Cognitive Development 48 (2018) 1–18

Contents lists available at ScienceDirect

Cognitive Development

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

The development of face expertise: Evidence for a qualitative change in processing

Peter J. Hills[a,*], Michael B. Lewis[b]

[a] Bournemouth University, Department of Psychology, Talbot Campus, Fern Barrow, Poole, Dorset, BH12 5BB, UK
[b] School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, UK

ARTICLE INFO

Keywords:
Face recognition
Development
Configural processing
Face-inversion effect

ABSTRACT

There is conflicting evidence regarding the development of expert face recognition, as indexed by the face-inversion effect (FIE; de Heering, Rossion, & Maurer, 2011; Young and Bion, 1981) potentially due to the nature of the stimuli used in previous research. The developmental trajectory of the FIE was assessed in participants aged between 5- and 18-years using age-matched and adult stimuli. Four experiments demonstrated that upright face recognition abilities improved linearly with age (presumably due to improved memory storage capacities) and this was larger than for inverted faces. The FIE followed a stepped function, with no FIE for participants younger than 9-years of age. These results indicate maturation of expert face processing mechanisms that occur at the age of 10-years, similar to expertise in other domains.

1. Introduction

While adult face recognition is one of the most impressive human visual skills given the ability to differentiate and recognise many thousands of faces (Ellis, 1986), the face recognition abilities of children are poorer (Adams-Price, 1992; Blaney & Winograd, 1978). Adult face processing is assumed to be based on some form of expert processing mechanism (Farah, Wilson, Drain, & Tanaka, 1998) that may well be specific to the processing of faces (Kanwisher, Tong, & Nakayama, 1998). Poorer face-recognition performance in children could be due to generally poorer cognitive, attentional, and perceptual systems (see e.g., Crookes & McKone, 2009) or a specific deficit in this expert face processing (e.g., Carey & Diamond, 1977).

Expert processing is typically referred to as configural processing and is made up of three components (Maurer, Le Grand, & Mondloch, 2002): processing of the first order relations (i.e., two eyes level, side-by-side and above the nose); the processing of second-order relational information (i.e., idiosyncratic deviations to the basic template; Carey & Diamond, 1994); and holistic processing, which is processing the face as a gestalt whole (Rossion, 2008), integrating the multiple sources of information (Farah et al., 1998; Searcy & Bartlett, 1996). While researchers may not entirely understand what drives expertise in face recognition, there is consensus that faces are processed differently to objects and this is likely due to some form of configural processing (Piepers & Robbins, 2012). There are many sources of evidence to suggest that expert processing is not based on second-order relational information (Burton, Schweinberger, Jenkins, & Kaufmann, 2015) but is based on this final form of configural processing, known as holistic processing (Hole, George, Eaves, & Rasek, 2002; Mondloch & Desjarlais, 2010). Configural processing is usually contrasted with the featural coding, which is not indicative of expertise. Featural coding is typically defined as the processing of individual features in isolation (see Cabeza & Kato, 2000; Tanaka & Sengco, 1997). One method typically employed to assess configural coding is that of inversion (e.g., Freire, Lee, & Symons, 2000). Indeed, Sergent (1984) suggests that configural encoding is what is disrupted by

* Corresponding author.
E-mail addresses: phills@bournemouth.ac.uk (P.J. Hills), LewisMB@Cardiff.ac.uk (M.B. Lewis).

https://doi.org/10.1016/j.cogdev.2018.05.003
Received 29 September 2016; Received in revised form 28 May 2018; Accepted 31 May 2018
0885-2014/ © 2018 Published by Elsevier Inc.

inversion, whereas featural encoding is far less disrupted by inversion (see also Lewis & Glenister, 2003). Therefore, the face-inversion effect (FIE) is a reliable index of expert face processing (Edmonds & Lewis, 2007; Gauthier et al., 2000; Yin, 1969).

While there is no doubt that face recognition is expert in adults, there is a debate about when this expertise develops. One theory suggests that there is an early development of expert face processing mechanisms complete by the age of approximately 5-years (Crookes & McKone, 2009; Gilchrist & McKone, 2003; Want, Pascalis, Coleman, & Blades, 2003). While face recognition improves with age, this view suggests that age-related improvements in face recognition are explained by general improvements in the ability to attend and focus on the demands of the task (Crookes & McKone, 2009). These general improvements increase with age and continue to develop throughout childhood and adolescence (Betts, McKay, Maruff, & Anderson, 2006; Pastò & Burack, 1997; Skoczenski & Norcia, 2002). An alternative view is that the expert processing mechanisms do not develop until around 10 years of age (Carey & Diamond, 1977, 1994), consistent with the notion that many forms of perceptual expertise take approximately 10 years of practice and development (Akhtar & Enns, 1989; Brodeur & Enns, 1997; Enns & Brodeur, 1989; Ericsson, Krampe, & Tesch-Römer, 1993; Pearson & Lane, 1991).

Recently, a view was put forward that there might be differential effects for the development of face perception and face memory, with face memory developing late and face perception developing early (Weigelt et al., 2013). Wiegelt et al. have presented evidence highlighting that the mechanisms that control expert face processing are not necessarily the same as expert face memory. Memory for faces, apparently, develops later than the perceptual expertise for faces. Memory for faces can be revealed through an increase in hit rate and response bias (as hit rate represents more faces being stored in memory and more efficient encoding) without affecting false alarm rate (which better reflects poorer encoding and poorer access to memory: Hills, 2012). General memory (Chi, 1977; Dempster, 1981; Kail, 1992) and memory for faces (Flin, 1980), does improve with increased age.

Consistent with the view that the FIE is a measure of expert face perception, then there should be sufficient evidence to establish whether face perception develops early or late. If face perception expertise develops late, then one would expect that children would show a smaller FIE than adults. The evidence for this is mixed. Most authors agree that face recognition abilities improve approximately linearly with age, reaching an asymptote at the age of 12 (Feinman & Entwisle, 1976), 17 years (Ellis, Shepherd, & Bruce, 1973; Golarai et al., 2007; Lawrence et al., 2008; O'Hearn, Schroer, Minshew, & Luna, 2010), or well into adulthood (e.g., Germine, Duchaine, & Nakayama, 2011; Susilo, Germine, & Duchaine, 2013) depending on the stimuli set used.[1] However, Flin (1980, 1985) has reported a face recognition performance "dip" at age 11 years[2] (see also, Carey, 1978, 1981; Carey, Diamond, & Woods, 1980). The improvement in recognition for inverted faces may also be linear, but at a slower rate. Using a novel (for this field) statistical procedure, de Heering, Rossion, and Maurer (2012) found that performance on the Benton Face Recognition Test (Benton, Sivan, Hamsher, Vareny, & Spreen, 1983) correlated with age, between the ages of 6-years and 12-years. This correlation was stronger for upright than inverted faces indicating that the magnitude of the FIE also correlated with age. Such an improvement for upright faces over inverted faces potentially reflects that the expert face processing system has developed and there is a general improvement in task performance or face memory. Alternatively, this improvement may reflect a protracted development of expert face processing skills.

Data from matching tasks reveal that children younger than 10 years of age are more likely to be affected by paraphernalia and pose changes than children older than 10 years and adults (Diamond & Carey, 1977; Ellis, 1992a, 1992b; Freire & Lee, 2001; Saltz & Sigel, 1967). These results indicate that children are not coding faces in the most effective configural manner. Indeed, six- and eight-year-old children do not show the FIE when tested in matching paradigms (Carey & Diamond, 1977; Hay & Cox, 2000; Joseph et al., 2006; Schwarzer, 2000) or recognition paradigms (Goldstein, 1975) indicating a greater reliance on featural processing (Schwarzer, 2000). In these studies, the FIE was found by some ten-year-old participants indicating some individual difference in the development of expert face processing which may sometimes mask effects when development is tested cross-sectionally. These results indicate a qualitative shift in the way children code faces at age 10 from an inexpert to expert mechanism (Baudouin, Gallay, Durand, & Robichon, 2010; Mondloch, Leis, & Maurer, 2006).

However, other authors have reported that the FIE is apparent in three- (Carey, 1981), five- (Fagan, 1972; Flin, 1983), or seven-year-old children (Young and Bion, 1981, 1982) leading to parallel improvements in recognition skills (Itier & Taylor, 2004). Proponents of the view that the FIE does not increase with age highlight that the studies that fail to show an FIE in younger participants suffer from floor effects (Young and Bion, 1981). Nevertheless, an age-by-orientation interaction is often found in studies that claim there is an FIE in younger participants,[3] indicating that the magnitude of the FIE increases with age (Brace et al., 2001; Carey, 1981; Carey & Diamond, 1994; Flin, 1983; Goldstein & Chance, 1964). Any effect of age on the magnitude of the FIE would indicate that children rely more on featural rather than configural coding (e.g., Hay & Cox, 2000).[4]

There are a number of methodological and statistical issues with the studies on children's face recognition. Firstly, most of the

[1] Ellis et al. (1973), Germine et al. (2011), and Susilo et al. (2013) used adult faces in their experiments, whereas Feinman and Entwisle (1976) used children's faces. These results are consistent with the own-age bias in face recognition (Anastasi & Rhodes, 2006) rather than demonstrating the developmental trajectory of face recognition.

[2] However, if a linear function is fitted to Flin's data, the *d'* measure does not show a significant blip at age 11 years. A blip in performance is observed only if a curvilinear function is fitted.

[3] For example, the significant age-by-orientation effect observed by Flin (1985) was put down to the recognition "blip" rather than a change in the magnitude of the FIE with age, even though the FIE was 1.5 times larger in her oldest age group than the youngest age group she tested.

[4] Using the parts and wholes test (Tanaka & Farah, 1993), Pellicano and Rhodes (2003) found evidence that children as young as four-years old use holistic processing. Similarly, Carey and Diamond (1994) have shown that 6- and 10-year-old show a composite face effect of a similar magnitude to adults. However, performance on these tasks may be unrelated (Konar et al., 2010; Rezlescu, Susilo, Wilmer, & Caramazza, 2017; Wilhelm et al., 2010), at least in children, to performance in the FIE.

studies conducted on face perception are cross-sectional. There are significant individual differences in face recognition ability (Li et al., 2010), in the amount of holistic processing participants engage in (Wang, Li, Fang, Tian, & Liu, 2012), and in terms of how faces are encoded (Bobak, Parris, Gregory, Bennetts, & Bate, 2017; Mehoudar, Arizpe, Baker, & Yovel, 2014). This means that, potentially, effects reported in the literature are due to cohort effects which may be unduly influenced by individual differences in studies with relatively small sample sizes. For example, the recognition blip observed by Flin (1985) and the change in FIE observed by Schwarzer (2000) may reflect cohort effects. A longitudinal study of face perception exploring the development of expert processing mechanisms has yet to be conducted. A longitudinal study would rule out such cohort effects.

Secondly, many of the studies that explore the FIE in children are underpowered. When testing multiple age-groups, the necessary increase in error degrees of freedom mean that it becomes much more difficult to detect significant differences in the FIE due to age. A within-subjects (and thereby longitudinal design) would improve the statistical power of such studies. Only when these issues are addressed can studies adequately address the mechanisms of face processing employed by children.

Thirdly, there are several statistical issues with existing work on the development of face recognition. Certain tasks do not adequately control for floor and ceiling effects. Floor effects cause a task to be too difficult for younger children to complete. This makes distinguishing any effect of inversion very difficult (a similar argument has been made by Weigelt et al., 2013). One reason for floor and ceiling effects potentially is the use of age-inappropriate stimuli. Given that the own-age bias exists in face perception, in which participants show a larger FIE for own-age than other-age faces (Anastasi & Rhodes, 2005; Harrison & Hole, 2009; Hills & Lewis, 2011; Kuefner, Macchi Cassia, Picozzi, & Bricolo, 2008),[5] in order to avoid floor effects, faces should be age-matched to the participants. In adults, the processing of other-group faces has been theoretically linked to not using the most expert configural processing system (Hugenberg & Corneille, 2009; Michel, Caldara, & Rossion, 2006; Michel, Rossion, Han, Chung, & Caldara, 2006), which lowers performance in such tasks. This problem means that standardised tests of children's face recognition performance, such as the Cambridge Face Memory Test – Children (CFMT-C; Croydon, Pimperton, Ewing, Duchaine, & Pellicano, 2014) might underestimate performance.

Finally, even when tasks are sensitive enough to detect differences, there is a further statistical issue: overall performance in younger children is lower than that of older children. This means that any effects of inversion may be harder to detect in younger children. In order to address this, a relative measure of performance needs to be considered (Goldstein, 1965). A relative measure takes into account the fact that children's overall performance will be lower than that of adults and therefore allows for smaller differences in performance in children to be equated to larger differences observed in adulthood. Even with a relative measure of the FIE, there is a potential issue with a younger children showing a more limited range of performance than older participants. This can be assess by ensuring that the variances in performance are equivalent for all groups of participants (which is an assumption of parametric data in any case). An alternative method to control for this is to match performance of upright faces in all children by presenting different numbers of stimuli. While matching performance addresses the issues of poorer performance in children, it creates a confound: face recognition is made up of face perception and face memory, therefore manipulating the number of stimuli prevents an analysis of face memory. It is for this reason, a relative measure of the FIE is the more appropriate technique for measuring face recognition performance in children.

This paper presents a solution to these problems in order to establish whether children show the FIE to a similar level as adults and, by extrapolation, utilise expert face processing. Here, the FIE, as a measure of expertise, was assessed using a standard old/new recognition paradigm in children (from 5- to 15-years-old) and adults. Three possible developmental trends are possible: The magnitude of the FIE may increase with age as a product of experience (developmental induction); The magnitude of the FIE may be constant throughout development if the effect is not based upon experience (due to early maturation); Finally, developmentally-late maturation may occur in which the FIE appears at a particular critical age. These trajectories lead to the increasing inversion effect, the constant inversion effect, and the "all or none" hypotheses. The first two hypotheses can be explained by an early maturation of expert face processing mechanisms and the final hypothesis is derived from a late maturation of expert face processing mechanisms. In this study, a relative measure of the FIE was used in order to control for poorer general cognition in younger children controlling for lower absolute performance of the younger children and therefore avoids floor effects. Absolute floor and ceiling effects were avoided by choosing a manageable number of stimuli for the youngest children tested that produced sufficient errors in the adult participants (Ellis, 1992b). While there remains the possibility that younger children's performance was more restricted than that of older participants which might obscure results, our data indicate that there was roughly equivalent variance in performance for the younger children and the older children suggesting that we did not have floor effects in our study. Floor effects were also avoided by testing own-age faces for all participants in Experiments 1 and 2. In order to show that development is occurring within participants (avoiding cohort effects) we also conducted a longitudinal study of face recognition by testing the same children at different ages in Experiments 2 and 4.

2. Experiment 1

Experiment 1 was a cross-sectional experiment executed in a similar manner to de Heering et al. (2012). Participants ranged from 5- to 15-years of age and an adult sample. The primary purpose of Experiment 1 was to understand the developmental trajectory of face recognition and expert face recognition as measured by the FIE. In order to do this, correlations and curve fitting was conducted

[5] There is evidence that newborns show a bias toward recognising young adult faces (Macchi Cassia, 2011) which may extend into childhood, however, most of the evidence indicates that once children are in school, they show an own-age bias (Rhodes & Anastasi, 2012).

Table 1
Participant characteristics (age range in years,months; mean age in years,months; SD age; and number of females in the group) for Experiment 1.

Age Group/Range	Mean Age	SD Age	Number Female
5,7–6,6	6,1	0.28	23
6,7–7,6	7,2	0.27	23
7,7–8,6	8,2	0.27	25
8,7–9,6	9,3	0.31	19
9,7–10,6	10,2	0.31	22
10,7–11,6	11,2	0.28	25
11,7–12,6	12,3	0.26	19
12,7–13,6	13,2	0.26	16
13,7–14,6	14,2	0.30	22
14,7–15,6	15,2	0.27	19
15,7–16,6	16,2	0.28	20
18,2–23,1	21,0	1.37	28

for upright and inverted face recognition and the FIE separately. This will distinguish the developmental trend of the face processing.

2.1. Methods

2.1.1. Participants

Participants were 440 children (198 male) aged from 5,7 years to 15,4 years and 40 adults (aged 18–23 years; 12 male). See Table 1 for participant details. The age groups chosen were linked to the school age they were studying in rather than strictly their age (see Hills & Lewis, 2011). Adult participants were recruited through a university whereas the children were recruited through local mainstream schools. All participants had normal or corrected vision and were ethnically White as indicated by parent-report (or self-report in the adult group). All of the children were considered typically developing by their schools. No participants were familiar with any of the faces used in this experiment.

2.1.2. Materials

Frontal-view photographs of White children's (aged from 5- to 18-years) faces were collected by a research assistant. Two images of each child were collected: one presented during the learning phase and one presented during the test phase. The images were quite similar, taken a few moments apart, but the facial expression was slightly different. This was done to reduce pictorial recognition (Bruce, 1982). Photographs of 32 (16 female) children were collected in each age group. All parents provided consent for these photographs to be used in this research project. The faces all had similar hairstyles, positioned in a frontal view in neutral or mildly happy expressions (this was randomised). All extraneous paraphernalia and the background were masked using Adobe™ Photoshop™. The stimuli were collected from local schools that were not used for the experimental testing. Participants only saw faces of their own age. To confirm that faces of one age category were not more dissimilar to those in another category, a pixelwise similarity comparison was made between the faces in each age group (Haushofer, Livingstone, & Kanwisher, 2008). This was not significant ($p > .56$). Similarly, attractiveness ratings did not differ across stimuli types ($p > .39$). While this cannot rule out stimulus differences across ages, it indicates that differences are not substantial. The faces were presented 100 mm by 110 mm dimensions in 72 dpi resolution in the learning phase and 150 mm by 165 mm in the test phase. An inverted version of each face was created using the rotate function in Adobe™ Photoshop™. These were presented using Superlab Pro 2™ Research Software using a Toshiba Tecra M4™ Tablet PC.

2.1.3. Design

Twelve participant groups were tested as determined by their school year group. Participants viewed both upright and inverted faces. The faces were counterbalanced between participants such that each face was a target as often as it was a distracter. The faces were counterbalanced such that they appeared upright as frequently as they appeared inverted. Faces were presented in a random order (i.e., there was no blocking of face type). The dependent variables were recognition accuracy, measured in terms of the Signal Detection Theory (SDT e.g., Swets, 1966) measure d' and response bias, measured in terms of the SDT measure, C. Reaction time could not accurately be measured due to the experimenter keying the responses (see procedure). The relative FIE was established using the formula:

$$\mathrm{FIE} = (d'_{\mathrm{U}} - d'_{\mathrm{I}})/(d'_{\mathrm{U}} + d'_{\mathrm{I}}) \tag{1}$$

where d'_{U} is the recognition accuracy of upright faces and d'_{I} is the recognition accuracy of inverted faces. This measure produces a relative measure of the FIE and controls for differences in overall accuracy between participant groups.

2.1.4. Procedure

Participants were tested individually in a quiet brightly-lit room in their school (or in a quiet laboratory for adults). Participants sat 50 cm from the computer screen. This screen was positioned away from the Experimenter such that he could not see the contents of the screen. Participants responded verbally and the experimenter entered the responses into a standard computer keyboard. Thus, the Experimenter could not influence the participants' responses because the stimuli were presented in a random order, and he was unaware of what was on the screen, preventing demand characteristics. Additionally, this ensured that the Experimenter could ensure the participants followed the instructions appropriately.

A standard old/new recognition paradigm was employed involving three consecutive phases: learning, distraction, and test. In the learning phase, participants were shown half of the set of faces of their own-age (N = 16) of which half were upright and half were inverted. Faces were presented centrally in a sequential random order. Participants were instructed to rate each face for how attractive they thought the face was using a 1–9 Likert-type scale, with the anchor points "ugly" and "beautiful" (Light, Hollander, & Kayra-Stuart, 1981). If a participant did not understand the scale, it was explained to them using alternative synonyms. The presentation of each face was response terminated: There were no differences in presentation duration across participant group (see Table 3, $F(11, 468) = 0.24$, $MSE = 1464959$, $p = .995$, $\eta_p^2 = 0.01$). There was a 150 ms blank screen inter-stimulus interval between each face.

Immediately after this presentation, participants were given some control questions. These were: "What is your first name?" "What is your surname?" "What is your gender?" "How old are you?" "When is your birthday?" "What school do you go to?" "What school year group are you in?" "Where were you born?". If the participant did not understand the question it was explained using simpler synonyms. Only participant age, birthday, and gender were recorded. These questions took no longer than 60 s to administer.

Following this, the participants were given the test phase. In this, the participants saw all 16 target faces and 16 distractor faces sequentially in a random order and made an old/new recognition judgement to each face. Each presentation was in the centre of the screen and was response terminated. Half the faces were upright and half the faces were inverted. Orientation of the target faces was matched from learning to test. The participants were told to be as quick and as accurate as possible. Between each face there was a blank screen for 150 ms. Once this phase was completed, the participants were thanked and debriefed. Total testing time was no more than four minutes per participant.

2.2. Results

Old/new responses from participants were converted to hit and false alarm rates and these were used to calculate d' and C using the MacMillan and Creelman (2005) method. In addition to reporting the traditional null hypothesis significance tests, we also report Bayesian statistics throughout this work. Bayesian analysis has the advantage that it is not based on the evaluation of significance levels that can be interpreted incorrectly (especially regarding non significant results, see Rouder, Speckman, Sun, Morey, & Iverson, 2009). The Bayes Factor (B_{10}) provides the likelihood ratio of the experimental hypothesis being true compared to the null hypothesis being true (Dienes, 2011). A Bayes Factor of between 1/3 and 3 only provides anecdotal evidence; a value between 3 and 10 provides substantial evidence for the hypothesis; a value above 10 provides strong evidence for the hypothesis; a value between 0.1 and 1/3 provides substantial evidence for the null hypothesis and; a value less that 0.1 provides strong evidence for the null hypothesis (Jeffreys, 1961). We used the Bayes calculator provided by Dienes (2015).

2.2.1. Recognition accuracy (d')

Zero counts in misses and false alarms were replaced with 0.01 (there were no zero counts in hits or correct rejections). With the number of stimuli in this experiment, d' ranged from 0 (chance recognition) to 4.48 (perfect performance). Fig. 1 presents the mean d' for upright and inverted faces grouped according to age and highlights, crucially, that there were no floor or ceiling effects in any condition.

These data were first subjected to a 12 × 2 mixed ANOVA with the factors: participant age and orientation. Critically, there was strong evidence for a significant interaction between orientation and participant age, $F(11, 468) = 13.19$, $MSE = 0.53$, $p < .001$, $\eta_p^2 = .24$, $B_{10} = 3.18 \times 10^{38}$. To decompose this interaction, two univariate ANOVAs were run: one for upright faces and one for inverted faces. The effect of age was larger for upright faces, $F(11, 468) = 23.60$, $MSE = 0.75$, $p < .001$, $\eta_p^2 = .36$, $B_{10} = 6.53 \times 10^{45}$, than for inverted faces, $F(11, 468) = 2.36$, $MSE = 0.27$, $p = .008$, $\eta_p^2 = .05$, $B_{10} = 401679$. Curve estimations were conducted for recognition accuracy, summarised in Table 2. The relationship between recognition accuracy and age was equally well represented by a linear, quadratic, and cubic function. Given that all three were significant, and that they all represent the data well, it is clear that this pattern indicates an overall linear improvement in recognition accuracy, but with a step change mid way through development, indicated by the quadratic and cubic functions. The correlations for upright faces were compared against the inverted faces using Fisher's r-to-z transformation. These revealed that the correlations were significantly stronger for upright faces than inverted faces (see Table 2).

An analysis was conducted on the mean-relative FIE presented in the bottom panel of Fig. 1, and revealed strong evidence that the FIE depended on age, $F(11, 468) = 6.84$, $MSE = 0.16$, $p < .001$, $\eta_p^2 = .14$, $B_{10} = 4445395$. Pairwise comparisons revealed that there was no significant difference in the FIE form participants aged 5- to 10-years nor between 9-years and adults (all $ps > .18$). Five- to 8-year old participants showed a significantly smaller FIE than participants age 11 and older (all $ps < .030$). Eleven-year old and older participants showed a significantly larger FIE than children under the age of 8-years of age. The FIE was compared at each age using a series of one-sample t-tests, with the significance levels reported in Fig. 1. Curve estimations show while the FIE increased with age according to an overall linear trend, the significant quadratic and cubic functions indicate a step change in the FIE during development.

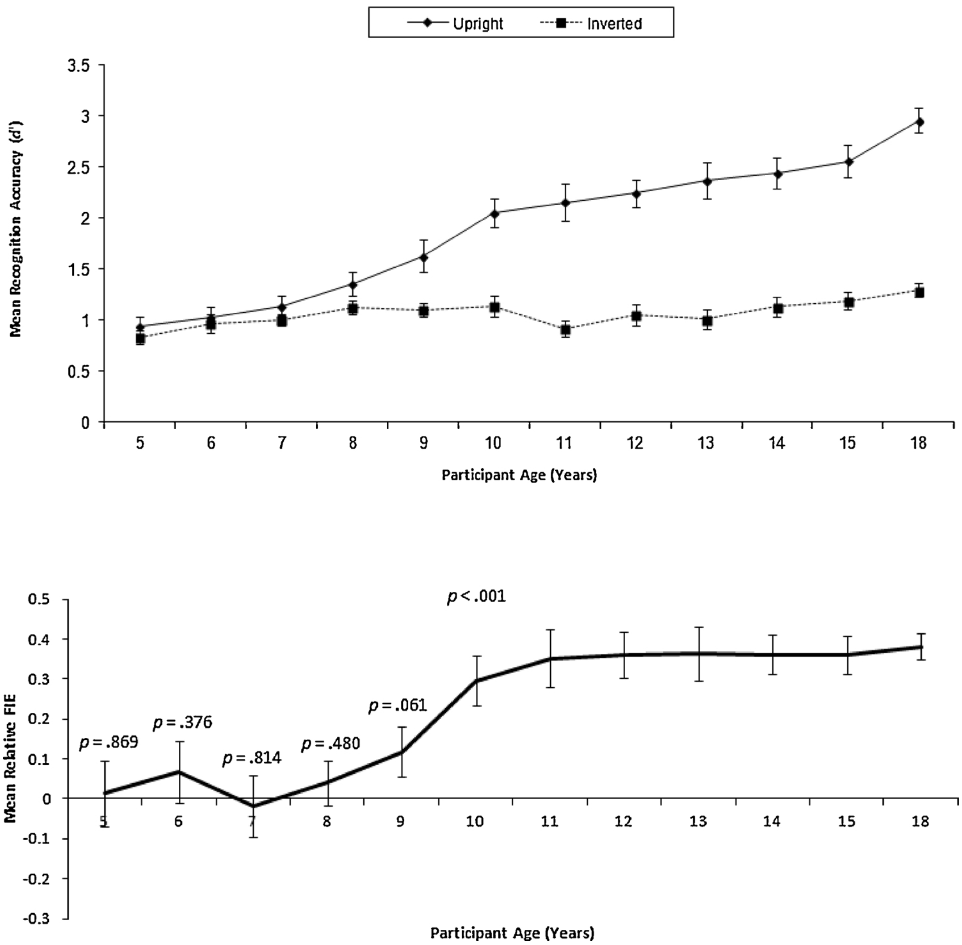

Fig. 1. Top panel: Mean recognition accuracy (d') for upright and inverted age-matched faces split by participant age (Experiment 1). Bottom panel: Mean FIE for age-matched faces split by participant age (Experiment 1). Error bars show standard error.

Table 2
Curve estimations for recognition accuracy (d'), hit rate, false alarm rate, response bias (C), and the FIE for upright and inverted age-matched faces (Experiment 1).

Variable	Orientation	Linear Estimation	Quadratic Estimation	Cubic Estimation
Recognition Accuracy (d')	*Upright*	$R = .58, F = 248.56, p < .001$	$R = .59, F = 127.78, p < .001$	$R = .59, F = 85.08, p < .001$
	Inverted	$R = .17, F = 13.89, p < .001$	$R = .17, F = 6.98, p = .001$	$R = .19, F = 6.13, p < .001$
	Difference	$Z = 7.58, p < .001$	$Z = 7.81, p < .001$	$Z = 7.50, p < .001$
Hit Rate	*Upright*	$R = .40, F = 91.87, p < .001$	$R = .41, F = 46.77, p < .001$	$R = .41, F = 31.88, p < .001$
	Inverted	$R = .04, F = 0.77, p = .382$	$R = .10, F = 2.49, p = .084$	$R = .17, F = 4.49, p = .004$
	Difference	$Z = 5.92, p < .001$	$Z = 5.18, p < .001$	$Z = 4.08, p < .001$
False Alarm Rate	*Upright*	$R = .37, F = 77.24, p < .001$	$R = .39, F = 41.82, p < .001$	$R = .40, F = 29.46, p < .001$
	Inverted	$R = .19, F = 17.30, p < .001$	$R = .20, F = 9.71, p < .001$	$R = .20, F = 6.48, p < .001$
	Difference	$Z = 3.03, p = .002$	$Z = 3.23, p = .001$	$Z = 3.41, p = .001$
Response Bias (C)	*Upright*	$R = .03, F = 0.43, p = .515$	$R = .10, F = 2.30, p = .101$	$R = .12, F = 2.30, p = .076$
	Inverted	$R = .08, F = 3.25, p = .072$	$R = .14, F = 4.64, p = .010$	$R = .16, F = 4.26, p = .006$
	Difference	$Z = 0.77, p = .441$	$Z = 0.63, p = .529$	$Z = 0.63, p = .529$
Face-Inversion Effect		$R = .33, F = 57.15, p < .001$	$R = .34, F = 31.84, p < .001$	$R = .35, F = 22.13, p < .001$

2.2.2. Response bias, C

A final analysis was conducted on the response bias data (shown in Table 3), calculated using the Macmillan and Creelman (2010) method. This measures participants' tendency to respond with an "old" response. The score typically ranges from -1 to $+1$, where 0 is a neutral bias. A positive number indicates participants are less likely to response with an "old" respond when they are unsure and a

Table 3
Mean RT to make attractiveness ratings (ms), hit rate, false alarm rate, and response bias (*C*) for upright and inverted age-matched faces split by participant age for Experiment 1. Standard error is shown in parenthesis.

Participant Age	RT to make Attractiveness Rating		Hit rate		False Alarm rate		Response Bias	
	Upright faces	Inverted faces	Upright faces	Inverted faces	Upright faces	Inverted faces	Upright faces	Inverted faces
5-years	3091 (142)	3243 (142)	.63 (.03)	.68 (.02)	.28 (.02)	.37 (.01)	0.14 (0.04)	−0.08 (0.04)
6-years	2807 (137)	3176 (141)	.56 (.02)	.67 (.02)	.22 (.02)	.32 (.02)	0.35 (0.05)	−0.03 (0.04)
7-years	2806 (139)	3180 (142)	.65 (.02)	.70 (.01)	.25 (.02)	.33 (.01)	0.16 (0.04)	−0.04 (0.02)
8-years	2807 (148)	3161 (149)	.58 (.01)	.71 (.02)	.16 (.01)	.31 (.02)	0.41 (0.07)	−0.03 (0.04)
9-years	2937 (143)	3279 (142)	.62 (.02)	.72 (.02)	.15 (.02)	.32 (.02)	0.45 (0.07)	−0.05 (0.05)
10-years	2850 (133)	3193 (131)	.62 (.02)	.70 (.02)	.11 (.02)	.30 (.02)	0.56 (0.12)	0.00 (0.05)
11-years	2953 (148)	3237 (147)	.74 (.03)	.67 (.02)	.17 (.03)	.33 (.01)	0.16 (0.12)	−0.00 (0.04)
12-years	2938 (136)	3205 (141)	.80 (.02)	.64 (.03)	.15 (.02)	.28 (.02)	0.13 (0.09)	0.12 (0.05)
13-years	2786 (135)	3335 (148)	.75 (.03)	.65 (.02)	.14 (.03)	.29 (.02)	0.28 (0.11)	0.08 (0.05)
14-years	2764 (132)	3408(148)	.75 (.02)	.67 (.03)	.12 (.02)	.27 (.02)	0.30 (0.11)	0.08 (0.05)
15-years	2751 (142)	3104 (142)	.78 (.02)	.71 (.02)	.09 (.02)	.28 (.02)	0.32 (0.09)	0.01 (0.04)
18-years	2795 (141)	3250 (155)	.88 (.01)	.74 (.02)	.08 (.01)	.28 (.02)	0.16 (0.06)	−0.04 (0.04)

negative number indicates a tendency for participants to respond with an “old” response when they are unsure.

There was strong evidence for a significant interaction between orientation and participant age, $F(11, 468) = 2.60$, $MSE = 0.19$, $p = .003$, $\eta_p{}^2 = .06$, $B_{10} = 1.69 \times 10^{15}$. To decompose this interaction, two univariate ANOVAs were run: one for upright faces and one for inverted faces. For upright faces, the effect of age was larger with a lesser tendency to respond with a "new" response for older participants than younger ones, $F(11, 468) = 2.78$, $MSE = 0.28$, $p = .002$, $\eta_p{}^2 = .06$, $B_{10} = 232.64$, than for inverted faces, $F(11, 468) = 2.10$, $MSE = 0.07$, $p = .019$, $\eta_p{}^2 = .05$, $B_{10} = 82.48$. Curve estimations were conducted for response bias, summarised in Table 2.

2.3. Discussion

The results from Experiment 1 indicate that face recognition improves with age. This is entirely consistent with previous research (e.g., Blaney & Winograd, 1978; Carey et al., 1980; Flin, 1985). Curve fitting revealed that upright face recognition improved more reliably with age than inverted face recognition. All measures of face recognition improved with age, suggesting that participants memory for faces improved with age (as indicated by hit rate) and the coding of faces becomes more accurate (as indicated by a reduction in false alarm rate). These results indicate that face memory and face perception are developing.

The pattern of data observed here also indicates that there is a step-change in the FIE, whereby participants under the age of 9-years do not show the FIE, whereas participants 10-years and older do show the FIE. In this analysis, we have fitted curves to the data. These curves show a linear, cubic, and quadratic trajectory for the development of the FIE suggesting a qualitative shift in the use of configural coding. If the development of configural processing was due to developmental induction, we would only expect to see a linear relationship. Only a hypothesis predicting that there would be no FIE until a certain age, followed by a full-strength FIE is compatible with the data. This is as predicted by the late maturation of configural processing account.

3. Experiment 2

In order to fully explore this step-change in the use of configural coding, a second experiment was conducted. This was a longitudinal study in which the same 9-year-old participants were followed up for two years following the original testing. We chose to follow these participants due to the observation in Experiment 1 that indicated a step change in the FIE between the ages of 9 and 11 years. This acts as an internal replication of the initial findings and critically offers a longitudinal approach to understand how face recognition develops. The longitudinal design also offers a window to explore development of a particular group of participants. Many of the drawbacks of cross-sectional designs are avoided in this approach as it allows researchers to see change in behaviour due to age. This is also one of a very small handful of longitudinal studies applied to face recognition.

3.1. Method

The same nine-year-old participants in Experiment 1 were tested two further times at approximately (within one month) yearly intervals. Testing took place in an identical manner as in Experiment 1 with the same stimuli used for the ten- and 11-year-old children in Experiment 1 for the second and third testing times respectively. Therefore, the participants were viewing own-age faces at age 9-, 10-, and 11-years. This, therefore, led to a 2 × 3 within-subjects design with the factors: orientation of the face and participant age (9-, 10-, and 11-years of age).

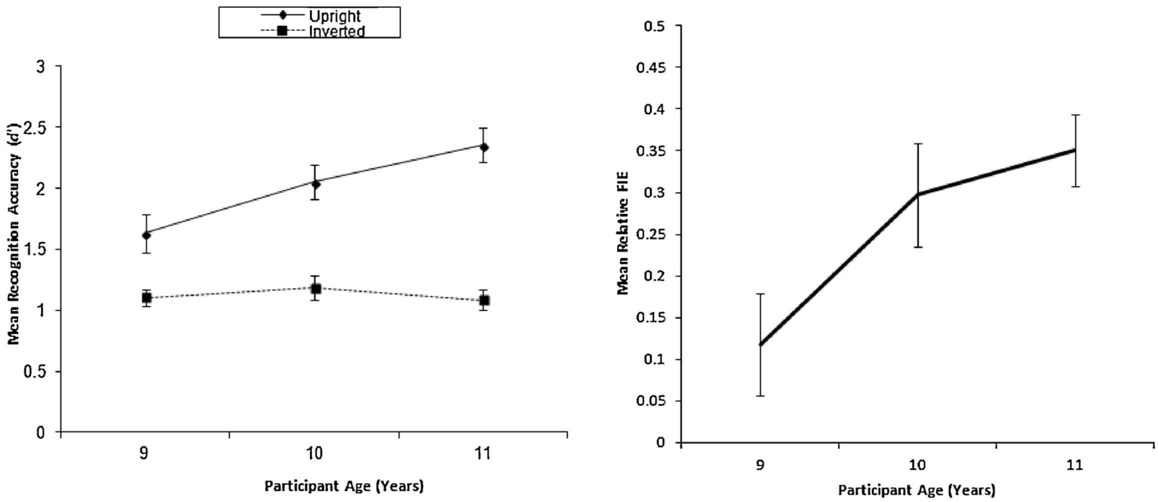

Fig. 2. Left panel: Mean recognition accuracy (*d*') for upright and inverted age-matched faces split by participant age (Experiment 2). Right panel: Mean FIE for age-matched faces split by participant age (Experiment 2). Error bars show standard error.

3.2. Results

The data were treated in the same way as in Experiment 1 and are presented in Fig. 2 and Table 4. The recognition accuracy (*d*') data were subjected to a 2 × 3 within-subjects ANOVA with the factors orientation and age. Critically, these effects interacted, $F(2, 78) = 4.20$, $MSE = 0.66$, $p = .018$, $\eta_p^2 = .10$, $B_{10} = 6.68 \times 10^8$. This interaction was decomposed by conducting two one-way ANOVAs, one for upright faces and one for inverted faces. The improvement in recognition accuracy for upright faces was significant, $F(2, 78) = 5.58$, $MSE = 0.94$, $p = .005$, $\eta_p^2 = .13$, $B_{10} = 37.85$, but it was not for inverted faces, $F(2, 78) = 0.36$, $MSE = 0.32$, $p = .699$, $\eta_p^2 = .01$, $B_{10} = 0.09$. A trend analysis was conducted and this revealed that the improvement in face recognition accuracy for upright faces was linear, $F = 11.32$, $MSE = 0.92$, $p = .002$, $\eta_p^2 = .23$, with no other significant developmental trajectories (all *ps* > .769).

A one-way ANOVA was run on the relative measure of the FIE, revealing a significant effect of age, $F(2, 78) = 4.05$, $MSE = 0.15$, $p = .021$, $\eta_p^2 = .09$, $B_{10} = 18.77$. One-sample *t*-tests showed that there was strong evidence the FIE was not present in the 9-year old participants, $t(39) = 1.93$, $p = .061$ effect size $r = .30$, $B_{10} = 0.24$, but was present in 10-year old, $t(39) = 4.27$, $p < .001$, effect size $r = .56$, $B_{10} = 1606.76$, and 11-year old participants, $t(39) = 8.22$, $p < .001$, effect size $r = .80$, $B_{10} = 5.00 \times 10^{13}$. A linear trend was observed, $F = 9.19$, $MSE = 0.12$, $p = .004$, $\eta_p^2 = .19$. No other contrast pattern was observed (all *ps* > .44).

Finally, a parallel ANOVA was run on the response bias (*C*) data, revealing a significant interaction, $F(2, 78) = 7.09$, $MSE = 0.19$, $p = .001$, $\eta_p^2 = .15$, $B_{10} = 1.27 \times 10^{11}$. The change in response bias with age (with a lesser tendency to respond with a "new" response for older participants than younger ones) was significant for upright faces, $F(2, 78) = 9.80$, $MSE = 0.28$, $p < .001$, $\eta_p^2 = .20$, $B_{10} = 1.10 \times 10^{90}$, but not for inverted faces, $F(2, 78) = 0.43$, $MSE = 0.09$, $p = .649$, $\eta_p^2 = .01$, $B_{10} = 0.02$. This change in response bias followed a linear pattern, $F = 11.92$, $MSE = 0.27$, $p = .001$, $\eta_p^2 = .23$, and a quadratic pattern, $F = 7.87$, $MSE = 0.30$, $p = .008$, $\eta_p^2 = .17$.

3.3. Discussion

We have demonstrated a significant improvement in the recognition of upright age-matched faces with development. This

Table 4
Mean hit rate, false alarm rate, recognition accuracy (*d*'), response bias (*C*) for upright and inverted age-matched faces and relative FIE split by participant age for Experiment 2. Standard error is shown in parenthesis.

		Participant Age		
		9-years	10-years	11-years
Hit Rate	Upright Faces	.62 (.03)	.67 (.03)	.79 (.03)
	Inverted Faces	.72 (.02)	.72 (.03)	.71 (.02)
False Alarm Rate	Upright Faces	.15 (.02)	.10 (.02)	.16 (.02)
	Inverted Faces	.32 (.02)	.32 (.02)	.32 (.02)
Response Bias (C)	Upright Faces	.45 (.07)	.55 (.07)	.05 (.11)
	Inverted Faces	−.05 (.05)	−.05 (.05)	.04 (.04)

Table 5
Participant characteristics (age range in years, months; mean age in years, months; SD age; and number of females in the group for Experiment 3.

Age Group/Range	Mean Age	SD Age	Number Female
5,7–6,6	6,2	0.29	12
6,7–7,6	7,1	0.24	13
7,7–8,6	8,3	0.34	14
8,7–9,6	9,3	0.24	10
9,7–10,6	10,2	0.25	9
10,7–11,6	11,1	0.27	10
11,7–12,6	12,3	0.30	11
12,7–13,6	13,1	0.24	12
13,7–14,6	14,3	0.30	13
14,7–15,6	15,2	0.30	11
15,7–16,6	16,1	0.27	12
18,2–23,1	21,3	1.22	6

development is much greater for upright faces than for inverted faces. The effect of this is that the FIE is not observed when our participants were 9-years of age, but was present when they were older. This developmental study indicates that a step change in the emergence of the FIE. This important advancement in our understanding of the development of face recognition.

4. Experiment 3

In order to be able to fully compare these results to previously published data, it is important to assess whether the developmental trends we have found exist when recognising adult faces. This is especially important given the use of the CFMT-C (Croydon et al., 2014) to assess deficits in children's face recognition (Bennetts, Murray, Boyce, & Bate, 2017) which use adult faces. If the results we found in Experiments 1 and 2 generalise to adult faces, then there is no concern with using such adult tests in children. To this end, we used a similar method to that use in Experiments 1 and 2, with different participants and different faces: specifically, adult faces.

4.1. Method

4.1.1. Participants, design, and procedure

Participants were 242 children (115 male) aged from 5,7 years to 16,2 years and 22 adults (aged 18–22 years; 6 male). A participant summary is shown in Table 5. All other participant details were the same as in Experiment 1. The design and procedure of this Experiment was identical to that in Experiment 1.

4.1.2. Materials

Two images of 32 (16 female) adult faces from the Minear and Park (2004) database were used in this study. These were of adults in their late 20 s (and therefore did not match the age of any of our participants). One image displayed a happy expression and the other displayed a neutral expression. One of these was presented during learning and the other during test (this was randomised). Using two images of the same face reduces pictorial recognition (Bruce, 1982). The face images in this database were of males and females, all with similar hairstyles, positioned in a frontal view. All extraneous paraphernalia and the background were masked using Adobe™ Photoshop™. The faces were presented in the same way as in Experiment 1 and in the same dimensions.

4.2. Results

The results for this Experiment were analysed in the same way as Experiment 1.

4.2.1. Recognition accuracy (d')

Recognition accuracy data are summarised in Fig. 3 and were subjected to a 12 × 2 mixed ANOVA, with the factors: participant age and orientation. This revealed strong evidence for a significant interaction between orientation and participant age, $F(11, 252) = 8.66$, $MSE = 0.44$, $p < .001$, $\eta_p^2 = .27$, $B_{10} = 1.45 \times 10^{18}$. To decompose this interaction, two univariate ANOVAs were run: one for upright faces and one for inverted faces. For upright faces, the effect of age was larger, $F(11, 252) = 13.52$, $MSE = 0.76$, $p < .001$, $\eta_p^2 = .37$, $B_{10} = 1.88 \times 10^{14}$, than for inverted faces, $F(11, 252) = 2.30$, $MSE = 0.54$, $p = .011$, $\eta_p^2 = .09$, $B_{10} = 70.47$. Curve estimations were conducted for recognition accuracy, summarised in Table 6. The relationship between recognition accuracy and age was equally well represented by a linear, quadratic, and cubic function. The correlations for upright faces were compared against the inverted faces using Fisher's r-to-z transformation. These revealed that the correlations were significantly stronger for upright faces than inverted faces (see Table 6).

An analysis was conducted on the mean-relative FIE presented in Fig. 3, and revealed that the FIE depended on age, $F(11, 252) = 3.11$, $MSE = 0.17$, $p = .001$, $\eta_p^2 = .12$, $B_{10} = 8.79$. However, no pairwise comparisons between participant ages were significant following Bonferroni-Šidák correction. This highlights the critical point about using age-inappropriate stimuli and a lack of

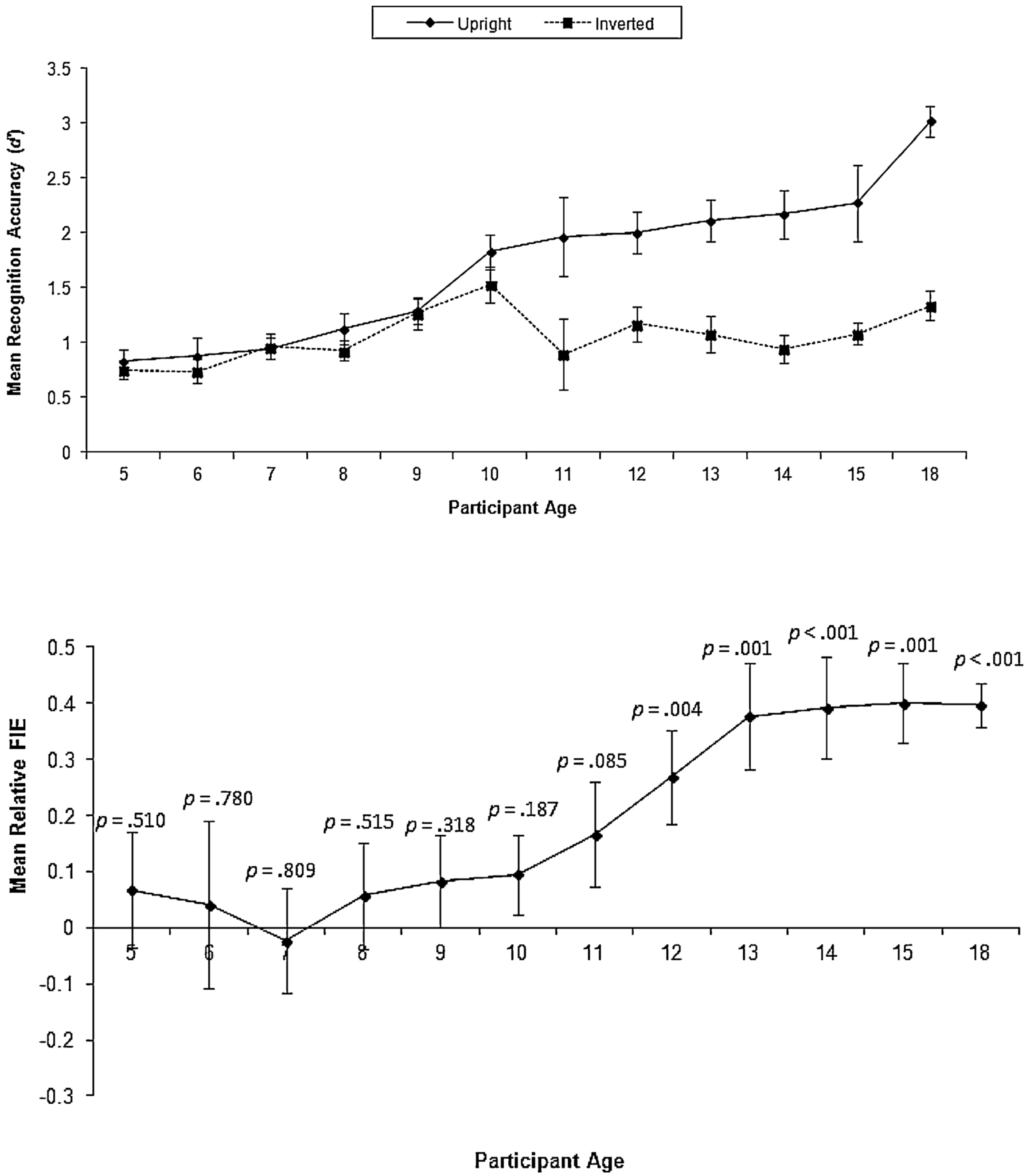

Fig. 3. Top panel: Mean recognition accuracy (d') for upright and inverted adult faces split by participant age (Experiment 3). Bottom panel: Mean FIE for adult faces split by participant age (Experiment 3). Error bars show standard error.

experimental power may hide real effects of development in the FIE. One-sample *t*-tests were used to compare the magnitude of the FIE at each age, shown in Fig. 3. Curve estimations show that changes in the FIE follow a linear, quadratic, and cubic function.

4.2.2. Response bias, C

A final analysis was conducted on the response bias data (see Table 7). There was a significant interaction between orientation and participant age, $F(11, 252) = 0.70$, $MSE = 0.15$, $p = .737$, $\eta_p^2 = .03$, $B_{10} = 45.72$. To decompose this interaction, two univariate ANOVAs were run: one for upright faces and one for inverted faces. For upright faces, the effect of age was smaller and with anecdotal evidence, $F(11, 252) = 1.88$, $MSE = 0.20$, $p = .042$, $\eta_p^2 = .08$, $B_{10} = 2.15$, compared to the strong evidence for inverted faces, $F(11, 252) = 3.33$, $MSE = 0.16$, $p < .001$, $\eta_p^2 = .13$, $B_{10} = 14.73$. Older participants demonstrated a lesser tendency to respond with an "old" response than younger ones. Curve estimations were conducted for response bias, summarised in Table 6.

Table 6
Curve estimations for recognition accuracy (d'), hit rate, false alarm rate, response bias (C), and the FIE for upright and inverted adult faces (Experiment 3).

Variable	Orientation	Linear Estimation	Quadratic Estimation	Logistic Estimation
Recognition Accuracy (d')	*Upright*	R = .60, F = 145.02, p < .001	R = .60, F = 72.25, p < .001	R = .60, F = 47.99, p < .001
	Inverted	R = .15, F = 6.30, p = .013	R = .17, F = 3.76, p = .025	R = .22, F = 4.41, p = .005
	Difference	Z = 6.05, p < .001	Z = 5.82, p < .001	Z = 5.24, p < .001
Hit Rate	*Upright*	R = .41, F = 53.75, p < .001	R = .41, F = 26.96, p < .001	R = .43, F = 19.92, p < .001
	Inverted	R = .10, F = 2.74, p = .099	R = .11, F = 1.59, p = .206	R = .19, F = 3.23, p = .023
	Difference	Z = 3.74, p < .001	Z = 3.63, p < .001	Z = 2.99, p = .003
False Alarm Rate	*Upright*	R = .42, F = 57.13, p < .001	R = .42, F = 28.57, p < .001	R = .43, F = 19.13, p < .001
	Inverted	R = .10, F = 2.80, p = .096	R = .11, F = 1.69, p = .186	R = .12, F = 1.33, p = .265
	Difference	Z = 3.88, p < .001	Z = 3.76, p = .001	Z = 3.79, p < .001
Response Bias (C)	*Upright*	R = .07, F = 1.17, p = .280	R = .07, F = 0.70, p = .496	R = .12, F = 1.33, p = .266
	Inverted	R = .00, F = 0.00, p = .990	R = .06, F = 0.40, p = .670	R = .08, F = 0.60, p = .619
	Difference	Z = 0.78, p = .435	Z = 0.11, p = .912	Z = 0.45, p = .653
Face-Inversion Effect		R = .32, F = 29.14, p < .001	R = .32, F = 14.51, p < .001	R = .34, F = 11.43, p < .001

Table 7
Mean hit rate, false alarm rate, and response bias (C) for upright and inverted faces split by participant age for Experiment 3. Standard error is shown in parenthesis.

Participant Age	RT to make Attractiveness Rating		Hit rate		False Alarm rate		Response Bias	
	Upright faces	Inverted faces	Upright faces	Inverted faces	Upright faces	Inverted faces	Upright faces	Inverted faces
5-years	3106 (151)	3213 (241)	.56 (.03)	.57 (.03)	.27 (.03)	.30 (.03)	0.26 (0.06)	0.20 (0.07)
6-years	3027 (158)	3369 (232)	.57 (.04)	.63 (.03)	.27 (.03)	.36 (.03)	0.34 (0.05)	0.03 (0.07)
7-years	2888 (167)	3083 (224)	.65 (.03)	.67 (.03)	.30 (.01)	.32 (.02)	0.05 (0.04)	0.01 (0.05)
8-years	2966 (189)	3045 (226)	.62 (.03)	.62 (.03)	.23 (.03)	.28 (.03)	0.22 (0.06)	0.15 (0.06)
9-years	2892 (144)	3057 (191)	.69 (.03)	.68 (.04)	.24 (.02)	.26 (.02)	0.11 (0.06)	0.07 (0.10)
10-years	2978 (172)	3610 (281)	.68 (.03)	.60 (.03)	.15 (.03)	.16 (.03)	0.42 (0.12)	0.50 (0.11)
11-years	2924 (156)	3450 (295)	.75 (.06)	.64 (.06)	.22 (.05)	.38 (.05)	0.08 (0.11)	−0.10 (0.08)
12-years	2923 (142)	3542 (287)	.80 (.03)	.68 (.04)	.20 (.04)	.28 (.03)	0.08 (0.10)	0.06 (0.07)
13-years	2829 (116)	3612 (265)	.75 (.03)	.60 (.05)	.15 (.03)	.27 (.05)	0.24 (0.12)	0.21 (0.14)
14-years	2849 (157)	3612 (269)	.74 (.04)	.63 (.04)	.14 (.03)	.30 (.03)	0.26 (0.13)	0.11 (0.08)
15-years	2955 (62)	3740 (272)	.68 (.08)	.61 (.02)	.11 (.01)	.25 (.02)	0.47 (0.17)	0.24 (0.01)
18-years	2749 (76)	3457 (277)	.89 (.01)	.73 (.02)	.07 (.02)	.26 (.03)	0.21 (0.08)	0.03 (0.06)

4.3. Discussion

The results from this Experiment are consistent with Experiment 1: face recognition improves with age and more reliably so for upright faces than inverted faces. The step-change in the FIE, whereby participants under the age of 9-years do not show the FIE, whereas participants 11-years and older do show the FIE, was also found. Curve-fitting showed a linear, cubic, and quadratic trajectory for the development of the FIE suggesting a qualitative shift in the use of configural coding suggesting late maturation of configural processing.

5. Experiment 4

Similar to Experiment 2, we retested the 9-year old participants in Experiment 3 one and two years later in order to show a developmental improvement in the recognition of faces from a longitudinal study.

5.1. Method

The same procedure described in Experiment 2, except it was conducted on the 9-year old participants we tested in Experiment 3. The one change to the method was that we used two additional sets of faces from the Minear and Park (2004) database to ensure that the faces were unfamiliar to our participants. These were of the same age-range as those described in Experiment 3. This was a 2 × 3 within-subjects design with the factors: orientation of the face and participant age (9-, 10-, and 11-years of age).

5.2. Results

The data was treated in the same way as in Experiment 1 and are presented in Table 8. The recognition accuracy (d') data were

Table 8
Mean hit rate, false alarm rate, recognition accuracy (*d*'), response bias (*C*) for upright and inverted age-matched faces and relative FIE split by participant age for Experiments 2 and 4. Standard error is shown in parenthesis.

		Participant Age		
		9-years Adult Faces	10-years Adult Faces	11-years Adult Faces
Recognition Accuracy (d')	Upright Faces	1.29 (0.12)	1.70 (.11)	1.99 (0.21)
	Inverted Faces	1.27 (0.15)	1.28 (.09)	0.90 (0.10)
Hit Rate	Upright Faces	.69 (.03)	.75 (.02)	.81 (.02)
	Inverted Faces	.68 (.04)	.72 (.02)	.64 (.03)
False Alarm Rate	Upright Faces	.24 (.02)	.19 (.02)	.21 (.04)
	Inverted Faces	.26 (.02)	.26 (.02)	.31 (.03)
Response Bias (C)	Upright Faces	.11 (.06)	.14 (.08)	.00 (.09)
	Inverted Faces	.07 (.10)	.04 (.04)	.07 (.07)
FIE		.08 (.08)	.13 (.04)	.35 (.09)

subjected to a 2 × 3 within-subjects ANOVA with the factors orientation and age. Critically, these effects interacted, $F(2, 42) = 8.90$, $MSE = 0.36$, $p = .001$, $\eta_p^2 = .30$, $B_{10} = 3.22 \times 10^{11}$. This interaction was decomposed by conducting two one-way ANOVAs, one for upright faces and one for inverted faces. The improvement in recognition accuracy for upright faces was significant, $F(2, 42) = 6.07$, $MSE = 0.44$, $p = .005$, $\eta_p^2 = .22$, $B_{10} = 14.67$, but smaller for inverted faces, $F(2, 42) = 3.24$, $MSE = 0.32$, $p = .049$, $\eta_p^2 = .13$, $B_{10} = 2.32$. A trend analysis was conducted and this revealed that the improvement in face recognition accuracy for upright faces was linear, $F = 9.24$, $MSE = 0.57$, $p = .006$, $\eta_p^2 = .31$, with no other significant developmental trajectories (all $ps > .689$).

A one-way ANOVA was run on the relative measure of the FIE, revealing a significant effect of age, $F(2, 42) = 3.32$, $MSE = 0.13$, $p = .046$, $\eta_p^2 = .14$, $B_{10} = 3.27$. One-sample *t*-tests showed that the FIE was not present in the 9-year old participants, $t(21) = 1.02$, $p = .545$, effect size $r = .22$, $B_{10} = 0.29$, but was present in 10-year old, $t(21) = 3.18$, $p = .005$, effect size $r = .57$, $B_{10} = 16.65$, and 11-year old participants, $t(21) = 3.74$, $p = .001$, effect size $r = .63$, $B_{10} = 251.31$. A linear trend was observed, $F = 4.83$, $MSE = 0.16$, $p = .039$, $\eta_p^2 = .19$. No other contrast pattern was observed (all $ps > .35$) (Fig. 4).

Finally, a parallel ANOVA was run on the response bias (*C*) data, revealing no effect of age, $F(2, 42) = 0.33$, $MSE = 0.13$, $p = .718$, $\eta_p^2 = .02$, $B_{10} = 0.37$. There was also no effect of orientation, $F(1, 21) = 0.16$, $MSE = 0.11$, $p = .695$, $\eta_p^2 = .01$, $B_{10} = 0.21$. The interaction was not significant, $F(2, 42) = 0.56$, $MSE = 0.13$, $p = .578$, $\eta_p^2 = .03$, $B_{10} = 0.25$.

5.3. Discussion

In this Experiment, we have demonstrated a significant improvement in the recognition of adult faces with development. This development is much greater for upright faces than for inverted faces. The FIE was not observed when our participants were 9-years of age, but was present when they were older. This longitudinal study replicates the step change in the emergence of the FIE around the age of 10-years.

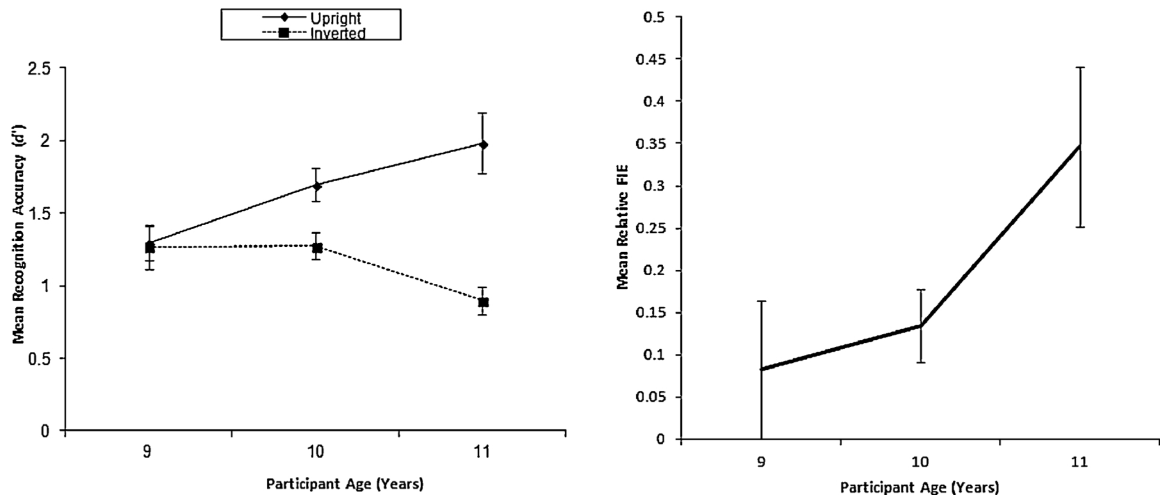

Fig. 4. Top panel: Mean recognition accuracy (*d*') for upright and inverted adult faces split by participant age (Experiment 4). Bottom panel: Mean FIE for adult faces split by participant age (Experiment 4). Error bars show standard error.

6. Comparison between experiments

While there are many limitations regarding the comparison across different studies, especially considering subtle methodological differences across the studies,[6] we compared the face recognition performance across Experiments 1 and 3 in a 12 × 2 × 2 Mixed ANOVA with the factors: age of participant, orientation of stimuli, and type of stimuli (age-matched, Experiment 1; adult faces, Experiment 3). Prior to this analysis, we compared the average stimulus pixelwise complexities for the children's faces and the adult faces and found that they were not significantly different ($p > .32$). Similarly, attractiveness ratings did not differ significantly across the adult and the children's faces ($p > .16$).

Figs. 1 and 3 appear to show similar patterns that indicate the recognition accuracy of upright adult and child faces increases faster than that of inverted faces. In this comparative analysis, there are two main effects of interest: the main effect of stimuli type and the interaction between stimuli type and orientation of the stimuli. We ran this analysis on the d' recognition accuracy measure first. This revealed a main effect of stimuli type, $F(1, 720) = 6.07$, $MSE = 0.62$, $p = .014$, $\eta_p^2 = .01$, $B_{10} = 3.01$. On average, performance with age-matched stimuli ($M = 1.48$, $SE = 0.03$) was approximately 10% greater than adult faces ($M = 1.38$, $SE = 0.03$) replicating the own-age bias (Anastasi & Rhodes, 2006). This effect interacted with orientation of the face, $F(1, 720) = 6.63$, $MSE = 0.50$, $p = .010$, $\eta_p^2 = .01$, $B_{10} = 5497$, replicating Hills (2012). This interaction was revealed through a larger face-inversion effect on own-age faces (mean difference = 0.84, $SE = .04$), $F(1, 479) = 252.57$, $MSE = 0.67$, $p < .001$, $\eta_p^2 = .35$, $B_{10} = 1.45 \times 10^{100}$, than other-age faces (mean difference = 0.65, $SE = 0.06$), $F(1, 263) = 95.14$, $MSE = 0.58$, $p < .001$, $\eta_p^2 = .27$, $B_{10} = 1.15 \times 10^{24}$. All these results indicate that the use of other-age faces when studying face recognition in children may well be underestimating performance.

Finally, the same interaction between stimuli type and orientation was observed in the response bias data, $F(1, 720) = 16.56$, $MSE = 0.17$, $p < .001$, $\eta_p^2 = .02$, $B_{10} = 2.79 \times 10^{17}$. The effect of stimuli type was larger for inverted (mean difference = .12, $SE = .03$), $F(1, 742) = 22.49$, $MSE = 0.11$, $p < .001$, $\eta_p^2 = .03$, $B_{10} = 10425.45$, than upright faces (mean difference = .07, $SE = .04$), $F(1, 742) = 2.76$, $MSE = 0.26$, $p = .097$, $\eta_p^2 < .01$, $B_{10} = 0.75$.

7. General discussion

Through four Experiments we have shown that there was a significant improvement in the recognition of upright faces with development (both longitudinally and cross sectionally) consistent with previous research (Blaney & Winograd, 1978; Carey & Diamond, 1977; Carey et al., 1980; Diamond & Carey, 1977; Goldstein, 1965; Goldstein & Chance, 1964; Saltz & Sigel, 1967). The longitudinal data showed numerically consistent results with the cross-sectional data for participants within three years. This indicates that recognition improvements over three years are relatively small but detectable with a suitable experiment. Here we shall first summarise the findings across the experiments before relating the findings to theoretical discussion presented in the introduction.

Across all measures, the improvement in the recognition of upright faces was approximately linear (Carey, 1978; Flin, 1980, 1985), though had significant quadratic and cubic functions. This pattern is indicative of a slow gradual overall improvement in face recognition, but with a step change causing the cubic and quadratic functions. The statistics highlight that that there is less improvement in face recognition during mid-adolescence. This is consistent with findings of other cognitive abilities (e.g., Flin, 1983). Furthermore, there is quick rise in face recognition abilities between the age of 9- and 12-years. To highlight this step change in abilities between these ages, we have shown that the improvement in face recognition at these ages is exclusively linear (in the longitudinal data, Experiment 2). This indicates that this is the age driving the cubic function. These findings are interesting as they present an often ignored aspect within developmental work: while the age-related developmental trajectory may produce an overall linear pattern, other functions are present indicating that factors exist that may enhance or inhibit such development. These factors maybe individual difference variables, relating to hormonal, maturational, or environmental changes surrounding the individual.

This linear improvement in the recognition accuracy of upright faces was significantly greater than the approximately linear improvement in the recognition of inverted faces, consistent with de Heering et al. (2012). In other words, the improvements to the recognition of upright faces were greater than the improvements in the recognition of inverted faces. This is entirely consistent with the notion of expertise developing for the recognition of faces that are encountered in the most common orientation. Relating to the general models of perceptual and skill development presented in the introduction, this is akin to developmental induction due to experience with upright faces outweighing experience with inverted faces and therefore leading to enhanced learning how to process these stimuli.

The FIE, measured in relative terms, showed a different pattern of development. This showed no increase in the FIE until the age of 9-years, followed by an increase in the FIE over two years before plateauing. A close inspection of de Heering et al.'s (2012) data reveals a similar pattern. These data are consistent with an "all-or-none"' late maturational model of the FIE (Carey et al., 1980; Itier & Taylor, 2004) with a slower development of configural coding than featural coding (Mondloch, Le Grand, & Maurer, 2002). Here, there is statistical evidence for this assertion: the change in the FIE with development followed primarily a cubic function. We present this assertion against a background that the measure of FIE we used should have enhanced the likelihood of us finding it in the

[6] Crucially, in Experiment 3 the change in image from learning to test involved recognising individual faces across changes of expression. Whereas in Experiment 1, the change in images was smaller. It could be also that matching face identity across expression changes is more difficult for children than adults similar to the effects of paraphernalia changes (Frieire & Lee, 2001). Such differences could explain the observed effects in this study.

youngest age groups. Since we did not, it clearly shows that configural coding (as measured by the FIE) is not employed for faces until about the age of 10-years.

The present data also indicate that there is a general increase in memory for faces with development (see e.g., Brown, 1975; Chi, 1977; Dempster, 1981; Flavell, 1977; Kail, 1992) given that hit rate increased with age. However, memory for inverted faces did not improve with age. This is more consistent with Weigelt et al.'s (2013) data indicating a domain-specific increase in face memory with age rather than a general improvement in cognitive functioning (Crookes & McKone, 2009; McKone & Boyer, 2006).

In the introduction, a number of potential explanations were presented for the FIE. Our data indicate a sharp increase in the FIE between the ages of 9 and 10 years. This indicates an increase in expertise in face recognition during this period. This expertise is likely to be the form of configural processing known as holistic processing (Maurer et al., 2002a). One theory of holistic processing is that inversion disrupts the perceptual field such that the entire upright face cannot be sampled from a single central fixation (Rossion, 2008, 2009). These effects mirror those observed in other areas of perceptual expertise that require many years of experience to achieve (Charness, Reingold, Pomplun, & Stampe, 2001; Ericsson, Krampe, & Tesch-Römer, 1993; Kundel & Nodine, 1975). An inverted face, on the other hand, cannot be processing from a central fixation: this has been found in eye-tracking evidence (Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006; Hills, Sullivan, & Pake, 2012; Xu & Tanaka, 2013). This suggests that with development, the perceptual field widens when viewing faces until, around the age of 10-years, it is sufficient to encode an entire face with one fixation. This hypothesis can be easily tested by exploring the eye-movements of children. There is evidence that the features used by children to process faces changes with age. Campbell, Walker, and Baron-Cohen (1995) has shown an external feature advantage at age 7-years that shifts to an internal feature advantage by 9-years of age. Further eye-tracking evidence supports a late development of this internal feature advantage (Kelly et al., 2011; Meaux et al., 2014; Senju, Vernetti, Kikuchi, Akechi, & Hasegawa, 2013). In addition, younger children should show more fixations with wider distribution over the features than older children and adults (Hills, Willis, & Pake, 2013). Indeed, Ge et al. (2008) have shown that older children tend to be able to recognise faces based on the most diagnostic features (the eyes for White faces and the nose for East Asian faces; Ge et al., 2008; Kelly et al., 2011; Liu et al., 2013). Younger children, on the other hand, rely on less diagnostic features. The refinement of eye-movements appears to occur at around 10 years of age.

The development of eye-movements has been interpreted as evidence of the development of face expertise (Ge et al., 2008; Tanaka et al., 2014). It is thought that the progression to more sustained fixation on the diagnostic features may help the development of face processing expertise such that it becomes a more rapid and automatic process (Kelly et al., 2011) consistent with notion that face expertise involves a change in processing style and cognitive encoding (Diamond & Carey, 1977), from local to holistic (Hole, 1994; Tanaka & Farah, 1993) and configural processing (Leder & Bruce, 2000).

The expert perceptual field theory has a great deal of support for it from other areas of congitive, attentional, and perceptual development. Expertise in many domains is associated with greater ability to "chunk" information and create more stable schemas (Braune & Foshay, 1983; Goldstein, 1975) with increased systematisation of knowledge (Karmiloff-Smith & Inhelder, 1974). Such systematised knowledge leads to improved memory when consistent with internal schemas (Simon & Barenfeld, 1969; Simon & Gilmartin, 1973 see also Chase & Simon, 1973a, 1973b; DeGroot, 1965, 1966; DeGroot & Gobet, 1996; Gobet & Simon, 1996a, 1996b). Wider perceptual fields are observed in expert footballers (Williams & Davids, 1997) and quicker processing from fewer fixations and greater chunking is observed in expert chess players (Charness et al., 2001; Reingold, Charness, Pomplun, & Stampe, 2001; Reingold, Charness, Schiltetus, & Stampe, 2001) and radiologists (Kundel & Nodine, 1975). Indeed, the automatisation of perceptual processing (Vurpillot, 1968) and attentional shifting requires less effort than inexpert processing as it is based on an unconscious and automatic process. This takes circa ten years to develop (e.g., Akhtar & Enns, 1989; Brodeur & Enns, 1997; Enns & Brodeur, 1989; Pearson & Lane, 1991) or occurs at the age of 10 years. The data fit with neuroscientific evidence that indicates the pattern of face-specific recruitment of cortical regions is not observed prior to the age of ten years (Aylward & Meltzoff, 2005 but see Golarai, Liberman, & Grill-Spector, 2015) or that the size of face-specific regions is smaller in children than adolescents or adults (Golarai et al., 2007) and that the magnitude of ERPs associated with face perception are different for children than adolescence and adults (Kuefner, De Heering, Jacques, Palmero-Soler, & Rossion, 2010).

The key finding here is that the FIE appears to follow a maturational development (as highlighted by the cubic function), whereas the development of recognising upright faces appears to follow an induction pattern. Of course, the notion that there is ten years of development that is required for face recognition expertise is an overly simplistic argument. It may be that this trend appears at the age of 10 years due to biological and/or social constraints that are vital for expertise to develop rather than experience. The cause of expertise development at the age of 10 years could be due to hormonal changes altering the functioning of the attentional spotlight and eye-movements. Puberty is associated with a number of hormonal changes known to affect cognitive functioning especially those associated with frontal lobe functioning (Blakemore & Choudhury, 2006) due to the significant synaptic pruning that occurs in this region during puberty (Woo, Pucak, Kye, Matus, & Lewis, 1997; Zecevic & Rakic, 2001). Indeed, development during this period involves the development of attentional shifting, abstract reasoning, and inhibition (Yurgelun-Todd, 2007). These factors indicate that at this age, the perceptual system can utilise a more robust schema that inhibits irrelevant dimensions and focuses on the most diagnostic information for recognising faces.

While there is hormonal and biological maturation occurring at this time, there are significant environmental changes that occur at the same age. School environmental changes at age 10–11 from small classes in smaller schools in primary education to larger classes at larger secondary schools (college). This results in exposure to more faces than previously encountered and typically a large number of new faces. This sudden increase in the amount of exemplar's entering the face-space may result in its refinement. Whatever the cause, this mechanism clearly results in a potentially unique and special processing for faces at this age.

There is, of course, a limitation of the present work. By the age of 5-years, children will have been exposed to a significant number of faces right from birth. Indeed, there is evidence that the early visual system is set up to process faces (de Heering et al., 2008). Newborn visual acuity makes faces the single most important visual stimulus leading to early engagement with faces (Coulon, Guellai, & Streri, 2011). This explains why there are implicit measures of face processing in infants, measured using ERPs (de Haan, Johnson, & Halit, 2003) that appear to show differentiation between faces and objects (Peykarjou, Pauen, & Hoehl, 2016; Peykarjou, Wissner, & Pauen, 2017) and even between upright and inverted faces (Halit, De Haan, & Johnson, 2003; Peykarjou & Hoehl, 2013). While these measures do not show the ability to individuate faces, such data indicates significant learning about faces and the development of mechanisms to process faces prior to the age of participants tested here. The influence of this on subsequent face recognition has not been considered in the present study. Infants will not be exposed to as many faces as adults, and are less likely to encounter own-age faces than other-age faces given that infants spend more time with their parents and family members than at schools or nurseries. The method of discriminating between a small number of faces encountered is likely to be based on simple mechanisms, especially if these faces are more varied in terms of age (we are considering an environment where an infant has frequent contact with both parents, older siblings, and grandparents). In these situations, simpler age cues may be sufficient to identify and discriminate faces. However, when older, the number and types of faces that are encountered are likely to be more homogenous and therefore the methods used to discriminate these necessarily need to be more sophisticated.

We used inverted faces in the present study as way of assessing development of inexpert processing. There is no reason to believe that the developmental trajectory for the recognition of objects would be different to that of inverted faces given that inverted faces are processed using the inexpert featural processing manner that is afforded to objects. We have also shown the effects of inversion to be greater when testing faces of one's own age relative to other-age faces. However, since we did not compare the recognition of own-age and other-age faces in the same children, we cannot rule out potential participant and stimuli effects for this difference.

One caveat with the explanations presented here is that these results are specific to the FIE during recognition. A general theory that explains all effects of inversion would require a great deal more specification than is currently available. Furthermore, there is no reason to believe that inversion affects all aspects of face processing and configural coding in the same way. Another limitation with the present study is that it cannot address what the maturational mechanism actually is, or indeed if it is enhancement, facilitation, or maintainence (Coren, Ward, & Enns, 1999). Previously proposed mechanisms are the qualitative shift from featural to holistic processing (Carey & Diamond, 1994). This is entirely possible, but the cause of this shift is not clear. Similarly, it could be that there is a shift from using certain (presumably featural) dimensions to other (more configural) dimensions of face-space (Valentine, 1991). Indeed, the data presented by Hills, Holland, and Lewis (2010) is consistent with this presumption: children under 10-years of age can be adapted to facial distortions (a single eye moved) that adults cannot because they are coding faces according to each feature individually rather than the features together. Alternatively, it is plausible that it requires ten years for schema (in this case, prototype face) to become more fixed or robust. All of these interpretations lack the precise description of the mechanism and cause for this change.

In conclusion, the expert processing mechanism (potentially configural or holistic processing) develops. However, there is a maturational component to this development, suggesting a stage where it is not largely utilised (prior to 10-years of age) to a stage where it is largely employed. This coding switch idea has been hypothesised before (Carey & Diamond, 1977), but adequate tests of it have not been forthcoming until now. This is not to suggest that configural processing cannot be done prior to the age of 10, but that its reliable and constant use is not likely before that age. Indeed, the effects of inversion may be unrelated to other measures of "configural" coding (such as the parts and wholes test; Konar, Bennett, & Sekuler, 2010; Wilhelm et al., 2010). Nevertheless, the results from this study are reliable given the internal replication and the robust statistical procedures employed.

References

Adams-Price, C. (1992). Eyewitness memory and aging: Predictors of accuracy in recall and person recognition. *Psychology and Aging, 7*, 602–608.

Akhtar, N., & Enns, J. T. (1989). Relations between covert orienting and filtering in the development of visual attention. *Journal of Experimental Child Psychology, 48*, 315–334.

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review, 12*(6), 1043–1047.

Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology, 8*, 237–253.

Aylward, E. H., & Meltzoff, A. N. (2005). Brain activation during face perception: Evidence of a developmental change. *Journal of Cognitive Neuroscience, 17*, 308–319.

Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35*(8), 1089–1105. http://dx.doi.org/10.1068/p5547.

Baudouin, J. Y., 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.

Bennetts, R. J., Murray, E., Boyce, T., & Bate, S. (2017). Prevalence of face recognition deficits in middle childhood. *The Quarterly Journal of Experimental Psychology, 70*(2), 234–258.

Benton, A. L., Sivan, A. B., Hamsher, K., Vareny, N. R., & Spreen, O. (1983). Facial recognition: Stimulus and multiple choice pictures. In A. L. Benton, A. B. Sivan, K. D. S. Hamsher, N. R. Varney, & O. Spreen (Eds.). *Contributions to neuropsychological assessment*New York: Oxford University Press pp. 30–40.

Betts, J., McKay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: The effect of age and task load. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 12*(3), 205–221. http://dx.doi.org/10.1080/09297040500488522.

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*, 296–312.

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

Bobak, A. K., Parris, B. A., Gregory, N. J., Bennetts, R. J., & Bate, S. (2017). Eye-movement strategies in developmental prosopagnosia and "super" face recognition. *The Quarterly Journal of Experimental Psychology, 70*(2), 201–217.

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 investigte face recognition. *Perception, 30*, 85–94. http://dx.doi.org/10.1068/p3059.

Braune, R., & Foshay, W. R. (1983). Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. *Instructional Science, 12*, 121–145.

Brodeur, D. A., & Enns, J. T. (1997). Lifespan differences in covert visual orienting. *Canadian Journal of Experimental Psychology, 51*, 20–35.
Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.). *Advances in child development and behaviour*. New York: Academic Press.
Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73*, 105–116.
Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science, 10*(4), 482–496.
Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science, 11*, 429–433.
Campbell, R., Walker, J., & Baron-Cohen, S. (1995). The development of differential use of inner and outer face features in familiar face identification. *Journal of Experimental Child Psychology, 59*(2), 196–210. http://dx.doi.org/10.1006/jecp.1995.1009.
Carey, S. (1978). A case study: Face recognition. In E. Walker (Ed.). *Explorations in the biology of language*. Montgomery, VT: Bradford Books.
Carey, S. (1981). The development of face perception. In G. Davies, H. D. Ellis, & J. Shepherd (Eds.). *Perceiving and remembering faces*. London: Academic Press.
Carey, S., & Diamond, R. (1977). From piecemeal to configurational representations of faces. *Science, 195*, 312–314.
Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*, 253–274.
Carey, S., Diamond, R., & Woods, B. (1980). The development of face recognition – A maturational component? *Developmental Psychology, 16*, 257–269.
Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memoryy & Cognition, 29*, 1146–1152.
Chase, W. G., & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology, 4*, 55–81.
Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.). *Visual information processing*. London: Academic Press.
Chi, M. T. H. (1977). Age differences in memory span. *Journal of Experimental Child Psychology, 23*, 266–281. http://dx.doi.org/10.1016/0022-0965(77)90104-7.
Coren, S., Ward, L. M., & Enns, J. T. (1999). *Sensation and perception.* Boston, MA: Harcourt College Publishing.
Coulon, M., Guellai, B., & Streri, A. (2011). Recognition of unfamiliar talking faces at birth. *International Journal of Behavioral Development, 35*(3), 282–287.
Crookes, K., & McKone, E. (2009). Early maturity of face recognition. No childhood development of holistic processing, novel face encoding, or face = space. *Cognition, 111*, 219–247.
Croydon, Abigail, Pimperton, Hannah, Ewing, Louise, Duchaine, Brad C., & Pellicano, Elizabeth (2014). The Cambridge Face Memory Test for Children (CFMT-C): a new tool for measuring face recognition skills in childhood. *Neuropsychologia, 62*, 60–67.
de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology, 51*(1), 45–58.
de Heering, A., Turati, C., Rossion, B., Bulf, H., Goffaux, V., & Simion, F. (2008). Newborns' face recognition is based on spatial frequencies below 0.5 cycles per degree. *Cognition, 106*(1), 444–454.
de Heering, A., Rossion, B., & Maurer, D. (2012). Developmental changes in face recognition during childhood: Evidence from upright and inverted faces. *Cognitive Development, 27*, 17–27.
DeGroot, A. D. (1965). *Thought and choice in chess.* The Hague: Mouton.
DeGroot, A. D. (1966). Perception and memory versus thought. In B. Kleinmuntz (Ed.). *Problem solving.* New York: Wiley.
DeGroot, A. D., & Gobet, F. (1996). *Perception and memory in chess: Heuristics of the processional eye.* Assen: Van Gorcum.
Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63–100. http://dx.doi.org/10.1037/0033-2909.89.1.63.
Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology, 23*, 1–22.
Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science, 6*(3), 274–290.
Dienes, Z. (2015). *Making the most of your data with Bayes.* Available at: . Accessed: 27th February 2016 http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/Bayes.htm.
Edmonds, A. J., & Lewis, M. B. (2007). The effect of rotation of configural encoding in a face-matching task. *Perception, 36*, 446–460.
Ellis, H. D. (1986). Processes underlying face recognition. In R. Bruyer (Ed.). *The neuropsychology of face perception and facial expression*. Hillsdale, NJ: Erlbaum.
Ellis, H. D. (1992a). The development of face processing skills. *Philosophical Transactions of the Royal Society B, 335*, 105–111.
Ellis, H. D. (1992b). The development of face processing skills. In V. Bruce, A. Cowey, A. W. Ellis, & D. I. Perrett (Eds.). *Processing the facial image*. Oxford: Clarendon Press.
Ellis, H. D., Shepherd, J., & Bruce, A. (1973). The effects of age and sex upon adolescent's recognition of faces. *Journal of Genetic Psychology, 123*, 173–174.
Enns, J. T., & Brodeur, D. A. (1989). A developmental study of covert orientation to peripheral visual cues. *Journal of Experimental Child Psychology, 48*, 171–189.
Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review, 100*(3), 363.
Fagan, J. F. (1972). Infants' recognition memory for faces. *Journal of Experimental Child Psychology, 14*, 453–476.
Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.
Feinman, S., & Entwisle, D. R. (1976). Children's ability to recognize other children's faces. *Child Development, 47*, 506–510.
Flavell, J. H. (1977). *Cognitive development.* Englewood Cliffs, NJ: Prentice-Hall.
Flin, R. H. (1980). Age effects in children's memory for unfamiliar faces. *Developmental Psychology, 16*, 373–374. http://dx.doi.org/10.1037/0012-1649.16.4.373.
Flin, R. H. (1983). *The development of face recognition. Unpublished PhD thesis.* University of Aberdeen.
Flin, R. H. (1985). Development of face recognition: An encoding switch? *British Journal of Psychology, 76*, 123–134. http://dx.doi.org/10.1111/j.2044-8295.1985.tb01936.x.
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. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*, 159–170.
Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12*, 495–504.
Ge, L., Anzures, G., Wang, Z., Kelly, D. J., Pascalis, O., Quinn, P. C., ... Lee, K. (2008). An inner face advantage in children's recognition of familiar peers. *Journal of Experimental Child Psychology, 101*(2), 124–136. http://dx.doi.org/10.1016/j.jecp.2008.05.006.
Germine, L., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition, 118*, 201–210.
Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: Local and relational distinctiveness effects in 7-year-olds. *Visual Cognition, 10*(7), 769–793. http://dx.doi.org/10.1080/13506280344000022.
Gobet, F., & Simon, H. A. (1996a). Recall of rapidly present random board positions is a function of skill. *Psychonomic Bulletin and Review, 3*, 159–163.
Gobet, F., & Simon, H. A. (1996b). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1–40.
Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L., Gabrieli, J. D., & Grill-Spector, K. (2007). Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience, 10*(4), 512–522. http://dx.doi.org/10.1038/nn1865.
Golarai, G., Liberman, A., & Grill-Spector, K. (2015). Experience shapes the development of neural substrates of face processing in human ventral temporal cortex. *Cerebral Cortex, 27*, 1229–1244.
Goldstein, A. G. (1965). Learning of inverted and normally orientated faces in children and adults. *Psychonomic Science, 3*, 447–448.
Goldstein, A. G. (1975). Recognition of inverted photographs of faces in children and adults. *Journal of Genetic Psychology, 127*, 109–123.
Goldstein, A. G., & Chance, J. E. (1964). Recognition of children's faces. *Child Development, 35*, 129–136.
Halit, H., De Haan, M., & Johnson, M. H. (2003). Cortical specialisation for face processing: Face-sensitive event-related potential components in 3-and 12-month-old infants. *Neuroimage, 19*(3), 1180–1193.
Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review, 16*(2), 264–269.
Haushofer, J., Livingstone, M. S., & Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS*

Biology, 6(7), e187.
Hay, D. C., & Cox, R. (2000). Developmental changes in the recognition of faces and facial features. *Infant and Child Development, 9*, 199–212.
Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology, 48*, 499–508.
Hills, P. J., & Lewis, M. B. (2011). The own-age face recognition bias in children and adults. *The Quarterly Journal of Experimental Psychology, 64*, 17–23. http://dx.doi.org/10.1080/17470218.2010.537926.
Hills, P. J., Holland, A. M., & Lewis, M. B. (2010). After effects for face attributes with different natural variability: Children are more adaptable than adolescents. *Cognitive Development, 25*, 278–289.
Hills, P. J., Sullivan, A. J., & Pake, J. M. (2012). Aberrant first fixations when looking at inverted faces in various poses: The result of the centre-of-gravity effect? *British Journal of Psychology, 103*(4), 520–538. http://dx.doi.org/10.1111/j.2044-8295.2011.02091.x.
Hills, P. J., Willis, S. F., & Pake, J. M. (2013). Eye-tracking face recognition in children: Effects of age, familiarity, and orientation. *Presented at the BPS joint cognitive and developmental sections conference*.
Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*(1), 65–74.
Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception, 31*(10), 1221–1240. http://dx.doi.org/10.1068/p3252.
Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science, 3*, 1173–1181. http://dx.doi.org/10.1111/j.1551-6709.2009.01048.x.
Itier, R. J., & Taylor, M. J. (2004). Effects of repetition and configural changes on the development of face recognition processes. *Developmental Science, 7*, 469–487.
Jeffreys, H. (1961). *Theory of probability.* Oxford University Press.
Joseph, J. E., Gathers, A. D., Liu, X., Corbly, C. R., Whitaker, S. K., & Bhatt, R. S. (2006). Neural developmental changes in processing inverted faces. *Cognitive, Affective, and Behavioral Neuroscience, 6*, 223–235.
Kail, R. (1992). Processing speed, speech rate, and memory. *Developmental Psychology, 28*, 899–904. http://dx.doi.org/10.1037/0012-1649.28.5.899.
Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68*, B1–B11.
Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition, 3*, 195–212.
Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental Science, 14*(5), 1176–1184. http://dx.doi.org/10.1111/j.1467-7687.2011.01067.x.
Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science, 21*(1), 38–43.
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*(4), 811.
Kuefner, D., De Heering, A., Jacques, C., Palmero-Soler, E., & Rossion, B. (2010). Early visually evoked electrophysiological responses over the human brain (P1, N170) show stable patterns of face-sensitivity from 4 years to adulthood. *Frontiers in Human Neuroscience, 3*, 67.
Kundel, H. L., & Nodine, C. F. (1975). Interpreting chest radiographs without visual search. *Radiology, 116*, 527–532.
Lawrence, K., Bernstein, D., Pearson, R., Mandy, W., Campbell, R., & Skuse, D. (2008). Changing abilities in recognition of unfamiliar face photographs through childhood and adolescence: Performance on a test of non-verbal immediate memory (Warrington RMF) from 6 to 16 years. *Journal of Neuropsychology, 2*, 27–45.
Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 53*(2), 513–536.
Lewis, M. B., & Glenister, T. E. (2003). A sideways look at configural encoding: Two different effects of face rotation. *Perception, 32*, 7–14.
Li, J., Tian, M., Fang, H., Xu, M., Li, H., & Liu, J. (2010). Extraversion predicts individual differences in face recognition. *Communicative & Integrative Biology, 3*(4), 295–298.
Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality and Social Psychology Bulletin, 7*, 269–276.
Liu, S., Anzures, G., Ge, L., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2013). Development of recognition of face parts from unfamiliar faces. *Infant and evelopment, 22*(2), 165–179.
Macchi Cassia, V. (2011). Age biases in face processing: The effects of experience across development. *British Journal of Psychology, 102*(4), 816–829.
MacMillan, N. A., & Creelman, C. D. (2010). *Detection theory: A user's guide.* New York: Cambridge University Press.
Maurer, D., Le Grand, R., & Mondloch, C. J. (2002a). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255–260.
McKone, E., & Boyer, B. I. (2006). Sensitivity of 4-year-olds to featural and second-order relational changes in face distinctiveness. *Journal of Experimental Child Psychology, 94*, 134–162.
Meaux, E., Hernandez, N., Carteau-Martin, I., Martineau, J., Barthélémy, C., Bonnet-Brilhault, F., & Batty, M. (2014). Event-related potential and eye tracking evidence of the developmental dynamics of face processing. *European Journal of Neuroscience, 39*(8), 1349–1362. http://dx.doi.org/10.1111/ejn.12496.
Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*(7), 6.
Michel, C., Caldara, R., & Rossion, B. (2006). Same-face faces are perceived more holistically than other-race faces. *Visual Cognition, 14*, 55–73.
Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17*, 608–615. http://dx.doi.org/10.1111/j.1467-9280.2006.01752.x.
Minear, M. E., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods. Instruments and Computers, 36*, 630–633. http://dx.doi.org/10.3758/BF03206543.
Mondloch, C. J., & Desjarlais, M. (2010). The function and specificity of sensitivity to cues to facial identity: An individual-differences approach. *Perception, 39*(6), 819–829.
Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*(5), 553–566.
O'Hearn, K., Schroer, E., Minshew, N., & Luna, B. (2010). Lack of developmental improvement on a face memory task during adolescence in autism. *Neuropsychologia, 48*(13), 3955–3960. http://dx.doi.org/10.1016/j.neuropsychologia.2010.08.024.
Pastò, L., & Burack, J. A. (1997). A developmental study of visual attention: Issues of filtering efficiency and focus. *Cognitive Development, 12*(4), 523–535. Retrieved from http://www.scopus.com/inward/record.url?eid = 2-s2.0-0037926152&partnerID = tZOtx3y1.
Pearson, D. A., & Lane, D. M. (1991). Visual attention movements: A developmental study. *Child Development, 61*, 1779–1795.
Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14*, 618–622.
Peykarjou, S., & Hoehl, S. (2013). Three-month-olds' brain responses to upright and inverted faces and cars. *Developmental Neuropsychology, 38*(4), 272–280.
Peykarjou, S., Pauen, S., & Hoehl, S. (2016). 9-month-old infants recognize individual unfamiliar faces in a rapid repetition ERP paradigm. *Infancy, 21*(3), 288–311.
Peykarjou, S., Wissner, J., & Pauen, S. (2017). Categorical erp repetition effects for human and furniture items in 7-month-old infants. *Infant and Child Development, 26*(5).
Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology, 3*, 559. http://dx.doi.org/10.3389/fpsyg.2012.00559.
Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science, 12*, 48–55.
Reingold, E. M., Charness, N., Schiltetus, R. S., & Stampe, D. M. (2001). Perceptual automaticity in expert chess players: Parallel encoding of chess relations. *Psychonomic Bulletin & Review, 8*, 504–510.
Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 43*(12), 1961.
Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin, 138*(1), 146.
Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*, 274–289. http://dx.doi.org/10.1016/j.actpsy.2008.02.003.

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132*, 300–312. http://dx.doi.org/10.1016/j.actpsy.2009.08.002.
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237.
Saltz, E., & Sigel, I. E. (1967). Concept overdiscrimination in children. *Journal of Experimental Psychology, 73*, 1–8.
Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child Development, 71*, 391–401.
Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 904–915.
Senju, A., Vernetti, A., Kikuchi, Y., Akechi, H., & Hasegawa, T. (2013). Cultural modulation of face and gaze scanning in young children. *PLoS One, 8*(8), http://dx.doi.org/10.1371/journal.pone.0074017.
Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*(2), 221–242.
Simon, H. A., & Barenfeld, M. (1969). Information-processing analysis of perceptual processes in problem solving. *Psychological Review, 76*, 473–483.
Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5*, 29–46.
Skoczenski, A. M., & Norcia, A. M. (2002). Late maturation of visual hyperacuity. *Psychological Science, 13*(6), 537–541.
Susilo, T., Germine, L., & Duchaine, B. (2013). Face recognition ability matures late: Evidence from individual differences in young adults. *Journal of Experimental Psychology: Human Perception & Performance, 39*, 1212–1217.
Swets, J. A. (1966). *Signal detection and recognition by human observers.* New York: John Wiley & Sons Inc.
Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A*, 225–246.
Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition, 25*, 583–592.
Tanaka, J. W., Quinn, P. C., Xu, B., Maynard, K., Huxtable, N., Lee, K., & Pascalis, O. (2014). The effects of information type (features vs. configuration) and location (eyes vs. mouth) on the development of face perception. *Journal of Experimental Child Psychology, 124*(1), 36–49. http://dx.doi.org/10.1016/j.jecp.2014.01.001.
Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race on face recognition. *Quarterly Journal of Experimental Psychology, 43A*, 161–204.
Vurpillot, E. (1968). The development of scanning strategies and their relation to visual differentiation. *Journal of Experimental Child Psychology, 6*, 632–650.
Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*(2), 169–177.
Want, S. C., Pascalis, O., Coleman, M., & Blades, M. (2003). Face facts: Is the development of face recognition in middle childhood really so special? In O. Pascalis (Ed.). *The development of face processing in infancy and early childhood: Current perspectives*Hauppauge NY: Nova Science Publishers pp. 207–221.
Weigelt, S., Koldewyn, K., Dilks, D. D., Bala, B., McKone, E., & Kanwisher, N. (2013). Domain-specific development of face memory but not face perception. *Developmental Science,* 1–12.
Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., & Sommer, W. (2010). Individual differences in perceiving and recognizing faces—One element of social cognition. *Journal of Personality and Social Psychology, 99*(3), 530.
Williams, A. M., & Davids, K. (1997). Visual search strategy, selective attention, and expertise in soccer. *Reseach Quarterly for Exercise and Sport, 68*, 111–128.
Woo, T. U., Pucak, M. L., Kye, C. H., Matus, C. V., & Lewis, D. A. (1997). Peripubertal refinement of the intrinsic and associational circuitry in monkey prefrontal cortex. *Neuroscience, 80*, 1149–1158.
Xu, B., & Tanaka, J. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. *Journal of Vision, 13*, 2. http://dx.doi.org/10.1167/13.2.22.
Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.
Young, A. W., & Bion, P. J. (1981). Identification and storage of line drawings presented to the left and right cerebral hemispheres of adults and children. *Cortex, 17*, 459–463.
Yurgelun-Todd, D. (2007). Emotional and cognitive changes during adolescence. *Current Opinion in Neurobiology, 17*, 251–257.
Zecevic, N., & Rakic, P. (2001). Development of layer I neurons in the primate cerebral cortex. *Journal of Neuroscience, 21*, 5607–5619.