Neuroscience and Biobehavioral Reviews 36 (2012) 1060–1084

Contents lists available at SciVerse ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review

Face identity recognition in autism spectrum disorders: A review of behavioral studies

Sarah Weigelt*, Kami Koldewyn, Nancy Kanwisher

Department of Brain and Cognitive Science and McGovern Institute for Brain Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 2 September 2011
Received in revised form
23 November 2011
Accepted 14 December 2011

Keywords:
Autism
Autism spectrum disorders
Asperger's syndrome
Pervasive developmental disorders
Neurodevelopmental disorders
Face perception
Face recognition
Identity perception
Face inversion effect
Behavioral studies
Psychophysics

ABSTRACT

Face recognition – the ability to recognize a person from their facial appearance – is essential for normal social interaction. Face recognition deficits have been implicated in the most common disorder of social interaction: autism. Here we ask: is face identity recognition in fact impaired in people with autism? Reviewing behavioral studies we find no strong evidence for a *qualitative* difference in *how* facial identity is processed between those with and without autism: markers of typical face identity recognition, such as the face inversion effect, seem to be present in people with autism. However, *quantitatively* – i.e., *how well* facial identity is remembered or discriminated – people with autism perform worse than typical individuals. This impairment is particularly clear in face memory and in face perception tasks in which a delay intervenes between sample and test, and less so in tasks with no memory demand. Although some evidence suggests that this deficit may be specific to faces, further evidence on this question is necessary.

Contents

* Corresponding author. Tel.: +1 617 258 0670; fax: +1 617 258 8654.
E-mail addresses: weigelt@mit.edu (S. Weigelt), kamik@mit.edu (K. Koldewyn), ngk@mit.edu (N. Kanwisher).

0149-7634/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neubiorev.2011.12.008

1. Introduction

Deficits in face recognition are widely argued to be important for or even at the core of the social impairments of people with autism spectrum disorders or ASD (Dawson et al., 2005; Schultz, 2005). However, the literature on this topic (Dawson et al., 2005; Golarai et al., 2006; Jemel et al., 2006; Marcus and Nelson, 2001; Pierce and Courchesne, 2000; Sasson, 2006; Simmons et al., 2009) is mixed. Here, we attempt to make sense of the apparent inconsistencies in the literature, focusing on one aspect of face perception, the recognition of facial identity (for a recent review on the recognition of emotional expressions see Harms et al., 2010). Potential deficits in face identity recognition are of particular importance because (in contrast to deficits in facial emotion recognition) they are not part of the diagnostic criteria for autism. Thus, here we ask: does the current scientific literature demonstrate a true deficit in face identity recognition in people with autism compared to typical individuals? And if so, what is the precise nature of that deficit? Is it specific to faces, or, is it part of a more general deficit in object recognition? Is it a deficit in perceptual discrimination, or memory, or both?

The answers to these questions are important for understanding the etiology and treatment of autism. If recognition of faces is indeed impaired in autism, it will be important to determine whether these deficits play a fundamental causal role in autism by leading to other aspects of the cognitive phenotype. That is, might deficits in face recognition in part cause broader deficits in social cognition? If so, then training better face recognition skills may lead to improvements in other cognitive abilities. Alternatively, deficits in face recognition might be consequences, not causes, of other cognitive deficits. In this case, training face recognition will not address the broader cognitive profile in autism. The first question, though, is whether the widely claimed deficit in face recognition is real, and if so, what is the precise nature of that deficit.

To answer these questions we distinguish between qualitative and quantitative differences in face identity recognition (McKone et al., 2009) between people with and without ASD. Qualitative differences refer to *how* facial identity is remembered or discriminated in people with and without ASD. To evaluate whether people with ASD process faces in the same way as typical participants, we first report on studies that test whether people with autism show the typical "face markers" – known psychophysical effects that are found only or more strongly for faces than for non-face objects, and that indicate "typical" face identity recognition. Examples of these face markers are the face inversion effect (Yin, 1969) and the composite effect (Young et al., 1987). Absent or reduced face markers in people with ASD would indicate qualitative differences in face identity recognition. Second, we review studies that test for quantitative differences in face identity recognition, i.e., *how well* facial identity is discriminated or remembered in people with and without ASD compared to typical individuals. Note that qualitative and quantitative differences in face recognition may occur independently from one another or jointly.

1.1. Inclusion/exclusion criteria for the review

Most prior studies and reviews on face recognition in ASD cite only a subset of the relevant literature. To avoid this problem, we invested considerable effort to conduct a comprehensive review. Specifically, we searched for all studies published through April 2011 that contained experiments on face identity recognition (including the face markers) on participants with ASD on PubMed, the Web of Science and the database of The National Autistic Society. Although this review is intended to be all-encompassing, we excluded studies or subparts of them if either: (1) they targeted participants who did not meet all criteria for autism spectrum disorder or ASD or (2) they lacked proper statistics or prerequisites for such statistics. Throughout the review we explain the appropriate statistical tests necessary for demonstrating group differences between participants with versus without ASD, and we evaluate studies accordingly. Appendix A lists all the excluded studies and reasons for their exclusion. We included studies testing individuals with all diagnoses along the whole range of ASD including autism, Asperger's syndrome, and pervasive developmental disorder not otherwise specified (PDD-NOS). We realize that the criteria for diagnosing ASD have changed (and are changing, see the current debate on the revision of the DSM http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=94). Therefore, we report participants' diagnoses and the diagnostic criteria in extensive tables. Additionally, these tables provide detailed information on the experiments such as the criteria used for matching ASD and typical groups, number and age of the participants, the type of experimental paradigm, an assessment of the memory and perceptual demands of the tasks, and statistical results. The main findings of this review are shown in summary Figs. 1 and 2.

2. Face markers

In this section, we ask *how* people with ASD process faces, or more precisely if people with ASD process faces in the same way typical people do. Several well-known behavioral phenomena, such as the face inversion effect (discussed below) have been treated as "signatures" of typical face identity recognition. Other "face markers" demonstrate that we typically integrate information from the different face parts in a way that leads to the perception of a face "as a whole". Some researchers have argued that it is this holistic or configural aspect of face perception that might be compromised in ASD (Dawson et al., 2005; Gauthier et al., 2009). Because face markers are stronger for faces compared to non-face object categories, and developmental studies have shown that all face markers are present at the earliest age measured (McKone et al., 2009), the

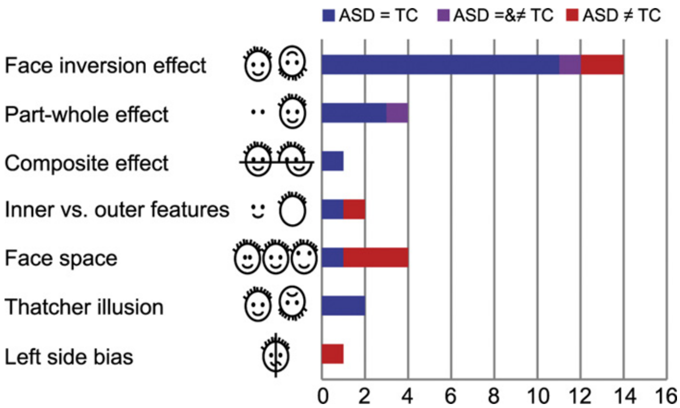

Fig. 1. Studies on markers of typical face processing. Number of studies finding same (blue), mixed (purple) or worse (red) performance of people with ASD in comparison to typical individuals for seven prominent markers of typical face processing. Few studies find differences between ASD and typical participants; important hallmarks of typical face perception – e.g. the face inversion effect – seem to be intact in most people with ASD. The results of one additional study for the face inversion effect were uninterpretable. Please refer to the text and Tables 1 and 2 for detailed information on these studies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

presence of face markers is indicative of a typically functioning face processing system. When typical face processing is impaired – as in developmental or acquired prosopagnosia – face markers are weak or absent (Busigny and Rossion, 2010; Ramon et al., 2010). Evaluating whether face markers are present in people with ASD will thus allow us to assess whether their face perception system works in the same way as the face perception system in typical individuals.

2.1. The face inversion effect

The best-known hallmark of typical face perception is the face inversion effect: face recognition is more accurate when faces are presented (and tested) upright than inverted. The inversion effect is larger for faces than for any other object category (Yin, 1969), and hence is most accurately called a “disproportionate inversion effect”. A number of studies have tested participants with ASD on upright and inverted faces (see Table 1). We divided these studies into 1) studies that set out to investigate the face inversion effect directly (Hobson et al., 1988; Lahaie et al., 2006; Rose et al., 2007; Scherf et al., 2008), 2) studies that investigated other aspects of face identity perception (e.g. sensitivity to spacing changes), but in that used upright and inverted faces (Faja et al., 2009; Joseph and Tanaka, 2003; Langdell, 1978; Nishimura et al., 2008; Riby et al., 2009; Rouse et al., 2004; Rutherford et al., 2007), and 3) studies that tested the effect of face inversion on other tasks, e.g. on emotional labeling (Bar-Haim et al., 2006; Falck-Ytter, 2008; Rosset et al., 2008; Tantam et al., 1989; van der Geest et al., 2002). Although these latter studies did not investigate face identity recognition, we included them because others have cited them as addressing the face inversion effect. One additional study was excluded from the review (Bookheimer et al., 2008; see Appendix A).

2.1.1. Statistical criteria

To demonstrate that participants with ASD do not show a face inversion effect (or do so to a significantly lesser extent) than typical individuals (T), a statistically significant interaction between group (T, ASD) and face orientation (upright, inverted) is necessary. Even stronger evidence would come from a triple interaction, the observation of a larger disproportionate inversion effect for faces compared to other object categories (e.g. shoes), for T versus ASD.

2.1.2. Studies investigating the face inversion effect directly

Four studies have investigated the face inversion effect directly; three of these showed an inversion effect in ASD, and one did not. In one study, Hobson et al. (1988) used a card-sorting task, in which participants had to match either face identity or emotional expression. Participants were tested on upright faces during their first session. At a second session a few days later, they were tested on inverted faces and performed the same matching task on the same stimuli. Both adolescents with ASD and the CA- and NVMA-matched typical individuals showed a face inversion effect in within-group statistics. However, authors do not report whether there is an interaction between group and stimulus orientation.

Lahaie et al. (2006) investigated the disproportionate inversion effect, using Greebles and faces and testing adults with and without autism. Between-group statistics on the accuracy (but not RT) data found a significant 3-way interaction of group × orientation × type of stimulus. This finding, however, appears to primarily result from the typical participants not showing a face inversion effect in accuracy (although they show an inversion effect in the RT data). Nonetheless, the adult ASD participants do show the expected disproportionate inversion effect, both in accuracy and RT. Therefore, Lahaie et al.’s study does not support a qualitative difference in face recognition between people with and without autism.

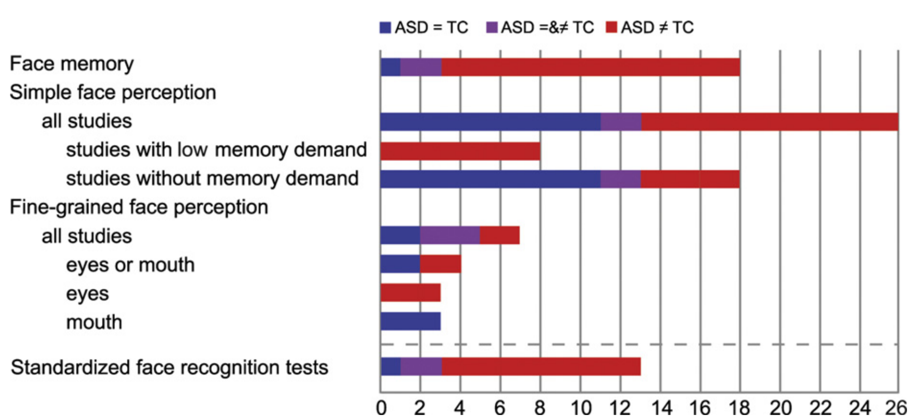

Fig. 2. Face identity recognition. Number of studies finding same (blue), mixed (purple) or worse (red) performance of people with ASD in comparison to typical individuals in tasks on face identity recognition. Evident are deficits in face memory (demonstrated in the experimental as well as standardized face recognition tests and in perceptual tests when they contain a stronger memory demand) and fine-grained perceptual discrimination of eyes. Note that studies using standardized face recognition tests are counted in the respective sections “Face memory” and “Simple face perception” as well as reported separately (marked by the dashed line). Note also that for fine-grained face perception, the mixed results can be resolved by looking at eye versus mouth perception separately. Please refer to the text and Tables 3 and 4 for detailed information on these studies. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1
Studies on the face inversion effect.

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Between-group-statistics
Bar-Haim et al. (2006)	ASD (DSM-IV, ADI-R, ADOS)	12	CHI	10.2	Detecting a dot on a face	–	–	Group (ASD, T) × orientation (upright, inverted) interaction was not significant.
	T (CA, VMA, NVMA)	12		10.2				
Faja et al. (2009)	ASD (DSM-IV, ADI-R, ADOS)	39	ADU	24.0	Delayed 2AFC match-to-sample	Low; 1 s ISI	Low; identical image (or part of image)	Group (ASD, T) × orientation (upright, inverted) was not significant (p nr). Main effect of orientation (upright > inverted) was significant ($p < .001$).
	T (CA, VMA, NVMA)	33		24.6				
Falck-Ytter (2008)	ASD (ADI-R, ADOS)	15	CHI	5.2	Passive viewing	No task	No task	Group (ASD, T) × orientation (upright, inverted) was not significant (p nr). ASD: main effect orientation (upright > inverted) was marginally significant ($p = .056$). T: main effect orientation (upright > inverted) was marginally significant ($p = .054$).
	T (CA)	15		4.9				
Hobson et al. (1988)	ASD (Rutter)	17	ADO/ADU	19.0	Sorting	Negligible; simultaneous presentation	Medium; across emotional expression	No between-group statistics were performed. ASD: main effect of orientation (upright > inverted) was significant ($p < .01$). T: main effect of orientation (upright > inverted) was significant ($p < .001$).
	T	17		18.1				
Joseph and Tanaka (2003)	ASD (ADI-R)	22	CHI	10.9	Delayed 2AFC match-to-sample	Low; 0 ISI	Low; identical image (or part of image)	Group (ASD, T) × orientation (upright, inverted) interaction was not significant (p ns), but main effect of orientation (upright > inverted) was significant ($p < .001$). Group (ASD, T) × orientation (upright, inverted) × feature (eye, mouth) interaction was significant ($p < .01$). ASD: main effect orientation (upright > inverted) was significant for mouths ($p < .01$), but not eyes (p ns). T: main effect orientation (upright > inverted) was significant for mouths ($p < .05$), and eyes ($p < .01$)
	T (CA, VMA, NVMA)	20		10.8				
Lahaie et al. (2006)	ASD (ADI-R, ADOS)	16	ADO/ADU	20.7	Delayed 2AFC match-to-sample	Low; 24 ms ISI	Low; identical image	Group (ASD, T) × orientation (upright, inverted) × type of stimulus (faces, greebles) was significant in accuracy ($p < .025$,) but not in RT ($p > .09$). Group (ASD, T) × orientation (upright, inverted) was significant for faces ($p < .045$), but not for greebles ($p > .3$). ASD: main effect orientation (upright > inverted) was significant ($p < .045$). T: main effect orientation (upright > inverted) was not significant ($p > .1$).
	T (CA, VMA, NVMA)	16		20.3				
Langdell (1978)	ASD (Rutter)	10	CHI	9.8	Identify peer	Not controlled; probably very high depending on when peer was last seen	High: 'real' peer vs. photograph	Not performed
	T (CA/MA/CA, MA)	10		5.6/9.6/9.8				
	ASD (Rutter)	10	ADO	14.1				
	T (CA/MA/CA, MA)	10		8.1/13.6/13.7				
Nishimura et al. (2008)	ASD (ADI-R, ADOS)	17	ADU	20.6	Simult. same–different	Negligible; simultaneous presentation	From low to high depending on displacements	Group (ASD, T) × orientation (upright, inverted) was not significant ($p = .25$), but main effect of orientation (upright > inverted) was significant ($p < .01$).
	T (CA, VMA, NVMA)	17		21.6				
Riby et al. (2009)	ASD (DSM-IV, CARS)	20	CHI/ADO	14.8	Simult. 2AFC match-to-sample	Negligible; simultaneous presentation	Medium; across pose	Group (ASD, T1, T2, T3) × orientation (upright, 90°, inverted) was not significant (p nr). Main effect of orientation (upright > 90° > inverted) was significant (p .001). Main effect group was significant: ASD performed worse than T ($p < .001$).
	T1 (CA)	20		14.9				
	T2 (VMA)	20		6.5				
	T3 (NVMA)	20		7.9				

Table 1 (*Continued*)

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Between-group-statistics
Rose et al. (2007)	ASD (DSM-IV)	16	CHI	10.3	Delayed same–different	Low; 500 ms ISI	Low; identical image	Group (ASD, WS, T) × orientation (upright neutral, upright affect, inverted neutral) interaction was significant ($p < .0001$). ASD: main effect orientation (upright > inverted) was not significant (*p* nr). WS: main effect orientation (upright > inverted) was significant (*p* nr). T: main effect orientation (upright > inverted) was significant (*p* nr).
	WS (VMA)	19	ADU	26.2				
	T (CA, NVMA)	17	CHI	10.0				
Rosset et al. (2008)	ASD (DSM-IV, ADI-R)	20	CHI	9.5	Categorize picture as happy/not-happy	No	Difficult to determine	Group (ASD, T1, T2) × orientation (upright, inverted) interaction was not significant (*p* nr), main effect group (ASD, T1, T2) was not significant ($p > .05$), but main effect orientation (upright > inverted) was significant ($p < .001$) ASD/T1/T2: all main effects of orientation (upright > inverted) were significant ($p < .001$).
	T1 (CA)	20		9.6				
	T2 (MA)	20		8.0				
Rouse et al. (2004)	ASD (ICD-10)	11	CHI	9.6	Odd-one-out	Negligible; simultaneous presentation	Difficult to determine	Group (ASD, T) × orientation (upright, inverted) interaction was not significant (*p* nr). Orientation (upright, inverted) × stimulus type (face, house) was significant ($p < .01$). Main effect of orientation (upright > inverted) was significant ($p < .01$).
	T (CA, NVMA)	15		9.4				
Rutherford et al. (2007)	ASD (ADI-R, ADOS)	16	ADO/ADU	19.6	Odd-one-out	Negligible; simultaneous presentation	From low to high; depending on displacements	Group (ASD, T) × orientation (upright, inverted) was not significant for eyes (*p* < ns) nor for mouths (*p* < ns).
	T (CA, NVMA)	19		24.3				
Scherf et al. (2008)	ASD (ADI-R, ADOS)	15	CHI	11.0	Delayed 2AFC match-to-sample	Low; 1 s ISI	Low; identical image	Group (ASD children, ASD adults, T children, T adults) × orientation (upright, inverted) interaction was not significant (*p* nr), only main effects of orientation (upright > inverted), age (adults > children) and group (ASD, T) were significant: ASD performed worse than T.
	T (CA, VMA, NVMA)	15	CHI	12.0				
	ASD (ADI-R, ADOS	15	ADU	32.0				
	T (NVMA)	15	ADU	22.0				
van der Geest et al. (2002)	ASD (DSM-IV, ADI-R)	16	CHI	10.8	Passive viewing	No task	No task	Group (ASD, T) × orientation (upright, inverted) interaction was significant ($p < .05$).
	T (CA, VMA, NVMA)	13		9.9				

Note: Mean age is in years. Memory demand was assessed on a five-point scale from negligible (simultaneous presentation), over low (0–2 s delay), medium (2 s to 1 min delay), high (1–5 min delay), to very high (>5 min delay). Perceptual demand was assessed on a three-point scale from low (matching identical images), over medium (matching across one dimension, e.g. pose), to high (matching across two dimensions, e.g. pose and lighting). Abbreviations (other than in Appendix A): ADO: adolescents; ADU: adults; CHI: children; ISI: interstimulus interval; *p* nr: *p* not reported; *p* ns: p not significant.

Consistent with this conclusion, Scherf et al. (2008) also found robust face inversion effects in children and adults with ASD and their respective CA- and IQ-matched typical groups. The interaction between group and orientation was not significant.

Only one published study – Rose et al. (2007) – reports the lack of a face inversion effect in children with ASD. Participants were instructed to perform a sequential same/different discrimination task where the two faces were presented either upright or inverted. Between-group statistics revealed a significant group × orientation interaction. However, the "group" factor consisted of three groups, a group of children with ASD, a group of typical children, and a group of adults with Williams syndrome and the "orientation" factor consisted of three orientation conditions, upright neutral faces, upright emotional faces and inverted neutral faces. These complications render the interaction difficult to interpret. Indeed, direct comparisons revealed that neither the upright nor the inverted conditions showed a significant difference between ASD and typical participants. Only the within-group statistics showed that children with ASD performed as well in the inverted as they did in the upright conditions (thus showing no face inversion effect), while typical children and adults with Williams syndrome performed worse in the inverted than the upright condition.

In summary, three out of four studies directly studying the face inversion effect, including one study on the disproportionate inversion effect, show that participants with ASD do exhibit face inversion effects. The one study that claimed not to find a face inversion effect in their ASD sample (Rose et al., 2007) did not report strong statistical tests of that claim.

2.1.3. Studies investigating other aspects of face identity recognition, but in that used upright and inverted faces

Seven studies investigated the effect of inversion on other aspects of face identity recognition (Faja et al., 2009; Joseph and Tanaka, 2003; Langdell, 1978; Nishimura et al., 2008; Riby et al., 2009; Rouse et al., 2004; Rutherford et al., 2007).

Some of these studies are also reviewed in subsequent sections of this paper because they are relevant to other face markers. Here we summarize their findings with respect to the effect of inversion only.

Two studies have been cited as revealing a qualitative difference in face recognition between people with and without ASD. In one of the most widely cited studies to test the perception of inverted faces in people with autism, Langdell (1978) tested children and adolescents with autism and both CA-matched typical and MA-matched typical individuals. This author first showed the participants inverted photographs of ten peers and asked them to identify them. After some intervening trials with different experimental conditions, but the same face photographs, Langdell showed the participants the same photographs upright. Because of the other experimental conditions potentially interfering with the recognition process, the order confound (always presenting the inverted face first), and the fact that the results for the upright condition were not reported, and no statistical tests of the face inversion effect were performed, the data from this study are impossible to interpret. This study simply does not provide evidence for or against a face inversion effect in ASD that is often attributed to it.

Joseph and Tanaka (2003) explored the effect of inversion on the part-whole effect. Between-group comparisons revealed that the simple group × orientation interaction was not significant. However, a significant group × orientation × feature (eyes, mouth) interaction showed that while typical children had an inversion effect for both eyes and mouths, ASD children only showed an inversion effect for mouths, but not for eyes. This finding is particularly interesting as it hints toward a deficit in processing eye information in those with ASD – a hypothesis we will follow in the later Section 3.3.

Besides these two papers, all other papers show similar face inversion effects in participants with and without ASD. Brief descriptions of each of these other studies follow.

Faja et al. (2009) investigated the effect of inversion on part-whole processing. They found a significant group × orientation × configuration (whole, part) interaction, while the simple group × orientation interaction was not significant. Because they did not perform post hoc tests to explore these interactions further, their results are difficult to interpret. From the reported means, it seems that the interaction might result from typical participants showing an inversion effect for both whole and part configurations, while ASD participants show an inversion effect more strongly for whole than for part configurations. In other words, ASD show the expected pattern of responses, while T do not.

Rutherford et al. (2007) investigated the effect of inversion on sensitivity to spacing (between face parts). The group × orientation interaction was not significant and both ASD and typical participants showed an inversion effect for the eye-to-eye-spacing condition. Neither T nor ASD participants showed inversion effects for the mouth-to-nose spacing condition.

Nishimura et al. (2008) also investigated the effect of inversion on sensitivity to spacing (on stimuli that consisted of both eye-to-eye and mouth-to-nose differences). The interaction between group and orientation was not significant, but the main effect of orientation was: both ASD and T groups showed inversion effects.

Rouse et al. (2004) investigated the effect of inversion on the Thatcher illusion (Thompson, 1980) using both face and house stimuli. ASD participants and CA- and NVMA-matched typical participants showed the expected inversion effect, and the group × orientation interaction was not significant. However, typical participants were at ceiling in the upright face condition, so the magnitude of the inversion effect may have been underestimated in the T group. Nonetheless, these authors also found evidence for a disproportionate inversion effect in a Thatcher-illusion paradigm (in both groups) based on an orientation × stimulus type (face, house) interaction.

Riby et al. (2009) investigated a group of children with ASD and three groups of typical individuals (either CA-, or VMA-, or NVMA-matched). Participants performed a 2AFC match-to-sample task on either the upper or lower halves of faces that were either presented upright, rotated 90° or fully inverted. The overall cost of inversion was significant (with a main effect of orientation in which upright was the easiest condition followed by 90° rotation and then inverted) and this effect apparently did not differ between groups (the group (ASD, T1, T2, T3) × orientation (upright, 90°, inverted) interaction was not significant). There was also a main-effect of group, with the ASD participants performing worse than any of the typical groups. However, no within-group comparisons were made to evaluate whether there was a significant face inversion effect in each group.

In summary, much like the studies on the pure face inversion effect, five out of seven studies reviewed here demonstrate similar inversion effects in their ASD and T samples. One of the other two studies is impossible to interpret (Langdell, 1978) and the other found inversion effects only for the mouth, not the eyes, in people with ASD (Joseph and Tanaka, 2003).

2.1.4. Studies using upright and inverted faces in other experimental paradigms

Upright and inverted faces have been used in other experimental paradigms that did not test face identity recognition. Although these studies are not of primary relevance to this review, we discuss them here because some of the studies are widely cited as evidence for the absence of a face inversion effect in people with ASD (e.g. Tantam et al., 1989). It is important to note that the classic face inversion effect concerns face *identity* recognition, so its presence or absence in other paradigms (such as facial expression recognition) does not allow for any firm conclusions about the functionality of the face recognition system.

Two studies report an absence of a face inversion effect in their ASD sample. Tantam et al. (1989) (see Appendix A) is widely cited as evidence for an absence of the face inversion effect in children with ASD in a task on emotional labeling. However, performance in the ASD children was at floor for the upright face condition, leaving no room for further reduction of performance for the inverted face condition, and thus making the data impossible to interpret. Thus, this study cannot be used as evidence that the face inversion effect is absent in children with ASD. The other study reporting an apparent absence of the face inversion effect in children with ASD measured looking behavior toward upright and inverted faces (van der Geest et al., 2002). They found a significant group × orientation interaction, with typical children spending more time looking at upright than inverted faces, while children with ASD spent the same amount of time looking at upright and inverted faces. In direct contrast, Falck-Ytter (2008) did not find any differences in the looking behavior of very young (M = 5.2 years) children with and without ASD: both groups spent more time looking at upright than inverted faces.

Two more studies did not find differences between their ASD and typical participants. Rosset et al. (2008) studied emotion recognition of upright and inverted faces, and found no between-group differences, but a clear main effect of orientation with better performance for upright than inverted faces. Finally, Bar-Haim et al. (2006) asked a group of ASD children and CA- and IQ-matched typical children to detect a briefly presented dot that was superimposed on either upright or inverted faces, with the dot presented either between the eyebrows or beneath the mouth. The group × orientation interaction was not significant. Inversion effects, however, were only found when the dot was presented between the eyebrows, with better performance for upright than inverted faces.

Thus, even in paradigms not addressing face identity recognition, the majority of studies show a similar face inversion effect in people with ASD and typical individuals.

Summarizing all studies on the face inversion effect (see also Fig. 1) excluding the two studies that are uninterpretable (Langdell, 1978; Tantam et al., 1989) only two of 14 studies (Rose et al., 2007; van der Geest et al., 2002) report an absence of face inversion effects in children with ASD – only one of which actually measured face identity recognition (Rose et al., 2007). Of these studies one has weak statistical evidence (Rose et al., 2007), and the other (van der Geest et al., 2002) was directly contradicted by a subsequent study (Falck-Ytter, 2008). Thus, the majority of the 14 studies (12 out of 14) demonstrate similar face inversion effects in people with autism and typical individuals (Bar-Haim et al., 2006; Faja et al., 2009; Falck-Ytter, 2008; Hobson et al., 1988; Joseph and Tanaka, 2003; Lahaie et al., 2006; Nishimura et al., 2008; Riby et al., 2009; Rouse et al., 2004; Rosset et al., 2008; Rutherford et al., 2007; Scherf et al., 2008). We conclude that people with ASD do not demonstrate qualitative differences in the best-known hallmark of typical face recognition – the face inversion effect.

2.2. The part-whole effect

An important marker for holistic/configural processing is the part-whole effect or "whole advantage" (Tanaka and Farah, 1993). This refers to the finding that participants are better at discriminating eyes or mouths when these parts are embedded in a full face than when they are shown in isolation. The whole advantage has been found only for upright faces, not for inverted faces. It is stronger for faces than non-face stimuli. Four studies have investigated the part-whole effect in participants with ASD (Faja et al., 2009; Joseph and Tanaka, 2003; Lopez et al., 2004; Wolf et al., 2008; see Table 2), one additional study was excluded from the review (Annaz et al., 2009; see Appendix A).

2.2.1. Statistical criteria

Evidence that participants with ASD do not show the part-whole effect would require a significant interaction between group (ASD, T) and feature (part, whole) as well as a subsequent within-group test showing that typical participants show a whole-advantage when viewing faces, but ASD participants do not.

Two studies found no differences in the part-whole effect between ASD and typical participants. The first study by Joseph and Tanaka (2003) tested four groups of children: 9-year-old and 11-year-old typical children, a group of ASD children and a CA- and IQ-matched group with language impairments or delays. First, both the 9- and 11-year-old typical groups demonstrated a whole advantage. Second, children with ASD also showed a whole advantage that was statistically equivalent to that shown by the typical children: between-group comparisons between the ASD group and the control group of children with language impairments found that the group × configuration interaction was not significant. Furthermore, children with ASD showed the expected orientation (upright, inverted) × feature (whole, part) interaction, further corroborating that they were, in fact, showing the whole advantage.

Wolf et al. (2008) extended Joseph and Tanaka (2003)'s finding by confirming the presence of a whole advantage in a large number of children: 66 with ASD and 68 CA- and IQ-matched typical children. A main effect of configuration revealed the expected whole advantage and this effect did not differ between groups (as evidenced by the lack of a group × configuration interaction).

Faja et al. (2009) studied adult ASD participants on a similar part-whole paradigm. ASD participants demonstrated a whole advantage, as did typical participants – the group × configuration interaction was not significant. A significant 3-way interaction between group × configuration × orientation (p = .05) seemed to be even driven by a *larger* whole advantage in the upright condition in the ASD participants than in the typical participants (this observation is based on the reported means as no post hoc within-group tests were conducted to follow-up on the interaction).

Lopez et al. (2004) investigated the effect of cuing on the part-whole effect. In the cued condition a written cue (e.g. "Look at the mouth") informed participants on which feature to base their matching. Lopez et al. (2004) found a whole advantage for their children and adolescents with ASD in the cued condition only. In contrast, CA- and IQ-matched T showed a whole advantage in both cued and un-cued conditions. These results suggest that individuals with ASD may attend differently to faces by default, but can – if instructed – attend to faces the same way typical participants do, and when this occurs, they show the typical whole advantage.

In summary, the results on the part-whole effect in ASD argue strongly for a typical part-whole effect in ASD (see also Fig. 1): two studies with strong statistical evidence found the expected part-whole effect in those with ASD (Joseph and Tanaka, 2003; Wolf et al., 2008), another one seems to show an even stronger part-whole effect in ASD than T (Faja et al., 2009), while one study found it when the relevant feature was cued (Lopez et al., 2004).

2.3. The composite effect

Another marker for holistic/configural processing is the composite effect: participants find it harder to identify one half of a composite face (e.g. top half of Barack Obama's face with bottom half of Will Smith's) if the inconsistent other half-face is spatially

Table 2
Studies on face markers.

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Between-group-statistics
The part-whole effect								
Faja et al. (2009)	ASD (DSM-IV, ADI-R, ADOS)	39	ADU	24.0	Delayed 2AFC match-to-sample	Low; 1 s ISI	Low; identical image (or part of image)	Group (ASD, T) × configuration (whole, part) approached significance ($p=.05$) – but this effect was driven by a larger whole advantage in ASD than in T.
	T (CA, VMA, NVMA)	33		24.6				
Joseph and Tanaka (2003)	ASD (ADI-R)	22	CHI	10.9	Delayed 2AFC match-to-sample	Low; 0 ISI	Low; identical image (or part of image)	Group (ASD, T) × configuration (whole, part) interaction was not significant (*p* ns). ASD: orientation (upright, inverted) × configuration (whole, part) interaction was significant ($p<.05$). T: orientation (upright, inverted) × configuration (whole, part) interaction was not significant ($p=.16$).
	T (CA, VMA, NVMA)	20		10.8				
Lopez et al. (2004)	ASD (DSM-III-R, DSM-IV, ICD-10)	17	CHI/ADO	13.0	Delayed 2AFC match-to-sample	Low; 500 ms ISI	Low; identical image (or part of image)	Group (ASD, T) × configuration (whole, part) × cue (cued, uncued) interaction was significant ($p<.05$). ASD: main effect of configuration (whole > part) only in the cued condition significant ($p<.01$), not in the uncued condition ($p>.05$). T: main effect of configuration (whole > part) was significant in both cued and uncued conditions (both $p<.05$).
	T (CA, VMA, NVMA)	17		13.1				
Wolf et al. (2008)	ASD (DSM-IV, ADI-R, ADOS)	66	CHI	11.9	Delayed 2AFC match-to-sample	Low; 0 ISI	Low; identical image (or part of image)	Group (ASD, T) × configuration (whole, part) interaction was not significant (*p* ns).
	T (CA, VMA, NVMA)	68		11.9				
The composite effect								
Gauthier et al. (2009) [measures a congruency, not the composite effect]	ASD (ADI-R, ADOS)	21	CHI/ADO	12.8	Delayed same–different	Low; 1430 ms ISI	Low: identical image (or part of image)	Congruency effect: Group (ASD, T) × alignment (aligned, misaligned) × congruency (congruent, incongruent) interaction was significant ($p=.04$). ASD: alignment (aligned, misaligned) × congruency (congruent, incongruent) interaction was not significant ($F<1$). T: alignment (aligned, misaligned) × congruency (congruent, incongruent) interaction was significant ($p<.001$).
	T (CA, VMA, NVMA)	21		12.0				
Nishimura et al. (2008)	ASD (ADI-R, ADOS)	17	ADU	20.6	Delayed same–different	Negligible; simultaneous presentation	From low to high depending on displacements	Group (ASD, T) × orientation (upright, inverted) was not significant ($p=.25$), but main effect of orientation (upright > inverted) was significant ($p<.01$).
	T (CA, VMA, NVMA)	17		21.6				
Inner vs. outer features effect								
Rondan et al. (2003)	ASD (DSM-IV, CARS)	14	CHI	10.1	Simult. 2AFC match-to-sample	Negligible; simultaneous presentation	Low; identical image (or part of image)	Group (ASD, T1, T2) × feature (inner, outer) interaction was significant ($p<.05$). ASD: main effect feature was not significant ($p>.5$). T1: main effect feature was significant ($p<.05$). T2: main effect feature was significant ($p<.001$).
	T1 (CA)	14		10.1				
	T2 (VMA)	14		7.1				
Face space								
Wilson et al. (2007)	ASD (DSM-IV, CARS)	17	CHI	8.6	Familiar vs. unfamiliar categorization	Not controlled; probably very high depending on when peer was last seen	High; 'real' peer vs. photograph	Group (ASD, DS, T) × feature (full, inner, outer) interaction was not significant ($p>.05$).
	C (CA, VMA)	17		8.7				
	T (CA)	17		8.2				

Table 2 (*Continued*)

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Between-group-statistics
Gastgeb et al. (2009)	ASD (ADOS, ADI-R)	27	CHI	10.8	"Encoding": looking at 14 faces; "pseudo-retrieval": 2AFC which ones looks more familiar?	Medium; 28 s ISI	High; subtle differences	Whole sample: main effect group (ASD, T) was significant: T select the prototype significantly more often than ASD ($p < .05$). Childen: main effect group (ASD, T) was not significant ($p = .12$); Adults: main effect group (ASD, T) was marginally significant ($p = .055$).
	T (CA, VMA, NVMA)	26		10.9				
	ASD (ADOS, ADI-R)	24	ADU	26.9				
	T (CA, VMA, NVMA)	23		28.5				
Gastgeb et al. (2011)	ASD (ADOS, ADI-R)	20	ADU	22.0	"Encoding": looking at 16 faces (×4 sets); "pseudo-retrieval": 2AFC which one looks more familiar?	Medium; 48 s ISI	High; subtle differences	Main effect group (ASD, T) was significant: T select the prototype significantly more often than ASD ($p < .01$).
	T (CA, VMA, NVMA)	20		25.5				
Pellicano et al. (2007)	ASD (DSM-IV, ADI-R)	14	CHI	11.0	Decide which team a person belongs to	No	High; subtle differences	Main effect group (ASD, T): adaptation was significantly reduced in ASD in comparison to T ($p < .05$). ASD: main effect adaptation was significant ($p < .001$). T: main effect adaptation was significant ($p < .001$).
	T (CA, VMA, NVMA)	15		11.1				
Thatcher illusion								
Riby et al. (2009)	ASD (DSM-IV, CARS)	20	CHI/ADO	14.8	Odd-one-out	Negligible; simultaneous presentation	Difficult to determine	Group (ASD, T1, T2, T3) × orientation (upright, inverted, 90°) was not significant ($p = .57$).
	T1 (NVMA)	20		7.9				
	T2 (VMA)	20		6.5				
	T3 (CA)	20		14.9				
Rouse et al. (2004)	ASD (ICD-10)	11	CHI	9.6	Odd-one-out	Negligible; simultaneous presentation	Difficult to determine	Main effect group (ASD, T) was not significant (*p* nr).
	T (CA, NVMA)	15		9.4				
Left-side bias								
Ashwin et al. (2005)	ASD (APA)	16	ADU	26.8	Decide which of two chimeric faces looked more like the original face	Negligible; simultaneous presentation	Difficult to determine	Main effect group (ASD, T) was significant: laterality bias was stronger in T than ASD ($p < .05$).
	T (CA, VMA, NVMA)	16		28.3				

Notes: Wilson et al. (2007): C group consisted of children with developmental delay.
Mean age is in years. Memory demand was assessed on a five-point scale from negligible (simultaneous presentation), over low (0–2 s delay), medium (2 s to 1 min delay), high (1–5 min delay), to very high (>5 min delay). Perceptual demand was assessed on a three-point scale from low (matching identical images), over medium (matching across one dimension, e.g. pose), to high (matching across two dimensions, e.g. pose and lighting). Abbreviations (other than in Appendix A): ADO: adolescents; ADU: adults; CHI: children; ISI: interstimulus interval; *p* nr: *p* not reported; *p* ns: *p* not significant.

aligned with the target half rather than misaligned (Young et al., 1987). One study has investigated the composite effect in participants with ASD (Nishimura et al., 2008; see Table 2), and we discuss another related study (Gauthier et al., 2009). One additional study was excluded from the review (Teunisse and de Gelder, 2003, see Appendix A).

2.3.1. Statistical criteria

The composite effect is evident in a main effect of alignment: it is easier to identify the component faces in a chimeric face when the two parts are misaligned rather than aligned. Furthermore, the composite effect is weaker or absent when faces are inverted, yielding a significant interaction between alignment and orientation. To make the case that participants with ASD do not show the composite effect, authors would at least have to show an interaction between group (ASD, T) and alignment (misaligned, aligned) for upright faces.

Nishimura et al. (2008) found that adult participants with ASD exhibit the composite effect. In an ANOVA on group (ASD, T) and alignment (misaligned, aligned), they found a significant main effect of alignment, thus showing a composite effect across both groups. There was no interaction between group and alignment, suggesting that the composite effect is intact in people with ASD.

In another related study, Gauthier et al. (2009) reported a lack of composite effect in adolescents with ASD. However, this study did not test the classic composite effect but instead a "congruency effect", in which either both halves of the test face were the same or different with respect to the study face ('congruent') or only one half of the test face was the same or different with respect to the study face ('incongruent'). A significant group × alignment × congruency interaction was found, as well as significant within-group results: the interaction between congruency and alignment – the "congruency effect" – was significant in T, but not in ASD. It has been argued that the congruency effect reflects generic visual processing not specific to faces; indeed, the congruency effect has been found with non-face objects and hence should not be used as a marker of face-specific processing (McKone and Robbins, 2011). Thus, although the results of Gauthier et al. (2009) are of some interest, this finding may reflect differences in generic attention, rather than differences in a classic hallmark of face processing.

In summary, the only study of the classic composite effect did not find differences between ASD and typical participants (see also Fig. 1).

2.4. Inner versus outer features

Another basic behavioral property of face identity perception is the inner versus outer features effect. For familiar face recognition, people rely more strongly on inner face regions (e.g. eyes, mouth) than on outer face regions (e.g. hair), whereas for unfamiliar faces, this pattern is reversed (Ellis et al., 1979). Two studies tested the inner versus outer features effect in children with ASD (Rondan et al., 2003; Wilson et al., 2007; see Table 2).

2.4.1. Statistical criteria

To determine which feature (inner or outer) is more important for correct identification, it is necessary to include both familiar and unfamiliar faces so as to test for the expected double-dissociation. Unfortunately, neither study tested both familiar and unfamiliar faces. Here we use a more relaxed criterion: to demonstrate that people with ASD differ in their reliance on particular features (for either familiar or unfamiliar faces), authors need to show a significant interaction between group (ASD, T) and feature (inner, outer). In addition, post hoc tests should be carried out to confirm that the expected pattern is present in the typical group and absent in the ASD group.

Rondan et al. (2003) tested one group of ASD children and two groups of typical children using unfamiliar faces. The task was to match a simultaneously presented full target face at the top of the screen to one of two test faces at the bottom of the screen. The two test faces were either shown only with their inner features or only with their outer features. The authors found a significant group (ASD, T1, T2) × feature (inner, outer) interaction. Within-group testing revealed that in contrast to the CA-matched and VMA-matched typical groups, children with ASD did not perform better in the outer features condition than the inner features condition. Thus, it seems that children with ASD do not rely more strongly on outer features than inner features when recognizing unfamiliar faces.

Wilson et al. (2007) also tested one group of ASD and two groups of typical children but used familiar faces. Two faces were presented side by side, one of which was a staff member at the children's school. Children were required to "touch the one you know from school". Faces were shown either fully, shown with only inner features or shown with only outer features. ASD children, similar to both CA-matched typical children and CA- and NVMA-matched children with developmental delay, were best at discriminating familiar from unfamiliar faces when the face was shown fully, and better when provided with inner features rather than outer features. Thus, participants with ASD show the typical pattern and use inner features more than outer features when recognizing familiar faces. The authors found no significant interaction between group (ASD, T1, T2) and feature (full, inner, outer). They did not perform between-group or within-group testing to compare inner versus outer features directly, but from the raw scores it seems that all three groups had a slight advantage for distinguishing familiar from unfamiliar faces based on inner features in comparison to outer features.

In summary (see also Fig. 1), there is evidence that children with ASD show the expected reliance on inner features when recognizing familiar faces. However, they do not show the expected reliance on outer features when recognizing unfamiliar faces. These findings need to be confirmed, particularly in a design where both familiar and unfamiliar faces are tested.

2.5. Face space—face adaptation aftereffects

Face adaptation aftereffects are a marker for the representation of faces in "face space". In the face-identity aftereffect, prolonged exposure or adaptation to a particular face, (e.g. Jack's face) biases perception of an average face toward the opposite identity (e.g., an "Anti"-Jack's face). Pellicano et al. (2007) tested children with ASD and CA- and NVMA-matched typical children on face-identity aftereffect. While both groups showed significant aftereffects, ASD children showed significantly reduced adaptation. Further, the amount of adaptation correlated with current symptom severity: more severely affected children experienced smaller face-identity aftereffects. Thus, Pellicano et al. (2007) conclude that adaptive face-coding mechanisms are compromised in ASD. However, because the study did not test adaptation of non-face stimuli or any other adaptive mechanisms such as motion adaptation, it is not clear whether adaptive mechanisms are selectively impaired for the category of faces, or more generally for all objects.

Another way to address the organization of face space is by studying how people process face prototypes, as these can be seen as the average face, i.e., the center of face space. Two studies by Gastgeb et al. (2009, 2011) investigated facial prototype formation in ASD, first using line drawings (2009) and then photographs (2011). Both studies found that adults with ASD have difficulties forming facial prototypes. (The first study also looked at children with ASD and found them to perform similarly to their typical peers.) However, impairments in prototype formation in those with autism do not seem to be limited to faces as they also occur with animal-like stimuli (Klinger and Dawson, 2001; although these stimuli also displayed faces and it is difficult to determine on which aspects the prototype formation was based upon).

Thus, although the idea that face space might be differently organized in those with autism is intriguing, it is yet unclear whether adaptive mechanisms in general are compromised.

2.6. The Thatcher illusion

In the Thatcher illusion, the mouth and eyes are cut out, inverted, and pasted back into a face. When upright, the resultant face appears grotesque, but when inverted it appears normal, or nearly so (Thompson, 1980).

2.6.1. Statistical criteria

To show that participants with ASD do not experience the Thatcher illusion or do so to a lesser extent than typical individuals, it is necessary to demonstrate a significant interaction between group (ASD, T) and orientation (upright, inverted), such that the normal inversion effect for "thatcherized" faces is greater in Ts than in ASDs.

Two studies investigated the Thatcher illusion in ASD (Riby et al., 2009; Rouse et al., 2004; see Table 2), and we excluded one other study (Nakahachi et al., 2008). In the two studies, two faces were presented simultaneously. One of the faces was "thatcherized" and participants were asked to "report the one that looks funny or strange" (Riby et al., 2009; Rouse et al., 2004). Both studies showed that ASD participants were sensitive to the Thatcher illusion: ASD participants as well as T were less likely to detect a thatcherized face when the faces were inverted than when they were upright (Riby et al., 2009: group (ASD, T1, T2, T3) × orientation (upright, 90°, inverted) was not significant; Rouse et al., 2004: group × orientation was not significant, but a main effect of inversion was significant). In summary, both studies show that participants with ASD are sensitive to the Thatcher illusion just as typical participants are (see also Fig. 1).

2.7. The left side bias

By splitting a face photograph in the middle and mirroring the left side onto the right side or the right side onto the left side one can create chimeric faces consisting solely of the left side of a face or the right side of a face. It has been found that chimeric faces made out of the *right* side of a face are judged to be more similar to the original face than chimeric faces made out of the *left* side of a face, which has been described as a "left visual field bias for faces" (Gilbert and Bakan, 1973). Ashwin et al. (2005) investigated the left visual field bias for faces in adults with ASD. The laterality scores were significantly higher in CA- and IQ-matched T than in ASD participants. However, laterality scores were overall very low: 0.12 for ASD adults and 0.29 for CA- and IQ-matched T, with 0 standing for no visual field bias and +1 for a strong left visual field bias, and it remains an open question whether the finding of a lesser left visual field bias in ASD adults is due to a different face scanning strategy (as any bias is only elicited when subjects strongly focus on the midline of a face) or actually reflects differences in brain asymmetry.

2.8. Summary of part 2: face markers

Summarizing all studies on face markers it appears that the classic hallmarks of typical face identity recognition such as the face inversion effect and the part-whole effect are evident in people with ASD (see also Fig. 1). Furthermore, those with ASD are sensitive to the Thatcher illusion. In the case of the composite effect, there are two studies with conflicting results, most probably due to the heterogeneity in the experimental designs. For the inner versus outer feature effect and the left side bias more studies are needed to confirm or refute the findings. The same is true with regard to face space, where the key unanswered question is whether the weaker adaptation aftereffects are specific to faces, or reflect a more general adaptation deficit in autism. On balance, the existing literature seems to find no compelling evidence for a qualitative difference between people with and without autism in how they process facial identity.

3. Face identity memory and perception

In the following we review studies looking at quantitative differences in face identity recognition – that is, *how well* one remembers or discriminates facial identities – between those with autism versus typical individuals. Although these studies differ considerably in experimental approach (see also Table 3), we were able to group them into (a) studies on face memory, in which a delay of at least half a minute occurs between encoding and retrieval, (b) studies on simple face identity perception (e.g. discriminating between two people), (c) studies on more fine-grained face perception (e.g. discriminating between two faces in which only the eyes differ), and (d) studies using standardized face recognition tests. The interval between the stimuli in the 'perceptual tasks' was always lower than 2 s. For this part of the review, the most crucial questions are: does empirical evidence exist for a face-specific recognition deficit in ASD, or is a deficit possibly based on domain-general impairments that affect other object recognition as well? If there is a face-specific recognition deficit, is it evident in both memory and perceptual discriminations?

3.1. Face memory

Here we review six studies, in chronological order, containing seven experiments investigating face memory in people with ASD compared to typical individuals. We define 'memory' as entailing a delay between encoding and recognition of at least 30 s, but most of the studies reviewed here have delays of several minutes, so they clearly tap into face memory. All studies found a face memory deficit in people with ASD. One additional study (Krebs et al., 2011) also investigated face memory, but ASD as well as typical participants were at ceiling on the task, thus making their results impossible to interpret (see Appendix A).

de Gelder et al. (1991) showed a group of ASD children and a group of typical children 16 faces one after the other, presented for 5 s each. Their memory was then tested with a 2AFC old–new recognition test. During encoding, faces were presented from a front view while during the memory test they were presented in 3/4 view. The children with ASD were significantly worse in remembering faces across viewpoint than the typical children.

Boucher and Lewis (1992) tested two groups of children and adolescents with ASD in two experiments on face and house memory. In Experiment 1, participants viewed 30 faces one after the other, presented for 7.5 s each. After all faces had been shown, children were presented with a 2AFC old–new recognition test. Boucher and Lewis (1992) found a significant main effect of group (ASD, C1, T2) with ASD participants performing worse than both control groups. In Experiment 2, participants either viewed 28 faces and 28 houses one after the other, presented for 10 s each, or they did a match-to-sample task on these stimuli. Subsequently, they were tested with a 2AFC old–new recognition test. The authors found a significant group (ASD, T) × type of stimulus (face, house) interaction with the ASD participants performing worse than the typical individuals on faces, but not houses. (The results of the

Table 3
Studies on face identity recognition.

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Finding
Face memory								
Boucher and Lewis (1992) (Experiment 1)	ASD (Rutter	10	CHI/ADO	13.2	Encoding: remember 30 faces, retrieval: 2AFC old–new	High; memory delay: 3 min 45 s, high number of stimuli	Difficult to determine	Main effect group (ASD, C, T) was significant: ASD performed worse than either C or T group (both $p < .01$).
	C (CA, NVMA)	10		13.3				
	T (CA)	10		13.2				
Boucher and Lewis (1992) (Experiment 2)	ASD (Rutter)	16	CHI/ADO	13.5	Encoding: either remember 28 faces/buildings or perform match-to-sample on 28 faces/buildings, retrieval: 2AFC old–new	High; timed condition: memory delay: 4 min 40 s, match-to-sample condition: uncontrolled, depended on performance	Difficult to determine	Group (ASD, C) × type of stimulus (face, house) interaction was significant ($p < .01$): ASD performed worse than C for faces ($p < .05$), but not houses (p ns).
	C (CA, VMA)	16		11.2				
Boucher et al. (1998)	ASD1 (Wing's criteria)	7	CHI	9.2	Retrieval: sort images of school staff and school building into familiar and unfamiliar	Not controlled; probably very high, depending on when staff members, school building were last seen	Difficult to determine	All children were 100% correct on the buildings; there was a main effect group (ASD1 + ASD2, C1 + C2) for faces: ASD performed worse than C ($p = .05$).
	ASD2 (Wing's criteria)	12		7.4				
	C1 (CA, VMA)	10		7.9				
	C2 (CA, VMA)	10		7.8				
de Gelder et al. (1991)	ASD (Rutter)	17	CHI/ADO	10.9	Encoding: remember 16 faces, Retrieval: 2AFC old–new	High; memory delay: 1 min 20 s, medium number of stimuli	Medium; across pose	Main effect group (ASD, T) was significant: ASD performed worse than T ($p < .01$).
	T (NVMA)	17		8.5				
Hauck et al. (1998)	ASD (DSM-III-R)	24	CHI	9.6	Encoding: match-to-sample, Retrieval: 2 AFC old–new	Very high; 20 min	Low; identical image	Group (ASD, T) × type of stimulus (face, object) was significant ($p < .05$): ASD performed worse than T on faces, but not objects; and group (ASD, T) × type of task (matching, memory) was significant ($p < .05$): ASD performed worse than T on the memory task, but not the matching task.
	T (VMA)	34		4.7				
Langdell (1978)	ASD1 (Rutter)	10	CHI	9.8	Identify peer	Not controlled; probably very high depending on when peer was last seen	High: 'real' peer vs. photograph	Group (ASD1, C1a, C1b, C1c) × condition (various face masking conditions) interaction was significant ($p < .001$). Group (ASD2, C2a, C2b, C2c) × condition (various face masking conditions) interaction was significant ($p < .001$).
	T1a-c (CA/MA/CA, MA)	10		5.6/9.6/9.8				
	ASD2 (Rutter)	10	ADO	14.1				
	T2a-c (CA/MA/CA, MA)	10		8.1/13.6/13.7				
McPartland et al. (2011)	ASD (DSM-IV, ADI-R, ADOS)	15	ADO	12.0	Encoding: remember 12 faces (2 sets); retrieval: old–new	Medium; 41 s	Low; identical image	Main effect group (ASD, T) was significant: ASD performed worse than T ($p < .01$).
	T (CA, NVMA)	17		13.2				
Simple face perception								
Behrmann et al. (2006)	ASD (ADI-R, ADOS)	14	ADU	34.5	Simultaneous same–different	Negligible; simultaneous presentation	Low; either within gender (more difficult), or across gender (more easy)	Accuracy: neither group (ASD, T) × condition (same, different within gender, different across gender) interaction nor main effect group (ASD, T) were significant (p nr). RT: main effect of group (ASD, T) was significant: ASD performed worse than T ($p < .001$).
	T (CA)	27		Not reported				

Table 3 (*Continued*)

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Finding
Boucher and Lewis (1992) (Experiment 2)	ASD (Rutter)	16	CHI/ADO	13.5	Simultaneous match-to-sample	Negligible; simultaneous presentation	Difficult to determine	Time on task: Group (ASD, C) × type of stimulus (face, house) was significant ($p < .05$) with C taking longer than ASD for houses ($p < .05$), but as long for faces (p ns). Error rate, ASD performed better than C on houses ($p < .05$), but similar on faces (p ns).
	C (CA, VMA)	16		11.2				
Davies et al. (1994) (Experiment 1)	ASD1 (LF, DSM-III-R)	10	CHI/ADO	13.9	Simultaneous matching a third picture to two other pictures	Negligible; simultaneous presentation	High; concept learning	Main effect group (ASD1, C) was not significant (p nr). Main effect group (ASD2, T) was significant: ASD performed worse than T ($p < .005$).
	C (LF: CA, VMA)	10		13.7				
	ASD2 (HF, DSM-III-R)	10		14.4				
	TC (HF: CA, VMA)	10		14.7				
Davies et al. (1994) (Experiment 2)	ASD1 (LF, DSM-III-R)	10	CHI/ADO	13.9	Sorting	Negligible; simultaneous presentation	Medium; across pose, across emotional expression	Main effect group (ASD1, C) was not significant (p nr). Main effect group (ASD2, T) was significant: ASD performed worse than T ($p < .05$).
	C (LF: CA, VMA)	10		13.7				
	ASD2 (HF, DSM-III-R)	10		14.4				
	TC (HF: CA, VMA)	10		14.7				
Deruelle et al. (2004)	ASD (DSM-IV, CARS)	11	CHI	9.3	Simultaneous 2AFC match-to-sample	Negligible; simultaneous presentation	Medium; across pose or emotional expression	Main effect group (ASD, T1, T2) was not significant ($p > .1$).
	T1 (VMA)	11		6.5				
	T2 (CA)	11		9.4				
Gepner et al. (1996) (Task 1)	ASD (DSM-II-R)	7	CHI	11.3	Delayed 4AFC match-to-sample	Medium; 2 s ISI	Low; identical image	Group (ASD, T1) × stimulus type (face, shoe) was not significant ($p > .05$), but a main effect group (ASD, T): ASD performed worse than T1 ($p < .025$). ASD performed worse than T1 on faces ($p < .01$), but not shoes ($p > .05$). Group (ASD, T2, DS) × stimulus type (face, shoe) was not significant ($p > .05$), but a main effect of group (ASD, T2, DS): ASD performed worse than T2 ($p < .025$) but similar to DS ($p > .05$).
	T1 (VMA)	7		5.6				
	T2 (NVMA)	7		5.9				
	DS	7		15.9				
Gepner et al. (1996) (Task 3a)	ASD (DSM-II-R)	7	CHI	11.3	Sorting task	No	Medium; across lip movements	Main effect group (ASD, T1) was not significant (p ns). Main effect group (ASD, T2, DS) was not significant (p nr).
	T1 (VMA)	7		5.6				
	T2 (NVMA)	7		5.9				
	DS	7		15.9				
Gepner et al. (1996) (Task 3b)	ASD (DSM-II-R)	7	CHI	11.3	Sorting task	No	Medium; across emotional expression	Main effect group (ASD, T1) was not significant (p nr). Main effect group (ASD, T2, DS) was not significant (p nr).
	T1 (VMA)	7		5.6				
	T2 (NVMA)	7		5.9				
	DS	7		15.9				
Gepner et al. (1996) (Task 4)	ASD (DSM-II-R)	7	CHI	11.3	Sorting task	No	Medium; across pose	ASD performed worse than T2 on unfamiliar faces ($p < .05$). All other comparisons were not significant.
	T1 (VMA)	7		5.6				
	T2 (NVMA)	7		5.9				
	DS	7		15.9				
Hauck et al. (1998)	ASD (DSM-III-R)	24	CHI	9.6	Simultaneous match-to-sample	Negligible; simultaneous presentation	Medium; across pose or clothes change	Group (ASD, T) × stimulus type (face, object) was not significant.
	T (VMA)	34		4.7				

Table 3 (*Continued*)

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Finding
Ozonoff et al. (1990) (Identity sort)	ASD (CARS)	14	CHI	6.4	Sorting	No	Medium; across emotional expression	Main effect group (ASD, T) was not significant (*p* ns).
	T (MA)	14		3.0				
Ozonoff et al. (1990) (Matching task)	ASD (CARS)	14	CHI	6.4	Simultaneous 4AFC match-to-sample	Negligible; simultaneous presentation	Medium; across pose	Main effect group (ASD, T) was not significant (*p* ns).
	T (MA)	14		3.0				
Riby et al. (2008)	ASD (CARS)	20	CHI/ADO	12.0	Simultaneous 2AFC match-to-sample	Negligible; simultaneous presentation	High; across pose and emotional expression	Main effect group (ASD, C1, C2) was not significant ($p = .73$).
	C1 (VMA)	20		7.5				
	C2 (NVMA)	20		8.9				
Riby et al. (2009)	ASD (DSM-IV, CARS)	20	CHI/ADO	14.8	Simultaneous 2AFC match-to-sample	Negligible; simultaneous presentation	High; across pose and only on face halves	Group (ASD, T1, T2, T3) × face part (upper, lower) was significant ($p < .001$), as was the main effect group (ASD, T1, T2, T3): ASD performed worse than any of the T ($p < .001$).
	T1 (NVMA)	20		7.9				
	T2 (VMA)	20		6.5				
	T3 (CA)	20		14.9				
Robel et al. (2004) (The Identity Match 2 Task)	ASD (DSM-IV)	20	CHI	8.4	Delayed same–different	Low; 0 ISI	Medium; across emotional expression	Main effect group (ASD, T) was significant: ASD performed worse than TC ($p < .008$).
	T (CA)	20		7.9				
Scherf et al. (2008)	ASD (ADI-R, ADOS)	15	CHI	11.0	Delayed 2AFC match-to-sample	Low; 1 s ISI	Low; identical image	Group (ASD, T) × stimulus type (face, Greeble, common object) was significant ($p < .05$): ASD performed worse than T on faces ($p < .005$) and Greebles ($p < .05$), but not on common objects (*p* nr).
	T (CA, VMA, NVMA)	15		12.0				
	ASD (ADI-R, ADOS	15	ADU	32.0				
	T (NVMA)	15		22.0				
Serra et al. (2003)	ASD (DSM-IV)	26	CHI	8.8	Delayed identification of target face in matrix of 4 faces	Low; 500 ms ISI	Low; identical image	Group (ASD, T) × stimulus type (face, pattern) interaction was significant ($p < .002$): ASD performed worse than T on faces ($p < .001$), but not on patterns (*p* nr).
	T (CA)	65		8.7				
Tantam et al. (1989)	ASD (Rutter)	10	CHI	12.1	Odd-one-out	Negligible; simultaneous presentation	Medium; across emotional expression	Main effect group (ASD, T) was significant: ASD performed worse than T ($p < .01$).
	T (CA, NVMA)	10		12.2				
Wallace et al. (2008)	ASD (ICD-10, ADI-R)	26	ADU	32.0	Delayed same–different	Low; 350 ms ISI	Low; identical image	Group (ASD, T) × stimulus type (face, car) interaction was significant ($p < .01$): ASD performed worse than T on faces ($p < .01$), but not on cars (*p* nr).
	T (CA, VMA, NVMA)	26		31.0				
Wilson et al. (2010a)	ASD (DSM-IV, ADI-R, ADOS)	13	CHI	10.1	Delayed 2AFC match-to-sample	Low; 0 ISI	High; across pose and illumination	Main effect group (ASD, T) was significant: ASD perform worse than T ($p < .001$)
	T (CA, NVMA)	14		10.7				
Wilson et al. (2010b)	ASD (DSM-IV, ADI-R, ADOS)	20	CHI	9.7	Delayed 2AFC match-to-sample	Low; 0 ISI	High; across pose and illumination	In comparison to standardized scores from typical children (data not presented), ASD performed relatively worse in face matching in comparison to shoe matching ($p < .03$).
Wolf et al. (2008) (Matching identity across emotional expression)	ASD (DSM-IV, ADI-R, ADOS)	66	CHI	11.6	Simultaneous 3AFC match-to-sample	Negligible; simultaneous presentation	Medium; across emotional expression	Main effect group (ASD, T) was significant: ASD performed worse than T ($p < .001$).
	T (CA, VMA, NVMA)	66		11.6				

Table 3 (*Continued*)

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Finding
Wolf et al. (2008) (Matching identity across masked features)	ASD (DSM-IV, ADI-R, ADOS)	67	CHI	11.7	Simultaneous 3AFC match-to-sample	Negligible; simultaneous presentation	High; across pose, but with either masked mouth, masked eyes, or no mask	Main effect group (ASD, T) was significant: ASD performed worse than T ($p < .001$).
	T (CA, VMA, NVMA)	67		11.7				
Wolf et al. (2008) (Immediate face memory)	ASD (DSM-IV, ADI-R, ADOS)	66	CHI	11.8	Delayed 3AFC match to sample	Low; 0 ISI	Medium; across pose	Main effect of group (ASD, T) was significant: ASD performed worse than T ($p < .001$).
	T (CA, VMA, NVMA)	67		11.8				
Fine-grained face perception								
Faja et al. (2009)	ASD (DSM-IV, ADI-R, ADOS)	39	ADU	24.0	Delayed same–different	Low; 1 s ISI	Low to medium; depending on displacements	Main effect group (ASD, T): ASD performed worse than T ($p < .001$).
	T (CA, VMA, NVMA)	33		24.6				
Joseph et al. (2008)	ASD (DSM-IV, ADI-R, ADOS)	20	CHI/ADO	12.5	Delayed 2AFC match-to-sample	Low; 0 ISI	Low; identical image	Group (ASD, T) × feature (eyes, mouth) interaction, nor main effect group (ASD, T) were significant.
	T (CA, VMA, NVMA)	20		11.9				
Nishimura et al. (2008)	ASD (ADI-R, ADOS)	17	ADU	20.6	Simultaneous same–different	Negligible; simultaneous presentation	Medium; depending on displacements	Main effect of group (ASD, T) was not significant ($p = .55$), nor was any of the interactions involving group significant.
	T (CA, VMA, NVMA)	17		21.6				
Riby et al. (2009)	ASD (DSM-IV, CARS)	20	CHI/ADO	14.8	Simultaneous same–different	Negligible; simultaneous presentation	Medium; depending on displacements	Group (ASD, T1, T2, T3) × condition (featural, configural) interaction was not significant ($p = .43$), but group (ASD, T1, T2, T3) × feature (eyes, mouth) interaction was significant ($p < .001$) as was the main effect group (ASD, T1, T2, T3) ($p < .001$): while all T groups performed better on eyes than mouth, ASD performed better on mouths than eyes.
	T1 (CA)	20		14.9				
	T2 (VMA)	20		6.5				
	T3 (NVMA)	20		7.9				
Rutherford et al. (2007)	ASD (ADI-R, ADOS)	16	ADO/ADU	19.6	Odd-one-out	Negligible; simultaneous presentation	Low to high; depending on displacements	Main effect group (ASD, T): ASD performed worse than T on eye-to-eye spacing ($p < .01$), but not on mouth-to-nose spacing (p ns).
	T (CA, NVMA)	19		24.3				
Wallace et al. (2008)	ASD (ICD-10, ADI-R)	26	ADU	32.0	Delayed same–different	Low; 350 ms ISI	Medium; depending on displacements	Group (ASD, T) × stimulus type (face, house) interaction was significant ($p < .001$): ASD performed worse than T on faces ($p < .0001$), but not on houses (p nr).
	T (CA, VMA, NVMA)	26		31.0				
Wolf et al. (2008) (Face and house dimensions)	ASD (DSM-IV, ADI-R, ADOS)	67	CHI	11.9	Simultaneous same–different	Negligible; simultaneous presentation	Medium; depending on displacements	Group (ASD, T) × feature (eye, mouth) interaction was significant ($p < .001$): ASD performed worse than T on eye-to-eye spacing ($p < .001$), but not on mouth-to-nose spacing (p nr).
	T (CA, VMA, NVMA)	66		11.9				

Notes: Boucher and Lewis (1992): C group consisted of children and adolescents with mental retardation; Riby et al. (2008): C group consisted of children and adolescents with developmental delay.
Mean age is in years. Memory demand was assessed on a five-point scale from negligible (simultaneous presentation), over low (0–2 s delay), medium (2 s to 1 min delay), high (1–5 min delay), to very high (>5 min delay). Perceptual demand was assessed on a three-point scale from low (matching identical images), over medium (matching across one dimension, e.g. pose), to high (matching across two dimensions, e.g. pose and lighting). Abbreviations (other than in Appendix A): ADO: adolescents; ADU: adults; CHI: children; ISI: interstimulus interval; *p* nr: *p* not reported; *p* ns: *p* not significant.

discrimination task are discussed in Section 3.2.) Boucher and Lewis (1992) thus provide first evidence for a memory impairment that is specific to faces in participants with ASD.

Similarly, Hauck et al. (1998) tested children with ASD on an interesting combination of tasks: during the first part of the experiment participants performed a match-to-sample task on faces and houses, then came a 20 min distraction period after which the children were confronted with a surprise old–new memory test on the target items of the match-to-sample task from before. Between-group statistics revealed a significant group (ASD, T) × type of task (matching, memory) interaction with ASD children performing worse than typical children on the memory task, but not on the matching task. Also a group (ASD, T) × type of stimulus (face, house) interaction was significant with ASD children performing worse than typical children on faces, but not houses. In fact, typical children performed best in the face memory task compared to the other three tasks (face perception house memory, house perception), while ASD children had the same scores on all tasks. Furthermore, performance in the face memory task correlated with social ability as measured by several questionnaires in children with ASD. Hauck et al. (1998) thus not only provide evidence for a face-selective recognition deficit in children with ASD, but also that such a deficit might be only evident in memory, not in perception.

Even more evidence for a face-specific memory deficit comes from a study by McPartland et al. (2011) testing adolescents with ASD and T on a non-face object recognition test – the Cambridge Neuropsychological Test Automated Battery Pattern Recognition Subtest – and an analogous task for faces. Participants had to remember 12 patterns/faces, presented for 3 s each and were then confronted with a 2AFC old–new test. Participants with ASD performed worse than T on the face memory test, but not on the pattern memory test.

Two studies tested children's skills at identifying their peers based on photographs. This is arguably a very different task to the ones described before as it is based on familiar face recognition in contrast to unfamiliar face recognition. Also the memory delay is uncontrolled, because it depends on when the participants saw their peers last.

The first study on face recognition in children with ASD was conducted by Langdell (1978). He asked participants to identify their peers from photographs. There were nine different experimental conditions, one full face presented inverted, and one upright, and seven other upright conditions created by obscuring different parts of the face pictures, e.g. showing the eyes only. Langdell (1978) studied two groups with ASD, a younger and an older group of children. In addition, he tested three different typical groups matched to each ASD group (one typical age-matched group, one typical mental-age matched group, and one group with learning disabilities, age- and mental-age matched) for a total of eight subject groups. With so many groups and experimental conditions, analyses become confusing – we report only on the strongest effects (the inversion effect is discussed in Section 2.1.3 above). For both younger and older groups there were significant group (ASD, T1, T2, T3) × experimental condition (7 different masking conditions and inverted presentation) interactions. Older and younger control children did not differ from each other, and all performed better on the upper than the lower half of the face. Younger ASD children performed worse on the upper half, but better on the lower half of the face than all their typical groups. Older ASD children also performed better on the lower half of the face than their typical groups, but they performed as well as their typical groups on the upper half of the face. In summary, Langdell (1978) provided the first evidence that children with ASD might process faces differently than typical children, specifically by showing better performance on lower halves of faces, and worse performance on upper halves.

Boucher et al. (1998) asked four groups of participants (one group of ASD and typical children each from two schools) to sort pictures of school staff and buildings into familiar/unfamiliar 'post boxes'. With regard to the school staff pictures, ASD children performed significantly worse than typical children. However, all children performed at ceiling on the school buildings; thus we are not able to count this data as evidence for a face-specific effect.

In summary, all seven experiments on face memory revealed impairments in the participants with ASD in comparison to T. Most notably, all these studies that also used a within-category object control condition (e.g. distinguishing between different cars or houses) found that the memory deficits were specific to faces. Additionally, of eleven experiments using standardized face recognition tests that are reviewed in a later Section 3.4 and that also employed face memory paradigms, eight also show impairments in the participants with ASD in comparison to T (see also Fig. 2 and Tables 3 and 4). Together, these studies provide strong evidence for a quantitative difference in face-specific memory between those with autism and typical individuals.

Our review of the literature on face memory suggests the hypothesis that people with autism have specific difficulties in remembering faces. Although the next two sections contain studies focusing on face perception – i.e., the perceptual discrimination of faces – we pay particular attention toward the memory demand imposed by the perceptual tasks because it is clear from our above review that face memory is impaired in people with ASD.

3.2. Simple face perception

The twenty-four very heterogeneous experiments testing simple face perception, that is, the perceptual discrimination between two faces with little or no delay between their presentations, in people with ASD, not surprisingly, have very mixed results. Half of the experiments ($N = 11$) find that people with ASD perform worse than typical individuals, while the other half ($N = 11$) does not find any differences (two studies have mixed results and two additional studies using standardized face recognition tests reviewed in Section 3.4. also find impairments; see Fig. 2 and also Table 3). However, based on our prior review of the studies on face memory (see the section above), we followed the hypothesis that tasks with a higher memory demand might be especially challenging for people with ASD. Most interestingly, we found a clear dissociation between memory demand and performance outcome in the studies on simple face perception. Most studies with no memory demand, that is, with simultaneous presentation of the items to be discriminated, do not find differences between those with autism and typical individuals. In contrast, most studies with memory demand, that is, with consecutive presentation of the items to be discriminated, do find differences between those with autism and typical individuals (see Fig. 2). Note that in these latter studies the memory demand is still low and considerably smaller than in the previously described studies on face memory.

One study (Gepner et al., 1996) is particularly important, because it investigated both, perceptual and memory performance, using sorting tasks with no memory demand, and a sequential 4AFC match-to-sample task with memory demand. In accordance with our hypothesis, these authors found no differences between children with ASD and typical children on the tasks with no memory demand, but children with ASD performed worse than typical children on the sequential match-to-sample task. Specifically, Gepner et al. (1996) administered four tasks on simple face identity perception: a delayed 4AFC match-to-sample task with both faces and shoes (Task 1), and three sorting tasks, in which photographs of people were sorted based on identity either across lip movements (Task 3a), emotional expression (Task 3b) or pose (Task 4). The latter sorting task contained both familiar and unfamiliar

Table 4
Studies using standardized face recognition tests.

Study	Diagnosis	*N*	Age group	Mean age	Task	Memory demand	Perceptual demand	Between-group-statistics
Annaz et al. (2009)	ASD (DSM-IV, ADOS) [LF, CARS]	16	CHI	8.5	Benton	Negligible; simultaneous presentation	High; across pose and illumination	Main effect group (ASD [LF], T) was significant: ASD [LF] performed worse than T ($p < .001$). Main effect of group (ASD [HF], T) was significant: ASD [HF] performed worse than T ($p = .014$). Main effect of group (DS, T) was significant: DS performed worse than T ($p < .001$). Main effect of group (WS, T) was not significant ($p > .5$).
	ASD (DSM-IV, ADOS) [HF, CARS]	17		8.4				
	DS (∼ CA)	15		9.5				
	WS (CA)	15		8.8				
	T (CA)	18		8.4				
Blair et al. (2002)	ASD (DSM-IV)	12	ADU	29.9	Warrington	High to very high; 2.5 min up to 30 min	Low; identical image	Main effect group (ASD, C) was significant: ASD performed worse than T ($p < .05$).
	C (CA, VMA)	12		31.1				
Faja et al. (2009)	ASD (DSM-IV, ADI-R, ADOS)	39	ADU	24.0	Wechsler	High to very high; 2.5 min up to 30 min	Low; identical image	Main effect of group (ASD, T) was significant: ASD performed worse than T ($p < .01$).
	T (CA, VMA, NVMA)	33		24.6				
de Gelder et al. (1991)	ASD (Rutter)	17	CHI/ADO	10.9	Kaufmann	Medium; several seconds	High; across pose and emotional expression	Main effect of group (ASD, T) was significant: ASD performed worse than T ($p < .02$).
	T (NVMA)	17		8.5				
Kirchner et al. (2011)	ASD (DSM-IV, ADI-R)	20	ADU	31.9	CFMT	Medium to very high: 3 s up to several minutes	Medium to high; across, across pose and illumination; across pose, illumination and with added noise	Main effect of group (ASD, T) was significant: ASD performed worse than T ($p < .01$).
	C (CA, VMA)	21		31.8				
Klin et al. (1999)	ASD	34	CHI	7.4	Kaufmann	Medium; several seconds	High; across pose and emotional expression	Main effect group (ASD, C) was significant ($p < .0001$). Main effect group (PDD-NOS, C) was not significant (*p* ns).
	PDD-NOS (CA, NVMA)	34		6.6				
	C (CA, NVMA)	34		6.3				
McPartland et al. (2011)	ASD (DSM-IV, ADI-R, ADOS)	15	ADO	12.0	Children's Memory Scale Faces Subtest	Medium; 32 s	Low; identical image	Main effect of group (ASD, T) was not significant ($p = .08$).
	T (CA, NVMA)	17		13.2				
O'Hearn et al. (2010)	ASD1 (DSM-IV, ADI-R, ADOS)	8	CHI	11.6	CFMT	Medium to very high: 3 s up to several minutes	Medium to high; across, across pose and illumination; across pose, illumination and with added noise	Main effect group (ASD1, T1) was not significant (*p* ns). Main effect group (ASD2, T2) was not significant (*p* ns). Main effect group (ASD3, T3) was significant ($p < .001$).
	T1 (CA, VMA, NVMA)	8		11.6				
	ASD2 (DSM-IV, ADI-R, ADOS)	12	ADO	15.4				
	T2 (CA, VMA, NVMA)	12		15.4				
	ASD3 (DSM-IV, ADI-R, ADOS)	14	ADU	23.1				
	T3 (CA, VMA, NVMA)	14		22.7				
Wallace et al. (2008)	ASD (ICD-10, ADI-R)	26	ADU	32.0	Benton	Negligible; simultaneous presentation	High; across pose and illumination	Main effect of group (ASD, T) was significant ($p < .001$).
	T (CA, VMA, NVMA)	26		31.0				
Williams et al. (2005)	ASD (DSM-IV, ADI-R, ADOS)	29	ADU	28.7	Wechsler	High to very high; 2.5 min up to 30 min	Low; identical image	Main effect group (ASD, T) was significant ($p < .001$).
	T (CA, VMA, NVMA)	34		26.5				

Notes: Blair et al. (2002): C group consisted of participants with learning disabilities; Klin et al. (1999): diagnostic criteria for ASD unclear, C group is non-pdd, mainly mental retardation and language disorders; Williams et al. (2005): statistics performed on average of part 1 and part 2 of WMS-III.
Mean age is in years. Memory demand was assessed on a five-point scale from negligible (simultaneous presentation), over low (0–2 s delay), medium (2 s to 1 min delay), high (1–5 min delay), to very high (>5 min delay). Perceptual demand was assessed on a three-point scale from low (matching identical images), over medium (matching across one dimension, e.g. pose), to high (matching across two dimensions, e.g. pose and lighting). Abbreviations (other than in Appendix A): ADO: adolescents; ADU: adults; CHI: children; ISI: interstimulus interval; *p* nr: *p* not reported; *p* ns: *p* not significant.

faces. Seven children with ASD were tested, along with three control groups, each containing 7 children (a group of children with Down syndrome and two typical groups, one matched on VMA, the other matched on NVMA). The authors compared the children with ASD to both the NVMA-matched T and the children with Down syndrome in a first analysis and to the VMA-matched T in a second analysis. For the delayed match-to-sample task children with autism performed worse than the typical children on faces, but not on shoes reflected in simple effect analyses and a main effect of group in both analyses (although the interactions between group (ASD, T1) and (ASD, T2, DS) and stimulus type (face, shoe) did not reach significance). The authors did not find differences between groups for the sorting tasks based on identity across lip movements or emotional expression. For the sorting task based on identity across pose the only difference between groups was that the ASD children performed worse than the NVMA-matched typical children on unfamiliar faces, all other comparisons revealed no differences between groups. Thus, this study found little difficulties in ASD in a simultaneous sorting task, but a deficit in a task with a short memory delay.

In what follows we review all experiments on simple face perception to evaluate the hypothesis that face deficits are found whenever sample and test stimuli are not present at the same time on the screen (even if the test stimulus is presented immediately at the offset of the sample stimulus with zero delay), but no deficits are found when stimuli are presented simultaneously. First we consider studies that fit this hypothesis: (1) experiments that contained no memory demand and also did not find differences between their participants with ASD and typical individuals, and (2) experiments that imposed a memory demand and in turn found differences between their participants with and without ASD. Finally, we describe five experiments that are not consistent with our hypothesis.

3.2.1. No memory demand—no difference in simple face perception between ASD and typical individuals

First, we briefly describe eight experiments that used simultaneously presented stimuli and thus no memory demand and hence found no differences in simple face perception between ASD and typical participants.

Ozonoff et al. (1990) tested a group of children with ASD on a face identity sorting and a face-identity-matching task with simultaneous presentation. ASD children performed no different on these tasks compared to VMA-matched T. (They then retested the same ASD children against a group of NVMA-matched T and here the children with ASD performed worse – however, we disregard this finding because of the retesting confound.)

Boucher and Lewis (1992) gave participants a simultaneous match-to-sample task as part of the encoding phase in a face memory task (we report on the data of the face memory task in Section 3.1). They did not find differences between their groups of children and adolescents with and without ASD on the matching task.

Hauck et al. (1998) gave children with ASD a match-to-sample task, in which participants were asked to match one stimulus at the top of the page (a face or an object) with one of four stimuli at the bottom of the page. The matching was across a change in either pose or clothing for the faces or across location or a change in state (e.g. complete doughnut, broken doughnut) for the objects. This task, which was part of a larger study (see Section 3.1), concluded that ASD children performed no different from typical children in matching faces or objects.

Deruelle et al. (2004) asked their participants to perform a simultaneous matching task. Although participants with ASD performed worse than typical participants when matching emotional expressions, gaze direction, gender and lip-reading, they performed just as well as individuals when matching facial identity.

Behrmann et al. (2006) used a simultaneous same–different task in which pairs of faces, objects, and Greebles were presented side by side. Stimulus pairs were displayed until response and accuracy was very high – in fact, there was no significant difference in accuracy between participants with and without ASD. Thus, the main dependent measure was reaction time. RTs were much higher for individuals with ASD than for typical participants for all three stimulus categories, and this effect was larger for the more difficult compared to easier discriminations for each stimulus type. However, these data do not provide evidence for a specific deficit in face perception in ASD although the authors argue that these tests show a face recognition deficit: the higher RTs in all three tasks could reflect either a general impairment in shape perception, or a general increase in RT for participants with ASD.

Riby et al. (2008) used a simultaneous 2AFC match-to-sample task either with full faces or with internal features only. ASD participants performed similarly to control participants with developmental delay; there was no significant effect of group. (However, all children performed at chance on the task with internal features only – thus this task was clearly too difficult for the children.)

In summary, studies that used simultaneous presentation of sample and test stimuli in simple face perception tests have not found differences in performance between their participants with ASD and typical individuals.

3.2.2. Memory demand—differences in simple face perception between ASD and typical individuals

Next we briefly describe seven experiments that imposed a small delay between sample and test faces and found face discrimination deficits in participants with ASD.

Serra et al. (2003) tested children with PDD-NOS on a paradigm in which they were asked to say if a target face was present or not in a subsequently presented array of four test faces. They also used a similar task with patterns. Between-group statistics revealed that the children with PDD-NOS performed significantly worse than T on the face task than on the pattern task.

Robel et al. (2004) used two versions of a same–different matching task. When discriminating identical images of a face from faces of two different people, children with ASD performed as well as their CA-matched T group (the main effect of group was not significant) – but both groups were at ceiling, thus we are disregarding this finding. However, when discriminating the same person but with a different emotional expression from two different faces, ASD children performed worse than typical children (and there were no ceiling/floor effects).

Using a sequential 2AFC match-to-sample task, Scherf et al. (2008) tested identity discrimination either within gender (‘individual or exemplar trials’) or across gender (‘gender or subordinate trials’) in both a group of children and a group of adults with ASD and their respective typical groups. Furthermore, they tested human faces as well as Greebles and common objects as stimuli. They found that the participants with ASD performed worse on faces than on Greebles or common objects, supported by a significant interaction of group (ASD, TD) × stimulus type (faces, Greebles, common objects). When statistical analyses were performed on the data from individual stimulus types, participants with ASD also performed worse than T in Greeble recognition (but performed similarly on common objects). Still, effect sizes clearly show that face recognition is disproportionately affected in ASD.

Wallace et al. (2008) had their participants perform simple identity discriminations with both faces and cars. ASD participants performed worse than T on faces, but not on cars as shown in a significant interaction between group (ASD, T) and stimulus type (face, car). Thus, Wallace et al. present additional evidence for a selective impairment of face recognition.

Wolf et al. (2008) tested 'immediate memory for faces': participants saw a target face in frontal view for 1 s and then immediately afterward three sample faces in 3/4 views; they were asked to match the target to one of the three samples. They found a significant main effect of group (ASD, T) with the 66 children with ASD performing worse than the 67 CA- and IQ-matched typical children. Interestingly, Wolf et al. (2008) also tested a subgroup on a similar paradigm with cars – here they did not find differences between groups, suggesting that their findings are face-specific.

In two recent studies, Wilson et al. (2010a,b) investigated two groups of children with ASD on a sequential matching task with faces and shoes. In the first study, children with ASD were impaired in both face matching and shoe matching in comparison to CA-matched individuals (Wilson et al., 2010a). When matching faces, the ASD group performed close to floor (but had above chance performance when matching shoes). In the second study (Wilson et al., 2010b), which was mainly designed to study parents of children with ASD, the children with ASD were impaired in face matching, but not in shoe matching relative to scores from typical individuals taken from the previous study.

Thus, many experiments (19 out of 24) fit our hypothesis that if a simple face perception task has no memory demand it will show no deficits in people with ASD, but if it has even a small memory demand it will show deficits in people with ASD. Five studies described next do not fit neatly into this scheme. All of them used simultaneously presented stimuli and find differences between their ASD and typical participants. Here we review them in more detail.

In the first study, Tantam et al. (1989) asked children with ASD to find "the odd-one" among four pictures, either showing three people with the same and one with a different emotional expression (emotion task), or three instances of the same person each with a different emotional expression and one different person (identity task). They found a main effect of group, with ASD children performing worse than typical children, and a main effect of task, with the identity task being easier than the emotion task. The interaction group × task (emotion, identity) was not significant. Unfortunately, no simple tests were run comparing the performance of the ASD and T on the identity task only (despite the interaction in the ANOVA not being significant, this would still be worth knowing particularly because of the main effect of task). Thus, it is not entirely clear if ASD participants with ASD performed worse than typical individuals on the face identity task. Reported means suggest that the ASD children performed much worse on the emotion task than the typical children, but only slightly worse on the identity task.

In the second study, Davies et al. (1994) administered various face tasks, always comparing a group of low-functioning ASD children to a group of low-functioning children with learning disabilities, and a group of high-functioning ASD children to a group of typical children. In Experiment 1, participants had to pick out a third picture to match two other simultaneously presented face pictures based on a common feature; either an incidental feature (e.g. glasses), the emotional expression or the person's identity. The authors found no differences between the low-functioning groups, but a main effect of group for the high-functioning groups, with typical children outperforming the ASD children. Experiment 2 consisted of four matching tasks, in which one face stimulus had to be matched to one of four simultaneously presented stimuli. Face matching was either based on identity across pose, identity across emotional expression, or emotional expression across identity. In addition, they also performed a pattern-matching task across a change in location. As in Experiment 1, Davies et al. (1994) found no performance differences between their low-functioning groups, but a significant difference between their high-functioning groups. The authors did not test group differences for each task separately, thus their result can only be interpreted as a very general deficit in (mainly face) perception in high-functioning ASD children in comparison to their typical peers, but not in low-functioning ASD children in comparison to a group of low-functioning children with learning disabilities. Thus, although stimuli were presented simultaneously, authors still found differences between high-functioning (but not low-functioning) ASD and typical individuals.

Riby et al. (2009) had participants perform a simultaneous 2AFC match-to-sample task on either upper or lower halves of faces. A significant interaction between group (ASD, T1, T2, T3) and condition (upper half, lower half) as well as a significant main effect of group showed that those with autism performed worse than any of the CA-, or VMA-, or NVMA-matched control groups. Additionally, while all typical groups performed better on the upper than the lower half of the face, there was no difference in performance for the upper versus the lower half of the face in the ASD group.

Wolf et al. (2008) had participants perform two simultaneous 3AFC match-to-sample tasks: matching identity across emotional expression, or matching identity across pose either without mask or with either the mouth or the eyes of the test faces masked. In both matching tasks ASD children performed significantly worse than CA- and VMA- and NVMA-matched typical children.

Thus, of the five experiments that do not fit neatly with our hypothesis (that if a simple face perception task has no memory demand it will show no deficits in people with ASD, but if it has even a small memory demand it will show deficits in people with ASD), the exact results with regard to the face identity task of one study are unclear (Tantam et al., 1989) and one study (with two experiments) finds differences only in high-functioning, but not in low-functioning children (Davies et al., 1994). The three other studies form an interesting class as they all test simple face perception – the discrimination between two individuals – with simultaneous presentations but with high perceptual demands: discriminations had to be made on face halves and across pose (Riby et al., 2009) or across emotional expression with 3 alternatives (Wolf et al., 2008), or across pose and with masked features and with 3 alternatives (Wolf et al., 2008). We speculate that these three studies thus tapped into more fine-grained perceptual discriminations. As we will discuss below, subjects with ASD are impaired in fine-grained perceptual discrimination on faces even if simultaneous presentation of the stimulus material does not impose a memory demand on the task. Thus, although these studies apparently do not fit our hypothesis in its first definition they provide first hints to a refinement of this hypothesis that we will discuss next.

Summarizing the studies on simple face perception, the conflicting results from the heterogeneous set of studies can be resolved in part by assessing the memory demand imposed by the task (see Fig. 2). Even a slightly higher memory demand – presenting stimuli sequentially rather than simultaneously – seems to account for the impairments in face discrimination reviewed here. Furthermore, five experiments also tested other object categories. Of these five studies, four found the recognition deficits to be specific for faces. We thus conclude that the evidence strongly favors a face-specific recognition impairment as long as the task entails some memory demand.

3.3. Fine-grained face perception

Fine-grained face perception task require discriminations between two (or more) faces with only subtle differences between their features (rather than two entirely different identities). Here we find that the relationship between memory demand and study outcome is influenced by a third factor, namely the feature to be discriminated (eyes only, mouth only, or either eyes or mouth combined; see also Fig. 2). Next, we summarize the seven studies testing participants' discrimination ability of "featural" and "configural" changes: "Featural changes" are changes in features per se, i.e., changing the identity or the size of the eyes only, but keeping the rest of the face, while "configural changes" are changes in the spacing between features, e.g. moving the eyes further away from each other. The latter is also commonly called "sensitivity to spacing".

Three studies (Faja et al., 2009; Nishimura et al., 2008; Wallace et al., 2008) used stimuli that contained both changes in eyes and mouths, giving the participants the opportunity to base their discriminations on either of these two features (without of course being able to control which ones they actually chose). Just as we saw for the studies on simple face perception, the study with memory demand finds worse performance in those with ASD in comparison to T (Faja et al., 2009), while the study without memory demand (Nishimura et al., 2008) does not.

Specifically, using paradigms in which stimuli were presented simultaneously and for unlimited time, Nishimura et al. (2008) did not find performance differences between their adult ASD participants and the CA- and IQ-matched typical participants (the main effect of group was not significant): neither in a combined analysis of the three tasks testing sensitivity to spacing, featural changes or changes in face contour, nor in a separate analysis of only the task testing sensitivity to spacing. ASD participants were however overall slower to respond than typical individuals. In contrast to Nishimura et al., Wallace et al. (2008) used a sequential same–different task with brief presentation times and did find differences between ASD and typical participants. They investigated feature discrimination and sensitivity to spacing together in face (e.g. eye-to-eye distance) as well as in house stimuli (e.g. window-to-window distance). ASD participants performed worse than T on faces, but not on houses as reflected in a significant interaction between group (ASD, T) and stimulus type (face, house). Faja et al. (2009) also used a sequential matching task, and found that their adult ASD participants performed less accurately than CA- and IQ-matched typical participants (main effect of group was significant).

The other four studies investigating fine-grained face perception contrasted discrimination abilities on eyes versus mouth. Unanimously, they all find that participants with ASD perform worse than T on eye discrimination, but perform similarly on mouth discrimination. This finding held independent of whether these studies addressed featural or configural changes, separately or combined. And importantly, this finding also held independent of the experimental approach. Thus, these studies strongly suggest an eye-selective recognition deficit in ASD.

Investigating sensitivity to spacing, Rutherford et al. (2007) tested adolescent and adult ASD participants and CA- and IQ-matched typical participants in a very thoroughly conducted experiment. Researchers not only varied either eye-to-eye or mouth-to-nose distance (each at five different displacement levels), they also tested the effect of inversion (upright, inverted) and display duration (1 s, 2 s and 4 s). Because many ASD participants did not even reach a threshold performance level (set at 67% correct) for more difficult displacement levels, the authors performed a between-group analysis only at the greatest displacement level (10 px): they found impairment in sensitivity to eye spacing changes in their participants with ASD in comparison to T (main effect group). In contrast, participants with ASD were as sensitive as T for spacing changes involving the mouth. Most interestingly, this study also found evidence for two distinct subgroups in their ASD participants: those who performed similar to the typical participants on the eye spacing task, and those who performed worse. The two groups did not differ in their performance on the mouth spacing task. This finding suggests another potential underlying cause for the heterogeneity seen in ASD, namely, the existence of subgroups.

Looking at featural changes alone, Joseph et al. (2008) used a sequential 2AFC match-to-sample task in which either the eyes or the mouth was varied. They did not find any differences between their children and adolescents with ASD and CA- and IQ-matched T (although the authors report a main effect of group, with ASD participants performing marginally [$p < .1$] worse than typical participants when the eyes were varied, but not when the mouths changed).

Investigating both featural and configural changes together, Riby et al. (2009) found a main effect of group, with the ASD group performing significantly worse than any of their either CA-, VMA-, or NVMA-matched control groups as well as a significant interaction between group (ASD, T1, T2, T3) and feature (eyes, mouth). Within-group analyses revealed that while all typical groups performed better on eyes than on mouth trials, ASD participants performed better on the mouth than the eye trials.

Similarly, Wolf et al. (2008) studied both configural (eye-to-eye or mouth-to-nose distance) and featural (size of eyes or mouth) discrimination in a large group of participants. They found a group (ASD, T) $\times$ feature (eye, mouth) interaction showing that their the 67 ASD participants were less sensitive to eye-to-eye distance and eye size, but just as sensitive to mouth-to-nose distance and mouth size as their 66 CA- and IQ-matched typical participants. Interestingly, Wolf et al. (2008) also tested sensitivity to spacing for house stimuli: here the ASD participants outperformed the typical participants in sensitivity for window-to-window distances (main effect of group); however, these results are difficult to interpret as the typical children performed at chance.

In summary, studies on fine-grained face perception suggest eye-specific recognition deficit in ASD (independent of memory demand; see also Fig. 2 and Table 3). When discrimination is not dependent on the eyes, participants with ASD show a recognition deficit only if the memory demand is high.

3.4. Standardized face recognition tests

Having found that people with ASD show a deficit in face memory tasks in comparison to typical participants we now ask the question: "Are individuals with ASD as impaired at face recognition as prosopagnosic patients are or are their deficits significant but subtle?" In acquired prosopagnosia, damage to the posterior right hemisphere regions leads to severe and selective deficits in face identity recognition. To compare ASD participants' performance in face recognition with scores obtained from a large typical population as well as with scores obtained from patients with prosopagnosia, researchers make use of standardized face recognition tests. Furthermore, the use of standardized face recognition tests allows for the comparison of face identity recognition performance across laboratory sites.

All but one study employing standardized face recognition tests found deficits in face processing in ASD participants compared to typical participants (as evidenced by between-group statistical differences; see Table 4): Benton facial recognition test (Annaz et al., 2009; Wallace et al., 2008), Warrington recognition memory for faces (Blair et al., 2002), Kaufman face recognition test for children (de Gelder et al., 1991; Klin et al., 1999), face subtest of the Wechsler Memory Scale (Faja et al., 2009; Williams et al., 2005) and Cambridge Face Memory Test (Kirchner et al., 2011; O'Hearn et al., 2010). McPartland et al. (2011) found their adolescents with ASD to be marginally worse ($p = .08$) on the Children's Memory Scale Faces Subtest in comparison to CA- and VMA-matched T. Two studies found only some, but not all, of their participants with ASD, were compromised on these tasks (Klin et al., 1999; O'Hearn et al., 2010, see below).

One of the most interesting findings among these studies is the result from O'Hearn et al. (2010). Unlike many of the other face recognition tests that have been criticized for a lack of validity (e.g. the Benton facial recognition test can be performed using a pixel-matching strategy rather than a face-identity-matching strategy, Duchaine and Nakayama, 2004) the test administered in this study is the Cambridge Face Memory Test (CFMT; Duchaine and Nakayama, 2006), which requires matching faces across changes in pose, illumination and with added noise. While children and adolescents with ASD were unimpaired in the CFMT in comparison to their typical peers, adults with ASD performed significantly worse than typical adults. In fact, adult participants with ASD performed no differently from patients with prosopagnosia! (Caveat: also typical children seem to perform at the level of patients with prosopagnosia, which raises the question whether the CFMT might underestimate typical face performance in children.) The authors interpret their findings as evidence that the typical developmental improvement in face recognition during childhood and adolescents is disrupted in ASD. (Note, that the study is not a longitudinal, but a cross-sectional study and thus might be confounded by cohort effects. For example, it is conceivable that the ASD children and adolescents nowadays received more extensive training than the adults (when they were children/adolescents) to compensate for their deficits and thus perform on an equal level with their peers). The findings by O'Hearn et al. were partly confirmed in another study by Kirchner et al. (2011), who also found their adults participants with ASD to perform worse on the CFMT compared with CA- and VMA-matched typical individuals. Wilson et al. (2010b) investigated parents of children with ASD using the CFMT. Fathers' scores but not mothers' scores, were significantly lower than average. Importantly, the scores were not so low as to support a diagnosis of prosopagnosia; only subtle face perception impairments were found in parents of children with ASD consistent with the Broader Autism Phenotype (Bailey et al., 1998; Bolton et al., 1994). Unfortunately, in all the studies cited above, no equivalent task involving objects was administered, thus it remains an open question whether these deficits are specific to faces, or affect shape perception more generally.

Of all studies using standardized face recognition tests, only one compared performance on a face recognition test with other object recognition tests. Blair et al. (2002) showed that ASD participants performed below the 50th percentile on the Warrington recognition memory for faces and these scores were significantly worse in comparison to VMA-matched individuals (the CA-matched individuals did not perform the memory task). Furthermore, although ASD participants performed similar to CA-matched individuals in leaf and building memory (but performed better than VMA-matched individuals), they performed worse in cat, horses, and motorbike memory in comparison to CA-matched individuals (but performed similarly to VMA-matched individuals in these categories). These results neither suggest a recognition deficit that is face-specific deficit nor a very general object recognition deficit.

Consistent with the idea that object memory may be generally impaired in ASD, Williams et al. (2005) as well as O'Hearn et al. (2010) used a spatial span memory test and found their ASD participants to be impaired in comparison to typical individuals. In contrast, Klin et al. (1999) tested their participants on a spatial memory test and found no differences between groups.

Thus, some domain-general memory deficit might contribute to the face recognition deficits found with the standardized tests. This is particularly the case as most of the standardized face recognition tests (with the exception of the Benton facial recognition test) impose a substantial memory demand.

In conclusion, almost all studies using standardized face recognition tests found ASD participants to be impaired in face recognition (see also Fig. 2). Notably, the majority of the standardized tests imposes a strong memory demand and at the same time shows face recognition deficits in ASD participants, thus strengthening our finding of face memory impairments in people with ASD. The exception to this rule is the Benton facial recognition test, in which images are presented simultaneously and on which ASD participants still were impaired. Particularly the studies using the Cambridge Face Memory Test seem to confirm the hypothesis that people with ASD are as impaired in face memory as patients with prosopagnosia. Because of its high validity and reliability, it would be most informative to test people with and without ASD on the CFMT and a similar test with object stimuli to further rule out the alternative explanation of a more generalized memory deficit.

3.5. Summary of part 3: face identity memory and perception

This review of studies on face identity memory and perception (see also Fig. 2) shows that face memory is impaired in participants with ASD compared to typical individuals. This finding is evident in the tasks that investigated face memory directly as well as in the standardized facial recognition tasks, of which most rely heavily on memory. The impact of a memory demand on face recognition also emerges in the studies on simple face perception that made use of sequential matching tasks, where people with autism also showed worse performance than typical individuals. In contrast, most studies that used simultaneous presentation of stimuli did not reveal differences in performance between those with and without autism. Thus, even if the memory demand is only minimal, it seems to have a sufficient impact on autistic participants' performance in face recognition. The studies on fine-grained face perception extend the findings on face memory and simple face perception by highlighting the fact that people with ASD seem to have particular difficulties discriminating eyes (but perform similarly to typical individuals on the mouth). This pattern is evident even when there is no memory demand.

Having established a deficit in face identity recognition in ASD, a crucial question is whether this deficit is specific to faces. If not, these findings would suggest a domain-general impairment in e.g. memory or attention. Fifteen experiments (in 12 articles) investigated face identity recognition and compared it to the recognition of other visual objects such as patterns (Davies et al., 1994; McPartland et al., 2011; Serra et al., 2003), cars (Wallace et al., 2008; Wolf et al., 2008), buildings (Boucher and Lewis, 1992; Hauck et al., 1998; Wallace et al., 2008; Wolf et al., 2008), Greebles (Scherf et al., 2008), common objects (Scherf et al., 2008) and shoes (Gepner et al., 1996; Wilson et al., 2010a,b). Blair et al. (2002) tested face memory against memory for cats, horses, motorbikes, leaves and buildings. All studies found face recognition deficits in their participants with ASD, and further 11 of these 15 studies found that their participants with ASD performed worse in face recognition, but not in the recognition of other object types, thus showing face-specific deficits. In summary, a majority of studies show face-specific recognition deficits in people with ASD.

4. Grand summary

At first glance, the literature reviewed here presents a frustratingly inconsistent picture: of ninety experiments investigating face identity recognition in ASD, about half found that people with ASD perform worse than typical individuals ($N = 46$) and about a half found them to perform the same ($N = 44$). Upon closer inspection, however, we find a number of systematic patterns in the literature.

First, our review finds that face identity perception is *qualitatively* similar between people with ASD and typical participants. That is, we do not find compelling evidence for differences in *how* people with ASD process facial identity. Indeed, the available evidence indicates that three of the classic hallmarks of face-specific processing – the inversion effect, the part-whole effect, and the composite effect – are all present in people with ASD. Most of the other face markers, such as face space and the Thatcher effect, are only investigated in a few studies, so it will be important in future research to determine which of these other markers of face processing are present in ASD.

Second, turning to *quantitative* differences in face perception (how well people with ASD process faces), we find that two different aspects of face identity processing are specifically impaired in ASD. In one, people with ASD perform less accurately in face *memory* tasks, even if the memory demand is only minimal (e.g. as in consecutive versus simultaneous presentation of stimuli). Current evidence further suggests that this memory impairment is quite selective for faces per se, and does not reflect a broader memory problem for any visual material. Further, this deficit is much stronger when a delay (even a very short delay) intervenes between the two presentations of a stimulus, indicating that this deficit reflects more of a problem with in face memory than in face perception. The idea that face memory may be a distinct mental function, dissociable from both visual memory for non-face objects, and from perceptual processing of faces, dovetails with prior work on selective deficits in face memory (or "prosopamnesia"). Specifically, two neurological patients have been reported (Tippett et al., 2000; Williams et al., 2007), who are impaired at remembering faces, but not other object categories, while at the same time showing no deficit in perceptual discrimination of faces. Additionally, Lawrence et al. (2003) show that women with Turner syndrome have impairments in face memory, but normal face perception, perhaps because of reduced emotional significance of faces – a hypothesis that is interesting to consider for the case of autism as well given that deficits in emotion recognition are part of the autism symptomatology (Harms et al., 2010). Thus, the evidence from autism adds to this prior neuropsychological evidence that the face memory system may be segregated from other memory systems, as well as from a face perception system.

The results of this review naturally raise the question of what brain differences in ASD underlie the face identity recognition deficits reviewed here. A full treatment of this controversial question is beyond the scope of the current review, but we note here that two key claims have been made: first, early reports of a hypoactivation of the fusiform face area in individuals with ASD (Schultz et al., 2000; Pierce et al., 2001) do not replicate consistently, e.g. when face scanning strategies are controlled (Dalton et al., 2005; Hadjikhani et al., 2004, 2006; Perlman et al., 2011) or individual data analyses are performed (Scherf et al., 2010). Second, the connectivity of face processing regions may be abnormal in ASD (Kleinhans et al., 2008; Koshino et al., 2008; Thomas et al., 2011), and selective disruptions of the connections of face-processing regions with memory structures could potentially underlie the face memory deficits reviewed here.

Beyond face memory deficits, we also found evidence that people with ASD have specific deficits discriminating eyes, even if the task poses no memory demand. It might be that a deficit in eye-specific discrimination can lead to low performance in face memory, as the eye region is one part of the face that differs notably between individuals. Thus, a key question for future research is whether the face memory deficit results from a specific deficit in eye perception, and/or whether it persists when the key differences between faces are found in other parts of the face beyond the eyes.

More generally, it will be important in future to understand the causal role of face processing deficits in the etiology of the rest of the autism phenotype. Do fundamental, early-developing deficits in face memory lead to other aspects of the autism phenotype, such as impaired social attention and social cognition? Or does the causality run the other way around, with deficits in face processing resulting from earlier-developing deficits in social attention or social cognition? Longitudinal studies of these abilities – particularly of children at risk for autism – might shed some light on which (if either) plays the greater causal role. The answer to this question will have important implications for the potential impact of any remediation of face processing. On the former story, early training in face processing might lead to across-the-board reductions of autism symptomatology, whereas on the latter hypothesis, training of face processing will affect only face processing.

One particular causal hypothesis is that atypical eye contact (cf. Jones et al., 2008) might lead to deficits in eye discrimination and hence face recognition. There is some evidence for a connection between atypical eye contact and the recognition performance of facial expressions (Spezio et al., 2007), which is not that surprising as a substantial amount of information on the type of emotion is in the eyes. However, for the case of face *identity* recognition a relationship between eye contact and performance has not been investigated so far. Some evidence against this hypothesis comes from studies comparing Western Caucasian and East Asian participants: while the former focus more on the eyes and the latter on the nose, the two groups perform similarly on face identity recognition tasks (Blais et al., 2008; Caldara et al., 2010). Of course the fixation patterns might still differ between East Asian people and people with ASD, in which case it would still be possible that differential fixation patterns are themselves sufficient to lead to the face processing deficits we find here in ASD.

In sum, the current literature indicates that face processing is qualitatively similar in ASD and typical participants, but that people with ASD perform quantitatively worse (on average) than typical participants on tasks tapping face memory and eye perception. Importantly, these deficits in face processing appear to be specific to faces, rather than reflecting broader impairments in visual processing. Three important questions remain for future research. First, are deficits in face processing causes or consequence of other aspects of the autism phenotype? For example, face recognition deficits could result from deficits in social attention and/or the social reward system. The answer to this question will determine the potential impact of any effective methods for remediation of face processing. Second, which aspects of the neural phenotype of autism underlie the differences in face processing described here (see brief discussion above)? That is, do differences in face processing in ASD result from differences in the fusiform face area, other brain regions, or their respective connections? Third, are the deficits in face perception reported here generally true of most people with ASD, or are they found in a distinct subgroup of people with ASD, and if so how does this subgroup differ cognitively, neurally, and genetically from other individuals with ASD? The answers to these three questions will be crucial for the ultimate future goal of mapping the causal chain that leads from autism genes to autism symptoms.

Acknowledgments

This research was funded by the Ellison Medical Foundation (NK) and the Simons Foundation (KK and SW). The authors wish to thank Yuhong Jiang and Elizabeth Pellicano for comments on the manuscript and Kirsten O'Hearn for sharing her data.

Appendix A. Appendix

Abbreviations

ADI-R	autism diagnostic interview, revised
ADOS	autism diagnostic observation schedule
AFC	alternative forced choice
ASD	autism spectrum disorder
CA	chronological age
CARS	childhood autism rating scale
CFMT	Cambridge Face Memory Test
FIE	face inversion effect

DS Down syndrome
DSM diagnostic and statistical manual of mental disorders
HF high-functioning
ICD International Classification for Diseases
LF low-functioning
M the mean
MA mental age
NVMA non-verbal mental age
PDD-NOS pervasive developmental disorder not otherwise specified
RT reaction time
T typical individual(s)
VMA verbal mental age
WS Williams syndrome

Excluded studies

Subject sample, who did not meet ASD criteria:

Barton, J.J., Cherkasova, M.V., Hefter, R., Cox, T.A., O'Connor, M., Manoach, D.S., 2004. Are patients with social developmental disorders prosopagnosic? Perceptual heterogeneity in the Asperger and socio-emotional processing disorders. Brain 127 (August (Pt 8)), 1706–1716.

Barton, J.J., Hefter, R.L., Cherkasova, M.V., Manoach, D.S., 2007. Investigations of face expertise in the social developmental disorders. Neurology. 69 (August (9)), 860–870.

Hefter, R.L., Manoach, D.S., Barton, J.J., 2005. Perception of facial expression and facial identity in subjects with social developmental disorders. Neurology. 65 (November (10)), 1620–1625.

These studies investigate people with "social developmental disorders", who have similar symptoms as people with ASD, but not the same.

Missing statistics:

Teunisse, J.P., de Gelder, B., 1994. Do autistics have a generalized face processing deficit? Int. J. Neurosci. 77 (July (1–2)), 1–10.

This study does not contain any statistical analysis.

Confounds:

Bookheimer, S.Y., Wang, A.T., Scott, A., Sigman, M., Dapretto, M., 2008. Frontal contributions to face processing differences in autism: evidence from fMRI of inverted face processing. J. Int. Neuropsychol. Soc. 14 (November (6)), 922–932.

This study sets out to investigate the face inversion effect with functional magnetic resonance imaging, while participants performed a simultaneous match-to-sample task on faces or forms. Unfortunately, in the inverted condition, only the target face, but not the two test faces was presented upside-down – and thus the task required mental rotation on top of processing the inverted face. This additional cognitive component makes the data uninterpretable with regard to the standard face inversion effect.

Floor/ceiling effects:

Krebs, J.F., Biswas, A., Pascalis, O., Kamp-Becker, I., Remschmidt, H., Schwarzer, G., 2011. Face processing in children with autism spectrum disorder: independent or interactive processing of facial identity and facial expression? J. Autism Dev. Disord. September 14.

Krebs et al. (2011) first trained a group of children and adolescents with ASD and a group of typical children and adolescents to identify two men from photographs. Participants then had to identify which man was shown (across changes in emotional expression). Groups did not differ in either accuracy or reaction time. However, both groups were at ceiling on this very easy task, which makes an interpretation of the data thus impossible. (This study also contains a task on emotional expression, which we are not discussing here.)

Tantam, D., Monaghan, L., Nicholson, H., Stirling, J., 1989. Autistic children's ability to interpret faces: a research note. J. Child Psychol. Psychiatry. 30 (July (4)), 623–630. Note, we are excluding only the part of this study that investigates the inversion effect (Experiment 2), another part (Experiment 1) is reviewed in Section 3.2.

Tantam et al. (1989) investigated children with ASD and their CA- and NVMA-matched peers and asked them to match emotional labels like "happy" to one of six faces, with these faces presented either upright or inverted. A significant group × orientation interaction was found: typical children were more successful at labeling upright faces than inverted faces (thus showing a face inversion effect), while children with ASD showed the same performance for both upright and inverted faces (thus not showing a face inversion effect). However, the ASD children were at floor in both upright and inverted conditions.

Failure to replicate well-known effects in the typical population (making data of ASD population impossible to interpret):

Annaz, D., Karmiloff-Smith, A., Johnson, M.H., Thomas, M.S., 2009. A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. J. Exp. Child Psychol. 102 (April (4)):456–486.

This study failed to replicate the well-established part-whole effect in the typical participants. Note, however, that we are reviewing this study with respect to of the results from the Benton Facial Recognition test – see Section 3.4.

Several issues:

Nakahachi, T., Yamashita, K., Iwase, M., Ishigami, W., Tanaka, C., Toyonaga, K., et al., 2008. Disturbed holistic processing in autism spectrum disorders verified by two cognitive tasks requiring perception of complex visual stimuli. Psychiatry Res. 159 (June (3)), 330–338.

This study compared a group of 10 participants with ASD across a very wide age range to a group of 15 participants, who were neither matched on chronological age, verbal mental age, or nonverbal mental age. Their also only reporting on part of their accuracy data so that overall an assessment of their results is impossible.

Teunisse, J.P., de Gelder, B., 2003. Face processing in adolescents with autistic disorder: the inversion and composite effect. Brain Cogn. 52 (August (3)), 285–294.

This study has several issues: clinical and typical groups were not matched, and the behavioral tasks differed (e.g. in stimulus presentation and thus difficulty), there are no between-group statistics, and they were not able to replicate the composite effect in the typical children.

References

Annaz, D., Karmiloff-Smith, A., Johnson, M.H., Thomas, M.S., 2009. A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. J. Exp. Child Psychol. 102, 456–486.
Ashwin, C., Wheelwright, S., Baron-Cohen, S., 2005. Laterality biases to chimeric faces in Asperger syndrome: what is 'right' about face-processing? J. Autism Dev. Disord. 35, 183–196.
Bailey, A., Palferman, S., Heavey, L., Le Couteur, A., 1998. Autism: the phenotype in relatives. J. Autism Dev. Disord. 28, 369–392.
Bar-Haim, Y., Shulman, C., Lamy, D., Reuveni, A., 2006. Attention to eyes and mouth in high-functioning children with autism. J. Autism Dev. Disord. 36, 131–137.
Behrmann, M., Avidan, G., Leonard, G.L., Kimchi, R., Luna, B., Humphreys, K., Minshew, N., 2006. Configural processing in autism and its relationship to face processing. Neuropsychologia 44, 110–129.
Blair, R.J., Frith, U., Smith, N., Abell, F., Cipolotti, L., 2002. Fractionation of visual memory: agency detection and its impairment in autism. Neuropsychologia 40, 108–118.
Blais, C., Jack, R.E., Scheepers, C., Fiset, D., Caldara, R., 2008. Culture shapes how we look at faces. PLoS One 3, e3022.
Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., Bailey, A., Rutter, M., 1994. A case-control family history study of autism. J. Child Psychol. Psychiatry 35, 877–900.
Bookheimer, S.Y., Wang, A.T., Scott, A., Sigman, M., Dapretto, M., 2008. Frontal contributions to face processing differences in autism: evidence from fMRI of inverted face processing. J. Int. Neuropsychol. Soc. 14, 922–932.
Boucher, J., Lewis, V., 1992. Unfamiliar face recognition in relatively able autistic children. J. Child Psychol. Psychiatry 33, 843–859.
Boucher, J., Lewis, V., Collis, G., 1998. Familiar face and voice matching and recognition in children with autism. J. Child Psychol. Psychiatry 39, 171–181.

Busigny, T., Rossion, B., 2010. Acquired prosopagnosia abolishes the face inversion effect. Cortex 46, 965–981.
Caldara, R., Zhou, X., Miellet, S., 2010. Putting culture under the 'spotlight' reveals universal information use for face recognition. PLoS One 5, e9708.
Dalton, K.M., Nacewicz, B.M., Johnstone, T., Schaefer, H.S., Gernsbacher, M.A., Goldsmith, H.H., Alexander, A.L., Davidson, R.J., 2005. Gaze fixation and the neural circuitry of face processin in autism. Nat. Neurosci. 8, 519–526.
Davies, S., Bishop, D., Manstead, A.S., Tantam, D., 1994. Face perception in children with autism and Asperger's syndrome. J. Child Psychol. Psychiatry 35, 1033–1057.
Dawson, G., Webb, S.J., McPartland, J., 2005. Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. Dev. Neuropsychol. 27, 403–424.
de Gelder, B., Vroomen, J., van der Heide, L., 1991. Face recognition and lip-reading in autism. Eur. J. Cogn. Psychol. 3, 69–86.
Deruelle, C., Rondan, C., Gepner, B., Tardif, C., 2004. Spatial frequency and face processing in children with autism and Asperger syndrome. J. Autism Dev. Disord. 34, 199–210.
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.
Ellis, H.D., Shepherd, J.W., Davies, G.M., 1979. Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. Perception 8, 431–439.
Faja, S., Webb, S.J., Merkle, K., Aylward, E., Dawson, G., 2009. Brief report: face configuration accuracy and processing speed among adults with high-functioning autism spectrum disorders. J. Autism Dev. Disord. 39, 532–538.
Falck-Ytter, T., 2008. Face inversion effects in autism: a combined looking time and pupillometric study. Autism Res. 1, 297–306.
Gastgeb, H.Z., Rump, K.M., Best, C.A., Minshew, N.J., Strauss, M.S., 2009. Prototype formation in autism: can individuals with autism abstract facial prototypes? Autism Res. 2, 279–284.
Gastgeb, H.Z., Wilkinson, D.A., Minshew, N.J., Strauss, M.S., 2011. Can Individuals with autism abstract prototypes of natural faces? J. Autism Dev. Disord. 41, 1609–1618.
Gauthier, I., Klaiman, C., Schultz, R.T., 2009. Face composite effects reveal abnormal face processing in autism spectrum disorders. Vision Res. 49, 470–478.
Gepner, B., deGelder, B., deSchonen, S., 1996. Face processing in autistics: evidence for a generalised deficit? Child Neuropsychol. 2, 123–139.
Gilbert, C., Bakan, P., 1973. Visual asymmetry in perception of faces. Neuropsychologia 11, 355–362.
Golarai, G., Grill-Spector, K., Reiss, A.L., 2006. Autism and the development of face processing. Clin. Neurosci. Res. 6, 145–160.
Hadjikhani, N., Joseph, R.M., Snyder, J., Chabris, C.F., Clark, J., Steele, S., McGrath, L., Vangel, M., Aharon, I., Feczko, E., Harris, G.J., Tager-Flusberg, H., 2004. Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. Neuroimage 22, 1141–1150.
Hadjikhani, N., Joseph, R.M., Snyder, J., Tager-Flusberg, H., 2006. Abnormal activation of the social brain during face perception in autism. Hum. Brain Map. 28, 441–449.
Harms, M.B., Martin, A., Wallace, G.L., 2010. Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol. Rev. 20, 290–322.
Hauck, M., Fein, D., Maltby, S., Waterhouse, L., Feinstein, C., 1998. Memory for faces in children with autism. Child Neuropsychol. 4, 187–198.
Hobson, R.P., Ouston, J., Lee, A., 1988. What's in a face? The case of autism. Br. J. Psychol. 79 (Pt 4), 441–453.
Jemel, B., Mottron, L., Dawson, M., 2006. Impaired face processing in autism: fact or artifact? J. Autism Dev. Disord. 36, 91–106.
Jones, W., Carr, K., Klin, A., 2008. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. Arch. Gen. Psychiatry 65, 946–954.
Joseph, R.M., Ehrman, K., McNally, R., Keehn, B., 2008. Affective response to eye contact and face recognition ability in children with ASD. J. Int. Neuropsychol. Soc. 14, 947–955.
Joseph, R.M., Tanaka, J., 2003. Holistic and part-based face recognition in children with autism. J. Child Psychol. Psychiatry 44, 529–542.
Kirchner, J.C., Hatri, A., Heekeren, H.R., Dziobek, I., 2011. Autistic symptomatology, face processing abilities, and eye fixation patterns. J. Autism Dev. Disord. 41, 158–167.
Kleinhans, N.M., Richards, T., Sterling, L., Stegbauer, K.C., Mahurin, R., Clark Johnson, L., Greenson, J., Dawson, G., Aylward, E., 2008. Abnormal functional connectivity in autism spectrum disorders during face processing. Brain 131, 1000–1012.
Klin, A., Sparrow, S.S., de Bildt, A., Cicchetti, D.V., Cohen, D.J., Volkmar, F.R., 1999. A normed study of face recognition in autism and related disorders. J. Autism Dev. Disord. 29, 499–508.
Klinger, L.G., Dawson, G., 2001. Prototype formation in autism. Dev. Psychopathol. 13, 111–124.
Koshino, H., Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., Just, M.A., 2008. fMRI investigation of working memory for faces in autism: visual coding and underconnectivity with frontal areas. Cereb. Cortex 18, 289–300.
Krebs, J.F., Biswas, A., Pascalis, O., Kamp-Becker, I., Remschmidt, H., Schwarzer, G., 2011. Face processing in children with autism spectrum disorder: independent or interactive processing of facial identity and facial expression? J. Autism Dev. Disord. 41, 796–804.
Lahaie, A., Mottron, L., Arguin, M., Berthiaume, C., Jemel, B., Saumier, D., 2006. Face perception in high-functioning autistic adults: evidence for superior processing of face parts, not for a configural face-processing deficit. Neuropsychology 20, 30–41.
Langdell, T., 1978. Recognition of faces: an approach to the study of autism. J. Child Psychol. Psychiatry 19, 255–268.
Lawrence, K., Kuntsi, J., Coleman, M., Campbell, R., Skyse, D., 2003. Face and emotion recognition deficits in Turner syndrome: a possible role for X-linked genes in amygdala development. Neuropsychology 17, 39–49.
Lopez, B., Donnelly, N., Hadwin, J.A., Leekam, S.R., 2004. Face processing in high-functioning adolescents with autism: evidence for weak central coherence. Visual Cogn. 11, 673–688.
Marcus, D.J., Nelson, C.A., 2001. Neural bases and development of face recognition in autism. CNS Spectr. 6, 36–59.
McKone, E., Crookes, K., Kanwisher, N., 2009. The cognitive and neural development of face recognition in humans. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. , 4th ed. MIT Press, Cambridge, MA, pp. 467–482.
McKone, E., Robbins, R., 2011. Are faces special? In: Calder, A., Rhodes, G., Johnson, M., Haxby, J. (Eds.), Oxford Handbook of Face Perception. Oxford University Press, Oxford, UK.
McPartland, J.C., Webb, S.J., Keehn, B., Dawson, G., 2011. Patterns of visual attention to faces and objects in autism spectrum disorder. J. Autism Dev. Disord. 41, 148–157.
Nakahachi, T., Yamashita, K., Iwase, M., Ishigami, W., Tanaka, C., Toyonaga, K., Maeda, S., Hirotsune, H., Tei, Y., Yokoi, K., Okajima, S., Shimizu, A., Takeda, M., 2008. Disturbed holistic processing in autism spectrum disorders verified by two cognitive tasks requiring perception of complex visual stimuli. Psychiatry Res. 159, 330–338.
Nishimura, M., Rutherford, M.D., Maurer, D., 2008. Converging evidence of configural processing of faces in high-functioning adults with autism spectrum disorders. Visual Cog. 16, 859–891.
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.
Ozonoff, S., Pennington, B.F., Rogers, S.J., 1990. Are there emotion perception deficits in young autistic children? J. Child Psychol. Psychiatry 31, 343–361.
Pellicano, E., Jeffery, L., Burr, D., Rhodes, G., 2007. Abnormal adaptive face-coding mechanisms in children with autism spectrum disorder. Curr. Biol. 17, 1508–1512.
Perlman, S.B., Hudac, C.M., Pegors, T., Minshew, N.J., Pelphrey, K.A., 2011. Experimental manipulation of face-evoked activity in the fusiform gyrus of individuals with autism. Soc. Neurosci. 6, 22–30.
Pierce, K., Courchesne, E., 2000. Exploring the neurofunctional organization of face processing in autism. Arch. Gen. Psychiatry 57, 344–346.
Pierce, K., Mueller, R.-A., Ambrose, J., Allen, G., Courchesne, E., 2001. Face processing occurs outside the fusiform 'face area' in autism: evidence from functional MRI. Brain 124, 2059–2073.
Ramon, M., Busigny, T., Rossion, B., 2010. Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. Neuropsychologia 48, 933–944.
Riby, D.M., Doherty-Sneddon, G., Bruce, V., 2008. Exploring face perception in disorders of development: evidence from Williams syndrome and autism. J. Neuropsychol. 2, 47–64.
Riby, D.M., Doherty-Sneddon, G., Bruce, V., 2009. The eyes or the mouth? Feature salience and unfamiliar face processing in Williams syndrome and autism. Q. J. Exp. Psychol. (Colchester) 62, 189–203.
Robel, L., Ennouri, K., Piana, H., Vaivre-Douret, L., Perier, A., Flament, M.F., Mouren-Simeoni, M.C., 2004. Discrimination of face identities and expressions in children with autism: same or different? Eur. Child Adolesc. Psychiatry 13, 227–233.
Rondan, C., Gepner, B., Deruelle, C., 2003. Inner and outer face perception in children with autism. Child Neuropsychol. 9, 289–297.
Rose, F.E., Lincoln, A.J., Lai, Z., Ene, M., Searcy, Y.M., Bellugi, U., 2007. Orientation and affective expression effects on face recognition in Williams syndrome and autism. J. Autism Dev. Disord. 37, 513–522.
Rosset, D.B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., Deruelle, C., 2008. Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. J. Autism Dev. Disord. 38, 919–925.
Rouse, H., Donnelly, N., Hadwin, J.A., Brown, T., 2004. Do children with autism perceive second-order relational features? The case of the Thatcher illusion. J. Child Psychol. Psychiatry 45, 1246–1257.
Rutherford, M.D., Clements, K.A., Sekuler, A.B., 2007. Differences in discrimination of eye and mouth displacement in autism spectrum disorders. Vision Res. 47, 2099–2110.
Sasson, N.J., 2006. The development of face processing in autism. J. Autism Dev. Disord. 36, 381–394.
Scherf, K.S., Behrmann, M., Minshew, N., Luna, B., 2008. Atypical development of face and greeble recognition in autism. J. Child Psychol. Psychiatry 49, 838–847.
Scherf, K.S., Luna, B., Minshew, N., Behrmann, M., 2010. Location, location, location: alterations in the functional topography of face- but not object- or place-related cortex in adolescents with autism. Front. Hum. Neurosci. 22 (4), 26.
Schultz, R.T., Gauthier, I., Klin, A., Fulbright, R.K., Anderson, A.W., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D.J., Gore, J.C., 2000. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Arch. Gen. Psychiatry 57, 331–340.

Schultz, R.T., 2005. Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. Int. J. Dev. Neurosci. 23, 125–141.
Serra, M., Althaus, M., de Sonneville, L.M., Stant, A.D., Jackson, A.E., Minderaa, R.B., 2003. Face recognition in children with a pervasive developmental disorder not otherwise specified. J. Autism Dev. Disord. 33, 303–317.
Simmons, D.R., Robertson, A.E., McKay, L.S., Toal, E., McAleer, P., Pollick, F.E., 2009. Vision in autism spectrum disorders. Vision Res. 49, 2705–2739.
Spezio, M.L., Adolphs, R., Hurley, R.S., Piven, J., 2007. Abnormal use of facial information in high-functioning autism. J. Autism Dev. Disord. 37, 929–939.
Tanaka, J.W., Farah, M.J., 1993. Parts and wholes in face recognition. Q. J. Exp. Psychol. A 46, 225–245.
Tantam, D., Monaghan, L., Nicholson, H., Stirling, J., 1989. Autistic children's ability to interpret faces: a research note. J. Child Psychol. Psychiatry 30, 623–630.
Teunisse, J.P., de Gelder, B., 2003. Face processing in adolescents with autistic disorder: the inversion and composite effects. Brain Cogn. 52, 285–294.
Thomas, C., Humphreys, K., Jung, K.J., Minshew, N., Behrmann, M., 2011. The anatomy of the callosal and visual-association pathways in high-functioning autism: a DTI tractography study. Cortex 47, 863–873.
Thompson, P., 1980. Margaret Thatcher: a new illusion. Perception 9, 483–484.
Tippett, L.J., Miller, L.A., Farah, M.J., 2000. Prosopamnesia: a selective impairment in face learning. Cogn. Neuropsychol. 17, 241–255.
van der Geest, J.N., Kemner, C., Verbaten, M.N., van Engeland, H., 2002. Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. J. Child Psychol. Psychiatry 43, 669–678.
Wallace, S., Coleman, M., Bailey, A., 2008. Face and object processing in autism spectrum disorders. Autism Res. 1, 43–51.
Williams, D.L., Berberovic, N., Mattingley, J.B., 2007. Abnormal fMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. Curr. Biol. 17, 1259–1264.
Williams, D.L., Goldstein, G., Minshew, N.J., 2005. Impaired memory for faces and social scenes in autism: clinical implications of memory dysfunction. Arch. Clin. Neuropsychol. 20, 1–15.
Wilson, C.E., Brock, J., Palermo, R., 2010a. Attention to social stimuli and facial identity recognition skills in autism spectrum disorder. J. Intellect. Disabil. Res. 54, 1104–1115.
Wilson, C.E., Freeman, P., Brock, J., Burton, A.M., Palermo, R., 2010b. Facial identity recognition in the broader autism phenotype. PLoS One 5, e12876.
Wilson, R., Pascalis, O., Blades, M., 2007. Familiar face recognition in children with autism; the differential use of inner and outer face parts. J. Autism Dev. Disord. 37, 314–320.
Wolf, J.M., Tanaka, J.W., Klaiman, C., Cockburn, J., Herlihy, L., Brown, C., South, M., McPartland, J., Kaiser, M.D., Phillips, R., Schultz, R.T., 2008. Specific impairment of face-processing abilities in children with autism spectrum disorder using the Let's Face It! skills battery. Autism Res. 1, 329–340.
Yin, R., 1969. Looking at upside-down faces. J. Exp. Psychol. 81, 141–145.
Young, A.W., Hellawell, D., Hay, D.C., 1987. Configurational information in face perception. Perception 16, 747–759.