Orientation effects support specialist processing of upright unfamiliar faces in children and adults.

Louise Ewing[1,2], Inês Mares[3], S. Gareth Edwards[1], & Marie L. Smith[3]

[1] School of Psychology, University of East Anglia

[2] School of Psychological Science, University of Western Australia

[3] School of Psychological Sciences, Birkbeck College, University of London

Corresponding author:

Louise Ewing

School of Psychology

University of East Anglia, Norwich Research Park

NR4 7TJ, United Kingdom

l.ewing@uea.ac.uk

Data supporting these findings is available at

https://osf.io/8qbv7/?view_only=2b89be08cbd949dcb169b267eb48860d

Abstract

It is considerably harder to generalize identity across different pictures of unfamiliar faces, compared to familiar faces. This finding hints strongly at qualitatively distinct processing of unfamiliar face stimuli – for which we have less expertise. Yet the extent to which face selective vs. generic visual processes drive outcomes during this task has yet to be determined. To explore the relative contributions of each, we contrasted performance on a version of the popular 'Telling Faces Together' unfamiliar face matching task, implemented in both upright and inverted orientations. Furthermore, we included different age groups [132 British children aged 6 to 11-years (69.7% Caucasian), plus 37 British Caucasian adults] to investigate how participants' experience with faces *as a category* influences their selective utilisation of specialised processes for unfamiliar faces. Results revealed that unfamiliar face matching is highly orientation-selective. Accuracy was higher for upright compared to inverted faces from 6 years of age, which is consistent with selective utilisation of specialised processes for upright vs. inverted unfamiliar faces during this task. The effect of stimulus orientation did not interact significantly with age, and there was no graded increase in the magnitude of inversion effects observed across childhood. Still, a numerically larger inversion effect in adults compared to children provides a degree of support for developmental changes in these specialised face abilities with increasing age/experience. Differences in the pattern of errors across age groups are also consistent with a qualitative shift in unfamiliar face processing that occurs some time after eleven years of age.

Public Health Significance Statement

This study supports a contribution of specialized processing mechanisms for unfamiliar face processing that operate selectively for faces presented upright, compared to inverted in a similar manner to those seen for familiar faces. Observed orientation effects did not increase between 6 and 11 years of age, but were much stronger in adults, which provides a degree of support for developmental changes in specialized unfamiliar face abilities with increasing age/experience.

Although humans have remarkably sophisticated face processing abilities, not all faces are processed equally well. Errors with *unfamiliar* faces are particularly common; with performance compromised by superficial changes in the image, e.g., pose, expression, and/or lighting (see Bruce et al., 1982; Burton et al., 2010). This fragility contrasts starkly with our more robust judgments of familiar faces (Burton et al., 1999), which has prompted lively debate about the extent to which our widely assumed status as face 'experts' extends only to *familiar* face stimuli (see Young & Burton, 2018; Rossion, 2018). Elucidating any distinct mechanisms utilised for familiar vs unfamiliar faces, despite them constituting equivalent visual inputs, is critically important for a comprehensive understanding of the processes that underpin our ability to recognise faces.

The elegant "Telling Faces Together" matching task (Andrews, Jenkins, Cursitor & Burton, 2915; Jenkins et al., 2011) highlights the challenges posed when trying to 'recognise' faces that are unfamiliar. Participants must determine how many distinct identities are present in a large set of unique face images. Matching/sorting images of each different identity into sets requires participants to differentiate within-person variability (different images of the same person) from between-person variability (images of different people). Matching performance is typically poor, with errors reflecting overestimation of the number of identities present (thinking two different images of the same person are separate identities), rather than misidentification (confusing two identities as a single person).

Nevertheless, we *are* able to process and recognize unfamiliar faces to some degree (if not as well as familiar faces), and the extent to which we utilise specialised processes for these stimuli remains unclear. The extent to which image variability disrupts unfamiliar face processing is suggestive of a reliance upon low-level stimulus characteristics (e.g., pictorial matching) - although this does not preclude sensitivity to higher-level stimulus properties (see Burton & Jenkins, 2011). The current study seeks to clarify the processes driving performance on this influential task. We probe the relative contributions of specialist (i.e., face selective) vs more generic mechanisms for this task by including a comparison condition that is not predicted to recruit the former: inverted faces.

Inverted faces constitute an ideal match for upright faces with respect to low level visual features, and crucially our reduced experience with this non-canonical orientation permits testing of expert mechanisms that operate selectively for upright faces (Rhodes, Brake & Atkinson, 1993). Pronounced performance differences with upright and inverted stimuli that signal divergent processing for the two orientations have been observed across different standardised memory and perceptual discrimination paradigms for familiar and unfamiliar identities (e.g., Kramer, Jenkins, Young & Burton, 2017; Valentine, 1988). Elucidating the extent to which inversion effects extend also to the ‘telling faces together’ paradigm - which presents such a challenge for unfamiliar face processing – would strengthen reports of specialist processing for unfamiliar faces in behavioural and neuroimaging research (e.g., composite effects, other race effects, neural response differences, inversion effects; see Rossion, 2018).

To our knowledge, only one study has previously explored orientation effects in this paradigm. A between-groups experiment with adults observed that participants’ sensitivity to extrapersonal variation (telling identities apart) was significantly lower when faces were presented inverted cf. upright, while their response criterion did not change (Balas, Gable & Pearson, 2019). Here, we build on this initial finding with a larger adult sample, completing both orientation conditions, and present targeted error analysis to clarify the extent to which orientation effects reflect differences in ability to tell different people apart vs. telling the same person together when presented upright or inverted.

We also investigate how experience influences matching ability for upright and inverted unfamiliar faces. A link between experience and ability is a defining feature of expertise, yet targeted efforts to train (even highly motivated) individuals to improve unfamiliar face recognition and discrimination have had limited success (Towler et al., 2019). Still, there are indications that differences in one's broad experience-base with faces can influence these abilities. For example, individuals from smaller communities (<1000 people/faces) make more accuracy-related errors on a Telling Faces Together task than those from larger communities (Balas & Saville, 2017). Similarly, Laurence, Zhou and Mondloch (2016) demonstrated that adults' relatively greater experience with own-race (cf. other-race) faces selectively boosts unfamiliar face matching abilities for that population. Such findings support some degree of tuning of unfamiliar face processing mechanisms with experience.

Adopting a developmental perspective may provide key insights on this point. Exposure to faces - as a category –increases systematically as we age. Assuming this experience supports the ongoing refinement of face abilities (see McKone, Crookes, Jeffery & Dilks, 2012 for a contrasting view) we can clarify these effects by exploring expertise in children of different ages. Lab studies consistently confirm that face ability improves with age, but clear evidence of concurrently changing processing mechanisms remains elusive. Many propose that the qualitative hallmarks of adult face expertise are present from the youngest ages tested, e.g., configural/holistic processing (Hayden, Bhatt, Reed, Corbly & Joseph, 2007; Macchi Cassia, Turati & Schwarzer, 2011; Turati, Di Giorgio, Bardi & Simion, 2010) and adaptive norm-based coding (Burton, Jeffery, Skinner, Benton, & Rhodes, 2013; Jeffery et al., 2010), which may indicate that specialised processes for faces mature early in development. Yet other work supports a more protracted developmental course for key indices of expertise, e.g., the face inversion effect (Hills & Lewis, 2018). EEG studies have also reported age-related changes in face-selective neural responses (e.g., increasing orientation selectivity), consistent with experience-related development of expertise across childhood (Mares et al., 2020).

Only a small body of research has targeted the typical development of unfamiliar face matching abilities using the Telling Faces Together Task. Consistent with the idea that experience matters for unfamiliar face processing, Neil et al. (2016) observed considerable individual variability in performance across middle childhood (6– 14 years). Overall children found the task more challenging than adults with a general pattern of age-related improvements in the number of perceived identities, i.e., ‘piles’ produced (children perceived significantly more identities than adults). There was also a graded negative developmental shift in the number of misidentification errors made (rare in adults).

Two additional studies used an alternate form of the unfamiliar identity matching task, which features modifications particularly appropriate for children. Rather than sorting a large set of cards/faces into piles, participants consider face stimuli individually; placing those of the target identity (e.g., Jane) inside a ‘house’ that belongs to her, and images of anyone else ‘outside’. Both studies observed significant improvements in matching performance between 5 and 12 years, and adulthood (Baker et al., 2017; Laurence & Mondloch, 2016). Again, such age-related improvements support developmental tuning of the processing mechanisms underpinning unfamiliar face matching – though it remains unclear whether such processes are specialised for faces, or support visual outcomes more generally.

The current study extends prior work by measuring the performance costs associated with *face inversion* in children (6 to 11 years of age) and adults to further probe the processing mechanisms underpinning unfamiliar face matching. Our findings will shed light on the relative balance of specialised and more generic processing mechanisms being utilised for unfamiliar face processing in this influential task, among children and adults. To the extent that attenuated matching performance for inverted (cf. upright) faces indexes the selective utilisation of specialised processes for upright faces, the magnitude of inversion effects across age groups will also reveal how experience with faces *as a category* influences this critical metric. To permit exploratory linkage between the current findings and the broader field of developmental face processing research, we track concurrent changes in another standard measure of face ability: the Cambridge Face Memory Test: Child or Adult version, as appropriate (Croydon et al., 2014; Duchaine & Nakayama, 2006).

Method

Participants

All participants provided verbal assent and written consent was provided by adult participants and children's parents/guardians. Five children were not included in the analysis due to mistakes in upright face catch trials (all 6 and 7 years; see below for more details on catch trials). The final sample comprised 132 children aged 6 to 11 years and 37 adults aged 18 - 34 years, see Table 1. We selected these age groups primarily because of our interest in developmental changes during the early childhood years, when studies widely report changes in performance on lab-based measures of face identity processing (e.g., Croydon et al, 2014). Further, the lower age limit reflected our confidence in participants ability to follow task instructions in the harder inverted faces task, while the higher age limit was set by the maximal age in English primary school (11 years).

Adults were recruited from an undergraduate psychology program and the wider community. All adults were Caucasian, to limit any contribution of other race effects accumulated across the lifespan upon task performance (Laurence, Zhou & Mondloch, 2016; Zhou and Mondloch, 2016). Children were recruited from metropolitan London and Kent classrooms, where it was less practical to constrain ethnicity (69.7% Caucasian, 30.3% other, including Afro Caribbean, Indian, and Pakistani). We report results with the full cohort here and present Caucasian-only results in Supplementary Materials 4, Tables S2 and S3, which did not differ substantially. Data from 14 additional children were excluded due to experimenter error. This work was approved by the University of East Anglia Ethics Committee (Project Refs: 161756/161757). This study was not preregistered.

Table 1. Demographic details of the final sample

	6-7 years N=51	8-9 years N=40	10-11 years N=41	Adults N=37
Mean age: Years (SD)	6.7 (0.5)	8.6 (0.5)	10.4 (0.5)	21.0 (3.6)
Gender: % Female	58.8%	57.5%	61.0%	81.0%
Ethnicity: % Caucasian	76.5%	60.0%	70.3%	100%

Stimuli

Participants determined whether each of 26 test faces matched a target reference in a game similar to those reported in Laurence and Mondloch (2016) and Baker, Laurence and Mondloch (2017). This task was completed with target and test faces upright and inverted. Participants viewed a different image set for each task, each comprising 9 unique images of a target identity and 9 unique images of a similar-looking distractor identity (18 unique images, all female). Photographs were naturalistic: taken on different days and varying in hairstyle, make-up, expressions, and lighting (see Supplementary Materials Figure S1). Images were converted to greyscale, cropped to show just the head and printed on laminated cards measuring 38 x 50mm. Two types of ‘control’ test images were included in each image set for ‘catch trials’ to confirm that participants understood the instructions. Four identical copies of the target image (an easy match) and four identical images of a very different identity (contrasting from the target in age and/or hairstyle: an easy mismatch). In total, participants saw images from 6 different identities across the two tasks with: 2 target identities, 2 distractor identities and 2 mismatch control identities. Assignment of specific target, distractor and mismatch grouping to orientation condition was counterbalanced across participants.

Procedure

In an initial training phase, participants were given a target reference image of Buzz Lightyear and nine cards comprising five different pictures of Buzz (one identical to the target), and four different pictures of another identity (Noddy). They were to place the pictures of Buzz inside his rocket and keep everyone else out. In the rare case of any sorting mistakes, this training was repeated until it was accomplished without error.

For the main task, participants were told that a target identity (e.g., ‘Alice’) lives in a toy house, as indicated by her photograph (target reference) on the roof. They were given a set of cards showing the test faces (order pseudorandomised by experimenter shuffling) and told that even though Alice could look different from day to day, they should place all the pictures of Alice inside her house while leaving other people outside. After completing this task with upright faces, participants performed the same task with a new image set, all presented inverted. They were told that the task would be exactly the same, only now the pictures they had to sort would be upside down. No rotation of the stimuli was permitted.

Most participants[1] also completed the Cambridge Face Memory Test (CFMT; Duchaine and Nakayama 2006) or Cambridge Face Memory Test for Children (CFMT-C, Croydon et al, 2014). In both cases, participants learn to recognise a series of Caucasian adult male face identities and then discriminate these individuals from similar-looking foils under increasingly difficult conditions, e.g., with changes in viewpoint and the additional of visual noise. Together the two tasks took approximately 25-35 minutes to complete. Adults received participation credits for their undergraduate course or a small fee for their time. Children were tested in school and received a certificate and stickers.

Data supporting these findings is available at https://osf.io/8qbv7/?view_only=2b89be08cbd949dcb169b267eb48860d

Results

Telling Faces Together Task

Matching ability was estimated using signal detection theory. Analysis was restricted to trials featuring the target identity and similar-looking distractor. ‘Catch trials’ were inspected only to confirm that all participants sorted these images correctly when upright signalling understanding of the task requirements[2]. Correctly matching an image of the target face with the reference (i.e., “putting the correct person in the house”) constituted a hit, failing to match the target identity (“leaving the correct person outside the house”) was a

[1] Some adults did not have scores for the CFMT (n=23) because they completed the matching task as part of a different testing battery.

[2] A small number of catch trial errors were observed in the inverted task (13% of 6-7 year olds, 10% of 8-9 year olds, 9% of 10-11 year olds, 2% of adults). No exclusions were made based on these errors, which reflected the difficulty of inverted face matching, rather than misunderstanding task instructions.

miss, and incorrectly matching the distractor identity (“putting the incorrect person in the house”) was a false alarm. The dependent variable for unfamiliar face processing ability was d’ sensitivity for upright and inverted stimuli. Despite the difficulty of the task, particularly in the inverted condition, performance was generally above floor for all age groups (see Supplementary Materials 2, Table S1).

Given the age divide between children and adults we focused the main developmental analysis on specific groups: a two-way mixed ANOVA investigated the effects of orientation (upright, inverted) and participant age (6-7 years, 8-9 years, 10-11 years, Adults) on matching sensitivity. For comparability to past research we also present a complementary analysis examining the effects of age as a continuous variable (restricted to children).

A significant main effect of orientation ($F(1,165)=30.37$, $p<0.01$, $\eta_p^2=0.15$) indicated that overall participants performed significantly better on the upright ($M=0.69\pm0.75$) than in the inverted task ($M=0.29\pm0.72$). This divergence points against the utilization of a consistent/generic processing mechanism for the canonical and non-canonical viewpoints. Critically, this effect of inversion did not increase across childhood (see Figure 1, and Supplementary Materials 3, Figure S2 for complementary violin plots). We observed no significant interaction between face orientation and age group (including adults, $F(3,165)=1.69$, $p=0.17$, $\eta_p^2=0.03$). Power analysis (conducted retrospectively because control over sample size was limited by the availability of schools and children that agreed to take part) indicated our study was 82.3% powered to detect an interaction of this size: above the typically targeted level of 80% (G*Power; Faul, Erdfelder, Lang, & Buchner, 2007).

A significant main effect of age group ($F(3,165)=6.12$, $p<.01$, $\eta_p^2=0.10$), reflected superior matching performance overall in adults compared to all three child groups (all ts $\geq$ 3.18, $ps \leq 0.01$). Unexpectedly, no significant developmental differences were observed across the range of childhood ages tested here. The 6-7 year olds, 8-9 year olds, and 10-11 year olds did not differ from each other (all *t*s ≤ 0.25, $p \geq 0.80$). To confirm the lack of developmental improvement in children, we ran an additional Bayes analysis looking at performance (d') across the three child age groups. Bayes Factors (*BF*) signal the relative strength of evidence for two competing hypotheses (e.g., the null and an alternative). They constitute the likelihood ratio of one hypothesis being true over another being true, with a value of 1 indicating that both are equally likely. Our $BF_{10} = 0.076$ indicates that the null is 13.18 times as likely as the alternative, and can be considered *strong* evidence that the null hypothesis is true (Dienes, 2014). Further, we observed no direct association between age (in months) and d' accuracy for upright faces (N=129[3], $r = 0.005$, $p = 0.95$; $BF_{10} = 0.11$, substantial evidence for the null) or inverted faces (N=129, $r = 0.026$, $p = 0.77$, $BF_{10} = 0.12$, substantial evidence for the null) in the child participants.

Given the centrality of orientation effects for our research question, we investigated the magnitude of inversion effects across groups (upright minus inverted d'). There was no significant difference in the cost associated with inversion across the children's age groups, all *t*s<1, *p*s >.36, or association with age in months ($r = -0.01$, $p = 0.87$, $BF_{10} = 0.11$, substantial evidence for the null). Additional Bayes analysis further supports the lack of a main effect of age across childhood (*BF*=0.102, substantial evidence for the null). Comparisons with adults (M=0.69±1.02) indicated that mature face inversion effects were larger in adults relative to children: this effect was significant for the 10-11 year olds (M=0.24±0.94, $t(76) = 2.05$, $p < .05$, Cohens $d = 0.45$) and numerical for 6-7 year olds (M=0.30±1.01, $t(86) = 1.8$, $p = 0.07$, Cohens $d = 0.38$), and 8-9 year olds (M=0.43±0.90, $t(75) = 1.23$, $p = 0.22$, Cohens $d = 0.27$).

[3] Age in months was not available for 3 participants.

Matched analysis of a response bias index of performance (C) revealed no main effects of orientation ($F(1,165)=0.43$, $p = 0.52$, $\eta_p^2 = 0.01$) or age group ($F(3,165)=0.28$, $p = 0.84$, $\eta_p^2 = 0.01$) and no significant interaction between these variables ($F(3,165) = 1.50$, p = 0.21, $\eta_p^2 = 0.03$). These findings seem to rule out any criterion-related explanation of the observed accuracy results.

Error Analysis

Two additional ANOVAs investigated the effects of orientation and age group on Misses (analogous to identity overestimation errors in the 'piles' version of the task) and False Alarms (misidentification errors), each reported as proportions.

Misses. There were significantly more misses overall in the inverted ($M=0.69\pm0.24$) than upright condition ($M=0.63\pm0.23$; $F(1,165)=14.13$, $p<0.01$, $\eta_p^2=0.08$). There was also a trend for a main effect of age group ($F(3,165)=2.18$, $p=0.09$, $\eta_p^2=0.04$) which reflected significantly fewer misses by adults compared to all three child groups (all ts > 2.13, ps < 0.03; and no difference across child age groups, all ts < 0.28, ps > 0.77). Critically however, this was mediated by a significant age group x orientation interaction ($F(3,165)=3.69$, $p<.05$, $\eta_p^2=0.06$). Follow up tests indicated that the effect of orientation on misses was selective to adults ($t(36) = 4.84$, p<.01) and was not significant for any of the child groups (6-7yrs: $t(50) = 0.9$, $p=0.38$; 8-9yrs: $t(39) = 1.6$, $p=0.11$; 10-11yrs: $t(40) = 0.22$, $p=0.83$). Indeed, Figure 1 highlights that adults' error-advantage was selective to the upright condition: it was only there that they made significantly fewer misses than any of the child groups (all ts >2.74, ps < 0.01), which did not differ from each other (all ts < 0.83, ps > 0.40). By contrast in the inverted condition, the number of misses did not differ across all four age groups (all ts < 0.84, ps > 0.39).

False Alarms. A significant main effect of orientation (*F*(1,165)=10.42, $p<.01$, $\eta_p^2=0.06$) indicated that overall participants made fewer false alarms when completing the task upright (*M*=0.17±0.18) compared to inverted (*M*=0.23±0.21). Here, there was no significant main effect of age (*F*(3, 165) = 0.49, $p=0.69$, $\eta_p^2= 0.01$) or interaction involving age (*F*(3,165)=0.07, $p=0.98$, $\eta_p^2=0.01$).

INSERT FIGURE ONE ABOUT HERE

Figure 1. Performance on the matching task across age groups: Figures indicate d' matching sensitivity (left) as well as the rates of Misses (top right) and False Alarms (bottom right).

Recognition Memory

Percent correct on the CFMT-C/CFMT indexed participants' recognition memory ability (see Table 2). Differences between the child and adult forms of the test prevent statistical comparisons across the full participant sample. One-way ANOVA confirms a significant main effect of age group for the children (*F*(2,131)=8.10, $p<0.01$, $\eta_p^2=0.11$). Performance was significantly poorer in the 6-7 year olds relative to both older groups (ts<2.71, ps <.01), which did not differ, but may have been approaching ceiling levels of performance (*t*(79)=-1.11, $p=0.26$, $d=-0.24$).

Table 2. Recognition accuracy across age groups.

	6-7 years[a]	8-9 years[a]	10-11 years[a]	Adults[b]
Mean % correct	75.5	81.6	83.9	80.7
(SD)	(11.6)	(8.8)	(10.1)	(7.7)
Range	50.0 – 95.0	61.7 – 96.7	61.7 – 100	63.0 – 93.0

a. Cambridge Face Memory Test for Children. b. Cambridge Face Memory Test, N=23

Upright face matching (d') was positively (albeit non-significantly) correlated with face memory for adults (N=23; $r = 0.39$, $p = 0.07$). There was no such correlation for upright face memory and matching among children when controlling for age (N= 126; $r = 0.05$, $p = 0.55$) nor for the difference variable (upright minus inverted d') for either group: children ($r = 0.009$, $p = 0.92$), adults ($r = 0.08$, $p = 0.72$).

Discussion

This study investigated the processes used by children and adults when completing the 'Telling Faces Together' unfamiliar identity matching task. Inclusion of an inverted condition allowed us to test the orientation-selectivity of performance outcomes (matching sensitivity, errors), which is informative regarding the relative contributions of specialist vs. more generic mechanisms. Results confirmed a clear performance cost associated with stimulus inversion across all participant groups. The drop in matching sensitivity for inverted, compared to upright faces supports a divergence in processing that could be quantitative (participants use the same specialist mechanisms for both categories: but are better at applying them with upright faces) or qualitative (certain mechanisms are used selectively for upright faces). In either case, we add to the evidence base for specialist processing of faces not being limited to familiar identities (see Rossion, 2018).

We contrasted effects across developmental time to explore how experience modulates this perceptual profile. Inversion effects did not interact significantly with age. In broad accordance with face expertise being refined through experience, we observed a numerically larger inversion effect in adults compared with children. Yet there was no evidence to support a graded increase between 6 and 11 years. It is possible, of course, that uncontrolled differences in the demographics of the child and adult samples may contributed to apparent effect of age. Yet any such effects are likely to be subtle and we confirmed empirically, for example, that group differences in ethnicity could not explain the observed differences between children and adults.

These results add to a mixed body of findings regarding developmental changes in the effects of face inversion. The profile of differences observed between children and adults is consistent with other reports of developmental increases in the size (e.g., Brace et al., 2001) and face-selectivity (e.g., Picozzi, Macchi Cassia, Turati & Vescovo, 2009) of inversion effects. Yet the absence of an increase across childhood contrasts with clear and selective improvements in upright (cf. inverted) face recognition recently reported in the same age range (Hills & Lewis, 2018). This difference in the timing of the developmental shift may reflect the fact that those experimenters tested participants' face memory: enlisting mechanisms of expertise that develop later than those for perception (Weigelt et al., 2014). Indeed, age-related changes in children's recognition memory were observed in our sample.

Error analysis provides clues regarding what is driving developmental differences in matching sensitivity. The significant inversion effects observed from the youngest ages tested supported differences in the processing of unfamiliar upright and inverted faces from 6 years. Yet this orientation selectivity need not reflect the same processing profile across age groups. The effect is more pronounced in adults, and age-related differences in the pattern of misses (signalling overestimation of identities) and false alarms (signalling identity confusions) that underpin matching sensitivity suggest qualitative differences in how children and adults approached the task across orientations. When completing the task upright (cf. inverted), only adults showed an enhanced ability to avoid Misses – the type of error to which this task is particularly sensitive. No such selective developmental difference was observed in the rate of false alarms upright or inverted. These results extend previous adult research by demonstrating, for the first time, that this performance profile is selective to the canonical orientation. The same ‘boosted’ ability to avoid misses when upright is not observed in children and does not appear to be driving their orientation selectivity of matching sensitivity. Instead, children’s superior ability to match upright (cf. inverted) faces is principally accounted for by reduced false alarms.

Unlike unfamiliar face matching sensitivity, children’s upright face recognition memory improved significantly during the targeted ages. This finding may indicate that the specialised processes supporting unfamiliar face matching emerge early and are not *closely* tuned with ongoing experience (McKone et al., 2012). Alternatively, our age range might simply have been too short to capture the developmental changes as they occurred (studies supporting experience-related effects in adults investigate the impact of *decades* rather than years as in the current study, e.g., Laurence et al., 2016). Of note, two other studies have reported graded increases in face matching sensitivity in children in the current age range using a measure similar to the current task (with upright faces only: Baker, Laurence & Mondloch, 2017; Laurence & Mondloch, 2016). Differences in the variability of the ambient face images within each stimulus set might have contributed to this difference between the findings. Indeed, the relatively high miss rate observed - even in adults - suggests that our task may have been particularly difficult.

Unfamiliar face matching was not robustly related to recognition memory performance. In adults, we observed a marginally significant correlation with performance in the upright condition, which adds to a mixed set of extant results using more classic ‘piles’ style matching tasks with adults. Fysh et al. (2020) observed a positive relationship between matching performance (a composite metric capturing errors telling faces apart and together) and accuracy on the long form CFMT+ (see Russell, Duchaine & Nakayama, 2009) whereas Stacchi et al. (2020) using the same measures did not, and Balas and Saville (2017) identified links only with a matching performance metric selectively indexing errors telling faces together. Here, in children, we observed no association with broader face ability when controlling for age – though strong conclusions are undermined by high levels of performance on the CFMT-C in the older children.

Together, our findings provide new insights into the development of unfamiliar face processing. We present novel evidence that inversion significantly impairs unfamiliar face matching from 6 years of age. There is clear evidence of quantitative if not qualitative differences in how participants tell faces together when they are presented upright vs inverted. Numerically larger inversion effect in adults compared to children provide a degree of support for developmental changes in these specialised face abilities with increasing age/experience. Moreover, the distinctive patterns of errors across groups (i.e., selectively reduced misses in the upright condition only in adults) is consistent with a qualitative shift in unfamiliar face processing that occurs some time after eleven years of age.

Acknowledgements

We would like to thank Sarah Laurence, and Cathy Mondloch for sharing their stimuli with us. We further thank Sarah and Cathy for helpful discussions regarding the work. We also greatly appreciate the contribution of the schools and their pupils that kindly participated and the research assistants and volunteers that assisted with data collection: Susan Scrimgeour, Emily Fergusson, Michaela Viragova, Pari Patel, and Natasha Fletcher. This research was supported by funds from the Leverhulme Trust awarded to MLS and LE (RPG-2016-021).

References

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology, 68*(10), 2041-2050. https://doi.org/10.1080/17470218.2014.1003949

Baker, K. A., Laurence, S., & Mondloch, C. J. (2017). How does a newly encountered face become familiar? The effect of within-person variability on adults' and children's perception of identity. *Cognition, 161*, 19-30. https://doi.org/10.1016/j.cognition.2016.12.012

Balas, B., & Saville, A. (2017). Hometown size affects the processing of naturalistic face variability. *Vision Research, 141,* 228-236. https://doi.org/10.1016/j.visres.2016.12.005

Balas, B., Gable, J., & Pearson, H. (2019). The effects of blur and inversion on the recognition of ambient face images. *Perception, 48*(1), 58-71. https://doi.org/10.1177/0301006618812581

Brace, N. A., Hole, G. J., Kemp, R. I., Pike, G. E., Van Duuren, M., & Norgate, L. (2001). Developmental changes in the effect of inversion: Using a picture book to investigate face recognition. *Perception, 30*(1), 85-94. https://doi.org/10.1068/p3059

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73(1),* 105-116. https://doi.org/10.1111/j.2044-8295.1982.tb01795.x

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods, 42(1),* 286-291. https://doi.org/10.3758/BRM.42.1.286

Burton, A. M., Wilson, S., Cowan, M. and Bruce, V. (1999). Face recognition from poor quality video: evidence from security surveillance. *Psychological Science, 10,* 243–248. https://doi.org/10.1111/1467-9280.00144

Burton, N., Jeffery, L., Skinner, A. L., Benton, C. P., & Rhodes, G. (2013). Nine-year-old children use norm-based coding to visually represent facial expression. *Journal of Experimental Psychology: Human Perception and Performance, 39(5),* 1261. https://doi.org/10.1037/a0031117

Burton, A. M., & Jenkins, R. (2011). Unfamiliar face perception. *The Oxford Handbook of Face Perception*, *28,* 287-306. https://doi.org/10.1093/oxfordhb/9780199559053.013.0015

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B. C., & Pellicano, E. (2014). The Cambridge Face Memory Test for Children (CFMT-C): A new tool for measuring face recognition skills in childhood. *Neuropsychologia, 62*, 60-67. https://doi.org/10.1016/j.neuropsychologia.2014.07.008

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers In Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44(4),* 576-585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Ewing, L. (2022, February 3). Unfamiliar face matching task data. Retrieved from osf.io/8qbv7

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191. https://doi.org/10.3758/bf03193146

Hayden, A., Bhatt, R., Reed, A., Corbly, C., & Joseph, J. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology, 97,* 85 – 98. https://doi.org/10.1016/j.jecp.2007.01.004

Hills, P. J., & Lewis, M. B. (2018). The development of face expertise: Evidence for a qualitative change in processing. *Cognitive Development*, *48,* 1-18. https://doi.org/10.1016/j.cogdev.2018.05.003

Jeffery, L., McKone, E., Haynes, R., Firth, E., Pellicano, E., & Rhodes, G. (2010). Four-to-six-year-old children use norm-based coding in face-space. *Journal of Vision, 10(5),* 18-18. https://doi.org/10.1167/10.5.18

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121(3),* 313-323. https://doi.org/10.1016/j.cognition.2011.08.001

Kramer, R. S., Jenkins, R., Young, A. W., & Burton, A. M. (2017). Natural variability is essential to learning new faces. *Visual Cognition*, *25*(4-6), 470-476. https://doi.org/10.1080/13506285.2016.1242522

Laurence, S., & Mondloch, C. J. (2016). That's my teacher! Children's ability to recognize personally familiar and unfamiliar faces improves with age. *Journal of Experimental Child Psychology, 143,* 123-138. https://doi.org/10.1016/j.jecp.2015.09.030

Laurence, S., Zhou, X., & Mondloch, C. J. (2016). The flip side of the other-race coin: They all look different to me. *British Journal of Psychology, 107(2),* 374-388. https://doi.org/10.1111/bjop.12147

Mares, I., Ewing, L., Farran, E. K., Smith, F. W., & Smith, M. L. (2020). Developmental changes in the processing of faces as revealed by EEG decoding. *NeuroImage, 211,* 116660. https://doi.org/10.1016/j.neuroimage.2020.116660

Macchi Cassia, V., Turati, C., & Schwarzer, G. (2011). Sensitivity to spacing changes in faces and nonface objects in preschool-aged children and adults. *Journal of Experimental Child Psychology, 109*, 454 – 467. https://doi.org/10.1016/j.jecp.2011.03.003

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology, 29(1-2),* 174-212. https://doi.org/10.1080/02643294.2012.660138.

Neil, L., Cappagli, G., Karaminis, T., Jenkins, R., & Pellicano, E. (2016). Recognizing the same face in different contexts: Testing within-person face recognition in typical development and in autism. *Journal of Experimental Child Psychology*, *143,* 139-153. https://doi.org/10.1016/j.jecp.2015.09.029

Picozzi, M., Cassia, V. M., Turati, C., & Vescovo, E. (2009). The effect of inversion on 3-to 5-year-olds' recognition of face and nonface visual objects. *Journal of Experimental Child Psychology, 102*(4), 487-502. https://doi.org/10.1016/j.jecp.2008.11.001

Rhodes G, Brake S, Atkinson AP (1993) What's lost in inverted faces? *Cognition 47*, 25–57. https://doi.org/10.1016/0010-0277(93)90061-Y

Rossion, B. (2018). Humans are visual experts at unfamiliar face recognition. *Trends in Cognitive Sciences*, *22(6),* 471-472. https://doi.org/10.1016/j.tics.2018.03.002

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review, 16(2),* 252–257. https://doi.org/10.3758/PBR.16.2.252.

Stacchi, L., Huguenin-Elie, E., Caldara, R., & Ramon, M. (2020). Normative data for two challenging tests of face matching under ecological conditions. *Cognitive research: Principles and implications*, *5*(1), 1-17. https://doi.org/10.1186/s41235-019-0205-0

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month old infants, and adults: Evidence from the composite face effect. *Child Development, 81,* 1894 – 1905. https://doi.org/10.1111/j.1467-8624.2010.01520.x

Towler A, Kemp RI, Burton AM, Dunn JD, Wayne T, Moreton R, et al. (2019) Do professional facial image comparison training courses work? *PLoS ONE 14(2):* e0211037. https://doi.org/10.1371/journal.pone.0211037

Valentine, T. (1988) Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471-491, https://doi.org/10.1111/j.2044-8295.1988.tb02747.x

Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. *Developmental Science, 17*(1), 47-58. https://doi.org/10.1111/desc.12089

Young, A. W., & Burton, A. M. (2018). Are we face experts? *Trends in Cognitive Sciences, 22,* 100 - 110. https://doi.org/10.1016/j.tics.2017.11.007

Zhou, X., & Mondloch, C. J. (2016). Recognizing “Bella Swan” and “Hermione Granger”: No own-race advantage in recognizing photos of famous faces. *Perception, 45(12)*, 1426-1429. https://doi.org/10.1177/0301006616662046