Developmental Science 9:5 (2006), pp 530–543

PAPER

Developmental changes in perceptions of attractiveness: a role of experience?

Philip A. Cooper,[1] Sybil S. Geldart,[1,2] Catherine J. Mondloch[3] and Daphne Maurer[1]

1. *Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada*
2. *Department of Psychology, Laurentian University, Sudbury, Ontario, Canada*
3. *Department of Psychology, Brock University, St Catharines, Ontario, Canada*

Abstract

In three experiments, we traced the development of the adult pattern of judgments of attractiveness for faces that have been altered to have internal features in low, average, or high positions. Twelve-year-olds and adults demonstrated identical patterns of results: they rated faces with features in an average location as significantly more attractive than faces with either low or high features. Although both 4-year-olds and 9-year-olds rated faces with high features as least attractive, unlike adults and 12-year-olds, they rated faces with low and average features as equally attractive. Three-year-olds with high levels of peer interaction, but not those with low levels of peer interaction, chose faces with low features as significantly more attractive than those with high-placed features, possibly as a result of their increased experience with the proportions of the faces of peers. Overall, the pattern of results is consistent with the hypothesis that experience influences perceptions of attractiveness, with the proportions of the faces participants see in their everyday lives influencing their perceptions of attractiveness.

Introduction

From a few days of age infants look preferentially towards faces previously judged attractive by adults, despite their minimal experience with faces (Slater, Bremner, Johnson, Sherwood, Hayes & Brown, 2000; Slater, Von der Schulenburg, Brown, Badenoch, Butterworth, Parsons & Samuels, 1998). Similarly, adults from different cultural and racial backgrounds give surprisingly similar ratings of facial attractiveness, even when they rate faces from a racial group with which they have had limited experience (Rhodes, Yoshikawa, Clark, Lee, McKay & Akamatsu, 2001; Perrett, Lee, Penton-Voak, Rowland, Yoshikawa, Burt, Henzi, Castles & Akamatsu, 1998; Cunningham, Roberts, Wu, Barbee & Druen, 1995; Perrett, May & Yoshikawa, 1994). The level of consistency between newborns and adults and among culturally diverse groups of adults suggests that judgments of attractiveness are stable over the course of development and insensitive to variations in experience. Some details of infants' looking preferences, however, diverge from adults' judgments of attractiveness and raise the possibility that there are developmental changes in the perception of attractiveness that might be related to developmental changes in experience.

One area of divergence is in reactions to the averageness of a face. Adults rate computer-generated average faces as more attractive than most of the individual faces used in their creation (Langlois & Roggman, 1990; Rhodes & Tremewan, 1996). The proportions, features and colouring of these average faces are the average values calculated from a large number of individual faces. Consistent with the cross-cultural research on attractiveness of individual faces, Caucasian, Japanese and Chinese adults rate average Caucasian, average Japanese, and average Chinese faces as more attractive than most of the individual faces used in their creation (Perrett *et al.*, 1998; Rhodes *et al.*, 2001). Early research suggested that 6-month-old infants show a preference for average faces because they look longer at an average face than at a face previously judged to be unattractive by adults (Rubenstein, Kalakanis & Langlois, 1999). However, unlike adults, 5- to 8-month-old infants failed to demonstrate a preference for average faces when tested with

Address for correspondence: Daphne Maurer, McMaster University, Department of Psychology, Neuroscience & Behaviour, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1; e-mail: maurer@mcmaster.ca

individual faces altered to have proportions and features more or less similar to those of an average face (Rhodes, Geddes, Jeffery, Dziurawiec & Clark, 2002): overall, they looked equally long at the two versions of each face, except for a slight tendency to look longer at the face altered to be *less* average. Although the authors suggest that the infants' looking preferences may reflect a novelty preference for bizarre-looking faces rather than an aesthetic preference, the results nevertheless suggest that the preference for average faces may change postnatally.

A second area of divergence between adults and infants became apparent in our previous research investigating the influence of the vertical location of internal facial features on adults' judgments of attractiveness and infants' looking preferences (Geldart, Maurer & Henderson, 1999). Adults and 5-month-old infants were tested with faces altered to have facial features placed in Average, Low, and High locations (the population mean ± two standard deviations) according to measurements of facial physiognomy of North Americans of Northern European descent (Farkas, 1994). Adults rated the schematic and photographed faces with features in Low and Average locations as more attractive than faces with features in the High location, and schematic faces with features in the Average location as more attractive than faces with features in the Low location (Geldart *et al.*, 1999). As with faces altered to be more or less like average (Rhodes *et al.*, 2002), infants did not show the adult-like preference for faces with features in the average location, but instead looked equally long at faces of the three types, save for a slight looking preference for faces with features in the High location over faces with features in the Low location, a pattern opposite that shown in adults' ratings of attractiveness (Geldart *et al.*, 1999). We have speculated that infants' preferences reflect their experience looking up at adults' faces. This perspective results in infants having experience with faces that have the appearance of high features by greatly accentuating the size of the chin while foreshortening the size of the forehead. One purpose of the current experiments was to trace the emergence of the adult-like pattern of attractiveness judgments for faces with varying heights of internal features. To do so, we collected judgments of attractiveness from children of a variety of ages for adult faces with their features in a Low, Average, and High location. We used six examples of female faces varying in feature height in order to collect ratings of a number of different examples while holding sex of face constant. Previous studies of adults have found that the effects of vertical location of features on adults' judgments of attractiveness are either more consistent with female faces (e.g. McArthur & Apatow, 1983–1984) or equivalent for male and female faces (Geldart *et al.*, 1999).

We hypothesized that attractiveness judgments would vary as a function of experience and thus concentrated on ages at which there are changes in the types of faces that children experience: puberty, when the faces of their peers change to have adult-like feature placement (Experiment 1), and the preschool years, when children start to have frequent contact not only with adults, but also with age-mates, whose faces have features that are relatively low, with an average location corresponding to an adult location 0.75 standard deviations below the adult mean (Experiments 2–3). These changes in experience are accompanied by changes in the perspective from which adults' faces are viewed that are related to changes in behaviour (e.g. from crawling to walking to sitting at the dining table) and in the child's height (e.g. growing after puberty to a perspective almost at eye level with adults even when standing) and that have the effect of removing the distortion caused by looking up at adult faces. Collectively, these changes may lead children no longer to find faces with high features – which have the proportions neither of adult nor children's faces viewed at or near eye level – to be attractive. Instead we hypothesize that their judgments will be influenced by the proportions of the faces of peers and adults they see near eye level and hence that the judgments will change as they gain more exposure with peers (e.g. age 3–4) and when the proportions of their peers' faces change (puberty).

Adults' preference for average faces appears to arise, at least in part, from a perceptual process favouring stimuli near the average of a category, be it a human face, a dog, a fish, or a bird (Langlois, Kalakanis, Rubenstein, Larson, Hallam & Smoot, 2000; Halberstadt & Rhodes, 2000, 2003). It can be explained by the representation of a category of stimuli in an n-dimensional space centred on a norm or prototype (Rhodes, Robbins, Jaquet, McKone, Jeffery & Clifford, 2005; Langlois *et al.*, 2000). Although familiarity can increase the attractiveness of individual faces, nevertheless, unfamiliar faces closer to the norm (i.e. more typical faces) are rated as substantially more attractive than familiar faces farther from the norm (i.e. familiarized atypical faces) (Peskin & Newell, 2004). The relevant norm is based on the individual's experience with exemplars from the category and can be altered temporarily by biased experience in the lab (e.g. Rhodes, Jeffery, Watson, Clifford & Nakayama, 2003). For example, after adaptation to 10 compressed faces, adults judge 10 other faces to be most normal and most attractive when they are slightly compressed. Note that in this example, the preferred face shape shifts only *slightly* toward the biased shape of the recently experienced faces, presumably because the years of experience seeing faces centred on a true average carry more weight. The cross-cultural similarity in finding average faces attractive, under this account, arises from the similarity of facial physiognomy across cultures (but see Rhodes, Lee, Palermo, Weiss, Yoshikawa, Clissa, Williams, Peters, Winkler & Jeffery, 2005). By 3 months of age infants possess the requisite cognitive mechanism: they treat the average of four faces with which they have been familiarized as being more familiar than the individual faces with which they have been familiarized (de Haan, Johnson, Maurer & Perrett, 2001). Thus we hypothesize that adult-like preferences will emerge around puberty, when children are exposed largely to faces with average-placed features and that younger children's preferences will change as a function of experience.

The studies were conducted with 9-year-olds, 12-year-olds, and adults (Experiment 1), 4-year-olds (Experiment 2), and 3-year-olds (Experiment 3). Because of changes in cognitive ability across this age range, it was necessary to use different methods at different ages. For each age, we adapted the methods to provide a sensitive measure of judged attractiveness for that age group. Adults, 12-year-olds, and 9-year-olds rated the attractiveness of each face using a 5-point Likert scale, but pilot testing indicated that the preschool children were unable to understand that task. For 4-year-olds, we simplified the rating scale to three levels, symbolized by three cups of increasing size, and practised use of the scale with Disney cartoon characters. For 3-year-olds, whose grasp of seriation was not sufficient to use even a 3-point rating scale (Inhelder & Piaget, 1964), we asked the child to choose which of two faces on the screen was prettier, after introducing the task with pairs of obviously ugly and pretty objects (e.g. a chewed pencil versus a shiny gold pen). The strength of this approach is that we were able to adapt our methods to provide sensitive measures of perceptions of attractiveness from participants of five different age groups. Had outwardly identical methods been administered to all participants, we would nevertheless have been using cognitively different methods across the different age groups. This is an issue in all developmental studies; in the general discussion, we consider the impact of these changes in methodology on the interpretation of the results for the preschool children.

Experiment 1

In Experiment 1, we investigated whether there are developmental changes around puberty in the effect of feature height on ratings of attractiveness of adult faces with features in an average versus low position. Such a change would be predicted by the hypothesis that age-related experience is associated with changes in what is judged most attractive. Adults experience faces with features in all possible locations, with the central tendency of their experience being a face with features in an average location. They also have some experience with children's faces, which have lower features, both as adults and during their own development. Thus, an influence of familiar proportions would lead them to rate faces with average proportions as most attractive and faces with low features as more attractive than faces with high features. Infants' predominant experience is looking up at adults' faces, a viewing angle that exaggerates chin size while minimizing the apparent size of the forehead, especially when the eyebrows are raised in the common expressive faces adults adopt when interacting with infants. Thus, infants' preferential looking and smiling towards faces with high features may reflect an influence of familiar proportions (Geldart *et al.*, 1999). If the developmental changes are related to the proportions of faces that are familiar, then we can predict that children will not begin to show the adult-like pattern of attractiveness judgments (i.e. prefer average to low- and high-placed features) until around puberty, at which point the faces of their peers mature to more adult-like proportions (Farkas, 1994). Before puberty, adult faces with features in low and average locations may both look attractive because they resemble, respectively, the proportions of the faces of peers and the faces of adults seen near eye level while seated. Based on these hypotheses, we measured 9-year-olds', 12-year-olds', and adults' judgments of the attractiveness of faces varying in the height of the internal features.

Methods

Participants

Participants were 36 9-year-old children (±3 months), 36 12-year-old children (±3 months) and 36 adults (aged 18 to 25); half the participants at each age were female. Children were recruited from names on file of mothers who had volunteered to be contacted about future studies shortly after the birth of the child and from a day camp for grade-school children that took place at McMaster University. Adults volunteered in response to posted ads and received either payment or course credit in psychology. All participants were of Caucasian descent. One additional 9-year-old was excluded because he gave the lowest rating to every stimulus.

Stimuli

The stimuli were the same as in our previous study of adults and infants (Geldart *et al.*, 1999). The 18 photographic stimuli were created using Adobe Photoshop to

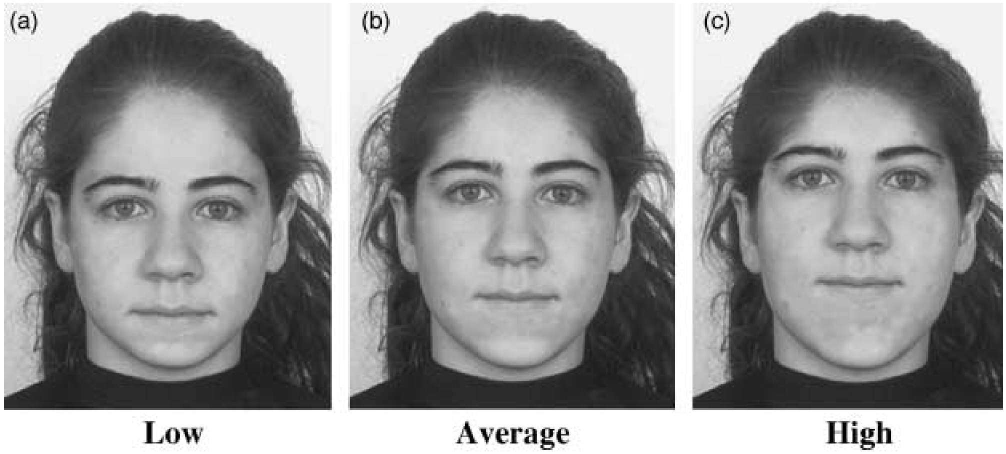

Figure 1 *An example of a female face with the features moved to (a) the low height, (b) the average height, and (c) the high height. The originals were approximately life size and in colour. The alterations were made to the faces of six adult females.*

digitally alter the colour facial photographs of six adult Caucasian females to create three different heights of internal features. The locations of the features used were calculated from published measurements of the size of the forehead and chin of the faces of adult Caucasian females (Farkas, 1994), and expressed in terms of ratios. Three vertical locations were used: Average, produced by altering faces to have proportionally average foreheads and chins (forehead-to-chin ratio of 1.4); High, produced by moving the features up two standard deviations of forehead from the average location (forehead-to-chin ratio of 0.87); and Low, produced by moving the features down two standard deviations of forehead from the average location (forehead-to-chin ratio of 2.4) (see Figure 1 for examples). The features of all stimulus faces were moved the necessary number of pixels up or down on the face to achieve the desired forehead-to-chin ratio.

The stimuli were projected one at a time onto a rear projection screen from a Kodak Ektagraphic III A carousel slide projector located 75 cm behind the screen, and were roughly the size of real faces (12 cm wide by 16 cm high or 9 by 12 visual degrees when viewed from 75 cm). The only other light in the room came from a nightlight behind the participant. Eighteen orders were created such that one example of each location of features, and the faces of three different females, were presented within the first three trials and the remaining 15 stimuli were in a pseudo-randomized order, with no two versions of the same female face being presented consecutively. Between trials, a non-face stimulus was presented, so as to minimize adaptation effects between faces and so as to match the conditions used in a previous study of 5-month-olds and adults (Geldart *et al.*, 1999). Each order was presented to a single male and single female participant from each age group.

Procedure

After the procedures were explained, we obtained written consent from a parent of the children or from the adult participant. The participant sat approximately 75 cm in front of the screen. The participant was asked to rate each stimulus' attractiveness on a 5-point Likert scale, with 1 representing very unattractive (very ugly), 3 representing average (neither pretty nor ugly) and 5 representing very attractive (very pretty). For children, the experimenter first explained the scale, and then had the child explain it back to him to check the child's understanding. Adults and 12-year-olds indicated their choice for each face by circling the appropriate number on a piece of paper. To aid understanding of the Likert scale, 9-year-olds were given a series of five increasingly large cups numbered from 1 to 5 to represent the ordinal nature of the attractiveness scale. They were asked to point to the cup that represented the attractiveness of each face, and the experimenter recorded their responses. The McMaster Research Ethics Board approved the protocol and procedures of all three experiments.

Results

The 18 attractiveness ratings supplied by each participant were collapsed into three means representing the mean attractiveness ratings for faces with features in the low, average, and high locations. The mean scores for each age group, broken down by sex of participant, are shown in Table 1. To allow comparisons across age in the influence of feature height on relative judgments of attractiveness, each participant's three mean ratings were transformed into *z*-scores based on the mean overall rating for all 18 faces and its standard deviation across all

Table 1 *Mean ratings and standardized ratings (expressed as z-scores) of attractiveness for faces with their features at low, average and high positions. The upper panel shows data for all 18 trials for 9-year-olds and 12-year-olds and for adults (Experiment 1) overall and for each sex of participant. The lower panel shows the data for the first three trials and all 18 trials for 4-year-olds (Experiment 2). The means from Experiment 1 are based on a 5-point scale; those from Experiment 2 are based on a 3-point scale*

Participants		Mean attractiveness ratings						Standardized attractiveness ratings					
		Low		Average		High		Low		Average		High	
Age	Sex	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Adults†	M	2.56	0.50	3.03	0.43	2.03	0.45	−0.020	0.51	0.448	0.43	−0.564	0.46
	F	2.81	0.60	2.97	0.51	2.11	0.60	0.223	0.61	0.392	0.51	−0.048	0.60
12-year-olds†	M	2.32	0.41	2.55	0.31	2.09	0.46	−0.065	0.21	0.207	0.37	−0.327	0.54
	F	2.41	0.53	2.73	0.49	2.13	0.50	0.044	0.62	0.425	0.58	−0.283	0.58
9-year-olds†	M	2.42	0.84	2.45	0.64	2.08	0.77	−0.101	0.75	−0.068	0.58	−0.400	0.69
	F	2.82	0.43	2.86	0.39	2.54	0.57	0.265	0.39	0.298	0.35	0.007	0.51
4-year-olds*	M	2.15	0.34	2.28	0.41	2.06	0.44	0.077	0.40	0.231	0.48	−0.033	0.52
(18 trials)	F	2.15	0.57	2.06	0.62	1.82	0.63	0.077	0.68	−0.033	0.74	−0.319	0.75
4-year-olds*	M	1.89	0.93	2.44	0.73	1.56	0.88	−0.105	1.05	0.526	0.83	−0.484	1.00
(1st 3 trials)	F	2.56	0.73	2.00	0.87	1.44	0.73	0.653	0.83	0.021	0.98	−0.611	0.83

† Scores in each cell are based on ratings of attractiveness on a 5-point Likert scale from 18 participants.
* Scores in each cell are based on ratings of attractiveness on a 3-point Likert scale from 9 participants..

participants of the same age. This approach eliminated the systematic differences between ages in mean rating of attractiveness while maintaining the pattern of ratings within each age group and allowed us to compare data based on a 5-point scale (Experiment 1) and a 3-point scale (Experiment 2).

The data were analyzed using a three-way repeated-measures ANOVA with stimulus set, sex of participant, and age of participant (9-year-olds, 12-year-olds and adults) as factors. The main effects of sex ($F_{(1,102)} = 4.45$, $p < .05$) and stimulus set ($F_{(2,204)} = 83.9$, $p < .01$) were both significant, as was the interaction of age and stimulus set ($F_{(4,204)} = 6.04$, $p < .01$); all other main effects and interactions were non-significant. Analysis of simple effects revealed that the main effect of stimulus set was significant for all three age groups (*p*s < .01). As illustrated in Figure 2, Fisher's post-hoc tests demonstrated that although the 9-year-olds, 12-year-olds and adults all rated faces with features at low and average locations as being significantly more attractive than those with high features (*p*s < .01), only the 12-year-olds and adults (both *p*s < .01), and not the 9-year-olds ($p > .5$), rated faces with features in an average location as being more attractive than faces with low features. The significant main effect of sex was driven by the higher overall ratings of attractiveness given by female participants (see Table 1).

Discussion

At all ages, participants rated the faces with high features as much less attractive than faces with their features in average or low positions. This result replicates previous findings for adults asked to judge the relative attractiveness of paired presentations of the same photographs (Geldart *et al.*, 1999) and adults asked to rate the attractiveness of drawings of infants' faces (Maier, Holmes, Slaymaker & Reich, 1984; Hildebrandt & Fitzgerald, 1979; Sternglanz, Gray & Murakami, 1977). Faces with high features may be perceived as least attractive because they diverge most from the faces experienced regularly by either children or adults.

Adults and 12-year-olds also rated faces with features in an average location as slightly but significantly more attractive than faces with low features. The results for adults replicate those found previously for pairs of schematic drawings of male and female faces rated by Caucasian (Geldart *et al.*, 1999) and Korean adults (McArthur & Berry, 1987). In contrast, 9-year-olds rated faces with features in average and low positions as equally attractive. These findings suggest that the adult-like pattern for judgments of the attractiveness of faces with features at different heights emerges between 9 and 12 years of age.

The pattern of results matches an experience-based explanation in which novel faces are compared to a norm that is constantly updated as faces are experienced, and in which, all else being equal, faces near the norm are rated as more attractive than faces farther from the norm. At every age, the norm should have features located near the central tendency of the individual's experience; just as adaptation in the lab appears to alter the norm (Rhodes *et al.*, 2003; Webster, Kaping, Mizokami &

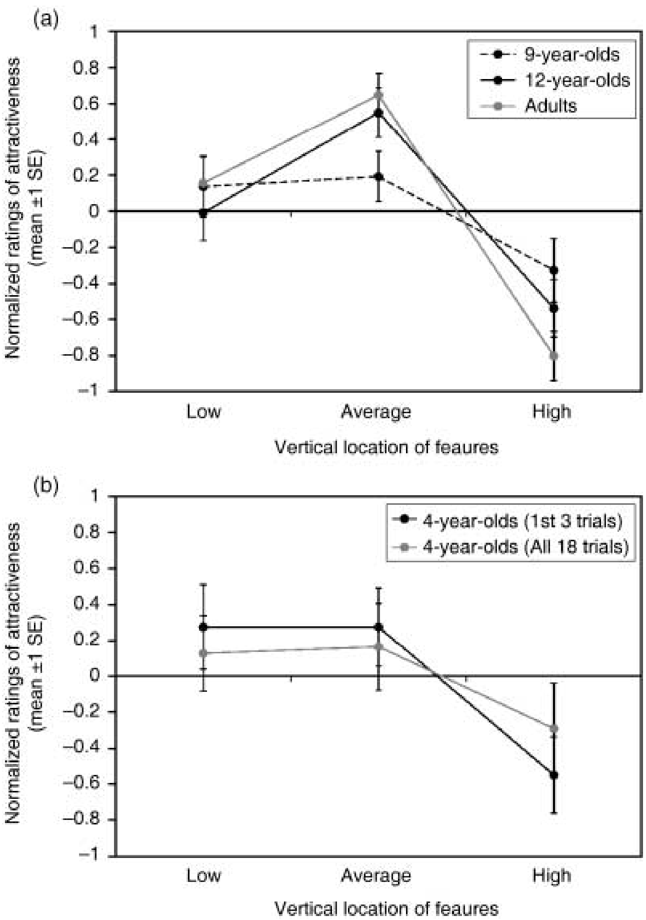

Figure 2 *Mean ratings of attractiveness (+1 SE) expressed as z-scores for faces with their features at low, average and high positions. Top Panel: Shown are data for all 18 trials for 9-year-olds and 12-year-olds and for adults (Experiment 1). Bottom Panel: Shown are data for the first three trials and all 18 trials for 4-year-olds (Experiment 2).*

Duhamel, 2004), so should changes in the proportions of faces encountered every day. Just such a change occurs between 9 and 12 years of age, as the features on their peers' faces move to the adult location (Farkas, 1994). At the same time the child's increased height means that adult faces are viewed at or near eye level not only when sitting but also when standing. Perhaps as a result of either or both of these changes, they start to rate adults' faces with features in an average location as more attractive than faces with low features.

Although the change between 9 and 12 years of age is consistent with an experience-based account, the explanation is based on a correlation and there are other possible explanations based on the many changes that happen at puberty. One change that appears particularly relevant is the improvement in children's ability to use small differences in the spacing among features (their second-order relations) to recognize individual identity between 10 years of age and adulthood (Mondloch, Le Grand & Maurer, 2002). Those improvements may also reflect the development of a better differentiated face space in which the dimensions better distinguish among adults' faces and are centred on a more veridical norm. Thus, the same changes in experience may lead to shifts in judgments of attractiveness and to increased sensitivity to differences between individuals in the spacing of their features. Those changes might be enhanced, in turn, by hormonal changes at puberty that may trigger enhanced interest in faces and their attractiveness. However, we cannot rule out the possibility that the hormonal changes themselves induce the changes in judgments of attractiveness, independent of any changes in norm-based coding.

The main effect of sex did not interact with age or the height of features in the stimuli. It reflects the fact that female participants gave more generous ratings of attractiveness than male participants. The reason may be different for the children and the adults. Boys near puberty may be more reluctant than girls to report positive reactions about adults. For adults, judgments of attractiveness may be influenced by whether they are for same- or opposite-sex faces (DeBruine, 2004). In any event, feature height had the same effect on perceptions of attractiveness regardless of sex of participant.

Experiment 2

In Experiment 2, we asked 4-year-olds to rate the prettiness of the adult faces with features at low, average, and high locations, using a simplified version of the 5-point Likert scale used in Experiment 1. Previous studies indicate that children as young as 3 can give verbal aesthetic judgments (Dion, 1973; Langlois & Stephan, 1977; Martin, Eisenbud & Rose, 1995) and that they already share with adults the 'beauty is good' stereotype, attributing positive qualities (e.g. friendliness, intelligence) to individuals they judge more attractive and negative qualities to individuals they judge less attractive (e.g. Dion, 1973; Dion, Berscheid & Walster, 1972). Their global aesthetic judgments also are similar to those of adults: for example, when asked to point to the face in a pair that they judge as more pretty, 3- to 6-year-olds point to the face previously judged by adults as more attractive (Dion, 1973).

Here we examined how the height of internal facial features influences the aesthetic judgments of 4-year-olds. If aesthetic judgments are affected by experience, we predicted that 4-year-olds might show a pattern similar to 9-year-olds, because their exposure to facial proportions is similarly biased. While seated, 4-year-olds see adults' faces from slightly below eye level – a perspective that makes the features look near their actual location. Exposure to adults' faces at nearly eye level may make faces with features near the average adult position seem attractive. (Note that unlike the case of an infant looking up at an adult's face, the 4-year-olds' perspective would in no way approximate the high placement of features used here.) On the other hand, 4-year-olds also have many face-to-face interactions at eye level with same-aged peers, who, like 9-year-olds, have faces with features located lower than do adults (Farkas, 1994; Enlow, 1982; see also Alley, 1981). Exposure to peers' faces with low features may cause faces with those proportions to seem normal to 4-year-olds because of their resemblance to a norm. Just as adults find faces – and other stimuli – that resemble their putative prototype to be attractive (Langlois & Roggman, 1990; Halberstadt & Rhodes, 2003), 4-year-olds may find the faces with low features, which are similar to those of their peers, and the faces with features in average location, which are similar to those of adults, to be prettier than the faces with High features. Like 9-year-olds and unlike adults, we did not expect 4-year-olds to find faces with features in an average location to be prettier than faces with low features, because faces with low features resemble the faces of their peers. The procedure was similar to Experiment 1 except that we simplified the Likert scale to three levels and preceded it by a training phase with Disney characters. These changes were made to account for 4-year-olds' limited understanding of seriation (e.g. Inhelder & Piaget, 1964).

Methods

Participants

Participants were 18 4-year-olds (±3 months), half of whom were female. Participants were recruited from names on file of mothers who had agreed at birth to be contacted about their child's participation in research.

Stimuli and procedure

The stimuli and apparatus were the same as in Experiment 1, except that the door to the testing room was left open so that a parent could stand in the doorway. This arrangement increased participant cooperation, decreased parental intrusions and allowed children a clear view of both the stimuli and the 'rating cups' placed before them. Parents were asked not to look at the stimulus faces. Overhead lights were turned off.

After informed consent was received from a parent, the child was seated approximately 75 cm in front of the screen. The child was asked to rate each face's attractiveness on a 3-point Likert scale, with 1 representing unattractive (*not at all pretty*), 2 representing average (*so-so/okay*) and 3 representing attractive (*very pretty*). A series of three increasingly large cups in front of the child represented the ordinal nature of the attractiveness scale. Children were told they would see a face and be asked if it was *very pretty*, *not at all pretty* or *so-so/okay*. They were also told that sometimes they would see a pattern instead of a face and they did not have to say anything about it. Prior to beginning the experiment proper, children were given practice using the cup rating scale by asking them to judge the attractiveness of four to six cartoon characters taken from Disney picture storybooks (e.g. Cinderella, Snow White, Wicked Witches, Winnie-the-Pooh), each presented as a black-and-white drawing with both face and body.

Results

The data were collapsed as in Experiment 1 except that the means were based on a 3-point scale instead of a 5-point scale. Table 1 shows the results for the raw data based on all 18 trials. To allow comparison with Experiment 1, the data were transformed into *z*-scores based on the mean and standard deviation for all 18 faces for the 4-year-olds (see Figure 2). We also analyzed the data from just the first three trials for each child, which contained one example of each of the three feature heights from three different faces, with different exemplars chosen for each child (see procedure of Experiment 1). We did so because pilot work indicated that 4-year-olds respond more systematically toward the beginning of the procedure, a pattern confirmed in the current findings: three of the children gave an identical rating for each of the faces shown during the last six trials.

The *z*-scores based on all 18 trials and on the first three trials were analyzed with a repeated-measures ANOVA with stimulus set (Average, Low and High) as a within-subject factor and sex of participant as a between-subjects factor. For all 18 trials, the main effect of stimulus set was marginally significant ($F_{(2,32)} = 3.20$, $p = .055$) but the main effect of sex ($F_{(1,16)} = 0.53$, $p = .47$) and interaction between stimulus set and sex of participant ($F_{(2,32)} = 0.86$, $p = .43$) were both non-significant. For the first three trials, the main effect of stimulus set was significant ($F_{(2,34)} = 5.39$, $p < .001$), the main effect of sex of participant and interaction between stimulus set and sex of participant were both non-significant (ps $>$ .05). Fisher's post-hoc tests for the first three trials demonstrate that the 4-year-olds, like adults and older children, rated faces with features at the low and average locations as significantly more attractive than those with high features (both ps < .01). However, unlike adults and 12-year-olds, but like 9-year-olds, they did not rate faces with features in an average location as more attractive than faces with low features ($p > .10$). As illustrated in the bottom panel of Figure 2, the pattern was similar for all 18 trials and for the first three trials.

Discussion

The results for 4-year-olds are identical to those found for 9-year-olds in Experiment 1: like adults, they rated faces with high features as less attractive than faces with features in an average or low position and, unlike adults and 12-year-olds, they did not differentiate between faces with features in an average and low position (see Figure 2). On the face of it, it is surprising not to find any change between 4 and 9 years of age in sensitivity to the spatial layout of faces. Between 4 and 8 years of age, there are massive improvements in memory for individual faces, including increased sensitivity to the small differences among individuals in the spacing among internal features, a cue called second-order relations that is believed to underlie adults' expertise at face recognition (Mondloch *et al.*, 2002; Maurer, Le Grand & Mondloch, 2002; Freire & Lee, 2001): 4-year-olds are at chance in detecting a change in identity based on the spacing of features in their own and peers' faces when the changes stay within natural limits (Mondloch, Leis & Maurer, 2006; but see Bhatt, Bertin, Hayden & Reed, 2005, and Pellicano, Rhodes & Peters, 2006, for evidence of detection under different circumstances), while 8-year-olds' accuracy is well above chance on similar tasks (Mondloch *et al.*, 2002; Mondloch, Dobson, Parson & Maurer, 2004). Our results suggest that those changes in processing of facial identity may not be relevant to judgments of attractiveness. What may matter for attractiveness is whether the face looks 'right' because its proportions match the central tendency of experience. For both 4-year-olds and 9-year-olds, the relevant experience is of adults with features near the average adult location and of peers with features near the low adult location. An additional influence may be the proportions of faces seen in children's books and cartoons, which typically have low features. Thus, as predicted by an experience-based model, 4-year-olds rated as equally attractive faces with low features – like those of their peers – and faces with features in an average location – like those of the typical adult face. They rated faces with high features – which deviate greatly from the low features they observe in their peers' faces – as least attractive.

This is the first study to obtain Likert ratings of attractiveness from children as young as 4. Although the results for all 18 trials were only marginally significant, they parallel the results from the first three trials. In our design, 4-year-olds initially practiced rating attractiveness of pictures of cartoon characters, many of which were familiar and humorous to the children. However, the Disney characters used in the warm-up may have had the unintended effect of calling attention to faces with particular locations of features and affected the child's subsequent judgments of attractiveness. A better procedure would be to design a warm-up in which the stimuli do not have faces at all. Nevertheless, we note that any such influence of the cartoon characters mimics an influence prevalent in the child's home environment, and that some of the cartoon characters had high features (The Wicked Witch), while others had low features (Cinderella).

Experiment 3

In Experiment 3, we developed a method to examine 3-year-olds' ratings of the 'prettiness' of faces varying in the height of their internal features. Our previous study (Geldart *et al.*, 1999) showed that 5-month-olds look and smile most at faces with high features, the faces which adults rate least attractive. This may be related to the foreshortened perspective from looking up at adults' faces – which emphasizes the chin and shrinks the forehead. Experiment 2 suggests there is a change sometime between 5 months and 4 years of age. In Experiment 3, we examined whether it has occurred by 3 years of age. We expected that there might be a change by 3 years of age because of physical changes (e.g. sitting upright, standing, becoming taller, sitting in a booster seat at the dining table), greater mobility (crawling, walking), and face-to-face contact with peers in preschool and playgroups. Even so, like 4-year-olds, 3-year-olds' relevant face-to-face experiences at eye level will be biased toward faces with their features at a low height because the features of children's faces are lower on the face than is true in adults (Enlow, 1982; see also Alley, 1981; Farkas, 1994). They are also typically low on the faces of dolls and stuffed animals. Thus, we suspected that by 3 years of age, children may perceive faces with their features at a low height, yielding a larger forehead and a smaller chin, as attractive or 'pretty' because such faces resemble the central tendency of the faces to which they have been exposed at eye level. We also expected that, like 4-year-olds, they might see adult faces with features in the average location as more attractive than faces with features in the high location because of the increasing exposure to adult faces at almost eye level – a perspective that could make faces with features in an average location, or very slightly higher, appear attractive.

Testing 3-year-olds also provided an opportunity to look for individual differences. By age 4, most children are in either a preschool or junior kindergarten programme and, as a result, interacting with multiple children on a daily basis. In contrast, there is wide variability among 3-year-olds in the amount of time spent with peers. Some 3-year-olds are in preschool or junior kindergarten, but many others are home with an adult caregiver whose face has features in an average location, which they see from variable perspectives: below (when both are standing) and at nearly eye level (when both are seated and the child is in a booster seat). If experience drives aesthetic preferences because it changes the average face, then 3-year-olds with more peer interactions should prefer low and, possibly, average features whereas those who spend less time with peers should not.

Because 3-year-olds do not yet grasp the concept of seriation (Inhelder & Piaget, 1964), we used a different method to assess their judgments of attractiveness, adapted from a study of preschoolers' ratings of the prettiness of adults' faces (Dion, 1973). The 3-year-olds viewed pairs of colour photographs of female faces with their features at two different heights, and for each pair, pointed to the face they perceived as more pretty. Based on pilot testing on the maximum number of trials during which most children remained attentive, each child received only three trials, one with each possible pairing of feature height, and the specific pairings were counterbalanced across children. To give them practice with the task, children were first asked to judge the relative prettiness of non-face objects (e.g. dolls' dresses, balls) which differed in colouring, style, and in prettiness as judged by 3-year-olds and adults in pilot testing.

Methods

Participants

The participants were 24 (12 male) 3-year-olds (±2 months) recruited in response to announcements within the department and from a pool of mothers who had volunteered their babies at birth for later study. Three additional children were tested but excluded from analyses because they failed the screening pre-test (see Procedure). All subjects were Caucasian.

Stimuli and apparatus

The face stimuli were the same as in Experiments 1 and 2. The six female faces were divided into two sets of three, with half of the subjects viewing one set and the other half viewing the second set. Within each set, we formed pairs of faces differing only in feature height, with their inner edges 5.5 cm apart. Each child saw three types of pairings (i.e. Low with High, Low with Average, and Average with High) in a counterbalanced order, with each pairing illustrated with a different face. The faces used to illustrate each type of pairing and their order were counterbalanced across subjects. The trials were separated by the same non-facial images used in Experiments 1 and 2 and in our previous study with 5-month-olds.

There were four pairs of non-face objects designed to give children practice with the task, and to assess whether they understood the instructions. The objects within each pair were of the same class: (1) a blue-coloured ball paired with a red ball decorated with coloured, plastic gems forming a pattern of vertical stripes; (2) a doll's dress made from green cotton fabric and containing a simple design paired with a doll's ball gown made of light blue satin, white lace, and sequins; (3) a black comb paired with a shiny-gold comb with a handle covered with coloured jewels; and, (4) an old, chewed yellow pencil paired with a new, gold-trimmed blue pen. These pairs were formed on the assumption that one member would be perceived as more pretty, an assumption which was later confirmed in pilot testing with 3-year-olds and adults. The order of presentation of the four pairs of non-face objects, and which item appeared on the left side, was randomized across subjects.

Each child sat in a child's chair in front of a child's table containing a small wooden frame that held the projection screen. A Kodak Carousel projector, from which the faces were rear-projected, was placed on a platform behind the screen. The tester advanced the slides using a remote control device.

For the pre-test with non-face objects, the child's chair was turned so that he/she faced the tester rather than the screen. The four pairs of objects were held in a bag and then brought out in front of the child one pair at a time. The tester held both objects by the tips of her fingers so that they were about 40 cm apart and 50 cm from the child's eyes.

Procedure

After the procedures were explained, a parent signed an informed consent form. The procedure began with the practice task in which the child viewed four pairs of non-face objects, and for each pair, was asked to point to the object perceived as more pretty. To be included in the analyses a subject was required to choose, for at least three of the four pairings, the member of each pairing that had been preferred by most adults and children in pilot tests. Three children failed to meet this criterion (see Participants).

For the main task, the child faced the projection screen and the experimenter faced the child so that she was unaware of the faces shown on each trial. When the child appeared ready, the tester advanced the slide projector to the first pair of faces, and said: 'These are the faces of two women. Show me the face you think is more pretty by pointing to her.' Following the child's pointing response, the tester coded it on the scoring sheet, advanced the projector to present a non-facial pattern, and said, 'This is not a face, is it? Let's go on and look at more faces.' She then presented the next pair of faces. This procedure continued until the child provided ratings for all three types of face pairings. No child took more than 5 minutes to complete the ratings.

After the child was tested, the tester asked the parent questions about his/her child's social contacts and recorded the responses on a scoring sheet. The parent was asked to indicate the number and ages of the child's siblings, and the frequency (e.g. hours, days per week) with which the child participated in social activities outside of the home, including school (e.g. Bible school, nursery school, daycare), group lessons (e.g. swimming, dance), and playgroups. The parent was also asked to list the approximate age(s) and numbers of children who participated in each activity as well as the ages of any siblings. The parent was told about the purpose of the questions and our hypothesis after completing the questionnaire.

We rank-ordered the children on the basis of the quantity of their interaction with other children and then split the sample at the median and classified the groups as either 'more exposed' or 'less exposed' to peers. A second observer, who independently rank-ordered the subjects, agreed with this classification for 21 of the 24 subjects. The two rank orderings were significantly correlated (Spearman correlation, $p < .05$). We then reviewed the ranks together and came to an agreement about the discrepancies.

Results

The data consisted of the number of 3-year-olds choosing each member of a pair, for the three different pairings of feature height. For each pairing, we compared the distribution of choices to a chance value of 50:50 using a two-tailed binomial test. We also used a chi-square analysis to test whether the distribution differed for the children with more versus less exposure to peers.

The binomial tests revealed that 3-year-olds' ratings of prettiness were distributed randomly between the two versions of each face, when judging pairs of faces

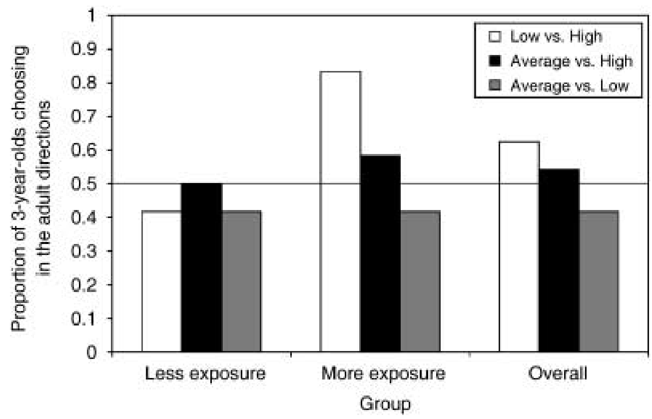

Figure 3 *The proportion of 3-year-olds who chose as prettier the member of a pair matching adults' preferences: the low member from the low/high pairing; the average member from the low/average pairing, and the average member from the average/high pairing. Shown are data from the high and low peer interaction groups, both separately and combined (Experiment 3).*

with Low and High features, Low and Average features, and with Average and High features (all $ps > .10$) (see Figure 3). However, when judging faces with Low versus High features, children in the group with more peer interaction rated the faces with Low features as prettier than the faces with High features, while for those in the group with less exposure to peers there was no effect of the height of the internal features ($\chi^2(1) = 4.44$, $p = .035$). Ten of the 12 children in the high exposure group chose the Low features as prettier, compared to five of the 12 children in the low exposure group. There was no significant effect of peer exposure for the remaining pairings (see Table 1).

Discussion

In general, the height of the internal features did not influence 3-year-olds' ratings of faces' prettiness. It is unlikely that the results are an artifact of the task being too difficult for the 3-year-olds. They merely had to point to the face of a pair that they perceived as more pretty, and all of those included in the final study had performed well during the practice trials with non-face objects. Moreover, Dion (1973) has shown that children as young as 3 can make relative judgments of attractiveness when faces differ greatly in attractiveness as judged by adults. It is possible that the children in this experiment did not look at the facial stimuli as carefully as the objects in the practice trials and/or had trouble making decisions about two similar-looking faces differing only in the features' height. Like any negative result with young children, the results of Experiment 3 may underestimate the impact of variations in feature height on 3-year-olds' aesthetic judgments. Note, however, that the method was sensitive enough to reveal that 3-year-olds with more exposure to peers rate faces with their features at a low height – as is true of their peers' faces – as significantly prettier than faces with their features at a high height. This finding shows that 3-year-olds are capable of making aesthetic judgments based on changes in feature height, that our method was sensitive enough to reveal the effect, and that their judgments in this respect are like those observed previously in adults and older children (Experiments 1 and 2), and unlike those revealed by looking time and smiling in 5-month-olds (Geldart *et al.*, 1999).

The pattern of results from the two groups of children supports the hypothesis that exposure to peers' faces plays a role in the development of aesthetic preferences for faces with their features at a particular height. Frequent exposure to the young-looking faces of other children (see Enlow, 1982, for facial characteristics of young children) may cause 3-year-olds to perceive faces with low features, including a larger forehead and a smaller chin, as familiar and appealing. By age 4, all children may have had enough exposure to the faces of peers to have developed this preference, as was found in Experiment 2. With increasing exposure to peers and to adults' faces from nearly eye level after infancy, the initial infant preference for adult faces with high features (matching how they look when seen from below) may be overcome and then reversed.

Both groups of 3-year-olds differed from the 4-year-olds in Experiment 2 in choosing the faces with features in an average and high position as equally pretty. The developmental difference may be related to the need for additional experience looking at adults' faces near eye level before faces with average proportions come to look 'right' and more attractive. Between 3 and 4 years of age, more frequent eye-level interactions with a wider variety of adults may lead children to develop a norm closer to the typical proportions of adults' faces (i.e. faces with an average sized forehead and chin). That experience may also help to offset the impact of experience with young-looking faces of play dolls and young peers (see Langlois & Roggman, 1990), making low and average placed features equally familiar and, thus, equally attractive. Alternatively, the changes we observed between ages 3 and 4 might be related to differential sensitivity of the methods used in Experiments 2 and 3 or to changes in general face processing skills that may allow 4-year-olds, the youngest age group that has been shown to be sensitive to natural variations in the spacing of facial features (Pellicano *et al.*, 2006), to pick up more information from a face and to more easily discriminate among the three vertical locations of features (see Brace, Hole, Kemp, Pike, Van Duuren & Norgate, 2001). Instead they might be related to more general cognitive changes that are not likely to be influenced by experience with specific faces, such as understanding that stimuli can differ on more than one dimension (Inhelder & Piaget, 1964). One could differentiate these hypotheses by examining the effect of exposure in the lab to numerous examples of the faces of peers and/or of adults from an en face perspective on 3-year-olds' aesthetic ratings tested with the same method used in Experiment 3.

General discussion

The results suggest that, although on the surface judgments of attractiveness based on the height of the internal features change during development, the underlying mechanism does not. At all ages tested, infants, children, and adults rate faces with the most familiar proportions as most attractive. Four-year-olds (Experiment 2) and 9-year-olds (Experiment 1) rated faces with features in low and average locations as equally attractive, despite showing sensitivity to variations in feature height by, like adults, rating faces with high features as least attractive. Three-year-olds were indifferent to feature height, except for finding faces with low features prettier than faces with high features if they were in the group that had had more contact with peers (Experiment 3). In contrast to the younger children, 12-year-olds, like adults (Experiment 1), rated faces with features in an average location as more attractive than faces with low features. Combined with evidence from infants, the results support an experiential influence on developmental changes in judgments of attractiveness, with participants finding faces with the proportions matching the central tendency of their previous experience most attractive.

As predicted, children developed the adult pattern of responses at 12, the age when their experience with adult facial proportions increases dramatically. Children aged 4 and 9 showed a different pattern: they rated faces close to the proportions of their peers as equal in attractiveness to faces with the average adult proportions, both of which they rated as more attractive than faces with high features. The beginnings of that pattern were evident in 3-year-olds with extensive contact with peers. The pattern of results across experiments is consistent with the hypothesis that experience plays an important role in the development of standards of attractiveness. That explanation also fits with the looking and smiling preferences of 5-month-olds for the same set of faces varying in feature height (Geldart *et al.*, 1999).

Evidence that experience plays a role in the development of judgments of attractiveness is rare in the literature. This may reflect the fact that studies beyond the infant period are rare, no doubt because of demonstrations that infants look longer at faces rated attractive by adults than at faces rated unattractive (Slater *et al.*, 2000; Rubenstein *et al.*, 1999; Slater *et al.*, 1998; Samuels, Butterworth, Roberts, Graupner & Hole, 1994). However, these studies have demonstrated only that infants are sensitive to extreme differences in overall attractiveness; they have not tested whether infants are sensitive to more subtle differences nor delineated which cues infants respond to. Like two previous studies of infants (Rhodes *et al.*, 2002; Geldart *et al.*, 1999), the results of the current experiments demonstrate that there are also changes in the perception of attractiveness after infancy, and even until puberty.

Although our results indicate that there are developmental changes in attractiveness judgments that are consistent with an experiential hypothesis, alternative explanations exist. The most obvious are the concurrent developmental changes in sensitivity to cues to facial identity (reviewed in Brace *et al.*, 2001; Mondloch, Geldart, Maurer & Le Grand, 2003; Mondloch *et al.*, 2002), and the concurrent hormonal changes between ages 9 and 12 that might affect judgments of attractiveness. Although we cannot rule out a role for hormonal changes, we note that the developmental changes in sensitivity to facial identity may reflect the development of a better-differentiated face space centred on an increasingly accurate prototype – exactly the mechanism we have speculated alters judgments of attractiveness. Another possibility is that the methods we used are better at revealing the aesthetic preferences of older children and that there is in fact no developmental change. Although we cannot rule out this possibility – as is true in all developmental research – we note that we succeeded in adapting the instructions and methods to each age group so as to reveal *an* effect of feature height at every age and so as to reveal a consistent pattern of developmental change (compare Figures 2 and 3).

The experiential hypothesis developed in this paper is consistent with evidence that parental age, eye colour and hair colour are correlated with later mate choice (Little, Penton-Voak, Burt & Perrett, 2003; Perrett, Penton-Voak, Little, Tidderman, Burt, Schmidt, Oxley, Kinloch & Barrett, 2002). It is also consistent with the finding that adults rate faces manipulated to resemble their own face as being both more average and more attractive than faces altered to resemble unfamiliar adults, possibly as a result of frequent experience with their own face and those of close relatives (DeBruine, 2002, 2004). Similarly, brief exposure to a biased sample of faces in the lab can alter adults' subsequent judgments of faces, including their sex, ethnicity, facial expression, normality and attractiveness (Webster *et al.*, 2004; Rhodes, Jeffery *et al.*, 2003; Webster & MacLin, 1999). Several authors have speculated that our everyday encounters with faces act in a similar way to adjust the norm of our face representations to take account of changes in our experience (Leopold & Bondar, 2005; MacLin & Webster, 2001; Rhodes *et al.*, 2003; Webster *et al.*, 2004). The resulting changes in judged attractiveness appear to act through alterations in the overall norm, rather than in the representation of the particular faces seen (e.g. Peskin & Newell, 2004; Rhodes, Halberstadt, Jeffery, & Palermo, in press). A similar process may underlie the developmental changes we documented in the effect of feature height on judgments of attractiveness. Furthermore the finding that average dogs, wristwatches and birds, created in the same manner as the average faces, are also rated as more attractive than the individual items used in their creation (Halberstadt & Rhodes, 2000, 2003), suggests that the influence of experience on attractiveness may not be limited to faces.

The experiential explanation of our results, if correct, suggests that children might represent the faces of adults and of their peers in the same *n*-dimensional face space, with judgments of the attractiveness of adults' faces being influenced by the proportions they experience in faces of both adults and peers. Future studies could test this inference directly by contrasting the influence of feature height on children's judgments of the faces of their peers versus adults. Future studies might also examine whether there are predictable differences in the aesthetic judgments of adults with experience viewing faces from a perspective different from eye level – caused by being substantially taller or shorter than average adult height.

Acknowledgement

This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant to Daphne Maurer.

References

Alley, T.R. (1981). Head shape and the perception of cuteness. *Developmental Psychology*, **17**, 650–654.

Bhatt, R.S., Bertin, E., Hayden, A., & Reed, A. (2005). Face processing in infancy: developmental changes in the use of different kinds of relational information. *Child Development*, **76**, 169–181.

Brace, N.A., Hole, G.J., Kemp, R.I., Pike, G.E., Van Duuren, M., & Norgate, L. (2001). Developmental changes in the effect of inversion: using a picture book to investigate face recognition. *Perception*, **30**, 85–94.

Cunningham, M.R., Roberts, A.R., Wu, C.H., Barbee, A.P., & Druen, P.B. (1995). 'Their ideas of beauty are, on the whole, the same as ours': consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology*, **68**, 261–279.

DeBruine, L.M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society of London, B*, **269**, 1307–1312.

DeBruine, L.M. (2004). Facial resemblance increases the attractiveness of same-sex faces more than other-sex faces. *Proceedings of the Royal Society of London, B*, **271**, 2085–2090.

De Haan, M., Johnson, M., Maurer, D., & Perrett, D. (2001). Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. *Cognitive Development*, **16**, 659–678.

Dion, K.K. (1973). Young children's stereotyping of facial attractiveness. *Developmental Psychology*, **9**, 183–188.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, **24**, 285–290.

Enlow, D.H. (1982). *Facial growth*. Philadelphia, PA: Saunders Publisher.

Farkas, L.G. (1994). *Anthropometry of the head and face in medicine* (2nd edn.). 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.

Geldart, S., Maurer, D., & Henderson, H. (1999). Effects of the height of internal features of faces on adults' aesthetic ratings and 5-month-olds' looking times. *Perception*, **28**, 839–850.

Halberstadt, J., & Rhodes, G. (2000). The attractiveness of non-face averages: implications for an evolutionary explanation of attractiveness of average faces. *Psychological Science*, **11**, 285–289.

Halberstadt, J., & Rhodes, G. (2003). It's not just average faces that are attractive: computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychonomic Bulletin and Review*, **10**, 149–156.

Hilderbrandt, K.A., & Fitzgerald, H.E. (1979). Facial feature determinants of perceived infant attractiveness. *Infant Behavior and Development*, **2**, 329–339.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child: Classification and seriation*. London: Routledge and Kegan Paul.

Langlois, J.H., Kalakanis, L.E., Rubenstein, A.J., Larson, A.D., Hallam, M.J., & Smoot, M.T. (2000). Maxims and myths of beauty: a meta-analytic and theoretical review. *Psychological Bulletin*, **126**, 390–423.

Langlois, J.H., & Roggman, L.A. (1990). Attractive faces are only average. *Psychological Science*, **1**, 115–121.

Langlois, J.H., & Stephan, C. (1977). The effects of physical attractiveness and ethnicity on children's behavioral attributions and peer preferences. *Child Development*, **48**, 1694–1698.

Leopold, D.A., & Bondar, I. (2005). Adaptation to complex visual patterns in humans and monkeys. In C. Clifford & G. Rhodes (Eds.), *Fitting the mind to the world: Adaptation and after-effects in high-level vision* (pp. 189–211). London: Oxford University Press.

Little, A.C., Penton-Voak, I.S., Burt, D.M., & Perrett, D.I. (2003). Investigating an imprinting-like phenomenon in humans: partners and opposite-sex parent have similar hair and eye colour. *Evolution and Human Behaviour*, **24**, 43–51.

McArthur, L.Z., & Apatow, K. (1983–1984). Impressions of babyfaced adults. *Social Cognition*, **4**, 315–342.

McArthur, L.Z., & Berry, D.S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, **18**, 165–192.

MacLin, O.H., & Webster, M.A. (2001). The influence of adaptation on the perception of distortions in natural images. *Journal of Electronic Imaging*, **10**, 100–109.

Maier, R.A., Holmes, D.L., Slaymaker, F.L., & Reich, J.N. (1984). The perceived attractiveness of preterm infants. *Infant Behavior and Development*, **7**, 403–414.

Martin, C.L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development*, **66**, 1453–1471.

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

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

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

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

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

Perrett, D.I., Lee, K.J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castles, D.L., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, **394**, 884–887.

Perrett, D.I., May, K.A., & Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature*, **368**, 239–242.

Perrett, D.I., Penton-Voak, I.S., Little, A.C., Tidderman, B.P., Burt, D.M., Schmidt, N., Oxley, R., Kinloch, N., & Barrett, L. (2002). Facial attractiveness judgements reflect learning of parental age characteristics. *Proceedings of the Royal Society of London B*, **269**, 873–880.

Peskin, N., & Newell, F.N. (2004). Familiarity breeds attraction: effects of exposure on the attractiveness of typical and distinctive faces. *Perception*, **33**, 147–157.

Rhodes, G., Geddes, K., Jeffery, L., Dziurawiec, S., & Clark, A. (2002). Are average and symmetric faces attractive to infants? Discrimination and looking preferences. *Perception*, **31**, 315–321.

Rhodes, G., Halberstadt, J., Jeffery, L., & Palermo, R. (in press). The attractiveness of average faces is not a generalized mere exposure effect. *Social Cognition*.

Rhodes, G., Jeffery, L., Watson, T.L., Clifford, C.W.G., & Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness after-effects. *Psychological Science*, **14**, 558–566.

Rhodes, G., Lee, K., Palermo, R., Weiss, M., Yoshikawa, S., Clissa, P., Williams, T., Peters, M., Winkler, C., & Jeffery, L. (2005). Attractiveness of own-race, other-race, and mixed-race faces. *Perception*, **34**, 319–340.

Rhodes, G., Robbins, R., Jaquet, E., McKone, E., Jeffery, L., & Clifford, C. (2005). Adaptation and face perception: how after-effects implicate norm based coding of faces. In C. Clifford & G. Rhodes (Eds.), *Fitting the mind to the world: Adaptation and after-effects in high-level vision* (pp. 213–240). London: Oxford University Press.

Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, **7**, 105–110.

Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., & Akamatsu, S. (2001). Attractiveness of facial averageness and symmetry in non Western cultures: in search of biologically based standards of beauty. *Perception*, **30**, 611–625.

Rubenstein, A.J., Kalakanis, L., & Langlois, J.H. (1999). Infant preferences for attractive faces: a cognitive explanation. *Developmental Psychology*, **35**, 848–855.

Samuels, C.A., Butterworth, G., Roberts, T., Graupner, L., & Hole, G. (1994). Facial aesthetics: babies prefer attractiveness to symmetry. *Perception*, **23**, 823–831.

Slater, A., Bremner, G., Johnson, S.P., Sherwood, P., Hayes, R., & Brown, E. (2000). Newborn infants' preference for attractive faces: the role of internal and external facial features. *Infancy*, **1**, 265–274.

Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, **21**, 345–354.

Sternglanz, S.H., Gray, J.L., & Murakami, M. (1977). Adult preferences for infantile facial features: an ethological approach. *Animal Behavior*, **25**, 108–115.

Webster, M.A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, **428**, 557–561.

Webster, M.A., & MacLin, O.H. (1999). Figural after-effects in the perception of faces. *Psychonomic Bulletin and Review*, **6**, 647–653.

Received: 21 December 2004
Accepted: 16 November 2005