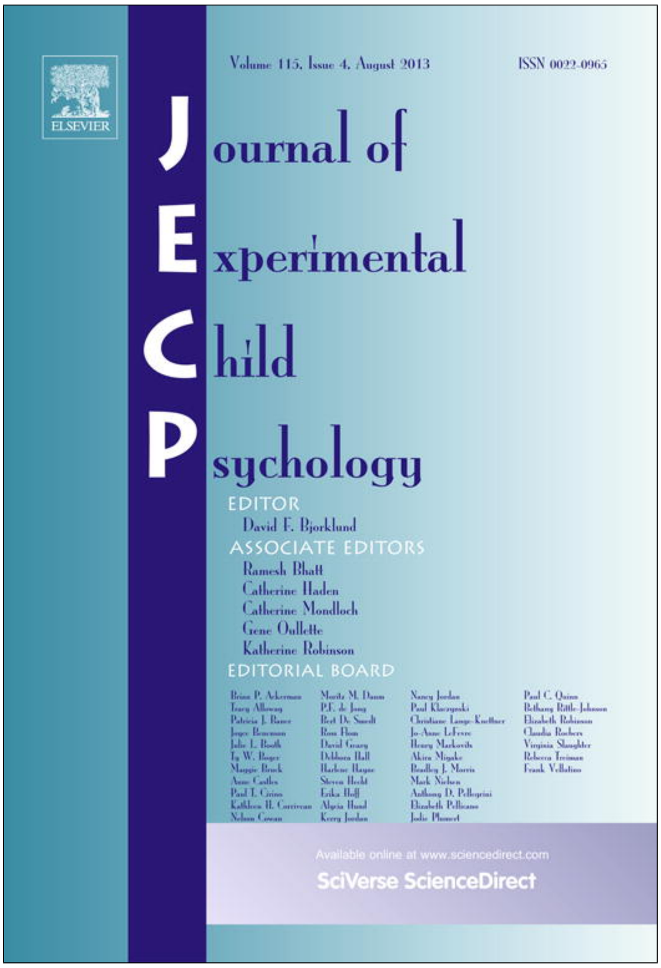
ELSEVIER
Volume 115, Issue 4, August 2013
ISSN 0022-0965
Journal of Experimental Child Psychology
EDITOR
David F. Bjorklund
ASSOCIATE EDITORS
Ramesh Bhatt
Catherine Haden
Catherine Mondloch
Gene Ouellette
Katherine Robinson
EDITORIAL BOARD
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Journal of Experimental Child Psychology 115 (2013) 624–639

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The influence of averageness on children's judgments of facial attractiveness

Larissa Vingilis-Jaremko, Daphne Maurer *

Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada L8S 4K1

ARTICLE INFO

Article history:
Received 24 September 2012
Revised 27 March 2013
Available online 23 May 2013

Keywords:
Face processing
Attractiveness
Averageness
Development
Children
Face space

ABSTRACT

We examined developmental changes in the influence of averageness on judgments of facial attractiveness by showing adults and children pairs of individual faces in which one face was transformed 50% toward its group average, whereas the other face was transformed 50% away from that average. In one comparison, adults and 5-year-olds rated the more average faces as more attractive whether the faces were of adult females, 5-year-old boys, or 5-year-old girls. The influence of averageness, however, was weaker in 5-year-olds than in adults. In another comparison, a new group of adults and 9-year-olds rated the more average faces as more attractive for male and female faces of adults, 9-year-olds, and 5-year-olds. The influence of averageness was again weaker for children than for adults, although the strength of 9-year-olds' preference was greater than that of 5-year-olds. Developmental changes may reflect the refinement of an average face prototype as children are exposed to more faces, increased sensitivity as visual perception develops, and/or the greater salience of attractiveness after puberty.

Introduction

Contrary to popular belief, there is high agreement among adults across cultures in the relative attractiveness of different faces (Bernstein, Lin, & McClellan, 1982; Cunningham, Roberts, Barbee, & Druen, 1995; Johnson, Dannenbring, Anderson, & Villa, 1983; Langlois et al., 2000; Perrett, May, & Yoshikawa, 1994), and developmentally infants look longer at faces judged by adults to be attractive

* Corresponding author.
E-mail address: maurer@mcmaster.ca (D. Maurer).

0022-0965/$ - see front matter
http://dx.doi.org/10.1016/j.jecp.2013.03.014

than those judged to be unattractive (Langlois et al., 1987; Samuels, Butterworth, Roberts, Graupner, & Hole, 1994; Slater, Quinn, Hayes, & Brown, 2000; Slater et al., 1998). Adults can appraise the attractiveness of a face in as little as a glance (Olsen & Marshuetz, 2005), and these quick judgments can influence social interactions because attractive people are judged to have more positive traits than those judged as unattractive (the "what is beautiful is good" stereotype; Dion, Berscheid, & Walster, 1972). One influence on facial attractiveness judgments in adults is a face's proximity to the population average.

In 1878, Sir Francis Galton published the observation that averaged composite faces are attractive. Using composite photography, Galton exposed the portraits of several individuals consecutively onto the same photographic plate, creating an average of the individual faces. He noted that the "composites are better looking than their components" (Galton, 1878, p. 98). Similarly, Langlois and Roggman (1990) found that averaged faces are attractive when they created averaged composites of digital images using 2, 4, 8, 16, and 32 faces by mathematically averaging the luminance values of individual pixels across the images. Adults rated the 16- and 32-face composites as more attractive than the mean rating of the original faces used in their creation. Moreover, composites created from greater numbers of original faces were rated as more attractive than those created from fewer original faces. These findings suggested that faces approximating the population mean are attractive. In addition, because the 16- and 32-face composites looked very similar to one another regardless of which original faces were used (Langlois & Roggman, 1990), and both were more attractive than the mean of their component faces, the average of 16 faces may be a good approximation of a population mean.

Although Langlois and Roggman's (1990) averaging method artificially smoothed skin texture, which could have led to the enhanced attractiveness of the composite over the component faces (Alley & Cunningham, 1991; Benson & Perrett, 1992), others have replicated the finding when they manipulated shape and texture separately. They did so by outlining the features and external contour of each face with landmark points, which can then be used to calculate an average face shape (Rowland & Perrett, 1995; Tiddeman, Burt, & Perrett, 2001). Individual faces can then be transformed relative to the average, such that the spatial configuration and shape changes, whereas the texture remains that of the original face. Using male faces, Little and Hancock (2002) found that separate manipulations that averaged texture or shape each increased attractiveness independently. Moreover, adults judge line drawings of faces, the shape of which have been transformed closer to their group average, to be more attractive than line drawings that have been transformed away from their group average despite the fact that line drawings remove the influence of skin tone and texture completely (Rhodes & Tremewan, 1996). These findings provide evidence that average face shape is attractive independent of average skin texture.

Whereas averageness and symmetry are confounded, because faces nearer to average are also more symmetrical, averageness remains attractive when the effects of symmetry and averageness are examined separately. For example, faces photographed in profile, where direct cues to bilateral symmetry are absent, are judged by adults to be more attractive after having been transformed toward their group average rather than away from their group average (Valentine, Darling, & Donnelly, 2004). In addition, faces that are nearer their group average are judged by adults to be more attractive than faces that are farther from their group average even when all faces have been made bilaterally symmetrical by blending each face with its mirror image (Jones, DeBruine, & Little, 2007; Rhodes, Sumich, & Byatt, 1999). Thus, averageness influences attractiveness judgments independent of symmetry. These studies, along with evidence that averaged faces are attractive across cultures (see Rhodes, Harwood, Yoshikawa, Nishitani, & McLean, 2002 for a review), and that faces naturally sitting closer to the population average are judged to be more attractive than more distinctive faces (Light, Hollander, & Kayra-Stuart, 1981), provide strong evidence that facial averageness is attractive.

From an evolutionary perspective, facial averageness may be attractive because of stabilizing selection, in which evolutionary pressures act against extremes of a trait in favor of average faces or the most common or average features (Dobzhansky, 1982). For many heritable traits, the average signals heterozygosity or having dissimilar gene pairs for heritable characteristics (Fink & Penton-Voak, 2002; Thornhill & Gangestad, 1993). Heterozygosity can signal an outbred individual with greater genetic diversity and resistance to parasites (Thornhill & Gangestad, 1993, 1999), and such individuals may carry fewer harmful mutations, all of which could lead to a mate preference (Dobzhansky, 1982).

Evolution may have also selected for cognitive mechanisms that facilitate processing of faces near the population average. It is hypothesized that faces are represented within a multidimensional face space centered on a norm, or average face, formed based on our accumulated experience with faces (Rhodes, 2006; Valentine, 1991). In this system of norm-based coding, individual faces are represented as unique multidimensional vectors defined by their differences and distances from the prototype; faces near the prototype may be processed more fluently, with greater speed and efficiency, and consequently preferred (Valentine, 1991; Winkielman, Halberstadt, Fazendeiro, & Catty, 2006). Indeed, random dot patterns closer to a prototype of random dots presented to adults during a training phase are processed more fluently, and rated as more attractive, than less prototypical patterns (Winkielman et al., 2006). In addition, dogs, wristwatches, and birds rated by adults as more prototypical are also rated as more attractive (Halberstadt & Rhodes, 2000). Prototypical patterns, objects, and faces may be processed more fluently, leading to a preference.

Developmentally, there is evidence that infants can form cognitive prototypes of faces by 3 months of age because 3-month-olds (but not 1-month-olds) show evidence of recognizing a composite of four faces after being familiarized to the four faces individually (de Haan, Johnson, Maurer, & Perrett, 2001; see Rubenstein, Kalakanis, & Langlois, 1999, for evidence in 6-month-olds). In that experiment, even without familiarization, female 3-month-olds also showed a looking preference for the composite face over the individual component faces (de Haan et al., 2001). By 5 years of age, children show evidence of norm-based coding for processing facial identity; adaptation to a face identity leads to a shift in the recognition of other faces, and by 8 years of age (youngest age tested) the shift is specific to faces that lie on a trajectory passing through an average face (Jeffery et al., 2011; 8-year-olds: Nishimura, Maurer, Jeffery, Pellicano, & Rhodes, 2008; Pimperton, Pellicano, Jeffery, & Rhodes, 2009). By 7 years of age, these identity aftereffects are stronger for adapters that lie farther from the average (Jeffery et al., 2011), consistent with the predictions of norm-based coding. There is similar evidence at a younger age for figural distortions; by 4 to 6 years of age, adaptation to faces that have been contracted or expanded (or have had the eyes moved up/down) leads to a shift in children's perception of an average face, and the shift is greater for distortions farther from the norm (Jeffery et al., 2010). Such distortions also shift children's judgments of attractiveness as early as 5 years of age (Short, Hatry, & Mondloch, 2011; see Anzures, Mondloch, & Lackner, 2009, for similar evidence in 8-year-olds); their attractiveness judgments shift in the adapted direction, a result suggesting that, like adults (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003), their judgments of attractiveness are based on a prototype that is constantly being updated as they encounter new faces. Children's accuracy at recognizing faces improves from 6 to 10 years of age (Diamond & Carey, 1977), presumably as the prototype becomes more refined.

However, there have been no studies of whether averageness affects attractiveness judgments in children to the same extent as in adults. At 6 months of age, infants look longer at average faces than at faces rated by adults to be unattractive (Rubenstein et al., 1999). At 5 to 8 months of age, infants do not look longer at faces transformed toward average than those transformed away from average; indeed, the longest look was toward the less average face (Rhodes, Geddes, Jeffery, Dziurawiek, & Clark, 2002). By adolescence, children do select faces that have been transformed toward their group average to be more attractive than the original versions of the faces (Saxton, Debruine, Jones, Little, & Roberts, 2009, 2011; Saxton et al., 2010). However, because these studies did not have adult comparison groups, it is not known when during development averageness becomes as strong an influence as in adults.

The purpose of this study was to explore the influence of averageness on judgments of facial attractiveness during mid-childhood because we are aware of no published data between 8 months of age and adolescence. We tested 5-year-olds, an age at which norm-based coding is used for processing of facial identity (Jeffery et al., 2010) and the youngest age able to complete enough trials to calculate reliable individual data. We also tested 9-year-olds, a pre-adolescent age at which most aspects of basic vision are adult-like (Adams & Courage, 2002; Ellemberg, Lewis, Liu, & Maurer, 1999; Lewis et al., 2004). Adults were tested for comparison. We showed children and adults pairs of individual faces that had been warped 50% toward and away from their group averages. Participants selected which face from each pair they found to be more attractive. We used faces of children and adults because children may have more experience with faces of their own age than of adults and because children and adults have own-age biases in processing faces (Anastasi & Rhodes, 2005; Hills & Lewis, 2011; but see Macchi Cassia, 2011, for evidence that children have a processing advantage for adult faces). We used photographs of young adults, 4- and 5-year-olds, and 8- and 9-year-olds, reflecting the recent experience of the participants. To shorten the task, we presented 5-year-olds and the first group of adult participants with the three face categories with which they should have most experience: faces of 5-year-old girls, 5-year-old boys, and women because the full set of six face categories would have been challenging for 5-year-olds to complete in a single test session and because any experience-based influence of averageness should be most likely to be manifest for these face categories. The 9-year-olds and another group of adults saw, in addition, faces of 9-year-old girls and boys (their age mates) and faces of men. This allowed us to evaluate whether the influence of averageness on children's judgments of attractiveness is weaker for men's faces than for women's faces, which they experience most beginning during infancy (Rennels & Davis, 2008) and which could continue into childhood. The use of all six face categories in the second comparison allowed us to evaluate the extent to which the influence of averageness on attractive judgments is invariant across face age and sex, at least for adults and 9-year-olds. The six face categories were presented in separate counterbalanced blocks.

Method

Participants

There were two subgroups of participants. Subgroup A consisted of 36 5.5-year-olds (±3 months) and 36 adults (17–28 years of age). Subgroup B, which completed a longer procedure, consisted of a second group of 36 adults (18–33 years of age) and 36 9-year-olds (±3 months). All participants were White,[1] and half at each age were male. Adults were undergraduate psychology students at McMaster University and participated in exchange for course credit. Children were recruited from the names of mothers who volunteered shortly after the births of their children to be contacted about future studies, and they received a toy or book for participation. All participants had normal or corrected-to-normal vision and met age-appropriate visual screening criteria. Adults and 9-year-olds had normal stereoacuity of 40 arc seconds of disparity on the Titmus test of Stereoacuity and a Snellen acuity of 20/20 or better in each eye, measured on a Lighthouse eye chart, and 5-year-olds had normal stereoacuity of 100 arc seconds of disparity on the Titmus test of Stereoacuity and a Snellen acuity of 20/30 or better in each eye, measured with the Cambridge Crowding Cards, because vision is still immature at this age (Adams & Courage, 2002; Ellemberg et al., 1999).

An additional 5 adults, 12 9-year-olds, and 9 5-year-olds were tested but excluded from the data because they failed our visual screening requirements (3 adults, 7 9-year-olds, and 3 5-year-olds), were outside of our age range (2 adults and 1 5-year-old), had their identical twin tested (1 9-year -old and 1 5-year-old), were inattentive (2 9-year-olds and 2 5-year-olds), had corrupted data (1 5-year-old), or were not White (2 9-year-olds and 1 5-year-old).

Stimuli

Stimuli were color full-face photographs with neutral expressions of young adult women, young adult men, 8- to 9-year-old girls, 8- to 9-year-old boys, 4- to 5-year-old girls, and 4- to 5-year-old boys. There were 16 faces in each of the six face categories for a total of 96 faces. Adult models, and parents of children, gave consent for their photographs to be used and manipulated in research. Faces were evenly lit and directly faced the camera. The brush tool in Adobe Photoshop CS was used to remove major blemishes. Each face was manually delineated with 189 landmark points, outlining the face

[1] McMaster University is a very multicultural university. Because others have reported better recognition and discrimination of faces of one's own race than other races in adults (see Meissner & Brigham, 2001, for a meta-analysis) and children (Chance, Turner, & Goldstein, 1982; Corenblum & Meissner, 2006; Goldstein & Chance, 1980; Sangrigoli & de Schonen, 2004), we wanted to ensure that participants in all age groups had a similar type of face experience. As such, we used White participants and White face stimuli.

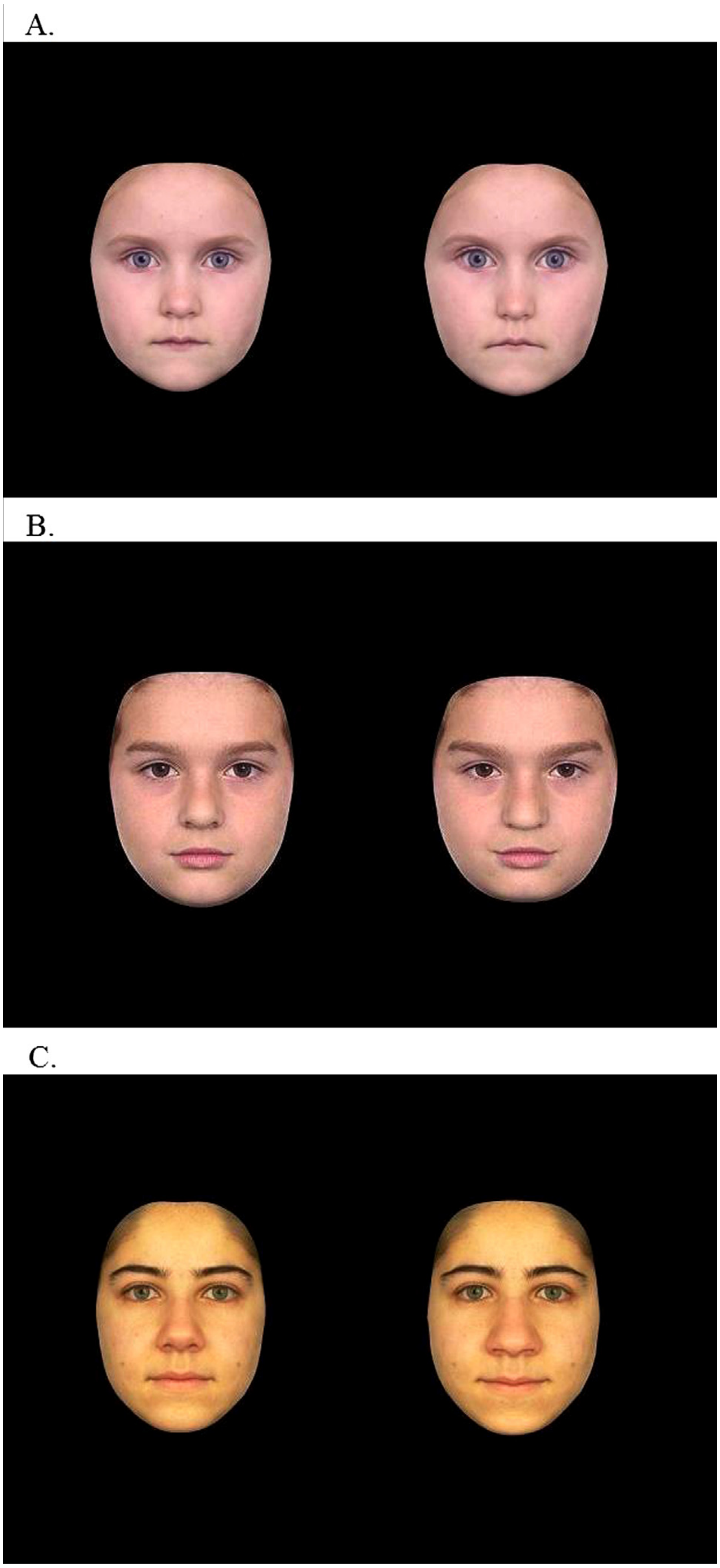

Fig. 1. More average (left) and less average (right) versions of a 5-year-old girl's face (A), a 9-year-old boy's face (B), and an adult woman's face (C).

shape and features, using PsychoMorph. An average face was then created for each face category (woman, man, 8- to 9-year-old girl, 8- to 9-year-old boy, 4- to 5-year-old girl, and 4- to 5-year-old boy) by moving the points on each of the 16 faces to their mean location. Each original face was then transformed 50% toward and away from its group average by moving each of its delineated points halfway toward the group average, and the same distance away from the group average (see Tiddeman et al., 2001), creating pairs of faces that differed in averageness based on shape but maintained the texture of the original face.

Because faces become more symmetrical when they are moved closer to average, all faces were made perfectly symmetrical by averaging each face with its mirror image to remove the influence of symmetry. We symmetrized both the shape and texture of the faces because bilateral differences in shadow and skin pigmentation can provide cues to symmetry, another influence on judgments of attractiveness (Perrett et al., 1999). Because this symmetrizing procedure leads to the reflection of two flashes in each eye, one of the reflections was removed from each eye with the brush tool in Adobe Photoshop CS. Because both faces from each pair were symmetrized, there were no major textural differences between the two versions of the face. To remove any influence from distortions in the hairstyle, external features and hair were removed by placing a black background outside the outline of each face while maintaining the original face shape (Fig. 1). Although it is not typical to view faces without hair, this precaution maximized the salience of internal physiognomic cues and, hence, made it less likely that we would underestimate children's sensitivity to the cue of averageness. Faces were standardized for size based on interpupillary distance. Subgroup A (5-year-olds and one group of adults) viewed the faces on a Dell Trinitron P1330, 21-inch CRT monitor. Subgroup B (9-year-olds and the other group of adults) viewed the faces on an HP 20555 SH249, 22-inch LCD monitor. Images had 2048 horizontal by 3072 vertical pixel resolution. From a viewing distance of 50 cm, faces subtended a visual angle of approximately 12° in height and 7° in width.

To ensure that the six face categories were equivalent in normality, a separate group of 24 White adults (17–27 years of age, 12 male and 12 female) rated how normal each of the unmanipulated faces looked on a 5-point scale (1 = *very normal* to 5 = *very unusual*). Each face was presented on a gray background without hair but retaining face shape. Faces were presented individually in blocks counterbalanced for age and sex of face. Participants viewed the faces on an HP 20555 SH249, 22-inch LCD monitor, and responses were made on a keyboard. A Cronbach's α of .873 indicated internal consistency in the ratings. The mean normality ratings did not differ among the six face categories (no main effect of category), $F(5,115) = 1.625$, $p = .159$, $\eta^2_p = .066$.

Design

We used a within-participant design, with the 5-year-olds and corresponding adults in Subgroup A viewing the three face categories (16 face pairs per face category) blocked for a total of 48 trials per participant. The 9-year-olds and corresponding adults in Subgroup B viewed the six face categories for a total of 96 trials per participant. Stimulus presentation was randomized within each block, and the side of the more average face was randomized across trials. The order of blocks was counterbalanced across participants in each age group.

Procedure

This study received ethics clearance from the institutional research ethics board. After explaining the procedure, we obtained verbal consent from child participants (and written consent from their parents) and from adult participants. We then explained the instructions in a game-like format:

> I'm having a birthday party, and I have invited all of my friends! Among my friends, pairs of brothers and pairs of sisters will be attending the party. These brothers and sisters have been trying very hard to look their very best. They are trying to make themselves look nice because a clown at the party will be giving a red balloon to the brother or sister from each pair who looks cuter, prettier, or more handsome today. Every time you see two faces on the screen, they are either sisters or brothers, and it will be your job to figure out who looks better, nicer, or more attractive today and will receive the red balloon from the clown!

A number of words were used to describe the concept of attractiveness, including *prettier, more handsome, nicer looking, better looking, cuter,* and *more attractive.* Children as young as 3 years can provide reliable attractive judgments in the same direction as those of adults for children's and adults' faces using the words *pretty, cute,* and *handsome* (Cooper, Geldart, Mondloch, & Maurer, 2006; Dion, 1973; Langlois & Stephan, 1977). Participants were asked whether they understood the procedure and were then presented with criterion trials containing nine pairs of non-face stimuli that differed in attractiveness (e.g., a new book and an old tattered book) and were asked to select which object looked better or nicer. All participants completed the criterion trials with 100% accuracy and moved onto the main experiment. Trials were self-paced, and participants made selections by clicking on the image with a mouse, which initiated the next trial. The 5-year-old participants were visually screened after the first block of faces and took a break after the second block of faces. The 9-year-old participants were visually screened after the third block of faces, and the adult participants were visually screened at the beginning of the experiment. All groups took additional breaks as needed. Following the experiment, participants were debriefed with information about the purpose of the experiment and about the "what is beautiful is good" stereotype (Dion et al., 1972).

Results

Data analysis

For each participant and group of faces, we calculated the mean proportion of trials on which the more average face was selected. We replaced outliers greater than 3 standard deviations from the mean with new participants (Subgroup A: 2 adults males and 1 5-year-old boy; Subgroup B: 5 adults [1 male and 4 females] and 4 9-year-olds [2 boys and 2 girls]). To assess whether participants of each age selected the more average faces more frequently than chance, we performed one-tailed one-sample t tests comparing the mean for each face category with chance (0.5) for each age group, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$ for Subgroup A and $\alpha = .008$ for Subgroup B. To assess whether attention declined during the experiment, we did a preliminary analysis with the within-participant factor of block order (Blocks 1–3 for Subgroup A and Blocks 1–6 for Subgroup B) and the between-participant factor of participant age. We followed up with paired-samples t tests, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$ for Subgroup A and $\alpha = .008$ for Subgroup B.

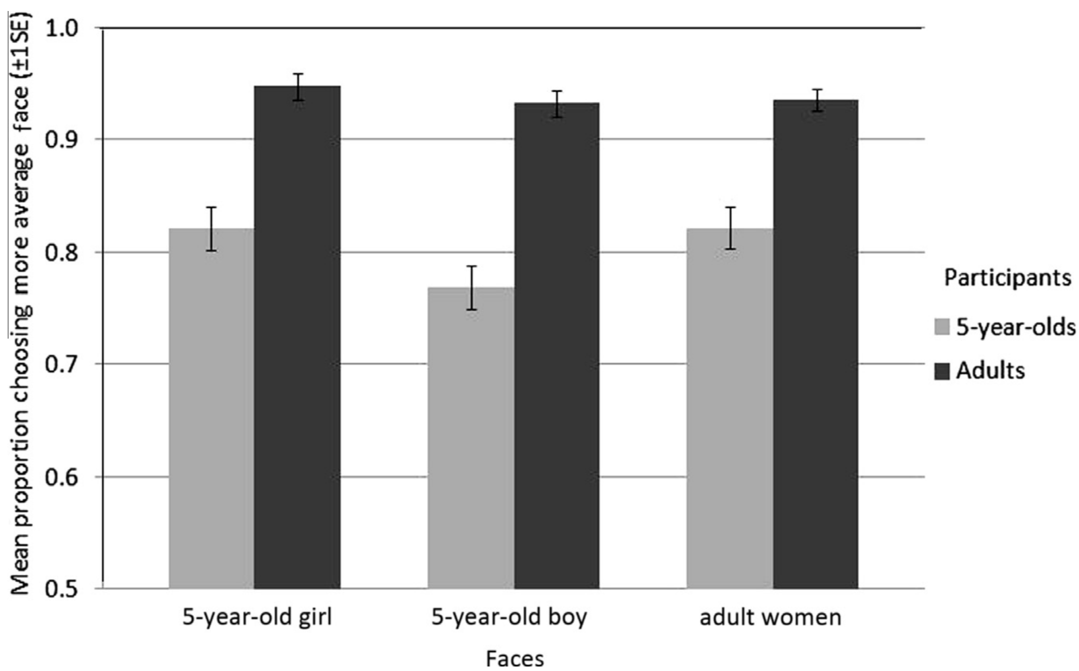

Fig. 2. Mean proportions of trials on which 5-year-old (gray bars) and adult (dark bars) participants selected the more average face for each face category. Bars represent between-participant standard errors.

To compare the strength of the influence of averageness across face categories and participant groups, we performed repeated-measures analyses of variance (ANOVAs). For Subgroup A, we used the between-participant factors of participant age (5 years or adult) and participant sex and the within-participant factor of face category (5-year-old girl, 5-year-old boy, or woman). For Subgroup B, we used the between-participant factors of participant age (9 years or adult) and participant sex and the within-participant factors of face age (5 years, 9 years, or adult) and face sex. We performed an additional ANOVA comparing 5- and 9-year-old participants on the three face categories on which they were both tested with the between-participant factors of participant age and participant sex and the within-participant factor of face category (5-year-old girl, 5-year-old boy, or woman).

We also conducted complementary item analyses across faces rather than across participants by calculating the proportion of participants of each age (5 years, 9 years, or adult) selecting the more average face for each group of faces (16 face pairs by three face categories for Subgroup A and six face categories for Subgroup B). For each participant age, this produced 16 preference scores for each of the groups of faces. The null hypothesis states that the mean of the preference scores should be 50%, with half of the raters selecting the more average face. Each item analysis followed the same sequence as the main analyses, as described above.

Subgroup A: 5-year-olds and corresponding adults

Adult and 5-year-old participants selected the more average face more frequently than chance for all face categories (all ps < .001; see Fig. 2). A preliminary analysis indicated a main effect of block order, $F(2,140) = 3.54$, $p = .032$, $\eta_p^2 = .048$, and no interaction between block order and participant age, $F(2,70) = 0.050$, $p = .952$, $\eta_p^2 = .001$. Paired-samples t tests indicated that averageness influenced attractiveness judgments more strongly in the third block ($M = .875$, $SD = .132$) than the second block ($M = .832$, $SD = .171$), $t(71) = -2.48$, $p = .016$, $d = 0.282$, and without Bonferroni correction more strongly in the third block than the first block ($M = .844$, $SD = .140$), $t(71) = -2.02$, $p = .047$, $d = 0.228$, suggesting that attention did not decline from the first block to the last block.

Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 7.382$, $p = .025$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.906$). The ANOVA revealed main effects of face category, $F(1.81,123.15) = 3.55$, $p = .036$,

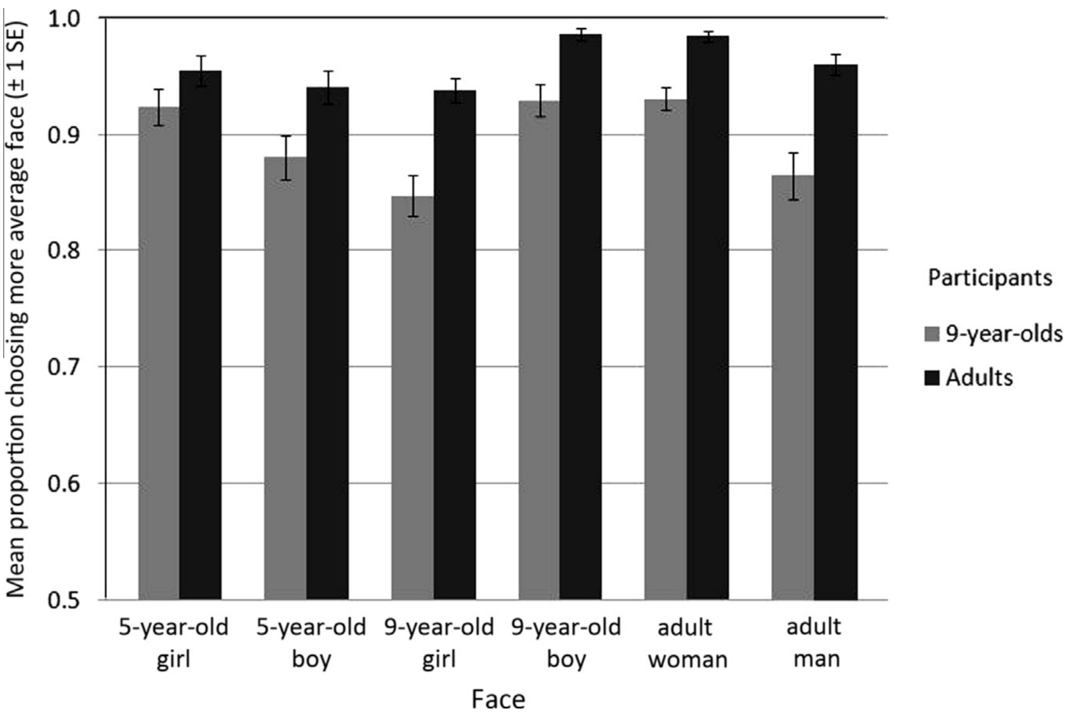

Fig. 3. Mean proportions of trials on which 9-year-old (gray bars) and adult (dark bars) participants selected the more average face for each face category. Bars represent between-participant standard errors.

$\eta_p^2 = .050$, and participant age, $F(1,68) = 67.96$, $p < .001$, $\eta_p^2 = .500$, with adults selecting the more average faces more frequently than 5-year-olds. None of the interactions was significant (all *ps* > .063). We followed up the main effect of face category with two-tailed paired-samples *t* tests comparing the means for the three face categories, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. Participants selected the more average faces marginally more frequently among 5-year-old girl faces ($M = .867$, $SD = .139$) than 5-year-old boy faces ($M = .826$, $SD = .162$), $t(71) = 2.42$, $p = .018$, $d = 0.272$. No other comparisons were significant (both *ps* > .095).

The item analysis indicated that for every face category, more 5-year-old and adult participants than expected by chance selected the more average faces for all three face categories (all *ps* < .001). The ANOVA on the item analysis revealed main effects of participant age, $F(1,60) = 139.036$, $p < .001$, $\eta_p^2 = .699$, and face category, $F(2,120) = 3.218$, $p = .043$, $\eta_p^2 = .051$, with more adults selecting the more average faces than 5-year-olds. We followed up with two-tailed paired-samples *t* tests comparing means for the 5-year-old girl, 5-year-old boy, and woman faces, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. There were no differences in the proportion of participants selecting the more average faces between face categories. Without Bonferroni correction, more participants selected the more average faces in the 5-year-old girl faces ($M = .867$, $SD = .052$) than the 5-year-old boy faces ($M = .826$, $SD = .06$), $t(15) = 2.37$, $p = .032$, $d = 0.730$.

Subgroup B: 9-year-olds and corresponding adults

Adult and 9-year-old participants selected the more average face more frequently than chance for all face categories (all *ps* < .001; see Fig. 3). A preliminary analysis indicated a main effect of block order, $F(5,350) = 4.60$, $p < .001$, $\eta_p^2 = .062$, and no interaction between block order and participant age, $F(5,70) = 1.65$, $p = .146$, $\eta_p^2 = .023$. Paired-samples *t* tests indicated that averageness influenced attractiveness judgments less strongly in the first block ($M = .898$, $SD = .095$) than the fifth block ($M = .944$, $SD = .070$), $t(71) = -4.22$, $p < .001$, $d = 0.551$, or sixth block ($M = .948$, $SD = .083$), $t(71) = -2.64$, $p < .001$, $d = 0.561$. Without Bonferroni correction, averageness influenced attractiveness judgments less strongly in the first block than all other blocks (all *ps* < .035). Averageness also influenced attractiveness judgments less strongly in the second block ($M = .924$, $SD = .088$) than the sixth block, $t(71) = -2.59$, $p = .012$, $d = 0.281$. These results indicate that attention did not decline during the experiment and that averageness influenced attractiveness judgments more strongly in later blocks than earlier blocks. As with Subgroup A, the increasing effect of averageness over blocks did not interact with age and, hence, cannot explain the age differences in the strength of the influence.

The ANOVA revealed a main effect of participant age, $F(1,68) = 25.96$, $p < .001$, $\eta_p^2 = .276$, with adults selecting the more average faces more frequently than 9-year-olds, and a face sex by face age interaction, $F(2,136) = 25.65$, $p < .001$, $\eta_p^2 = .274$. To understand whether there are differences between the different face ages because of possible processing advantages for same-age faces (Anastasi & Rhodes, 2005; Hills & Lewis, 2011) or adult faces (Macchi Cassia, 2011), we followed up the interactions with separate ANOVAs for male and female faces with the within-participant factor of face age (5 years, 9 years, or adult) and found a main effect of face age for male faces, $F(2,142) = 10.71$, $p < .001$, $\eta_p^2 = .131$. Mauchly's test indicated that the assumption of sphericity had been violated in the female face ANOVA, $\chi^2(2) = 8.71$, $p = .013$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.90$). The analysis revealed a main effect of face age, $F(1.79,127.13) = 18.48$, $p < .001$, $\eta_p^2 = .207$. We followed up for each face sex with two-tailed paired-samples *t* tests comparing each face age, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. The interaction arose because participants selected the more average faces less frequently for 9-year-old girl faces ($M = .892$, $SD = .097$) than 5-year-old girl faces ($M = .939$, $SD = .090$), $t(71) = 3.72$, $p < .001$, $d = 0.502$, or women's faces ($M = .958$, $SD = .054$), $t(71) = -7.01$, $p < .001$, $d = 0.840$. The reverse was true in male faces, with participants selecting the more average faces more frequently for 9-year-old boy faces ($M = .958$, $SD = .069$) than 5-year-old boy faces ($M = .911$, $SD = .104$), $t(71) = -4.36$, $p < .001$, $d = 0.533$, and men's faces ($M = .912$, $SD = .105$), $t(71) = 4.21$, $p < .001$, $d = 0.518$. There were no differences between adult and 5-year-old faces for either sex (both *ps* > .102).

It is also of interest to understand whether there is an influence of face sex for each face age because of a potential processing advantage for female faces (Quinn, Yahr, & Kuhn, 2002; Ramsey, Langlois, &

Marti, 2005; Ramsey-Rennels & Langlois, 2006); therefore, we broke down the interaction a second way by doing two-tailed paired-samples t tests comparing male and female faces at each face age, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. Participants selected the more average faces more frequently in the faces of adult females ($M = .958$, $SD = .054$) than adult males ($M = .912$, $SD = .105$), $t(71) = 4.56$, $p < .001$, $d = 0.551$, and more frequently in the faces of 9-year-old boys ($M = .958$, $SD = .069$) than 9-year-old girls ($M = .892$, $SD = .097$), $t(71) = -6.10$, $p < .001$, $d = 0.784$. Participants selected the more average faces marginally more frequently without Bonferroni correction in the faces of 5-year-old girls ($M = .939$, $SD = .090$) than of 5-year-old boys ($M = .911$, $SD = .104$), $t(71) = 1.97$, $p = .052$, $d = 0.288$.

In the complementary analysis by item, more 9-year-old and adult participants than expected by chance selected the more average faces for each of the six face categories (all $ps < .001$). Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 8.68$, $p = .013$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.88$). The ANOVA revealed a main effect of participant age, $F(1,60) = 46.89$, $p < .001$, $\eta_p^2 = .439$, and a face sex by face age interaction, $F(2,120) = 15.48$, $p < .001$, $\eta_p^2 = .205$. Separate ANOVAs for male and female faces with the within-participant factor of face age (5 years, 9 years, or adult) revealed a main effect of face age for male faces, $F(2,126) = 7.89$, $p = .001$, $\eta_p^2 = .111$. Mauchly's test indicated that the assumption of sphericity had been violated in the female face ANOVA, $\chi^2(2) = 20.53$, $p < .001$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.78$). The analysis revealed a main effect of face age, $F(2,126) = 9.22$, $p = .001$, $\eta_p^2 = .128$, among female faces. We followed up for each face sex with two-tailed paired-samples t tests comparing each face age, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. More participants selected the more average faces for 9-year-old boy faces ($M = .958$, $SD = .03$) than 5-year-old boy faces ($M = .911$, $SD = .059$), $t(15) = -2.94$, $p = .01$, $d = 1.00$, or without Bonferroni correction men's faces ($M = .912$, $SD = .065$), $t(15) = 2.37$, $p = .032$, $d = 0.908$. No other comparisons were significant (all $ps > .067$). To get a more complete picture of the interaction, we also did two-tailed paired-samples t tests comparing male and female faces at each face age, controlling for multiple comparisons with Bonferroni correction at $\alpha = .017$. More participants selected the more average faces in 5-year-old girl faces ($M = .939$, $SD = .042$) than 5-year-old boy faces ($M = .911$, $SD = .059$), $t(15) = 2.81$, $p = .013$, $d = 0.547$. Without Bonferroni correction, more participants selected the more average faces in adult female faces ($M = .958$, $SD = .041$) than adult male faces ($M = .912$, $SD = .065$), $t(15) = 2.47$, $p = .026$, $d = 0.846$, and marginally more participants selected the more average faces in 9-year-old boy faces ($M = .958$, $SD = .030$) than 9-year-old girl faces ($M = .892$, $SD = .112$), $t(15) = -2.07$, $p = .056$, $d = 0.805$.

Comparison of 5-year-old and 9-year-old data for overlapping face categories

Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 7.10$, $p = .029$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity ($\varepsilon = 0.91$). The ANOVA revealed main effects of face category, $F(2,136) = 6.39$, $p = .003$, $\eta_p^2 = .086$, and participant age, $F(1,68) = 43.208$, $p < .001$, $\eta_p^2 = .389$, with 9-year-olds selecting the more average faces more frequently than 5-year-olds. We followed up with paired-samples t tests comparing the three face categories, controlling for multiple comparisons with Bonferroni correction at $\alpha = .016$. Participants selected the more average faces more frequently for 5-year-old girl faces ($M = .855$, $SD = .140$) than 5-year-old boy faces ($M = .800$, $SD = .158$), $t(71) = 2.94$, $p = .004$, $d = 0.368$, and more frequently for adult female faces ($M = .854$, $SD = .145$) than 5-year-old boy faces, $t(71) = -2.89$, $p = .005$, $d = 0.356$. There were no differences between 5-year-old girl faces and adult female faces, $t(71) = 0.061$, $p = .952$, $d = 0.356$.

Discussion

We found that averageness influences 5-year-olds', 9-year-olds', and adults' judgments of facial attractiveness of both children's and adults' faces. This is the first evidence, to our knowledge, of the influence of averageness on children's judgments of facial attractiveness between infancy and adolescence, and it suggests a long developmental trajectory, with the influence emerging before 5 years and maturing after 9 years of age. Other studies have found that infants as young as 3 months can recognize an average of faces to which they were familiarized (de Haan et al., 2001). By 6 months, infants look longer at average faces than unattractive faces (Rubenstein et al., 1999); however, when presented with stimuli like those in the current study, infants look equally long at the two faces, and the longest look is to the less average face (Rhodes, Geddes, et al., 2002). Our results indicate that by 5 years of age, averageness has clearly emerged as an influence on children's judgments of attractiveness. This pattern is consistent with the influence of average feature height on children's facial attractiveness judgments, which is present by 5 years and not yet adult-like at 9 years of age (Cooper et al., 2006). Averageness, however, influences children's judgments of attractiveness earlier than bilateral symmetry, which has little to no influence during infancy (Rhodes, Geddes, et al., 2002) or even at 5 years of age (Vingilis-Jaremko & Maurer, 2013). When the influence becomes adult-like is unknown because the studies that documented that adolescents find averageness to be attractive did not have adult comparison groups (Saxton et al., 2009, 2010, 2011). This study is also the first demonstration that averageness influences attractiveness judgments of children's faces, a result suggesting that both adults and children have enough experience with children's faces to form a prototype[2] or that a prototype of adult faces is sufficiently similar to a prototype of children's faces to exert an influence on judgments of attractiveness. Our results suggest that averageness affects children's judgments of attractiveness in everyday life by 5 years of age, at least for categories with which they have had the most experience, namely own age and adult female faces. Of course, it is possible that the influence is weaker at 5 years of age for less familiar categories not tested here such as adult male faces. In any event, the influence for even familiar categories was weaker than in 9-year-olds and adults.

As with any study in developmental psychology, we cannot rule out differences in attention and motivation. However, all participants passed the nine criterion trials with 100% accuracy, suggesting that the children understood the task and that they were capable of adult-like performance when examples were sufficiently dissimilar. In addition, performance did not decline for any age group from the first block to the last block. There are several possible visual and cognitive explanations for the weaker influence of averageness on children's than adults' judgments of facial attractiveness. First, children have had less experience with faces than adults, a difference that could lead to a less stable prototype and, hence, more variability in judgments of attractiveness based on averageness. Differences in exposure in the real world have been shown to alter judgments of attractiveness in both adults (Apicella, Little, & Marlowe, 2007; Geldart, 2008) and children (Cooper et al., 2006). For example, the Hadza, an isolated hunter–gatherer society in Africa, do not find averageness to be attractive in British faces, a group with whom they have little to no experience, whereas they do find averageness to be attractive in faces of the Hadza. Conversely, Western individuals, who presumably have more experience with faces from diverse cultures, find averageness to be attractive in both British and Hadza faces (Apicella et al., 2007). Additional evidence comes from studies of the influence of feature height on judgments of attractiveness in adults (Geldart, 2008) and children (Cooper et al., 2006). Shorter adults, who typically look up at faces, prefer faces with a larger chin and smaller forehead, matching the foreshortening with which they typically see faces (Geldart, 2008). That preference is present during infancy (Geldart, Maurer, & Henderson, 1999), and changes to a preference for faces with a smaller chin, such as those of peers, when children begin to interact more with their peers at eye level after entering preschool (Cooper et al., 2006).

Short-term shifts of a similar type can be created in the laboratory by biased exposure to faces with distorted feature height (Cooper & Maurer, 2008) or distorted shape (Anzures et al., 2009). For example, showing adults faces with features at a high (or low) height shifts their subsequent judgments of the ideal feature height in the adapted direction (Cooper & Maurer, 2008). Showing adults faces that are expanded or compressed shifts their attractiveness judgments in that direction (Rhodes et al., 2003). Similar attractiveness aftereffects have been shown in children (Anzures et al., 2009) and occur

[2] Although some believe that an exemplar-based model better accounts for preferential responses near the average, recent evidence more strongly favors a prototype-based model for face perception in adults (Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006; Robbins, Maurer, Hatry, Anzures, & Mondloch, 2007; Webster & MacLin, 1999) and children (Jeffery et al., 2010; Jeffery & Rhodes , 2011; Jeffery et al., 2011; Nishimura et al., 2008). However, it should be noted that the interpretation of the preferences found here remains the same whether based on prototype- or exemplar-based models.

even after exposure to very bizarre faces that have no effect on adults' judgments of oddness or attractiveness (Hills, Holland, & Lewis, 2010). Such shifts raise the possibility that the prototype might not be as refined in children as in adults, and that might explain why averageness has a weaker influence on their judgments of attractiveness.

A second possible explanation for the increase in the influence of averageness on judgments of attractiveness from childhood to adulthood is the immaturity of the visual system. Although both 5-year-olds and 9-year-olds could discriminate between the faces, they may have had greater difficulty in perceiving some of the finer differences that were visible to adults. Children's sensitivity to subtle differences in the location of faces' internal features improves from 6 years to 10 or 11 years of age (Baudouin, Gallay, Durand, & Robichon, 2010; Mondloch, Le Grand, & Maurer, 2002), and their ability to detect global structure from dot patterns, moving dots, and biological motion improves after 6 years until as late as 12 to 14 years of age for moving patterns (Hadad, Maurer, & Lewis, 2011; Lewis et al., 2004). In addition, acuity and contrast sensitivity become adult-like around 6 years and 7 to 9 years of age, respectively (Adams & Courage, 2002; Ellemberg et al., 1999). These visual immaturities could affect children's ability to detect as many subtleties between the faces as adults and thereby limit the influence of averageness on their judgments of attractiveness.

A related possibility is that a less stable norm and/or worse visual sensitivity could limit the processing advantage for faces nearer the prototype. It is hypothesized that adults prefer faces closer to the population average because those faces are processed more quickly and easily than faces that are farther from the average (Winkielman et al., 2006). If children have higher levels of internal noise, which can originate from random fluctuations in neural response, information loss during neural transmission, and receptor sampling errors, and/or if children have less efficient processing, defined as a lesser ability to detect the signal despite the internal noise (see Lu and Dosher 1999) , than do adults, judgments based on processing fluency could be affected. It is possible to estimate processing efficiency and levels of internal noise by superimposing external noise (i.e., random variations in luminance across pixels) onto stimuli to be detected (see Levi, 2005). Jeon, Maurer, and Lewis (2012) did so by measuring contrast sensitivity for gratings embedded in external noise. They found improvements in the ability to detect the signal in the noise from 5 to 9 years and from 9 years to adulthood. Computer modeling suggested that the improvements were caused by both decreases in internal noise and increases in efficiency. These developmental differences in internal noise and efficiency could reduce children's preferences based on processing fluency, including the influence of averageness on judgments of facial attractiveness.

Finally, it is possible that the influence of averageness on judgments of facial attractiveness is not adult-like by 9 years of age because pre-pubescent children might not be attuned to decisions relevant to mate choice. Averageness could be attractive to adults because of stabilizing selection (Dobzhansky, 1982) and/or because it is a signal of heterozygosity (Fink & Penton-Voak, 2002; Thornhill & Gangestad, 1993). The influence of averageness on judgments of attractiveness might not mature to adult levels until after puberty, when mate choice decisions are more relevant. Although averageness influences adolescents' judgments of attractiveness (Saxton et al., 2009, 2010, 2011), and the strength of the influence does not change from early to mid-adolescence (Saxton et al., 2009, 2011), we do not know whether the influence strengthens after adolescence because there were no adult comparison groups in these studies and we do not know whether it changes between 9 years of age and the onset of puberty.

We found no evidence of an own-age bias because there was no interaction between face age and participant age. Such a bias might be expected based on differential experience, which could lead to a more well-defined and stable prototype for own-age faces. Indeed, adults and children are better at recognizing faces of their own age than faces of other ages, consistent with an experiential account (Harrison & Hole, 2009; Anastasi & Rhodes, 2005, 2006; Hills, 2012; Hills & Lewis, 2011; Kuefner, Macchi Cassia, Picozzi, & Bricolo, 2008). In addition, differential experience appears to affect attractiveness judgments in some cases (e.g., Anzures et al., 2009; Apicella et al., 2007; Cooper et al., 2006). Such experiential differences might affect attractiveness through the formation of prototypes for different categories of faces that are somewhat distinct. Evidence for at least partially distinct prototypes comes from studies using contingent adaptation with adults. For example, adults can be adapted in opposite directions by male and female faces distorted in opposite ways (Little, DeBruine, & Jones, 2005) or by own- and other-race faces manipulated in opposite ways (Jacquet, Rhodes, & Hayward, 2008). Because such opposite aftereffects have not been tested with child and adult faces, we do not know whether distinct prototypes exist for faces of different ages. However, when adults are adapted to one sex of child face, the adaptation partially transfers to adult faces, a result suggesting a partially shared representation of sex across face ages (Barrett & O'Toole, 2009). In the current study, there was no evidence of an influence of an own-age effect on children's judgments of attractiveness; the influence of averageness was not consistently strongest for faces matching participants' ages (5-year-old faces viewed by 5-year-olds, 9-year-old faces viewed by 9-year-olds, and adult faces viewed by adults). However, it has been suggested recently that children actually get more exposure to adult faces than child faces throughout development and, hence, may be best at processing adult faces (Macchi Cassia, 2011). The predicted effect of a stronger influence of averageness on children's judgments of adult faces than children's faces also was not observed. We speculate that the influence of the well-formed face prototype from familiar categories may affect the reaction to less familiar categories because there may be similarities in the average face shape across categories. Alternatively, some minimal level of exposure may be sufficient to form a prototype that resembles an average face for that category.

We found interactions between age and sex of face; for 5-year-old and adult faces, the influence of averageness was generally stronger for female faces than male faces. It was not the case, however, that participants had a stronger preference for averageness in female faces in general, with 9-year-olds and adults selecting the more average faces more frequently for 9-year-old boy faces than 9-year-old girl faces (5-year-old participants were not tested with this face category). Others have reported a processing advantage for female faces over male faces in infants (see Ramsey et al., 2005; Ramsey-Rennels & Langlois, 2006), possibly because infants typically have more experience with female females than male faces (Rennels & Davis, 2008). This advantage for processing female faces could persist if children continue to see more female faces than male faces and/or if early experience contributes to later face processing (see Macchi Cassia, Kuefner, Picozzi, & Vescovo, 2009). Although this could explain our findings of a greater influence of averageness on women's faces than men's faces, it does not account for why averageness affected attractiveness judgments more strongly in 9-year-old boy faces than 9-year-old girl faces. Alternately, it is possible that the interaction resulted from differences in the typicality of the particular 16 faces we chose for each category, although we note that there were no differences in the mean normality ratings for the original unmanipulated faces among the six categories, as rated by a separate group of adults.

Although averageness is attractive, it is not the peak of attractiveness; an averaged composite face created only from faces rated by adults as attractive is judged to be more attractive than a composite created from a wider selection of faces (Perrett et al., 1994). Exaggerating the differences between the two composites can lead to faces judged to be even more attractive (DeBruine, Jones, Unger, Little, & Feinberg, 2007; Perrett et al., 1994). Future directions should include exploring whether children similarly find these faces along the "attractiveness dimension" to be attractive and testing the influence of averageness on attractiveness judgments with older children and adolescents with adult comparison groups to determine the age at which these preferences mature.

In summary, averageness influences judgments of attractiveness as early as 5 years of age and does so reliably enough that it is likely to influence everyday social interactions. After this age, its influence increases as children build more stable face prototypes and become sensitive to more subtle differences among faces, both of which may allow a stronger effect of processing fluency. By the time children begin to think about mate choice after puberty, the preference is well established.

Acknowledgements

We thank Kathleen Lee, Sarah Lyon, and Julian Ghloum for help with the collection of data, and David Feinberg for contributing a subset of the face photographs that we used to create stimuli, for advice on creating stimuli and for comments on an earlier draft of the manuscript. We thank Cathy Mondloch for contributing another subset of the face photographs that were used to create stimuli. This research was supported by a grant from the Canadian Natural Sciences and Engineering Council (9797) to Daphne Maurer. Larissa Vingilis-Jaremko was supported by an NSERC CGS D.

References

Adams, R. A., & Courage, M. L. (2002). Using a single test to measure human contrast sensitivity from early childhood to maturity. *Vision Research, 42*, 1205–1210.

Alley, T. R., & Cunningham, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science, 2*, 123–125.

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review, 12*, 1043–1047.

Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology, 8*, 237–252.

Anzures, G., Mondloch, C. J., & Lackner, C. (2009). Face adaptation and attractiveness aftereffects in 8-year-olds and adults. *Child Development, 80*, 178–191.

Apicella, C. L., Little, A. C., & Marlowe, F. W. (2007). Facial averageness and attractiveness in an isolated population of hunter–gatherers. *Perception, 36*, 1813–1820.

Barrett, S. E., & O'Toole, A. J. (2009). Face adaptation to gender: Does adaptation transfer across age categories? *Visual Cognition, 17*, 700–715.

Baudouin, J., Gallay, M., Durand, K., & Robichon, F. (2010). The development of perceptual sensitivity to second-order facial relations in children. *Journal of Experimental Child Psychology, 107*, 195–206.

Benson, P., & Perrett, D. (1992). Face to face with the perfect image. *New Scientist*, 32–35.

Bernstein, I. H., Lin, T., & McClellan, P. (1982). Cross- vs. within-racial judgments of attractiveness. *Perception & Psychophysics, 32*, 495–503.

Chance, J. E., Turner, A. L., & Goldstein, A. G. (1982). Development of differential recognition for own- and other-race faces. *Journal of Psychology, 112*, 29–37.

Cooper, P. A., Geldart, S. S., Mondloch, C. J., & Maurer, D. (2006). Developmental changes in perceptions of attractiveness: A role of experience? *Developmental Science, 9*, 530–543.

Cooper, P. A., & Maurer, D. M. (2008). The influence of recent experience on perceptions of attractiveness. *Perception, 37*, 1216–1226.

Corenblum, B., & Meissner, C. A. (2006). Recognition of faces of ingroup and outgroup children and adults. *Journal of Experimental Child Psychology, 93*, 187–206.

Cunningham, M. R., Roberts, A. R., 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.

de Haan, M., Johnson, M. H., 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.

DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1420–1430.

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

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.

Dobzhansky, T. (1982). *Genetics and the origin of species*. New York: Columbia University Press.

Ellemberg, D., Lewis, T. L., Liu, C. H., & Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vision Research, 39*, 2325–2333.

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11*, 154–158.

Galton, F. J. (1878). Composite portraits. *Nature, 18*, 97–100.

Geldart, S. (2008). Tall and good-looking? The relationship between raters' height and perceptions of attractiveness. *Journal of Individual Differences, 29*, 148–156.

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

Goldstein, A. G., & Chance, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology, 105*, 47–59.

Hadad, B., Maurer, D., & Lewis, T. L. (2011). Long trajectory for the development of sensitivity to global and biological motion. *Developmental Science, 14*, 1330–1339.

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

Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review, 16*, 264–269.

Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology, 48*, 499–508.

Hills, P. J., Holland, A. M., & Lewis, M. B. (2010). Aftereffects for face attributes with different natural variability: Children are more adaptable than adolescents. *Cognitive Development, 25*, 278–289.

Hills, P. J., & Lewis, M. B. (2011). The own-age face recognition bias in children and adults. *Quarterly Journal of Experimental Psychology, 64*, 17–23.

Jacquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition, 16*, 734–753.

Jeffery, L., McKone, E., Haynes, R., Firth, E., Pellicano, E., & Rhodes, G. (2010). Four- to six-year-old children use norm-based coding in face-space. *Journal of Vision, 10*, 18.

Jeffery, L., & Rhodes, G. (2011). Insights into the development of face recognition mechanisms revealed by face aftereffects. *British Journal of Psychology, 102*, 799–815.

Jeffery, L., Rhodes, G., McKone, E., Pellicano, E., Crookes, K., & Taylor, E. (2011). Distinguishing norm-based from exemplar-based coding of identity in children: Evidence from face identity aftereffects. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1824–1840.

Jeon, S. T., Maurer, D., & Lewis, T. L. (2012). Characterizing the mechanisms behind improvements in visual sensitivity during childhood [abstract]. *Journal of Vision*, 12, art. 1366.

Johnson, R. W., Dannenbring, G. L., Anderson, N. R., & Villa, R. E. (1983). How different cultural and geographic groups perceive the attractiveness of active and inactive feminists. *Journal of Social Psychology, 119*, 111–117.

Jones, B. C., DeBruine, L. M., & Little, A. C. (2007). The role of symmetry in attraction to average faces. *Perception & Psychophysics, 69*, 1273–1277.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 811–817.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or 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., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype? *Developmental Psychology, 23*, 363–369.

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., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after effects. *Nature Neuroscience, 4*, 89–94.

Levi, D. M. (2005). Perceptual learning in adults with amblyopia: A reevaluation of critical periods in human vision. *Developmental Psychobiology, 46*, 222–232.

Lewis, T., Ellemberg, D., Maurer, D., Dirks, M., Wilkinson, F., & Wilson, H. (2004). A window on the normal development of sensitivity to global form in glass patterns. *Perception, 33*, 409–418.

Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality & Social Psychology Bulletin, 7*, 269–276.

Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B, 272*, 2283–2287.

Little, A. C., & Hancock, P. J. B. (2002). The role of masculinity and distinctiveness in judgments of human male facial attractiveness. *British Journal of Psychology, 93*, 451–464.

Lu, Z., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America, 16*, 764–778.

Macchi Cassia, V. (2011). Age biases in face processing: The effects of experience across development. *British Journal of Psychology, 102*, 816–829.

Macchi Cassia, V., Kuefner, D., Picozzi, M., & Vescovo, E. (2009). Early experience predicts later plasticity for face processing evidence for the reactivation of dormant effects. *Psychological Science, 20*, 853–859.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*, 3–35.

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

Nishimura, M., Maurer, D., Jeffery, L., Pellicano, E., & Rhodes, G. (2008). Fitting the child's mind to the world: Adaptive norm-based coding of facial identity in 8-year-olds. *Developmental Science, 11*, 620–627.

Olsen, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*, 498–502.

Perrett, D. I., Burt, M. B., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior, 20*, 295–307.

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

Pimperton, H., Pellicano, E., Jeffery, L., & Rhodes, G. (2009). The role of higher level adaptive coding mechanisms in the development of face recognition. *Journal of Experimental Child Psychology, 104*, 229–238.

Quinn, P., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception, 31*, 1109–1121.

Ramsey, J. L., Langlois, J. H., & Marti, C. N. (2005). Infant categorization of faces: Ladies first. *Developmental Review, 25*, 212–246.

Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infants' differential processing of female and male faces. *Current Directions in Psychological Science, 15*, 59–62.

Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior and Development, 31*, 665–678.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.

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

Rhodes, G., Harwood, K., Yoshikawa, S., Nishitani, M., & McLean, I. (2002). The attractiveness of average faces: Cross-cultural evidence and possible biological basis. In G. Rhodes & L. A. Zebrowitz (Eds.), *Facial attractiveness: Evolutionary, cognitive, and social perspectives* (pp. 35–58). Westport, CT: Ablex.

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46*, 2977–2987.

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., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science, 10*, 52–58.

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

Robbins, R. A., Maurer, D., Hatry, A., Anzures, G., & Mondloch, C. J. (2007). Effects of normal and abnormal visual experience on the development of opposing aftereffects for upright and inverted faces. *Developmental Science, 15*, 194–203.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications, 15*(5), 70–76.

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.

Sangrigoli, S., & de Schonen, S. (2004). Effect of visual experience on face processing: A developmental study of inversion and non-native effects. *Developmental Science, 7*, 74–87.

Saxton, T. K., Debruine, L. M., Jones, B. C., Little, A. C., & Roberts, S. C. (2009). Face and voice attractiveness judgments change during adolescence. *Evolution and Human Behavior, 30*, 398–408.

Saxton, T. K., Debruine, L. M., Jones, B. C., Little, A. C., & Roberts, S. C. (2011). A longitudinal study of adolescents' judgments of the attractiveness of facial symmetry, averageness, and sexual dimorphism. *Journal of Evolutionary Psychology, 43*, 43–55.

Saxton, T. K., Kohoutova, D., Roberts, S. C., Jones, B. C., DeBruine, L. M., & Havlicek, J. (2010). Age, puberty, and attractiveness judgments in adolescents. *Personality and Individual Differences, 49*, 857–862.

Short, L. A., Hatry, A. J., & Mondloch, C. J. (2011). The development of norm-based coding and race-specific face prototypes: An examination of 5- and 8-year-olds' face space. *Journal of Experimental Child Psychology, 108*, 338–357.

Slater, A., Quinn, P. C., Hayes, R., & Brown, E. (2000). The role of facial orientation in newborn infants' preference for attractive faces. *Developmental Science, 3*, 181–185.

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

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty, averageness, symmetry, and parasite resistance. *Human Nature, 4*, 237–269.

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*, 452–460.

Tiddeman, B., Burt, D. M., & Perrett, D. (2001). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications, 21*(5), 42–50.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology A, 43*, 161–204.

Valentine, T., Darling, S., & Donnelly, M. (2004). Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychonomic Bulletin & Review, 11*, 482–487.

Vingilis-Jaremko, L., & Maurer, D. (in press). The influence of symmetry on children's judgments of facial attractiveness. *Perception, 42*(3), 302–320.

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

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science, 17*, 799–806.