

The Cambridge Mindreading (CAM) Face-Voice Battery: Testing Complex Emotion Recognition in Adults with and without Asperger Syndrome

Ofer Golan,^{1,2,*} Simon Baron-Cohen,¹ and Jacqueline Hill¹

Adults with Asperger Syndrome (AS) can recognise simple emotions and pass basic theory of mind tasks, but have difficulties recognising more complex emotions and mental states. This study describes a new battery of tasks, testing recognition of 20 complex emotions and mental states from faces and voices. The battery was given to males and females with AS and matched controls. Results showed the AS group performed worse than controls overall, on emotion recognition from faces and voices and on 12/20 specific emotions. Females recognised faces better than males regardless of diagnosis, and males with AS had more difficulties recognising emotions from faces than from voices. The implications of these results are discussed in relation to social functioning in AS.

KEY WORDS: Emotion recognition; complex emotions; face perception; voice perception; theory of mind; Asperger Syndrome; autism spectrum; psychometrics.

The ability to understand people's emotional and other mental states, also known as "theory of mind" or "mindreading" (Wellman, 1992) underlies fundamental social skills. This ability is also sometimes referred to as empathising (Baron-Cohen, 2002; Baron-Cohen, 2003). Individuals with autism spectrum conditions have difficulties understanding others' mental (including their emotional) states, which is thought to play a major role in their social and communicational difficulties (Baron-Cohen, 1995). The measurement and analysis of such difficulties may shed light on the phenotype of the autistic

spectrum, and be clinically useful in diagnosis and treatment. The present study reports a new test battery to assess recognition of complex mental states.

The standard test of emotion recognition (Ekman & Friesen, 1971) focuses on the ability to recognise six basic emotions that are recognised cross-culturally: happiness, sadness, anger, fear, surprise, and disgust. Some studies report that people with autism spectrum conditions have difficulties recognising such emotions from photographs (Bormann-Kischkel, Vilsmeier, & Baude, 1995), films (Yirmiya, Sigman, Kasari, & Mundy, 1992) and voice recordings (Loveland, Tunali Kotoski, Chen, & Brelsford, 1995) as well as with matching faces and voices of these emotions (Buitelaar, Van der Wees, Swabb Barneveld, & Van der Gaag, 1999; Hobson, 1986a, b). However, other studies (Grossman, Klin, Carter, & Volkmar, 2000; Loveland *et al.*, 1997) have found that children and adolescents with high functioning autism or Asperger Syndrome (AS) have no difficulties in basic emotion

¹ Department of Psychiatry, Autism Research Centre, Cambridge University, Douglas House, 18b Trumpington Road, CB2 2AH, Cambridge, UK.

² Correspondence should be addressed to: Ofer Golan, Department of Psychiatry, Autism Research Centre, Cambridge University, Douglas House, 18b Trumpington Road, CB2 2AH, Cambridge, UK; Tel: 44 (0)1223 746062; Fax: 44 (0)1223 746033; e-mail: og211@cam.ac.uk

recognition. Adults with autism spectrum conditions may also not differ from typically developed controls in the ability to recognise the 6 basic emotions (Adolphs, Sears, & Piven, 2001). These groups also passed first-and second-order theory of mind tasks, which *could* be taken to indicate no theory of mind difficulty in autism.

This lack of difficulty in basic emotion recognition can be explained in different ways. For example, one can distinguish situation-based emotions (such as sadness and happiness) from belief-based emotions (such as surprise) (Baron-Cohen, Spitz, & Cross, 1993). According to the theory of mind hypothesis, individuals with autism have specific deficits in recognising belief-based emotions, compared to situation-based emotions. The latter study found that, compared to learning disabled and typically developing controls, matched on mental age, children with autism had no difficulty in recognising sadness or happiness, but they found it harder to recognise surprise. Similarly, children with high functioning autism (HFA) found it harder to describe situations in which they felt embarrassed and to recognise embarrassment (a belief-based emotion) from photographs, relative to a matched task involving sadness or happiness (Capps, Yirmiya, & Sigman, 1992). These findings support the distinction between “situational” and “cognitive” emotions (Harris, 1989).

The lack of difficulty in passing first- and second-order theory of mind tasks by people with HFA and AS can be explained by the fact this competence is usually achieved by typically developing children as young as 6 years old. The fact that people with AS or HFA pass these tasks may simply indicate that their understanding of mental states is at least equal to that of an average 6 year-old.

The success with basic emotion recognition by people with AS may also reflect that such individuals (who have normal intelligence) may be using compensation strategies to bypass their earlier difficulties. For example, Grossman *et al.* (2000) showed children with AS pictures of 5 “basic” emotions, with matching or mismatching labels. The children with AS had no problem recognising these emotions or identifying the emotions when labelled with matching words, but had difficulties in recognising the emotions in the mismatching labels condition. This result suggests that, instead of recognising the emotions in the face, children with AS were using the written label to answer the question. The authors concluded that individuals with AS use verbal mediation as a compensatory

strategy, which may mask their deficits under certain circumstances.

For these reasons, in order to examine mental state recognition abilities in people with autism spectrum conditions of normal intelligence, there is a need for more a fine grained test battery, examining more subtle emotions.

To address this need, the Strange Stories Test was developed (Happé, 1994). This assesses the ability to provide context-appropriate mental state explanations for non-literal statements made by story characters (e.g. ironic or sarcastic statements). Happé’s study with adolescents found specific deficits on this task, a result that was later replicated with adults with AS or HFA (Jolliffe & Baron-Cohen, 1999).

A different approach involves the “Reading the Mind in the Eyes” test, which measures the ability to identify complex mental states from still pictures of the eye region (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). This approach has been extended in two studies, using tasks in which the participant is required to identify mental states from recordings of brief spoken phrases (Kleinman, Marciano, & Ault, 2001; Rutherford, Baron-Cohen, & Wheelwright, 2002). All these tasks showed that individuals with AS or HFA, despite having at least normal intelligence, have specific difficulties in identifying subtle mental states.

However, there are a few methodological and conceptual issues that these tasks do not address:

1. The “Reading the Mind in the Eyes” (the Eyes task) and the “Reading the Mind in the Voice” (the Voice task) tasks test different complex emotions and mental states, but they do not systematically examine the recognition of particular emotions. Rather, these tests include a variety of stimuli, which generate one overall score of emotion recognition.
2. The Eyes task, like most visual tasks of emotion recognition, uses *still* pictures rather than motion. The result is a relatively narrow range of emotions which can be studied, as the distinction between many emotions requires motion (e.g. relief). In addition, this makes the test less naturalistic, and therefore may not assess an individual’s actual ability to identify emotions from moving faces.
3. No tasks have studied understanding of complex mental states via both visual *and* auditory channels. This has been tested with

the basic emotions (Hobson, 1986a, b; Love-land *et al.*, 1995).

In the construction of The Cambridge Mindreading (CAM) Face-Voice Battery, we address these issues. The test was designed to assess the wide emotional repertoire of adults and to examine each emotion thoroughly through both visual and auditory modalities, using motion in the visual task.

The test is based on a new taxonomy of emotion (Baron-Cohen, Golan, Wheelwright, & Hill, 2004). This model comprises 412 unique emotion concepts, including all the emotion terms in the English language, as well as epistemic mental states with an emotional dimension (e.g. doubting). Mental states that could be a purely bodily state (e.g. hungry) and epistemic states with no emotional dimension (e.g. reasoning) are not included.

These 412 emotions are grouped into 24 mutually exclusive emotion groups (such as the happy group, the thinking group, the sneaky group, etc). A list of the 24 emotion groups is shown in Fig. 1. Developmental testing resulted in the 412 emotion concepts being further subdivided into six different levels, on the basis of word frequency in the English language and verbal comprehension. The six levels represent an age range from preschool through to adulthood. The full list of emotions, according to emotion groups and developmental levels can be found elsewhere (Baron-Cohen, Golan, Hill & Wheelwright, submitted; Baron-Cohen *et al.*, 2004).

Using this taxonomy, and together with a multimedia production company, we created a multimedia database on DVD-ROM, in which actors of both sexes, different age groups and ethnicities enacted each of the 412 emotions. These were captured using video of their facial expression, and using audio of their vocalisation. Each of the 412 emotions therefore has 6 films and 6 voice recordings expressing it

Afraid	Excited	Liked	Surprised
Angry	Fond	Romantic	Thinking
Bored	Happy	Sad	Touched
Bothered	Hurt	Sneaky	Unfriendly
Disbelieving	Interested	Sorry	Unsure
Disgusted	Kind	Sure	Wanting

Fig. 1. The 24 groups included in the emotion taxonomy (adapted from Baron-Cohen *et al.*, 2004).

(Baron-Cohen, Hill, Golan, & Wheelwright, 2002). Every film and audio recording of every emotion was validated, using a panel of 10 volunteers from the general population. If at least 8 out of 10 judges agreed that the label given to a face or voice could be a reasonable description of the particular emotion, the film or voice recording was included in the database. This database is used in a piece of educational software, for teaching emotion recognition (for details, see www.jkp.com/mindreading).

The CAM battery evaluates a selection of 20 emotion concepts, taken from the above mentioned taxonomy, representing 18 of the 24 emotion groups. This covers a good variety of emotions and mental states, while remaining relatively brief for administration. Coverage of all 412 emotion concepts would have made the test itself unfeasible in a single session. The battery includes two tasks: emotion recognition in the face and emotion recognition in the voice. Each of these tasks has fifty questions, in which the participant is either watching 3–5 second silent clips of actors portraying an emotion (facial task), or listening to short sentences, spoken in a particular emotional intonation (vocal task). After watching the clip/listening to the voice recording, the participant is presented with four adjectives and is asked to “choose the word that best describes how the person is feeling”.

In order to make sure that the chosen concepts are taken from the adult emotional repertoire, they were all selected from the higher levels of the taxonomy; 6 concepts from level 4 (concepts understood by typical 15–16 year olds), 13 concepts from level 5 (understood by typical 17–18 year olds) and one concept from level 6 (words understood by less than 75% of typical 17–18 year olds). A detailed description of the level validation process is available elsewhere (Baron-Cohen *et al.*, submitted).

We included mental states that are “positive” in valence, such as *empathic* and *intimate*, as well as concepts that are “negative”, such as *guarded* and *insincere*. An attempt was also made to include emotions of varying intensity, i.e.—subtle emotions on the one hand (e.g.—*uneasy*, *subdued*) and intense ones on the other (e.g.—*appalled*, *mortified*). Concepts were coded for intensity and valence when creating the taxonomy by 3 independent judges (Baron-Cohen *et al.*, submitted). In total, the CAM battery has 5 “positive” concepts, 12 “negative”, and 3 “neutral”. 3 concepts were coded as having high intensity. The twenty concepts with their level, valence and intensity coding are listed in Table I.

Table I. The 20 Emotional Concepts Included in the CAM, their Emotion Group, Developmental Level, Valence and Intensity^a

Concept	Emotion Group	Level	Valence	Intensity
Appalled	Surprised	4	Negative	Strong
Appealing (asking for)	Wanting	4	Neutral	
Confronted	Hurt	5	Negative	Strong
Distaste	Disgusted	5	Negative	
Empathic	Kind	5	Positive	Strong
Exonerated	Happy	6	Positive	
Grave	Sad	5	Negative	Strong
Guarded	Disbelieving	4	Negative	
Insincere	Sneaky	5	Negative	Strong
Intimate	Romantic	4	Positive	
Lured	Interested	5	Neutral	Strong
Mortified	Sorry	5	Negative	
Nostalgic	Touched	5	Neutral	Strong
Reassured	Liked	4	Positive	
Resentful	Unfriendly	5	Negative	Strong
Stern	Unfriendly	5	Negative	
Subdued	Sad	5	Negative	Strong
Subservient	Unsure	5	Negative	
Uneasy	Afraid	4	Negative	Strong
Vibrant	Excited	5	Positive	

Note: ^aconcepts with high emotional intensity are marked as “strong”.

Each of the 20 emotion concepts tested was expressed by 5 examples (or items). Our criterion for passing a concept was correct recognition of at least 4 out of 5 items. Achieving 4 or more out of 5 would represent above chance recognition of the concept (Binomial test, $p < .05$). Of the 20 concepts, 10 concepts were measured by the following combination of 5 items: 3 faces and two voices. The other 10 concepts were measured by a variation on this: 3 voices and two faces. This design allowed us to keep the task brief overall.

There are four different scores that can be derived from the CAM:

1. *An overall emotion recognition score*: defined as the sum of all the correctly answered questions, ranging from 0 to 100, describing overall emotion and mental state recognition. Any score greater than 35 is above chance at the $p < .01$ level (Binomial test).
2. *Facial emotion recognition score*: defined as the sum of all items answered correctly from the facial items, ranging from 0 to 50. Any score greater than 20 is above chance at the $p < .01$ level (Binomial test).
3. *Vocal emotion recognition score*: defined as the sum of all items answered correctly in the vocal items, ranging from 0 to 50. Any score

greater than 20 is above chance at the $p < .01$ level (Binomial test).

4. *Concepts correctly recognised*: The concepts can be studied in two ways: The sum of concepts correctly recognised, ranging from 0 to 20 (any score greater than 2 is above chance at the $p < .01$ level, according to Binomial test), or the particular concepts correctly answered, analysed individually/by subgroups (e.g.—positive/negative).

In this study, we predicted that participants with AS would show deficits, relative to matched controls, on all 4 scores above.

Another comparison we carried out was between the ability to recognise emotions from the face and from the voice. Such a comparison is possible since the CAM includes the same emotions in the facial and vocal scale. We also tested for a group difference on the face-voice comparison (i.e. a group by modality interaction).

Lastly, we predicted that CAM scores would positively correlate with the Revised “Reading the Mind in the Eyes” test (Baron-Cohen *et al.*, 2001) and with a revised version of the “Reading the Mind in the Voice” test (Golan, Baron-Cohen, Rutherford, & Hill, submitted), which includes 4 foils for each question instead of two in Rutherford *et al.*’s original

task (Rutherford *et al.*, 2002). The CAM scores were also predicted to correlate negatively with scores of the Autism Spectrum Quotient (AQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001).

METHOD

Participants

Group 1 comprised twenty-one adults (15 males and 6 females), aged 18–50 (Mean age = 30.2, SD = 10.5). Participants had all been diagnosed with Asperger Syndrome in specialist centres using established criteria (APA, 1994; World-Health-Organisation, 1994). They were recruited from several sources, including a local clinic for adults with AS, support organisations, and colleges for individuals with autism spectrum conditions around the UK. All participants were given the Wechsler Abbreviated Scale of Intelligence (WASI), comprising the vocabulary, similarities, block design and matrix reasoning tests. The WASI produces verbal, performance and full scale IQ scores, with correlations of .88, .84 and .92, respectively, with the full Wechsler scales (Wechsler, 1999). All participants scored above 70 on both verbal and performance scales. Mean verbal IQ score was 114.4 (SD = 9.02) and mean performance IQ score was 107.71 (SD = 15.95).

Group 2 comprised 21 adults recruited from a local employment agency. After screening for autistic spectrum conditions using the AQ (Baron-Cohen *et al.*, 2001), four participants were excluded for scoring above cut-off of 32. The remaining seventeen, 12 males and 5 females, matched the clinical group in age (range = 18–51, mean age = 27.1, SD = 10.3) verbal IQ (mean = 118.47, SD = 7.42) and performance IQ (mean = 110.82, SD = 8.57). They spanned an equivalent range of socio-economic classes and educational levels as that seen in the clinical group. Chi-square test for sex and *t*-tests for age, verbal and performance IQ revealed no

significant differences between the groups at the $p < .05$ level (see Table II).

Instruments

The CAM: Test Development

Twenty-five concepts were selected from the emotion taxonomy (Baron-Cohen *et al.*, submitted). Six of the twenty-five were taken from level 4, 15 from level 5 and four from level 6, which are all adult levels. A pilot study carried out with adult participants with AS using the basic emotions resulted in ceiling effects. This confirmed that choice of the adult level emotion concepts was appropriate.

Selection of the 25 concepts followed 3 principles: concepts should be (a) selected from all 24 emotion groups, (b) mainly subtle, and (c) important for everyday social functioning. Although most of the 24 emotion groups were represented in the CAM by one emotion concept, the larger emotion groups (*unfriendly* and *sad*) were represented by 2 emotion concepts each. For each concept, six test questions were created, using a computer programme which randomly selected three video films and three audio voice recordings out of the six existing on the multimedia database, and matched them with foil words from levels 4 and 5 of the emotion taxonomy, ensuring that foils were not from the same emotion group as the target answer. Although choosing foils from other emotion groups possibly makes the tasks easier, it was believed that foils taken from the same categories as targets might be too similar and increase the dependency of performance on verbal ability (i.e. the ability to distinguish one emotion from another within the same group). Some of the groups might still provide quite difficult foils to other emotion groups with a similar valence and theme. For example—emotions from the *unfriendly* group were used with targets from the *angry* group, emotions from the *sad* group were used with targets from the *hurt* group, etc.

Table II. Means, Standard Deviations and Ranges of Chronological Age and WASI Scores for the AS and Control Groups

	AS group ($n = 21$)			Control group ($n = 17$)			<i>t</i>
	Mean	SD	Range	Mean	SD	Range	
Age	30.23	10.50	17.9–49.9	27.08	10.25	17.6–51.2	.93
VIQ	114.38	9.02	100–127	118.47	7.42	104–130	1.50
PIQ	107.71	15.95	72–134	110.82	8.57	92–125	.72
FSIQ	112.48	12.26	84–131	116.29	6.71	106–128	1.22

Note: $p > .1$ for all *t*-tests.

Two tasks (face recognition and voice recognition) were created and run on an IBM compatible computer, using the experimental software DMDX (Forster & Forster, 2003). Each task started with an instruction slide, followed by two practice items, to ensure understanding of the task. In both tasks the test items were presented in a random order. The face task comprised silent clips of adult actors, both male and female, and of different ethnicities, expressing the emotions in the face (though with shoulders visible). Example questions showing one frame from two of the clips are shown in Fig. 2. The voice task comprised recordings of short sentences expressing various emotional intonations. An attempt was made to use as neutral content as possible for the voice recordings (e.g. "Where did you get that?"), so that recognition had to rely primarily on intonation rather than semantic content.

In both tasks four adjectives, numbered from 1 to 4, were presented after playing each stimulus. Participants were asked to press 1, 2, 3 or 4 on a keyboard to select their preferred answer. After choosing an answer the next item was presented. No feedback was given during the task. A handout of definitions of all the adjectives used in the task was



Fig. 2. Examples of questions from the emotion recognition in the face task (showing only one frame out of the full clip).

available for the participants at the beginning of the assessment. Participants were encouraged to go through it and make sure they were familiar with all the words, and to use the handout in case of any doubt during the task. Hence, response time was unrestricted and not measured.

Item validation was conducted before carrying out any group analysis. The data from the 21 adults in the control group was first analysed as follows: An item was considered valid if at least 11 out of 21 (> 50%) of these participants selected the target word and no more than 6 (< 33%) selected any one of the foils. Using these criteria, 5 of the concepts (3 from level 6 and 2 from level 5) were excluded from the battery. Eight other concepts had one invalid item each, and these items were removed. In order to keep the same number of items for all the concepts and to keep the battery brief to administer, one item was then randomly removed from each of the remaining 12 concepts, so that the final battery comprised 20 concepts with 5 items for each concept. The number of concepts having 3 face and 2 voice items or 3 voice and 2 face items were counterbalanced.

Autism Spectrum Quotient (AQ, Baron-Cohen et al., 2001): The AQ is a self-report questionnaire, which measures the degree to which any adult of normal IQ possesses traits related to the autistic spectrum. Scores range from 0 to 50, and the higher the score, the more autistic traits a person possesses.

Reading the Mind in the Eyes Task [Revised, adult version] (Baron-Cohen et al., 2001): The task has 36 items, in which participants are presented with a photograph of the eyes region of the face and must choose one of four adjectives or phrases to describe the mental state of the person pictured. A definition handout is provided at the beginning of the task and a practice item precedes the first trial. In the present study, the pictures and adjectives were presented on the computer screen (using DMDX software), to avoid possible difficulties due to communication with a human examiner (Ozonoff, 1995). Items were presented in a random order. There was no time limit for answering.

Reading the Mind in the Voice Task (Revised). We used a revised version of the original task (Rutherford et al., 2002). In the original task, 40 segments of speech, taken from BBC drama series, were played on audio tape to the participants, who were asked to choose one out of two possible answers, describing the speaker's mental state in each item. The task was revised as follows: Each of the test items was allocated two more foils, taken from the

same level, one level above or one level below the correct answer (based on the emotion taxonomy). Foils were selected to match the content of the verbalisations but not the intonation, thus making the task harder to answer. This avoided ceiling effects that the original version of the test is prone to. Seven items were removed since the authors found the correct answer inappropriate to the verbalisation. Eight more items were excluded after validation by a sample of 15 typically developing adults. The final task included 25 items with 4 possible answers for each of them (For a full description of the revision and validation, see Golan *et al.*, submitted). The test items were “cleaned” from background noise as far as possible and played on a computer (using DMDX software) in random order, preceded by an instruction slide and two practice items. Participants were again given a definition handout before the beginning of the task. There was no time limit for answering.

Procedure

Participants were tested either at the Autism Research Centre in Cambridge, or at local support centres and colleges for individuals with autism spectrum conditions. Participants were seated in front of IBM compatible computers with 15 inch monitors and were given headphones for the voice tasks. The CAM took about 45 minutes to complete.

RESULTS

A calculation of facial, vocal and overall scores was made by counting the number of correct answers in each of the scales and across the whole battery. Furthermore, a tally of concepts correctly recognised was made. A minimum of 4 correct answers out of 5 items was considered successful recognition of a concept.

All participants scored above chance ($p < .01$, Binomial test) on the vocal scale and on the overall score. All except for two participants from the AS group scored above chance on the facial scale ($p < .01$). These two participants scored just below the threshold (20 correct answers out of 50), one of whom scored above chance at the $p < .05$ level. They were therefore included in the analysis.

Univariate analyses of variance (ANOVA) were performed on the facial scale score, the vocal scale score, the CAM overall score and the number of concepts passed. Group (AS/control) and Sex were

the independent variables. Verbal IQ, performance IQ, and age were entered as covariates.

A main effect of group was found in all four analyses. The scores of the AS group were significantly lower than those of the control group on the facial scale ($F[1,31] = 15.61$, $p < .001$), the vocal scale ($F[1,31] = 21.26$, $p < .001$), the overall score ($F[1,31] = 25.32$, $p < .001$) and the number of concepts passed ($F[1,31] = 15.67$, $p < .001$).

In addition, a main effect of sex was found for the facial scale ($F[1,31] = 5.02$, $p < .05$). Females were significantly better than males at recognising emotions from faces in both groups, regardless of diagnosis. No other effect or covariate contribution was significant. The means and standard deviations of the 4 scores for males and females in the AS and control group are shown in Table III and in Fig. 3.

Table III. Facial, Vocal and Overall Mean Scores and Number of Concepts Passed for Males and Females in the AS & Control Groups

		AS ^c	Control ^c	Total
<i>a. Facial scale (Max = 50)</i>				
Females	Mean	35.33	46.00	40.18 ^b
	SD	6.95	2.55	7.60
Males	Mean	31.13	42.50	36.19 ^b
	SD	8.24	4.17	8.78
Total	Mean	32.33 ^a	43.53 ^a	37.34
	SD	7.96	4.03	8.55
<i>b. Vocal scale (Max = 50)</i>				
Females	Mean	33.17	44.80	38.45
	SD	9.62	2.59	9.27
Males	Mean	36.73	41.92	39.04
	SD	4.22	3.96	4.81
Total	Mean	35.71 ^a	42.76 ^a	38.87
	SD	6.19	3.78	6.29
<i>c. Overall score (Max = 100)</i>				
Females	Mean	68.50	90.80	78.64
	SD	15.76	3.63	16.28
Males	Mean	67.87	84.42	75.22
	SD	10.32	5.85	11.92
Total	Mean	68.05 ^a	86.29 ^a	76.21
	SD	11.69	5.99	13.19
<i>d. Number of concepts passed (Max = 20)</i>				
Females	Mean	11.50	18.40	14.64
	SD	5.68	.89	5.43
Males	Mean	10.47	16.08	12.96
	SD	4.32	2.50	4.56
Total	Mean	10.76 ^a	16.76 ^a	13.45
	SD	4.62	2.39	4.81

Notes: ^aTotal difference between AS and control groups is significant for all 4 measures ($p < .001$).

^bTotal difference between males and females is significant for the facial scale ($p < .05$).

^cGroup sizes: 6 females and 15 males in the AS group (total = 21), 5 females and 12 males in the control group (total = 17).

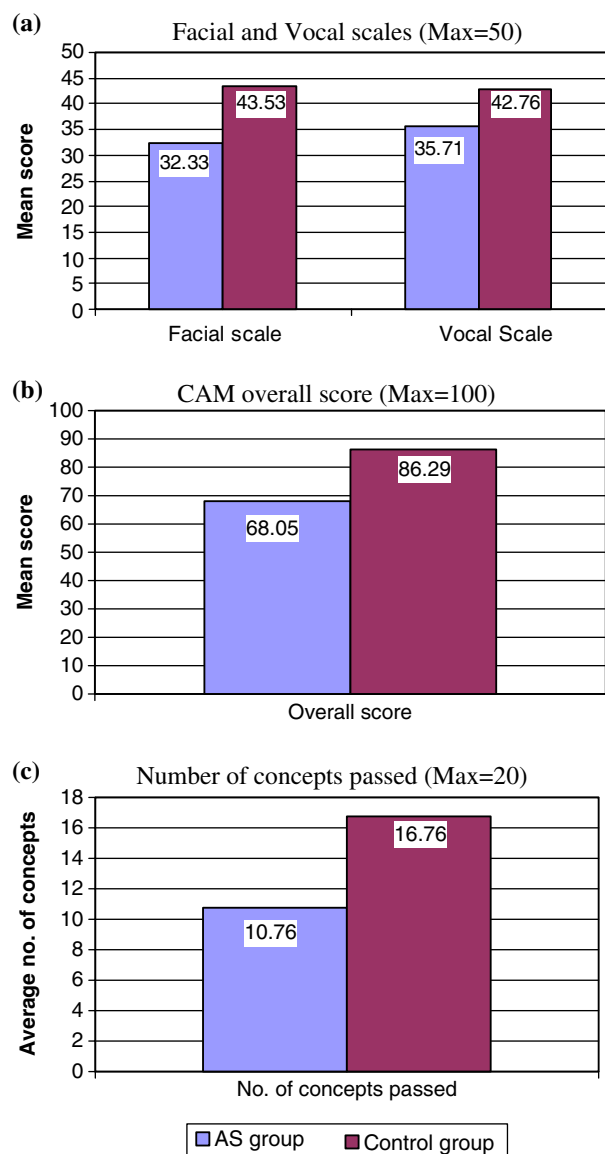


Fig. 3. Mean scores of AS and control groups on the four measures of the CAM.

To test for differences between the vocal and facial scale scores among the groups (modality effect) and between them (a modality by group interaction), a multivariate ANOVA for repeated measures was performed with the CAM facial and vocal scale scores as the dependent variables, group (AS/control) and sex as independent variables and verbal IQ, performance IQ and age as covariates. No effect of modality was found ($F_{\text{wilks}}[1,31] = 2.63$, n.s.), nor an interaction of group by modality ($F_{\text{wilks}}[1,31] = 0.39$, n.s.). A significant interaction of modality by sex was

found ($F_{\text{wilks}}[1,31] = 5.06$, $p < .05$), suggesting that females found it easier to recognise emotions from faces than from voices, whereas males showed the opposite preference. However, the analysis of simple main effects was not significant.

A three-way interaction of modality by group by sex was also found to be significant ($F_{\text{wilks}} [1,31] = 4.45$, $p < .05$). Analysis of simple main effects revealed that while there were no differences between recognition of emotions from faces and voices in the control group and among females with AS, there was indeed a significant difference among males with AS ($t[14] = 2.69$, $p < .05$), who found the recognition of emotions from faces ($M = 31.1$, $SD = 8.24$) significantly harder than recognition from voices ($M = 36.7$, $SD = 4.22$).

Next, in order to compare the recognition of individual concepts in the two groups, a 18 by 2 multivariate analysis of variance was performed for the proportions of participants who correctly recognised each concept in the two groups. Due to the large number of dependent variables and since groups were matched on sex, age and IQ, these variables were not included. Two concepts—*uneasy* and *appalled*, which all of the participants of the control group passed, were excluded from the analysis. The analysis yielded a significant overall effect of group ($F_{\text{wilks}}[18,19] = 2.60$, $p < .05$). Individual concept analyses revealed that the AS group scored significantly lower than the control group in the recognition of 11 of the 18 concepts. These concepts were: *intimate* and *reassured* from level 4 of the taxonomy; *distaste*, *insincere*, *lured*, *mortified*, *nostalgic*, *resentful*, *subservient* and *grave* of level 5; and *exonerated* of level 6. The two concepts, which were excluded from the analysis, were then analysed separately with a goodness of fit test. The proportion of participants with AS who correctly recognised *uneasy* was significantly lower than that of the control group ($\chi^2[1] = 8.2$, $p < .01$). There was no difference between the groups in the recognition of *appalled* ($\chi^2[1] = 0.831$, N.S.). A Wilcoxon signed ranks test was held to check the difference between proportion of concepts passed from levels 4 and 5 of the taxonomy (level 6 was excluded since it was represented by one concept only). No significant difference was found for either group ($Z = .47$ for the AS group and $Z = .12$ for the control group, $p > .5$ for both). The proportion of participants of the two groups who passed each of the 20 concepts is shown on Table IV.

Table IV. Proportion of Participants who Correctly Recognised the CAM 20 Concepts

Concept	AS group (<i>n</i> = 21) %	Control group (<i>n</i> = 17) %	
Appalled	95.2	100.0	^a
Appealing (asking for)	52.4	76.5	
Confronted	71.4	82.4	
Distaste	57.1	94.1	**
Empathic	76.2	76.5	
Exonerated	33.3	94.1	**
Grave	42.9	82.4	*
Guarded	42.9	52.9	
Insincere	28.6	88.2	**
Intimate	42.9	94.1	**
Lured	42.9	82.4	*
Mortified	66.7	94.1	*
Nostalgic	66.7	94.1	*
Reassured	42.9	82.4	*
Resentful	61.9	94.1	*
Stern	47.6	52.9	
Subdued	66.7	70.6	
Subservient	28.6	70.6	**
Uneasy	61.9	100.0	** ^a
Vibrant	81.0	94.1	

Notes: ***p* < .01.

**p* < .05.

^aThe analysis of these concepts was done using goodness of fit test due to a ceiling effect in the control group.

Power calculations for the different scales (with $\alpha = 0.01$) show that the scales are quite powerful in differentiating the group with AS from the control group. Power scores were $1-\beta = .996$ for the facial scale, .939 for the vocal scale, .989 for the tally of correctly recognised concepts and .999 for the CAM overall score. With a more conservative alpha level

($\alpha = .001$) the respective power scores were: .972 (facial), .761 (vocal), .933 (number of concepts recognised), and .993 (overall).

An examination of the correlation matrix, shown in Table V reveals that as predicted, the CAM is strongly and positively correlated with the “Reading the Mind in the Eyes” task (Baron-Cohen *et al.*, 2001) and the revised version of the “Reading the Mind in the Voice” Task (Golan *et al.*, submitted; Rutherford *et al.*, 2002). All of the CAM’s scores correlated positively with these external criteria. Unsurprisingly, the facial scale had a stronger correlation with the Eyes task ($r = .74, p < .001$) compared to the Voice task ($r = .49, p < .01$). Similarly, the vocal scale of the CAM correlated more strongly with the Voice task ($r = .62, p < .001$) than with the Eyes task ($r = .32, p < .05$). The two scales of the CAM maintained a strong correlation between themselves ($r = .57, p < .001$), as they did with the overall score and the number of correctly recognised concepts.

All of the CAM’s measures were negatively correlated with the AQ score, which means that the more autism spectrum characteristics one possesses, the lower one’s CAM scores. All correlations of the CAM scores with IQ or age of the participants were found to be non-significant, which suggests that the CAM measures are independent of both verbal and performance IQ, as well as chronological age.

DISCUSSION

This study introduces the CAM—a new battery for testing recognition of complex emotions and mental states in the face and the voice. It expands previous work that found difficulties

Table V. Correlation of the CAM Scores with Each Other, with External Criteria and with Background Parameters

	CAM—overall	Facial scale	Vocal scale	No. of concepts passed
Facial scale	.920**			
Vocal scale	.847**	.571**		
No. of concepts passed	.975**	.904**	.816**	
Reading the mind in the eyes	.632**	.737**	.324*	.647**
Reading the mind in the voice-R	.611**	.489**	.616**	.609**
AQ	-.574**	-.472**	-.563**	-.512**
AGE	-.039	-.122	.084	-.037
Verbal IQ	.235	.257	.143	.271
Performance IQ	.186	.184	.140	.176

Notes: **Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

in this domain among adults of normal intelligence, diagnosed with autism spectrum conditions (Baron-Cohen *et al.*, 1997; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen *et al.*, 2001; Rutherford *et al.*, 2002). However, unlike previous studies, the CAM battery allowed a test of the recognition of specific emotions and mental states as well as overall performance, and recognition in the two perceptual channels separately. It also tested recognition of complex emotions and mental states using films of faces rather than still pictures.

Results showed that individuals with Asperger Syndrome (AS), when compared to general population controls, had more difficulties in recognising mental states from both faces and voices. In addition, participants with AS recognised fewer mental state concepts than controls. In twelve out of the twenty emotions and mental states tested in the CAM, a significantly lower number of participants with AS successfully recognised the concept, compared to age- and IQ-matched controls.

The fact that controls were matched on chronological, verbal and nonverbal mental age, and the lack of correlations between the CAM scores and these factors, suggests the independence of complex emotions and mental state recognition from verbal and non-verbal ability. The strong negative correlation of the CAM scores with the participants' Autism Spectrum Quotient (AQ) score (Baron-Cohen *et al.*, 2001) supports the relevance of emotion and mental state recognition difficulties in high-functioning adults with autism spectrum conditions. These are important components of empathising (Baron-Cohen, 2003). Despite their ability to recognise basic emotions, such adults still find it hard to "mindread" complex mental states from faces and voices. The relatively high correlation of the CAM facial and vocal scales with the "Reading the Mind in the Eyes" (Baron-Cohen *et al.*, 2001) and "Reading the Mind in the Voice-R" (Golan *et al.*, submitted) tasks, respectively, provides the task with important measures of external validity. The CAM goes beyond these two tasks by using motion in the facial scale items and by allowing the opportunity to analyse individual concepts. Its power levels show it is sensitive to group differences across all scales and scores.

The CAM tests recognising emotions independent of weak central coherence (Frith, 1989) or executive function (Ozonoff, Pennington, & Rogers, 1991) because there is minimal context or planning, which burden working memory. However, like

almost everything, responses do of course require some minimal inhibition (the ability to go through all the answers and choose the best one) and working memory.

A review of the mental states with which the AS group had significant difficulties reveals no clear pattern: Of the concepts included in the CAM, the groups did not find positive emotions easier to recognise than negative ones; concepts known to 15–16 year olds were not recognised more easily than concepts known to adults. There was even no preference for recognition of more intense emotions in comparison to more subtle ones: a group difference was found for *distaste* and *mortified* which were both marked as "strong", whereas other, more subtle, mental states such as *subdued* or *empathic* did not produce a group difference. However, emotion valence and subtlety were not systematically studied in the CAM, and these could be studied in their own right in future studies.

The lack of a developmental level effect may be attributed to the AS group members having a verbal IQ within the normal range. In other words, individuals who are older than 18 (as the participants of this study were) would be expected to be familiar with these concepts. An alternative explanation is the relative similarity in complexity between level 4 and level 5 concepts (which the CAM comprises). In order to test for an effect of level, i.e. differential recognition of complex and simpler mental states, the task should compare concepts from the higher levels (4–6) with lower levels (1–3). A comparison of this kind, using still pictures of faces and pictures of the eyes and mouth region (Baron-Cohen *et al.*, 1997) revealed that whilst adults with autism and AS were able to detect basic mental states in the whole face, they were impaired at recognising complex mental states, and especially at recognising such mental states from the eyes alone.

Despite the small number of female participants in this study, a sex difference was found on the facial scale in this study, and this was independent of diagnosis: females recognised emotions in faces significantly better than males. This has been found in typically developed controls using the "Reading the Mind in the Eyes" (Baron-Cohen *et al.*, 2001) or the "Profile for Non verbal Sensitivity" (Baron-Cohen, 2003; Hall, 1984). The lack of studies investigating females with autism spectrum conditions calls for a thorough investigation of their profile. The absence of a sex difference on the vocal scale mirrors the lack of such a difference on the

“Reading the Mind in the Voice” task (Rutherford *et al.*, 2002).

The reason for the difference between facial and vocal emotion recognition when testing sex differences might actually derive from the linguistic nature of the vocal task. Our results show that males with AS had higher vocal scores than their facial ones. It might be that they used the content to guess the emotions, to compensate for their difficulties in emotion recognition (Grossman *et al.*, 2000).

Although the dependent variables in this study measured number of correctly recognised emotions, it is nevertheless interesting to examine the errors made by the groups. Due to the wide variety of emotions used as distracters, it was not possible to locate specific error patterns. However, there were some interesting anecdotal examples of errors made only by participants in the AS group. These errors were mainly about missing subtleties in face items and answering according to the content while ignoring the intonation in voice items. Though in most cases chosen distracters were of the same valence as targets, in some items participants in the AS group preferred a cross valence distracter upon the target answer (i.e. preferred a negative distracter when target was positive and vice versa). All examples of the errors quoted in the discussion section below were chosen by at least 30% of the participants with AS.

When looking at the group differences on individual concept recognition, the deficit among the AS group in recognising *insincerity* is most striking. Less than a third of the AS group members recognised this mental state successfully, which supports evidence from other studies, showing specific difficulty in understanding deception (Baron-Cohen, 1992; Sodian & Frith, 1992). One study found that high-functioning adults with autism had difficulties assessing the trustworthiness and approachability of people from pictures of their faces (Adolphs *et al.*, 2001). Their participants tended to evaluate people in the photos as being more trustworthy and approachable than controls' evaluations. Understanding of hidden intent, masked behind an incongruent facial expression, is a major area of difficulty for individuals with autistic spectrum conditions. In her book *Thinking in Pictures*, Temple Grandin described her experience as a high functioning adult with autism: “It is easy for me to understand the concept of deception when it involves playing tricks...but understanding the social cues that indicate an insincere person is much more difficult” (Grandin, 1995, p. 137). Other studies have also reported a specific

difficulty in matching incongruent faces and emotion labels among children with AS (Grossman *et al.*, 2000). It is possible that on the CAM items for *insincere*, the participants with AS were distracted by the content of what was said in the vocal items, rather than judging the intonation, which caused them to make errors. Similarly, they might have used more obvious cues in the face (e.g. the fact that the person in the film was smiling) to answer the facial items while missing other, more subtle facial cues (e.g. gaze that was incongruent with the smile). An example of such an error is the preference of the distracter “spellbound” when asked to recognise insincerity in one of the face items. The participants might have interpreted the smile and avoiding gaze as a spellbound rather than an insincere emotional state. Since deception is tested in the false belief tasks (which adults with AS passed with no difficulty), the CAM might provide a more subtle alternative to these tasks.

Another emotional concept the AS group had particular difficulty recognising was *subservient*. For example, two thirds of the AS group preferred the label “miffed” for a subservient face item (comparing to 11% in the control group). This could reflect their confusion between dominant and submissive characters in a social situation. Since dominance hierarchies are widely recognised in social primates (De Waal, 1998) it is surprising that people with AS should find this emotion difficult. However, it may be that non-human primates rely on other cues to judge dominance and subservience (e.g. physical size or success in conflict). It is likely that people with AS would have no difficulty in understanding social hierarchies from non-emotional cues (such as rank). It may therefore be that their deficit arises only when the cues are from emotional expression. Such misunderstanding of hierarchical human systems and social relations might, for example, lead to the use of an inappropriate attitude towards authority.

A similar problem might arise for the misunderstanding of *intimacy*—another clearly interpersonal emotion. The AS group had difficulties spotting the interpersonal aspect in this mental state. More than 40% of participants with AS mislabelled *intimate* face items as “determined” and “carefree”. Similarly, 30% of them mislabelled an intimate voice item as “subservient”, possibly relying on its content (“ask me anytime”) while not picking up the intonation. It is easy to imagine how such a cognitive deficit could lead to difficulties in relationships and to difficulties distinguishing genuine closeness and interest on one

hand, from politeness or even boredom in an encounter on the other. The AS group's difficulties in recognising *resentful*, *grave* or *mortified* also may reflect longstanding difficulties in reading emotions and mental states. The lack of group difference for mental states like *guarded* or *stern* may reflect these items being hard for both groups, since both groups scored relatively low on them. Further examination of these mental states and the groups they represent will be required. This applies especially to *guarded* and its group *disbelieving* due to its relevance to theory of mind and the understanding of deception.

Another mental state which the AS group had significant difficulties with was *exonerated*. A third of the participants in the AS groups mistook a face item of *exonerated* for "remote", which is a cross valence distracter. Similarly, they confused the positive *exonerated* for a negative label, "resigned" in a voice item, again using the content only ("now I can get down to work"). This concept was the only remaining representative of level 6 in the taxonomy (emotion words which were known to less than 75% of adult participants in the validation study (Baron-Cohen *et al.*, submitted)). Nevertheless, analysis revealed no significant effect of verbal IQ in the ability to recognise this emotion. This mental state (which concerns relief from blame, like *reassured* which also had a significant group effect), requires the use of a theory of mind for its interpretation, which might have made its recognition especially hard for the AS group.

One of the emotional concepts which was recognised by all members of the control group but only 62% of the AS group is *uneasy*. This emotion, expressing very subtle degrees of fear, is a good example of the difficulty individuals with AS might have with picking up subtle emotional cues. A new fMRI study of the amygdala, a key brain area underlying the detection of fear in others, involved showing pictures of people expressing fear at different intensities to individuals with AS and controls (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, submitted). Amygdala activation among the participants with AS was significantly reduced relative to controls. Furthermore, the AS group showed no difference in amygdala activation between intense, moderate and mild degrees of fear. This study supports the amygdala theory of autism (Adolphs *et al.*, 2001; Baron-Cohen *et al.*, 2000; Howard *et al.*, 2000) and shows that the brain in high-functioning adults with autistic spectrum conditions does not respond differentially to expressions

of subtle fear. Our study is consistent with these findings. Future studies should investigate amygdala activity in individuals with AS not only when watching but also when listening to fear items of different intensity (e.g. uneasy, afraid, terrified).

Another significant difference in recognition between the AS and control groups was with the concept *distaste* from the *disgusted* group in the taxonomy. This high-intensity concept, was recognised by nearly all of the controls (94.1%) but only by a little more than half of the participants with AS (57.1%). More than a third of the participants in the AS group mislabelled a face item of this emotion as 'offended' and two thirds of them mislabelled a voice item as "battered" ("you've done it again"). Surprisingly, not many studies have studied recognition of disgust. Those which have, found difficulties in disgust recognition among severely autistic adolescents with mental retardation (Hobson, Ouston, & Lee, 1988) and no difficulty in its recognition among high-functioning adults with autistic spectrum conditions (Adolphs *et al.*, 2001). The group difference we found in the recognition of faces and voices expressing *distaste* (which is slightly more subtle than *disgust*) suggests that even high-functioning individuals with AS might have problems recognising this in faces and voices. So far, no brain mapping studies have assessed the recognition of disgust by individuals with autism spectrum conditions. However, studies of patients with insula and basal ganglia lesions report specific difficulties with the recognition and experience of disgust, suggesting brain specificity for this emotion (Calder, Keane, Manes, Antoun, & Young, 2000; Calder, Lawrence, & Young, 2001).

Interestingly, there was no difference between the AS and control groups in the recognition of empathy (the concept *empathic*), and this was not predicted. The reason for this might have been the easy foils which appeared with this concept. It will be interesting to explore other emotions in the *kind* group, to test this concept more thoroughly. Equally surprising was the lack of difference in recognising mental states such as *appalled*, *vibrant*, *confronted* and *subdued*. These results suggest that in such adults with AS, their "mindblindness" is by no means total.

The CAM presents both visual and vocal stimuli, including motion in the face task and requiring word labels for the emotion. As these different kinds of stimuli activate different brain areas, one might wonder which of these underlies the difficulties of the AS group. Such studies were mainly done with typically developed and with brain damaged

participants. In a recent study assessing a patient with bilateral amygdala lesions, Adolphs and colleagues found that while the patient could not recognise the 6 “basic” emotions from still faces, he could easily recognise them when they were expressed by a live model. They argued that still and motion emotional stimuli activate different areas of the brain: while the fusiform gyrus, as well as the posterior and superior temporal cortex and the amygdala and insula are involved in recognising emotions from still faces, emotion recognition from faces in motion is related to middle temporal and middle superior temporal areas of the brain, together with parietal and frontal sectors (Adolphs, Tranel, & Damasio, 2003).

Kilts and colleagues conducted a PET imaging study assessing intensity of anger and happiness from still and video stimuli. They found that judgment of anger in dynamic expressions was associated with increased right-lateralised activity in the medial, superior, middle, and inferior frontal cortex and cerebellum, while judgments of happiness were associated with relative activation of the cuneus, temporal cortex, and the middle, medial, and superior frontal cortex. In contrast, the perception of anger or happiness in static facial expressions activated a motor, prefrontal, and parietal cortical network (Kilts, Egan, Gideon, Ely, & Hoffman, 2003).

As with visual stimuli, there are different findings with regards to emotion recognition from the voice. In a study involving over a hundred participants with focal brain damage, Adolphs and colleagues found that lesions in right cortical regions impair recognition of emotion in prosody. Phillips and colleagues found in a functional MRI study that the amygdala, as well as the superior temporal gyrus, were involved with recognition of fear and disgust in both visual and vocal stimuli (Phillips *et al.*, 1998). However, Adolphs and Tranel found no difficulty in recognising emotions from prosody among amygdala lesioned patients (Adolphs & Tranel, 1999). Evidently, this calls for further inquiry of brain structures involved in emotion recognition from voices.

The above mentioned studies assessed emotion recognition using the 6 “basic” emotions. Further investigation into activation of brain regions in recognition of complex emotions from still, dynamic and vocal stimuli, such as those that appear in the CAM is still required. Since this study was purely behavioural, no conclusions could be drawn of relevant brain regions involved in the recognition of the CAM’s different emotions. Work is under way in our lab to investigate the neural basis underlying

dynamic emotion recognition using the videos from the CAM in a functional MRI experiment involving adults with and without autism.

The CAM includes only 20 out of the full set of 412 emotions and mental states in the emotion taxonomy. Our team is currently working on a version for children (Golan & Baron-Cohen, in preparation). Future studies using the taxonomy and the multimedia database could explore each particular emotion group in depth, and compare certain emotions in different intensities. Temple Grandin mentioned the visual collection of social “clips” she holds in her mind for reference, which help her cope with interpersonal situations (Grandin, 1995). Matching facial and vocal expressions of mental states could be used to teach individuals with autism spectrum conditions to improve their emotion recognition. We are currently engaged in such an intervention study.

We conclude by suggesting that this new test of subtle emotion recognition in the face and the voice reveals that adults with HFA or AS have residual difficulties in recognising complex emotions and mental states, independent of IQ, language, central coherence, or executive function.

ACKNOWLEDGMENTS

OG was supported by the Corob Charitable Trust, the Cambridge Overseas Trust and the National Alliance for Autism Research (NAAR). SBC and JH were supported by the Shirley Foundation, Medical Research Council (MRC), and the Three Guineas Trust. We are grateful to our colleagues at Red Green and Blue Ltd, Autism West Midlands, Autism London, European Services for People with Autism (ESPA), and to Chris Ashwin, Sally Wheelwright, Yael Golan and Sarah Johnson.

REFERENCES

- Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience*, *13*, 232–240.
- Adolphs, R., & Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia*, *37*, 1285–1292.
- Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, *52*, 61–69.
- American Psychiatric Association. (1994). *DSM-IV Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: American Psychiatric Association.

- Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (submitted). Differential activation of the social brain during emotion processing in adults with autism.
- Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology and Psychiatry*, 33, 1141–1155.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6, 248–254.
- Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a “language of the eyes”? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition*, 4, 311–331.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5–17.
- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.
- Baron-Cohen, S. (2003). *The Essential Difference: Men, women and the extreme male brain*. London: Penguin.
- Baron-Cohen, S., Golan, O., Hill, J. J., & Wheelwright, S. (submitted). Developmental changes in the emotion lexicon: a research note.
- Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). *Mindreading: The interactive guide to emotions*. London: Jessica Kingsley Limited (www.jkp.com).
- Baron-Cohen, S., Hill, J. J., Golan, O., & Wheelwright, S. (2002). Mindreading Made Easy. *Cambridge Medicine*, 17, 28–29.
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry*, 38, 813–822.
- Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. (2000). The amygdala theory of autism. *Neuroscience and Biobehavioral Reviews*, 24, 355–364.
- Baron-Cohen, S., Spitz, A., & Cross, P. (1993). Can children with autism recognize surprise? *Cognition and Emotion*, 7, 507–516.
- Baron-Cohen, S., Wheelwright, S., Hill, J. J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241–251.
- Bormann-Kischkel, C., Vilsmeier, M., & Baude, B. (1995). The development of emotional concepts in autism. *Journal of Child Psychology and Psychiatry*, 36, 1243–1259.
- Buitelaar, J. K., Van derVan der Wees, M., Swabb Barneveld, H., & Van derVan der Gaag, R. J. (1999). Theory of mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and Psychopathology*, 11, 39–58.
- Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, 3, 1077–1078.
- Calder, A. J., Lawrence, A. D., & Young, A. W. (2000). Neuropsychology of fear and loathing. *Nature Neuroscience*, 2, 352–363.
- Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology and Psychiatry*, 33, 1169–1182.
- De Waal, F. B. M. (1998). *Chimpanzee politics: power and sex among apes* (Rev. ed.). Baltimore, London: Johns Hopkins University Press.
- Ekman, P., & Friesen, W. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124–129.
- Forster, K. L., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods Instruments & Computers*, 35, 116–124.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.
- Golan, O., Baron-Cohen, S., Rutherford, M. D., & Hill, J. J. (submitted). The Reading the Mind in the Voice test–Revised: a study of complex emotion recognition in adults with and without Autism Spectrum Conditions.
- Grandin, T. (1995). *Thinking in pictures*. New York: Doubleday.
- Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal bias in recognition of facial emotions in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 369–379.
- Hall, J. A. (1984). *Non-verbal sex differences: Communication, accuracy and expressive style*. Baltimore; London: Johns Hopkins University Press.
- Happe, F. (1994). An advanced test of theory of mind: Understanding of story characters’ thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129–154.
- Harris, P. L. (1989). *Children and emotion: The development of psychological understanding*. Oxford: Blackwell.
- Hobson, R. P. (1986a). The autistic child’s appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry*, 27, 321–342.
- Hobson, R. P. (1986b). The autistic child’s appraisal of expressions of emotion: A further study. *Journal of Child Psychology and Psychiatry*, 27, 671–680.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: co-ordinating faces and voices. *Psychological Medicine*, 18, 911–923.
- Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., & Roberts, N. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport*, 11, 2931–2935.
- Jolliffe, T., & Baron-Cohen, S. (1999). The Strange Stories Test: A replication with high-functioning adults with autism or Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 29, 395–406.
- Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage*, 18, 156–168.
- Kleinman, J., Marciano, P. L., & Ault, R. L. (2001). Advanced theory of mind in high-functioning adults with autism. *Journal of Autism and Developmental Disorders*, 31, 29–36.
- Loveland, K. A., Tunali Kotoski, B., Chen, R., & Brelsford, K. A. (1995). Intermodal perception of affect in persons with autism or Down syndrome. *Development and Psychopathology*, 7, 409–418.
- Loveland, K. A., Tunali Kotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., & Gibes, M. C. (1997). Emotion recognition in autism: Verbal and non-verbal information. *Development and Psychopathology*, 9, 579–593.
- Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, 9, 491–500.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, 32, 1081–1105.
- Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., Williams, S. C., Bullmore, E. T., Brammer, M., & Gray, J. A. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society Series B-Biological Sciences*, 265, 1809–1817.
- Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults

- and adults with Asperger syndrome and high functioning autism. *Journal Autism and Developmental Disorders*, 32, 189–194.
- Sodian, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Journal of Child Psychology and Psychiatry*, 33, 591–605.
- Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence (WASI), *The Psychological Corporation*. UK.
- Wellman, H. M. (1992). *The child's theory of mind*. Cambridge, Mass/London: The MIT Press.
- World Health Organisation. (1994). *ICD-10—International classification of diseases* (10th ed.). Geneva, Switzerland: World Health Organisation.
- Yirmiya, N., Sigman, M. D., Kasari, C., & Mundy, P. (1992). Empathy and cognition in high-functioning children with autism. *Child Development*, 63, 150–160.