VISUAL COGNITION, 1994, *1* (2/3), 253–274

Are Faces Perceived as Configurations More By Adults than by Children?

Susan Carey and Rhea Diamond

Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

Adult face recognition is severely hampered by stimulus inversion. Several investigators have attributed this vulnerability to the effect of orientation on encoding relational aspects of faces. Previous work has also demonstrated that children are less sensitive to orientation of faces than are adults. This has been interpreted as reflecting an increasing reliance on configural aspects of faces with increasing age and expertise.

Young, Hellawell, and Hay (1987) demonstrated that for adults the encoding of relations among facial parts is, indeed, sensitive to orientation. When chimeric faces are upright, the top half of one face fuses with the bottom half of the other, making the person depicted in the top half difficult to recognize. This effect (the composite effect) is not seen when the faces are inverted. The present study obtained the composite effect for 6-year-old and 10-year-old children, just as for adults. The composite effect was found to an equal degree at all ages tested and was seen both in tasks involving highly familiar faces and in those involving newly learned, previously unfamiliar faces. Thus, these data provided no support for the hypothesis of increasing reliance on configural aspects of faces with increasing age, at least in the sense tapped by this procedure.

However, the data did confirm an Age × Orientation interaction. In recognizing both familiar and previously unfamiliar faces, 6-year-olds were less affected by inversion than were 10-year-olds, who, in turn, were less affected than were adults. Increasing vulnerability to inversion of faces with age was independent of the composite effect. Apparently, there are two distinct sources to the large effect of inversion that characterizes adult face encoding: one seen throughout development and one acquired only with expertise.

Requests for reprints should be sent to S. Carey, Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

This work was supported by grant R01-HD221666 from the National Institute of Health. We thank J. Hamilton and K. Nigam for technical assistance in these studies.

In spite of the impressive capacity for face encoding that infants display (Fagan, 1979; Johnson & Morton, 1991), young children are dramatically worse than adults at encoding and subsequently recognizing unfamiliar faces. Marked improvement between ages 2 and 10 is observed on simple recognition tasks in which a set of facial photographs is inspected, later to be picked out from distractors (Carey, 1981; Carey & Diamond, 1977; Carey, Diamond, & Woods, 1980; Flin, 1980.) Dramatic improvement with age is also apparent on tasks that place no demands on memory. For example, Benton and Van Allen required subjects to match a target photograph with photographs of the same person taken under different lighting or from a different point of view or with different facial expressions. On this simultaneous matching task, 6-year-olds performed at a level associated with right-hemisphere damage in adults (Benton & Van Allen, 1973; Carey et al., 1980). Other data also show that 6-year-olds are severely limited at recognizing whether or not two photographs (simultaneously presented) depict the same person if the photos differ in angle of view, expression, or clothing worn (Diamond & Carey, 1977; Ellis, 1992; Flin, 1980; Saltz & Sigel, 1967). Such data indicate that the child's problems are in the initial encoding of new faces, not just in memory or retrieval. Young children differ from adults, then, in the ability to encode a new face in terms of distinguishing features that ensure it is recognized and differentiated from other faces.

One indication that children are doing something *different* from adults, rather than just less of what adults do, is the evidence that children are less affected by the orientation of the face during encoding and recognition than are adults.

For adults, encoding individual faces is more affected by inversion than is encoding of individual members of almost any other class studied to date: houses, bridges, stick figures of men, buildings, landscapes, dog faces (Diamond & Carey, 1986; Scapinello & Yarmey, 1970; Yin, 1969, 1970). In these studies the stimuli are usually presented in the same orientation during both inspection and recognition; inverted stimuli are first seen upside down and are also presented for recognition upside down. Thus, the difficulty is in forming an adequate representation of an inverted face, not in coping with a mismatch in orientation between inspection and recognition. Typically, for faces there is a 20–30% decrement in recognition accuracy associated with the inverted condition, whereas there is only a 0–10% inversion decrement for stimuli from the other classes.

Two results emerge from studies of the developmental history of the inversion effect on face encoding. (1) As long as ceiling and floor effects are controlled for, face encoding is affected by orientation at every age, even in infancy (Carey, 1981; Fagan, 1979; 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. (2) There is often an Age × Orientation interaction, which, of course, is also sensitive to floor and ceiling effects. That is, the magnitude of the inversion effect increases with age

(Carey, 1981; Flin, 1983; Goldstein & Chance, 1964). Thus, if we understood the large inversion effect on face encoding, we might begin to understand what changes with development.

One approach to this problem is to probe more deeply the circumstances under which the large inversion effect obtains. As reviewed above, early studies showed face recognition to be more affected by inversion than was recognition of any other class of stimuli. However, Diamond and Carey (1986) showed that the large inversion effect is observed in at least one additional case: sporting dog experts are affected by inversion in encoding whole body profiles of such dogs as much as they (and other normal adults) are in encoding faces. Dog novices, in contrast, show the usual Stimulus Class × Orientation interaction, being affected by inversion in encoding faces far more than in encoding dog profiles. Thus, the Age × Orientation interaction for face encoding can be taken as an Expertise × Orientation interaction. 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 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!

What do expert face encoding and expert sporting dog encoding have in common? Faces share a 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, houses, buildings, or landscapes. It is true, however, of a randomly chosen set of sporting-dog profiles. 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 produced by connecting the same set of points defined on each face (see Rhodes, Brennan, & Carey, 1987). Diamond and Carey dubbed features that are distinctive variations of a shared configuration "second-order relational features" and hypothesized that the ability to encode individuals in terms of such features requires expertise and that it is particularly affected by inversion.

Diamond and Carey (1986) offered no direct evidence for this hypothesis, supporting it only with the finding that it accounts for the large inversion effect in both adult face encoding and expert dog encoding. Rhodes, Brake, and Atkinson (1993) directly manipulated the features by which faces differed (e.g. by adding a moustache or eye glasses, by changing a nose or eyes, by varying the internal spacing of parts) and explored whether detection of changes in these different types of features was differentially affected by inversion. They found, as hypothesized, that changes in internal spacing were among the transformations most affected by orientation, but so, too, were changes in single features (e.g. just the eyes). As Rhodes et al. pointed out, this latter finding is not inconsistent with the hypothesis that inversion affects encoding of second-order relational features, for when the eyes within a face are changed, so are the relations between points on the eyes and other features of the face.

There is ample additional evidence that adults use the spatial relations among the parts of a face as a basis for individuating and recognizing faces. For example, in a series of studies, Haig showed that manipulating the vertical spacing of parts of faces (e.g. raising the nose relative to the rest of the face) affected recognition rates (e.g. Haig, 1984).

There is also ample additional evidence that the encoding of the spatial relations among the parts of a face is sensitive to inversion. For instance, if the eyes and mouth are inverted within the face, the result is monstrous, but the grotesque appearance of such faces is perceived only if the face as a whole is upright (Thompson, 1980). In a related finding, Bartlett and Searcy (1993) showed that faces could be made to look grotesque by manipulating the spatial relations among their parts. Subjects judged such transformed faces less grotesque when inverted than when upright. Apparently, if the face is upside down, subjects cannot detect the subtle deviations from the normal configuration introduced by such transformations.

Tanaka and Farah (1993) and Farah, Tanaka, and Drain (in press) have offered an alternative analysis of why face recognition is so sensitive to inversion and, thus, what developmental changes in face representations underlie the orientation effect characteristic of adults. They distinguish between holistic and parts-based representations, arguing that faces are more likely to be encoded holistically than are other classes of stimuli. They further suggest that holistic encoding is more sensitive to inversion than is parts-based encoding.

In their papers, Tanaka and Farah (1993; Farah et al., in press) provide two distinct characterizations of holistic encoding. First, holistic representations are those in which the parts of the stimulus are not explicitly represented. In the case of faces, this would mean that a particular face would not be represented in terms of the identities of parts—such as the nose, eyebrows, mouth, for example—but, rather, in terms of a template-like representation of the whole. In such a representation, individual parts (e.g. Bob's nose) should be more difficult to recognize in isolation than in the context of the whole face, and, indeed, this is what Tanaka and Farah (1993) found. Evidence that holistic representations are sensitive to orientation was provided by the finding that the advantage for recognizing Bob's nose in the whole face disappeared when the stimuli were turned upside down.

This first characterization of holistic encoding raises the question of what is meant by "explicitly represented". One interpretation is that the parts are less accessible to analysis and report than is the whole, much in the same sense in which the syllable is more accessible than the phoneme in phoneme/syllable monitoring tasks (Fodor, Bever, & Garrett, 1975). Here "explicitly represented" means consciously accessible, or subject to attentional monitoring. A second interpretation is that the parts of a face, such as eyes, noses, and mouths, are not the atoms of face representations; that the face template is not built from these parts. These two interpretations are certainly different, for although phonemes are less accessible than syllables, nobody would deny that phonemes are the atoms of syllabic representations. The Tanaka and Farah demonstration actually supports the accessibility interpretation; subjects can, after all, recognize Bob's nose in isolation—but they do so more slowly than when it is in the context of the whole face.

In their second characterization of holistic encoding, Farah et al. (in press) state that in holistic representations, the spatial relations among the parts are more important in specifying an individual object than are the representations of the individual parts themselves. Note that this characterization is very different from the first, for in this second characterization there is no claim that the individual parts are not explicitly represented, in the sense of being the atoms of facial representations. Indeed, as Farah et al. admit, under this characterization, the distinction between holistic representations and configural representations becomes blurred to the point of disappearing.

In the remainder of this paper, we use "holistic encoding" when referring to the accessibility hypothesis—that representations of whole faces are more easily and quickly accessed than are representations of parts, and use "configural encoding" when referring to the hypothesis that the spatial relations *among* parts are especially important in face encoding.

Young et al. (1987) have provided a striking demonstration of one sense in which upright faces are encoded configurally, whereas inverted faces are not. They created composites of the top half of one person's face and the bottom half of another's (see Figure 1). In a typical experiment, the stimuli were photographs of famous faces, and the subject's task was to name the person depicted in the

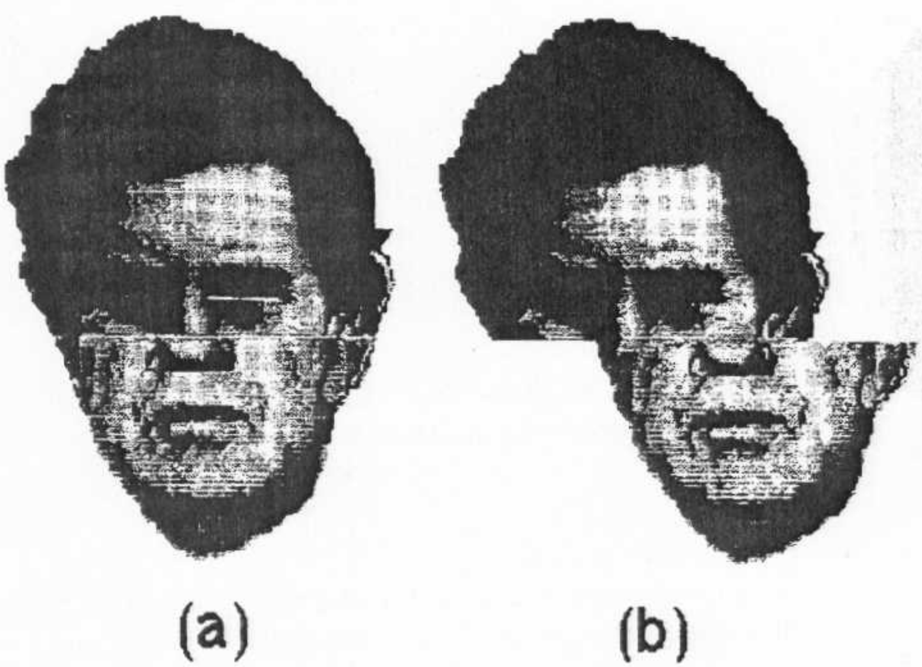

FIG. 1. Sample stimuli. 1a. Composite Nixon/Prince Charles. 1b. Non-composite Nixon/Prince Charles.

top half, ignoring the bottom half. To control for possible interference from recognition of the bottom half, Young et al. also created non-composite faces (see Figure 1). They found that the reaction time (RT) to recognize the top half was slower for upright composites than for upright non-composites. Apparently, the two halves of the composites fused into a new face, making it much more difficult to identify the old top half, whereas the bottom half of the non-compesites could be ignored. Crucially, there was no difference between the composites and non-composites when the faces were inverted. Integration of the top and bottom halves so that it is difficult to disentangle them occurs only when the faces are upright. Young et al. also found this pattern of results when adults encoded unfamiliar faces—evidence that both unfamiliar and familiar faces are encoded configurally.

The Young et al. demonstration admits of interpretation in terms of both the hypothesis that inversion interferes with holistic encoding and that inversion interferes with configural encoding. With respect to the configural encoding hypothesis, note that many relational features are changed when the top half of one face is fused with the bottom half of the other. And as regards the holistic encoded hypothesis, note that the task requires accessing a part of a face (the top half) from within a whole face. The developmental course of the composite effect, then, will bear on evaluating the hypothesis that the Age × Orientation interaction in face encoding reflects increasing reliance on relational features and/or increasing reliance on holistic representations of faces.

Whereas adults encode both familiar and unfamiliar faces configurally in the sense tapped by Young et al., in other ways familiar and unfamiliar faces are encoding differently. For example, adults base recognition of familiar faces more on their inner features (eyes, nose, mouth, cheeks—those that reflect the bone structure of the face) whereas recognition of unfamiliar faces is based more on the outer features of the face (hair and overall face shape; Ellis, Shepherd, & Davies, 1979). Also, a more robust right-hemisphere advantage in recognition is found for familiar than for unfamiliar faces (Levine, 1981). Furthermore, there is some evidence that children encode familiar faces differently from unfamiliar faces, in the sense of relying on configural features of familiar faces but on relatively piecemeal features of unfamiliar faces. Levine (1981) found a right-hemisphere advantage at age 8 for recognition of familiar faces, but not for unfamiliar ones, and Diamond and Carey (1977) found that piecemeal, misleading cues (hats, eyeglasses), were *ignored* by 6-year-olds when the faces depicted were familiar, whereas recognition was *based* on these cues when the faces depicted were unfamiliar. Finally, although most experiments to date reveal an Age × Orientation interaction in the case of encoding of unfamiliar faces, there is conflicting evidence concerning the effect of orientation on recognition of familiar faces. In one study, Goldstein (1975) found that 6- to 10-year-olds were little affected by inversion of photographs of highly familiar peers (12% errors), and that the sensitivity to orientation increased markedly through the age range of 13–14 (19% errors), 17–18 (24% errors), and 19–20 (38% errors). However, in an earlier study, Brooks and Goldstein (1963) found that 6-year-olds made many errors (33%) in identifying inverted faces they could recognize upright, whereas error rates fell to 0 by age 10.

The present studies exploit the Young et al. paradigm to ask three questions: (1) In the sense tapped by this procedure, are there developmental changes in configural or holistic encoding of *familiar* faces? If children are less sensitive to the configuration of the face than are adults, they should show less difference between composites and non-composites in the upright condition. At all ages, we would expect no difference between composites and non-composites in inverted faces. Therefore, the development of sensitivity to facial configuration would be revealed by a three-way interaction between age, stimulus type (composite vs. non-composite), and orientation. (2) Are there developmental increases in configural encoding of *unfamiliar* faces? (3) What is the effect of orientation on children's encoding of familiar faces—is there or is there not an Age × Orientation interaction in the encoding of familiar faces? The answers to these questions will permit us to consider the relations between two phenomena: integration of the parts of an upright face into a unit (the Young et al. demonstration) and the developmental increase in the recognition advantage afforded the upright face.

EXPERIMENT 1

Configural Encoding of Familiar Faces

Method

Subjects.

Subjects were 20 first-graders (mean age 7;1), 20 fifth-graders (mean age 10;9) and 12 adults (mean age 28;3). At each grade, the 20 children were drawn from two separate classes (10 from each class). The adults were graduate students and research assistants in our department.

Stimuli.

Adults: Six males and six females from among the graduate students and staff of MIT's Department of Brain and Cognitive Science were photographed. The photographs were scanned into MacPaint files. Two stimulus sets were made by pairing the top half of each photograph with the bottom half of the photographs of all five other people of the same sex. Each pairing was made in two versions: Composite (with the two halves aligned) and non-composite (the two halves offset.) See Figure 1 for examples made from famous faces (Prince

Charles and Richard Nixon). Thus, there were 30 composite faces and 30 non-composite faces in each set. The 60 stimuli were randomized, subject to the constraint that runs of no more than 3 stimuli of the same type (composite or non-composite) were allowed, and runs of no more than 2 stimuli with the same person in the top half were allowed. A single random order of stimuli was used for the male and female sets.

Children: Six boys and six girls from each class were photographed, and stimulus sets for each class were constructed as for the adults. Thus, there were 2 male sets for each age and 2 female sets for each age, as there were two classes at each age.

Procedure

All children who served as subjects were photographed, and each was given a polaroid photograph of themselves, plus two printouts—one of their own face and one a composite with half of their own face and half of another child's face. These printouts served to motivate the children to participate in the study and to ensure they understood the construction of the stimuli. Subjects were tested on the two stimulus sets made from children in their own classes.

The task was run on a Macintosh Computer. Stimuli were presented until the subject responded with the name. A voice-key terminated the trial, and the experimenter recorded whether or not the response was correct. Each subject participated in two versions of the experiment—one upright and one inverted. Whether the male face set or the female face set was seen first, and whether upright or inverted faces were seen first was counterbalanced within each age group.

Subjects were instructed to respond loudly, as quickly as possible, avoiding errors. They were first presented with the whole faces (three runs through the set), which they were to name as quickly as possible. This gave them practice with the voice key. They were then shown the top halves of the faces upright and asked to name them as quickly as possible (one time each). Subjects at all ages found it very easy to name the top halves alone (see also Goldstein & Mackenburg, 1966; Chance, Goldstein, & Schict, 1967). The task proper was then explained to them. They were told to name the top half of the face (the part containing the eyes and forehead—the part they had just practised on) and to ignore the bottom halves. If they were in an inverted condition, they were first shown the relevant half faces alone upside down for one additional practice trial.

Results

Errors.

Error rates were around 10% at each age: First graders 8%; fifth graders, 8%; adults, 10%. Separate analyses of variance (ANOVA) on error rates at each age revealed that at all ages there was a main effect for stimulus type (more errors on composites than on non-composites) and that at no age was there a Stimulus Type × Orientation interaction. Fifth graders made more errors on inverted than on upright faces; this was the only main effect of orientation. Errors will not be discussed further.

Reaction Times.

Figure 2 shows the RT data for correct responses. An ANOVA was carried out, with age (6, 10, adult) as a between-subject variable and orientation (upright, inverted) and condition (composite, non-composite) as within-subject variables. There was a significant main effect for age, $F(2, 49) = 14.47$, $p < 0.001$. Six-year-olds were slower (1339 msec) than 10-year-olds (1023 msec), who, in turn, were slower than adults (980 msec). There was also a significant main effect for condition, $F(1, 49) = 37.71$, $p < 0.001$. RTs to composites were slower (1193 msec) than those to non-composites (1075 msec). Also, the Orientation × Condition interaction that constitutes Young et al.'s effect was significant, $F(1, 49) = 31.15$, $p < 0.001$. That is, RTs for upright composites (1222 msec) were slower than for upright non-composites (1024 msec), whereas RTs for inverted composites (1165 msec) did not differ from those for inverted non-composites (1126 msec). The two effects of theoretical significance for the issues addressed in this paper are: (1) there is no three-way Age × Condition × Orientation interaction—that is, the Young et al. effect is seen equally at each

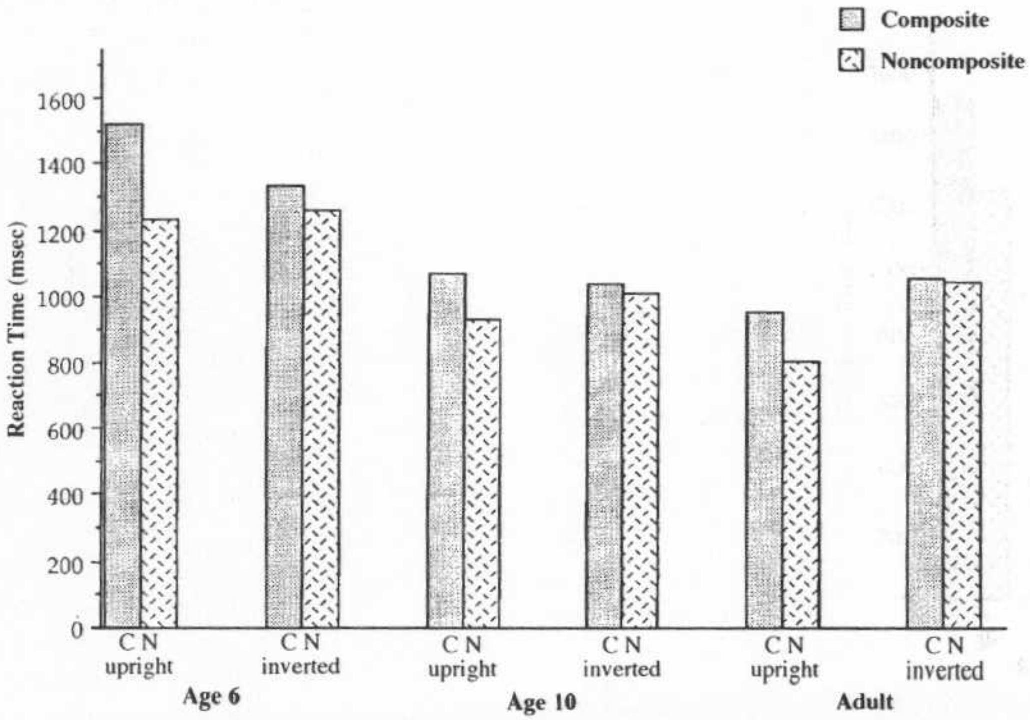

FIG. 2. Experiment 1: RT for correct responses

age (see Figure 2). (2) The Age × Orientation interaction was significant, $F(2, 49) = 4.56$, $p < 0.025$. Let us examine these two effects in turn.

The major question asked here is whether children, like adults, would be more influenced by the mismatch between the top half and the bottom half of the face in upright composites than in upright non-composites, and would show no effect of stimulus type when the stimuli were inverted. The absence of a three-way-interaction indicates that there is no developmental change in these effects. As a check, we analysed each age group's data separately. The Orientation × Stimulus Type interaction was significant at the 0.05 level at every age, except for adults, for whom it approached that level: first-graders, $F(1, 18) = 25.1$, $p < 0.001$; fifth-graders, $F(1, 18) = 9.94$, $p < 0.006$; adults, $F(1, 11) = 4.42$, $p < 0.06$. Apparently, children fuse the halves of two different familiar upright faces into a new face, just as adults do. And, just as adults, children can ignore the interference from the bottom half of the face when the stimuli are inverted.

First-Graders. The subjects from each class showed the same pattern of results; there was no main effect of class and no interaction of class with any other variable. Neither was there any effect of sex of stimuli, nor of condition order (upright first or inverted first). The only significant effects were for stimulus type, $F(1, 18) = 25.66$, $p < 0.001$, composites RT = 1428 msec, non-composites RT = 1250 msec, and the interaction depicted in Figure 2. Notably, there was no main effect of orientation; the reaction time for upright faces was 1379 msec, and that for inverted faces was 1298 msec.

Fifth-Graders. As for the first-graders, there were no effects involving class, sex of stimuli, or condition order. The only significant effects were for stimulus type, $F(1, 18) = 14.48$, $p < 0.001$, composites RT = 1064 msec, non-composites RT = 982 msec, and the interaction depicted in Figure 2. Again, there was no main effect of orientation; the RT for upright faces was 1006 msec, and that for inverted faces was 1091 msec.

Adults. As for children, there were no effects of condition order or of sex of stimuli. Adult RTs yielded a main effect for orientation: $F(1, 11) = 7.6$, $p < 0.02$; RTs were faster for upright faces (892 msec) than for inverted faces (1067 msec). Also seen was a main effect for stimulus type, $F, 1, 11) = 5.83$, $p < 0.04$; RTs for composites were slower (1018 msec) than RTs for non-composites (941 msec). The interaction depicted in Figure 2 just missed significance, presumably due to the smaller sample of adults.

Age Changes in the Effect of Orientation.

As can be seen in Figure 2, inversion slows performance increasingly with age. Figure 3 shows the difference in RT between inverted and upright stimuli. For both composites and non-composites, this difference becomes greater with

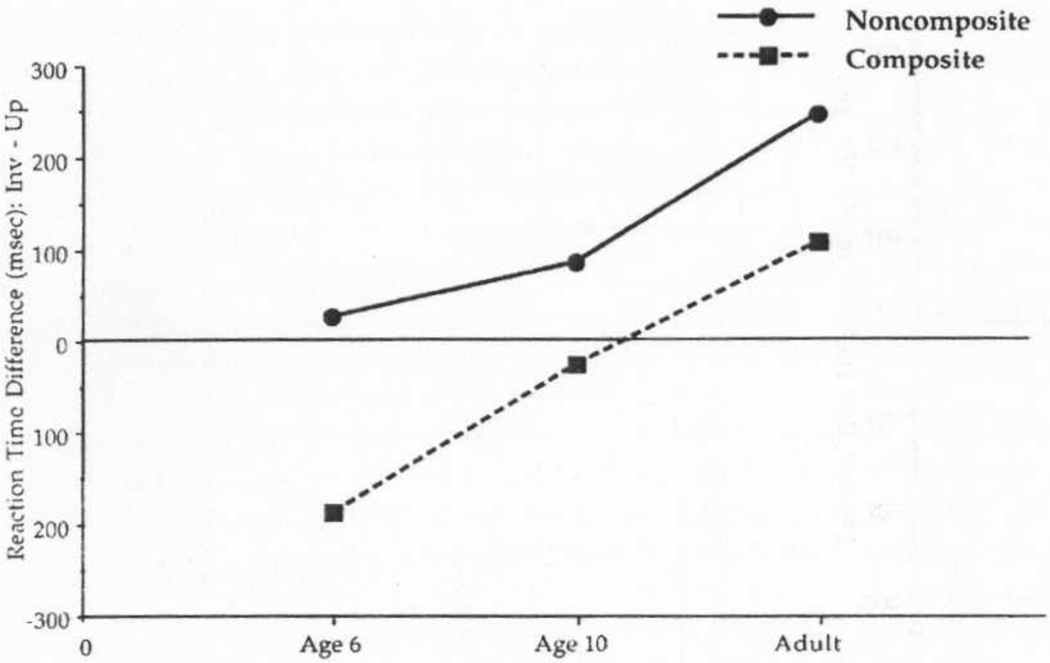

FIG. 3. Experiment 1: RT for inverted faces minus RT for upright faces.

increasing age, as confirmed by the Age × Orientation interaction reported above. Also evident from Figure 3 is the absence of a three-way Age × Orientation × Condition interaction. That is, the increased sensitivity to orientation with age was equally evident for composite and non-composite stimuli.

Averaged over composites and non-composites, first-graders are faster in the inverted condition, whereas adults are faster in the upright condition. However, all subjects are slowed by the mismatch of the bottom half of the composite face in the upright condition but not in the inverted condition, so RTs to composites do not give a pure indication of the effect of inversion. RTs to non-composites come closer to this, as in both orientations subjects are able to ignore the conflicting information and simply name the face whose top half is displayed. And as can be seen in Figure 3, 6-year-olds were equally fast on inverted (1263 msec) and upright (1237) non-composites, 10-year-olds were slower on inverted (1020 msec) than on upright (935 msec) non-composites, whereas adults were much slower on inverted (1063 msec) than on upright non-composites (818 msec).

Conclusions

The pattern of results reported by Young et al. is remarkably robust. Young et al.'s familiar faces were famous males; Experiment 1 extends their results to faces of familiar personal acquaintances of both sexes. In addition, the first- and fifth-grade groups each provided a complete replication of the pattern of data on two independent groups of subjects viewing different stimulus sets.

The data from Experiment 1 support two conclusions: (1) For familiar faces, there is no increase with age in configural or holistic encoding, at least in the sense tapped by the Young et al. procedure. That is, age had no effect on either the RT advantage for non-composites over composites, or on the restriction of this advantage to the upright condition. (2) The question of whether young children are as affected by the orientation of familiar faces as are adults receives an unequivocal answer: no. Older subjects were much more disrupted by stimulus inversion than were younger subjects. Indeed, first-graders were equally fast on the non-composite faces, whether they were right-side up or upside down. Adults, in contrast, were much slower when the faces were inverted, in spite of the fact that they had to differentiate only 6 highly familiar faces, each depicted in a single photograph seen many times in the experiment.

In sum, the important findings of Experiment 1 are that the Age × Orientation interaction typical of encoding unfamiliar faces is also seen here in the recognition of familiar faces, and that this effect is statistically independent of the orientation by composite effect that reveals configural or holistic encoding of faces.

As previously noted, there is reason to believe that children differ less from adults in their representations of highly familiar faces than in their capacity to encode unfamiliar faces (Diamond & Carey, 1977; see Carey, 1981, for a review). Thus, although there appear to be no changes between age 6 and adulthood in the degree to which familiar faces are encoded configurally or holistically, perhaps young children are less able to encode newly encountered faces in such a manner. In Experiment 2, we turn to the question of whether for unfamiliar faces there is evidence for increasing reliance on configural encoding with age.

EXPERIMENT 2

Configural Encoding of Unfamiliar Faces

Method

Subjects

The children who took part in Experiment 1 also took part in Experiment 2. Also, 18 young adults who had not taken part in Experiment 1 participated.

Materials

The two sets of faces (male and female) used with adults in Experiment 1 were used in Experiment 2. Neither the children nor the adults who served as subjects were familiar with these faces.

Procedure

Adults. For each set of same-sex faces, subjects first learned the first name of each person, by cycling through the set of whole faces displayed on the Macintosh screen. Usually adults learned the set of six names after 2 or 3 runs through the whole set. Following this, the procedure of Experiment 1 was followed: three practice trials with whole faces to familiarize subjects with the voice key, then naming the top half faces alone, and finally naming the top half faces in the composite and non-composite photographs (the experiment proper). For the inverted condition, the names were learned on upright faces; only in the final experimental condition were the faces displayed upside down. As in Experiment 1, whether subjects learned male faces first or female faces first and whether the first set was seen upright or inverted was counterbalanced.

Children. Pilot studies with 6-year-olds revealed that it was almost impossible for them to learn names for six faces under these conditions. Therefore, we modified the procedure as follows. Children began with two photographs, and learned the names of the two people depicted. Then a third photograph was added, and the set of three names practised until the children were fluent. This process was repeated until all six names were learned. At this point, additional practice was given with the stimuli presented on the Macintosh screen, until the children could produce all six names without hesitation. The procedure then continued as for adults. For first-graders, the initial learning process sometimes took two 20-minute sessions. Because of the extraordinary difficulty of learning to associate six names with six new faces, first-graders learned and were tested on only one set, presented in the upright orientation. Half the children learned the female set and half the male set. Fifth-graders learned both sets and were tested on one set upright and one inverted, counterbalanced as for the adults.

Results

Errors

Figure 4 shows the pattern of errors. The steep developmental function usually found on tasks involving the encoding of unfamiliar faces is reflected in these data. Both groups of children made substantial numbers of errors (first-graders, 21%, fifth-graders, 17%). These error rates were twice as high as those on familiar faces in Experiment 1 (8% at both ages). The adult error rate (10%) did not differ from that on familiar faces in Experiment 1 (10%).

The pattern of errors at all ages is consistent with configural encoding of unfamiliar faces. First-graders, who were tested only on upright faces, made twice as many errors on composites (28%) as on non-composites (14%), $F(1, 19) = 26.32$, $p < 0.001$. An ANOVA on the fifth-grade error data revealed a main effect for stimulus type, composites 21%, non-composites 14%, $F(1, 20) = 18.58, p < 0.001$, and a Stimulus Type × Orientation interaction (see Figure 4), $F(1, 20) = 10.45, p < 0.004$). An ANOVA on the adult error data revealed no effects of stimulus type or orientation, but the pattern was the same as that for the fifth-graders—more errors on composites than on non-composites upright, but no such difference for inverted faces (see Figure 4).

Reaction Times

An ANOVA on RT for correct responses was carried out at each age. Because of extremely long RTs (relative to other subjects of the same age) and high variability, two fifth-graders were removed from the analysis, leaving 18 of that age. As can be seen in Figure 5, subjects of each age were slower on upright composites than on upright non-composites. Adults and fifth-graders responded equally quickly to the two types when the stimuli were inverted. The Orientation × Stimulus Type interaction was significant for fifth-graders and approached this level for adults: fifth grade, $F(1, 17) = 5.58, p < 0.03$; adults, $F(1, 17) = 3.04, p < 0.09$.

First-Graders. Half of the youngest group of subjects learned the female faces and half the male faces. There was no effect of stimulus set. The difference between composites and non-composites (all upright) was significant, $F(1, 19) = 17.18, p < 0.001$).

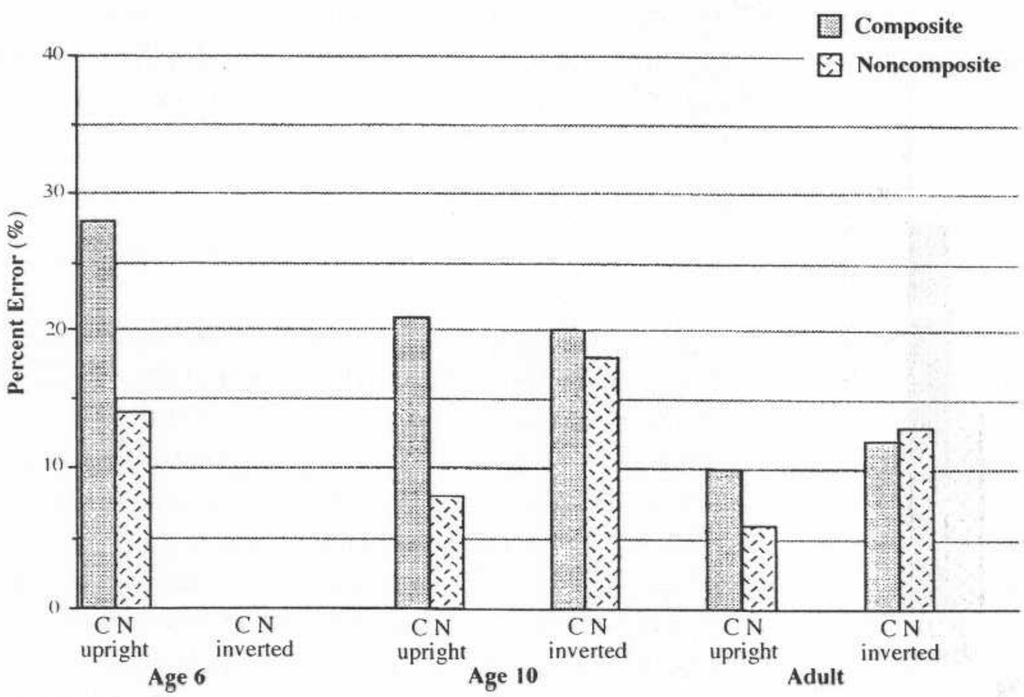

FIG. 4. Experiment 2: Error rates.

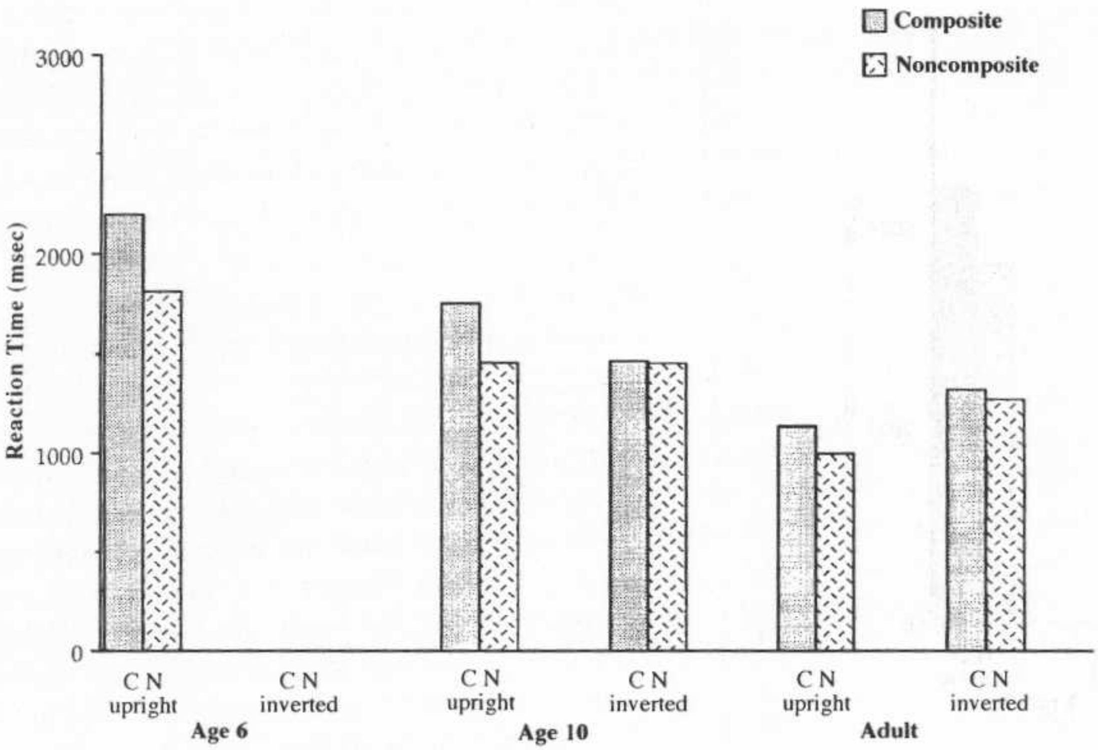

FIG. 5. Experiment 2: RT for correct responses.

Fifth-Graders. There were no effects of stimulus set or order (upright or inverted first). There was a main effect for stimulus type: $F(1, 17) = 10.93$, $p < 0.001$; RT for composites (1608 msec) was slower than for non-composites (1454 msec). The main effect for inversion approached significance, $p < 0.1$; RTs to upright faces (1602 msec) were *slower* than RTs to inverted faces (1460 msec). The only other significant effect was the Stimulus Type × Orientation interaction reported above (Figure 5).

Adults. There were main effects both for orientation and stimulus type. RT to upright faces (1065 msec) was faster than to inverted faces (1303 msec), $F(1, 17) = 9.11$, $p < 0.008$. RT to composites (1231 msec) was slower than to non-composites (1137 msec), $F(1, 17) = 10.56$, $p < 0.005$. The stimulus type by orientation interaction is depicted in Figure 5, and as mentioned above, approached significance. There were no significant effects involving any other variable.

Age Changes in the Effect of Orientation

Just as with familiar faces, inversion interferes with performance increasingly with age. Figure 6 shows the difference between RTs to upright and inverted stimuli. For both composites and non-composites, this difference score is greater for adults than for fifth-graders. Averaged over both stimulus types, fifth-graders

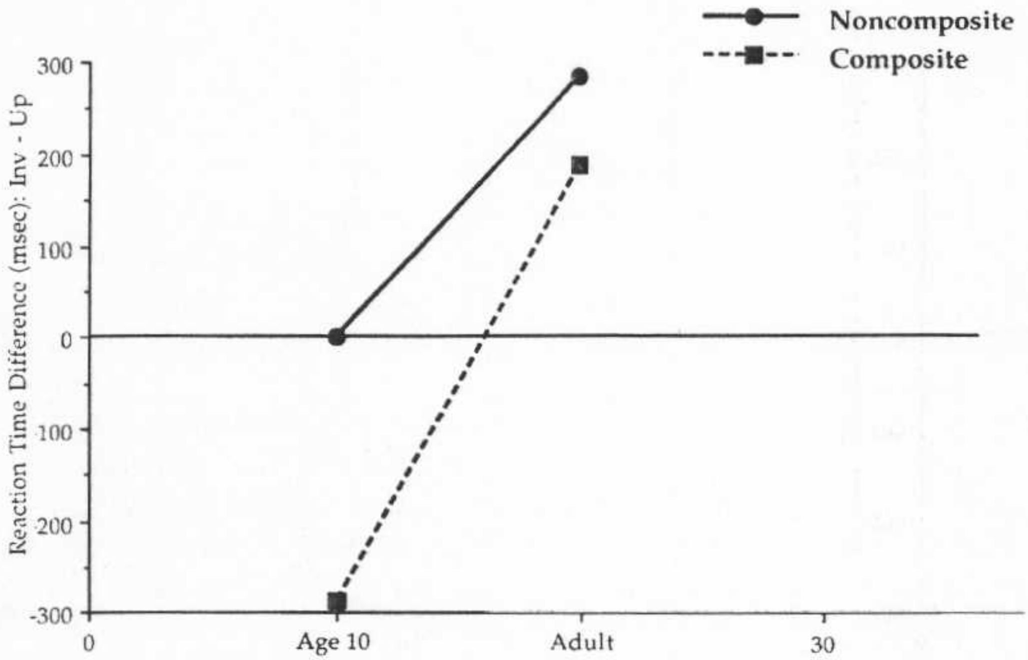

FIG. 6. Experiment 2: RT for inverted faces minus RT for upright faces.

are actually slower in the upright condition, whereas adults are faster in the upright condition. The adult advantage for the upright holds for composites as well as non-composites. This result confirms the magnitude of the adult advantage in recognizing faces presented upright over those presented inverted; this effect is so robust that it overwhelms the difficulty associated with disentangling the top half of upright composites. Just as for the familiar faces in Experiment 1, the best estimate of the true inversion effect comes from the non-composites alone: adults responded much faster on upright (993 msec) than on inverted (1279 msec) non-composites, whereas the fifth-graders were equally fast on the two (upright non-composites, 1464 msec; inverted non-composites, 1455 msec).

An overall ANOVA including the data from both fifth-graders and adults revealed a main effect for age, $F(1, 34) = 11.829$, $p < 0.005$. Children (1531 msec) responded more slowly than did adults (1184 msec). There was also a main effect for stimulus type, $F(1, 34) = 20.499$, $p < 0.001$, with composites being responded to more slowly than non-composites, and a Stimulus Type × Orientation interaction, $F(1, 34) = 8.265$, $p < 0.01$. There was no main effect of orientation, $F(1, 34) = 0.703$, n.s. Most important is the Age × Orientation interaction, $F(1, 34) = 11.334$, $p < 0.005$, depicted in Figure 6, and the absence of a three-way Age × Stimulus Type × Orientation interaction, $F(1, 34) = 2.007$, n.s. Thus these results present exactly the same pattern as those from Experiment 1.

Conclusion

These data confirm and extend findings in the literature concerning the developmental function for face encoding. (1) Six-year-olds find it extremely difficult to encode new faces from still photographs, as shown by the fact that some six-year-olds took 40 minutes to learn the names of six unfamiliar adults, and as shown additionally by the steep developmental functions for both error rates and RT. (2) These data show that gains in expertise are still being made after age ten, ten-year-olds make more errors and have slower RTs than adults. (3) These data confirm that this gain in expertise is reflected in increased sensitivity to inversion. Experiment 2 is the first study to show an Age × Orientation interaction in encoding unfamiliar faces over the range from age 10 to adulthood, presumably because other investigations have looked only at error rates in identification, not at more sensitive RT measures.

These data also confirm the robustness of the Young et al. pattern of findings. Composite uprights were more difficult to identify than were non-composites, as reflected both in error rates and RT, and there was no difference between the two types of stimuli when they were inverted.

The new finding from Experiment 2 parallels that from Experiment 1. There is no hint of a developmental change in configural or holistic encoding of unfamiliar faces. Six-year-olds made twice as many errors naming the top half of upright composites as they did naming upright non-composites, and their correct RTs were over 25% slower. Ten-year-olds, like adults, showed the pattern typical of the subjects in Young et al.'s study—interference from the composites when the stimuli were upright but not when they were inverted.

GENERAL DISCUSSION

Young et al.'s composite effect certainly reflects holistic encoding of faces; when upright, the parts of a face are less accessible than are whole faces. Anecdotally, subjects who knew these people very well reported that composite faces, such as that of Nixon and Prince Charles in Figure 1, resembled both familiar people, and that sometimes it was difficult to access which half was whose. The data from Experiments 1 and 2 show that 6-year-old children, no less than 10-year-olds and adults, encode faces holistically. The composite effect is independent of age, even in the presence of the effect that reflects acquisition of expertise, that is, the increased effect of inversion on face recognition. What this means is that the expertise effects are not the result of an increased reliance on holistic encoding.

Tanaka and Farah's (1993) task implicates holistic encoding in the same sense tapped by the Young et al. procedure—the whole face is more accessible than its parts. If expertise is not necessary for holistic encoding, then we should predict no developmental change over the same ages probed in the present studies. Tanaka (personal communication) has carried out a study with 6-, 8-, and 10-

year-olds with just this result. When faces were upright, at all ages equally it was easier to identify Bob's nose in the context of Bob's face than alone. There was no interaction with age, and we would predict no three-way Age × Condition (whole face/part) × Orientation interaction either.

Why is holistic coding sensitive to orientation of the face? That is, why does the whole face interfere with access to the parts only when faces are upright? The first obvious response—that the whole face is represented in memory with respect to a frame of reference (i.e. is oriented)—is only part of an answer, for we *can* recognize faces upside down, however badly. The second response is that, in the upright condition, the relations among the parts of the face create emergent features used in recognition, and these are less accessible (and thus less interfering) when faces are inverted. That is, the second response draws on the hypothesis that upright faces are configurally encoded.

Besides tapping holistic encoding, in the access sense, the Young et al. procedure reflects configural encoding. When the top half of Prince Charles' face is joined to the bottom half of Nixon's face, new relational features among parts of the face are created. If expertise is required for configural encoding, should we not have expected the three-way Age × Stimulus Type (composite/non-composite) × Orientation interaction? Only if children have *no* configural features in their representations of faces. As long as they represent some features of that type, then the interference due to the access sense of holistic encoding should be seen at all three ages, as it was. Thus, these data rule out an encoding switch from complete reliance on piecemeal features of faces to greater reliance on configural features (as proposed in Carey & Diamond, 1977) between ages 6 and adulthood; configural features and clearly represented at all levels of expertise tapped in the present study.

Apparently, there are at least two sources of the inversion effect for faces. First there is holistic encoding, as tapped by the composite effect. This is present throughout the age/expertise range studied here. And then there is something else that is gained with expertise. The something else is the mystery factor.

What is the mystery factor? It is very likely that it is just what we always thought—greater reliance on relational features with expertise. Once the top half of the face has been accessed, it must still be recognized. Adult (expert) recognition is based more on features that can be extracted more easily from upright faces, just as expert dog recognition is based much more on features than can be extracted more easily from upright dogs. A minimal reliance on such features is all that is required for holistic encoding in the access sense; but expertise at face encoding (or dog encoding) requires greater reliance on such features.

Several important pieces of the puzzle still need to fall into place. Exactly how are these relational features represented? Are they simply represented as ratios of distances among various points on a face? Many have suggested another possibility—that the features of an individual face are encoded with reference to a norm (e.g. see Rhodes & Tremewan, this issue). Expertise is not required for norm-based coding; Rhodes and Tremewan showed caricature effects for inverted faces that were equal in magnitude to those for upright faces, even in the presence of large effects of orientation. Ellis (1992) showed that young children judged caricatures of Kylie Monogue to be her best likeness. Whatever features, including relational features, are used to individuate faces, they may be encoded relative to a norm.

Still, within the framework of norm-based coding, a possible way to think about the increased reliance on relational features with expertise is that the representation of the norm changes with expertise. At all levels of expertise, it crudely reflects the shared overall configuration of faces, but with increasing expertise it becomes more and more completely specified. A simple metaphor is that it becomes specified in terms of more and more points. Thus, with increasing expertise, norm-based coding will engage many more points, and many more spatial relations among points. At all levels of expertise, distinctive configural features are encoded relative to the norm, but these become more adequate to distinguishing among highly similar faces as the norm becomes more fleshed out. As inversion interferes with norm-based coding of relational distinguishing features, improvement with expertise at encoding upright faces will be greater than improvement at encoding inverted faces (the Expertise × Orientation interaction).

Several well-known phenomena are consistent with this picture. Valentine (1991) showed that inversion interferes more with adult (expert) encoding of typical faces than with atypical faces. On the assumption that typical faces are closer to the norm, and thus that more subtle relational features are required to distinguish among them, typical faces will place higher demands on the norm-based coding mechanism that inversion disrupts. This analysis also predicts a Race × Orientation interaction for subjects who are experts at distinguishing among faces of only one racial group. That is, subjects should show a larger inversion effect for faces at which they are expert than for those from other racial groups. Rhodes, Tan, Brake, & Taylor (1989) obtained this result.[1]

Recent developmental findings from Ellis (1992) bear on the hypothesis that acquisition of expertise involves a fuller specification of the norm—the shared configuration among faces. He found a Distinctiveness × Age interaction in a face recognition task; that is, 6- to 7-year-olds showed no recognition advantage for atypical faces, and over the age range from 6 to 14 developmental improve-

[1]Bruce and Valentine (1986) found the opposite interaction—greater inversion effect for other race faces—but they did not have a full crossover design, so that their finding may reflect differences in difficulty between the two sets of faces. Also, the faces within each race were not as homogeneous as those of Rhodes and Tremewan, as Bruce and Valentine included stimuli with beards, moustaches, and glasses. Finally, Bruce and Valentine equated performance on the upright by giving subjects less time to encode same-race faces, and this manipulation may have interfered with the encoding of just those subtle relational features that are more easily encoded from upright faces.

ment at encoding atypical faces was greater than that for encoding typical faces. Typicality was determined by adult (expert) ratings. These data show that the youngest children's norm is not adequate to distinguishing what adults judge as typical faces from those adults judge atypical. As the norm becomes more fully specified, distinguishing features from atypical faces are the first to become employable.

As previously noted, the results from these studies argue against the hypothesis that there is an encoding shift between ages 6 and 10 from complete reliance on relatively more piecemeal distinguishing features to greater reliance on relatively more cnfigural distinguishing features of faces (Carey & Diamond, 1977). At all ages, children encode faces in terms of configural distinguishing features. At all ages, children are sensitive to the orientation of the face; at least from age 6 they gain access to the whole upright face faster than to its parts, and at least from age 4 they show caricature effects when only the configuration of the face is manipulated. However, it is possible that the Age × Orientation interaction that marks increasing expertise at face encoding reflects a fuller specification of the shared configuration of the face, so that the young child's configural encoding involves many fewer features than does the 10-year-old's or the adult's. In this sense only, then, is it likely that expertise reflects increasing reliance on configural distinguishing features of the face.

REFERENCES

Bartlett, J.C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25* (3), 281–316.

Benson, P.J., & Perrett, D.I. (1991). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology, 3,* 105–135.

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

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

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

Brooks, R.M. & Goldstein, A.G. (1963). Recognition by children of inverted photographs of faces. *Child Development, 34,* 1033–1040.

Bruce & Valentine (1986).

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

Carey, S. & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195,* 312–314.

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

Chance, J., Goldstein, A.G., & Schicht, W. (1967). Effects of acquaintance and friendship on children's recognition of classmates' faces. *Psychonomic Science, 7,* 223–224.

Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology, 23,* 1–22.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. Journal of Experimental Psychology: *General, 115* (2), 107–117.

Ellis, H.D. (1992). The development of face processing skills. In V. Bruce, A. Cowey, A.W. Ellis, & D.I. Perrett (Eds.), *Processing the facial image* (pp. 105–111). Proceedings of a Royal Society Discussion Meeting, July 9–10, 1991. Oxford: Clarendon Press.

Ellis, H.D., Shepherd, J.W., & Davies, G.M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8,* 431–439.

Fagan, J.F. (1979). The origins of facial pattern recognition. In M.H. Bornstein & W. Kessen (Eds.), *Psychological development from infancy: Image to intention.* (pp. 83–113). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Farah, M.J., Tanaka, J.W. & Drain, H.M. (in press). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance.*

Flin, R. (1980). Age effects in children's memory for unfamiliar faces. *Developmental Psychology, 16,* 373–374.

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

Goldstein, A.G. (1975). Recognition of inverted photographs of faces by children and adults. *Journal of Genetic Psychology, 127,* 109–123.

Goldstein, A.G., & Chance, J.F. (1964). Recognition of children's faces. *Child Development, 35,* 129–136.

Goldstein, A.G., & Mackenberg, E.J. (1966). Recognition of human faces from isolated facial features: a developmental study. *Psychonomic Science, 6,* 149–150.

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

Johnson, M.H., & Morton, J. (1991). *Biology and Cognitive Development: The Case of Face Recognition.* Oxford: Basil Blackwell.

Levine, S.C. (1981). Developmental changes in right hemisphere involvement in face recognition. In C.T. Best (Ed.), *Developmental neuropsychology and education: Hemisphere specialization and integration,* (pp. 157–191). New York: Academic Press.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Rhodes, G., Brake, S., & Atkinson, A. (1993). What's lost in inverted faces? *Cognition, 47,* 25–57.

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

Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology, 80,* 313–331.

Saltz, E., & Sigel, I.E. (1967). Concept overdiscrimination in children. *Journal of Experimental Psychology, 73,* 1–8.

Scapinello, K.F., & Yarmey, A.D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21,* 329–330.