Neuropsychologia 62 (2014) 60–67

Contents lists available at ScienceDirect

Neuropsychologia

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

The Cambridge Face Memory Test for Children (CFMT-C): A new tool for measuring face recognition skills in childhood

Abigail Croydon [a], Hannah Pimperton [b], Louise Ewing [c,d], Brad C. Duchaine [e], Elizabeth Pellicano [a,d,*]

[a] Centre for Research in Autism and Education (CRAE), Institute of Education, University of London, London, UK
[b] Institute of Cognitive Neuroscience, University College London, London, UK
[c] Department of Psychological Science, Birkbeck, London, UK
[d] School of Psychology, University of Western Australia, Perth, Australia
[e] Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

ARTICLE INFO

Article history:
Received 13 May 2014
Received in revised form
21 June 2014
Accepted 10 July 2014
Available online 20 July 2014

Keywords:
Face recognition
Face memory
Development
Children
Inversion effect

ABSTRACT

Face recognition ability follows a lengthy developmental course, not reaching maturity until well into adulthood. Valid and reliable assessments of face recognition memory ability are necessary to examine patterns of ability and disability in face processing, yet there is a dearth of such assessments for children. We modified a well-known test of face memory in adults, the Cambridge Face Memory Test (Duchaine & Nakayama, 2006, *Neuropsychologia*, 44, 576–585), to make it developmentally appropriate for children. To establish its utility, we administered either the upright or inverted versions of the computerised Cambridge Face Memory Test – Children (CFMT-C) to 401 children aged between 5 and 12 years. Our results show that the CFMT-C is sufficiently sensitive to demonstrate age-related gains in the recognition of unfamiliar upright and inverted faces, does not suffer from ceiling or floor effects, generates robust inversion effects, and is capable of detecting difficulties in face memory in children diagnosed with autism. Together, these findings indicate that the CFMT-C constitutes a new valid assessment tool for children's face recognition skills.

1. Introduction

Faces are critical to social interaction. They provide a wealth of information about an individual's gender, ethnicity, emotional state, direction of attention and, crucially, they uniquely identify the owner. The ability to identify persons from their facial appearance – face identity recognition – begins early in development (e.g., Bushnell, Sai, & Mullin, 1989; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995). Despite this early facility, the emergence of adult face expertise follows a protracted course of development, with performance on tests of face recognition not approaching maturity until well into adulthood (e.g., Germine, Duchaine, & Nakayama, 2011; Susilo, Germine, & Duchaine, 2013).

Much current research is focused on understanding the perceptual, cognitive, and neural mechanisms underlying this lengthy developmental trajectory (e.g., Crookes & McKone, 2009) and elucidating how such processes might develop differently in children with neurodevelopmental conditions, such as autism (e.g., Weigelt, Koldewyn, & Kanwisher 2012) and developmental prosopagnosia (Wilson, Palermo, Schmalzl, & Brock, 2010). Such research would be facilitated by standardized assessments of unfamiliar face identity recognition, providing tools useful for experimental and clinical settings and enabling direct comparison between individuals with and without neurodevelopmental conditions (see Dalrymple, Corrow, Yonas, and Duchaine (2012)).

Yet many of the existing standardized face recognition tests for adults (e.g. the Benton Facial Recognition Test: Benton, Sivan, Hamsher, Varney, & Spreen, 1983; the Recognition Memory Test for Faces: Warrington, 1984) and children (e.g. the Identity Matching Test: Bruce et al., 2000) suffer from significant shortcomings, where a score in the 'normal range' does not necessarily reflect typical face recognition skills (Duchaine & Nakayama, 2004; Duchaine & Weidenfeld, 2003). For example, in the Benton Facial Recognition Test both the target and the test faces are presented simultaneously, which means that participants can derive the correct responses by using a feature-matching strategy, while in Warrington's Recognition Memory Test for Faces, participants can use non-face information present in the stimuli to select the

* Corresponding author at: Centre for Research in Autism and Education, Department of Psychology and Human Development, Institute of Education, 25 Woburn Square, London WC1H 0AA, UK. Tel.: +44 207 331 5140.
E-mail address: l.pellicano@ioe.ac.uk (E. Pellicano).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.07.008
0028-3932/

correct response. Both these limitations are also present in the children's Identity Matching Test (Bruce et al., 2000).

The Cambridge Face Memory Test (CFMT) was developed both to capitalise on the strengths of the previous adult assessments and to overcome their limitations in order to provide researchers and clinicians with a standardised test of face recognition that would accurately and reliably measure face memory ability (Duchaine & Nakayama, 2006). In the CFMT, adults study, and are subsequently tested on, the facial identities of 6 men posing with neutral expressions in three distinct experimental stages. Stage 1 requires participants to identify the same learned face amongst two distractor images, when the test image is identical to the study face. Stage 2 calls for recognition of the same learned faces in novel viewpoints and/or lighting conditions. Stage 3 requires recognition of the learned faces from novel images degraded by the presence of visual noise in order to increase the difficulty and to force greater reliance on face-specific mechanisms (McKone, Martini, & Nakayama, 2001).

Duchaine and Nakayama (2006) established that the CFMT showed excellent psychometric properties: the test showed a good range of responses in typical adults ($n=50$), did not suffer from ceiling or floor effects, showed the classic face-inversion effect (a decrement in performance when the face is turned upside-down) and could reliably diagnose individuals with acquired prosopagnosia, who have profound face-specific memory deficits (see Wilmer et al. (2010), for further psychometric findings). Consequently, the test is now used widely to test face identity recognition (e.g., Banissy et al., 2011; Bowles et al., 2009; Di Simplicio, Massey-Chase, Cowen, & Harmer, 2009; Hedley, Brewer, & Young, 2011; Kirchner, Hatri, Heekeren, & Dziobek, 2011; O'Hearn, Schroer, Minshew, & Luna, 2010; Richler, Cheung, & Gauthier, 2011; Wilmer et al., 2010) – at least in adults. A study using the CFMT with typically developing children aged between 9 and 17 years showed that performance in the younger children (9- to 12-year-olds) was poor and that the test was not sufficiently sensitive to distinguish between children with autism and typically developing children at this age (O'Hearn et al., 2010).

In the current study, we adapted Duchaine and Nakayama's CFMT to create a developmentally sensitive test of face recognition for children aged between 5 and 12 years. Here, we report on the performance of a large sample of primary and secondary school children on the upright and inverted versions of this new test, the CFMT for children (CFMT-C). We also examined the validity of the CFMT-C by administering it to a group of children diagnosed with autism, who have marked difficulties in social interaction and for whom problems with face identity recognition have been consistently implicated (Weigelt et al., 2012).

2. Methods

2.1. Participants

Four hundred and one participants (202 females) aged between 5 and 12 years were recruited from primary and secondary schools in England, UK, to take part in this study. Two hundred and eighty-two children completed the upright version of the test, including 37 5-year-olds (13 females), 38 6-year-olds (20 females), 37 7-year-olds (21 females), 40 8-year-olds (20 females), 33 9-year-olds (17 females), 41 10-year-olds (24 females), 29 11-year-olds (13 females) and 27 12-year-olds (13 females). One hundred and nineteen children completed the inverted version, including 12 5-year-olds (3 females), 10 6-year-olds (4 females), 12 7-year-olds (6 females), 18 8-year-olds (9 females), 20 9-year-olds (12 females), 23 10-year-olds (12 females), 12 11-year-olds (7 females) and 12 12-year-olds (8 females).

2.2. Stimuli

The face stimuli were selected from those used in the adult version of the CFMT (Duchaine & Nakayama, 2006). The faces were greyscale photographs of men in early adulthood posing with neutral expressions. Each face was photographed in the same three poses and lighting conditions and cropped to remove the hairline and any facial blemishes (see Fig. 1 for example stimuli). Similar to the CFMT, the same faces were used in the upright and inverted versions of the CFMT-C, the difference being that in the inverted version all images were presented upside-down (rotated 180°).

Based on pilot testing with 16 children aged between 5 and 12 years, several modifications were made to the original CFMT to create the CFMT-C. First, the number of target faces that the children viewed was reduced from six to five. Asking children to remember six target faces resulted in a floor effect during the most difficult 'noise' section of the test, while four target faces induced a ceiling effect for older children. The face selected for elimination from the test was the one that produced the lowest percentage of correct responses out of the six target faces in the original test (Duchaine & Nakayama, 2006). Second, test items were altered from a three- to a two-alternative forced choice design, comprising one target face and one distractor face (see Fig. 1A). Based on error data from adult participants, the least distinct distractor (i.e., the one most similar to the target and therefore incorrectly selected most often by adult participants) was removed from each test item. Third, the exposure time of the target faces was increased from 3000 ms to 5000 ms in stage 1 of the test to give children more time to encode the faces. Finally, the wording of the on-screen instructions was made appropriate for children, including practice trials showing a popular cartoon face, and reinforcement slides (e.g., "Keep up the good work!") to provide encouragement and help maintain children's attention.

2.3. Procedure: upright version

As in the original CFMT, the CFMT-C was comprised of a short practice stage plus three progressively more difficult stages (same images, novel images and novel images with noise).

2.3.1. Practice

This stage was used to familiarise children with the task structure. In the study phase, children saw the face of a popular cartoon character presented three times sequentially for 5000 ms each shown from different viewpoints (left-facing, front view, right-facing). Children were instructed to look at the images very carefully because they would need to remember them later. In the test phase, children saw the face of the same cartoon character alongside another character. They were asked to select which of the two faces they had just seen and to make the corresponding keypress ('1' or '2'). There were 3 test trials, one for each of the three different viewpoints. If an error was made on one or more practice trials, the practice stage was repeated until the child achieved perfect performance.

2.3.2. Stage 1: same images

Children were told that they would now need to memorise the faces of five different (real) people. During the study phase (see Fig. 1A), three different images (left-facing, front view, right-facing) of each face were presented sequentially for 5000 ms – just like in the practice stage. Next, in the test phase (see Fig. 1B), each of the three images was presented alongside a distractor face and children were required to choose which face they had just seen by making the corresponding keypress ('1' or '2') themselves (older children) or by informing the experimenter of their answer, who pressed the corresponding key on their behalf (young children: 5- to 6-year-olds). No feedback was given. One point was given for each correct response (maximum$=$15).

2.3.3. Stage 2: novel images

In the study phase, children initially inspected a single screenshot showing front views of all the five target faces for 20 s (see Fig. 1C). They were instructed to look carefully at the faces and to try to memorise them. During the test phase, children completed 25 trials, each consisting of one of the target faces and a distractor face, which remained onscreen until a response (keypress: '1' or '2'). Children selected which of the two faces they thought was one of the five target faces they had been asked to memorise. Each target face was shown 5 times each in a fixed, randomized order. Children were given one point for each correct response (maximum$=$25). Each distractor face appeared several times throughout the test phase to avoid the possibility that participants might select the correct answer simply by choosing faces that looked familiar. All target faces presented during the test phase were novel images, that is, either the lighting or viewpoint was different from those used in stage 1 (same images) (see Fig. 1C; for further detail, see Duchaine & Nakayama, 2006).

2.3.4. Stage 3: novel images with noise

The final stage began with another study phase – a single presentation of all five target faces, which children were again instructed to review for 20 s (see Fig. 1D). The subsequent test phase consisted of 20 trials in which a novel image of each target face was presented together with a distractor face 4 times each. Critically, both target and distractor images were masked with a pre-specified degree of Gaussian noise (see Fig. 1D), making the judgment more difficult than in stage 2. Trials were presented in a fixed, random order and the faces remained onscreen until children indicated which of the two faces most closely resembled one of the five target faces. One point was given for each correct response (maximum$=$20).

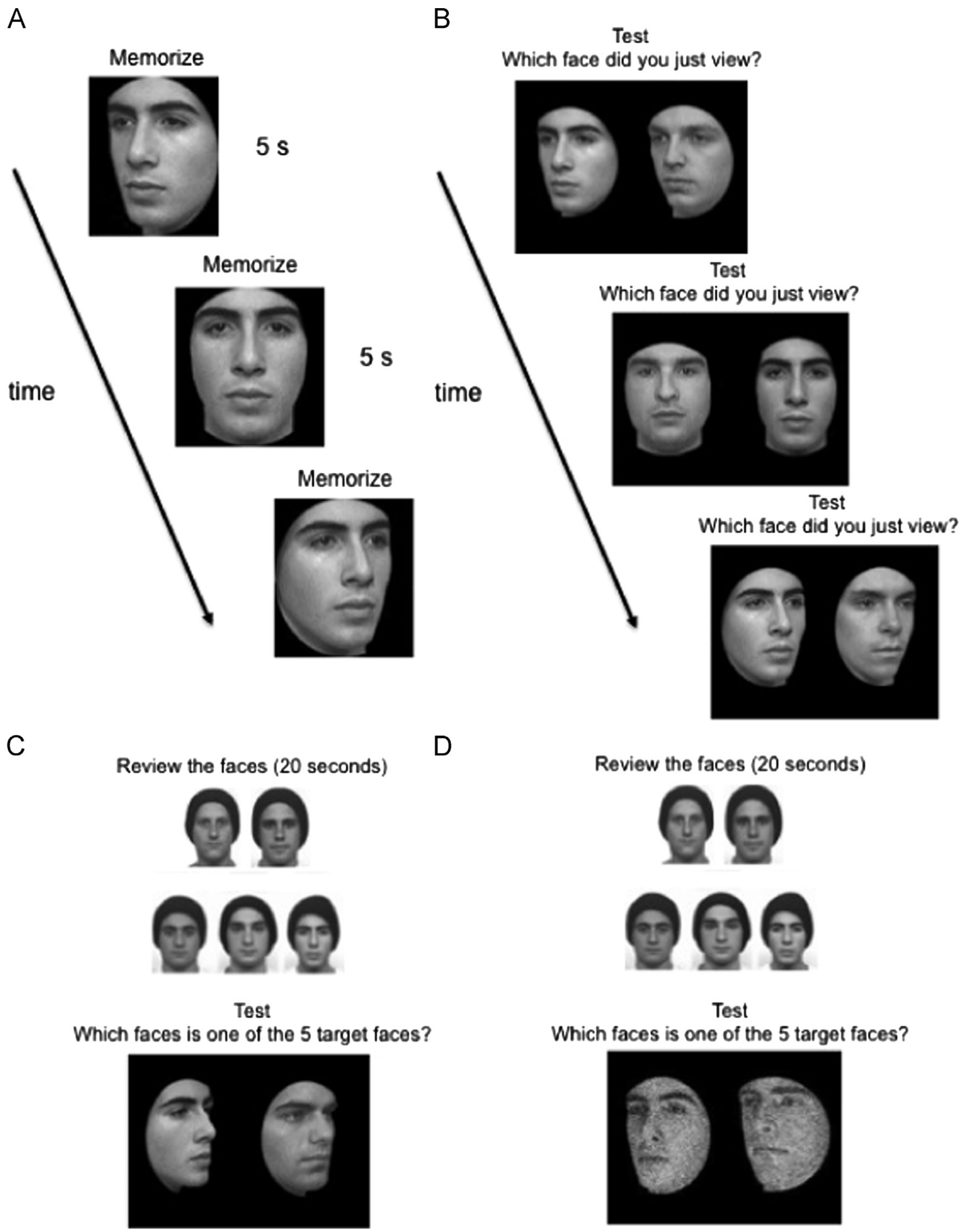

Fig. 1. Stimuli and procedure for the Cambridge Face Memory Test for Children (CFMT-C). A: Stage 1 (same images study phase). During the study phase, participants viewed 5 target faces, one at a time from three different viewpoints. B: Stage 1 (same images test phase). In the test phase, participants were then required to judge which of two faces was the one they had just seen. C: Stage 2 (novel images). During the study phase, participants were presented with all five target faces and asked to review them carefully. In the test phase, one of these faces was presented in a different viewpoint and children were asked to identify which of two faces (a novel image of the target face and a distractor face) was one of the 5 target faces. D: Stage 3 (novel images with noise). Similar to the study phase of stage 2, participants were asked to review the same 5 target faces. During the test phase, they were shown two faces comprising a novel image of target face and a distractor face, both masked with Gaussian noise.

Like in the adult CFMT, scores across all three stages were summed to yield a total recognition accuracy score for each child (out of 60).

2.4. Procedure: inverted version

The procedure for the inverted version of the CFMT-C was identical to the upright version. The exception to this was that children were told that there was "something a bit strange about the faces – they are going to be all upside-down!" They were encouraged to look at them carefully, just like in the upright version.

2.5. General procedure

Written informed consent was provided by the parents of all children prior to participation. All children were seen individually in a quiet space within their school. We used a 15-in. Macintosh Powerbook G4 running OSX, and children were seated at a distance of approximately 50 cm from the computer screen. The test took approximately 8–10 min to complete.

3. Results

To begin, we determine the presence of any floor or ceiling effects in the upright dataset. We then examine the relationship between age and total performance on the upright and inverted versions of the CFMT-C and also, using ANOVA, we examine age-related differences on the three stages of each test version (upright, inverted) and provide normative data for children of different ages. Finally, we establish the reliability and validity of the CFMT-C. Note that children's scores were converted to percentages to facilitate comparison across the three stages.

3.1. Floor/ceiling effects

One sample t tests (with Bonferroni correction for multiple comparisons; corrected $p<.004$) on children's upright performance showed that there were no floor and ceiling effects (see Table 1 for scores). Five- and 6-year-olds scored significantly above chance (50%) in all three stages (all $ps<.001$). Eleven- and 12-year-olds scored significantly below ceiling (100%) under all conditions (all $ps<.002$). The mean performance of the oldest children (12-year-olds; $n=27$) was 82.6% (SD=9.4%). This performance is comparable to Duchaine and Nakayama's (2006) adult participants ($M=80.42$%, SD=11.0), suggesting that test difficulty for the older children in the CFMT-C is closely matched to test difficulty for adults in the original test.

3.2. Age-related gains in overall CFMT-C performance

To examine the relationship between age and children's total CFMT-C performance, we conducted two separate regression analyses on children's total percentage correct for each version (upright, inverted) separately. Fig. 2 shows that there is much individual variation in performance on both tasks but also clear age-related improvements in face identity recognition. For the upright task, age accounted for 25% of the variance in children's performance, $F(1, 280)=94.9$, $p<.001$. Each additional birthday resulted in a 2.8% increase in children's total score. Similarly, for the inverted task, age accounted for a significant amount of variance ($R^2=.31$) in the model, $F(1,117)=52.10$, $p<.001$. Children improved in their performance on the inverted task by approximately 2.8% with each birthday.

3.3. Age-related differences on stages of the CMFT-C

Table 1 shows children's performance on each stage of the upright and inverted versions of the CFMT-C at each age. To examine age-related differences on the CFMT-C, we performed a repeated-measures ANOVA with age group (5 years, 6 years, 7 years, 8 years, 9 years, 10 years, 11 years, 12 years) and version (upright, inverted) as the between-participants factors and stage (same images, novel images, novel images with noise) as the within-participants factor. An additional ANOVA with gender as a factor (male, female) showed no main effect of gender or any interaction involving gender (all $ps<.08$).

There was a main effect of version, $F(1, 385)=136.06$, $p<.001$, $\eta_p^2=.26$. As expected, there was a significant inversion effect. Performance was significantly better in the upright ($M=76.3$%; SD=12.2) than the inverted ($M=65.6$%; SD=10.2) version of the CFMT-C. On average, the effect of inversion led to a 10.7% reduction in performance. There was no interaction between version and age group, $F<1$, indicating no developmental increase in the size of the inversion effect.

There was also a main effect of stage, $F(2, 770)=285.94$, $p<.001$, $\eta_p^2=.43$. These main effects were qualified by a stage × version interaction, $F(2, 770)=8.61$, $p<.001$, $\eta_p^2=.02$. In the upright version, performance in stage 1 (same images: $M=90.5$%; SD=12.3) was significantly better than in stage 2 (novel images: $M=76.2$%; SD=15.6), $t(281)=17.68$, $p<.001$, which in turn was significantly better than performance in stage 3 (novel images with noise: $M=67.2$%; SD=14.7), $t(281)=10.44$, $p<.001$. In the inverted version, performance in stage 1 (same images: $M=77.1$%; SD=16.6) was significantly better than in stage 2 (novel images: $M=63.2$%; SD=13.8), $t(118)=9.48$, $p<.001$, which in turn was significantly better than inverted performance in stage 3 (novel images with noise: $M=59.6$%; SD=11.9), $t(118)=2.39$, $p<.05$. The source of the interaction came from comparison of the *difference* between performance in the inverted stages. The difference in performance between stages 1 and 2 was significantly greater ($M=13.92$; SD=16.02) than the difference between stages 2 and 3 ($M=3.58$; SD=16.36), $t(118)=4.30$, $p<.001$, most likely because performance in stage 3 was approaching floor.

Table 1
Children's mean performance (% correct) on the different stages of the upright and inverted versions of the CFMT-C.

Version	Age group	CFMT-C stage			Total % correct
		Same images M (SD)	Novel images M (SD)	Novel images with noise M (SD)	M (SD)
Upright	5-year-olds ($n=37$)	80.0 (12.1)	62.3 (14.3)	60.5 (13.0)	66.1 (11.0)
	6-year-olds ($n=38$)	83.7 (13.3)	67.3 (15.6)	58.4 (13.0)	66.9 (11.3)
	7-year-olds ($n=37$)	88.4 (11.8)	73.5 (14.4)	67.3 (14.3)	75.2 (10.6)
	8-year-olds ($n=40$)	90.8 (10.5)	76.8 (13.1)	67.6 (12.3)	76.1 (9.8)
	9-year-olds ($n=33$)	94.5 (8.2)	82.2 (11.9)	71.3 (13.9)	81.6 (9.0)
	10-year-olds ($n=41$)	95.6 (7.1)	81.3 (15.6)	69.6 (15.3)	80.1 (12.0)
	11-year-olds ($n=29$)	97.0 (4.6)	85.1 (12.0)	70.5 (14.0)	83.2 (9.2)
	12-year-olds ($n=27$)	97.3 (7.4)	86.2 (9.0)	75.6 (14.4)	85.6 (8.2)
Inverted	5-year-olds ($n=12$)	66.1 (12.8)	47.3 (8.8)	52.9 (11.2)	53.4 (5.3)
	6-year-olds ($n=10$)	64.0 (12.3)	51.6 (13.1)	58.0 (8.9)	56.5 (7.2)
	7-year-olds ($n=12$)	73.9 (16.7)	59.3 (13.6)	54.2 (14.9)	64.0 (11.8)
	8-year-olds ($n=18$)	77.0 (16.4)	62.0 (9.4)	60.8 (14.3)	65.4 (7.3)
	9-year-olds ($n=20$)	81.7 (11.4)	67.6 (9.9)	58.8 (10.4)	67.9 (6.1)
	10-year-olds ($n=23$)	76.5 (18.9)	65.6 (13.5)	60.9 (9.6)	66.8 (9.9)
	11-year-olds ($n=12$)	86.7 (14.2)	74.7 (11.5)	64.2 (13.4)	74.2 (9.1)
	12-year-olds ($n=12$)	86.1 (17.2)	70.7 (12.8)	65.4 (9.4)	72.8 (10.1)

Note that chance performance on each stage of the CFMT-C is 50%.

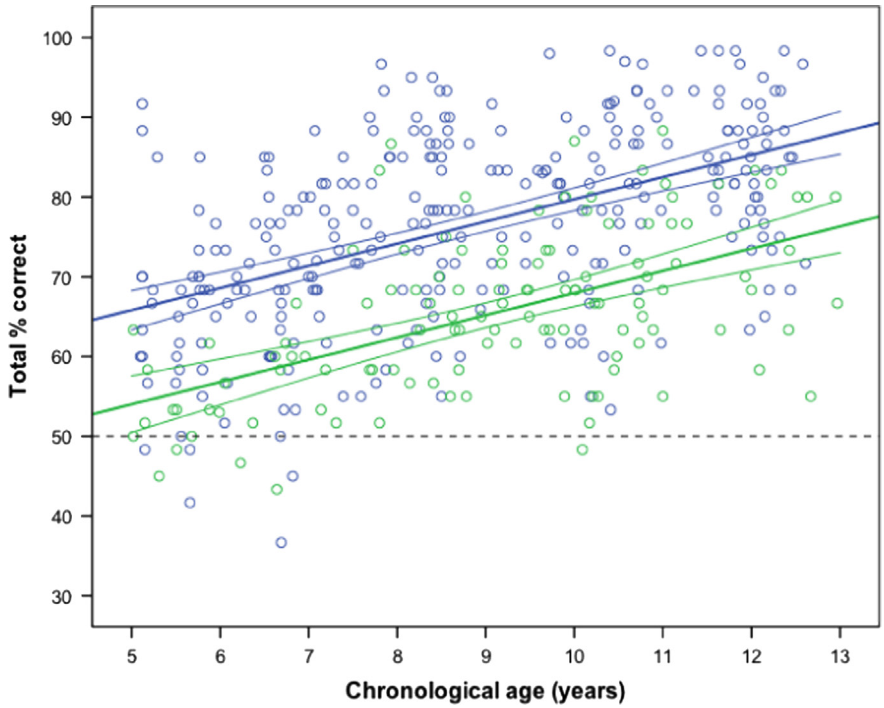

Fig. 2. Total percentage correct on the CFMT-C plotted against age for children who completed the upright (open blue circles) and inverted (open green circles) versions. The regression lines are shown for each relationship (upright: solid blue, $y=51.89+(2.78)$ age; inverted: solid green, $y=40.09+(2.79)$ age) and 95% confidence intervals. Dotted line represents chance performance (50%). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

There was a significant main effect of age group $F(7, 385)=18.47$, $p<.001$, $\eta_p^2=.25$. Post-hoc comparisons with Bonferroni correction confirmed a general pattern of age-related improvements in face identity recognition. The performance difference between adjacent ages, however, was not always significant (see Table 1 for scores). Five-year-olds performed significantly worse than all other age groups (all $ps<.001$) apart from 6-year-olds, 6-year-olds performed worse than all other groups (all $ps<.005$) apart from 5- and 7-year-olds, 7-year-olds performed significantly better than 5-year-olds ($p<.001$) and significantly worse than 11- and 12-year-olds ($ps<.001$), 8-year-olds performed better than 5- and 6-year-olds ($ps<.005$) but worse than 11- and 12-year-olds ($ps<.02$), both 9- and 10-year-olds performed better than 5- and 6-year-olds ($ps<.001$) and similar to all other age groups, and 11- and 12-year-olds obtained higher scores than all age groups ($ps<.002$) with the exception of 9-year-olds. There were no significant interactions involving age group ($ps>.09$). These results mirror the regression analyses showing a gradual age-related increase in face identity recognition.

3.4. Reliability

Following Duchaine and Nakayama (2006) and Bowles et al. (2009), we calculated the reliability of the upright version of the CFMT-C across all participants in two ways. First, we correlated children's performance on stage 2 of the test (novel images) with their performance on stage 3 (novel images with noise). Performance was significantly correlated across these conditions, $r(279)=.54$, $p<.001$. This moderately-sized correlation was not as strong as that reported by Duchaine and Nakayama (2006); $r=.74$; $n=50$ adults; $r=.75$ or Bowles et al. (2009), $n=124$ young adults and may be due to the possibility that children's performance is less stable with development than adults'.

Second, we examined the internal consistency of the upright version of the CFMT-C. The estimate of Cronbach's alpha was high ($\alpha=.88$) and comparable to that reported with the adult CFMT (Bowles et al., 2009; $\alpha=.89$). Like the CFMT, the CFMT-C therefore meets the standard reliability requirements for clinical tests (Cronbach's alpha $>.85$; Aiken, 2003).

3.5. Validity

To determine the validity of the CMFT-C, we administered the upright version on a population with known difficulties in face identity recognition (see Weigelt et al. (2012), for review). Forty-four children (9 girls) diagnosed with an autism spectrum condition aged between 7 and 12 years (M age=10.7 years; SD=1.6) took part in this validation study. All children had received independent clinical diagnoses of autism ($n=36$) or Asperger syndrome ($n=8$) and obtained a score of at least 15 or above (the cut-off for autism) on the Social Communication Questionnaire (SCQ; Rutter et al., 2003). Children with autism were compared to a subsample of typical children ($n=44$; 11 females) who had completed the upright version of the CFMT-C (M age=10.6 years; SD=1.7). All typical children fell well below the cutoff score of 15 on the SCQ, suggesting that they showed few behavioural features of autism (see Table 2).

The groups were of similar chronological age, $F(1,86)=.29$, $p=.87$, verbal IQ, $F(1,86)=.71$, $p=.40$, and performance IQ, $F(1,86)=.02$, $p=.89$, as measured by the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999) (see Table 2 for scores). All children completed the upright version of the CFMT-C in a single, individual session alongside the measure of intellectual functioning. Parents gave informed written consent for their child to take part.

To examine potential group differences in CFMT-C performance, an ANOVA with group (autism, typical) as the between-participants factor and stage (same images, novel images, novel images with noise) as the within-participants factor was performed on children's scores (percentage correct). There was a

Table 2
Descriptive statistics for chronological age, measures of intellectual functioning, autistic symptomatology and upright CFMT-C scores (% correct) in each group separately.

	Group	
	Children with autism (n=44)	Children without autism (n=44)
Age (in months)		
M (SD)	128.52 (19.42)	127.80 (20.80)
Range	90–155	91–155
Verbal IQ[a]		
M (SD)	99.27 (14.75)	101.45 (8.90)
Range	61–131	82–120
Performance IQ[a]		
M (SD)	98.36 (11.50)	98.77 (11.50)
Range	73–129	66–123
SCQ[b]		
M (SD)	25.23 (5.48)	3.73 (3.04)
Range	15–36	0–11
CFMT-C stage 1 (same images)		
M (SD)	85.15 (14.08)	94.54 (8.54)
Range	46.67–100	66.67–100
CFMT-C stage 2 (novel images)		
M (SD)	71.82 (16.25)	83.27 (12.02)
Range	32–100	52–100
CFMT-C stage 3 (novel images with noise)		
M (SD)	61.36 (15.19)	68.98 (14.08)
Range	35–90	35–95
CFMT-C total % correct		
M (SD)	71.67 (12.78)	81.32 (10.18)
Range	43.44–91.67	51.67–98.33

[a] Children's intellectual functioning was measured using the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999), standard scores reported here.
[b] SCQ: Social Communication Questionnaire (Rutter et al., 2003).

significant main effect of stage, $F(2, 172)=153.76$, $p<.001$, $\eta_p^2=.64$. Overall, participants performed better in stage 1 (same images: $M=89.8\%$; SD=12.5) than in stage 2 (novel images phase: $M=77.5\%$; SD=13.3), $t(87)=9.40$, $p<.001$, and stage 3 (novel images with noise: $M=65.2\%$; SD=15.1), $t(87)=8.46$, $p<.001$. There was also a main effect of group, $F(1,86)=15.63$, $p<.001$, $\eta_p^2=.15$. Children with autism ($M=71.7\%$; SD=12.8) obtained significantly lower scores than typical children ($M=81.3\%$; SD=10.2). There was no interaction between group and condition, $F<1$.

Autistic children's total performance on the CFMT-C was unrelated to their age, verbal IQ or performance IQ (all $ps>.11$). Yet there was a significant negative correlation between children's total % correct and their SCQ scores, $r(43)=-.40$, $p=.007$. Greater degrees of autistic symptomatology were related to worse face memory performance on the CFMT-C. Overall, these results suggest that the CFMT-C is sensitive for detecting atypicalities in face identity recognition in children with autism.

4. Discussion

Face identity recognition skills follow a lengthy trajectory and are at risk of developing atypically in individuals with neurodevelopmental conditions, such as autism and developmental prosopagnosia. There are, however, remarkably few tests that are appropriate for assessing these skills in children and those that do exist (e.g., Bruce et al., 2000) are limited in various ways. In this paper, we describe the development and application of a child-friendly version of the CFMT (Duchaine & Nakayama, 2006), a test that is increasingly being used by researchers as a valid and reliable assessment of face memory skills in adulthood. Here, we show in a large group of typically developing children aged between 5 and 12 years that the CFMT-C is sensitive to developmental differences in the recognition of unfamiliar upright and inverted faces, is sufficiently reliable to provide an accurate indication of a child's performance and is capable of detecting difficulties in face memory in children diagnosed with autism. Together, these findings suggest that the CFMT-C is a valid and reliable tool for assessing face recognition in middle childhood.

As expected, we observed gradual gains in face memory skills between 5 and 12 years of age. These findings are consistent with existing research showing that children's memory for faces follows a protracted developmental course both at the behavioural (e.g., Carey & Diamond, 1977; Johnston & Ellis, 1995) and neural levels (e.g., Golarai, Liberman, Yoon, & Grill-Spector 2009; Haist, Adamo, Han Wazny, Lee, & Stiles, 2013). We also found a significant effect of face inversion: upright faces were recognised more accurately than were inverted faces across all age groups.

The face inversion effect, that inversion disproportionately impairs the recognition of faces to a greater degree than the recognition of other classes of objects, is one of the most robust findings in the face processing literature and reflects the purportedly special status of faces (Diamond & Carey, 1986; Yin, 1969). In a seminal study, Carey and Diamond (1977) found that while 8- and 10-year-olds were much better at recognising the faces when they were presented upright than upside-down, 6-year-olds showed no such inversion effect. On this basis, Carey and Diamond (1977) proposed that young children process faces in terms of individual facial features until the age of 10 years, when there is a qualitative shift to a more adult-like processing style involving representations of the overall facial configuration. This proposal has since been vigorously debated in the literature (e.g., see Crookes and McKone (2009), Mondloch, Le Grand, and Maurer (2002), Pellicano and Rhodes (2003), and Pellicano, Rhodes, and Peters (2006)). Crookes and McKone (2009) have highlighted that when studies have reported differences in the magnitude of child and adult inversion effects (e.g., Brace et al., 2001; Carey & Diamond, 1977), these effects may be attributable to the presence of ceiling and floor effects. Indeed, when efforts are made to avoid such effects, children as young as 3 years show the classic inversion effect (Sangrigoli & de Schonen, 2004, Experiment 3; see also Pascalis, Demont, de Haan, and Campbell (2001)).

The absence of a significant interaction between version (upright, inverted) and age group in the current study provides further evidence of no age-related changes in the size of the inversion effect – at least between the ages of 5 and 12 years – and accords with other findings challenging qualitative differences in the way that younger and older children recognise faces (e.g., see Pellicano et al. (2006) and Crookes and McKone (2009)). Instead, the mechanisms responsible for face processing appear to be mature in early childhood, with developmental improvements in face memory skills potentially arising from more general gains in memory, attention and processing speed (Crookes & McKone, 2009; though see Mondloch, Le Grand, and Maurer (2010), for an alternative view).

We also assessed the validity of the CFMT-C by administering it to children with autism, a neurodevelopmental condition that affects the way an individual interacts with and experiences the world around them (American Psychiatric Association, 2013). Some key early indicators, including limited eye contact, poor social orienting, and reduced social responsiveness (Dawson, Webb, & McPartland, 2005; Zwaigenbaum et al., 2005), have led some researchers to suggest that face processing difficulties – and sociocognitive impairments more broadly – might be at the core of autism (Dawson et al., 2005; Schultz, 2005). A recent review of 90 experiments investigating face processing in autism suggested that, on average, autistic individuals perform significantly worse than typical individuals on tasks tapping face recognition (Weigelt et al., 2012). Our findings are consistent with Weigelt and colleagues' conclusions. We showed that 7- to 12-year-olds with autism performed significantly worse than typical children of similar age and ability – in fact, they performed on average one standard deviation below the mean of typical children. Furthermore, the absence of an interaction between CFMT-C stage and group supports the view that face identity recognition might be *qualitatively* similar in these children with and without autism (Weigelt et al., 2012).

Notably, our results are in contrast with the results of one study using the adult version of the CFMT. O'Hearn et al. (2010) reported face-processing difficulties in adolescents, but not in children, with autism and concluded that face memory difficulties emerge in autism only after adolescence. It is plausible, however, that these authors failed to identify such difficulties at earlier ages because the adult CFMT was not sufficiently sensitive to detect difficulties in children (with or without autism), many of whom seemed to perform at floor. The discrepancy between our findings using the modified CFMT-C and those of O'Hearn et al. serves to reinforce the importance of creating valid and developmentally appropriate measures of face identity recognition.

Importantly, the degree of face memory difficulties in children with autism is not so profound that many of them would be considered prosopagnosic. Rather, their face-memory difficulties are reasonably subtle. The CFMT-C has been used with a handful of young children with developmental prosopagnosia, who often show severe face recognition problems (Wilson et al., 2010). Future research should seek to validate further the CFMT-C with this group of children.

In sum, these results indicate that the Cambridge Face Memory Test – Children is a valid and reliable measure of unfamiliar face recognition ability that is sensitive to a wide range of abilities. The large number of typically developing children in this study warrants confidence in the results. We note here that adult face stimuli were used in the CFMT-C to allow comparison with the adult version of the CFMT (Duchaine & Nakayama, 2006). While evidence exists for an 'other age effect', where participants are better at recognising faces from their own rather than another age group, the impact of such an effect in children appears to be small (Hedges $g = .24$; Rhodes & Anastasi, 2012) and, in any case, such an effect is not an issue for this particular test because children's scores can be compared to the normative data presented here. Furthermore, the development of a version of the CFMT-C with children's faces (see Dalrymple, Gomez, & Duchaine, 2012) will allow us to test the degree of impact, if any, of the other age effect. One limitation of the CFMT-C, just like its adult counterpart, is that the same faces are used for the upright and inverted tasks, precluding the possibility of using the task to examine the degree of the inversion effect in individual participants. Nevertheless, the test, which is freely available for research purposes, will unquestionably prove useful for those wanting to examine patterns of ability and disability in face processing in children and to compare face identity recognition performance across laboratory sites and across populations.

Acknowledgements

This paper is in memory of Dr Andy Calder, a mentor and friend, whose dedication to the highest quality scientific research in face processing was unsurpassed. We are very grateful to the children, families and school staff who kindly took part in this research. Thanks also to Laura Dixon, Louise Edgington, Marlene Flögel, Emma Jaquet, Lydia King, Elena Klaric, Catherine Manning, Romina Palermo, Gill Rhodes, Erica Salomone, Martin Thirkettle and Ellie Wilson for help with data collection and to Marc Stears for comments on a previous version of this manuscript. Research at the Centre for Research in Autism and Education (CRAE) is also supported by The Clothworkers' Foundation and Pears Foundation.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Aiken, L. R. (2003). *Psychological testing and assessment* (11th ed.). Boston, MA: Allyn and Bacon.

Banissy, M. J., Garrido, L., Kusnir, F., Duchaine, B., Walsh, V., & Ward, J. (2011). Superior facial expression, but not identity recognition, in mirror-touch synesthesia. *The Journal of Neuroscience*, *31*, 1820–1824.

Benton, A., Sivan, A., Hamsher, K., Varney, N., & Spreen, O. (1983). *The Benton Facial Recognition Test*New York: Oxford University Press.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., et al. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant–stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, *26*, 423–455.

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

Bruce, V., Campbell, R., Doherty Sneddon, G., Import, A., Langton, S., McAuley, S., et al. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology*, *18*, 319–333.

Bushnell, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*, 3–15.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*, 312–314.

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.

Dalrymple, K. A., Corrow, S., Yonas, A., & Duchaine, B. (2012a). Developmental prosopagnosia in childhood. *Cognitive Neuropsychology*, *29*, 393–418.

Dalrymple, K. A., Gomez, J., & Duchaine, B. (2012b). CFMT-Kids: a new test of face memory for children. *Journal of Vision*, *9*, 492.

Dawson, G., Webb, S., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, *27*, 403–424.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Di Simplicio, M., Massey-Chase, R., Cowen, P., & Harmer, C. (2009). Oxytocin enhances processing of positive versus negative emotional information in healthy male volunteers. *Journal of Psychopharmacology*, *23*, 241–248.

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*, 576–585.

Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. *Neurology*, *62*, 1219–1220.

Duchaine, B. C., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, *41*, 713–720.

Germine, L., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition*, *118*, 201–210.

Golarai, G., Liberman, A., Yoon, J. M., & Grill-Spector, K. (2009). *Differential development of the ventral visual cortex extends through adolescence*. *Article 80*, 3 http://dx.doi.org/10.3389/neuro.09.080.2009.

Haist, F., Adamo, M., Han Wazny, J., Lee, K., & Stiles, J. (2013). The functional architecture for face-processing expertise: fMRI evidence of the developmental trajectory of the core and the extended face systems. *Neuropsychologia*, *51*, 2893–2908.

Hedley, D., Brewer, N., & Young, R. (2011). Face recognition performance of individuals with Asperger syndrome on the Cambridge Face Memory Test. *Autism Research*, *4*, 449–455.

Johnston, R. A., & Ellis, H. (1995). The development of face recognition. In: T. Valentine (Ed.), *Cognitive and computational aspects of face recognition: Explorations in face space* (pp. 1–23). London: Routledge.

Kirchner, J. C., Hatri, A., Heekeren, H. R., & Dziobek, I. (2011). Autistic symptomatology, face processing abilities, and eye fixation patterns. *Journal of Autism and Developmental Disorders*, *41*, 158–167.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 573–599.

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

Mondloch, C. J., Le Grand, R., & Maurer, D. (2010). Development of expertise in face recognition. In: I. Gauthier, M. J. Tarr, & D. Bub (Eds.), *Perceptual expertise: bridging brain and behavior* (pp. 67–106). Oxford, UK: Oxford University Press.

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*, 3955–3960.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant Behavior & Development*, *18*, 79–85.

Pascalis, O., Demont, E., de Haan, M., & Campbell, R. (2001). Recognition of faces of different species: A developmental study between 5 and 8 years of age. *Infant and Child Development*, *10*, 39–45.

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science*, *14*, 618–622.

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

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146–174.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, *22*, 464–471.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.

Sangrigoli, S., & de Schonen, S. (2004). Recognition of own-race and other-race faces by three-month-old infants. *Journal of Child Psychology and Psychiatry*, *45*, 219–1227.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, *23*, 125–141.

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.

Warrington, E. (1984). *Manual for the recognition memory test for words and faces*Windsor, UK: NFER-Nelson.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: a review of behavioural studies. *Neuroscience and Biobehavioral Reviews*, *36*, 1060–1084.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 5238–5241.

Wilson, C. E., Palermo, R., Schmalzl, L., & Brock, J. (2010). Specificity of impaired facial identity recognition in children with suspected developmental prosopagnosia. *Cognitive Neuropsychology*, *27*, 30–45.

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

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, *23*, 143–152.