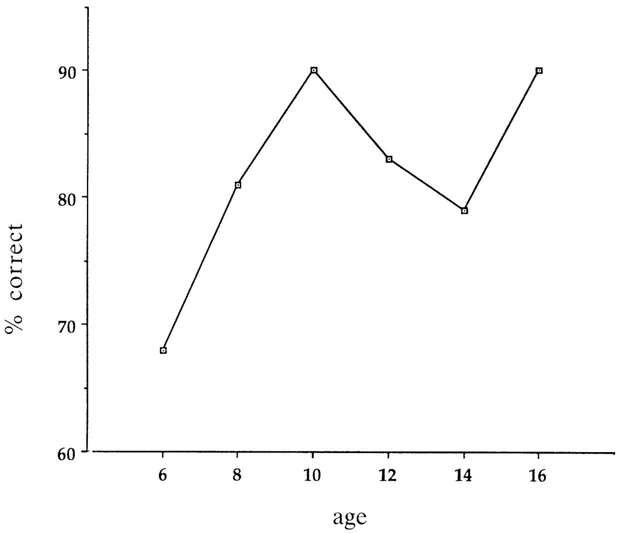

Figure 1. Developmental function: recognition memory for unfamiliar faces.

In this case, subjects were shown 36 photographs for 5 s each and then asked to discriminate these from new photos. Six-year-olds performed just barely better than chance, compared with the ceiling performance of adults. If the size of the set of faces to be encoded is varied, to ensure performance levels of 85% or better, children aged 3 years and under succeed only at a set size of one, whereas by age 10, children can manage sets of ten or more.

Several studies indicate that the improvement over the first decade in life is followed by a decline around age 12 (figure 1, see also Carey *et al.* (1980); Flin (1980)). Puberty is implicated in this disruption of performance; two studies have shown that girls in the midst of pubertal change perform worse than prepubescent or postpubescent controls matched for age (Diamond *et al.* 1983). Of course, this correlation is consistent with either a biological or a cognitive explanation for the disruption of face encoding at age 12. First (the biological): hormonal changes at puberty may directly affect mental processes. Alternatively (the cognitive): the child's reaction to the bodily changes at puberty might conceivably lead to a new interest in personal appearance and style, which in turn might lead to a reorganization of representations of faces. Flin (1983) found a disruption in performance for recognition of non-face stimuli as well, militating against any cognitive interpretation specific to the representation of faces. At any rate, the explanation of the decline in performance in early puberty remains to be worked out, and I do not speculate here. Rather, I concentrate on the issues raised by the improvement in face encoding skills during the first decade of life.

4. THE DESCRIPTIVE PROBLEM: WHAT IS DEVELOPING?

As children get older, we expect them to get better at just about anything. This is why the decline in performance at age 12 seems the aspect of the developmental function most in need of explanation. But the fact that we expect children to improve does not mean we are forgiven the task of explaining why they do. A first step in providing an explanation is locating the source of improvement. What is it about face recognition in young children that makes them worse at it than older children and adults?

Recognition memory for faces can be broken down into two components: the formation of a representation of a previously unseen face (the initial encoding) and the process of matching a current stimulus with a stored representation (recognition). Several lines of evidence converge on the conclusion that young children's problems concern the encoding of new faces, rather than the processes of recognition, *per se*. Most straightforwardly, several of the tasks that diagnose subjects' face encoding skills do not involve memory at all, instead requiring deciding whether two different photographs depict the same person or not. When the faces differ in expression, angle of view, direction of lighting, hairstyle, clothing, or even size (of photo), young children perform very badly (Benton & Van Allen 1973; Saltz & Sigel 1967; Diamond & Carey 1977; H. D. Ellis, this symposium). Such matches are mediated by representations, of course. Apparently, young children have difficulty encoding faces in terms of features that are invariant over such changes, just as they have difficulty encoding faces in terms of features that differentiate one face from another (figure 1).

5. IS THE CHILD LIMITED AT ENCODING ANY PATTERN?

The features that distinguish faces one from another must be learned from experience with them. We all have had the experience of encountering a new race of faces and having great difficulty telling people apart. Does children's poor face encoding skills derive from a similar lack of knowledge of faces from their own social group? Alternatively, perhaps the young child is worse than the adult at encoding any complex pattern, and the improvement at face encoding during the first decade of life derives from improvement at pattern encoding, in general. We are asking here whether there is anything developmental in the function on figure 1. That is, do information processing limitations of children contribute to their poor performance at face encoding, or are they simply novices at the task?

One can see how, in principle, this question should be addressed. One need only compare the developmental course for face encoding with that for some task which places comparable demands on a pattern encoder but for which the adult has no more experience than has the child. In practice, however, it is difficult to meet these desiderata. We have tried twice.

Our first attempt involved comparing the developmental course of recognition memory for upright faces with that for inverted faces. For adults, orientation markedly interferes with encoding success. Moreover, performance on upright faces is not even correlated with performance on inverted faces, suggesting that adults are not able to recruit all their knowledge of faces when encoding inverted faces (Phillips & Rawles 1979). In terms of pattern complexity, upright and inverted faces are identical, so as meaningless patterns, both place equal demands on a pattern encoder. Thus, a comparison of the developmental course of encoding upright and inverted faces would help tease apart the contribution of general improvement at pattern encoding skills (applicable to upright and inverted faces) from the contribution of acquisition of expertise at face encoding, *per se* (perhaps applicable to upright faces only). The results from the developmental studies are clear: as long as ceiling and floor effects are controlled for, face encoding is affected by orientation at every age tested, even in infancy (Fagan 1979; Carey 1981; Flin 1983). At least by age 5 months, new faces are being encoded relative to specific knowledge of faces, knowledge better exploited from upright than from inverted stimuli. Equally clear is an age by orientation interaction. That is, the magnitude of the inversion effect increases with age (see Carey 1981; Flin 1983).

What is not so clear is how to interpret this pattern of results with respect to the question at hand. That children improve more on upright faces than inverted faces shows, I would argue, that part of their improvement is due to acquisition of knowledge of faces, *per se*, knowledge that cannot be applied as efficiently to the encoding of inverted faces. Children's improvement on inverted faces could reflect acquisition of general pattern encoding skills that contribute to the emerging expertise at face encoding. But it could also reflect increasing ability to exploit what is known about upright faces in the encoding of inverted faces. Thus, while the age by orientation interaction indicates that part of the development of face encoding in the first decade of life is due to the acquisition of face specific expertise, these data leave open the question whether all is due to this source.

In our second attempt to address this issue, we studied the developmental course of encoding random dot patterns of the sort first studied by Posner & Keele (1968). Prototypical nine-dot patterns are randomly generated. For each prototype, a set of patterns is created, ranging from small distortions of the prototype to relatively large distortions. The subject is given the task of learning to categorize sets of distortions from a single prototype together, distinguishing them from the patterns derived from different prototypes. Because the patterns are randomly generated, no subject can have had any experience with the features which differentiate them. If the training set consists of large distortions from the prototype, this task places great demands on a pattern encoder. Finally, after a training criterion has been met, data from generalization trials allow a characterization of how the patterns have been encoded.

In three separate studies, using several different measures of encoding success, the following pattern emerged: 6- and 10-year-olds did not differ; both groups encoded the patterns less adequately than did 12-year-olds and adults. Between ages 6 and 10 years, the period of huge changes in face encoding success, the developmental function was totally flat (Diamond & Carey 1990)! This task requires the subject to extract a prototype, based on the configuration of several points, and to classify novel exemplars in terms of configural similarity to that prototype. All these are certainly aspects of the requirements face encoding places on a pattern encoder. We can tentatively conclude that development of these domain general pattern encoding skills do not underly improvement at face encoding in the years before age 10. These data are consistent with the conclusion that all improvement during these years is due to acquisition of specific expertise about faces. Naturally, the question is still open; future studies will falsify this conclusion if some other general pattern encoding skill applicable to faces is shown to improve over these years.

6. WHAT DO WE KNOW ABOUT THE ACQUISITION OF EXPERTISE, INDEPENDENT OF CONSIDERATIONS OF AGE?

We have arrived at a fairly trivial conclusion. Young children are less able than older children and adults to encode new faces in terms of features that support differentiation of that face from others, and this lack derives largely from lack of knowledge of faces, *per se*. Our task, therefore, is to specify what expertise at face encoding consists in, so as to characterize what the child must acquire. We seek hints about what is different in how young children (novices) encode faces from how adults (experts) do so.

One indication that children are doing something different from adults, rather than just less of what adults do, is the fact that children are less affected by inversion. If we understood the large inversion effect on face encoding, we might have the beginning of an understanding of what changes with acquisition of expertise.

Encoding individual faces is more affected by inversion than is the encoding of individuals from almost any other class studied to date: houses, bridges, stick figures of men, buildings, landscapes, dog's faces (Diamond & Carey 1986; Scapinello & Yarmey 1970; Yin 1969, 1970*a*). In these studies, the stimuli to be encoded are presented in the same orientation both during inspection and recognition; inverted stimuli are first seen upside down and also presented for recognition upside down. The difficulty is in forming an adequate representation of an inverted face, not in coping with a mismatch of orientation between test and recognition. Typically, one finds a 20–30% decrement in recognition accuracy for inverted compared with upright faces, whereas one finds only a 0–10% decrement when stimuli from the other classes are inspected and recognized upside down.

This result has been taken by some to indicate that faces are a unique stimulus class; they pose unique problems to a pattern encoder, and perhaps even have dedicated neural substrate for the solutions to these problems (e.g. Yin 1970*b*). No doubt there are dedicated neural underpinnings to the innate representation of faces, and no doubt there are many areas of

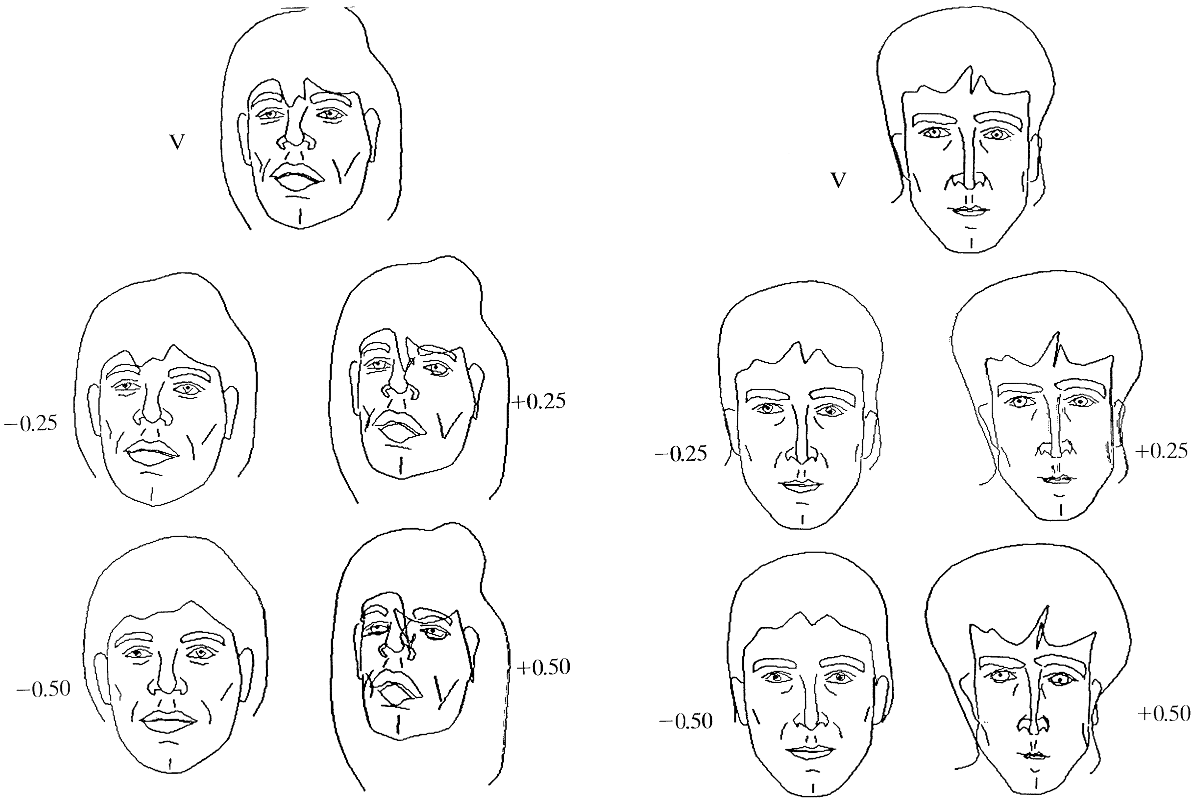

Figure 2. Examples of caricatures (+0.25, +0.5), veridical line drawings (V), and anticaricatures (−0.25, −0.5). Left: Mick Jagger; right: John Lennon.

the brain which subserve face recognition. But evidence for such neural specialization does not show that faces are processed in some ways uniquely. We attempted an analysis of what it is about faces that makes their encoding so vulnerable to inversion. Support for the analysis would be provided by successful prediction of other classes of stimuli that are similarly affected by inversion.

Faces share a basic configuration in a way that can be made precise: each face can be defined in terms of a fixed set of points, such that the average of a set of faces, so defined, is still recognizable as a face. This is not true of a randomly chosen set of bridges, or houses, or buildings, or landscapes. Furthermore, some of the features by which we individuate faces are distinctive variations of that basic configuration. This is seen by the recognizability of line drawings such as those on figure 2: these line drawings are all specified in terms of the same 169 points (consider, for now, only those marked 'V', that is, only the veridical line drawings; these are produced by locating 169 points on photographs, and instructing the computer to connect the appropriate points, smoothing the curves). We dubbed such features 'second-order relational features,' and hypothesized that extracting second order relational features was particularly affected by inversion, and that the ability to encode individuals in terms of such features required considerable experience with faces, i.e. required acquisition of expertise.

Dogs' faces also share a basic configuration, yet encoding of dogs' faces is affected only minimally by inversion (Scapinello & Yarmey 1970). But in Scapinello & Yarmey's study, the encoders were not dog experts. This analysis predicts that dog experts, encoding individual dogs, should be affected by orientation just as all adults (face experts) are in encoding individual faces. Two studies tested this prediction. American Kennel Club judges inspected a series of individual dogs, shown in profile, and then picked out those individuals from distractors they had not seen before. Non-experts were tested in two conditions: at the same series size as experts, and at series half the size, so as to equate performance on the upright. Two series of inspection and associated recognition items were prepared, so that each subject could be tested on upright and on inverted dogs. In the first study, the dog breeds were poodles, setters, Scotties. Experts were more affected by inversion (12%) than were novices (8% at the large set size; 3% at the small set size), but not significantly so. However, the experts complained bitterly at the task we had given them: the three breeds were from different categories of dogs (sporting dogs, working dogs), and only 'best of show' judges know all breeds. So in the second study we used only sporting dogs – setters, retrievers, spaniels – and only sporting-dog experts. These experts were equally affected by inversion encoding dog profiles (22%) and human faces (20%). Novices showed the usual stimulus by orientation interaction, that is they were more affected by orientation at encoding faces (23%) than dogs, at either the large set size (2%) or the small set size (−2%; Diamond & Carey 1986).

It appears that such perceptual expertise requires about 10 years to develop, whether one is a child or an adult. It is at age 10 years that children perform in the normal adult range on face encoding tasks. And the period of apprenticeship for becoming an American Kennel Club judge is 10 years!

Faces are not special in the sense of posing unique problems for a pattern encoder, at least not as reflected in unique sensitivity to inversion. These data support the hypothesis that the inversion effect for faces reflects individualization within a class of patterns that share a configuration in terms of second-order relational features, and that reliance on distinguishing features of this type requires considerable expertise.

7. A RELATED EFFECT OF EXPERTISE

Consider again the faces in figure 2. A set of faces represented by a fixed set of points can be averaged, yielding an average face. One method is to normalize the faces in the set by aligning the pupils of the two eyes, and then simply average the values of each of the other points. Caricatures can then be created as follows: Find the difference between each point on the face to be caricatured with the corresponding point on the average face, and multiply that difference by a fixed amount (say 50%). This operation has the effect of exaggerating more those aspects of faces that differ more from the average face (50% of a big difference is greater than 50% of a smaller difference). One can also create an anticaricature, by decreasing the difference between the face to be caricatured and the average face. Figure 2 shows a set of anticaricatures and caricatures of Mick Jagger and John Lennon (see Brennan (1985) for a full characterization of the caricature generator).

Distorting faces away from the average face is a nonlinear transformation; it is quite unlike stretching a face along a vertical or horizontal axis. Each point is moved in a unique direction, and by a unique distance. None the less, the caricaturing transformation does not reduce the recognizability of the face; indeed, in some experiments, a 50% caricature is recognized faster than the veridical drawing and a slight caricature is judged the 'best likeness' of the person depicted (see Rhodes *et al.* 1987). This contrasts with the anticaricature transformation, which markedly reduces the recognizability of the face.

The line drawings of figure 2 are very degraded representations of faces; it takes much longer to recognize them than to recognize the photograph from which they are traced, and subjects make many errors. Clearly, those aspects of the configuration that define a face captured in such drawings form only a small part of the basis for face recognition. In a recent series of studies, Benson & Perrett (1991) created caricatures of full photographs by exaggerating just the same aspects of difference from the norm, leaving information about hair and skin colour and texture, eye colour, etc., normally provided by photographs. They too found that subjects judged the slight caricature the best representation of the face, and that caricatures were recognized slightly faster. Thus, even in the context of much fuller information about a face, manipulations of the shared configuration depicted on figure 2 affected recognition similarly to when this configurational information is all that subjects had to go by.

An actual caricature advantage is counterintuitive. Our memory representations serve recognition of actual people, their real faces; why should there be an advantage for recognition of caricatures? Because caricatures exaggerate what is distinctive about a face, there could be an advantage in caricaturing the memory representation itself. Alternatively, the caricature advantage could arise in the recognition process itself, even if memory representations were veridical, because exaggerating what is distinctive decreases the similarity of the target to other faces stored in memory (see Rhodes *et al.* (1987) for a discussion of these two possibilities, and Tanaka (1990) for a connectionist model in which a caricature advantage arises, in spite of veridical representations of the stimuli to be recognized).

The lesson I wish to draw from the Rhodes *et al.* results does not depend on the actual caricature advantage. Here I wish to emphasize that in spite of the considerable distortion, caricatures are as well or better recognized than actual faces, much better than anticaricatures which are equally distorted from the veridical. That caricatures are better recognized than anticaricatures is unsurprising, as one can see from figure 2: anticaricatures are collapsed toward the average face, so they all begin to resemble the average face, and are very similar to each other. To ensure that the better recognizability of caricatures and veridicals, compared with anticaricatures, is not due only to this fact, we created a set of 'lateral caricatures'. In a lateral caricature, each point is moved the same distance from its origin as on the caricature or anticaricature, but at a right angle from the vector defined by the relation between the origin and its corresponding point on the average face. The lateral caricatures differ, then, from the veridicals exactly as much as do the caricatures and the anticaricatures, and they are not collapsed toward the average face. Nonetheless, they are not recognizable. In a study using famous faces, we found that subjects need 1901 ms to recognize the caricatures, 2130 ms for the veridicals (note the actual caricature advantage), 3322 ms the anticaricatures, and 4377 the laterals (Carey *et al.* 1992). Whereas 50% of the caricatures and veridicals were recognized, only 12% of the laterals were.

In creating these stimuli, we manipulate only second-order relational features, that is, only aspects of faces dependent upon the points that determine the shared configuration. Only caricatures, those representations that exaggerate what is distinctive in the configuration, are as well or better recognized than veridicals. These data establish that second-order relational features play a role in face recognition. Holistically distorting a face away from the average face by manipulating only aspects of the shared configuration creates a psychologically relevant vector in the face space.

As we saw above, other classes of stimuli require discrimination within a shared configuration. Rhodes & McLean (1990) recently showed a caricature advantage for recognition of passerines (song birds) from profile line drawings, but only on the part of expert bird watchers!

I have presented two reflections of expertise relevant to discriminating among stimuli which share a configuration: a disproportionately large effect of inversion, and a caricature advantage. As we try to understand what expertise in face encoding embodies, we must explain both of these effects.

8. THE EFFECT OF ORIENTATION ON CONFIGURAL ENCODING OF FACES

Several sources of data support the importance of configural information in expert face encoding. The relations among features can be directly manipulated on schematic line drawings of faces, and on more realistic faces that can be assembled in photo-fit type systems (as used by police departments in creating likenesses from witness description). Adults are sensitive to changes in the spatial relations among features, as well as to changes in the features themselves (e.g. Sergent 1984; Haig 1984). Furthermore, there is abundant evidence that inversion disproportionately disrupts the processing of configural information of this sort. For example, in a timed task involving same–different judgements of schematic faces, mismatches due to differences in internal spacing of features were processed differently in the upright and inverted conditions, whereas mismatches due to changes of eyes, or changes in overall face contour were processed the same in the two orientations (Sergent 1984).

Other demonstrations underline the importance of orientation to configural processing of faces. Consider the famous Thatcher illusion (Thompson 1980). Thompson inverted the then British Prime Minister Margaret Thatcher's mouth and the eyes, keeping the rotated features in their normal place within the photograph of her face. The resulting photograph looks grotesque if upright, but does not look particularly unusual if inverted. The grotesque appearance in the upright may be due to violations of constraints on second-order relations among the points that define the shared configuration, constraints that are only defined relative to upright faces (e.g. eyes don't slant that way, relative to the nose and forehead). Others have suggested different interpretations of the difference in monstrosity between the upright and upside-down faces (e.g. Parks *et al.* 1985); most probably there are several distinct sources of the illusion.

Sergent (1984) documented a second sense in which faces are encoded configurally when upright, but not when inverted. The dimensions on which her schematic faces varied were processed interactively in the upright, but independently when inverted. Others have reported similar results. For example, Maruyama & Endo (1984) showed that perception of eye gaze in schematic faces is more affected by the orientation of the profile when faces are upright then when they are inverted. Faces are processed more holistically when upright. In the course of constructing an integrated representation, one independently manipulated feature influences the contribution other independently manipulated features make to the final representation.

Young *et al.* (1987) provided one of the most striking demonstrations that inversion interferes with configural encoding of faces. Here I describe just one version of the effect they found. Note that the top halves of photographs of famous people, or familiar colleagues, are easily recognizable. Young *et al.* made two types of displays: composites and non-composites. In composite photographs the top half of one face (say John F. Kennedy) is perfectly aligned with the bottom half of another (say Richard Nixon), creating a photograph of what seems to be a new person who resembles both Kennedy and Nixon. In non-composite photographs, the top half of one face is displayed above, but offset from, the bottom half of another, so the two do not fuse into a new face. The subject's task is simply to name the person whose face comprises the top part of the composite or non-composite. Young *et al.* found that reaction times and errors were markedly greater for composites than non-composites, but only when the faces were upright. Performance on inverted composites did not differ from that on inverted non-composites.

9. ARE THERE AT LEAST TWO DISTINCT SENSES OF 'CONFIGURAL ENCODING?'

All of the results cited in the previous section have been taken as support for the proposition that encoding configural aspects of faces is less efficient if the face is not in its canonical orientation. What is not clear is how many different things are meant by 'configural encoding.' Nor is it clear, if there are distinct senses, which are involved in which particular effects of orientation. For example, how should we think about the split face result? When the top half of one face is combined with the bottom half of another, unique second-order distinctive features emerge; therefore, that inverted composites are not recognized less well than inverted non-composites is consistent with the hypothesis that inversion interferes with the encoding of second-order relational features. A second possible interpretation of the split-face results derives from the idea that inversion affects holistic encoding. A face processor is getting conflicting information about the identity of the face: Nixon from the top; Kennedy from the bottom. When the face is upright, and processed holistically, the processor cannot ignore the information from the bottom. When the faces are offset, the information is not interpreted as coming from a single face, so no interference occurs. Since inversion interferes with holistic processing, no interference occurs from composite upside-down faces.

Holistic encoding is genuinely distinct from encoding second-order relational features. Holistic encoding underlies such phenomena as the word superiority effect, in which the discrimination of two letters (e.g. 'a' from 'u') is faster when the letters are in the context of full words (e.g. 'cat' versus 'cut') than when they are alone. Words, however, do not share a configuration, and therefore are not distinguished on the basis of second-order relational features.

As just argued, once at least two distinct sources of the large inversion effect for face encoding are admitted, the question of which is the source of the difference in processing upright and inverted faces in each case arises. Another question also presents itself. Which, if either, is involved in the expertise by orientation interaction? Does expertise result in increased reliance on configural encoding of faces in both senses?

One way to address this question is to assess the developmental history of the effects of orientation on the tasks that reflect configural encoding reviewed in § 8, above. As an example of this research strategy, I will sketch a recent example from Carey & Diamond (1992). In several studies, we assessed the developmental course of the effects Young *et al.* documented with composite and non-composite faces. The patterns of results from all the studies are the same; I will present just one here. We prepared composite (aligned) and non-composite (offset) split-face photographs of children in first grade classes (age 6–7 years), children in fifth grade classes (age 10–11 years) and adults in the Department of Brain and Cognitive Sciences at Massachusetts Institute of Technology. The study was run at the end of the year; all subjects were very familiar with the people photographed. We then repeated the Young *et al.* procedure; subjects named as quickly as possible the faces depicted in the top halves of composite and non-composite photographs. Each subject saw one series in the upright and one inverted.

The results were totally clear. At each of the three ages, the pattern of results described by Young *et al.* obtained. That is, in the upright, the people in composite photographs were named much more slowly, with more errors, than those in non-composite photographs, whereas composites did not differ from non-composites if faces were inverted. There is no developmental increase, at least from age 6 years on, in the interference of the bottom half of a composite face to the naming of the top half, and at all ages this interference obtains only when faces are upright. However, there was a second, equally clear result: an overall age by orientation interaction. That is, 6-year-olds were overall slower on upright faces (averaged over composites and non-composites) than on inverted faces; 10-year-olds processed the upright and inverted faces approximately equally quickly, whereas adults were overall much faster on upright than inverted faces.

This pattern of results supports several conclusions. As far as this reflection of configural encoding of faces is concerned, there are no developmental changes over the ages of 6 years to adulthood. Therefore, this reflection of configural encoding does not underly the greater sensitivity to inversion that adults (experts) experience. This is shown by the fact that 6-year-olds experience the interference in upright composites, and by the fact that the normal age by orientation interaction is seen in this study, independent of the composite interference effect. Finally, these data confirm that there are at least two distinct, independent, sources of the effect of orientation on face encoding. Of course, exactly how the age (expertise) by orientation interaction is to be understood is left entirely open by these data.

11. A FEW CONCLUDING REMARKS

Locating the source of developmental change is not easy; nor is characterizing the differences between novices and experts at some tasks. These are the challenges for those attempting to understand the protracted course of acquisition of skill at face recognition.

The work reported here was supported by a National Institutes of Health grant, 'Face recognition: developmental questions' to Dr Rhea Diamond and myself. All of my research on face recognition has been in collaboration with Dr Rhea Diamond, whose contribution to the ideas developed here cannot be distinguished, in my mind, from my own. Recently, Dr Diamond and I have begun collaborating with Dr Gillian Rhodes, whose ideas also have had a large impact on my current views.

REFERENCES

Bahrick, H.P., Bahrick, O.O. & Wittlinger, R.P. 1975 Fifty years of memory for names and faces: a cross-sectional approach. *J. exp. Psychol.: Gen.* **104**, 54–75.

Benson, P.J. & Perrett, D.I. 1991 Perception and recognition of photographic quality facial caricatures: implications for the recognition of natural images. *Eur. J. cogn. Psychol.* **3**(1), 105–135.

Benton, A.L. & van Allen, M.W. 1973 Test of facial recognition. *Neurosensory Center Publication No. 287.* Iowa City, Iowa: University Hospitals, Department of Neurology.

Blaney, R.L. & Winograd, E. 1978 Developmental differences in children's recognition memory for faces. *Devl Psychol.* **14**, 441–442.

Brennan, S.E. 1985 The caricature generator. *Leonardo* **18**, 170–178.

Bushnell, I.W.R., Sai, F. & Mullin, J.T. 1989 Neonatal recognition of the mothers face. *Br. J. devl Psychol.* **7**, 3–15.

Carey, S. 1981 The development of face perception. In *Perceiving and remembering faces* (ed. G. Davies, H. Ellis & J. Shephard), pp. 9–38. New York: Academic Press.

Carey, S. & Diamond, R. 1992 Do young children rely less on configural information when processing faces than do adults? (In preparation.)

Carey, S., Diamond, R. & Woods, B. 1980 The development of face recognition – a maturational component? *Devl Psychol.* **16**(4), 257–269.

Carey, S., Rhodes, G., Diamond, R. & Hamilton, J. 1992 Comparing the recognizability of caricatures, anticaricatures and lateral caricatures. (In preparation.)

Diamond, R. & Carey, S. 1977 Developmental changes in the representation of faces. *J. exp. Child Psychol.* **23**, 1–22.

Diamond, R. & Carey, S. 1986 Why faces are and are not special: an effect of expertise. *J. exp. Psychol.: Gen.* **115**(2), 107–117.

Diamond, R. & Carey, S. 1990 On the acquisition of pattern encoding skills. *Cogn. Develop.* **5**, 345–368.

Diamond, R., Carey, S. & Back, K. 1983 Genetic influences on the development of spatial skills during early adolescence. *Cognition* **13**, 167–185.

Ellis, H.D. 1990 Developmental trends in face recognition. *The Psychologist* March issue, 114–119.

Fagan, J.F. 1979 The origins of facial pattern recognition. In *Psychological development from infancy: image to intention* (ed. M. H. Bornstein & W. Kessen), pp. 83–113. Erlbaum.

Flin, R. 1980 Age effects in children's memory for unfamiliar faces. *Devl Psychol.* **16**, 373–374.

Flin, R.H. 1983 The development of face recognition. Ph.D. thesis, Aberdeen University.

Goren, C.C., Sarty, M. & Wu, P.Y.K. 1975 Visual following and pattern discrimination of face-like stimuli by new born infants. *Pediatrics* **56**, 544–549.

Haig, N.D. 1984 The effect of feature displacement on face recognition. *Perception* **13**, 505–512.

Johnson, M., Dziurawiec, S., Ellis, H. & Morton, J. 1992 New borns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition.* (In the press.)

Maruyama, K. & Endo, M. 1984 Illusory face dislocation effect and configurational integration in the inverted face. *Tohoku Psychol. Folia* **43**(1–4), 150–160.

Parks, T.E., Coss, R.G. & Coss, C.S. 1985 Thatcher and the Cheshire cat: context and the processing of facial features. *Perception* **14**, 747–754.

Phillips, R.J. & Rawles, R.E. 1979 Recognition of upright and inverted faces: a correlational study. *Perception* **8**, 577–583.

Posner, M.I. & Keele, S.W. 1968 On the genesis of abstract ideas. *J. exp. Psychol.* **77**, 353–363.

Rhodes, G., Brennan, S. & Carey, S. 1987 Identification and ratings of caricatures: implications for mental representations of faces. *Cogn. Psychol.* **19**(4), 473–479.

Rhodes, G. & McLean, I.G. 1990 Distinctiveness and expertise effects with homogeneous stimuli: towards a model of configural coding. *Perception* **19**, 773–794.

Saltz, E. & Sigel, I.E. 1967 Concept overdiscrimination in children. *J. exp. Psychol.* **73**, 1–8.

Sergent, J. 1984 An investigation into component and configural processes underlying face perception. *Br. J. Psychol.* **75**, 221–242.

Tanaka, J.W. 1990 Caricature recognition in a neural network. *Proceedings of XII Annual Conference of the Cognitive Science Society*, 622–628.

Thompson, P. 1980 Margaret Thatcher: a new illusion. *Perception* **9**, 483–484.

Walton, G.E. & Bower, T.G.R. 1991 Newborn preference for familiar faces. Paper presented at SRCD, Seattle.

Yin, R.K. 1969 Looking at upside-down faces. *J. exp. Psychol.* **81**, 141–145.

Yin, R.K. 1970*a* *Face recognition: a special process?* Ph.D. thesis, Massachusetts Institute of Technology.

Yin, R.K. 1970*b* Face recognition by brain-injured patients: A dissociable ability? *Neuropsychologia* **8**, 395–402.

Young, A.W., Hellawell, D. & Hay, D.C. 1987 Configural information in face perception. *Perception* **16**, 747–759.

Discussion

S. de Schonen (*Cognitive Neuroscience Unit, L.N.F.1, C.N.R.S., Marseille, France*). I am not sure of the meaning of the changes with age Professor Carey found in children. In 4–9-month-old infants we have shown first that the right hemisphere has an advantage over the left in familiar face recognition (from photographs), and second, that in this age range, the right hemisphere has an advantage over the left hemisphere at discriminating and recognizing two faces that differ in the size of the eyes only or in the orientation of the eyes only (the two faces are in fact the same face, but in one either the original eyes have been enlarged or their orientation has been changed without changing their shape itself). However, the right hemisphere is unable to discriminate and recognize the two faces when the shape of the eyes has been changed without changing the overall size and orientation (the eyes of the original face are replaced by the eyes of a different face). On the contrary the left hemisphere has an advantage in this latter task but a disadvantage in the first task. The right hemisphere sensitivity to the relative position of the components of a pattern and the left hemisphere sensitivity to the shape of the components is found not only in face recognition but also in geometrical pattern recognition. This right hemisphere mode of processing faces and patterns is the main ingredient for a prototype building process. Now, we also showed that during this same age period, what is learned from a face by one hemisphere is not transmitted to the other hemisphere. Therefore, one of the changes with age might be related to the occurrence of a coordination between the two modes of processing faces and patterns. Other changes might be related to how hemispheric priority is organized. I am not convinced that the crucial change occurring in the age period you have been describing consists in the emergence of prototype building skills which are probably already in function. It might rather be related to a reorganization of time sharing between the hemispheres or to the emergence of a new system of hierarchical priority of one over the other, or to the emergence of a composition process between the modes of processing by the two hemispheres.

S. Carey. I find the data Dr de Schonen reports fascinating, and see them as consistent, in broad outline, with Dr Sergent's work on left and right hemisphere processes involved in face recognition. Of course, especially important is her finding of this pattern of results with infants!

Dr de Schonen misunderstood my argument. I presented data that there is no developmental change in spatial prototype building skills over the ages 6–10 years, agreeing with Dr de Shonen's conclusion that development of these skills could not underly the emergence of face recognition skills during these ages. It is certainly possible that the changes over this age range involve a new system of hierarchical priority of one hemisphere's processing over the other's, or some other change involving hemispheric specialization. It would be fruitful to try to bring evidence to bear on such hypotheses. But even if such a hypothesis were to be supported, we would be left with one of the developmental questions I addressed in my paper: is this a 'developmental' change, under maturational control? The data I offered comparing adult dog novices with dog experts, and that Rhodes & McLean (1990) offer comparing adult bird novices and adult bird experts, suggest that some of the changes underlying the developmental course of face encoding are not developmental in nature.

H. D. Ellis (*School of Psychology, University of Wales College of Cardiff, U.K.*). I was struck by the fact that in Professor Carey's data concerning split faces, children were equally good at upright and inverted stimuli. I wonder whether this is the result of using restricted sets of a small number of target faces. Would the same results occur with an infinite set size? In other words, could the apparent facility with inverted stimuli be attributed to a strategy of spotting identifying features which could only work for inverted stimuli when the target set is very small?

S. Carey. Six-year-old children were actually better at inverted than at upright faces in the split-face study; 10-year-olds performed equally well at both orientations. I am not arguing that children are not affected by orientation; we know they are. I offer these data simply as another case of the age by orientation interaction: adults are more affected by inversion.

The explanation Professor Ellis offers – that young children adopt a strategy of spotting identifying features that can be encoded from inverted stimuli but that serve to discriminate only small set sizes – is exactly the same as the explanation Dr Diamond and I favour. With an infinite set size, of course, adults would be at chance on both upright and inverted faces, as there would be another face arbitrarily close to the target on whatever dimensions underly face discrimination. But I agree with the thrust of Professor Ellis' comment: the question is why children adopt this strategy.

The period from age 6 to 10 years witnesses the flowering of metamemorial strategies: 10-year-olds are much better than 6-year-olds, and adults still better than them, at figuring out the demands of a task and adopting a relevant strategy. The set sizes were equally small for the 10-year-olds and adults. Adults and 10-year-olds also adopted strategies of relying on features adequate to discriminating this small set of faces: reaction times got much faster with age and errors decreased. But the features adults relied on could not be encoded as well from inverted faces whereas those children relied on could. Left unspecified by Professor Ellis' suggestion is what the difference is between features that work for inverted stimuli and those that don't. These are issues I touch upon in the paper.