Journal of Experimental Child Psychology 103 (2009) 355–375

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Exploring children's face-space: A multidimensional scaling analysis of the mental representation of facial identity

Mayu Nishimura [a,b,*], Daphne Maurer [b], Xiaoqing Gao [b]

[a] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[b] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ont., Canada L8S 4K1

ARTICLE INFO

Article history:
Received 28 October 2008
Revised 12 February 2009
Available online 1 April 2009

Keywords:
Face processing
Face recognition
Visual memory
Multidimensional scaling
Similarity judgments
Face-space

ABSTRACT

We explored differences in the mental representation of facial identity between 8-year-olds and adults. The 8-year-olds and adults made similarity judgments of a homogeneous set of faces (individual hair cues removed) using an "odd-man-out" paradigm. Multidimensional scaling (MDS) analyses were performed to represent perceived similarity of faces in a multidimensional space. Five dimensions accounted optimally for the judgments of both children and adults, with similar local clustering of faces. However, the fit of the MDS solutions was better for adults, in part because children's responses were more variable. More children relied predominantly on a single dimension, namely eye color, whereas adults appeared to use multiple dimensions for each judgment. The pattern of findings suggests that children's mental representation of faces has a structure similar to that of adults but that children's judgments are influenced less consistently by that overall structure.

Introduction

Adults have a remarkable ability to remember and recognize many faces in their everyday lives. Adults' recognition of faces is impaired when faces are inverted, and this "inversion effect" is greater for faces than for non-face objects such as houses and airplanes (e.g., Yin, 1969). Adults appear to process faces holistically; they process the individual features, such as the eyes and mouth, together as a unitary whole or Gestalt percept (e.g., Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). Adults also demonstrate acute sensitivity to changes in feature shape and second-order relations, which refer to the spatial relations among internal facial features, such as the distance between the two eyes, but

* Corresponding author. Fax: +1 412 268 2798.
E-mail address: mayunishimura@gmail.com (M. Nishimura).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2009.02.005

more for upright faces than for inverted faces (e.g., Brooks & Kemp, 2007; Collishaw & Hole, 2000; Freire & Lee, 2001; Mondloch, Le Grand, & Maurer, 2002).

Like adults, children (and to a certain extent even infants) demonstrate a face inversion effect (e.g., Brace et al., 2001; Itier & Taylor, 2004a; Mondloch et al., 2002; Schwarzer, 2000; Turati, Sangrigoli, Ruel, & de Schonen, 2004). Children as young as 4 to 6 years of age also demonstrate holistic processing of faces (Carey & Diamond, 1994; de Heering, Houthuys, & Rossion, 2007; Mondloch et al., 2007; Pellicano & Rhodes, 2003; Pellicano, Rhodes, & Peters, 2006; Tanaka et al., 1998). In addition, 6-year-olds can recognize the identity of faces based on changes in external contour as accurately as adults, and 8-year-olds are nearly as good as adults in detecting changes in the shapes of internal features such as eyes and mouth (Mondloch et al., 2002).

However, it takes many years for children's face processing abilities to reach adult levels. For example, 8-year-olds are worse than adults at recognizing facial identity across changes in point of view (Mondloch, Geldart, Maurer, & Le Grand, 2003) or noticing subtle changes in the second-order relations (Mondloch et al., 2002). Their immaturity in processing the second-order relations of upright faces is evident even when memory demands are eliminated as well as when the salience of feature shapes is reduced (Mondloch, Dobson, Parson, & Maurer, 2004). Improvements with age in sensitivity to second-order relations continue into adolescence (Mondloch, Le Grand, & Maurer, 2003). These behavioral immaturities are consistent with neuroimaging data. Adults demonstrate face-selective activation in the right anterior fusiform gyrus, an area referred to as the face fusiform area or the FFA (e.g., Kanwisher, McDermott, & Chun, 1997). Under many conditions, such FFA activation is not observed in children until 12 to 16 years of age (Aylward et al., 2005; Gathers et al., 2004; Passarotti et al., 2003; Passarotti, Paul, & Stiles, 2001; Scherf, Berhmann, Humphreys, & Luna, 2007). However, when age-related differences in blood-oxygenation-level-dependent (BOLD) signals are controlled carefully, face-selective right FFA activation has been shown in children as young as 7 to 11 years of age, albeit in an area that is significantly smaller than that in adults (Golari et al., 2007). Similar developmental trajectories are revealed through event-related potentials (Itier & Taylor, 2004b; Taylor, Edmonds, McCarthy, & Allison, 2001).

The source of the immaturities in face processing may lie in children's mental representation of faces. A useful framework that describes adults' mental representation of faces is face-space, a multidimensional space in which individual faces are represented by unique multidimensional vectors from the origin (Valentine, 1991). The origin of face-space represents the average of previously encountered faces such that the distance from the origin represents the distinctiveness of a face (more typical faces will be located closer to the average), and the direction from the origin represents *how* the face deviates from average along dimensions that are important for face recognition (Valentine, 1991). The average of face-space is updated continuously with experience, and there may be separate averages for different populations of faces such as male versus female faces or faces of different races (e.g., Baudouin & Gallay, 2006; Byatt & Rhodes, 1998; Levin, 1996; Little, DeBruine, & Jones, 2005; Valentine, 1991; Watson, Rhodes, & Clifford, 2006).

Adaptation paradigms provide supporting evidence that adults indeed encode faces relative to an average that is updated constantly. For example, when an observer adapts for several seconds to a face that has been digitally compressed, he or she judges a previously normal face as being expanded (Rhodes et al., 2003; Watson & Clifford, 2003; Webster & MacLin, 1999). Presumably, adaptation has shifted the observer's perception of normality toward the adapting face (i.e., the observer's mental "average" is more compressed than it was prior to adaptation), so that a previously "normal" face looks "expanded" relative to the updated mental average. Similarly, when a pair of "opposite" faces (relative to an averaged morph) is created through digital morphing, adapting to one face shifts the perception of identity toward the other face (e.g., Anderson & Wilson, 2005; Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006).

The face-space model is consistent with adults' faster classification judgments (i.e., judging whether a stimulus is a face or not) of typical faces than of distinctive faces but their slower individual recognition of faces (i.e., judging whether a face belongs to "Bob" or "Fred") for typical faces than for distinctive faces (e.g., Johnston & Ellis, 1995; Rhodes, Byatt, Tremewan, & Kennedy, 1997; Valentine, 1991; Valentine & Bruce, 1986). This effect is likely caused by the spatial gradient of face-space. Because the distance from the origin represents the distinctiveness of a face, and by definition there are more typical faces than distinctive faces, more faces will cluster around the origin, making the distance among neighboring faces smaller for typical faces than for distinctive faces. Therefore, activation of a single typical face may also activate neighboring faces, creating a stronger pooled signal that a stimulus is a face but also causing more misidentification errors. Consistent with this interpretation, caricatures created by exaggerating the difference between a face and the norm are easier to recognize and anticaricatures created by decreasing the difference are harder to recognize (e.g., Rhodes et al., 1997).

Very little research has examined the utility of the face-space model in describing children's face processing. Chang, Levine, and Benson (2002) showed that 6-year-olds, like older children (8- and 10-year-olds) and adults, chose caricatures as being more distinctive and anticaricatures as being more typical than their original faces. However, the magnitude of the caricature effect was much smaller in 6-year-olds than in older children and adults. One possible explanation for this age effect is that fewer faces are represented in the face-space of younger children such that the spatial gradient is weaker, producing weaker caricature effects. Consistent with this interpretation, in another study, 5- to 7-year-olds demonstrated weaker effects of typicality than did adults because the children did not identify distinctive faces better than typical faces, although like adults they were faster at classifying typical faces than distinctive faces as faces (Johnston & Ellis, 1995). However, children at least 9 years of age showed typicality effects like those of adults in both identification and classification tasks (Johnston & Ellis, 1995). Therefore, an adult-like face-space may be forming during 7 to 9 years of age. A recent study using an adaptation paradigm showed that 8-year-olds demonstrate an identity aftereffect of a magnitude similar to that of adults (Nishimura et al., 2008), suggesting that by 8 years of age children represent facial identity as deviations from the average.

Studies examining face identity aftereffects and the effects of typicality or distinctiveness provide some information about face-space but cannot elucidate whether children and adults use different dimensions in face-space. The theory of face-space, as originally proposed by Valentine (1991), did not specify what the dimensions of face-space are; rather, Valentine proposed only that the dimensions represent cues to identity that are important for discriminating faces. For example, a dimension may represent an easily identifiable feature (e.g., eye color) or a more complex combination of cues that cannot be verbalized easily (e.g., combined surface area of forehead, cheeks, and chin). One approach to describing the dimensions of face-space is to use multidimensional scaling (MDS), a statistical procedure that represents measurements of perceived similarity among pairs of objects as distances between points in a multidimensional space. MDS can often show regularities that remain hidden if the raw similarity judgments are examined directly (Borg & Groenen, 2005).

Using Caucasian male faces 20 to 25 years of age, Johnston, Milne, Williams, and Hosie (1997) applied MDS to adults' similarity ratings. The MDS solution revealed that faces that were rated independently as being more typical were located closer to the origin and that faces that were rated independently as being more distinctive were located farther away from the origin, a pattern consistent with the model of face-space. In addition, Johnston and colleagues found that three or four dimensions accounted well for adults' similarity ratings. For male Caucasian faces, these dimensions appeared to represent face width, perceived age, facial hair, and forehead size (Busey, 1998; Johnston et al., 1997).

In the current study, we used MDS to examine the characteristics of 8-year-olds' representation of faces because by 8 years of age children show evidence of norm-based coding (Nishimura et al., 2008) and adult-like caricature effects (Chang et al., 2002; Johnston & Ellis, 1995), but 8-year-olds are not yet adult-like in making identity judgments based on the spatial relations of facial features or across changes in viewpoint (e.g., Mondloch, Le Grand, & Maurer, 2002, 2003). The 8-year-olds' immaturities in face processing may reflect differences between children's and adults' representations in the number of dimensions, the nature and utility of those dimensions (e.g., children may have dimensions that represent unreliable cues to facial identity such as the presence of glasses), and the weights placed on each dimension, and/or the tuning of the dimensions (i.e., children's dimensions might not be well defined). One previous study by Pedelty, Levine, and Shevell (1985) used MDS to examine the representation of unfamiliar faces by 7-, 9-, and 12-year-olds and adults. Using 12 face stimuli and a child-friendly rating scale, they found no age differences in the number of dimensions needed to adequately account for the observed similarity ratings. A three-dimensional solution was sufficient, with the dimensions likely representing hair color, face width, and hairstyle. The results suggest that hair cues play an important role in how children and adults perceive the identity of unfamiliar faces and that children's and adults' representations are organized in a similar manner. However, individual results revealed that the younger children tended to rely more heavily on one or two of the dimensions, whereas adults used the three dimensions equally.

One limitation to the study by Pedelty et al. (1985) was the availability of hair color and hairstyle as major cues to identity in their small stimulus set, which may have led to an overestimation of the similarity between children's and adults' representations. Already by 5 years of age, children are able to use the outer contour of faces to recognize unfamiliar faces in a manner similar to adults (Want, Pascalis, Coleman, & Blades, 2003), whereas processing identity based on the spatial relations among the internal features of a face appears to improve into adolescence (Mondloch, Le Grand, & Maurer, 2002, 2003). In addition, children under 10 years of age are more likely to make errors in recognizing unfamiliar faces by basing their judgments on misleading cues from paraphernalia (Baenninger, 1994; Diamond & Carey, 1977; Freire & Lee, 2001). The salience of the external hair cues in the faces used by Pedelty et al. (1985) may have led children to rely on that feature alone, exaggerating their reliance on a single dimension. Adults also use external features, such as hairstyle and contour shape, as cues to identity when they are available (e.g., Johnston et al., 1997; Pedelty et al., 1985), but unlike children they are also able to discriminate faces reliably when these features are not available (e.g., Collishaw & Hole, 2000; Freire, Lee, & Symons, 2000; Haig, 1984; Mondloch et al., 2002; Rhodes, Brake, & Atkinson, 1993). In this study, we examined whether 8-year-olds' representations resemble those of adults when salient external features, such as hair color and hairstyle, are not available as cues to identity.

We developed a child-friendly paradigm to measure children's face representations using MDS. Although traditional MDS paradigms use rating scales to measure adults' perception of object similarity (e.g., Busey, 1998; Johnston et al., 1997; Lee, Byatt, & Rhodes, 2000; Pedelty et al., 1985), this method assumes that the observer uses the scale consistently across trials, which may be difficult for children. An "odd-man-out" paradigm can circumvent this problem because on each trial the observer simply needs to choose the stimulus that appears to be most different from the rest, requiring no reference to an internal rating scale and no memory of the values assigned to previous faces. The odd-man-out method has been used reliably to collect similarity judgments from adults (Kahana & Bennett, 1994; Romney, Brewer, & Batchelder, 1993; Weller & Romney, 1988; Yotsumoto, Kahana, Wilson, & Sekuler, 2007) and children (Miller & Gelman, 1983).

The purpose of Experiment 1 was to determine the smallest number of face stimuli that would produce an adequate account of adults' face representations because we wanted to minimize the number of trials for children. Specifically, we asked adults to rate the similarity of 630 pairs of faces from young adult Caucasian women. We limited the face stimuli to female faces because it is unclear whether sex of face is one dimension of a single face-space or whether there are different face-spaces for the two sexes (Baudouin & Gallay, 2006; Bruce, Burton, & Dench, 1994; Johnston et al., 1997; Little et al., 2005). We also used the same hair for all faces to tap into more subtle differences between children's and adults' representations than those examined in Pedelty et al. (1985) study, in which the 12 faces differed markedly in hairstyle and hair color. From the MDS solutions based on adults' similarity ratings, we determined the smallest number of faces that would produce stable solutions. Using this subset, we obtained similarity judgments from another group of adults using the odd-man-out method (Experiment 2a). Once the convergence of the two methods was confirmed, we used the odd-man-out paradigm to examine 8-year-olds' face representations (Experiment 2b).

Experiment 1

The purpose of Experiment 1 was twofold: to obtain similarity data from adults using a rating scale and to determine the smallest number of face stimuli that would produce a reliable MDS solution. In addition, according to the face-space hypothesis (Valentine, 1991) and previous findings (e.g., Johnston & Ellis, 1995; Rhodes et al., 1997; Valentine & Bruce, 1986), typical faces should cluster around the origin of face-space (i.e., the norm) and distinctive faces should be located in the periphery. So that we could test this hypothesis with the similarity data, a separate group of adults rated the distinctiveness of the faces.

Method

Participants

A total of 24 Caucasian undergraduate students (18 women and 6 men, 18–20 years of age, mean age = 18.13 years) rated the similarity of pairs of faces. An additional 24 Caucasian undergraduates (20 women and 4 men, 18–25 years of age, mean age = 19.20 years) rated the distinctiveness of faces. All students participated for bonus credit in an introductory psychology course and had normal or corrected-to-normal vision.

Stimuli

The test stimuli consisted of 36 colored photographs of female Caucasian faces chosen from a larger set of 97 female faces that were rated previously on attractiveness. The 18 most attractive and 18 least attractive faces with minimal makeup were chosen to ensure a diverse sample of faces. To eliminate hair color and hairstyle as diagnostic cues to identity, the hair (including small parts of the forehead and cheeks) from one woman was cropped and copied onto the faces of the remaining 35 women using Adobe Photoshop such that all 36 faces had the same hair (see Fig. 1). The skin tones of all faces were adjusted slightly to smooth the transition between the new hair outline and the original faces. Pilot testing confirmed that the digital modifications did not result in artifacts that made the faces appear to be unusual.

In addition to the 36 test stimuli, a set of 6 male faces was created in an analogous manner to serve as the practice set. All stimuli were 10.5×15.3 cm and were viewed from a distance of 100 cm ($6.00 \times 8.74°$ of visual angle).

Apparatus

Stimuli were presented in color on a 22-inch monitor using Superlab software running in Mac OS 9.

Procedure

This experiment was approved by the research ethics board of McMaster University. After being informed about the nature of the task, all participants provided written consent at the beginning of the session. Each participant was tested individually in a dimly lit room.

During the practice session, the experimenter explained the use of the rating scale from 1 to 7 (1 = *not very similar*, 7 = *very similar*). The practice session began by presenting 6 male faces sequentially (1000 ms per face) and then presenting all 6 faces on the screen simultaneously. The observer was asked to choose the pair that appeared to be most similar and was instructed that such a pair should receive a 6 or 7 on the rating scale. The observer then chose the pair that appeared to be most different and was instructed that such a pair should receive a 1 or 2 on the rating scale. The practice

Fig. 1. An example of a typical female face (left) and a distinctive female face (right) used as test stimuli.

session was used to explain the range of variability in the face stimuli and to encourage the full use of the scale.

The testing session began by first showing all 36 female test faces sequentially (1000 ms per face) to demonstrate the range of variability in the face stimuli. Following this presentation, pairs of faces were presented side by side on the screen on every trial, and the participant was asked to rate the similarity of the 2 faces on a scale from 1 to 7. All possible pairings and orderings (i.e., Face A on right and Face B on left; Face A on the left and Face B on the right) of the 36 faces produce 1260 trials. We divided the 1260 trials pseudorandomly into two groups (Orders 1 and 2), so that for each group there were 630 trials during which each face *pair* was shown only once and each *face* was shown approximately equally often on the left and right sides. Each participant saw either Order 1 or Order 2.

Separately, the 36 face stimuli were rated on distinctiveness by a different group of 24 participants. Each face was printed out on a 4×6-inch photo postcard using a Canon printer. Six poster boards were constructed, with each poster board showing 6 photographs attached with Velcro. The assignment of each photo to a particular poster board, as well as the positions of the photos on each poster board, was randomized across participants. Poster boards were used as a fun and interactive method that would be suitable for children if it became necessary to collect distinctiveness ratings for children and adults separately. Each adult observer was asked to stand 100 cm from the poster board and to select the most distinctive face (i.e., "the face that would stand out the most in a crowd"). The chosen face was then removed, and the observer was asked to choose the next most distinctive face. This process continued until all 6 faces had been removed from the poster board and was repeated for the remaining five poster boards until all 36 faces had been judged. The first face to be removed from each poster board (i.e., the most distinctive) was given a distinctiveness score of 6, followed by 5, 4, 3, 2, and 1 for subsequently chosen faces. These distinctiveness scores were averaged across participants to obtain a mean distinctiveness score for each face. The mean ratings ranged from 4.71 to 1.83 ($SD = 0.74$), with greater values indicating greater distinctiveness. The faces were then ranked based on their distinctiveness scores and divided into two categories such that the top 18 faces were categorized as distinctive (D1 = most distinctive, D18 = least distinctive, mean rating = 4.11) and the bottom 18 faces were categorized as typical (T1 = most typical, T18 = least typical, mean rating = 2.89).

Data analysis

The similarity ratings were recoded into distances such that pairs of faces given the highest similarity rating (7 on the rating scale) had a distance of 1 and pairs of faces with the lowest similarity rating (1 on the rating scale) had a distance of 7. Using these distance values from each participant, MDS was performed using the INDSCAL procedure in SPSS software (version 13.0) for Mac OS X. INDSCAL is an MDS procedure that produces a group solution as well as a measure that characterizes the individual differences within the group (Martens & Zacharov, 2000). The Euclidean distance model was selected, with the measurement level being specified as ordinal and the matrix shape being specified as symmetric, as in previous studies (Johnston et al., 1997; Yotsumoto et al., 2007). Five analyses were performed to assess the fit of solutions with two to six dimensions, as in previous studies (e.g., Busey, 1998; Johnston et al., 1997; Lee et al., 2000). We then examined the layout of the solution to assess whether typical faces were located closer to the origin than distinctive faces. After the validity of the MDS solution was confirmed based on the typicality effect, we examined the fit of the solutions using data from fewer stimuli by systematically removing the least distinctive and least typical faces to determine the minimum number of faces that were necessary to produce an MDS solution comparable to that based on the full stimulus set.

Results

The goodness-of-fit values for the two- to six-dimensional MDS solutions, measured by Kruskal's stress formula 1 (Kruskal & Wish, 1978), are shown in Fig. 2. Although there are no guidelines for assessing Kruskal's stress values (Giguere, 2006), the fits of the five-dimensional solution (stress = .21) and six-dimensional solution (stress = .18) are comparable to those reported previously as good fits in studies using MDS with face stimuli (e.g., stress = .14 in Johnston et al., 1997; stress = .19 in Lee et al., 2000; stress = .26 in Yotsumoto et al., 2007). We also examined whether the MDS solutions

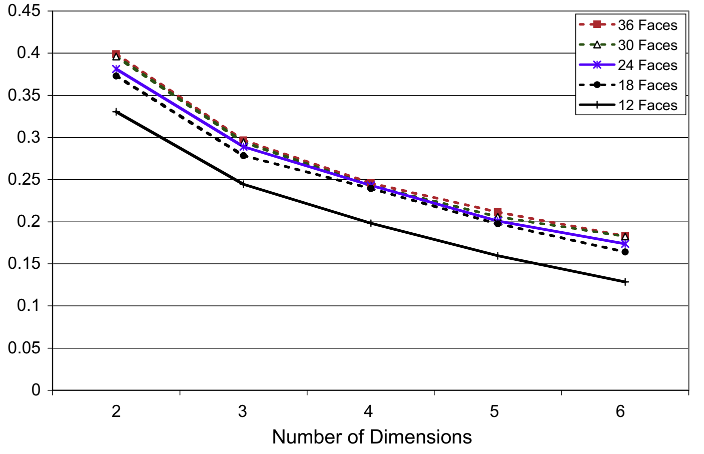

Fig. 2. Goodness-of-fit values (Kruskal's stress formula 1 [Kruskal & Wish, 1978]) of two- to six-dimensional solutions with varying numbers of face stimuli included in MDS (Experiment 1).

corresponded to the independent ratings of typicality, and *t* tests revealed that the distance to the origin for typical faces in each of the two- to six-dimensional MDS solutions was smaller than the distance for distinctive faces (Table 1). The effect sizes (Table 1) were large and comparable to those reported previously by Johnston et al. (1997), even though we had increased the homogeneity of our stimulus set by using the same hair on all faces.

To determine the smallest number of faces that would produce a comparable MDS representation of adults' similarity judgments, we conducted four additional sets of MDS analyses with smaller subsets of the face stimuli: 30, 24, 18, and 12 faces. The 3 least typical and 3 least distinctive faces were removed each time, so that in each subset half of the faces were the most typical faces and half were the most distinctive faces. The goodness-of-fit values of the two- to six-dimensional solutions for each subset are shown in Fig. 2. As can be seen from the figure, using a minimum of 18 or 24 faces produced a solution comparable to the solution using the full set of 36 face stimuli.

Discussion

The results showed that with additional dimensions, the fit of the MDS solution to adults' raw similarity data improved; however, with each additional dimension, the interpretability of the solutions decreased. The fits of the five- and six-dimensional solutions were comparable to those reported previously in studies of face-space (e.g., Johnston et al., 1997; Lee et al., 2000; Yotsumoto et al., 2007). The

Table 1
Comparison of the mean distances to the origin (and standard deviations) for distinctive faces versus typical faces for two- to six-dimensional MDS solutions based on adults' similarity ratings using a 7-point rating scale (Experiment 1).

	Number of dimensions				
	2	3	4	5	6
Mean distance to origin of distinctive faces	1.48	1.81	2.09	2.31	2.52
Standard deviation	(0.13)	(0.11)	(0.09)	(0.09)	(0.09)
Mean distance to origin of typical faces	1.32	1.63	1.90	2.15	2.37
Standard deviation	(0.09)	(0.08)	(0.07)	(0.08)	(0.06)
$t(23)$ =	3.56	4.56	5.60	5.03	5.05
p <	.01	.001	.001	.001	.001
Effect size (Cohen's d)	1.48	1.90	2.34	2.10	2.11
Effect size (Cohen's d) from data reported by Johnston et al. (1997)	1.84	2.05	2.06	1.64	1.97

Note. Standard deviations are in parentheses. Effect sizes (Cohen's *d*) are shown for the data from the current study as well as the data reported by Johnston et al. (1997) in a previous study using 36 male faces with individuating hair cues.

layout of the faces in the MDS solutions revealed that independently rated typical faces were closer to the origin than distinctive faces, as predicted by the face-space model and previous findings (e.g., Johnston et al., 1997). This result confirms that MDS is a useful method of analysis even when the homogeneity of the stimulus set is increased because all faces share the same hair. The nature of what the dimensions represent in the MDS solutions of Experiment 1 is discussed later together with the results from Experiment 2.

The results also suggested that using 18 to 24 face stimuli would be sufficient to produce MDS solutions comparable to those obtained with the full set of 36 faces. Because the fit of any MDS solution will improve with more dimensions relative to the number of stimuli used, we chose the more conservative method of using a minimum of 24 faces.

Experiment 2

The goal of Experiment 2 was to examine whether similarity judgments comparable to those obtained in Experiment 1 could be collected through the use of an odd-man-out paradigm rather than a rating scale and, if so, to compare 8-year-olds' and adults' representations using MDS. In an odd-man-out paradigm, observers are asked to choose the one item that looks most different from the rest of the array on each trial. No reference to a rating scale is required, making the method more suitable for children. The odd-man-out paradigm has been combined successfully with MDS in previous studies of adults' perception (e.g., Romney et al., 1993), including perceived similarity of facial identity (Yotsumoto et al., 2007). We minimized working memory load by presenting only 3 faces simultaneously, and the observer's task was to choose the face that looked most different from the other 2 faces in the array. Given the results of Experiment 1, we chose to use 25 face stimuli to have a balanced incomplete block design (Weller & Romney, 1988; Yotsumoto et al., 2007). This design allowed us to present the minimal number of necessary *triads* (arrays of 3 faces, e.g., A–B–C) to observers while ensuring that each face *pair* (e.g., A–B, A–C, B–C) was presented an equal number of times (a balanced incomplete block design is not possible with 24 stimuli). The 13 most distinctive and 12 most typical faces were chosen as the face stimuli because we aimed to provide sufficient variability in the internal facial characteristics among faces to reveal differences, if any, between children's and adults' similarity judgments.

We first collected similarity judgments from adults using the odd-man-out method (Experiment 2a) and verified the convergence of the MDS solutions between Experiments 1 and 2a. We then collected similarity judgments from 8-year-olds using the odd-man-out method (Experiment 2b) to compare children's and adults' similarity judgments and to evaluate the influence of typicality/distinctiveness on those judgments.

Method: Experiment 2a

Participants

Participants were 24 Caucasian undergraduate students (17 women and 7 men, 18–22 years of age, mean age = 19.90 years) participating for bonus credit in an introductory psychology course. All had normal or corrected-to-normal vision.

Stimuli

The 13 most distinctive and 12 most typical faces from Experiment 1 were chosen to form a balanced incomplete block design. Stimuli were 10.3×15.0 cm and were viewed from a distance of 100 cm ($5.90 \times 8.74^\circ$ of visual angle).

Apparatus

Stimuli were presented in color on a 19-inch monitor using custom software written with X code in Mac OS X.

Procedure

This experiment was approved by the research ethics board of McMaster University. All participants were informed of the nature of the task and provided written consent at the beginning of the session. Each participant was tested individually in a dimly lit room.

The procedure began by introducing the purpose of the study as a game in which "many women are looking for a partner to join a race." The observer's task was to help these women find their best partners "because partners who look alike are going to do the best in this race." To clarify the task, a demonstration trial showed 3 schematic faces on the screen (2 smiling faces and 1 frowning face) and the participant was asked to choose the face that appeared to be "most different."

Participants then performed six validity trials with photographs of vehicles of transportation to verify that they had understood the task of choosing the stimulus that was most different (i.e., odd-man-out method). On each trial, two photographs of the same vehicle type were shown (e.g., two transport trucks) and one photograph of a different vehicle (e.g., an airplane). Participants were asked to choose the object that was most different and were required to answer all six validity trials correctly to proceed to the next practice block. All participants passed this criterion on their first attempt.

Participants then completed six practice trials with faces. Face stimuli were chosen such that there was one obviously different face in each triad (e.g., two women's faces and one man's face), although each participant was free to choose the face that appeared to be most different to him or her (i.e., no feedback from the experimenter was given other than words of encouragement). On each trial, after the observer had chosen the most different face, the remaining two faces were shown in an animation with two people holding hands and forming a team to emphasize the purpose of the game.

Participants then completed 200 test trials. On each trial, 3 faces were shown in the same layout as practice trials, and participants were encouraged to look carefully at all 3 faces before choosing the face that appeared to be most different by either a verbal response or a mouse click (see Fig. 3). The formation of the triads was randomized across participants, with the constraint that each face pair was presented twice for every participant, because having only a single similarity judgment per face pair has been shown to produce unreliable results in a balanced incomplete block design (Burton & Nerlove, 1976). After every 15 to 30 trials, there was a "bonus" trial in which each participant was free to take a break as needed.

Data analysis

Data from the odd-man-out judgments were recoded such that, given a triad of A–B–C, if Face A was chosen as most different, Face Pairs A–B and A–C were given a distance score of 1 and Face Pair

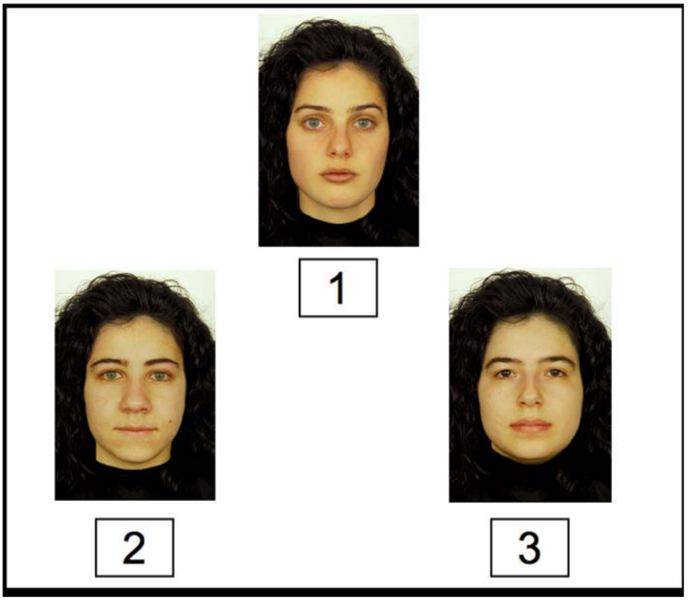

Fig. 3. An example trial of the odd-man-out paradigm. Face 1 = 3rd most distinctive (D3); Face 2 = 11th most distinctive (D11); Face 3 = 3rd most typical (T3).

B–C was given a score of 0 because similar faces should be closer together in face-space. The average distance for each face pair was calculated for each participant. Before performing MDS, we compared these distance values for each face pair with the raw similarity ratings from Experiment 1.

Using the distance values from each participant, MDS was performed using the INDSCAL procedure in SPSS software (as in Experiment 1). In addition, pairwise distance in the MDS solutions, equal to the square root of the sum of the intradimensional differences (Borg & Groenen, 2005), was calculated between all possible pairs of faces for the two- to six-dimensional solutions. These distances were compared with pairwise distances in the MDS solutions from Experiment 1.

To interpret what the MDS dimensions represent psychologically, for each MDS solution, the faces were ranked in terms of their distance from the origin along each dimension. For each dimension, the four most extreme faces at the positive end were morphed together and the four most extreme faces at the negative end were morphed together in an attempt to cancel out cues that were not diagnostic of that dimension. These morphed faces were then compared with each other and with the average face (a morph of all 25 faces) to infer the facial characteristic(s) represented by that dimension. The interpretability of these dimensions and the goodness-of-fit of the MDS solutions were used to choose the best MDS solution that captured adults' similarity judgments, which appeared to be the five-dimensional solutions. An independent group of adult observers viewed the faces at opposite ends of each dimension to characterize what the dimensions represent. Subsequently, cluster analysis (Sireci & Geisinger, 1992) was performed using SPSS (version 16.0) for Mac OS X on the coordinates of the faces in the five-dimensional MDS solutions to compare local clustering of faces in Experiments 1 and 2a.

Results

The correlation between the odd-man-out judgments and the raw similarity ratings from Experiment 1 was very high, $r(298) = .797$, $p < .001$. The goodness-of-fit values measured by Kruskal's stress formula 1 (Kruskal & Wish, 1978) for the MDS solutions with two to six dimensions are shown in Fig. 4 along with the goodness-of-fit values for the solutions from Experiment 1 (using only the same subset of 25 faces used in Experiment 2). The fits of the MDS solutions from the two experiments appear to be very similar (see overlapping lines in Fig. 4), a pattern again suggesting that similarity judgments collected through the use of a rating scale and the use of odd-man-out judgments are comparable. Five or six dimensions appear to account well for the similarity data.

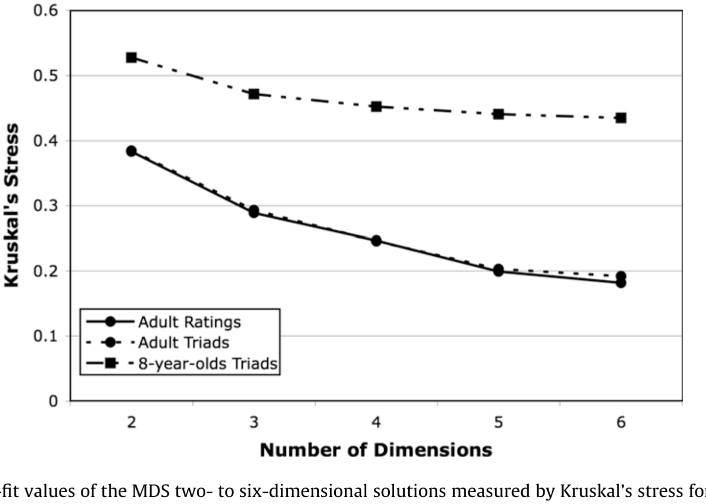

Fig. 4. Goodness-of-fit values of the MDS two- to six-dimensional solutions measured by Kruskal's stress formula 1 (Kruskal & Wish, 1978), for which smaller values indicate better fits. The lines represent fits to similarity judgments based on adults' ratings (Experiment 1), adults' odd-man-out judgments (Experiment 2a), and 8-year-olds' odd-man-out judgments (Experiment 2b).

The pairwise distances of faces in the MDS solutions of Experiment 2a (odd-man-out method) correlated highly with the pairwise distances in the corresponding MDS solutions of Experiment 1 (rating scale method) (see Table 2). The layouts of the two-dimensional solutions (for ease of representation) from Experiments 1 and 2a are shown in Fig. 5. The pattern of correlations, along with Kruskal's stress values, suggests that four to six dimensions account optimally for adults' similarity judgments.

Table 2
Correlation coefficients (Pearson's *r*) for the pairwise distances obtained from solutions using the rating scale method (Experiment 1) and the odd-man-out method (Experiment 2a).

	Number of dimensions				
	2	3	4	5	6
Correlation coefficient	.582	.774	.765	.752	.743
p =	.000	.000	.000	.000	.000

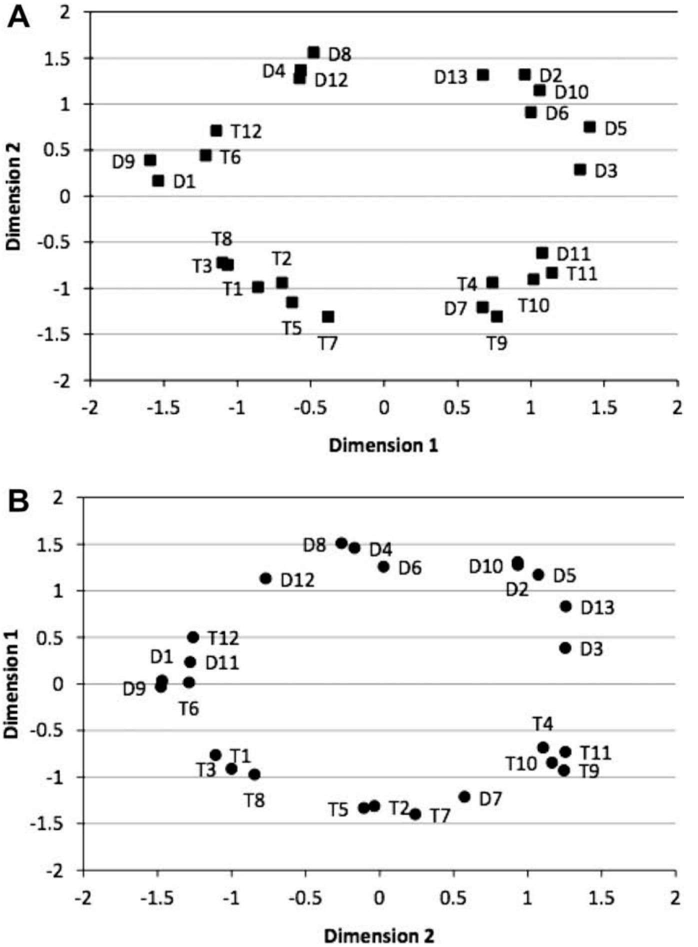

Fig. 5. Two-dimensional MDS solution of adults' similarity judgments of 25 faces based on pairs of faces using a 7-point rating scale (Experiment 1) (A) and odd-man-out decisions (Experiment 2a) (B).

As in Experiment 1, we also compared the distance to the origin for typical versus distinctive faces for each of the two- to six-dimensional MDS solutions. For every solution, typical faces were closer to the origin than distinctive faces (Table 3). Although the effect sizes were smaller than those observed in Experiment 1 and in a previous study (Johnston et al., 1997), they were still large by conventional standards (Cohen's $d > 0.70$). This finding suggests that a 7-point rating scale (Experiment 1) is more sensitive than the odd-man-out method (distances range only from 0 to 1) but that the odd-man-out judgments can still reveal robust typicality effects in the MDS solutions.

To examine what the MDS dimensions represent, and to determine the best solution, we morphed the faces with the most positive and most negative coordinates along each dimension in the four-, five-, and six-dimensional MDS solutions from Experiments 1 and 2a. On visual inspection of the morphs by the authors, the four- and six-dimensional solutions appeared to be less interpretable than the five-dimensional solution; therefore, subsequent analyses are based on the five-dimensional solutions. We asked an independent group of 24 observers (17 women and 7 men, 18–27 years of age, mean age = 20.80 years) to view these face morphs (Fig. 6A and B) and to describe the facial characteristics that best described what the dimensions represent (observers could report multiple cues). The results are summarized in Fig. 7A and B. As can be seen from Fig. 7, the variability in the descriptions within a dimension, as well as the overlap in the reported characteristics across dimensions, suggests that each dimension represents multiple facial characteristics. Nonetheless, there appears to be some overlap in the MDS dimensions from Experiments 1 and 2a. For example, the first two dimensions can be interpreted as coding mainly the eyes and lips/mouth.

Another method of comparing the MDS solutions based on adults' similarity ratings and odd-man-out judgments is to examine how the faces cluster together locally in the MDS solutions. Because it is impossible to visually inspect local clusters of faces in a five-dimensional solution, we conducted cluster analyses (e.g., Sireci & Geisinger, 1992) on the five-dimensional solutions to first determine which faces cluster together locally in the solutions from each of the two experiments. When faces were grouped so that single clusters contained no more than 4 faces (producing eight or nine clusters), 19 of 25 faces in the solution from Experiment 1 fell into the same clusters in Experiment 2a (Fig. 8A and B). This pattern of results suggests that the local clustering of faces (i.e., perceived similarity of groups of faces) was similar in the MDS solutions produced from adults' similarity ratings (Experiment 1) and odd-man-out judgments (Experiment 2a).

Discussion

The pattern of findings based on adults' odd-man-out judgments was consistent with the pattern from adults' similarity ratings collected in Experiment 1, verifying the validity of the odd-man-out method to obtain similarity data suitable for MDS. The raw similarity judgments collected through the two methods, as well as the pairwise distances based on the MDS solutions, all were highly correlated. The results from the cluster analysis revealed that similar faces clustered together in the MDS solutions from the two methods. Larger effects of typicality on the MDS solutions were observed for the similarity ratings, but robust effects were still observed in the odd-man-out data. These results

Table 3
Comparison of the mean distances (and standard deviations) to the origin for distinctive faces versus typical faces for two- to six-dimensional MDS solutions based on adults' similarity judgments made using the odd-man-out method (Experiment 2a).

	Number of dimensions				
	2	3	4	5	6
Mean distance to origin of distinctive faces	1.45	1.75	2.04	2.27	2.48
Standard deviation	(0.12)	(0.07)	(0.05)	(0.04)	(0.03)
Mean distance to origin of typical faces	1.37	1.71	1.96	2.20	2.42
Standard deviation	(0.08)	(0.05)	(0.05)	(0.05)	(0.04)
$t(23)$ =	1.88	2.00	3.67	4.06	3.89
p value	.07	.06	.001	.0001	.001
Effect size (Cohen's d)	0.78	0.83	1.53	1.69	1.62

Note. Standard deviations are in parentheses.

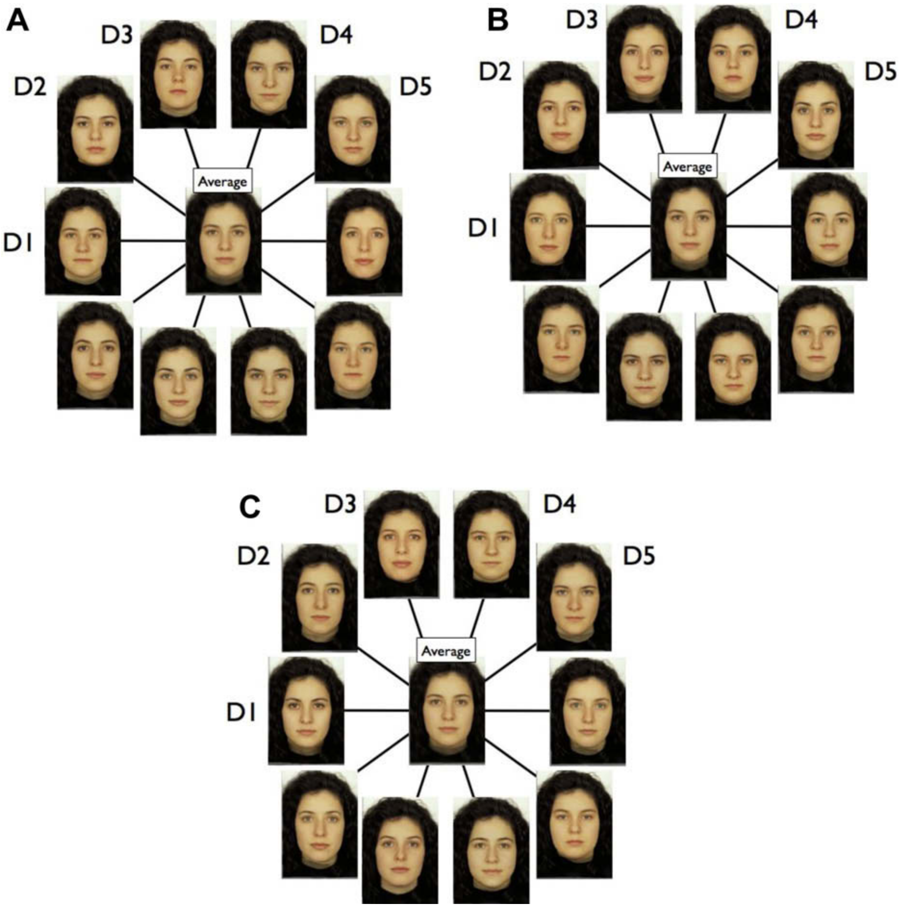

Fig. 6. A representation of the faces representing the most extreme coordinates along each of the five dimensions in the five-dimensional MDS solution for adults' ratings (A), adults' odd-man-out judgments (B), and 8-year-olds' odd-man-out judgments (C).

suggest that the odd-man-out method is a suitable method for collecting similarity judgments of facial identity. In Experiment 2b, we used the odd-man-out method to collect similarity judgments from 8-year-olds.

Method: Experiment 2b

Participants

A total of 24 Caucasian 8-year-olds (±3 months, 12 girls and 12 boys) with normal or corrected-to-normal vision participated in the study. Children were recruited from a database of names of mothers who had volunteered for developmental studies at the times of their children's births. Children were given snacks (cookies and/or juice) and stickers during breaks, as well as a toy (value of $1–2) from the toy box, to take home as a token of our appreciation at the end of the experiment. Parents did not receive any compensation other than free parking on campus during the experiment.

Stimuli and procedure

This experiment was approved by the research ethics board of McMaster University. Children provided informed verbal assent, and their parents provided informed written consent. The stimuli and

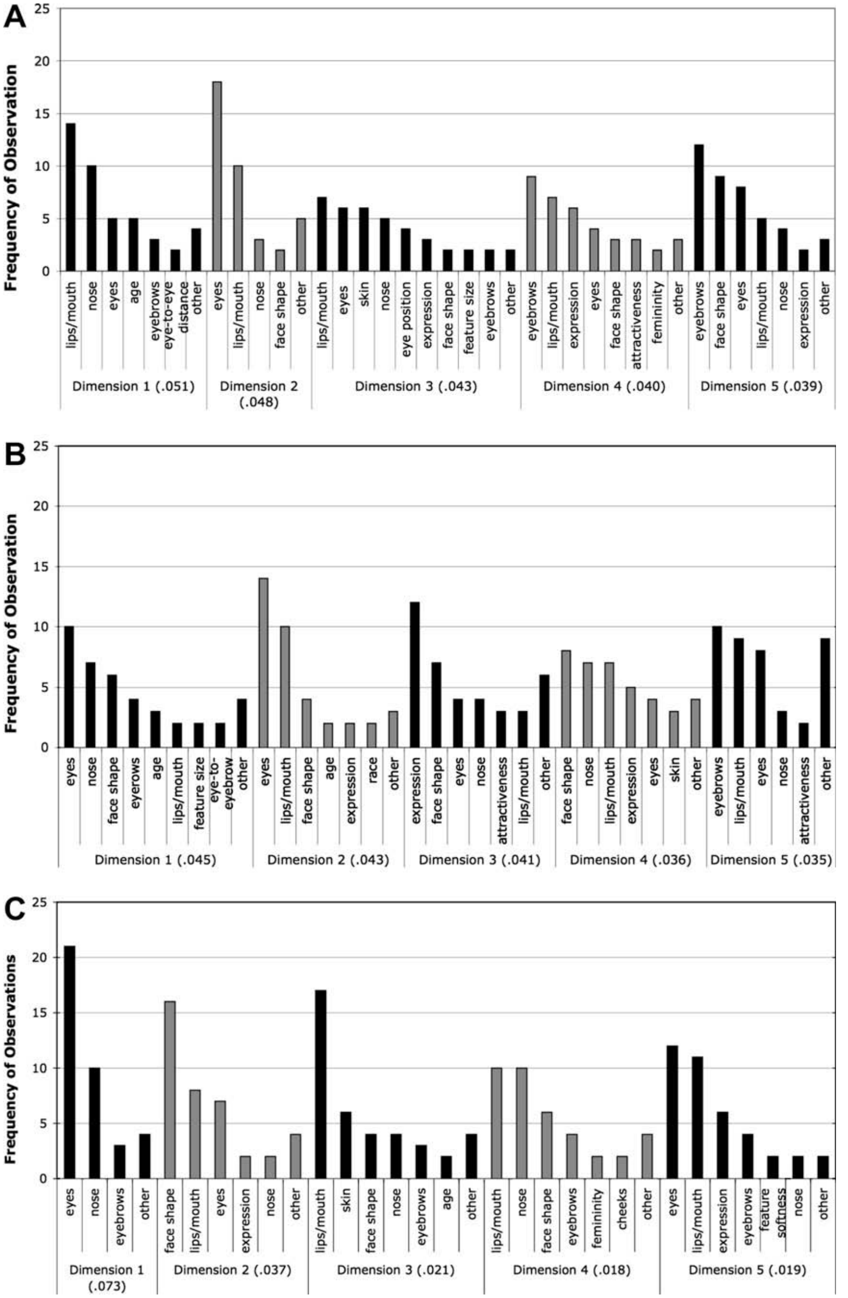

Fig. 7. (A) Independent reports of the facial characteristics that the dimensions likely represent in the MDS solution based on adult similarity ratings. The average weightings of the dimensions are shown in parentheses. (B) Independent reports of the facial characteristics that the dimensions likely represent in the MDS solution based on adult odd-man-out judgments. The average weightings of the dimensions are shown in parentheses. (C) Independent reports of the facial characteristics that the dimensions likely represent in the MDS solution based on 8-year-olds' odd-man-out judgments. The average weightings of the dimensions are shown in parentheses.

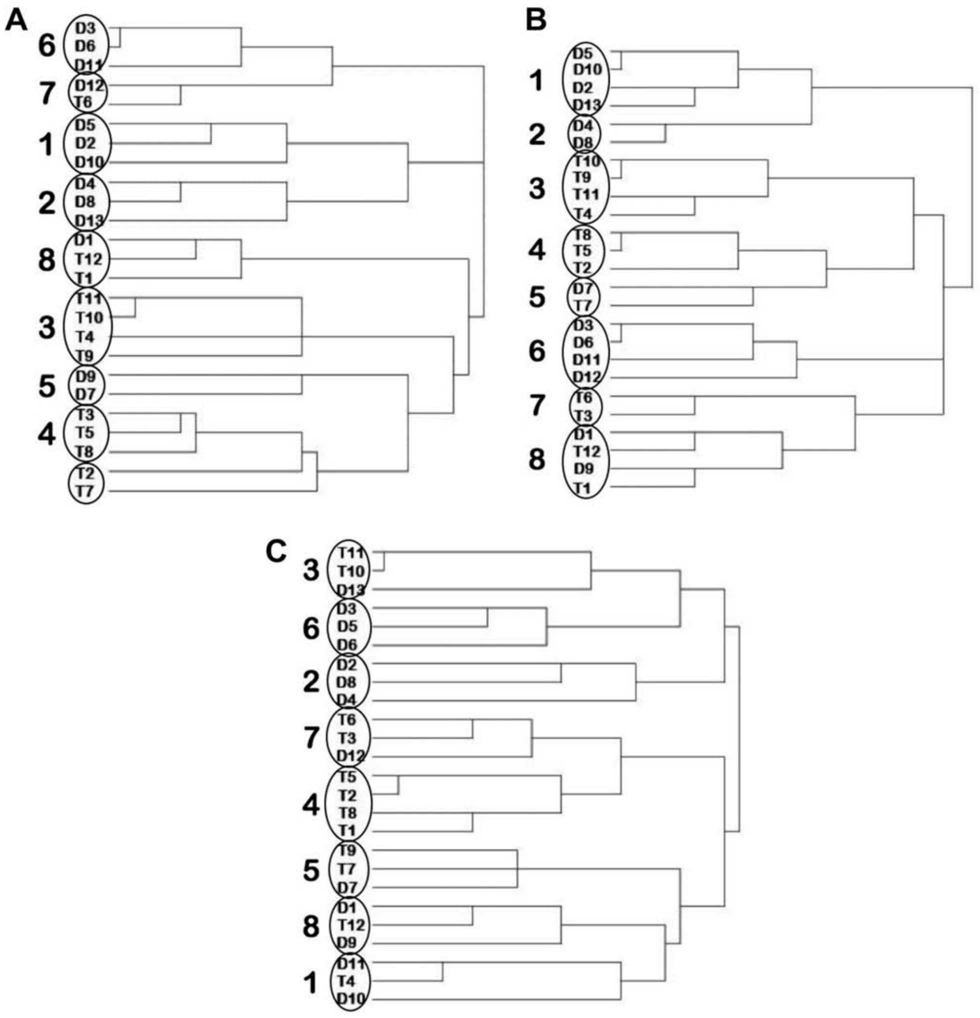

Fig. 8. Results of the cluster analysis using the coordinates of faces in the five-dimensional MDS solutions based on adults' similarity ratings (Experiment 1) (A), adults' odd-man-out judgments (Experiment 2a) (B), and 8-year-olds' odd-man-out judgments (Experiment 2b) (C). Local clusters were created by grouping together a maximum of four adjacent faces (shown in ovals) and were numbered 1 to 8 in the solution from adults' odd-man-out judgments (Fig. 7B) and compared with the corresponding local clusters in the solution for adults' similarity ratings and 8-year-olds' odd-man-out judgments.

procedure were identical to those in Experiment 2a except for the stickers and a snack break halfway through the session. All participants passed the criterion on their first attempt (i.e., choosing the most different vehicle correctly on all six validity trials).

Data analysis

Data analysis was performed in the same manner as in Experiment 2a.

Results

A comparison of the raw odd-man-out judgments made by 8-year-olds and adults revealed a significant correlation, $r(298) = .721$, $p < .01$. From the INDSCAL method, the goodness-of-fit values of the two- to six-dimensional solutions were assessed through Kruskal's stress formula (Kruskal & Wish, 1978) and are shown in Fig. 4. Similar to the adult data, five or six dimensions appeared to account best for children's similarity judgments. However, the stress values for 8-year-olds' solutions were relatively high—higher than any of the two- to six-dimensional MDS solutions for adults tested with the same method (Experiment 2a) and higher than what have been reported in previous studies of adult face perception (e.g., Johnston et al., 1997; Lee et al., 2000; Yotsumoto et al., 2007).

Despite children's higher stress values, the pairwise distances in the MDS solutions all were highly correlated with those of the adult MDS solutions (Table 4). The distance to the origin in children's MDS solutions was smaller for typical faces than for distinctive faces, similar to the results from adults' odd-man-out judgments in Experiment 2a, with comparable effect sizes (cf. Tables 3 and 5). However, the standard deviations of the distances to the origin appear to be higher (Table 5) than those observed in the adult MDS solutions (Table 3), suggesting greater variability in the locations of typical and distinctive faces in the MDS solutions of children than in those of adults.

Fig. 6C shows the morphed faces representing the dimensions of the five-dimensional MDS solution (face morphs based on four- and six-dimensional solutions were less interpretable, similar to the adult data). The results of the cluster analysis based on the five-dimensional MDS solution revealed that 16 of 25 faces in the 8-year-olds' solution fell into the same clusters as in the adult solution from Experiment 2a (Fig. 8C). This finding is similar to the comparison of the clusters between the two adult groups and reveals that local clustering of faces in the MDS solution from 8-year-olds was similar to that from adults.

The same group of observers as in Experiment 2a viewed the morphs of the most extreme faces along each dimension for the children (Fig. 6C) during the same session and verbally characterized each dimension. The results are summarized in Fig. 7C and show that many of the same facial characteristics observed in the adult solutions also appear to describe the dimensions in the children's MDS solution. One striking difference is the high agreement among observers that Dimension 1 of the children's solution appears to code for the eyes. Importantly, the average weightings of the dimensions reveal that children relied very heavily on this first dimension and relied much less on the other dimensions, whereas adults' use of various dimensions was more equally distributed. To examine this age difference further, we compared individual children's and adults' use of the dimensions when only two dimensions were imposed on the similarity data. Fig. 9 reveals that adults' weightings (Experi-

Table 4
Correlation coefficients for the pairwise distances from MDS solutions based on odd-man-out judgments by adults and 8-year-olds (Experiment 2).

	Number of dimensions				
	2	3	4	5	6
Correlation coefficient	.623	.681	.539	.606	.581
p =	.000	.000	.000	.000	.000

Table 5
Comparison of the mean distances (and standard deviations) to the origin for distinctive faces versus typical faces for two- to six-dimensional MDS solutions based on 8-year-olds' similarity judgments (Experiment 2b).

	Number of dimensions				
	2	3	4	5	6
Mean distance to origin of distinctive faces	1.47	1.80	2.09	2.33	2.57
Standard deviation	(0.18)	(0.20)	(0.26)	(0.26)	(0.26)
Mean distance to origin of typical faces	1.34	1.65	1.87	2.11	2.29
Standard deviation	(0.10)	(0.10)	(0.13)	(0.14)	(0.15)
$t(23)$ =	2.27	2.35	2.58	2.59	3.19
p <	.05	.05	.05	.05	.01
Effect size (Cohen's d)	0.94	0.98	1.08	1.08	1.33

Note. Standard deviations are in parentheses.

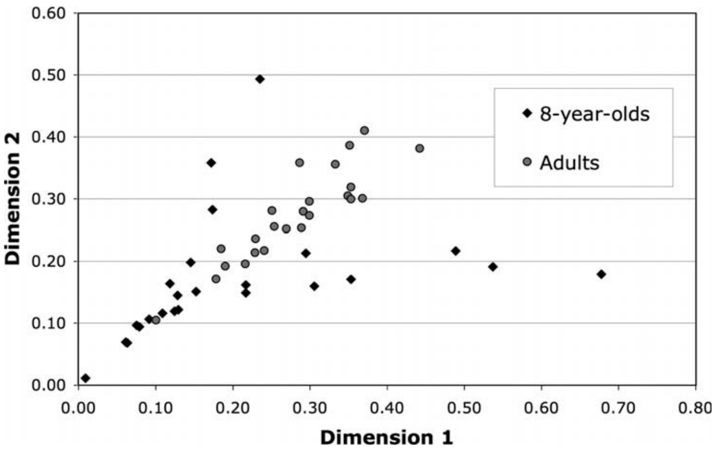

Fig. 9. Individual weightings of the two dimensions produced in the group INDSCAL two-dimensional solution for 8-year-olds and adults.

ment 2a) cluster around the diagonal, which represents roughly equal use of the two dimensions by the adult group. However, the weightings of roughly half of the 8-year-olds are biased heavily toward a single dimension. Moreover, for many 8-year-olds, the weightings are close to the origin. The clustering around the origin likely represents children switching their use of the dimensions across trials and/or relying on a different dimension(s) that was not captured in the group solution (especially the one child whose weightings for both dimensions are near zero). The variability and low weightings of the 8-year-olds confirm the poor fit of the MDS solutions to the raw data indicated by Kruskal's stress values (Fig. 4).

If the low weightings of the MDS dimensions demonstrated by roughly half of the 8-year-olds (Fig. 9) represent inconsistencies in children's use of the dimensions across trials, we should observe greater variability in children than in adults both across and within individuals. Interindividual variability was assessed by calculating the standard deviations of all pairwise similarity distances in the raw odd-man-out judgments from 8-year-olds and adults (each individual rated every pair twice). A paired-samples t test revealed that adults' standard deviations were smaller than those of 8-year-olds, $t(299) = 4.75$, $p < .001$, suggesting that there was more variability between 8-year-olds than between adults in their similarity judgments. Intraindividual variability was assessed by comparing how often 8-year-olds and adults chose the same pair as being most different when it was presented a second time with a different third face. As a group, 8-year-olds were more likely to change their answers on the second presentation than adults, $t(46) = 2.83$, $p < .01$.[1] This finding suggests that children were more variable in their odd-man-out choices than adults, consistent with the interpretation that children with low weightings on the MDS dimensions were using different dimensions on different trials.

Discussion

The results from Experiment 2 demonstrate a number of similarities between the judgments of 8-year-olds and those of adults. First, children's similarity judgments from an odd-man-out paradigm were highly correlated with adults' judgments using the same method. Second, a comparison of the clustering of faces in the MDS solutions revealed that 8-year-olds perceived the similarity of facial

[1] Consistency in responses at the individual level cannot be assessed adequately with the current method because although all observers saw each face pair twice, the third face forming the triad differed for every individual. For example, we cannot meaningfully compare whether Participant 1 was more consistent than Participant 2 because the two participants did not observe face pair A–B relative to the same third face. However, if adults and 8-year-olds were responding in a similar manner, we should not observe any difference at the group level because triads were formed randomly across participants in each age group.

identity in a manner similar to that of adults even when salient hair cues could not be used. These findings are consistent with previous reports suggesting that by 7 years of age children have a mental representation of facial identity similar to that of adults (e.g., Pedelty et al., 1985) and that by 8 years of age this representation is centered on an average face or norm (Nishimura et al., 2008). Third, the MDS solutions based on children's and adults' odd-man-out judgments revealed that typical faces, rated by an independent group of adult observers, were located closer to the origin than distinctive faces, a finding that is also consistent with previous research (e.g., Johnston et al., 1997) and suggests that by 8 years of age children perceive the typicality/distinctiveness of faces in a manner similar to that of adults (Chang et al., 2002; Johnston & Ellis, 1995). These findings suggest that the spatial layout of children's mental representation is similar to that of adults in some respects.

The most striking difference between the MDS solutions of 8-year-olds and adults was the fact that no MDS solution fit the raw data of 8-year-olds as well as those of adults (see Fig. 4). The poor fit may be accounted for, in part, by the greater variability in the odd-man-out choices made by 8-year-olds than by adults, which was revealed in the analyses of within- and between-participants variability. In addition, the average weightings (Fig. 7) of 8-year-olds' MDS dimensions revealed that children relied heavily on a single dimension, likely eye color, when making similarity judgments, whereas adults used all five dimensions more equally, a finding that is consistent with that of Pedelty et al. (1985). This age difference was not revealed by simply correlating the similarity judgments of 8-year-olds and adults and demonstrates the utility of MDS in examining subtle developmental differences. The interpretation that children relied heavily on eye color as a cue when making similarity judgments is consistent with previous research showing children's reliance on a single face attribute when categorizing faces (Schwarzer, 2000) and children's nearly adult-like sensitivity to changes in feature shapes by 8 years of age (Mondloch et al., 2002). Conversely, unlike the adult dimensions, there was no clear indication of a dimension in children's five-dimensional solution that coded the spatial position of facial features such as eye-to-eye distance and eye-to-eyebrow distance (see Fig. 7), consistent with the slow development of sensitivity to the spatial relations among internal facial features (Mondloch, Le Grand, & Maurer, 2002, 2003). Therefore, some of the immaturities in face processing reported in previous research may be attributed to the lack of dimensions that code the spatial relations of internal features in children's mental representation of facial identity.

In the current paradigm, each face pair was viewed twice but in reference to a different third face, and these triads differed across participants. Therefore, we could not assess the extent to which the variability observed in children's responses reflected general inconsistencies and/or the effect of triads (i.e., context). A future study could use the same triads for all observers to examine interindividual differences and repeat the triads to assess intraindividual inconsistencies. The heavy reliance by children on a single dimension may have contributed to the trial-by-trial variability in their responses. For example, when shown two faces with brown eyes and one face with blue eyes, children may rely solely on eye color to choose the odd-man-out, but when shown three faces with blue eyes, they may rely solely on another characteristic such as face shape.

The results are also consistent with the interpretation that children's face-space is more flexible either because children shift among adult-like dimensions or because they use a wider range of dimensions but with variable weightings that determine which dimension is used for a particular judgment. A more flexible face-space during development is consistent with the finding that when Asian children under 9 years of age move from an all-Asian environment to an all-Caucasian environment, as adults they demonstrate better recognition of Caucasian faces than of Asian faces, much like Caucasian adults (Sangrigoli et al., 2005). In contrast, adults may have a more rigid face-space and use the same amalgam of characteristics for every trial. This interpretation predicts that children's MDS solutions will vary more than adults' solutions across studies if different stimulus sets are used. Indeed, the MDS solutions from 7- to 11-year-olds appeared to be more similar to those of adults when hair cues were present (Pedelty et al., 1985) than when they were removed as in the current study. Future studies could compare consistencies in the MDS solutions for children and adults across face sets differing systematically in homogeneity, a variable that has been shown to affect adults' recognition memory for unfamiliar faces (Yotsumoto et al., 2007). For example, children's similarity judgments based only on typical faces (presumably sampling a small area of face-space) versus only distinctive faces (presumably sampling a wide area of face-space) should differ more than adults' judgments. It is unclear, however, whether children would appear to be more like adults when judging only typical or only distinctive faces because not all distinctive faces differ from the average in the same way, making different facial characteristics salient on different trials.

Collectively, the current findings suggest that 8 years of visual experience with faces is sufficient to set up a coarse mental representation of faces that is similar to that of adults but insufficient to establish a stable face-space that uses multiple dimensions equally and simultaneously. Children's reliance on a few salient features to make perceptual judgments does not appear to be limited to judgments of faces given that it has also been shown with house stimuli (Vurpillot, 1968) and, more recently, with highly similar novel objects that differed in both the shape of a salient part and its position (Mash, 2006). Therefore, part of what develops in children's face-space beyond 8 years of age may include general improvements in visual processing and cognitive capacities such as memory, attention, and problem solving.

In summary, the current study used a novel child-friendly paradigm to compare children's and adults' face-space. By creating a homogeneous set of face stimuli with the same hair, subtle differences between children's and adults' processing of facial identity could be examined because salient hair cues that could lead to an overestimation of the similarity between children's and adults' judgments were removed. In addition, we provided evidence to support the utility of an odd-man-out paradigm in obtaining similarity judgments instead of a rating scale that is quite abstract. The findings from this novel paradigm suggest that the layout of 8-year-olds' mental representation of facial identity is similar to that of adults but that children tend to favor the eyes as cues to identity and are more variable than adults in their use of the dimensions. This variability may contribute to immature performance on face processing tasks at 8 years of age.

References

Anderson, N. D., & Wilson, H. R. (2005). The nature of synthetic face adaptation. *Vision Research, 45*, 1815–1828.

Aylward, E. H., Park, J. E., Field, K. M., Parsons, A. C., Richards, T. L., Cramer, S. C., et al (2005). Brain activation during face perception: Evidence of a developmental change. *Journal of Cognitive Neuroscience, 17*, 308–319.

Baenninger, M. (1994). The development of recognition: Featural or configurational processing? *Journal of Experimental Child Psychology, 57*, 377–396.

Baudouin, J. Y., & Gallay, M. (2006). Is face distinctiveness gender based? *Journal of Experimental Psychology: Human Perception and Performance, 32*, 789–798.

Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2nd ed.). New York: Springer Science & Business Media.

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.

Brooks, K. R., & Kemp, R. I. (2007). Sensitivity to feature displacement in familiar and unfamiliar faces: Beyond the internal/external feature distinction. *Perception, 36*, 1646–1659.

Bruce, V., Burton, A. M., & Dench, N. (1994). What's distinctive about a distinctive face? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology A, 47*, 119–141.

Burton, M. L., & Nerlove, S. B. (1976). Balanced designs for triad tests: Two examples from English. *Social Science Research, 5*, 247–267.

Busey, T. A. (1998). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science, 9*, 476–483.

Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research, 38*, 2455–2468.

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*, 253–274.

Chang, P. P. W., Levine, S. C., & Benson, P. J. (2002). Children's recognition of caricatures. *Developmental Psychology, 38*, 1038–1051.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893–909.

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of Experimental Child Psychology, 96*, 57–70.

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

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.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*, 159–170.

Gathers, A. D., Bhatt, R., Corbly, C. R., Farley, A. B., & Joseph, J. E. (2004). Developmental shifts in cortical loci for face and object recognition. *NeuroReport, 15*, 1549–1553.

Giguere, G. (2006). Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS. *Tutorials in Quantitative Methods for Psychology, 2*, 27–38.
Golari, G., Ghahremani, D. G., Whitfield-Gabrieli, W., Reiss, A., Eberhardt, J. L., Gabrieli, J. D. E., et al (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience, 10*, 512–522.
Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception, 13*, 505–512.
Itier, R. J., & Taylor, M. J. (2004a). Face inversion and contrast-reversal effects across development: In contrast to the expertise theory. *Developmental Science, 7*, 246–260.
Itier, R. J., & Taylor, M. J. (2004b). Face recognition memory and configural processing: A developmental ERP study using upright, inverted, and contrast-reversed faces. *Journal of Cognitive Neuroscience, 16*, 487–502.
Johnston, R. A., & Ellis, H. D. (1995). Age effects in the processing of typical and distinctive faces. *Quarterly Journal of Experimental Psychology A, 48*, 447–465.
Johnston, R. A., Milne, A. B., Williams, C., & Hosie, J. A. (1997). Do distinctive faces come from outer space? An investigation of the status of a multidimensional face-space. *Visual Cognition, 4*, 59–67.
Kahana, M. J., & Bennett, P. J. (1994). Classification and perceived similarity of compound gratings that differ in relative spatial phase. *Perception and Psychophysics, 55*, 292–304.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.
Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.
Lee, K., Byatt, G., & Rhodes, G. (2000). Testing the face-space framework. *Psychological Science, 11*, 379–385.
Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*, 89–94.
Levin, D. T. (1996). Classifying faces by race. The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1364–1382.
Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face aftereffects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B, 272*, 2283–2287.
Martens, W. L., Zacharov, N. (2000). Multidimensional perceptual unfolding of spatially processed speech: I. Deriving stimulus space using INDSCAL. Paper presented at the meeting of the Audio Engineering Society, Los Angeles. Available from: <http://www.aes.org/e-lib/browse.cfm?elib=9114>.
Mash, C. (2006). Multidimensional shape similarity in the development of visual object classification. *Journal of Experimental Child Psychology, 95*, 128–152.
Miller, K., & Gelman, R. (1983). The child's representation of number: A multidimensional scaling analysis. *Child Development, 54*, 1470–1779.
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 to the spacing of facial features. *Journal of Experimental Child Psychology, 89*, 159–181.
Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology, 86*, 67–84.
Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553–566.
Mondloch, C., Le Grand, R., & Maurer, D. (2003). Early visual experience is necessary for the development of some––but not all––aspects of face processing. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood: Current perspectives* (pp. 99–117). New York: Nova Science.
Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adultlike holistic face processing. *Visual Cognition, 15*, 564–577.
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.
Passarotti, A., Paul, B., Bussiere, J., Buxton, R., Wong, E., & Stiles, J. (2003). The development of face and location processing: An fMRI study. *Developmental Science, 6*, 100–117.
Passarotti, A., Paul, B., & Stiles, J. (2001). Development affects the ventral and dorsal processing streams differently: An fMRI study on face and location processing in children, teenagers, and adults. *NeuroImage, 13*, S345.
Pedelty, L., Levine, S. C., & Shevell, S. K. (1985). Developmental changes in face processing: Results from multidimensional scaling. *Journal of Experimental Child Psychology, 39*, 421–436.
Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14*, 618–622.
Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science, 9*, 270–277.
Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*, 25–57.
Rhodes, G., Byatt, G., Tremewan, T., & Kennedy, A. (1997). Facial distinctiveness and the power of caricatures. *Perception, 26*, 207–223.
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 aftereffects. *Psychological Science, 14*, 558–566.
Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science, 4*, 28–34.
Sangrigoli, S., Pallier, C., Argenti, A-M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*, 440–444.
Scherf, K. S., Berhmann, M., Humphreys, K., & Luna, B. (2007). Visual category selectivity for faces, places, and objects emerges along different developmental trajectories. *Developmental Science, 10*, F15–F30.
Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child Development, 71*, 391–401.

Sireci, S. G., & Geisinger, K. F. (1992). Analyzing test content using cluster analysis and multidimensional scaling. *Applied Psychological Measurement, 16*, 17–31.

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

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

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *NeuroReport, 12*, 1671–1676.

Turati, C., Sangrigoli, S., Ruel, J., & de Schonen, S. (2004). Evidence of the face-inversion effect in 4-month-old infants. *Infancy, 6*, 275–297.

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., & Bruce, V. (1986). The effects of distinctiveness in recognizing and classifying faces. *Perception, 15*, 525–535.

Vurpillot, E. (1968). The development of scanning strategies and their relation to visual differentiation. *Journal of Experimental Child Psychology, 6*, 632–650.

Want, S. C., Pascalis, O., Coleman, M., & Blades, M. (2003). Recognizing people from the inner or outer parts of their faces: Developmental data concerning "unfamiliar" faces. *British Journal of Developmental Psychology, 21*, 125–135.

Watson, T. L., & Clifford, C. W. G. (2003). Pulling faces: An investigation of the face-distortion aftereffect. *Perception, 32*, 1109–1116.

Watson, T. L., Rhodes, G., Clifford, C. W. G. (2006). Norm-based coding of faces adapts to help us identify those around us. Poster presented at the meeting of the Vision Sciences Society, Sarasota, FL.

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

Weller, S. C., & Romney, A. K. (1988). *Systematic data collection (Qualitative Research Methods, Vol. 10)*. Newbury Park, CA: Sage.

Yin, R. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Yotsumoto, Y., Kahana, M., Wilson, H. R., & Sekuler, R. (2007). Recognition memory for realistic synthetic faces. *Memory and Cognition, 35*, 1233–1244.

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