Child Development, September/October 2008, Volume 79, Number 5, Pages 1513 – 1523

Limitations in 4-Year-Old Children's Sensitivity to the Spacing Among Facial Features

Catherine J. Mondloch and Kendra Thomson
Brock University

Four-year-olds' sensitivity to differences among faces in the spacing of features was tested under 4 task conditions: judging distinctiveness when the external contour was visible and when it was occluded, simultaneous match-to-sample, and recognizing the face of a friend. In each task, the foil differed only in the spacing of features, and spacing alterations were within normal limits. Children performed at chance levels in all but 1 task—match-to-sample, and in that task, only 10 of 18 children were correct on more than 50% of the trials. Sensitivity to the spacing of facial features in identity and distinctiveness tasks is very weak at 4 years of age—at least when the changes do not exceed ±2.5 *SD* of normal variability.

Adults are experts at recognizing facial identity: They can recognize thousands of faces, including, at least for a short time, ones they encounter briefly in the course of a laboratory experiment (see Bruce & Young, 1998, for a review). Adults use multiple cues to recognize the identity of individual faces, two of which have been studied extensively in children: the shape of individual features (e.g., eyes, mouth, and chin: *featural processing*) and the spacing of the facial features (e.g., the distance between the eyes, a type of configural processing called *second-order relational processing*; for a review, see Maurer, Le Grand, & Mondloch, 2002). Despite the early emergence of some face-processing skills, adult-like expertise is not achieved until some time during adolescence: Even in matching tasks, which eliminate memory demands, performance improves dramatically between 4 and 11 years of age (Bruce et al., 2000; see also De Sonneville et al., 2002).

Two lines of evidence suggest that children's poor performance on face recognition tasks can be attributed in part to immature sensitivity to second-order relations. First, children perform like adults on identity tasks that tap sensitivity to the appearance of individual features. For example, when asked to make same/different judgments about pairs of faces that differ only in the shape of the external contour, 6-year-olds are as accurate as adults. They are nearly like adults when making same/different judgments about pairs of faces that differ only in the shape of internal features (eyes and mouth), with no statistical difference by 10 years of age (Mondloch, Le Grand, & Maurer, 2002; see also Freire & Lee, 2001). Second, children aged 6 – 14 years make significantly more errors than adults when asked to make same/different judgments about pairs of faces that differ only in the spacing of features (Freire & Lee, 2001; Mondloch, Dobson, Parsons, & Maurer, 2004; Mondloch et al., 2002, 2003).

There is general consensus that children are less sensitive than adults to the spacing among facial features (but see Gilchrist & McKone, 2003; Pellicano, Rhodes, & Peters, 2006), but there is debate as to when this sensitivity first emerges. Although less accurate than adults, even 6-year-olds' accuracy is above chance levels in making same/different judgments about upright faces differing only in the spacing of their features (M = 0.69 correct vs. 0.82 correct for adults tested with the same task; Mondloch et al., 2002). In addition, like adults, 7-year-old children remember distinctive faces better than typical faces, regardless of whether the faces are made distinctive by manipulating features (e.g., bushy eyebrows) or their spacing (e.g., mouth up; Gilchrist & McKone, 2003).

To investigate when sensitivity to the spacing among facial features first emerges, Mondloch, Leis, and Maurer (2006) asked 4-year-old children to recognize each of three faces when paired with a foil that

This research was funded by an NSERC Discovery Grant to Catherine J. Mondloch. The original adult faces were taken from the MacBrain Face Stimulus Set. Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development. Please contact Nim Tottenham at tott0006@tc.umn.edu for more information concerning the stimulus set.

Correspondence concerning this article should be addressed to Catherine J. Mondloch, Department of Psychology, Brock University, St. Catharines, Ontario, Canada L2S 3A1. Electronic mail may be sent to cmondloch@brocku.ca.

 0009-3920/2008/7905-0019

differed from the target in one of three ways: the shape of the external contour, the appearance of individual features, or the spacing among features. Four-year-olds were familiarized with two of the target faces over a 2-week period by providing them with a storybook depicting two children on a visit to the farm. To encourage attention to second-order relations, Mondloch et al. varied the appearance of facial features by presenting the faces of the two children in the storybook from different points of view and with a variety of emotional expressions. The third target face was the child's own. Recognition was tested in the context of a computerized game in which 4-year-olds used a ''magic wand'' to put children on a bus or train during a trip to the zoo. Four-year-olds made no errors on control trials in which the foil was the face of another child or in which the foil had a different external contour; they performed above chance levels when the foil differed in the shape and color of the eyes and mouth. In contrast, 4-year-olds performed at chance levels when the foil differed from the target only in the spacing of facial features, both when tested with the two target faces from the storybook and when tested with their own face. Mondloch et al. concluded that sensitivity to second-order relations emerges after 4 years of age—at least for facial identity tasks when the spacing of features remains within normal limits

In the same year that Mondloch et al. (2006) published their manuscript, two other studies reported that 4-year-old children are sensitive to the spacing of facial features. Pellicano et al. (2006) tested children's sensitivity to the spacing of facial features using the *whole/part* paradigm. The *whole/part effect* is a measure of holistic processing (Tanaka & Farah, 1993) and refers to the finding that adults and children as young as 4 years of age recognize the features from an individual's face more easily in the context of the whole face (e.g., Larry's nose in Larry's face) than in isolation (Pellicano & Rhodes, 2003; Tanaka, Kay, Grinnell, Stansfield, & Szechter, 1998). The usual interpretation of the whole/part paradigm is that poor performance in the isolated features condition indicates that faces are processed holistically (i.e., as a gestalt) and that characteristics of individual features are encoded in the context of this gestalt representation (Tanaka & Farah, 2003). Disrupting holistic processing by presenting features in isolation impairs recognition of features, as does altering the gestalt by changing second-order relations (Tanaka & Sengo, 1997). Pellicano et al. showed that altering the spacing of facial features in one part of a face also impairs the ability of 4-year-olds to recognize facial features in the whole condition.

McKone and Boyer (2006) reported that 4-year-olds are sensitive to differences in the spacing of facial features when asked to judge which of two faces is more distinct. They manipulated facial distinctiveness by altering the appearance of individual features (e.g., by making the eyebrows bushy) or by altering the spacing among features. Based on adult ratings of the original and distorted faces viewed one at a time, McKone and Boyer created pairs of faces in which differences in distinctiveness between the original face and its altered version were large, medium, or small. Adults and 4-year-olds were then asked to indicate which member of each pair was more distinctive; adults' performance was above chance for all pairings, and 4-year-olds' performance was above chance when differences in distinctiveness were either large or medium.

The purpose of the present set of studies was to resolve these discrepant findings. It may be that sensitivity to differences among faces in the spacing of facial features emerges at different rates for recognizing a face's identity than for judging distinctiveness. In the current series of studies, we measured 4-year-olds' sensitivity to the spacing of facial features under several task conditions—when judging distinctiveness, in a simultaneous match-to-sample task, and when recognizing the faces of friends from a day care. In all these studies, the spacing changes were identical to those used by Mondloch et al. (2006). This allowed us to assess whether 4-year-olds would show sensitivity to those spacing changes under different task conditions. Furthermore, we made each task child friendly: The number of trials was minimized (four test trials plus eight control trials), and children were presented with a ''game'' in which they used a magic wand to put faces on a bus or put twins into a ''twin clubhouse.''

Experiment 1a

Previously, we reported that 4-year-olds were not able to recognize either faces from a storybook or their own face when targets were paired with foils that differed only in the spacing of facial features (Mondloch et al., 2006). This result is in contrast to that of McKone and Boyer (2006) who reported that 4-year-olds were able to select the more distinctive of two faces that differed in the spacing of features, as long as the differences in distinctiveness were either large or medium according to adult judgments. In Experiment 1a, we showed children four pairs of faces that differed in the spacing of features and asked them to use a magic wand to put the more unusual child on an ''unusual'' bus. Each of the spatially altered faces had been rated as more distinctive than its original version when adults viewed three versions of each of 26 faces sequentially (Mondloch et al., 2006); in that study, 4-year-olds performed at chance levels in an identity task. If the spacing of facial features influences children's judgments of distinctiveness by 4 years of age, they should select the more distinctive version of each face (Mondloch et al., 2006). We also tested a group of adults in the current study using the method designed for children. We expected that adults would perform without error, thus validating our stimulus set.

Method

Participants

Eighteen 4-year-old children (9 females; mean age = 4.43 years, range = 4.0 – 4.9 years) participated in this study. Twelve adults (mean age = 19.9 years, range = 19.0 – 23.6) with normal or corrected-to-normal vision served as a control group. All participants were Caucasian. Children in all the experiments were recruited from a database of children whose parents had volunteered for participation or from local day cares whose parents had provided informed consent. Adult participants received course credit.

Materials

Test stimuli. On each of the four test trials, we presented an original (i.e., unaltered) face paired with a foil that differed only in the spacing of internal features. We selected four face pairs from Experiment 1 of Mondloch et al. (2006). We used the original faces of the two storybook characters plus the faces from two of the participants in that study. Adults had previously rated each of these faces on a distinctiveness scale (Mondloch et al., 2006, Experiment 2b) and reported that the spatially altered versions were more distinct than the original faces (mean difference = 1.66, $SE = 0.22$). In this and all subsequent studies, the faces were close to life size: The distance between the chin and the hairline was 130 mm. Based on anthropomorphic norms (Farkas, 1994), we had moved the eyes up/down 21 pixels (where 1 pixel = 0.353 mm; 1.3 *SD*; 0.71° from the testing distance of 60 cm) or closer together or farther apart 14 pixels (2.8 *SD*; 0.47° from 60 cm) and we moved the mouth up/down 7 pixels (0.8 *SD*; 0.24° from 60 cm). Thus, these spacing changes covered most of the natural variability in the real world and consequently remained within normal limits. Two of the test faces had the eyes and mouth moved up and two had the eyes and mouth moved down. Figure 1a provides an example.

Criterion stimuli. Two sets of criterion stimuli were used to ensure that the 4-year-olds understood the task and that they continued to "play the game" throughout the procedure. One set consisted of four animals (e.g., a horse, a squirrel) paired with a "weird" version of that same animal (e.g., a horse with stripes, a squirrel with pigs' ears; Figure 1b). The other set consisted of grayscale photographs of adult Caucasian women paired with a spatially altered version in which the spacing of facial features was outside normal limits. Adults and 8-year-olds rate these spatially altered versions as grotesque (Mondloch et al., 2004), and unpublished data of Mondloch and Maurer (2003) show that 4-year-olds find these faces weird (Figure 1c).

Procedure

The child sat 60 cm away from a 58-cm computer screen. We began by presenting a picture of two buses, one of which was a typical school bus and one of which was weird. The experimenter said, "Look at these two buses. Do you think that one of these buses is a bit strange or unusual or weird? Which one? Why do you think that bus is weird? Good job! Now we are going to go for a ride with this bus and help some people and animals get back to their town. Do you think you can help me? Great!" Each trial consisted of two photos of animals, children's faces, or adult female faces. The experimenter asked, "Do you think one of these (animals, females, or children) is a bit strange, unusual, weird, or might stick out to you if you saw them on the playground? Good, now wave your magic wand to put the weird (animal, female, or child) on the weird bus!" After the child waved their magic wand over one of the pictures, a small picture of that stimulus appeared on the bus. The order of the trials was such that all participants received four criterion trials (two animals, two adult female faces), followed by two test trials (children's faces). Participants then received two more criterion trials (animals), followed by the remaining two test trials. We ended with two final criterion trials (adult female faces) to ensure that children were still "playing the game." An alternate version of the task was created so that the left – right positions of the stimuli were reversed. To ensure that the experimenter was blind to the location of the more distinctive stimulus, another laboratory member opened one of the two files on the computer for the experimenter. The experimenter viewed the stimuli on a laptop computer on which the pictures were occluded. This

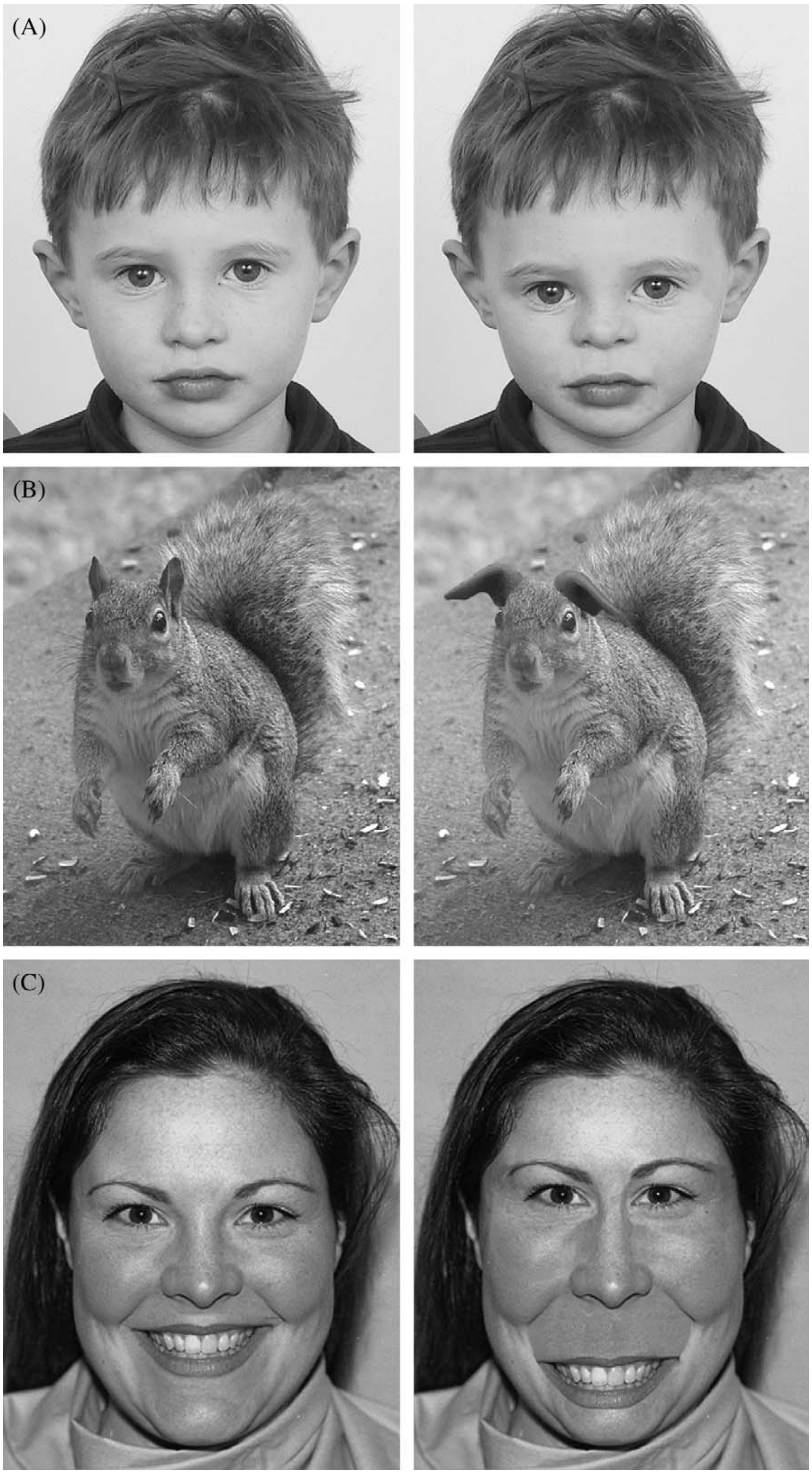

Figure 1. Examples of stimuli used during (a) test trials, (b) criterion trials in which animals were presented, and (c) criterion trials in which adult faces were presented. In each case, the original picture is shown on the left and the altered (i.e., distinctive) picture is shown on the right. The spacing changes made to adult faces were outside normal limits; spacing changes made to children's faces were within normal limits.

protocol was approved by the Research Ethics Board at Brock University.

Results and Discussion

Children made almost no errors when asked to find the more distinctive animal, mean correct (on four trials) = 3.90, $t(17) = 35.0$, $p < .001$, and when asked to find the more distinctive adult face, mean correct = 3.70, $t(17) = 14.6$, $p < .001$ (all tests were one-tailed). Children performed at chance levels, mean correct = 1.90, on trials in which the original version of a child's face was paired with its spatially altered version and the spacing changes were within normal limits. Only 7 of the 18 children tested were correct on at least three of the four test trials. A control group of 12 adults (7 females) performed at 100% accuracy on all trial types, confirming that the spatially altered faces are easily detected as the most distinctive by adults. These data suggest that 4-year-olds are not able to use information about the spacing of facial features to determine which of two faces is more distinctive when the spacing changes are within normal limits. This is in contrast to the findings of McKone and Boyer (2006) who reported above-chance performance in 4-year-olds—at least when differences in distinctiveness were large or medium in size. We note, however, that McKone and Boyer occluded the external contour of their face stimuli, whereas the hair was visible in our stimuli, albeit identical within each stimulus pair. Several studies have demonstrated that the external contour is a salient stimulus to young children (Campbell & Tuck, 1995; Campbell, Walker, & Baron-Cohen, 1995; Mondloch et al., 2006); to see whether the presence of hair interfered with children's performance, we replicated the procedure of Experiment 1a, but occluded the hair.

Experiment 1b

Method

Participants

Eighteen 4-year-old children (5 females; mean age = 4.38 years, range = 4.0–4.7 years) participated in this study. All participants were Caucasian. Experiment 1b was identical to Experiment 1a, with one exception: The hair was occluded on children's faces (participants were told that the children were peeking through windows).

Results and Discussion

No child made an error on criterion trials, but performance was at chance levels when children were asked to find the more distinctive face on test trials, mean correct = 2.20, $t(17) = 0.62$, $p > .2$. Only 6 of the 18 participants were correct on at least three of the four test trials. These data confirm that 4-year-olds do not attend to the spacing among facial features per se when making distinctiveness judgments.

At first glance, it appears that our results contradict those of McKone and Boyer (2006) who reported that 4-year-old children judge spatially altered faces as more distinctive than their original versions. However, in their study, 4-year-olds were much less accurate than adults even when the spacing changes were large (mean correct = 0.90 vs. 0.62 for adults and children, respectively). When differences in distinctiveness were small, 4-year-olds did not perform above chance levels, whereas adults were correct on approximately 75% of the trials. Thus, even in McKone and Boyer's study, 4-year-olds' sensitivity to the spacing of facial features was much weaker than that of adults and was most reliable when differences in distinctiveness were large according to adult ratings. Likewise, in the current study, 4-year-olds were able to use information about the spacing of facial features when the spacing changes were large and outside normal limits (i.e., when making distinctiveness judgments about the adult faces). The high level of accuracy shown for adult faces with large spatial distortions (12 children correct on four of four trials; 6 children correct on three of four trials) in contrast to chance performance for children's faces with spatial distortions that contrast more or less average spacing with the edge of natural limits suggests that 4-year-old children are sensitive to differences among faces in the spacing of features but only when those differences are large. One potential limitation of our study is that two training trials in which differences in distinctiveness were large preceded the first two test trials. We did this in order to familiarize children with the task. Although one possibility is that the spacing changes made to the children's faces were harder to detect after viewing the large spacing changes in adult faces, we reasoned that viewing the large spacing may direct 4-year-olds' attention to this cue to distinctiveness. Nonetheless, we note that this may have made the task more difficult for the children.

Experiment 2

In Experiment 2, we simplified the task for 4-year-old children by presenting them with a simultaneous match-to-sample task. On each of four test trials, they were shown a child's face (either original or spatially altered) and then shown a pair of faces, one of which was identical to the target and one of which differed only in the spacing of features. Children were asked which member of the face pair matched the target; they then placed the ''twins'' in a ''twin clubhouse.'' This task involved no memory demands and allowed the child to compare the target face to two test faces with no time restrictions. We also included criterion trials in which children were presented with sets of adult faces and animals from Experiment 1.

Method

Participants and Materials

Eighteen 4-year-old children (8 females; mean age = 4.44 years, range = 4.0–4.9 years) participated in this study. All participants were Caucasian. The test stimuli were identical to those used in Experiment 1a; they were printed and laminated to allow the child to handle the faces. We created a twin clubhouse into which children were invited to place identical pictures.

Procedure

Each participant sat at a table with the stack of cards containing stimulus pairs in front of him or her. The experimenter sat next to the participant with an additional stack of individual stimuli. The twin clubhouse was placed directly in front of the participants. First, participants were shown a drawing of three bears, two of which were identical and one that was only slightly different. The experimenter said, ''Do you know what twins are? Can you pick out the twins from these three bears?'' If participants did not understand the differences between the bears, the experimenter explained, ''Twins are people or animals that look exactly the same.'' Next, the experimenter told participants, ''Today I have a special job for you! You are going to be a 'twin detective'! Some kids, animals and adults are starting a special 'twin clubhouse' where only twins can go. Can you help me find twins to put in the twin clubhouse? Great! First I am going to flip over a card with an animal (or face) on it and then you can flip your card that will have two pictures of animals (or faces) on it. Your job is to pick the animal (or face) from your card that is exactly like the animal (or face) on my card (the animal's/person's twin)! Then we can put the twin animals/people in the twin clubhouse. Remember to look carefully at both of the pictures on your card!''

Results and Discussion

Children made no errors on the animal criterion trials, and only 3 children made a single error on the criterion trials in which we presented adult faces. Their performance was above chance, $M = 2.40$, $t(17) = 1.80$, $p < .05$, when tested with children's faces, though only 10 of the 18 participants were correct on at least three of the four test trials. These data suggest that the visual system of 4-year-old children is able to detect differences in the spacing of facial features, a finding that is consistent with report by Pellicano et al. (2006) of sensitivity to the spacing of features in the whole/part paradigm. Four-year-olds showed sensitivity to the spacing of facial features in our match-to-sample task in which they were required only to match facial identity; this is in contrast to when they were asked to recognize a face from memory (Mondloch et al., 2006) or to decide which of two faces was distinctive (Experiment 1). Nonetheless 8 of the 18 children tested were not correct on more than 50% of the trials, suggesting that 4-year-olds' sensitivity to the spacing of facial features is weak even in a matching task. Based on this, it is unlikely that they use the spacing of facial features to recognize faces in the real world.

Nonetheless, we may have underestimated children's sensitivity to second-order relations by testing them with unfamiliar faces. Carver et al. (2003) reported an effect of familiarity on the Nc component of the event-related potential waveform between 18 and 24 months of age; this raises the possibility that children may process highly familiar faces differently from unfamiliar faces. Mondloch et al. (2006) tested 4-year-olds with familiarized faces from a storybook and with their own faces. We reasoned, however, that 4-year-olds do not see their own face as often as that of people with whom they interact on a daily basis and they do not see their own face from a variety of perspectives and displaying a range of facial expressions. In Experiment 3, we tested the hypothesis that 4-year-olds may be sensitive to the spacing of facial features when asked to identify the ''real'' picture of a highly familiar face—that of a friend from their day care. To do so, we replicated the original method of Mondloch et al. (2006) but replaced pictures of the storybook character and the participant's own face with the faces of three friends from the participant's day care.

Experiment 3

Method

Participants

Eighteen 4-year-old children (6 females; mean age = 4.25 years, range = 4.0–4.9 years) participated in this study. All participants were Caucasian. All participants attended a day care and were highly familiar with the faces of children who served as models. Two additional children were tested but excluded from the analyses. One child was able to identify only two of his three friends when shown the initial array of stimuli, and the other was incorrect on all three contour trials. This poor performance is very atypical (Mondloch et al., 2006); therefore, we excluded these children to avoid underestimating 4-year-olds' sensitivity to the spacing of features.

Stimuli

For each participant, a day-care worker identified three children with whom the participant was highly familiar; these children served as models. After obtaining informed parental consent, we took three pictures of each model's face. We then created new versions of each of these pictures. On contour trials, we presented one original version paired with a foil in which we combined the external contour (hair, ears, and chin) of a different child with the internal features and the clothing of the original image. On test trials, we presented an original photograph paired with a spatially altered version. These spatial alterations were identical to those used by Mondloch et al. (2006). On half of the trials, the eyes and mouth were moved either up or down (as in Experiments 1 and 2); on the remaining trials, the eyes were moved 14 pixels closer together and the mouth was moved down 7 pixels or the eyes were moved 14 pixels farther apart and the mouth was moved up 7 pixels. A different original picture was presented on each spacing trial to ensure that participants were recognizing a face, rather than a particular photograph. On control trials, we presented one of the original pictures of the model paired with a picture of another 4-year-old child with whom the participant was not familiar.

Procedure

The child sat 60 cm away from a 58-cm computer screen; during the test, the experimenter faced the child so as to be blind as to on which side of the monitor the ''correct'' face was presented. We began by presenting an array of six pictures, each of which depicted a 4-year-old child and three of which were friends of the participant. The participant was asked to identify their three friends.

On the first three test trials, the child was shown one of the original faces paired with a face with the same internal features but a different external contour. The background picture consisted of a school bus. The child was told, ''Today we are going to pretend that some of your friends are going on a trip to the zoo! But the bus driver doesn't know what all the kids in the class look like, so maybe you can help him out. Here is a magic wand and when you wave the wand the children will get on the bus.'' Two faces appeared on the face at this time. ''Look! It's 'Sam'! And there's another boy who looks a little bit like Sam. Can you wave the magic wand over the real picture of Sam to put him on the bus?'' After the child pointed the magic wand to the face he or she thought was Sam's, a new display appeared that consisted of the interior of the bus and a small picture of Sam's face. On this and every subsequent trial, the correct response was placed on the bus (or train). Two additional contour trials followed in which the child was asked to help the bus driver by pointing to their two remaining friends. The purpose of these trials was to illustrate to the child that the members of each face pair would be quite similar, but that only one would be the real person.

The child was then shown a series of four animal pictures (e.g., a monkey, an elephant) and the experimenter continued the story about the trip to the zoo. Then, the game continued: ''Oh look, there is a train ride. The kids want to go on the train ride at the zoo, but the conductor doesn't know what the kids look like. Maybe you can help him out. But, it is going to be harder now. Just for today, Sam has a brother who looks a lot like him so it's going to be more difficult to tell them apart so look very carefully.'' Two faces were then presented on the screen; the faces differed only in the spacing of the facial features. ''Here's Sam and Sam's brother. They look a lot alike, but only one of them is really Sam. Can you wave your magic wand at Sam to put him on the train?'' When the child waved the magic wand, the experimenter presented a new picture that consisted of a picture of the train overlaid with a small picture of Sam's face. Two additional trials followed in which different friend's faces were presented. The child was then told that it was time for the children to go home but that the bus driver was sad because he could not remember what the children in the class looked like. A second set of spacing trials then ensued, one for each of the three original faces.

To ensure that the child was still attentive at the end of the procedure, three control trials were included. The child was told, ''So now the bus takes them all the way back to school. At school, their mommies and daddies are waiting to pick them up. Can you help Sam's mom pick out Sam?'' Two faces were presented side-by-side on the screen, but this time, Sam was presented next to a completely different boy. The background consisted of a cartoon version of a car.

After the child pointed to the face he or she thought belonged to Sam, Sam appeared in the passenger side of the car. The same test was then repeated for the other two friends. Any child who could not correctly identify these faces would have been excluded from the study. However, not one child made an error.

For each participant, we created two versions of the test in which the positions of the faces were reversed left to right. We then asked a laboratory member other than the experimenter to randomly open one of the two files to ensure that the experimenter was unaware of on which side the target face was presented.

Results and Discussion

One child made a single error on the contour trials, and no child made an error on control trials. In contrast, 4-year-olds performed at chance levels on the six spacing trials, $M = 3.39$, $t(17) = 1.13$, $p > .10$. Nine of the 18 participants were correct on at least four of the six test trials. Accuracy on the three trials in which the eyes were moved up and down ($M = 1.67$) did not differ from accuracy in which the eyes were moved horizontally, $M = 1.72$, $t(17) = 0.164$, $p > .50$. These results suggest that in previous studies, we had not underestimated 4-year-olds' sensitivity to the spacing among facial features by testing them with unfamiliar or familiarized faces. Even when tested with the faces of children with whom they interact on a daily basis, 4-year-old children do not use differences in the spacing of facial features for face recognition.

General Discussion

This set of studies indicates that 4-year-old children are not sensitive to small differences among faces in the spacing of facial features. They did not select spatially altered faces as more distinctive than the unaltered versions (Experiments 1a and 1b), and they did not select the unaltered versions as the real picture of their friends (Experiment 3). Across these three tasks, no more than half of the children performed better than a chance level of 50%. Their performance was above chance in the simultaneous match-to-sample task (Experiment 2), but only 10 of the 18 children were correct on more than 50% of the trials. Poor performance on these tasks cannot be attributed to 4-year-olds not understanding task instructions; their performance was almost errorless on control trials in which they were asked to select the more distinctive animal or adult female face in which the spacing changes were large and outside normal limits (Experiments 1a and 1b) and when asked to match the identity of these stimuli (Experiment 2). Similarly, only one child made a single error when asked to find the real picture of a friend's face when the foil differed in the appearance of the external contour. Poor performance is also unlikely due to the limited number of test trials. Our decision to include four test trials (plus eight control trials) in Experiments 1 and 2 was based on extensive pilot work showing that increasing the number of trials on a difficult task results in several children failing to complete the procedure and in random performance on control trials. A trial-by-trial analysis reveals no consistent improvement across trials. Across the first three studies, 27 of 54 children were correct on the first test trial and 31 of 54 children were correct on the last test trial. Rather, these results suggest that 4-year-old children are not sensitive to differences among faces in the spacing of facial features, at least when the spacing of features does not exceed ± 2.5 *SD* of normal variability.

The nature of our spacing changes should have facilitated 4-year-olds' ability to use second-order relations in each task. In our distinctiveness tasks and in the match-to-sample task, we always moved the eyes vertically. This decision was based on adult ratings of distinctiveness from a previous study (Mondloch et al., 2006). Goffaux and Rossion (2007) reported that inversion disrupts the ability to discriminate faces in which features have been moved vertically more than the ability to discriminate faces in which features have been moved horizontally or in which the appearance of individual features has been altered. They suggest, "It may be that vertical relations are more significant for upright face discrimination than horizontal relations" (p. 1001). This is consistent with evidence from the composite face effect of stronger perceptual integration across the top and bottom halves of faces than across the two sides (Hole, 1994). Thus, the use of vertical movements should have made this task easier for children. In the friends study (Experiment 3), we moved the eyes vertically in half of the trials and horizontally in the other half. When we moved the eyes farther apart, the mouth was moved up; when we moved the eyes closer together, the mouth was moved down. These combinations of movement should have maximally disrupted the aspect ratio of the triangular relations in the original faces (Barton, Zhao, & Keenan, 2003), thus increasing 4-year-olds' sensitivity to the changes in spacing.

Our results are different from those of two previous studies reporting that 4-year-old children are sensitive to small differences among faces in the spacing of facial features. McKone and Boyer (2006) reported that 4-year-olds were able to select the more distinctive of two faces that differed only in the spacing of features, at least when differences in distinctiveness were medium or large. To create three face sets, they altered the spacing among features and then used adult ratings of distinctiveness to form face pairs in which *differences* in distinctiveness were classified as small, medium, or large. After completing our studies, McKone provided us with the complete set of the stimuli used by McKone and Boyer. We resized McKone and Boyer's stimuli to match ours and then measured how far features were moved when creating the distinctive face. We note several differences between our stimuli and those of McKone and Boyer. First, McKone and Boyer moved the eyes of all but two faces horizontally; the eyes of some faces also were moved vertically. Second, the mean number of pixels by which eyes were moved was larger in all three face sets (small: $M = 16.73$, medium: 15.81, large: 19.38) relative to the horizontal movements used in our current studies (14 pixels). The mean number of pixels by which the mouth was moved was also larger in the McKone and Boyer face sets (small: $M = 8.91$, medium: 8.76, large: 8.13) than in the current study (7 pixels). Third, within each face set, the number of pixels by which features were moved varied (e.g., eyes in the "small differences" face set were moved between 0 and 28 pixels), and within each set, the eyes of some faces were moved at least 20 pixels (4 *SD*). In sum, the average horizontal eye movements used by McKone and Boyer were larger than ours, and some were especially large. We note that despite using larger spacing changes, 4-year-olds' accuracy was low in the McKone and Boyer's study for the large (mean correct = 62%) and medium ($M = 58\%$) face sets and at chance (50%) for the small face set.

The results of our present studies also appear to be inconsistent with those of Pellicano et al. (2006) who showed that altering the spacing of facial features in one part of a face (e.g., the eyes) impairs the ability of 4-year-olds to recognize facial features in the same region or in another part of the face (e.g., the mouth). One possibility is that different task demands lead to differences in children's sensitivity to the spacing among features. In addition, the spacing among facial features in the part/whole task were larger than those used in the current study. Pellicano et al. created two versions of each face (e.g., one version with the eyes moved out and one with the eyes moved in). Those spacing changes were within normal limits and were identical to those used in Experiment 3 of the current study when features were moved horizontally. However, Pellicano et al. did not present the original face on test trials; rather, the model face had eyes moved in one direction (e.g., farther apart) and the test face had eyes moved in the opposite direction (i.e., closer together; L. Pellicano & G. Rhodes, personal communication, November 2007). This is in contrast to our studies in which the original, unaltered version of a face was presented with a spatially altered version in which the eyes were moved in one direction. Consequently, differences in the spacing among facial features were twice as large in the study by Pellicano et al. (approximately 5 *SD*) than in the current study (L. Pellicano & G. Rhodes, personal communication, November 2007).

Our current results also may be surprising in light of recent evidence that 5-month-old infants show sensitivity to differences in the spacing of facial features that fall within normal limits (Hayden, Bhatt, Reed, Corbly, & Joseph, 2007). Hayden et al. (2007) used a visual expectancy procedure in which two stimuli are presented on a screen. The stimuli are stationary for a period of time and then one of the stimuli begins to move; during the first part of the trial, infants look longer at the stimulus that they predict will not move, presumably because they look longer at the moving stimulus during the last part of the trial. Evidence for 7- and 5-month-old infants being sensitive to differences among faces in the spacing of features comes from their successfully predicting which of two faces would move when the pair differed only in the spacing among facial features. Hayden et al. concluded that sensitivity to the spacing among facial features emerges during infancy and they suggested that the 4-year-olds in our previous study in which children were tested with their own face and with faces familiarized via a storybook may have failed because of memory load. This explanation seems unlikely in the current series of studies, however, because there was no memory load either in our distinctiveness tasks or in our matching task. It is possible that 4-year-olds' poor performance when asked to recognize their own face, faces from a storybook (Mondloch et al., 2006), or the face of a close friend (Experiment 3), is due, in part, to memory load. This is of great interest, however, given adults' exquisite memory for the spacing of facial features in highly familiar faces (Ge, Luo, Nishimura, & Lee, 2003); when shown various versions of Chairman Mao's face that differed only in the spacing of features, Mainland Chinese adults were extremely sensitive to small alterations. Their accuracy in reporting which photographs were authentic matched that of Eastern Asian adults who were asked to make same/different judgments about pairs of stimuli from the same set. Nonetheless, memory load cannot completely explain children's poor performance in this series of tasks.

We reconcile our results with those of Hayden et al. (2007) in the following way. The visual system of infants is able to detect differences in the spacing of facial features and infants are able to use that information to predict which of two stimuli is about to move. Presumably, then, the visual system of 4-year-old children is able to detect differences in the spacing of facial features, but 4-year-olds do not use that information to judge distinctiveness or identity even when the spacing of facial features deviates from a more or less average position by up to 3 *SD*—a deviation that samples most of the variability among the faces 4-year-olds will attempt to identify in the real world. We agree with McKone and Boyer (2006) that 4-year-olds' judgments of distinctiveness can be influenced by the spacing among features but only when the spacing changes are quite large. This conclusion is consistent with 4-year-olds' performing well on our criterion trials in which we presented adult female faces and with previous work showing that 8-year-olds' rate these adult faces as grotesque, but their ratings are much lower than those of adults (Mondloch et al., 2004).

Overall, our results provide additional evidence that adult-like sensitivity to small differences among faces in the spacing of features is slow to develop. The ultimate development of sensitivity to the spacing of features is dependent upon visual input during early infancy (Le Grand, Mondloch, Maurer, & Brent, 2001, 2003), but sensitivity to spacing in identity and distinctiveness tasks emerges after 4 years of age—at least when spacing changes are ≤ 2.5 *SD*. We have tested 4-year-olds' sensitivity to the spacing among facial features under a range of task conditions: judgments of distinctiveness, matching facial identity in the absence of memory demands, and recognizing the identity of personally familiar faces. Sensitivity to the spacing among facial features appears to allow adults to recognize a face from a new point of view (Mondloch, Geldart, Maurer, & Le Grand, 2003). Our results predict, therefore, that 4-year-old children may perform poorly when asked to recognize a face from a novel point of view and under other conditions that make featural cues less reliable (e.g., under poor lighting conditions). Under many conditions, however, 4-year-old children will recognize faces with great accuracy; they are very sensitive both to differences among faces in facial contour and in the appearance of individual features (e.g., the eyes and mouth; Freire & Lee, 2001; Mondloch et al., 2006)—other important cues to facial identity. The contribution of sensitivity to other cues known to influence adults' ability to recognize faces (e.g., surface cues such as color and texture; Lee & Perrett, 1997; Russell, Biederman, Nederhouser, & Sinha, 2007) should also be investigated in future studies.

References

Barton, J. J. S., Zhao, J. & Keenan, J. P. (2003). Perception of global facial geometry in the inversion effect and prosopagnosia. *Neuropsychologica*, *4*, 1703 – 1711.

Bruce, V., Campbell, R. N., Doherty-Sneddon, G., Import, A., Langton, S., McAuley, S., et al. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology*, *18*, 319 – 333.

Bruce, V., & Young, A. W. (1998). *In the eye of the beholder: The science of face perception*. Oxford, UK: Oxford University Press.

Campbell, R., & Tuck, M. (1995). Recognition of parts of famous-face photographs by children: An experimental note. *Perception*, *24*, 451 – 456.

Campbell, R., Walker, J., & Baron-Cohen, S. (1995). The development of differential use of inner and outer face features in familiar face identification. *Journal of Experimental Child Psychology*, *59*, 196 – 210.

Carver, L., Dawson, G., Panagiotides, H., Meltzoff, A. N., McPartland, J., Gray, J. et al. (2003). Age-related differences in neural correlates of face recognition during the toddler and preschool years. *Developmental Psychobiology*, *42*, 148 – 159.

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, *24*, 200 – 213.

Farkas, L. G. (1994). *Anthropometry of the head and face in medicine* (2nd ed.). New York: Elsevier.

Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year-olds: Processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology*, *80*, 347 – 371.

Ge, L., Luo, J., Nishimura, M., & Lee, K. (2003). The lasting impression of chairman Mao: Hyperfidelity of familiar-face memory. *Perception*, *32*, 601 – 614.

Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: local and relational distinctiveness effects in 7-year-olds. *Visual Cognition*, *10*, 769 – 793.

Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 995 – 1002.

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology*, *97*, 85 – 98.

Hole, G. (1994). Configural factors in the perception of unfamiliar faces. *Perception*, *23*, 65 – 74.

Lee, K. J., & Perrett, D. (1997). Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces. *Perception*, *26*, 733 – 752.

Le Grand, R., Mondloch, C., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *410*, 890 (Correction: 2001, *412*, 786).

Le Grand, R., Mondloch, C., Maurer, D., & Brent, H. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience*, *6*, 1108 – 1112 (Erratum: 2003, *6*, 1329).

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science*, *6*, 255 – 260.

McKone, E., & Boyer, B. L. (2006). Sensitivity of 4-year-olds to featural and second-order relational changes in face distinctiveness. *Journal of Experimental Child Psychology*, *94*, 134 – 162.

Mondloch, C. J., Dobson, K. S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds can't tell the difference between Steve Martin and Paul Newman: Factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology*, *89*, 159 – 181.

Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology*, *86*, 67 – 84.

Mondloch, C. J., Leis, A., & Maurer, D. (2006). Recognizing the face of Johnny, Suzy, and me: Insensitivity to the spacing among features at 4 years of age. *Child Development*, *77*, 234 – 243.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*, 553 – 566.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2003). Early visual experience is necessary for the development of some—but not all—aspects of face processing. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood* (pp. 99 – 117). New York: Nova.

Mondloch, C. J., & Maurer, D. (2003). [Four-year-old children's sensitivity to three manipulations of grotesqueness in faces: featural changes, thatcherization, and spatial alterations]. Unpublished raw data.

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science*, *14*, 618 – 622.

Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science*, *9*, 270 – 277.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research*, *47*, 157 – 165.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, Human Experimental Psychology*, *46a*, 225 – 245.

Tanaka, J.W., & Farah, M. J. (2003). Holistic face recognition. In M. Peterson & G. Rhodes (Eds.), *Analytic and holistic processes in the perception of faces, objects and scenes* (pp. 53 – 91). New York: Oxford University Press.

Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, T. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition*, *5*, 479 – 496.

Tanaka, J. W., & Sengo, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*, 583 – 592.