SCIENCE CHINA
Life Sciences

THEMATIC ISSUE: Autism
• RESEARCH PAPER •

October 2015 Vol.58 No.10: 1024–1035
doi: 10.1007/s11427-012-4337-4

A new approach to the diagnosis of deficits in processing faces: Potential application in autism research

WANG RuoSi[1], LIU Ling[1] & LIU Jia[1,2*]

[1]*National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China;*
[2]*Graduate University of Chinese Academy of Sciences, Beijing 100049, China*

Deficits in social communication are one of the behavioral signatures of autism spectrum disorder (ASD). Because faces are arguably the most important social stimuli that we encounter in everyday life, investigating the ability of individuals with ASD to process faces is thought to be important for understanding the nature of ASD. However, although a considerable body of evidence suggests that ASD individuals show specific impairments in face processing, a significant number of studies argue otherwise. Through a literature review, we found that this controversy is largely attributable to the different face tests used across different studies. Therefore, a more reliable and valid face test is needed. To this end, we performed a meta-analysis on data gleaned from a variety of face tests conducted on individuals with developmental prosopagnosia (DP) who suffer a selective deficit in face processing. Based on this meta-analysis, we selected an old/new face recognition test that relies on face memory as a standard diagnostic test for measuring specific face processing deficits. This test not only reliably reflects DP individuals' subjective experiences with faces in their daily lives, but also effectively differentiates deficits in face processing from deficits caused by other general problems. In addition, DP individuals' performance in this test predicts their performance in a variety of face tests that examine specific components of face processing (e.g., holistic processing of faces). Finally, this test can be easily administrated and is not overly sensitive to prior knowledge. In summary, this test can be used to evaluate face-processing ability, and it helped to resolve the controversy whether individuals with ASD exhibit face-processing deficits.

autism spectrum disorder, developmental prosopagnosia, face recognition, face discrimination, old/new face recognition test

Citation: Wang RS, Liu L, Liu J. A new approach to the diagnosis of deficits in processing faces: Potential application in autism research. Sci China Life Sci, 2015, 58: 1024−1035, doi: 10.1007/s11427-012-4337-4

ASD is a spectrum of neurodevelopmental disorders characterized by impairments in social interactions, communication deficits, restricted interests, and repetitive behaviors (ICD-10, World Health Organization, 1992). Deficits in social reciprocity are a major behavioral signature of ASD, as individuals with this disorder are impaired both in processing social information [1] and in establishing social relationships [2]. Because faces provide rich social information, such as one's identity, mood (through expression), social interests (through direction of eye gaze), age, and gender, the ability to process faces is necessary for social interactions and communication. In fact, individuals with better social skills are also better at recognizing faces [3,4], whereas deficits in processing faces may cause fear and avoidance of social situations [5]. Therefore, investigating the ability to process faces in ASD individuals may provide a specific window to understanding the mechanisms underlying their impairments in social activities.

A classic model on face processing proposed by Bruce and Young [6] involves multiple cognitive components that are hierarchically organized. A revised model based on recent neuroimaging findings [7] suggests that faces are pro-

*Corresponding author (email: liujia@bnu.edu.cn)

cessed by a distributed neural network containing a core system for extracting facial identity information and extended systems for processing facial expressions, the direction of eye gazes, and speech-related mouth movements. In this model, the core system is the most important node in the network because it provides identity information and sends necessary information for later analyses conducted in the extended systems. In this study, we therefore focused on how facial identity information is extracted and represented.

There is a considerable body of evidence that ASD individuals have difficulties in processing facial identity. Some of them even have symptoms that are similar to individuals with developmental prosopagnosia (a.k.a., face blindness) (DP), which is a lifelong impairment in face processing without any brain damage or deficits in sensory or intellectual functions (for reviews see [8,9]). For example, an autistic woman called T.G. reported having experiences with faces that were very similar to those reported by individuals with DP. An example of one of her reports is "I often get into embarrassing situations because I do not remember faces unless I have seen the people many times or they have a very distinct facial feature such as a big beard, thick glasses, or a strange hairstyle" [10]. In fact, ASD individuals are sometimes diagnosed as having DP [11]. Consistent with these subjective experiences, careful examination has further revealed that ASD individuals show a variety of abnormal behaviors in processing faces (for reviews see [12,13]), such as peculiar scanning paths while viewing faces [14,15], and paying little attention to faces from a very young age [16,17]. Moreover, there is an increasing body of neurological evidence that suggests that the brains of ASD individuals are inclined to use the object-processing system, rather than the specialized face system, to process faces [18,19].

However, other researchers failed to observe face-specific deficits in ASD individuals. For example, Deruelle et al. [20] reported that the performance of autistic children in matching faces is not different from that of the control group (see also [21–24]). Those researchers who did observe deficits in face processing showed that the deficit is not face-specific but extends to processing of non-face objects [25,26]. In short, whether ASD individuals suffer face-specific deficits is still hotly debated.

To resolve the controversy, we proposed a new approach that is a reliable and valid test of face processing that can be used to quantify deficits in face processing quickly and efficiently in individuals with ASD. To this end, we first reviewed studies on face processing in the ASD literature, and then suggested that the differential findings of face-processing deficits across studies are attributable to differences in the type of face paradigms adopted across those studies. Therefore, a reliable and valid face test is needed. Second, we selected such a test through a meta-analysis on findings in individuals who exhibit lifelong deficits in face-specific processing. The criteria reliability, validity, predictability, and ease-of-use were used to evaluate a variety of face tests. Accordingly, this study consists of a review that describes the need to select a standard test of face processing in ASD and a proposed test that was designed by examining a variety of face tests through a meta-analysis.

1 Review of studies on face processing in ASD

To extract identity information from a face, one must perceptually examine the visual properties of the face (i.e., structural encoding), and then compare the extracted information to a stored template in memory. Therefore, face tests can in principle be classified into two paradigm categories based on the involvement of face memory. If a test is mainly based on information available in the structural encoding stage, it is classified as a discrimination test. A typical discrimination test requires participants to discriminate simultaneously presented faces based on differences in their stimulus properties. In contrast, if a test requires making comparisons between face images that are currently being observed and face images stored in memory, it is a recognition paradigm. A typical recognition test requires participants to judge whether faces have been seen before (i.e., familiar or famous). Outside of these two types of paradigms, a small number of studies have focused on specific components of face processing, such as holistic face processing, which includes the face inversion effect or the whole part effect [21,27–29]. We did not review these studies because the total number of the studies is too small for a meta-analysis. Instead, in the present report, we reviewed studies on face processing in the ASD literature based on the type of face tests used.

Studies using face recognition tests unvaryingly demonstrate that ASD individuals are selectively impaired in recognizing faces, with their ability to recognize non-face objects largely intact. De Gelder et al. [30] asked autistic children to remember a set of novel faces presented during a learning phase. In the testing phase, the autistic children were significantly worse in judging whether they had seen the faces before than matched normal controls. In contrast, their ability to recognize buildings [31], ordinary objects [32], words [33], animals (e.g., cats and horse) and plants (e.g., leaves) [34] was normal or even better than normal. Similarly, ASD participants had difficulties recognizing people whom they had encountered repeatedly in everyday life, such as their schoolteachers [35], family members and friends [36], and famous people [37]. In contrast, their ability to recognize familiar objects was largely preserved [36]. The only one that failed to find these face-specific deficits with face recognition tests [38] adopted a forced-choice response method, which is known to improve behavioral performance significantly [29]. Accordingly, studies using face recognition tests have overwhelmingly shown that

ASD individuals are specifically impaired in processing faces.

In contrast, the findings from studies using face discrimination tests are mixed. On the one hand, several studies report that autistic children were significantly worse at differentiating a singleton face from exemplar faces from another individual [27,39,40]. On the other hand, some studies failed to find impairments in ASD participants when two or more faces were presented simultaneously [20–22,41] or sequentially [23,24]. Still other studies argue that these impairments, if observed, are not face-specific, because the same degree of impairment was also found in processing non-face objects [25,26].

The percentages of studies finding normal or abnormal face processing in individuals with ASD separated by whether the studies used recognition tests or discrimination tests are shown in Figure 1. Seven out of nine studies using recognition tests clearly showed face-specific deficits [30–36], and one study found that 2/3 of ASD participants showed these deficits [37]. Only one study did not find these deficits [38]. However, this level of consistency was not observed among the twelve studies using discrimination tests. Only five studies reported face-specific deficits [25–27,39,40], whereas the remaining studies did not [20–24,32,41]. In addition, among the five studies that did find these deficits, two studies argued that the deficits are not face-specific [25,26]. Direct evidence that experimental paradigms are critical to the results comes from a study where the same group of ASD participants were tested in both recognition and discrimination tests [32]. The results showed the same pattern that we reported in the review as ASD participants exhibited deficits in face recognition tests but not in face discrimination tests. Therefore, it is likely that this controversy in the ASD literature is at least partly because of the type of face tests used.

In summary, although ASD individuals demonstrate deficits in processing faces in everyday life, experimental studies show mixed results. We suggest that these contradictory conclusions may have largely resulted from differences in the experimental paradigms used in these studies. Therefore, a more reliable and valid face test is needed to examine deficits in processing faces in ASD individuals. As would be expected, recognition tests are more reliable than the discrimination tests, as described above. However, further quantification of this intuitively appealing conclusion is necessary for two reasons. First, there is no consensus that ASD individuals suffer deficits in face processing, and we therefore do not know whether the recognition or discrimination paradigms more accurately characterize the true nature of ASD. In other words, an external criterion is needed. Therefore, a meta-analysis of results from a special subject population who are known to be selectively impaired in face processing (i.e., developmental prosopagnosia, DP) will provide a valuable external criterion. Second, even if the recognition tests generally perform better, we do not know which test is better in particular. Therefore, we need to examine the reliability, validity, and predictability of these tests quantitatively through a meta-analysis. In the next section, we report the findings of our meta-analysis of DP studies, which we performed to provide an independent confirmation of the findings in this review and quantitatively evaluate the tests used for face studies.

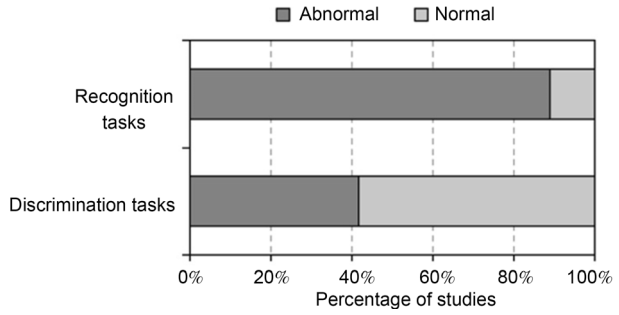

Figure 1 The percentage of studies that reported that ASD participants have deficits in face processing. For the studies using face recognition tests, eight out of nine studies (89%) described impaired face processing in ASD (i.e., abnormal, dark gray), whereas only five out of twelve studies using face discrimination tests (42%) found this deficit.

2 Quantification of face tests

Prosopagnosia refers to a selective deficit in face processing. It was first reported in patients with brain lesion [42] primarily in the occipitotemporal cortex [43,44]. Recently, prosopagnosia has also been found in subjects who do not have any known brain damage [45,46]. To contrast acquired prosopagnosia following brain lesions, it is called developmental prosopagnosia (DP). This relatively isolated deficit provides a rare opportunity to investigate mechanisms underlying face processing.

Previous studies have shown that individuals with DP are specifically impaired in face processing. First, DP participants are significantly worse in face processing, but their ability to process non-face objects are largely intact, suggesting that the deficit is face-specific [11,47,48] (but see [49]). Second, although DP individuals are unable to integrate face features into a whole face (i.e., holistic face processing) [50,51], they can process objects holistically, as they show normal sensitivities to global form and global motion [52], gestalt completion [53], and global-local interference [50] (but see [49]). In addition, they are unable to detect either spacing or partial changes in images of faces [50,52], but they can detect the same changes in images of houses [51]. Finally, although they are not able to recognize faces, they have no difficulties recognizing facial expressions or age or gender of the person in the face image [54–57], further suggesting that they are selectively impaired in extracting and representing facial identity infor-

mation. In short, the impairments observed in DP individuals are relatively pure, and therefore their performance in face tests can be viewed as a gold standard means of evaluating a variety of face tests.

Importantly, DP individuals share many abnormal and unusual characteristics in face processing with ASD individuals. First, irregular scanning paths while viewing faces are observed in both DP [58,59] and ASD participants [14,15]. Individuals in both groups predominantly scan regions containing non-face features, and fixate on external features (e.g., hair-style), whereas normal controls fixate on the center of facial features (e.g., eyes and mouth). Second, both groups process faces largely based on non-face information. For example, an autistic child successfully matched two faces because they had the same type of hair [25]. Similarly, DP participants often use non-face cues such as hair-style, clothing, and accessories rather than facial information to recognize people [60,61]. Third, both ASD [21,27,62] and DP participants [50,63,64] do not show the face inversion effect (FIE), which is a behavioral marker for face-specific processing [65,66], suggesting that both groups may process faces as separate parts rather than an integrated whole. Finally, DP individuals also suffer difficulties in social communication, and they often show fear and avoidance of social situations and even experience traumatic social interaction difficulties [5]. Given the similarities between DP and ASD individuals, it is likely that their deficits in face processing may be mediated by the same underlying mechanism. Therefore, a face test selected from DP studies may be also applicable to studying ASD individuals.

To quantify face tests, we used four criteria. First, we examined the reliability/generality of the face tests, which was indexed as the consistency between the behavioral measurement of a face test and a DP individuals' daily experience with faces. That is, if a subject had reported difficulties in processing faces in daily life, the subject should have failed this test. Second, we examined the validity/selectivity of the face tests, which was indexed as the capacity of a test to differentiate deficits in face processing from other general deficits (e.g., object agnosia). Third, we examined the predictability of the face tests, which was indexed as whether the results from a test predicted performance in tests focusing on specific components of face processing. Finally, we examined the ease-of-use of the face tests, which was indexed as whether performance in a test was influenced by the participants' prior knowledge, age, cultural background, education, and socio-economic status.

3 Materials and methods

3.1 Data for meta-analysis

Research articles used for the meta-analysis were identified from PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and ISI web of knowledge (http://www.isiknowledge.com) with the keywords "developmental prosopagnosia," "congenital prosopagnosia," and "hereditary prosopagnosia," published from 1990 to 2009. Studies that did not contain tests that measure the processing of facial identity[1)] or studies that tested brain-damaged participants were excluded.

Three criteria were used to collect the data in these studies. First, *Z* values or data suitable for conversion to *Z* values were available. Studies that provide nominal classification (e.g., normal or abnormal) of DP participants based on their behavioral performance were also included. Second, if a participant was tested more than once either within a study or in different studies, all results were included[2)]. On the other hand, if one measure was repeatedly reported in different studies, it was only included once[3)]. Third, results from face tests focusing on specific components of face processing (e.g., holistic processing) were not included in the meta-analysis. Instead, they were used to examine the predictability of the tests selected from the meta-analysis. Table 1 shows all 29 studies selected for the meta-analysis.

Based on the implication from our review of the ASD studies, we also classified the face tests used in studies on DP either as a face recognition test or as a face discrimination test. The former includes the famous or familiar faces test (FFT), the old/new recognition test (OldNew), and the Cambridge face memory test (CFMT) [83]. In the FFT test, participants are instructed to recognize familiar or famous faces (e.g., movie stars) by providing their identity information (e.g., names)[4)]. In the OldNew test, there is a learning phase where participants are instructed to remember a set of novel faces; then, in a testing phase where the learned faces are mixed with another set of novel faces and the participants are asked to judge which ones they have seen previously in the learning phase. The CFMT is a variant of the OldNew test, where participants are instructed to recognize

1) Several studies are excluded from the meta-analysis, such as those examining specific components of face processing (e.g., holistic processing [98]; visual mental imagery [97]), those focusing on processing social information (e.g., gaze perception [99]), and those investigating scan paths while viewing faces [58,59].

2) Studies that examined the ability of DP participants to recognize objects usually included more than one object category [11,48,56,77]. The performance of DP participants on each of the object recognition tests was included in the meta-analysis. In other studies [47,52,69,74], DP participants were tested more than once in the same test but with different procedures or face stimuli. The performance of DP participants in each test was also included.

3) Some studies contain results that were reported in other studies. For example, one DP participant (HV) in [76], three DP participants (EB, KL, and ML) in [80], two DP participants (BC and NM) in [52], and five DP participants (F2, F3, M1, M2, and M3) in [47] have already been reported in [75], [51], [53,69], and [51,69], respectively. These results were included as one measurement in the meta-analysis.

4) The results from the famous face test conducted in [54] are not included in the meta-analysis because the experimental procedure used is qualitatively different from the procedures used for the famous face test in the other studies, as the participants discriminated famous faces from novel faces without providing their identity information.

Table 1 The 29 studies on developmental prosopagnosia that were included in the meta-analysis. The data in the recognition tests, discrimination tests, and tests with overt cues columns are the passing rates in those tests in the format of xx/yy. That is, xx out of yy DP participants performed significantly worse in the test. For case studies, n (normal) and abn (abnormal) are used instead. Data within parentheses indicate that only nominal classifications were reported in those studies (i.e., no *Z* values). The check mark in the meta-analysis column indicates studies that were included in the corresponding analysis[a)]

Studies	Number of participants	Screening methods	Recognition tests				Discrimination tests			Tests with overt cues		Meta-analysis		
			Obj	FFT	CFMT	OldNew	Obj	CFPT	Fmatch	RMF	BFRT	R	V	P
Bentin et al. [67]	C.S.	self report	—	abn	—	—	—	—	—	(abn)	(n)	√	—	—
de Gelder & Rouw [68]	C.S.	self report	—	—	—	—	—	—	(abn)	(abn)	(abn)	√	—	—
Duchaine [53]	C.S.	self report	—	abn	—	abn	—	—	abn	(n)	(n)	√	—	—
Nunn et al. [63]	C.S.	self report	n	abn	—	abn	—	—	—	(n)	(n)	√	√	—
Duchaine et al. [69]	C.S.	self report	—	abn	—	2/2	—	—	n	(abn)	—	√	—	—
Duchaine et al. [11]	C.S.	self report	3/7	abn	—	abn	—	—	—	—	—	√	√	—
Duchaine [70]	C.S.	self report	(1/7)	abn	abn	(abn)	—	—	abn	—	—	√	—	—
Bentin et al. [71]	C.S.	self report	—	abn	abn	—	—	—	—	(abn)	(abn)	√	—	—
Li & Song [64]	C.S.	self report	—	—	—	—	—	—	(abn)	—	—	√	—	—
Steede et al. [72]	C.S.	self report	—	abn	abn	—	—	—	abn	(abn)	—	√	—	—
Bate et al. [73]	C.S.	self report	—	(abn)	abn	—	—	—	—	(abn)	(n)	√	—	—
Striemer et al. [48]	C.S.	self report	0/4	abn	—	abn	—	—	—	—	—	√	√	—
Grueter et al. [54]	8	self report	—	(7/7)	—	—	—	—	—	3/8	—	√	—	—
Duchaine et al. [56]	10	self report	6/14	8/10	10/10	4/7	—	8/9	—	—	—	√	√	CFPT
Minnebusch et al. [74]	4	self report	—	8/8	—	2/4	—	—	—	3/4	1/4	√	—	—
Righart & de Gelder [75]	4	self report	—	—	—	—	(0/4)	—	(1/4)	3/4	(4/4)	√	—	—
Van den Stock et al. [76]	3	self report	—	—	—	—	(0/2)	—	(1/2)	(2/2)	(1/2)	√	—	—
Lee et al. [77]	3	self report	3/18	2/3	3/3	2/3	—	1/3	—	—	—	√	√	—
Bowles et al. [78]	7	self report	—	6/7	6/7	—	—	1/7	—	—	—	√	—	CFPT
Duchaine & Nakayama [79]	11	testing	—	—	—	—	—	—	—	—	(4/11)	√	—	—
Harris et al. [80]	5	testing	—	2/2	—	2/2	—	—	—	—	—	√	—	—
Duchaine & Nakayama [47]	7	testing	(16/38)	(2/2)	—	(3/4)	—	—	—	—	—	√	—	—
Behrmann et al. [49]	5	testing	—	(9/10)	—	—	—	—	—	—	—	√	—	—
Yovel &Duchaine [51]	12	testing	—	12/12	12/12	12/12	—	—	—	—	—	√	—	SP(A)& SP(J)
Le Grand et al. [52]	5	testing	—	10/11	—	3/3	—	—	—	—	—	√	—	SP(J)
Duchaine et al. [50]	14	testing	—	13/14	14/14	—	—	8/14	—	—	—	√	—	CFPT& SP(A)
Behrmann et al. [81]	6	testing	—	(2/2)	—	—	—	—	—	—	—	√	—	—
Dobel et al. [55]	6	testing	—	6/6	—	—	0/5	—	0/5	—	2/5	√	—	—
Humphreys et al. [57]	3	testing	—	(2/3)	—	—	—	—	(3/6)	—	(1/3)	√	—	—
Bate et al. [82]	3	testing	—	3/3	3/3	—	—	1/3	—	—	—	√	—	—
Total measures				110	53	42		36	23	26	35			
# of DP showing deficits				102	52	35		19	10	17	15			

a) C.S., Case study; self report/testing, DP participants were selected either by their self-reports or by behavioral tests; Obj, object tests; FFT, famous or familiar face test; CFMT, Cambridge face memory test; OldNew, old/new recognition test; CFPT, Cambridge face perception test; FMatch, face matching test; SP(A), spacing-part test with Alfred face set; SP(J), spacing-part test with Jane face set. R, reliability; V, validity; P, predictability.

six learned faces at different levels of degradation and at different viewpoints.

The face discrimination tests include the face matching test (Fmatch), where participants are required to match facial identity based on stimulus properties, and the Cambridge face perception test (CFPT) [56], where participants sort a set of morphed faces based on similarity of face identity. Note that the CFPT, in our opinion, does not examine general face processing ability, but rather a specific component of face processing (i.e., holistic processing). However, we included it here because the authors claim that this test can be served as a diagnostic test for face-specific deficits. In addition, although the Warrington recognition memory test for faces (RMF) and the Benton face recognition test (BFRT) belong to recognition tests and discrimination tests respectively, we classified them into a new group because both contain prominent non-face features that confound the processing of face identity based on face features.

3.2 Statistic analysis

Because DP was previously thought to be a very rare syndrome, the number of DP participants tested in many studies is quite small and some of these studies were case studies.

Therefore, we could not calculate the effect size in these studies. Accordingly, non-parametric tests were performed in the meta-analysis. It is important to note that these non-parametric tests are more conservative and less affected by the sample size and extreme values than parametric tests.

Z values were normalized before the calculation of correlations between the tests, because of differences in the normal controls and experimental paradigms across the studies. To do this, the original *Z* values were transformed to normalized *Z* values by setting the mean of the *Z* values to zero for each study.

4 Results

The studies and data used in the meta-analysis are shown in Table 1 and Figure 2. Table 1 shows an overview of all 29 studies included in the meta-analysis, consisting of information on the number of DP participants, screening methods, passing rates in the tests, and the type of meta-analysis in which they were included. DP participants' *Z* values in each test are shown in Figure 2, with normal controls' performance as a baseline (i.e., zero). In the meta-analysis, we first examined the reliability of each test by comparing the performance of DP participants with their self-reports. Then, the validities of the tests that had high reliability were examined by contrasting the performance of DP participants in processing faces versus non-face objects. Finally, the tests with both high reliability and high validity were evaluated for how well they predicted the performance of DP participants in tests of specific components of face processing.

4.1 Reliability/generality

To investigate the reliability of each test, we measured the percentage of self-reported DP participants who also showed deficits in the test, where a deficit was defined as performance 1.65 standard deviations or more below the mean performance of normal controls ($P<0.05$, one-tail). The raw *Z* values for each test are shown in Figure 2, and the percentages of DP participants whose performance was below the normal range were averaged across studies using the same test. As shown in Figure 3A, DP participants performed worse in face recognition tests such as the FFT, CFMT, and OldNew, but not in face discrimination tests such as the Fmatch and CFPT.

This observation was further confirmed by a binomial test that examined whether the consistency between the DP participants' self-reports and their behavioral performance in a test was significantly higher than chance (i.e., 50%). We found that classifications based on the FFT, CFMT and OldNew results were significantly above the chance level (all $P<0.0001$, two-tails), whereas classifications based on the CFPT or Fmatch were not (CFPT, $P=0.87$; Fmatch, $P=0.68$). Besides, both the RMF and BFRT, which are contaminated by the inclusion of prominent non-face features

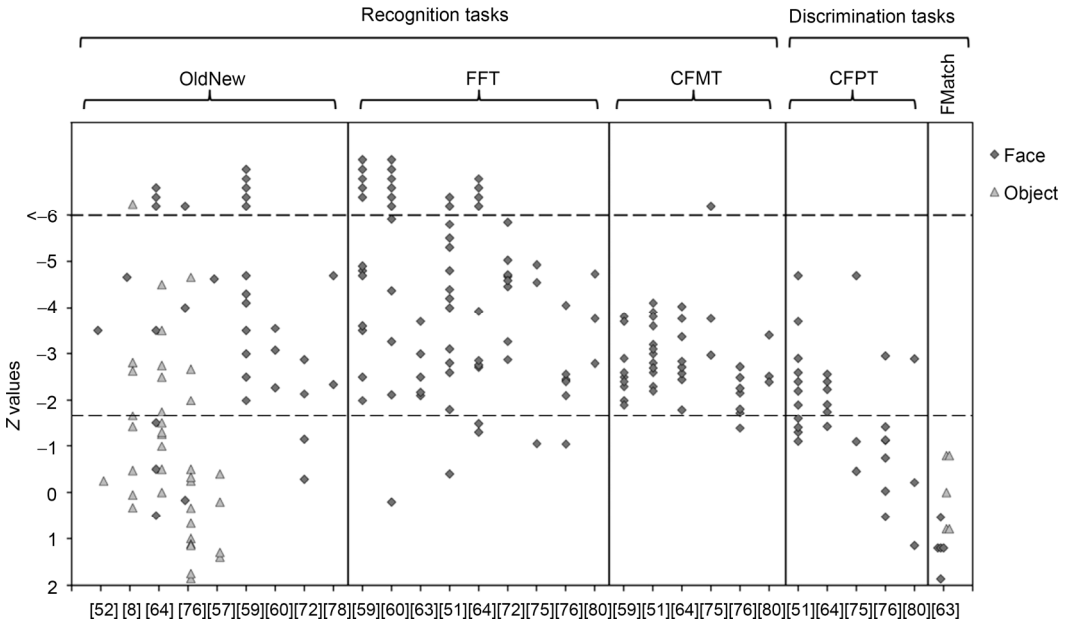

Figure 2 The performance of DP participants in a variety of face and object tests. Each point in a column represents one DP participant's performance in one test in a study (diamonds, face test; triangles, object test). Zero on the *y*-axis indicates the mean performance of normal controls, and two dashed lines indicate 1.65 and 6 standard deviations (STDs) from the mean of the normal control subjects' performance. Points above the bottom dashed line indicate that the DP participants represented by those points performed significantly worse than the control participants ($P<0.05$). The *x*-axis indicates the studies where the *Z* scores were derived from (in order by the type of tests and date of publication). Data from some case studies are not shown here because of the limited space. Note that the participants' performance in the OldNew and FFT is more widely distributed than that in the CFMT. OldNew, the old/new recognition test; FFT, the familiar or famous face test; CFMT, the Cambridge face memory test; CFPT, the Cambridge face perception test; Fmatch, the face matching test.

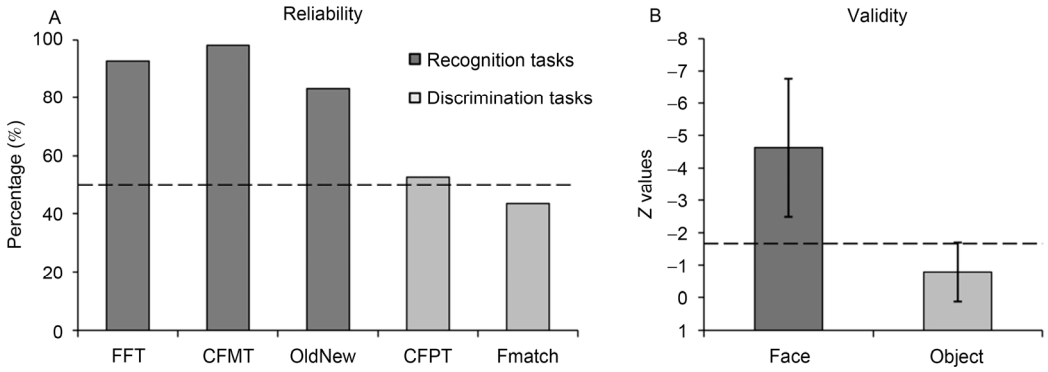

Figure 3 The reliability and validity of the face-processing tests. A, Reliability. The *y*-axis indicates the percentage of self-reported DP participants whose performance was significantly worse than the control participants on the tests listed on the *x*-axis. The dashed line represents the chance level (50%) of the nominal classification (normal versus abnormal). The face recognition tests (dark gray) show higher reliability (FFT, 93%; CFMT, 98%; OldNew, 83%), whereas the classification accuracy based on performance in the face discrimination tests (light gray) is around the chance level (CFPT, 53%; Fmatch, 43%). B, Validity. The performance of DP participants averaged across the face recognition tests (dark gray) versus the object recognition tests (light gray). Zero on the *y*-axis indicates the mean performance of normal control participants, and the dashed line indicates 1.65 STDs from the mean of the control participants' performance. The error bar indicates one STD of the distribution of the DP participants' performance.

[79,84], also failed to reach significance (RMF, P=0.17; BFRT, P=0.50) (not shown in the figure). This pattern suggests that the face recognition tests (i.e., the FFT, CFMT, and OldNew) have higher reliability than the face discrimination tests (i.e., the CFPT and Fmatch) or the tests containing non-face features (i.e., the RMF and BFRT).

To directly compare the reliability between the face recognition and the face discrimination tests, a Chi-square test was conducted with the nominal variables of experimental paradigm (recognition versus discrimination) and the performance of DP participants (normal versus abnormal). We found that the experimental paradigm was significantly related to the percentage of abnormal performance identified in the test ($\chi^2(1)=58.99$, $P<0.001$), and performance in the face recognition tests was more similar to the DP participants' self-reports than that in the face discrimination tests.

However, one may argue that because the DP participants were first screened by a face recognition test in some studies (i.e., participants that did not show deficits in the test were not included in the study, e.g., [51,79,80]), the data for the meta-analysis of the reliability were not completely independent. To rule out this alternative, we only included studies where DP participants were selected based on their self-reports. Using this procedure, the recognition tests again showed significantly higher reliability than the discrimination tests ($\chi^2(1)=15.26$, $P<0.001$).

4.2 Validity/selectivity

The meta-analysis of the reliability of the face tests shows that DP participants are more likely to show deficits in face recognition tests. Here, we further examined whether their performance was comparable to normal subjects in recognizing non-face objects with the same experimental paradigm. That is, do the recognition tests differentiate deficits in face processing from deficits in object processing? Five studies using the OldNew test for both face and object recognition were included in the analysis (Table 1).

We found that the DP participants' recognition of faces was significantly worse than their recognition of objects (Mann-Whitney U test, Z=–3.62, $P<0.001$), suggesting that the OldNew is capable of detecting face-specific deficits (Figure 3B). Similar analyses were unable to be conducted for the FFT and CFMT because they did not include non-face objects as the test stimuli. Nevertheless, the performance of DP participants in recognizing faces in both the FFT and CFMT tasks were positively correlated with their performance in the OldNew task (FFT and OldNew: spearman's rho=0.74, $P<0.0001$, n=19; CFMT and OldNew: rho=0.62, $P<0.01$, n=19), implying that they may have high validity as well.

4.3 Predictability

The face recognition tests we examined are designed to measure participants' ability to process whole faces, but they are not designed to measure processing of specific face components. On the other hand, a large number of experimental paradigms have been developed to investigate the specific components of face processing at different stages. Here, we examined whether the performance of DP participants in the face recognition tests can predict their performance in experiments examining specific face components. Five studies that used the CFPT and/or the spacing-part test [85,86] to measure holistic face processing were included (Table 1). Because of the limited number of subjects, we used a composite score, called the general face perception ability (GFPA), to index the performance of DP participants in tests of the holistic face processing to increase the statistical power of the analysis. The composite score was the average of the Z values from these component tests.

We found that the performance of DP participants in both the FFT (Figure 4A, left) and OldNew (Figure 4A, middle) tasks positively correlated with their performance in the tests focusing on holistic face processing (FFT and GFPA: rho=0.50, P<0.005, n=45; OldNew and GFPA: rho=0.73, P<0.005, n=14). This result implies that performance in both the FFT and OldNew tasks can predict performance at different stages of face processing. However, although the CFMT had both high reliability and validity, the correlational analysis suggests that it did not have high predictability (CFMT and GFPA: rho=0.11, P=0.52, n=37) (Figure 4A, right). This pattern is further confirmed by a more conservative 'leave-one-out' analysis, where correlation coefficients are calculated by leaving one study out for each repetition of the analysis to avoid the undue influence of a single study. Using this approach, a similar pattern was found as the results from the FFT were positively correlated with GFPA at every repetition (all rhos>0.40, all P<0.05), whereas the correlation between CFMT and GFPA never reached significance (all rhos<0.30, all P>0.1).

The lack of correlation between the CFMT and GFPA is not a result of insufficient statistical power, because the correlation coefficients between the FFT/OldNew and the GFPA were significantly higher than that between the CFMT and the GFPA (Mann-Whitney Test, FFT versus CFMT: Z=−2.44, P<0.05; OldNew versus CFMT: Z=2.37, P<0.05). Besides, there was no significant difference between the OldNew and the FFT in their relationships to the GFPA (Z=1.24, P=0.22). Moreover, since the meta-analysis of the predictability of the tasks is based on correlational analyses, one may argue that some general processing (e.g., attention, memory), rather than the face-specific processing, underlies the correlation between the FFT/OldNew and the GFPA. Therefore, the lack of correlation between the CMFT and the GFPA may be because the CMFT does not include this general processing. Though intuitive, this alternative is unlikely. The FFT was positively correlated with the CFPT only when faces were upright (rho=0.44, P<0.05, n=23) (Figure 4B, left), but not when faces were inverted (rho=0.03, P=0.89, n=23) (Figure 4B, right). In addition, no correlation was found between the CFMT and the CFPT, regardless of whether faces were upright (rho=0.31, P=0.15, n=23) or inverted (rho=0.13, P=0.55, n=23).

5 Discussion

In this study, we examined a variety of face tests through a meta-analysis to select a reliable and valid standard test for efficiently examining deficits in face processing in ASD. By performing the meta-analysis on the results from studies on DP, we examined the reliability, validity, and predictability of each face test. We found that tests that heavily rely on face memory (i.e., face recognition paradigm) not only faithfully reflected DP individuals' daily experience with faces, but also effectively determined whether their deficits are face-specific. In contrast, tests that mainly rely on the discrimination of difference among face images (i.e., face

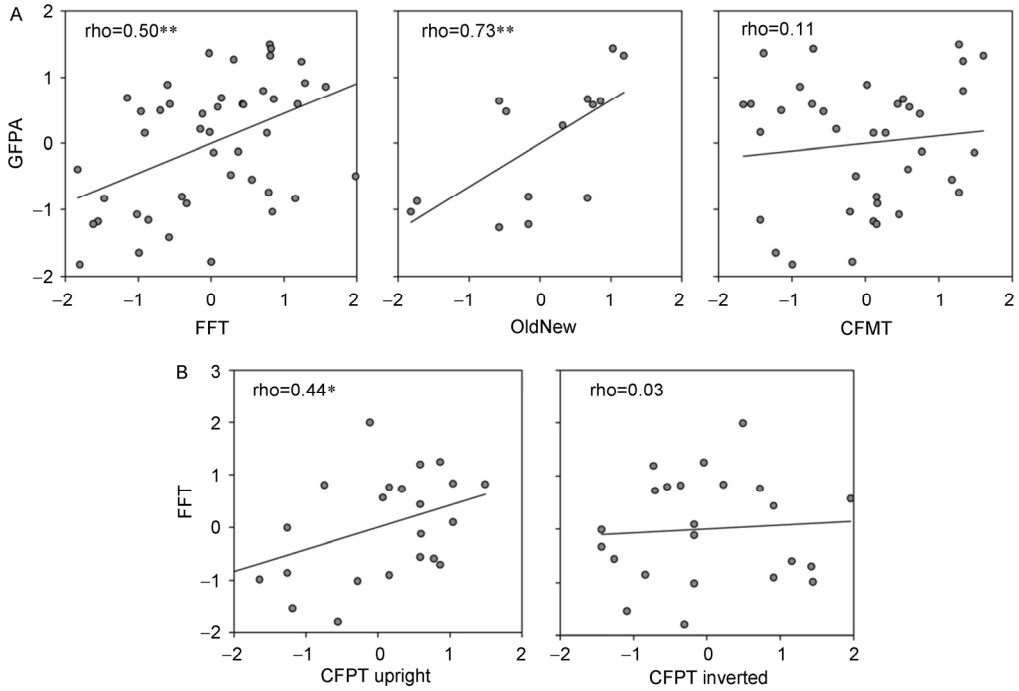

Figure 4 The predictability of the face-processing tests. A, Correlations between general face perception ability (GFPA) and performance in the FFT (left), OldNew (middle), and CFMT (right) tests. Each point represents one DP participant's performance in these tests. B, Correlation between the FFT and CFPT when the faces were presented upright (left) or inverted (right). Note that the Z values were normalized for the correlational analyses. **, P<0.01; *, P<0.05.

discrimination paradigm) did not exhibit either reliability or validity in detecting face-specific deficits. Finally, of all the face recognition tests we examined, we found that the Old-New test was both a reliable and valid measure of face processing ability and is capable of predicting the results of tests of specific face components. Therefore, we propose that the OldNew test should be used as a standard diagnostic of deficits in face processing in individuals with ASD.

Although our study is the first one that evaluates the pros and cons of a variety of face tests through a meta-analysis, researchers intuitively use face recognition tests more frequently than face discrimination tests in their studies. For example, most published articles on DP have adopted face recognition tests to examine deficits, possibly because the use of face discrimination tests is more likely to provide negative findings. Besides its application in detecting face-specific deficits in ASD and DP, face recognition tests are also suitable for exploring individual differences in the normal population. For example, by using the OldNew test, a recent study has found that individual differences in face memory are correlated with the extent to which subjects process faces holistically [87]. Another study on individuals who self-reported that they could remember faces seen only once and many years ago (a.k.a. super face recognizers) found that they outperformed normal controls in a variety of face recognition tests, including the FFT and CFMT [88]. In addition, face recognition tests have been widely used in studies that investigate the neural networks involved in face processing [81,89], the development of face-selective cortical regions [90], and genetic influences on face processing [91–93]. Therefore, because of their high reliability and validity, face recognition tests have become a gold standard tool for measuring face processing.

Several features in face recognition tests may account for their high reliability and validity. First, face recognition tests closely simulate our daily experiences with faces, so we are more likely to use the same strategy in carrying out face recognition tests. Therefore, a subject who has difficulties recognizing faces in social situations will likely fail face recognition tests conducted in laboratory (i.e., reliability). Second, because the fidelity of images stored in memory decays rapidly over time (e.g., [94]), we sometimes mistake friends as strangers, and vice versa. Previous studies have shown that the configural information of faces (both first-order and second-order face configurations) greatly facilitates the recognition of face parts (e.g., eyes, nose, and mouth) [95,96], so a good mnemonic trace likely relies on the holistic representation of faces [87], which differentiates face processing from general object processing. Therefore, face recognition tests may bias participants to process faces in a holistic fashion, which is distinct from the parts-based analysis used in object recognition (i.e., validity).

The principles that make face recognition tests a success are the same ones that make face discrimination tests a failure. In our daily lives, we seldom judge whether two faces are the same based on stimulus properties. Instead, the discrimination of differences between images usually exists in situations where non-face objects are involved (e.g., "which one is authentic: NOKIA or NOKLA?"). Therefore, face discrimination tests may bias participants to process faces in a parts-based fashion, like non-face objects. Furthermore, DP and ASD individuals are known to be less impaired in parts-based analyses and relatively sensitive to non-face cues to compensate for their inability to process faces [25,60,61,97]. Thus, the simultaneous or sequential (with a short delay) presentation of face images in discrimination tests obviously provides rich information on low-level image properties of non-face features.

However, not all tests using a face recognition paradigm are suitable to serve as a standard test. Through the meta-analysis, we found that the FFT, OldNew, and CFMT tests have both high reliability and high validity. However, results from the CFMT are less likely to predict performance in experiments that measure specific components of face processing. That is, although the CFMT can detect face-specific deficits in DP, it may fail to quantify the magnitudes of these deficits. One possibility is that the CFMT is too complicated to reflect our daily experience with faces. On the other hand, the FFT has high predictability, but has poor "ease-of-use." The FFT relies on the recognition of famous or familiar faces, which makes it less applicable in practice. First, it is hard to control for individual differences in exposure to the famous or familiar faces, especially for subjects with ASD. Second, it is difficult to find famous people who are known to individuals in different places and who have different ages. Therefore, it is less feasible to compare results across different studies.

The OldNew test, on the other hand, satisfies all four criteria (reliability, validity, predictability, and ease-of-use). Therefore, we propose that the OldNew test should be used as a quick and efficient means of examining face-specific processing. Several issues must be taken into account when applying the OldNew test. First, general cognitive abilities (e.g., general intelligence, attention, working memory) also contribute to face processing [87,92]. Therefore, stimuli such as scenes or flowers should be included in the test as a baseline control. Second, explicit non-face cues such as clothes, hairlines, beards, and moles should be removed from the face stimuli to prevent participants from adopting strategies other than face processing. Third, although a paper-and-pencil-based test is sufficient, the data may be more informative if both accuracy and response time are available. Fourth, the OldNew test is a fast and efficient measure of face processing. To comprehensively characterize the nature of deficits in face processing, a full battery of face tests, including face discrimination tests, for different aspects of face processing should be conducted. Finally, the OldNew test is not limited to studies of deficits in face processing. It can be also used in situations where the evaluation of facial recognition is needed, such as a job pre-requisite exam for

positions that require a strong ability to recognize faces (e.g., security surveillance).

Our meta-analysis of these studies on DPs is consistent with the conclusions of our initial review of studies on face processing in ASD. Through a literature review, we found that most studies using face recognition tests have reported significant differences in face processing between ASD participants and matched controls. Moreover, these deficits were restricted to the processing of faces. In contrast, when face discrimination tests were used, the results were mixed. This similarity in both ASD and DP studies suggests that the test chosen based on our meta-analysis of DP studies is likely to be reliable and valid when it is applied to ASD studies. That is, the OldNew test is likely to serve as an effective diagnostic test for detecting face-specific deficits in ASD. Future studies are needed to examine the reliability and validity of the OldNew test by testing ASD individuals.

Previous studies have shown that there is a close link between face processing and social communication. Thus, a reliable and valid face test of face processing may serve to identify causes that lead to deficits in social communication in ASD. In particular, it may help to resolve the current controversy regarding whether ASD individuals suffer deficits in face-specific processing. Besides, the old/new recognition test provides valuable information for diagnosing ASD, which can be used to supplement current diagnostic methods. Finally, the old/new face recognition test can be used to establish a normative distribution of face processing ability in ASD. Then, ASD individuals may be categorized into more homogeneous subgroups, which will allow for the tailoring of different treatments to the unique characteristics of each subgroup.

We thank Dr. Zhu Qi for constructive suggestions. This work was supported by the 100 Talents Program of Chinese Academy of Sciences, the National Basic Research Program of China (Grant Nos. 2010CB833903 and 2011CB505402), the National Natural Science Foundation of China (Grant No. 91132703), and the Fundamental Research Funds for the Central Universities (Grant No. 2009SD-3).

1 Baron-Cohen S, Cox A, Baird G, Swettenham J, Nightingale N, Morgan K, Drew A, Charman T. Psychological markers in the detection of autism in infancy in a large population. Br J Psychiatry, 1996, 168: 158–163
2 Volkmar F R, Klin A, Siegel B, Szatmari P, Lord C, Campbell M, Freeman BJ, Chicchetti DV, Rutter M, Kline W, Buitelaar J, Hattab Y, Fombonne E, Fuentes J, Werry J, Stone W, Kerbeshian J, Hoshino Y, Bregman J, Loveland K, Szymansky L, Towbin K. Field trial for autistic disorder in DSM-IV. Am J Psychiatry, 1994, 151: 1361–1367
3 Li JG, Tian MQ, Fang HZ, Xu M, Li H, Liu J. Extraversion predicts individual differences in face recognition. Commun Integr Biol, 2010, 3: 1–4
4 Cheung CHM, Rutherford H JV, Mayes LC, McPartland JC. Neural responses to faces reflect social personality traits. Soc Neurosci, 2010, 5: 351–359
5 Yardley L, McDermott L, Duchaine BC, Duchaine B, Nakayama K. Psychosocial consequences of developmental prosopagnosia: a problem of recognition. J Psychosom Res, 2008, 65: 445–451
6 Bruce V, Young A. Understanding face recognition. Br J Psychol, 1986, 77: 305–327
7 Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends Cogn Sci, 2000, 4: 223–233
8 Kress T, Daum I. Developmental prosopagnosia: a review. Behav Neurol, 2003, 14: 109–121
9 Duchaine BC, Nakayama K. Developmental prosopagnosia: a window to content-specific face processing. Curr Opin Neurobiol, 2006, 16: 166–173
10 Grandin T. Thinking in pictures: and other reports from my life with autism. New York, NY: Vintage, 1996
11 Duchaine BC, Nieminen-von W, New J, Kulomäki T. Dissociations of visual recognition in a developmental agnosic: evidence for separate developmental processes. Neurocase, 2003, 9: 380–389
12 Sasson NJ. The development of face processing in autism. J Autism Dev Disord, 2006, 36: 381–394
13 Dawson G, Webb SJ, McPartland J. Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. Dev Neuropsychol, 2005, 27: 403–424
14 Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. Arch Gen Psychiat, 2002, 59: 809–816
15 Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, Piven J. Visual scanning of faces in autism. J Autism Dev Disord, 2002, 32: 249–261
16 Dawson G, Webb S, Schellenberg GD, Dager S, Friedman S, Aylward E, Richards T. Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. Dev Psychopathol, 2002, 14: 581–611
17 Grelotti DJ, Gauthier I, Schultz RT. Social interest and the development of cortical face specialization: What autism teaches us about face processing. Dev Psychobiol, 2002, 40: 213–225
18 Pierce K, Muller RA, Ambrose J, Allen G, Courchesne E. Face processing occurs outside the fusiform 'face area' in autism: evidence from functional MRI. Brain, 2001, 124: 2059–2073
19 Schultz RT, Gauthier I, Klin A, Fulbright RK, Anderson AW, Volkmar F, Skudlarski P, Lacadie C, Cohen DJ, Gore JC. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Arch Gen Psychiat, 2000, 57: 331– 340
20 Deruelle C, Rondan C, Gepner B, Tardif C. Spatial frequency and face processing in children with autism and Asperger syndrome. J Autism Dev Disord, 2004, 34: 199–210
21 Hobson RP, Ouston J, Lee A. What's in a face? The case of autism. Br J Psychol, 1988, 79: 441–453
22 Braverman M, Fein D, Lucci D, Waterhouse L. Affect comprehension in children with pervasive developmental disorders. J Autism Dev Disord, 1989, 19: 301–316
23 Celani G, Battacchi MW, Arcidiacono L. The understanding of the emotional meaning of facial expressions in people with autism. J Autism Dev Disord, 1999, 29: 57–66
24 Chawarska K, Volkmar F. Impairments in monkey and human face recognition in 2-year-old toddlers with autism spectrum disorder and developmental delay. Dev Sci, 2007, 10: 266–279
25 Davies S, Bishop D, Manstead AS, Tantam D. Face perception in children with autism and Asperger's syndrome. J Child Psychol Psychiat, 1994, 35: 1033–1057
26 Behrmann M, Avidan G, Leonard GL, Kimchi R, Luna B, Humphreys K, Minshew N. Configural processing in autism and its relationship to face processing. Neuropsychologia, 2006, 44: 110–129
27 Tantam D, Monaghan L, Nicholson H, Stirling J. Autistic children's ability to interpret faces: a research note. J Child Psychol Psychiat, 1989, 30: 623–630
28 Teunisse JP, de Gelder B. Face processing in adolescents with autistic disorder: The inversion and composite effects. Brain Cogn, 2003, 52: 285–294
29 Joseph RM, Tanaka J. Holistic and part-based face recognition in children with autism. J Child Psychol Psychiat Allied Discip, 2003, 44: 529–542

30 de Gelder B, Vroomen J, van der Heide L. Face recognition and lip-reading in autism. Eur J Cogn Psychol, 1991, 3: 69–86
31 Boucher J, Lewis V. Unfamiliar face recognition in relatively able autistic children. J Child Psychol Psychiat, 1992, 33: 843–859
32 Hauck M, Fein D, Maltby N. Memory for faces in children withautism. Child Neuropsychol, 1998, 4: 187–198
33 Ellis HD, Ellis DM, Fraser W, Deb S. A preliminary study of right hemisphere cognitive deficits and impaired social judgments among young people with Asperger syndrome. Eur Child Adolesc Psychiat, 1994, 3: 255–266
34 Blair RJR, Frith U, Smith N, Abell F, Cipolotti L. Fractionation of visual memory: agency detection and its impairment in autism. Neuropsychologia, 2002, 40: 108–118
35 Boucher J, Lewis V, Collis G. Familiar face and voice matching and recognition in children with autism. J Child Psychol Psychiat Allied Discip, 1998, 39: 171–181
36 Dalton KM, Nacewicz BM, Johnstone T, Schaefer HS, Gernsbacher MA, Goldsmith HH, Alexander AL, Davidson RJ. Gaze fixation and the neural circuitry of face processing in autism. Nat Neurosci, 2005, 8: 519–526
37 Barton JJ, Cherkasova MV, Hefter R, Cox TA, O'Connor M, Manoach DS. Are patients with social developmental disorders prosopagnosic? Perceptual heterogeneity in the Asperger and socio-emotional processing disorders. Brain, 2004, 127: 1706–1716
38 Wilson R, Pascalis O, Blades M. Familiar face recognition in children with autism: The differential use of inner and outer face parts. J Autism Dev Disord, 2007, 37: 314–320
39 Klin A, Sparrow SS, de Bildt A, Cicchetti DV, Cohen DJ, Volkmar FR. A normed study of face recognition in autism and related disorders. J Autism Dev Disord, 1999, 29: 499–508
40 Szatmari P, Tuff L, Finlayson MA, Bartolucci G. Asperger's syndrome and autism: neurocognitive aspects. J Am Acad Child Adolesc Psychiat, 1990, 29: 130–136
41 Adolphs R, Sears L, Piven J. Abnormal processing of social information from faces in autism. J Cogn Neurosci, 2001, 13: 232–240
42 Bodamer J. Die Prosop-agnosie. Archiv fur Psychiatrie und Nervkrankheiten, 1947, 179: 6–47
43 Rossion B, Caldara R, Seghier M, Schuller AM, Lazeyras F, Mayer E. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain, 2003, 126: 2381–2395
44 Sorger B, Goebel R, Schiltz C, Rossion B. Understanding the functional neuroanatomy of acquired prosopagnosia. Neuroimage, 2007, 35: 836–852
45 McConachie HR. Developmental prosopagnosia. A single case report. Cortex, 1976, 12: 76–82
46 De Haan EH, Campbell R. A fifteen year follow-up of a case of developmental prosopagnosia. Cortex, 1991, 27: 489–509
47 Duchaine BK, Nakayama K. Dissociations of face and object recognition in developmental prosopagnosia. J Cogn Neurosci, 2005, 17: 249–261
48 Striemer C, Gingerich T, Striemer D, Dixon M. Covert face priming reveals a 'true face effect' in a case of congenital prosopagnosia. Neurocase, 2009, 15: 509–514
49 Behrmann M, Avidan G, Marotta JJ, Kimchi R. Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. J Cogn Neurosci, 2005, 17: 1130–1149
50 Duchaine B, Yovel G, Nakayama K. No global processing deficit in the Navon task in 14 developmental prosopagnosics. Soc Cogn Affect Neurosci, 2007, 2: 104–113
51 Yovel G, Duchaine B. Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental Prosopagnosia. J Cogn Neurosci, 2006, 18: 580–593
52 Le Grand R, Cooper PA, Mondloch CJ, Lewis TL, Sagiv N, de Gelder B, Maurer D. What aspects of face processing are impaired in developmental prosopagnosia? Brain Cogn, 2006, 61: 139–158
53 Duchaine BC. Developmental prosopagnosia with normal configural processing. Neuroreport, 2000, 11: 79–83
54 Grueter M, Grueter T, Bell V, Horst J, Laskowski W, Sperling K, Halligan PW, Ellis HD, Kennerknecht I. Hereditary prosopagnosia: the first case series. Cortex, 2007, 43: 734–749
55 Dobel C, Bolte J, Aicher M, Schweinberger SR. Prosopagnosia without apparent cause: overview and diagnosis of six cases. Cortex, 2007, 43: 718– 733
56 Duchaine BC, Germine L, Nakayama K. Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. Cogn Neuropsychol, 2007, 24: 419–430
57 Humphreys K, Avidan G, Behrmann M. A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. Exp Brain Res, 2007, 176: 356–373
58 Schwarzer G, Huber S, Grüter M, Grüter T, Gross C, Hipfel M, Kennerknecht I. Gaze behaviour in hereditary prosopagnosia. Psychol Res, 2007, 71: 583–590
59 Schmalzl L, Palermo R, Green M, Brunsdon R, Coltheart M. Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. Cogn Neuropsychol, 2008, 25: 704–729
60 Kennerknecht I, Grueter T, Welling B, Wentzek S, Horst J, Edwards S, Grueter M. First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). Am J Med Genet A, 2006, 140: 1617–1622
61 Grueter T, Grueter M, Carbon CC. Neural and genetic foundations of face recognition and prosopagnosia. J Neuropsychol, 2008, 2: 79–97
62 Langdell T. Recognition of faces: an approach to the study of autism. J Child Psychol Psychiat, 1978, 19: 255–268
63 Nunn JA, Postma P, Pearson R. Developmental prosopagnosia: Should it be taken at face value? Neurocase, 2001, 7: 15–27
64 Li XB, Song YY. First congenital prosopagnosia case reported in China. Prog Nat Sci, 2007, 17: 165–171
65 Farah MJ, Tanaka JW, Drain HM. What causes the face inversion effect? J Exp Psychol Hum Percept Perform, 1995, 21: 628–634
66 Maurer D, Grand RL, Mondloch CJ. The many faces of configural processing. Trends Cogn Sci, 2002, 6: 255–260
67 Bentin S, Deouell LY, Soroker N. Selective visual streaming in face recognition: evidence from developmental prosopagnosia. Neuroreport, 1999, 10: 823–827
68 de Gelder B, Rouw R. Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? Neuroreport, 2000, 11: 3145–3150
69 Duchaine BC, Parker H, Nakayama K. Normal recognition of emotion in a prosopagnosic. Perception, 2003, 32: 827–838
70 Duchaine BC. Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. Cogn Neuropsychol, 2006, 23: 714–747
71 Bentin S, De Gutis JM, D'Esposito M, Robertson LC. Too many trees to see the forest: Performance, event-related potential, and functional magnetic resonance imaging manifestations of integrative congenital prosopagnosia. J Cogn Neurosci, 2007, 19: 132–146
72 Steede LL, Tree JJ, Hole GJ. I can't recognize your face but I can recognize its movement. Cogn Neuropsychol, 2007, 24: 451–466
73 Bate S, Haslam C, Tree JJ, Hodgson TL. Evidence of an eye movement- based memory effect in congenital prosopagnosia. Cortex, 2008, 44: 806–819
74 Minnebusch DA, Suchan B, Ramon M, Daum I. Event-related potentials reflect heterogeneity of developmental prosopagnosia. Eur J Neurosci, 2007, 25: 2234–2247
75 Righart R, de Gelder B. Impaired face and body perception in developmental prosopagnosia. Proc Natl Acad Sci USA, 2007, 104: 17234–17238
76 Van den Stock J, van de Riet WA, Righart R, de Gelder B. Neural correlates of perceiving emotional faces and bodies in developmental prosopagnosia: an event-related fMRI-study. PLoS ONE, 2008, 3: e3195
77 Lee Y, Duchaine B C, Wilson H R, Nakayama K. Three cases of developmental prosopagnosia from one family: Detailed neuropsychological and psychophysical investigation of face processing. Cortex, 2009, 46: 949–964
78 Bowles DC, Mckone E, Dawel A, Duchaine B, Palermo R, Schmalzl

L, Rivolta D, Wilson CE, Yovel G. Diagnosing prosopagnosia: Effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. Cogn Neuropsychol, 2009, 26: 423–455
79 Duchaine BC, Nakayama K. Developmental prosopagnosia and the benton facial recognition test. Neurology, 2004, 62: 1219–1220
80 Harris AM, Duchaine BC, Nakayama K. Normal and abnormal face selectivity of the M170 response in developmental prosopagnosics. Neuropsychologia, 2005, 43: 2125–2136
81 Behrmann M, Avidian G, Gao F, Black S. Structural imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. Cereb Cortex, 2007, 17: 2354–2363
82 Bate S, Haslam C, Jansari A, Hodgson TL. Covert face recognition relies on affective valence in congenital prosopagnosia. Cogn Neuropsychol, 2009, 26: 391–411
83 Duchaine BC, Nakayama K. The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia, 2006, 44: 574–583
84 Duchaine BC, Weidenfeld A. An evaluation of two commonly used tests of unfamiliar face recognition. Neuropsychologia, 2003, 41: 713–720
85 Le Grand R, Mondloch CJ, Maurer D, Brent HP. Early visual experience and face processing. Nature, 2001, 410: 890
86 Riesenhuber M, Jarudi I, Gilad S, Sinha P. Face processing in humans is compatible with a simple shape-based model of vision. Proc R Soc Lond Ser B-Biol Sci, 2004, 271: 448–450
87 Wang R, Li J, Fang H, Tian M, Liu J. Individual differences in holistic processing predict face-recognition ability. Psychol Sci, 2012, 23: 169–177
88 Russell R, Duchaine BC, Nakayama K. Super-recognizers: People with extraordinary face recognition ability. Psychon Bull Rev, 2009, 16: 252–257
89 Thomas C, Avidan G, Humphreys K, Jung KJ, Gao F, Behrmann M. Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. Nat Neurosci, 2008, 12: 29–31
90 Golarai G, Liberman A, Yoon JMD, Grill-Spector K. Differential development of the ventral visual cortex extends through adolescence. Front Hum Neurosci, 2010, 3: 80
91 Savaskan E, Ehrhardt R, Schulz A, Walter M, Schächinger H. Post-learning intranasal oxytocin modulates human memory for facial identity. Psychoneuroendocrinology, 2008, 33: 368–374
92 Zhu Q, Song YY, Hu SY, Li X, Tian M, Zhen Z, Dong Q, Kanwisher N, Liu J. Heritability of the specific cognitive ability of face perception. Curr Biol, 2010, 20: 137–142
93 Wilmer JB, Germine L, Chabris, Chatterjee G, Williams M, Loken E, Nakayama K, Duchaine B. Human face recognition ability is specific and highly heritable. Proc Natl Acad Sci USA, 2010, 107: 5238–5241
94 Luck SJ, Hollingworth A. Visual Memory. Oxford: Oxford University Press, 2008
95 Tanaka JW, Farah MJ. Parts and wholes in face recognition. Q J Exp Psychol A, 1993, 46: 225–245
96 Tanaka JW, Sengco JA. Features and their configuration in face recognition. Mem Cognit, 1997, 25: 583–592
97 Grüter T, Grüter M, Bell V, Carbon CC. Visual mental imagery in congenital prosopagnosia. Neurosci Lett, 2009, 453: 135–140
98 Zhu Q, Li XB, Chow K, Liu J. The part task of the part-spacing paradigm is not a pure measurement of part-based information of faces. PLoS ONE, 2009, 4: e6239
99 Duchaine BC, Jenkins R, Germine L, Calder AJ. Normal gaze discrimination and adaptation in seven prosopagnosics. Neuropsychologia, 2009, 47: 2029–2036