Voice Processing Abilities in Children with Autism, Children with Specific Language Impairments, and Young Typically Developing Children

Jill Boucher

University of Warwick, Coventry, U.K.

Vicky Lewis

The Open University, Milton Keynes, U.K.

Glyn M. Collis

University of Warwick, Coventry, U.K.

It is well established that people with autism have impaired face processing, but much less is known about voice processing in autism. Four experiments were therefore carried out to assess (1) familiar voice-face and sound-object matching; (2) familiar voice recognition; (3) unfamiliar voice discrimination; and (4) vocal affect naming and vocal-facial affect matching. In Experiments 1 and 2 language-matched children with specific language impairment (SLI) were the controls. In Experiments 3 and 4 language-matched children with SLI and young mainstream children were the controls. The results were unexpected: the children with autism were not impaired relative to controls on Experiments 1, 2 and 3, and were superior to the children with SLI on both parts of Experiment 4, although impaired on affect matching relative to the mainstream children. These results are interpreted in terms of an unexpected impairment of voice processing in the children with SLI associated partly, but not wholly, with an impairment of cross-modal processing. Performance on the experimental tasks was not associated with verbal or nonverbal ability in either of the clinical groups. The implications of these findings for understanding autism and SLI are discussed.

Keywords: Autistic disorder, emotional expression, facial perception, specific language impairment, voice processing.

Abbreviations: SLI: specific language impairment.

General Introduction

Voices and faces are the most important sources of information enabling infants to learn about other people. Voices and faces are also the most important channels for nonverbal communication between infants and their carers. Not surprisingly, therefore, some voice and face processing abilities are functional in infants at birth, and others develop early. With regard to voice processing, neonates respond preferentially to voices (Eisenberg, 1976) and can recognise their mother's voice (DeCasper & Fifer, 1980). Week-old babies can recognise the intonational patterns of the language spoken by their mother (Mehler et al., 1988), and by the age of 6 months normal infants respond to the affective content of vocal tones as well as facial expressions (Walker-Andrews, 1988). From birth, infants signal different types of distress or discomfort by discriminable patterns of reflexive crying (Wolff, 1969). Two-month-old infants respond to their mother's face and voice with increased vocalisation, smiling, and body movement (Brazelton, Koslowski, & Main, 1974). By 6 months, vocal turn-taking and proto-conversations are well established, preceding the onset of speech-like babble in the second half of the first year (Locke, 1993).

In contrast to the typically developing child, the child with autism lacks normal social awareness and nonverbal communication skills, and it has been suggested that there is a fundamental impairment in the ability to process social stimuli, an impairment that is variously understood in terms of the perceptual processing of social stimuli (Tantam, 1992), or the sociocognitive mechanisms of intersubjectivity (Kanner, 1943; Hobson, 1993). Studies of face processing abilities in autism do in fact demonstrate perceptual encoding abnormalities, and impairments of recognition, expression analysis, facial speech analysis, and directed visual processing (Boucher & Lewis, 1992; Boucher, Lewis, & Collis, 1998; Davies, Bishop, Manstead, & Tantam, 1994; MacDonald et al., 1989). However, the impairments that have been demonstrated are not robust. For example, despite there being significant differences between children with autism and controls on tests of face recognition, there is considerable overlap between groups. Similarly, difficulties in interpreting basic facial expressions are only evident when children with autism are compared with controls

matched for nonverbal, as opposed to verbal, ability. However, difficulties in interpreting complex facial expressions such as guilt or embarrassment are evident regardless of the matching procedure used (Capps, Yirmaya, & Sigman, 1992; Fein, Lucci, Braverman, & Waterhouse, 1992).

Much less is known about voice processing abilities in autism. However, such research as has been carried out suggests that voice processing is impaired, possibly to a greater extent than face processing. For example, Klin (1991, 1992) found that young children with autism differed from young typically developing children and from children with learning difficulties in that they did not show a preference for listening to their mother's voice as opposed to other speech stimuli. Ricks (1975) found that very young children with autism have idiosyncratic patterns of emotionally toned crying, and several studies have shown that prosody is abnormal in autism (e.g., Baltaxe & Guthrie, 1987; Fine, Bartolucci, Ginsberg, & Szatmari, 1991). Hobson, Ouston, and Lee (1989) showed impaired ability to name vocally expressed emotion, and three studies have shown impaired ability to match vocally and facially expressed emotion (Hobson, Ouston, & Lee, 1988; Ozonoff, Pennington, & Rogers, 1990; VanLancker, Cornelius, & Kreiman, 1989). Consistent with these latter findings, de Gelder, Vroomen, and van der Heide (1991) showed abnormal visual and auditory integration in facial speech processing.

In an earlier study (Boucher et al., 1998), we found that school-age children with autism have significant difficulty relative to a mixed group of language-delayed children in matching familiar voices and faces. The children with autism were also impaired, relative to these controls, in recognising familiar faces. A subset of the children with autism were impaired in recognising familiar voices, relative to a subset of the controls who had specific language impairments rather than language delay associated with general learning difficulties (mental retardation). We concluded that impaired familiar voice-face identity matching resulted from a combination of impaired face recognition and impaired voice recognition.

However, it was not possible to rule out the possibility that a cross-modal processing deficit contributed to impaired voice-face matching in our experiment. There is some limited evidence of cross-modal processing impairments in autism, Bryson (1972) having reported less good performance by children with autism on a task in which a spoken word such as "red" had to be matched to a red object such as a brick, as compared to performance in a condition in which two red objects had to be matched with each other. A cross-modal processing deficit might also be contributing to the impairments in matching vocally and facially expressed affect that have been shown in several experiments, and to the impairment of facial speech analysis reported by de Gelder et al. (1991). It therefore seemed important to assess voice-face identity matching and voice-face affect matching again, using methods designed to assess the possible role of a cross-modal processing deficit. In addition, in our 1998 study only subsets of the experimental and control groups took part in the test of familiar voice recognition. This was because several children in the original groups failed to reach the training criterion, or showed negative reactions during testing and were therefore excluded. The groups tested were therefore small, and it seemed important to test familiar voice recognition again in larger groups of

children, using a simplified procedure. Finally, there have been no studies assessing unfamiliar voice discrimination in children with autism. Because perceptual or discrimination difficulties could be causing or contributing to the impairments of voice-face identity and affect matching demonstrated in other experiments, and to the impairment of familiar voice recognition demonstrated in our own previous work, it seemed important to extend earlier findings by including a test of unfamiliar voice discrimination.

In sum, the central aim of the present study was to replicate and extend earlier findings on voice processing in autism, and to clarify the interpretation of any impairments. To achieve this central aim we carried out four experiments, the specific aims of which are outlined below, with predictions concerning expected outcomes.

In Experiment 1 we repeated the assessment of familiar voice-face matching, and added an assessment of the ability to match nonsocial objects and sounds. The aims of this experiment were to (1) replicate the earlier findings on familiar voice-face matching in children with autism; (2) to assess the specificity of a voice-face matching impairment; and (3) to clarify whether or not a cross-modal matching impairment might contribute to this impairment. We predicted (1) that children with autism would be impaired at matching familiar voices and faces; (2) that voice-face matching would be more impaired in the children with autism, relative to controls, than sound-object matching. We made no prediction concerning (3) the presence/absence of evidence of a cross-modal matching impairment.

In Experiment 2 we repeated the assessment of familiar voice recognition. The aim of this experiment was to confirm our earlier finding using larger groups of participants and a simplified procedure. We predicted that the children with autism would show impaired recognition of familiar voices.

In Experiment 3 we assessed the ability to discriminate between unfamiliar voices. The aim of Experiment 3 was to plug a gap in the data set on voice processing in autism. We predicted that we would find no impairment of voice discrimination. This prediction rested on three observations. First, face discrimination is unimpaired in autism (Boucher & Lewis, 1992; Boucher et al., 2000). Second, there are frequent parallels between voice and face processing abilities in typically developing individuals (Mann, Diamond, & Carey, 1979). Third, at some stage of development many children with autism produce echolalia, which mimics vocal as well as speech characteristics quite accurately.

In Experiment 4 we assessed vocal-facial affect matching and vocal affect naming. The aims of this experiment were (1) to confirm previous findings; and (2) to assess the possible role of a cross-modal matching impairment as a cause, or contributory cause, of a voice-face affect matching impairment. On the basis of findings of earlier studies (Hobson et al., 1988, 1989) we predicted (1) that the children with autism would be impaired on vocal-facial affect matching and vocal affect naming. If a cross-modal matching impairment contributes to the impairment of voice-face affect matching, then voice-face affect matching in children with autism will be more impaired, relative to controls, than vocal affect naming. As with the test of the possible role of cross-modal matching impairments in familiar voice-face identity matching (Experiment 1), we made no prediction concerning (2) voice-face affect matching relative to vocal affect naming.

Two subsidiary aims of the overall study were to assess (1) relationships between task performance and the background variables of age, verbal ability, nonverbal ability, and number of terms in school within each of the experiments, and (2) relationships between task performances across the four experiments. The analyses of these relationships is reported in a single section following the reports of the individual experiments.

The Experiments

Participants

Experiments 1 and 2 constituted tests of children's ability to recognise familiar voices (and faces). There is no pool of famous voices and faces that can safely be assumed to be familiar to all children with autism, and in previous work we overcame this problem by selecting an experimental group and a language-matched control group from schools that cater both for children with autism and children with language impairments. This enabled us to use as stimuli the voices and faces of people within the school who were well known to both the children with autism and the children with language impairments. In the present study we used this method again. A consequence of this is, however, that a typically developing control group cannot be included, as there are no schools in the U.K. that cater for significant numbers of children with autism, children without autism but with low language abilities, and typically developing children. If typically developing children from a mainstream school were to be tested, any difference between the typically developing children's performance and the performance of the clinical groups would be conflated with possible differences in material. In Experiments 1 and 2, therefore, we did not include a typically developing control group. Experiments 3 and 4 did not, however, involve using familiar voices or faces as material, and in both these experiments we included a typically developing language-matched control group.

A second consequence of the methodological difficulty outlined above is that the choice of schools in which to work is extremely limited. We had to return to the same schools which had taken part in the previous study (but we did not test any of the same children). However, the intake policy of one of the schools had changed, and this school was phasing out its intake of children with non-specific language delay associated with overall learning difficulties, and concentrating its intake on children with autism and children with specific language impairments. The second school in which we had previously worked caters exclusively for children with autism and children with specific language impairments. In all four experiments reported in this paper, therefore, we compared children with autism to children with specific language impairments, rather than with the mixed group of children with poor language who took part in the previous study.

In sum, a group of 19 children with autism and a group of 19 children with specific language impairments (SLI), took part in all four experiments, and an additional group of 19 young mainstream children took part in Experiments 3 and 4.

The children with autism and the children with SLI came from two nonresidential special schools, School A and School B, which cater for children with communication difficulties, and in particular for children with autism or SLI. Seven of the children with autism attended School A, and 12 attended School B. Eight of the children with SLI attended School A, and 11 attended School B. Children in the two groups were equated for chronological age, for language ability as assessed on the Test for the Reception of Grammar (TROG) (Bishop, 1989), for nonverbal reasoning ability as assessed on the Coloured Progressive Matrices (Raven, 1990), and for the number of terms they had attended that school. No child was included in the study who had sensory or motor difficulties, or who was unwilling to cooperate. Details of participants are shown in Table 1.

The children with autism had been all been diagnosed as autistic by independent psychiatrists or psychologists, and were attending Schools A and B because of their autism. Their current behaviour met DSM-IV (American Psychiatric Association, 1994) criteria for autistic disorder. None of the children with autism met the criteria for Asperger's syndrome. However, the children's non-verbal reasoning ability was within the normal range (see Table 1).

The children with SLI had all been diagnosed by speech and language therapists as having specific speech and language impairments, and were attending Schools A and B because of their language acquisition difficulties. Both School A and School B had speech and language therapists working on site, and the children with SLI were selected in consultation with the therapists and Head Teachers of the schools. None of the children were considered by educational psychologists to have signifi-

Table 1
Participant Details

Groups	Male:Female	CA (yrs; mths) Mean	(SD)	TROG verbal MA (yrs; mths) Mean	(SD)	Matrices nonverbal MA (yrs; mths) Mean	(SD)	Terms in school Mean	(SD)
Autism (Expts 1–4)	16:3	9;7	(1.0)	5;11	(1.2)	8;9	(1.2)[a]	11;0	(4.3)
SLI (Expts 1–4)	12:7	9;0	(1.4)	6;1	(1.2)	8;5	(1.8)[a]	9;0	(5.3)
Mainstream (Expts 3 & 4)	10:9	6;4	(0.8)	5;10	(1.5)			3;0	(1.8)

[a] These means and standard deviations have been calculated by converting raw scores at the extreme ends of the distribution (for which mental ages are not given in the conversion table) using extrapolation.

cant learning difficulties. None of the children with SLI had a significant hearing loss, or socioemotional problems, according to their teachers' and speech/language therapists' reports. Children with so-called "semantic-pragmatic language disorder" were excluded. Children with severe comprehension difficulty were also excluded. Children included in the SLI group thus had a range of predominantly expressive phonological, grammatical, and semantic impairments.

The 19 young mainstream children who took part in Experiments 3 and 4 were recruited from the first two classes of a nonresidential mainstream primary school. They were selected according to whether or not parental permission for participation in the study was given and on the basis of their TROG scores. We aimed to select a group of children with average language ability, whose mean language ability and range of language abilities corresponded to the means and ranges of the two clinical groups. None of the mainstream children were perceived by their teachers as having any special problems.

Materials and Procedures

The children were tested individually in a quiet room in school after a minimum of two visits, one for familiarisation and one in which language ability and, for both clinical groups, nonverbal ability, were assessed. The four experiments were carried out over a number of weeks in the order in which they are reported, with breaks between each experiment. The children with autism and the children with SLI were tested by one person, and the young typically developing children by another.

In each experiment, a training phase preceded experimental testing. The materials used in training were selected so as to pose no challenge to the child, in order to enable us to assess their understanding of the task demands per se, independent of their ability to succeed when presented with the actual test material. Children had to reach specified training criteria in order to be included in the study.

In each experiment, some training items were inserted among the test items. Because the training items had been selected as being particularly easy to recognise/discriminate etc., and because children were only included in an experiment if they reached criterion on training, the insertion of training items among the test items provided a means of checking that the children remained on task. In addition, these items are likely to be easy for the children to respond to correctly, helping to maintain motivation and attention. Children's responses to the inserted training items were not included in test scores.

Experiment 1: Familiar Voice-face and Sound-object Matching

Method

For each school the following materials were prepared.
Training stimuli. Two sets of training stimuli were prepared, one for voice-face matching and one for sound-object matching. For voice-face training, three colour photographs measuring 10 cms × 12 cms and mounted on cards were prepared. The photographs were of highly familiar and/or distinctive individuals associated with the school. Individuals were photographed looking into the camera. Clothing was concealed by a dark cloth. However, accessories such as earrings and spectacles were not removed. In addition, audio tapes were prepared, on which each individual whose face appeared in the training set

was recorded talking about going shopping or what they would do if they won the Lottery. Each voice sample lasted approximately 30 seconds. For sound-object matching, training materials were prepared as for voice-face matching, using familiar objects from around the school. In addition a mounted photograph of the school was prepared.
Test stimuli. Two sets of test stimuli were prepared, one for voice-face matching and one for sound-object matching. For voice-face matching, 21 photographs similar to the training photographs but of different people were used as test stimuli. A separate audio tape was prepared for each individual whose photograph was used, recorded as for the training stimuli. For sound-object matching, the test stimuli consisted of 21 photographs and audio-recordings of familiar objects from around the school, excluding those used in training.
Procedure. The training and testing for the two conditions of voice-face matching and sound-object matching were completed in a single session, the presentation order of the two conditions being balanced within each group. The 21 test photographs with copies of the 3 training photographs interspersed among them were laid out on a table in a standard 6 × 4 display and covered by a cloth. The photograph of the child's school was placed face up on top of the cloth. The child was first asked to identify the photograph of the school. They were then asked to fix this photograph on the top of an empty cassette holder and was told that this was called "the (name) school box", with the name of the child's school being emphasised. The experimenter then placed the three training photographs (of either faces or objects) in front of the child and asked the child to name the person/object in each photograph. Help was given if necessary. The experimenter then explained to the child that she had tape recordings of the people talking/the noises that the objects made. The first recording was played and the child was instructed to point to the photo of the person talking/the object making the noise. If the child pointed to the correct photograph the next recording was played. If the child chose the wrong photograph the experimenter corrected the child and moved on to the next recording. This was continued, using a different order of voices/noises, until the child made three correct consecutive responses without guidance from the experimenter or until each voice/noise had been played three times. All children tested met the criterion of three consecutive correct responses within the nine presentations.

The experimenter then explained to the child that she had lots of photographs of different people/different objects from the (name) school and, pointing to a pile of cassette tapes, that she had tape recordings of the people talking/the noises that the objects made. She explained that when one of the tapes was played the child had to look at all the photographs and find the person who was talking/what was making the noise and put the photograph and tape into the (name) school box. The cloth was then removed from the array of 24 photographs. The child was helped to name the people/objects in the photographs and to say something about them to ensure that they recognised the people/objects in the photographs. Each tape was played in turn and the child's task was to select the person talking/the object making the noise. If the child made a mistake, a correct response was elicited, using prompts. Each tape and its corresponding photograph were removed and placed in turn in the (name) school box. A proportion of correct responses were praised, and comments such as "Here's an easy one", "Not many more to do" and so on were used to maintain the child's motivation. A record was kept of the child's initial correct and incorrect responses.

Results

All children were correct on all the inserted training items, indicating that they remained on task throughout testing. Responses to the final three test recordings were excluded from the analysis, since as the number of alternatives was reduced, it was increasingly likely that the

correct response could have been made by guessing. The number of correct matches out of the remaining 18 voice-face and 18 sound-object matches made by each child were noted. Mean scores for voice-face matching and sound-object matching in the autism group were, respectively, 13.21 (*SD* 2.9) and 13.79 (*SD* 2.55); and the corresponding scores in the SLI group were 12.68 (*SD* 3.33) and 14.95 (*SD* 1.78). A three-way analysis of variance (groups, schools, and conditions) showed no significant main effects of either group or school ($F < 1$ in both cases). The effect of condition was significant, $F(1, 34) = 8.4$, $p = .007$, sound-object matching being superior to voice-face matching. The interaction between group and condition approached significance, $F(1, 34) = 3.07$, $p = .089$. The superiority of sound-object matching over voice-face matching was clear for the children with SLI, $t(18) = 3.86$, $p = .001$, but not in the children with autism, $t(18) = 0.76$, n.s. No other interaction approached significance. In the SLI group, performance on voice-face matching correlated with performance on sound-object matching (Pearson $r = .65$, $p = .002$). Within the autism group the correlation was much smaller ($r = .25$, n.s.).

The results compared to the predictions. We had predicted (1) that we would replicate the earlier finding of impaired voice-face identity matching in children with autism; and (2) that voice-face matching would be more impaired in the children with autism, relative to controls, than sound-object matching. We made no prediction concerning (3) whether or not we might find evidence of a cross-modal matching impairment. In the event, prediction (1) was not confirmed: the children with autism were not impaired relative to children with SLI. Nor was prediction (2) confirmed, the children with SLI showing rather larger within-subject differences than the children with autism in favour of sound-object matching. With regard to prediction (3), we argued that if children's performance on a voice-face matching task is affected by cross-modal matching impairments, then performance on social and nonsocial matching tasks will be correlated. This was not the case for the children with autism, suggesting that these children do not have a cross-modal matching impairment. However, in the case of the children with SLI there was a significant correlation, suggesting that cross-modal processing impairments may be present. This finding will be considered in the overall Discussion.

Experiment 2: Familiar Voice Recognition

Method

Training stimuli. For each school 12 training tapes of the following voices/sounds were used: the three familiar voices used in training in Experiment 1; three distinctive unfamiliar voices not previously heard; the three familiar sounds used in training in Experiment 1; and three unfamiliar computer-generated sounds not previously heard.

Test stimuli. For each child 18 audio tapes were prepared, of which 9 tapes were of voices of people familiar to them in their own school (including their teacher, their classroom assistant(s), and classmates). The remaining nine voices were of unfamiliar people (staff and children from the other school). All the recordings were of people talking about going shopping or what they would do if they won the Lottery.

Procedure. Training began with the experimenter showing the child the cassette holder with a picture of their school on it which had been used in Experiment 1, and reminding the child that all the tapes of people and things in the child's school

belonged in the holder. A waste paper bin was then introduced as "the rubbish box", which was for tapes of people or sounds they did not know ("for 'rubbish' tapes"). The positions of "the (name) school box" and "the rubbish box" on the table to the child's right or left were alternated across children within each group. The experimenter then showed the child a pile of tapes she had on a table to one side, explaining that they had got muddled up and needed to be sorted out into the correct boxes. She asked the child to help her.

The training tapes were played one by one in a prespecified order. For the first six training tapes the experimenter guided the child as to whether the voice/sound was familiar and so belonged in the named school box, or unfamiliar and therefore belonged in the rubbish box. On the second six training trials the experimenter provided guidance (if necessary) as to whether the voice/sound was familiar or unfamiliar but gave no guidance on where to place the tapes. The training criterion was the correct placement of five out of the last six training tapes. If the child reached this criterion testing began. If the child failed to reach this criterion, the training procedure was repeated in a later session with the tapes being played in the reverse order. All children tested met the criterion either on the first or second presentation of the training tapes.

Testing was carried out in a later session, normally within 2 days of the training. At the outset of testing the child was reminded of the training session and the function of "the (name) school box" and "the rubbish box". The child was then shown a stack of 28 tapes which the experimenter needed help to sort out and place in the correct boxes. The 28 tapes consisted of the 18 test tapes and 10 of the training tapes interspersed among them. Each tape was then played in turn and the child was asked to place the tape in the correct box. Errors were not corrected although praise and encouragement were given as in the other experiments. The child's correct and incorrect placements were recorded.

Results

Responses on the inserted training items indicated that all children remained on task throughout testing. The numbers of voices, out of the 18 test voices, correctly identified as familiar or as unfamiliar by each child were noted. For the children with autism, the mean overall score was 14.58, *SD* 2.3 (mean 7.58, *SD* 1.5 for the familiar voices; mean 7.0, *SD* 1.86 for the unfamiliar voices). For the children with SLI, the overall mean score was 14.42, *SD* 2.9 (mean 7.26, *SD* 1.91 for the familiar voices, and mean 7.16, *SD* 1.86 for the unfamiliar voices). A two-way ANOVA (schools × groups) on the total numbers of voices correctly recognised showed no main effects (for both schools and groups, $F < 1$), and no interaction.

The results compared to the predictions. For this experiment we predicted that familiar voice recognition would be impaired in the children with autism relative to controls. This prediction was not supported, the two groups performing very similarly to each other.

Experiment 3: Unfamiliar Voice Discrimination

Method

Test stimuli. The test stimuli consisted of 20 sets of 3 tapes. Within each set, two of the tapes were of one person talking, the topics on each tape being different. The voice of this person was the target voice. The third tape in each set was of a different person of the same gender talking. This was the foil voice. The 40 voices were of males and females unknown to any of the participants, representing a range of ages, dialects, and ethnicities talking about going shopping and what they would do if they won the Lottery. Each voice sample lasted approximately

30 seconds. Each target voice was paired with a foil voice so that target/foil pairs varied in discriminability.

Training stimuli. Six sets of three tapes were prepared, as for the test stimuli. However, two sets were of nonhuman sounds (a car engine and an ambulance siren; a dog barking and a cat miaowing), and two sets were of highly discriminable voices (e.g., a male target voice paired with a female foil voice). The final two sets were of two male voices and two female voices respectively.

Procedure. Training and testing took place over three sessions. For the children with autism and the children with SLI the three sessions were spread over 2 days; for the young mainstream children all the sessions were on the same day. In each session the child sat at a table on which two sets of headphones (one red, one yellow), and cassette holders (one red, one yellow) were placed. The experimenter explained to the child that they were going to play a game called "find the sound". In the training session, the experimenter showed the child a tape and said "On this tape we're going to hear a car engine. Listen carefully". The tape of a car engine was then played using a single-track cassette recorder. The experimenter then pointed to two sets of headphones, one yellow and one red, saying "See if you can find the car engine on the red headphones or the yellow headphones". While the target sound was replayed, a multi-track recorder was used to play a second tape of the car engine through the red headphones, and the foil sound (of an ambulance siren) through the yellow headphones. The experimenter demonstrated how to listen to the headphones and then allowed the child to listen to both in turn. The experimenter then asked the child "Where was the car engine?" If the child pointed to the red headphones she or he was praised and the experimenter, removing the tape from the cassette player, said "Yes, the car engine was on the *red* headphones, so we are now going to put the tape in the *red* box". If the child pointed to the (incorrect) yellow headphones they were asked what they heard through those headphones. If they answered "An ambulance siren", the experimenter asked "So where did you hear the car engine?" If the child then indicated the red headphones they were praised, the tape was removed and placed in the red cassette holder. If the child did not respond the procedure was repeated using the same material. On the second training trial the sound of a dog barking was played and had to be discriminated from a recording of a cat miaowing using the same procedure as described above, but with the target sound being played through the yellow headphones. On the third training trial the experimenter said: "Now listen carefully to this lady talking. You don't know this lady. Just listen to her talking". The experimenter then said "Now see if you can find the lady on the red headphones or the yellow headphones. She'll be saying something different, but it's the same lady/same voice". In the fourth training trial the target voice was male, to be discriminated from a female voice, using the above procedure. The fifth and sixth training trials involved the discrimination of two male voices and two female voices respectively. The training criterion was correct discrimination on five of the six training trials including both trials involving male/female voice discrimination and at least one of the same-sex voice trials. One child with SLI failed to meet this criterion and was excluded from the experiment.

Testing took place in two subsequent sessions. At the beginning of the first test session the child was reminded of the training session and told they were going to do some more of the "find the sound" game. Testing was preceded by presentation of the two sets of nonhuman sounds, following the same procedure used in training. This was done to remind the child of the procedure learned in training. The child was then presented with 10 test sets, plus 3 inserted training sets. The child was asked if the target voice was of a man or a woman talking and was then asked to find the same man/woman talking on the headphones. For each trial the tape was placed in the box of the same colour as that of the headphones chosen, irrespective of whether the answer was correct or not. In the second test session the child was presented with the remaining 10 test sets, plus 1

inserted training set, using the same procedure as before. Children's responses were noted.

Results

Two children did not complete this experiment: the child with SLI who failed to reach the training criterion, and one child with autism who was absent on the test days. As evidenced by their performance on the inserted training items, those children completing the experiment remained on task throughout. For these children, the number of voices correctly matched to sample (out of 20, i.e. excluding the inserted training items) was counted. The mean scores were 12.33 (*SD* 3.29) for the autism group, 12.06 (*SD* 2.21) for the SLI group, and 13.68 (*SD* 2.24) for the young mainstream group. A one-way ANOVA showed no differences between the groups, $F(2, 52) = 2.06$, n.s. All three groups scored significantly above chance.

The results compared to the predictions. We predicted that children with autism would not show impaired voice discrimination, and this prediction was supported. The children with SLI also showed unimpaired voice discrimination, relative to typically developing controls.

Experiment 4: Vocal Affect Naming, and Vocal-facial Affect Matching

Method

Training stimuli. Six audio tapes were prepared of a male actor reciting the days of the week or the months of the year in ways appropriate to each of the following emotions: happiness, sadness, disgust, fear, anger, and surprise. Nonverbal vocalisations appropriate to the emotions were included, for example, laughter, or gasps of surprise. Six photographs of a man expressing each of these emotions were taken from Ekman and Friesen (1975).

Test stimuli. Eighteen audio tapes were made of an actress reciting the days of the week or the months of the year in ways appropriate to each of the six emotions used in training. Three copies of each of 6 photographs of a woman expressing these emotions were taken from Ekman and Friesen (1975), i.e. 18 photos in all. Three additional copies of Ekman and Friesen photographs were made of a woman expressing disgust, fear, and surprise. These were added to the set in order that the guessing rate on the final three trials would not be unacceptably high—in other words, they were foils only.

Procedure. Training and testing for both parts of the experiment took place in the same session. Training began by the experimenter introducing the child to the phrase "feeling like". She played the child a tape of the man emoting happiness, and said "Listen carefully and tell me what this man is feeling like". If the child responded with an appropriate label (e.g., "happy" or "laughing") the photograph of a happy man was placed in front of the child and the experimenter said "Yes, it's the … man", using whatever term the child had used. If the child did not respond or responded inappropriately, the experimenter said "What do you think has happened?" If there was still no response the experimenter said "It's the happy man" and placed the photograph of the happy man in front of the child. This same procedure was repeated with the materials corresponding to the other emotions, the experimenter attempting in each case to elicit the child's preferred term for the emotion depicted. In the small number of cases where a child did not respond, the experimenter provided the words sad, yuk, scared/frightened, angry/cross, surprised/shocked to label the remaining five emotions. Following this, the six training photographs were shuffled and placed in front of the child. The experimenter pointed at each photograph in turn and asked the child to name how the man was feeling. This procedure was

repeated until the child had appropriately named the feelings displayed in all the photographs. The experimenter then placed the six photographs in front of the child in a random order and said to the child "Listen to the tape and tell me how the man is feeling". When she had elicited a response, the experimenter said "Can you find his photograph? When you've found his photograph turn it over". The six tapes were played in turn and any errors made by the child were corrected. The training criterion was correct emotion naming, and correct identification of the matching photograph, for four out of the six emotions. The criterion was met by all the children.

For the test, 3 of the training photographs and corresponding tapes were inserted within the 18 test items, making a total of 21 items. The experimenter showed the 21 photographs of the woman to the child in turn, asking the child to name how the woman was feeling each time. Any errors were corrected. After they had been named, the 18 test photographs, plus the 3 additional foils and the 3 training photographs, were laid out in fixed position 6×4 array in front of the child. Each tape was then played, using equivalent instructions to those used in training. The experimenter noted the label given by the child to each vocally expressed emotion. The child then selected a photograph. Correctly identified photographs were removed from the array. The experimenter corrected any mistakes to ensure that the correct photograph was removed each time. The child's responses were noted.

Results

From their performance on the three inserted training items, it was evident that all the children remained on task throughout. The number of vocally expressed emotions correctly named (out of 18), and the number of vocally expressed emotions correctly matched to facial expressions (out of 18) by each child were counted. The mean scores for the autism group were 13.5 (SD 7.79) (naming) and 13.17 (SD 3.2) (matching); for the SLI group 11.11 (SD 3.02) (naming) and 9.58 (SD 3.37) (matching); and for the typically developing children 13.05 (SD 1.84) (naming) and 15.84 (SD 1.8) (matching).

A two-way ANOVA (group × condition) showed a significant main effect of group, $F(2, 53) = 12.41$, $p < .001$, and a significant group by condition interaction, $F(2, 53) = 38.42$, $p < .001$. The main effect of condition was not significant. Pairwise comparisons of the main effect means across groups showed that, overall, the children with SLI performed significantly worse than the other two groups, which did not differ from one another. The interaction arose because the SLI children did better on naming vocally expressed emotions than on matching them to facial expressions, whereas the mainstream children did better on matching than on naming (both comparisons significant with Bonferroni-corrected t tests, $p < .01$). Pairwise comparisons among the groups showed that, for naming emotions, the SLI children did significantly worse than both the autism and mainstream groups ($p < .05$ in both cases). For matching emotions to facial expression, all three groups differed significantly from one another ($p < .01$).

The results compared to the predictions. We predicted (1) that we would confirm existing findings showing that children with autism have impaired voice-face affect matching and vocal affect naming. This prediction was supported only insofar as the children with autism performed worse than the young typically developing controls on the affect matching task. They did not perform worse than the typically developing controls on the affect naming task, and they performed better on both tasks than the children with SLI. With regard to the

possibility that a cross-modal matching impairment contributes to an affect matching impairment in children with autism, there was no clear evidence of cross-modal problems in the children with autism: they performed very similarly on the tests of affect matching and affect naming. By contrast, the children with SLI performed worse on affect matching than on affect naming, providing further evidence (in addition to that obtained in Experiment 1) of cross-modal difficulties.

Analyses of Relationships between Task Performance and Background Variables

There were no significant correlations for the children with autism or for the children with SLI between task performance (on all the experiments) and background variables (chronological age, verbal ability, nonverbal ability, and number of terms in school), using a Pearson correlation matrix and, in view of the number of correlations examined, a probability value of .01. For the mainstream children, the only significant correlation between task performance (on Experiments 3 and 4) and background variables was the correlation between affect naming and age.

Analyses of Relationships between Task Performances

In the children with autism, there were significant correlations between object-sound matching (Experiment 1) and affect naming (Experiment 4) ($r = .704$, $p = .001$); between object-sound matching and affect matching ($r = .684$, $p = .002$); and between affect naming and affect matching ($r = .944$, $p < .001$). In the children with SLI, voice-face matching (Experiment 1) correlated with object-sound matching (Experiment 1) ($r = .653$, $p = .002$); with voice recognition (Experiment 2) ($r = .624$, $p = .004$); with voice discrimination (Experiment 3) ($r = .572$, $p = .01$); and with affect naming (Experiment 4)($r = .633$, $p = .004$).

Discussion

The main findings from this study are summarised in Table 2. The findings were to a large extent unexpected. We had assumed that children with SLI do not have impairments in processing socioemotional stimuli, at least not when language ability has been controlled for. On the basis of this assumption, and on the basis of existing evidence of impaired processing of socioemotional stimuli in autism, we predicted that children with autism would be impaired relative to children with SLI on a test of familiar voice-face identity matching (Experiment 1) and on a test of familiar voice recognition (Experiment 2). We further predicted that children with autism would be impaired relative to children with SLI and to language-matched typically developing children on tests of the ability to name vocally expressed emotions and to match vocally and facially expressed emotions (Experiment 4). However, we predicted that children with autism would not be impaired relative to controls on a test of the ability to discriminate between unfamiliar voices (Experiment 3).

In the event, only two predictions were supported: the children with autism were impaired relative to the young typically developing children on the test of affect match-

Table 2
Summary of Findings

Group	Expt 1 Familiar voice-face matching	Expt 1 Familiar object-sound matching	Expt 2 Familiar voice recognition	Expt 3 Unfamiliar voice discrimination	Expt 4 Affect matching	Expt 4 Affect naming
Autism (A)	Same	Same	Same	Same	Superior to SLI; impaired relative to MS	Superior to SLI; same as MS
SLI	Same	Same	Same	Same	Impaired relative to A and to MS	Impaired relative to A and to MS
Mainstream (MS)	Not tested	Not tested	Not tested	Same	Superior to SLI and to A	Superior to SLI; same as A

ing; and their unfamiliar voice discrimination was un-impaired. Unexpectedly, the children with SLI performed similarly to the children with autism on the tests of familiar voice-face identity matching (Experiment 1) and familiar voice recognition (Experiment 2). Moreover, the children with SLI were not only impaired relative to the typically developing children on the tests of voice-face affect matching and vocal affect naming, they actually performed significantly worse than the children with autism on both these tests (Experiment 4).

The striking results of the tests of affect processing suggest that children with SLI have previously unsuspected difficulties in interpreting vocally expressed affect. This conclusion is based mainly on the finding that they were impaired relative to language-matched typically developing controls on a test of affect naming. The children with SLI were also impaired on affect naming relative to the children with autism. However, it seems possible that the surprisingly good performance of the children with autism on this task resulted from the fact that children with autism are specifically taught emotion naming. Some of the children with autism taking part in the experiment were in fact being taught to name basic emotions in the same week as they were tested, and our pre-test procedure of eliciting names for emotions may have reinforced this teaching. The children with SLI were also impaired relative to the typically developing children and to the children with autism on a test of the ability to match vocally and facially expressed emotions. This probably resulted from a combination of factors. In the first place, it seems from the results of the affect naming test that children with SLI have difficulty in interpreting vocally expressed emotion. In the second place, it is known that some children with SLI have cross-modal processing impairments (Kamhi et al., 1984), and the positive correlation between voice-face and sound-object identity matching in the children with SLI in Experiment 1 tends to confirm this. The fact that affect matching was even more impaired than affect naming in the children with SLI provides further evidence suggesting that cross-modal processing impairments contributed to poor performance on the affect matching task. In the third place, it may be that children with SLI also have impaired ability to process faces. Face processing has not been investigated in children with SLI, and it is an open question as to whether their impaired vocal affect naming is associated with poor ability to process voices, to process affect, or to process social stimuli in general, including faces.

After completing the present study we came across an earlier finding by Courtright and Courtright (1983) of impaired matching of vocally and facially expressed

emotion by children with SLI. However, apart from their evidence, the only other evidence which we know of which might suggest that some children with SLI might have impaired ability to process socioemotional stimuli comes from a study by Rutter, Mawhood, and Howlin (1992). These authors reported that some adults who had severe receptive language impairments in childhood had social relating difficulties in adulthood resembling those seen in mild cases of autism. It is, in addition, well known that children with SLI are at risk for behavioural and psychiatric disorder (Beitchman et al., 1996; Cohen et al., 1998). In the past, this has been thought to result from the negative consequences of communication difficulties. However, both Beitchman et al. and Cohen et al. tentatively suggest that language impairments and behavioural/psychiatric problems might arise from a common cause, such as neurological immaturity or limited processing capacity.

This suggestion is strengthened by the fact that there exists a group of children who have semantic and pragmatic language difficulties and who may show some of the other behaviours that are symptomatic of autism (Brook & Bowler, 1992; Rapin, 1996). It is controversial as to whether or not these children should properly be described as autistic, rather than as language impaired (Boucher, 1998). However, regardless of the controversy over terminology, the existence of these children suggests that there is a continuum of language and communication related disorders with no clear demarcation between those that include clinically identified problems in social relating and those that do not. The children in our study specifically did not have pragmatic impairments. If our evidence on children with predominantly expressive phonological and grammatical impairments were to be confirmed it would suggest that social impairments extend further into the continuum of language impairments than has previously been suspected.

The failure to find any differences between the performance of children with SLI and children with autism on the tests of familiar voice-face matching and familiar voice recognition (Experiments 1 and 2) was also unexpected. Moreover, because we did not include a typically developing control group in these two experiments, it is not possible to decide whether these surprising findings result from unexpectedly poor performance by the children with SLI or unexpectedly good performance by the children with autism. However, a likely interpretation of the findings on each of these experiments can be arrived at by considering circumstantial detail. We will consider each experiment in turn.

Regarding the results of Experiment 1, we have shown impaired voice-face matching in children with autism in

two previous experiments, relative to children with poor language and mild mental retardation (Boucher & Lewis, 1992; Boucher et al., 1998). In addition, there is internal evidence of cross-modal processing difficulties in the children with SLI, such as would have almost certainly impaired their performance on this task relative to language-matched typically developing children, had we been able to make this comparison. It seems likely, therefore, that performance on the voice-face matching task was unexpectedly impaired in the SLI group rather than being unexpectedly spared in the children with autism. This interpretation is supported by the fact that neither group performed well in absolute terms in this experiment.

Regarding Experiment 2, we have shown in a previous study that children with autism are impaired relative to children with SLI on a test of familiar voice recognition (Boucher et al., 1998). However, because children with autism had found familiar voice recognition extremely difficult in the 1998 study (several failing to complete the test), in the present study we used fewer trials and included only highly familiar and discriminable voices. As a result, both groups of children performed well on this test, with one child with autism and four children with SLI performing at ceiling. It seems possible, therefore, that a reduction in the sensitivity of the test in this experiment caused the failure to demonstrate the expected impairment in the children with autism. Further tests are needed to clarify this point.

It might be suggested that the surprisingly poor performance of the children with SLI in Experiment 4, and possibly also in Experiments 1 and 2, is associated with low nonverbal ability. It might further be suggested that the below-average nonverbal ability of the children with SLI makes their diagnosis suspect. These two possibilities will be considered in turn. With regard to the suggestion that the poor performance of the children with SLI might result from low nonverbal ability, this suggestion does not hold because nonverbal ability in the children with SLI was no lower than in the children with autism (mean nonverbal IQ in the children with SLI was 94, and in the children with autism it was 91). In addition, it can be inferred that nonverbal ability in the two clinical groups was considerably higher than in the typically developing children given the younger age of the latter group (see Table 1).

With regard to the suggestion that the below-average nonverbal ability in the children with SLI makes their diagnosis suspect, it is well known that the performance of children with SLI on standard measures of nonverbal ability declines with age (see Leonard, 1997, for a discussion of reasons for this decline). The children with SLI who took part in this study had a mean chronological age of 9 years 0 months, *SD* 1.4 years, and some decline in performance on a standard measure of nonverbal ability, such as the Coloured Progressive Matrices, would be expected over this age range. The children tested had all been diagnosed at younger ages by experienced speech and language therapists as having SLI; the children's diagnoses were confirmed in their formal Statements of Special Educational Need, as a result of which they were attending schools for children with specific communication problems and not schools for children with overall mental retardation.

Turning now to the performance of the children with autism across all four of the experiments, the fact that the children with SLI performed unexpectedly poorly means

that we did not succeed in replicating our earlier findings of impaired voice-face identity matching (Experiment 1), or impaired familiar voice recognition (Experiment 2) in the children with autism. However, as argued above, it seems more likely that children with SLI resemble children with autism in having voice processing impairments, than that the children with autism were performing normally. The results of Experiment 1 provide evidence against there being cross-modal processing deficits in autism, and this strengthens our previous interpretation of impaired voice-face identity matching in terms of the additive effects of impaired voice processing and impaired face processing.

As predicted, the children with autism showed no impairment of unfamiliar voice discrimination, relative to a group of language-matched, typically developing controls (Experiment 3). This is in keeping with findings on face discrimination in children with autism (Boucher & Lewis, 1992; Boucher et al., 2000).

The ability of the children with autism to match vocally and facially expressed emotions was, as predicted, impaired relative to language-matched, typically developing controls (Experiment 4). The ability to name vocally expressed emotions was not impaired, which was surprising. However, as suggested above, this may have resulted from the fact that children with autism are often taught in school to name emotions, and our pre-test procedures may have reinforced this teaching. Nevertheless, even if explicit teaching facilitated emotion naming in the children with autism, the present finding must be taken as indicating that the ability to name vocally expressed basic emotions is not severely impaired in autism. This would be consistent with findings on the ability to interpret facial expressions of basic emotions. It would also be consistent with an interpretation of impaired vocal-facial affect matching in children with autism in terms of an additive effect of mildly impaired facial affect processing and mildly impaired vocal affect processing.

Subsidiary aims of the experiments were to assess relationships between task performance and the background variables of age, verbal ability, nonverbal ability, and number of terms in school within each of the experiments; and to assess within-group relationships between task performances across the four experiments. The negative results of correlation tests comparing the background measures with scores on the experimental tasks probably resulted from a lack of variability in the background measures. However, this cannot explain the lack of correlation between the number of terms in school (which varied considerably) and performance on the tests involving the recognition of the faces and voices of staff and children from the participants' schools. The lack of correlation here is surprising. However, it could reflect the fact that children change classes over the years, and forget the voices of staff or children with whom they no longer interact every day.

With regard to correlations between task performances, the significant correlations in the children with autism between object-sound matching and the two emotion processing tasks are difficult to explain, but could suggest that vocal expressions of emotion are perceived and processed as if they were environmental, nonsocial, sounds. The set of correlations in the children with SLI are more systematic. In particular, it is noteworthy that performances across all of the voice processing tasks (with the exception of the vocal-facial affect

matching task) were correlated. In addition, familiar voice-face identity matching correlated with object-sound matching, as would be expected if cross-modal processing impairments are affecting performance. Vocal-facial affect matching did not correlate with either of these tasks, as might also have been expected. However if, as suggested above, very poor affect matching by children with SLI results from a combination of impaired ability to identify vocally expressed emotions, plus impaired cross-modal processing, and possibly also some degree of impairment of face processing, then a potential correlation between all of the matching tasks may have been diluted by the effects of these other factors.

In conclusion, the results of the present study suggest that children with SLI not only have impaired cross-modal processing, but also that they have additional problems in processing vocally expressed affect and possibly, also, more wide-ranging impairments in processing social stimuli. However, they are able to discriminate between unfamiliar voices, which could suggest that any problems in processing social stimuli arise at the level of processing meaning. The results of the study further suggest that children with moderate to high-functioning autism do not have cross-modal processing impairments. Nevertheless, the results confirm that these children do have problems in matching vocally and facially expressed emotions. The children with autism, like the children with SLI, were able to discriminate between unfamiliar voices. This finding corresponds to Boucher and Lewis's (1992) finding that children with autism have no difficulty in discriminating between unfamiliar faces and suggests that for children with autism, as for children with SLI, the problem is not at the level of perceptual encoding, but rather at the level of encoding meaning.

Replication and extension of the findings of the present study, and investigation into the causes of impaired socioemotional processing in children with SLI or autism will be important for our understanding of both disorders. Specifically, if children with SLI are confirmed as having significant problems with processing voices, and possibly also faces, then descriptions of children with specific language impairments and approaches to working with such children will need to be radically altered. Equally, it would not be logical to suggest as some theorists have done that autism is caused, or partly caused, by impaired ability to process social stimuli such as faces and voices, or facially, vocally, or bodily expressed affect. It would not be logical because if this were the case then children with SLI *who also have these impairments* should also be autistic. The impairments in processing social stimuli that occur in autism may, of course, be of a different kind and have a different cause, or cause(s), from those which occur in SLI. Furthermore it may be that whatever causes impaired processing of social stimuli in children with autism also causes the other signs and symptoms of autism. If this is the case, however, it is inaccurate and misleading to say that impaired processing of social stimuli *causes* autism: rather impaired processing of social stimuli is one of the consequences of whatever fundamental psychological deficit(s) underlie autism.

Replication and extension of the findings of the present study, and investigation into the causes of impaired socioemotional processing in children with SLI or autism, will also be important for our understanding of ways in which SLI and autism are related. There are already multiple links between the two disorders, comorbidity between language disorders and autism being common, and language-related impairments commonly occuring in relatives of individuals with autism. Moreover, there appears to be a seamless continuum not just across a spectrum of autism-related "disorders", but from what may be termed Asperger syndrome, through pragmatic language impairments without structural language difficulties, through to semantic, syntactic, morphological, and phonological language impairments (in that order). The unexpected results of the present study suggest that further research into social cognition in children with SLI might show even closer links across the spectrum of language-related developmental disorders than has previously been envisaged.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental diseases* (4th ed.; DSM-IV). Washington, DC: Author.

Baltaxe, C., & Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic, and autistic children. *Journal of Autism and Developmental Disorders*, 17, 255–271.

Beitchman, J. H., Brownlie, E. B., Inglis, A., Wild, J., Ferguson, B., Schacter, D., Lancee, W., Wilson, B., & Mathews, R. (1996). Seven-year follow-up of speech/language impaired and control children: Psychiatric outcome. *Journal of Child Psychology and Psychiatry*, 37, 961–970.

Bishop, D. (1989). *Test for the Reception of Grammar*. Oxford: Thomas Leach Ltd.

Boucher, J. (1998). Semantic-pragmatic language disorder as a distinct diagnostic entity: Logical considerations and directions for future research. *International Journal of Language and Communication Disorders*, 33, 71–81.

Boucher, J., Cowell, P., Broks, P., Mayes, A., Roberts, N., & Howard, M. (2000). *A study of neuroanatomy, neuropsychology, and diagnostic profiles in adults with high functioning autism*. Paper presented at the Experimental Psychology Society meeting, Cambridge, July.

Boucher, J., & Lewis, V. (1992). Unfamiliar face recognition in relatively able autistic children. *Journal of Child Psychology and Psychiatry*, 33, 843–859.

Boucher, J., Lewis, V., & Collis, G. (1998). Familiar face and voice matching and recognition in children with autism. *Journal of Child Psychology and Psychiatry*, 39, 171–182.

Boucher, J., Lewis, V., Collis, G., & LePage, G. (1992). *Familiar face and voice recognition in children with autism*. British Psychological Society Developmental Section Conference, University of Edinburgh.

Brazelton, T., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother–infant interaction. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on the caregiver*. New York: Wiley.

Brook, S., & Bowler, D. (1992). Semantic and pragmatic impairments in children. *Journal of Autism and Developmental Disorders*, 22, 61–81.

Bryson, C. (1972). Short-term memory and cross-modal information processing in autistic children. *Journal of Learning Disabilities*, 5, 81–91.

Capps, L., Yirmaya, N., & Sigman, M. (1992). Understanding simple and complex emotions in nonretarded children with autism. *Journal of Child Psychology and Psychiatry*, 33, 1169–1182.

Cohen, N. J., Menna, R., Vallance, D. D., Barwick, M. A., Im, N., & Horodezky, N. (1998). Language, social cognitive processing, and behavioural characteristics of psychiatrically disturbed children with previously identified and unsuspected language impairments. *Journal of Child Psychology and Psychiatry*, 39, 853–864.

Courtright, J. A., & Courtright, I. C. (1983). The perception of nonverbal vocal cues of emotional meaning by language-disordered and normal children. *Journal of Speech and Hearing Research*, 26, 412–417.

Davies, S., Bishop, D., Manstead, A., & Tantam, D. (1994). Face perception in autistic children. *Journal of Child Psychology and Psychiatry*, 35, 1033–1058.

DeCasper, A., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mother's voice. *Science*, 208, 1174–1176.

Eisenberg, R. (1976). *Auditory competence in early life: The roots of communicative behavior*. Baltimore, MD: University Park Press.

Ekman, P., & Friesen, V. (1975). *Unmasking a face: A guide to recognising emotions from facial cues*. Englewood Cliffs, NJ: Prentice-Hall.

Fein, D., Lucci, D., Braverman, M., & Waterhouse, L. (1992). Comprehension of affect in context in children with pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 33, 1157–1168.

Fine, J., Bartolucci, G., Ginsberg, G., & Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 32, 771–782.

de Gelder, B., Vroomen, J., & van der Heide, L. (1991). Face recognition and lipreading in autism. *European Journal of Cognitive Psychology*, 3, 69–86.

Hobson, R. P. (1993). *Autism and the development of mind*. Hove, UK: Lawrence Erlbaum Associates.

Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: Co-ordinating faces and voices. *Psychological Medicine*, 18, 911–923.

Hobson, R. P., Ouston, J., & Lee, A. (1989). Naming emotion in faces and voices: Abilities and disabilities in autism and mental retardation. *British Journal of Developmental Psychology*, 7, 237–250.

Kamhi, A., Catts, H., Koenig, L., & Lewis, B. (1984). Hypothesis-testing and non-linguistic symbolic abilities in language-impaired children. *Journal of Speech and Hearing Disorders*, 49, 169–176.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterisation of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, 21, 29–42.

Klin, A. (1992). Listening preferences in regard to speech in four children with developmental disabilities. *Journal of Child Psychology and Psychiatry*, 33, 736–770.

Leonard, L. (1997). *Children with specific language impairments*. Cambridge, MA: MIT Press.

Locke, J. (1993). *The child's path to spoken language*. Cambridge, MA: Harvard University Press.

MacDonald, H., Rutter, M., Howlin, P., Rios, P., Le Couteur, A., Evered, C., & Folstein, S. (1989). Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology and Psychiatry*, 30, 865–877.

Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Psychology*, 27, 153–165.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halstead, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.

Ozonoff, S., Pennington, B., & Rogers, S. J. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry*, 31, 343–361.

Rapin, I. (1996). Developmental language disorders: A clinical update. *Journal of Child Psychology and Psychiatry*, 37, 643–656.

Raven, J. (1990). *The Coloured Progressive Matrices*. Oxford: Oxford Psychologists' Press.

Ricks, D. (1975). Vocal communication in preverbal normal and autistic children. In N. O'Connor (Ed.), *Language, cognitive deficits and retardation*. London: Butterworth.

Rutter, M., Mawhood, L., & Howlin, P. (1992). Language delay and social development. In P. Fletcher & D. Hall (Eds.), *Speech and language-impairments in children*. London: Whurr Press.

Tantam, D. (1992). Characterising the fundamental social handicap in autism. *Acta Paedopsychiatrica*, 55, 83–91.

VanLancker, D., Cornelius, C., & Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic and normal children. *Developmental Neuropsychology*, 5, 207–226.

Walker-Andrews, A. S. (1988). Infants' perception of the affordances of expressive behaviors. In C. Rovee-Collier (Ed.), *Advances in infancy research*, Vol. 5. Norwood, NJ: Ablex.

Wolff, P. (1969). The natural history of crying and other vocalisations in early infancy. In B. M. Foss (Ed.), *Determinants of infant–behaviour*, Vol. IV. London: Methuen.