BRITISH JOURNAL OF PSYCHIATRY (1999), 175, 444-451

Autism screening questionnaire: diagnostic validity

SIBEL KAZAK BERUMENT, MICHAEL RUTTER, CATHERINE LORD, ANDREW PICKLES and ANTHONY BAILEY

Background Good interview and diagnostic measures for autism and other pervasive developmental disorders (PDDs) are available but there is a lack of a good screening questionnaire.

Aims To develop and test a screening questionnaire based on items in the best available diagnostic interview – the Autism Diagnostic Interview – Revised (ADI–R).

Method A 40-item scale, the Autism Screening Questionnaire (ASQ), was developed and tested on a sample of 160 individuals with PDD and 40 with non-PDD diagnoses.

Results The ASQ has good discriminative validity with respect to the separation of PDD from non-PDD diagnoses at all IQ levels, with a cut-off of 15 proving most effective. The differentiation between autism and other varieties of PDD was weaker.

Conclusions The ASQ is an effective screening questionnaire for PDD.

Declaration of interest The study was supported by the Medical Research Council.

Several different diagnostic measures for autism have been developed over the years (Schopler & Mesibov, 1988), but the Autism Diagnostic Interview (ADI; Le Couteur *et al*, 1989; Lord *et al*, 1994) has become the generally accepted standardised diagnostic parental interview, with the Autism Diagnostic Observational Schedule (ADOS; Lord *et al*, 1989; Di Lavore *et al*, 1995) as the comparable accepted observational measure. Following Rimland's (1964) check-list, there have been several attempts to develop a reliable and valid screening instrument (e.g. Krug *et al*, 1980; Dahlgren & Gillberg, 1989; Baron-Cohen *et al*, 1992), but each is limited in its psychometric qualities (Parks, 1988; Volkmar *et al*, 1988), reliance on outmoded diagnostic criteria, range of behaviours covered, restriction to current functioning and/or age range covered.

Hence, there is a need for a reliable and valid screening instrument that is based on the current diagnostic criteria for autism and which can be used with all age groups. Such an instrument could be very useful for epidemiological research or when the aim of the research is to compare autism with other clinical groups in terms of autism-like features. The Autism Screening Questionnaire (ASQ) was developed to fulfil this need.

DEVELOPMENT OF THE AUTISM SCREENING QUESTIONNAIRE

The ASQ was designed by M.R. and C.L. to be completed by the primary care-giver on individuals who might have a pervasive developmental disorder (PDD). The selection of questions was based on the revised version of the ADI algorithm (ADI–R; Lord *et al*, 1994) used for ICD–10 (World Health Organization, 1992) and DSM–IV (American Psychiatric Association, 1994) diagnosis of autism. These provide an operational diagnosis which is based on the behavioural item scores in three areas of functioning: reciprocal social interaction; language and communication; and repetitive and stereotyped patterns of behaviours.

Care was taken in the design and choice of questions in order to focus on behaviours that care-givers are likely to have had the opportunity to observe, that involved concepts likely to be understandable to non-professionals, and that required a minimum of inference. In the design of the questionnaire, attention was also paid to simplicity, clarity and lack of ambiguity in wording. Following the approach of the ADI–R, the questions were designed to focus on qualitative deviance rather than developmental delay or impairment and, where the latter was likely to affect codings, the questions were focused on the four- to five-year age period. Otherwise the questions concern lifetime manifestations.

Hence, the ASQ consists of 40 questions that are based on the ADI–R but which have been modified into a form understandable by parents without further explanation. There are questions on reciprocal social interaction (such as social smiling, interest in other children, and offering comfort to others), language and communication (including the use of conventional gestures, reciprocal conversation, and stereotyped utterances), and repetitive and stereotyped patterns of behaviours (including circumscribed interests and unusual preoccupations). In addition, the ASQ includes a question about self-injurious behaviour and a question about the individual's current language functioning.

Two versions of the questionnaire were designed: one for individuals under six years of age and the other for individuals aged six years and over. A score of 1 is given for the presence of the abnormal behaviour and a score of 0 for its absence. Thus, the total score ranges from 0 to 39 (the item on current language level not being included in the summary score) for individuals with language. For those without language, the top score is 34 because the abnormal language items are inapplicable. In the ADI–R (Lord *et al*, 1994) there are separate algorithms for those with and without language, but for the ASQ it was decided to have just a single score, because it is desirable to have a simple scoring system for a screening

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

questionnaire. Empirical findings showed that the mean total scores for individuals with autism with and without language were broadly comparable (23.13 and 20.49, respectively), the sample size being 62 for those with and 21 for those without language. Although the mean score for those without language was significantly lower (P=0.03), the proportions with scores of 15 (the cut-off deriving from receiver operating characteristic analyses; see below) or above were similar (61/62 with *v.* 20/21 without language).

METHOD

The sample consisted of 200 individuals who had participated in previous studies. These studies included a family genetic study of autism (Bolton *et al*, 1994), a study of adolescents with clinically diagnosed Asperger syndrome or conduct disorder, a study of individuals with either the fragile X anomaly or Rett syndrome, and a study of the diagnosis of autism in young children presenting with developmental problems (further details available from the corresponding author upon request). There were 160 individuals with PDD (comprising 83 with autism, 49 with atypical autism, 16 with Asperger syndrome, seven with fragile X but not autism, and five with Rett syndrome; see Table 1). There were 40 individuals with non-PDD diagnoses (comprising 10 with conduct disorder, 7 with specific developmental language disorder, 15 with mental retardation and 8 with other psychiatric diagnoses such as anxiety disorders). As expected, the male preponderance in autism (2.8:1) and in other PDD (6.7:1) was greater than that in the non-PDD diagnoses group (1.7:1). However, because in neither the autism group nor the non-PDD group was there a substantial or statistically significant gender difference in ASQ score (25.2 *v.* 25.2 in the autism group and 11.1 *v.* 11.3 in the non-PDD group), gender differences are not considered further. The mean score for males in the other PDD group was higher than that for females (19.9 *v.* 13.9), but this was to be expected on the basis of the differential pattern of diagnoses (e.g. Rett syndrome only in females).

All the participants had previously been assessed on the ADI or ADI–R as part of other studies: 77 participants had been assessed on the original version of the ADI (Le Couteur *et al*, 1989) and 123 on the revised version (Lord *et al*, 1994). In most cases, however, these instruments had been administered several years ago. ADI–R scores can be translated into ADI scores, and that procedure was followed here. The differentiation between autism and atypical autism was made on the basis ADI algorithm criteria (because the clinical diagnostic conventions were changing over the period in which the samples were being collected); otherwise the diagnoses are those made by the clinicians. Of the 140 individuals with diagnoses of PDD, 115 had been clinically diagnosed as having autism, but only 83 of these met ADI criteria. In all cases the ASQ was sent as a postal questionnaire.

Table 1 Subject details

Diagnosis	*n*	Chronological age (years)	
		Mean (s.d.)	Range
Autism	83	23.08 (8.07)	4.01–40.03
Atypical autism	49	7.03 (7.01)	4.00–19.06
Asperger syndrome	16	17.03 (4.09)	9.08–30.00
Fragile X	7	13.04 (4.06)	8.10–21.04
Rett syndrome	5	11.04 (3.04)	8.02–16.11
Conduct disorder	10	14.03 (5.01)	7.04–20.00
Language delay	7	5.06 (0.10)	4.04–7.00
Other clinical diagnosis	8	11.01 (4.07)	6.10–20.06
Mental retardation	15	8.04 (6.11)	4.04–32.07

Design

Four steps were taken to assess the diagnostic validity of the ASQ. First, a factor analysis was performed to determine whether the scale provided a differentiation that reflected the conceptualisation of the three main domains of abnormality found in autism (reciprocal social interaction, communication, and repetitive stereotyped behaviour). Second, the combination of individual items was assessed by noting their correlation with the total ASQ score and the extent to which they differentiated PDDs (including autism) from other diagnoses. Third, the correlations between the ASQ and the ADI were calculated. Fourth, receiver operator curves were applied to determine the degree to which the ASQ differentiated PDD from other diagnoses. These analyses were repeated within IQ strata to check whether the differentiation was affected by IQ level.

RESULTS

Internal consistency of the ASQ

The ASQ's factor structure was explored (see Table 2). Evaluation of three- and four-factor solutions for the 39 items suggested that a four-factor model appeared to be the more meaningful. Principal component factoring with varimax rotation yielded four factors which explained 42.4% of the total variation of the ASQ data, with 24.3% accounted for by a social interaction factor (eigenvalue 9.7), 8.7% by a communication factor (eigenvalue 3.38), 5% by an abnormal language factor (eigenvalue 1.94) and 4.5% by a stereotyped behaviour factor (eigenvalue 1.74). The alpha reliability coefficient for the total scale was 0.90; for the first factor it was 0.91, for the second factor 0.71, for the third factor 0.79 and for the fourth factor 0.67. All the individual item to total score correlations were positive and mainly substantial, in the range 0.26–0.73 (23 of the 39 exceeding 0.50). The extent to which the four factors mapped onto the three key domains of autistic phenomena, as operationalised in the ADI–R algorithm criteria, is indicated by the domain designation of each item. The social interaction factor largely coincided with the social domain, and the stereotyped behaviour factor with the repetitive stereotyped behaviour domain. The communication domain items tended to divide between the other two factors, half being in factor 3, mainly reflecting communicative deficits, and half in factor 4, mainly reflecting abnormal language features, with some also in the first social factor.

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

Table 2 Factor loadings

Item number and domain designation[1]	Item	Correlation between item and total score	Factor loading factor 1, 'social'	Factor loading factor 2, 'communication'	Factor loading factor 3, 'abnormal language'	Factor loading factor 4, 'stereotyped behaviour'
29 (S)	Offering to share	0.73	0.73			
36 (S)	Interest in children	0.60	0.73			
40 (S)	Group play	0.58	0.71			
37 (S)	Response to other children's approaches	0.63	0.70			
34 (C)	Imitative social play	0.60	0.69			
31 (S)	Offering comfort	0.64	0.65			
28 (S)	Showing and directing attention	0.67	0.65			
30 (S)	Seeking to share enjoyment	0.61	0.61			
21 (C)	Imitation	0.63	0.60			
39 (S)	Imaginative play with peers	0.60	0.59			
22 (C)	Pointing to express interest	0.61	0.55			
27 (S)	Social smiling	0.55	0.53			
26 (S)	Eye gaze	0.58	0.52			
35 (C)	Imaginative play	0.58	0.52			
33 (S)	Range of facial expressions	0.52	0.51			
38 (–)	Attention to voice	0.54	0.50			
23 (C)	Gestures	0.39	0.47			
32 (S)	Quality of social overtures	0.48	0.46			
20 (S)	Friends	0.34	0.38			
18 (–)	Self-injury	0.37	0.38			
25 (C)	Head-shaking to mean 'no'	0.61		0.64		
3 (C)	Conversation	0.36		0.62		
24 (C)	Nodding to mean 'yes'	0.60		0.57		
2 (C)	Social chat	0.34		0.52		
10 (S)	Inappropriate facial expressions	0.39		0.51		
16 (R)	Hand and finger mannerisms	0.36		0.43		
4 (C)	Stereotyped utterances	0.64			0.80	
8 (R)	Verbal rituals	0.63			0.75	
5 (C)	Inappropriate questions	0.53			0.74	
7 (C)	Neologisms	0.57			0.69	
6 (C)	Pronoun reversal	0.45			0.51	
13 (R)	Repetitive use of objects	0.51				0.65
15 (R)	Unusual sensory interests	0.42				0.64
9 (R)	Compulsions and rituals	0.44				0.55
12 (R)	Unusual preoccupations	0.41				0.50
11 (S)	Use of other's body to communicate	0.30				0.50
17 (R)	Complex body mannerisms	0.26				0.38
19 (R)	Unusual attachment to objects	0.27				0.35
14 (R)	Circumscribed interests	0.28				0.32

1. S, social reciprocal interaction domain; C, communication domain; R, repetitive behaviour and stereotyped patterns domain; –, not in algorithm.

Validity of the individual ASQ items

Table 3 shows the extent to which individual ASQ items differentiated PDDs from other diagnoses, as given clinically on the basis of standardised interview and observation data in the samples from which the cases were drawn. Of the 39 items, 33 showed a statistically significant differentiation; four of the items that did not concerned abnormal language features (stereotyped utterances, inappropriate questions, pronoun reversal and neologisms). Each of these had a relatively high frequency in non-PDD children but, as shown in Table 2, had substantial correlations with the total score (0.64, 0.53, 0.45 and 0.57, respectively). Two items

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

Table 3 Item validity analysis

Item number	Item	χ^2	Percentage with abnormality – PDD	Percentage with abnormality – other
29	Offering to share	27.75****	76.3	32.5
36	Interest in children	24.34****	79.4	40.0
40	Group play	11.20***	80.5	55.0
37	Response to other children's approaches	26.12****	78.9	37.5
34	Imitative social play	19.20****	70.0	32.5
31	Offering comfort	32.79****	72.6	23.1
28	Showing and directing attention	25.97****	61.8	15.8
30	Seeking to share enjoyment	16.34****	63.0	27.5
21	Imitation	19.94****	70.7	32.5
39	Imaginative play with peers	29.21****	86.1	46.2
22	Pointing to express interest	25.14****	67.7	25.0
27	Social smiling	10.45***	51.8	23.1
26	Eye gaze	19.85****	65.0	25.6
35	Imaginative play	38.91****	73.4	20.0
33	Range of facial expressions	19.64****	54.1	15.0
38	Attention to voice	15.72****	62.4	27.5
23	Gestures	14.67****	66.9	33.3
32	Quality of social overtures	18.30****	40.8	5.0
20	Friends	5.56**	71.1	51.3
18	Self-injury	3.18†	40.3	25.0
25	Head-shaking to mean 'no'	26.35****	66.0	20.5
3	Conversation	5.05*	32.5	13.5
24	Nodding to mean 'yes'	26.17****	69.0	23.7
2	Social chat	7.40**	17.4	0.0
10	Inappropriate facial expressions	4.03*	27.8	12.5
16	Hand and finger mannerisms	33.38****	76.5	25.0
4	Stereotyped utterances	1.57	81.8	72.2
8	Verbal rituals	13.89****	69.2	34.2
5	Inappropriate questions	0.053	56.6	54.1
7	Neologisms	0.97	47.1	37.8
6	Pronoun reversal	0.002	52.5	52.9
13	Repetitive use of objects	9.40**	66.5	40.0
15	Unusual sensory interests	6.45**	53.5	30.8
9	Compulsions and rituals	5.24	69.6	50.0
12	Unusual preoccupations	20.97****	67.3	27.5
11	Use of other's body to communicate	8.91**	61.3	35.0
17	Complex body mannerisms	12.71****	61.4	30.0
19	Unusual attachment to objects	3.23†	22.8	10.0
14	Circumscribed interests	7.62**	54.4	30.0

*$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$, †$P=0.07$.
PDD, pervasive developmental disorder.

(self-injury and unusual attachment to objects) differentiated only at the 7% significance level; both showed only modest correlations with the total score (0.37 and 0.27, respectively).

Correlations between the ASQ and the ADI

Correlations between the ADI and the ASQ were calculated for the total score and the ADI domain (social, communication and repetitive behaviour) totals. Correlation coefficients were highly significant for all comparisons both within and across domains (see Table 4).

To assess the discriminant power of the ASQ, a series of receiver operating characteristic analyses (Hanley & McNeil, 1982; Fombonne, 1991) and *t*-tests was carried out. The area under the curve served as the index of accuracy (see Table 5 for detailed numerical results). It appears that the discriminant ability of the ASQ is high in differentiating PDD (including autism) from non-PDD conditions (including mental retardation). The ASQ similarly differentiated well between autism and mental retardation and between autism and non-PDD diagnoses other than mental retardation. The ASQ also showed a significant differentiation between autism and other PDDs, but there was substantial overlap and the differentiation was much less clear-cut.

The analyses were repeated using an ASQ score that did not include the six items that failed to differentiate groups at the 5% level of statistical significance. Some marginal improvement in discriminative validity was obtained, but that between autism and other PDD was slightly worse. The discriminative validity of the ASQ was compared with that of the ADI by contrasting the areas under the receiver operating curve: for PDD *v.* non-PDD they were respectively 0.88 and 0.87; for autism *v.* mental retardation they were 0.93 and 0.96; and for autism *v.* other PDD they were 0.73 and 0.74.

Because the groups differed in IQ distribution, it was possible that the ASQ diagnostic differentiation was an artefact, deriving from the IQ differences. In order to determine whether that was the case, the analyses were repeated within IQ bands (see Table 6). Inevitably, as the data derived from several different studies with different aims and covering different age ranges, the IQ scores derived from several different tests. For individuals of school age or older, the appropriate Wechsler Performance Scale (Wechsler, 1989, 1991) was usually available; most of the younger children had been tested with Raven's Matrices (Raven *et al*, 1991), Mullen's Scale of Early Learning or the Differential Abilities Scales (Mullen, 1989). Because of this test heterogeneity any detailed assessment of IQ effects would be inappropriate, but the data were adequate for the use of broad IQ

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

Table 4 Correlations between Autism Diagnostic Interview (ADI) total, ADI domain totals and Autism Screening Questionnaire (ASQ) total and ASQ domain totals based on ADI domains

	ASQ total	ASQ social interaction	ASQ language/ communication	ASQ repetition/ stereotyped behaviour
ADI total	0.71***	0.67***	0.61***	0.48***
ADI social interaction	0.57***	0.59***	0.51***	0.31***
ADI language/communication	0.61***	0.62***	0.55***	0.36***
ADI repetition/stereotyped behaviour	0.63***	0.48***	0.47***	0.59***

***$P < 0.0005$.

Table 5 Discriminative validity of the Autism Screening Questionnaire (ASQ)

		ASQ total			ASQ total after dropping non-significant items		
	n	Mean ASQ	*t*-test	AOC	Mean ASQ	*t*-test	AOC
PDD (incl. autism) *v.* non-PDD	160	22.28	8.73***	0.86	19.88	9.46***	0.88
(incl. mental retardation)	40	11.18			8.88		
Autism *v.* non-autism (excl.	83	25.24	11.01***	0.94	22.51	11.99***	0.95
mental retardation)	25	10.25			7.96		
Autism *v.* mental retardation	83	25.24	7.54***	0.92	22.51	8.28***	0.93
	15	12.75			10.41		
Autism *v.* other PDD	83	25.24	5.89***	0.74	22.51	5.66***	0.73
	77	19.09			17.04		

***$P < 0.0005$.
AOC, area under the curve; PDD, pervasive developmental disorder.

Table 6 *t*-tests and receiver operating characteristic indices of the Autism Screening Questionnaire (ASQ) for IQ-matched groups

	ASQ total			
	n	Mean ASQ (s.d.)	*t*-test	AOC (s.e.)
IQ≥70: PDD	56	21.18 (8.11)	6.21***	0.90 (0.04)
IQ≥70: Non-PDD	18	8.39 (5.63)		
IQ 50–69: PDD	46	22.53 (6.27)	5.14***	0.90 (0.05)
IQ 50–69: Non-PDD	10	11.40 (5.87)		
IQ 30–49: PDD	25	24.70 (5.61)	3.39***	0.79 (0.12)
IQ 30–49: Non-PDD	7	14.74 (10.43)		
IQ≥70: Autism	27	24.23 (7.11)	7.92***	0.95 (0.03)
IQ≥70: Non-PDD (excl. mental retardation)	18	8.39 (5.63)		
IQ 50–69: Autism	28	25.08 (5.64)	6.51***	0.96 (0.03)
IQ 50–69: Non-PDD (excl. mental retardation)	10	11.40 (5.87)		
IQ 30–49: Autism	18	26.57 (3.93)	4.20***	0.84 (0.11)
IQ 30–49: Non-PDD (excl. mental retardation)	7	14.74 (10.43)		

***$P < 0.0005$.
AOC, area under the curve; PDD, pervasive developmental disorder.

strata. The findings showed that in the non-autistic group the mean ASQ score was lowest (at 8.39) in the sub-group with an IQ above 70 and highest in the group with severe retardation (14.74) but did not vary by IQ within the PDD (including autism) group. The diagnostic differentiation within all IQ bands (including those with severe mental retardation) was highly significant, although it was clearest in the group with an IQ above 70.

Differentiation according to ASQ domain score

Finally, analyses were undertaken to determine whether the individual behavioural domains of the ASQ provided a better diagnostic differentiation than that obtained with the total score. Individual items on the ASQ were allocated to the three key domains of autistic symptoms as determined by the equivalent items on the ADI–R. The receiver operating characteristic analyses are summarised in Table 7. All three domains provided satisfactory differentiation of PDDs (including autism) from other diagnoses, with areas under the curve ranging from 0.79 to 0.83. However, the differentiation on the total score (area under the curve 0.90) was better, the difference for both the non-communication and repetitive behaviour domains being statistically significant (z=2.75 and z=2.17, respectively). The repetitive behaviour domain on its own was not very good at differentiating autism from mental retardation (area under the curve 0.70) or autism from PDD (area under the curve 0.59). The results indicate that the most satisfactory differentiation is provided by the total ASQ score.

Cut-off points for most effective diagnostic differentiations

Examination of the receiver operating curves for the total ASQ suggested scores of 15 or more as the standard optimal cut-off for differentiating PDDs (including autism) from other diagnoses. The sensitivity was 0.85, specificity 0.75, positive predictive value 0.93 and negative predictive value 0.55 in this sample. Other cut-offs may be preferable for general population samples and choice may also vary with the purpose, for example, screening for case detection *v.* case collection. The 15-or-more cut-off point gave a sensitivity of 0.96 and specificity of 0.80 for autism *v.* other diagnoses (with mental retardation

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

excluded), and a sensitivity of 0.96 and specificity of 0.67 for autism *v.* mental retardation.

As would be expected, a much higher cut-off (22 or more) was required to separate autism from other PDDs, the sensitivity being 0.75 and specificity 0.60 at that point.

DISCUSSION

Screening for PDD

The findings showed that the ASQ succeeded well in its aim to provide an effective screening for PDD. Thirty-three of the 39 items, considered individually, provided a significant diagnostic differentiation, and the majority of the items showed substantial correlations with the total score. The overall correlation between the ASQ total score and the ADI algorithm score was high (0.712). Receiver operating characteristic analyses showed that the total score provided a good differentiation between PDD and other diagnoses. This was particularly clear-cut when the other diagnoses did not include mental retardation, but it was also good when the comparison was with mental retardation alone. We conclude that the findings show the ASQ to be a highly effective screening instrument. Its success is likely to derive in large part from the fact that it was modelled on a diagnostic interview (the ADI–R) with a demonstrated highly reliability and validity, and from the choice of questions shown to provide a good differentiation when given in an interview format designed to elicit behavioural descriptions.

ASQ and ADI

Perhaps surprisingly, the ASQ proved to be as effective a discriminator as the ADI, except that the latter was marginally better at differentiating autism from mental retardation not associated with PDD. In that connection it is probably relevant that, in most cases, the ADI data available derived from the original version of the interview (Le Couteur *et al*, 1989) and not its revision (Lord *et al*, 1994) with its considerable improvements in item wording and its better coverage of relevant areas of abnormality. Also, in all cases the parents who completed the ASQ had received the ADI or ADI–R some years previously, and it may be that the interview experience sensitised the parents to the types of features covered. This may have enhanced the discriminative validity of the ASQ – although, because the interval between the interview and the ASQ was usually many years, it is most unlikely to have directly influenced the answers to specific ASQ questions. A prospective study in which the ASQ is *followed* by the ADI is needed to provide a more rigorous test of validity. The only data of this kind derive from a highly unusual sample of adoptees from Romanian orphanages (Rutter *et al*, 1999), but the findings suggest that the ASQ may provide valid differentiation under these more rigorous testing conditions. Very young children whose parents were not familiar with autism were not included in the present sample. The study of such a group is a priority for future research. Further data are also needed on interrater and test–retest reliability.

Despite the demonstrated success of the ASQ as a screening instrument, it is, like any other questionnaire, inappropriate for

Table 7 Diagnostic differentiation of the separate symptom domain scores of the Autism Screening Questionnaire

	n	Domain scores[1]	Mean	*t*-test	AOC (s.e.)
PDD	160	S – PDD	10.18	7.55***	0.83 (0.03)
Non-PDD	40	S – non-PDD	4.68		
		C – PDD	6.93	6.49***	0.79 (0.04)
		C – non-PDD	3.93		
		R – PDD	7.57	6.56***	0.79 (0.04)
		R – non-PDD	3.85		
Autism	82	S – autism	11.77	9.16***	0.91 (0.04)
Non-PDD (excluding mental retardation)	25	S – non-PDD	4.48		
		C – autism	7.89	7.71***	0.89 (0.04)
		C – non-PDD	3.80		
		R – autism	8.10	8.41***	0.91 (0.04)
		R – non-PDD	2.76		
Autism	83	S – autism	11.77	7.09***	0.90 (0.04)
Mental retardation	15	S – mental retardation	5.00		
		C – autism	7.89	5.38***	0.83 (0.07)
		C – mental retardation	4.15		
		R – autism	8.10	2.96***	0.70 (0.08)
		R – mental retardation	5.67		
Autism	83	S – autism	11.77	5.35***	0.72 (0.04)
PDD (excluding autism)	77	S – PDD (excluding autism)	8.47		
		C – autism	7.89	5.17***	0.72 (0.04)
		C – PDD (excluding autism)	5.90		
		R – autism	8.10	2.17*	0.59 (0.05)
		R – PDD (excluding autism)	7.00		

*$P < 0.05$, ***$P < 0.0005$.
AOC, area under the curve; PDD, pervasive developmental disorder.
1. S, social interaction; C, language/communication; R, repetition/stereotyped behaviour.

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

diagnosis at the individual level, which requires detailed descriptions of actual behaviour, the meaning of which can be evaluated by the clinician, and not just yes/no answers that rely on parents providing their own interpretation of the behaviour specified in the structured question. The ADI–R provides a standardised interview for that purpose. Diagnosis also requires direct observation of the person's behaviour during social and communicative interaction. The ADOS (Lord *et al*, 1989; Di Lavore *et al*, 1995) provides a standardised measure for that purpose. In addition, diagnostic evaluation requires assessment of the individual's cognitive level.

Value of the ASQ in different samples

For obvious reasons, the discriminative power of the ASQ will be influenced by the samples studied. With the exception of the small conduct disorder group, all the children with non-PDD diagnoses in our study had received the ADI because autism or some related developmental disorder had been suggested or suspected by the agency referring the child for clinical or research evaluation. Accordingly, some autistic features might be expected. That was exactly what was found. Accordingly, the mean ASQ score of the non-retarded non-autistic group was 11.2, and that of the group with mental retardation was 12.8. This contrasts sharply with the mean score of 5.2 found for a general population of UK children adopted in infancy (Rutter *et al*, 1999). It should be noted, too, that the mean score of the six non-autistic children with severe retardation in our study was 14.7, meaning that nearly half of them scored above the cut-off point of 15 that proved optimal for the detection of autism.

It should be noted that the ASQ was not particularly effective in differentiating autism from other varieties of PDD. In that, of course, it is no different from other questionnaires or, indeed, from the best diagnostic instruments; such differentiation remains an important challenge for the future.

Diagnostic value of different features

Our findings are also informative about the diagnostic value of different behavioural features. The findings are clear-cut in showing that each of the three main domains of symptoms (social deviance, communication deficits, and repetitive behaviours) serves to differentiate PDD, but that the best differentiation is provided by the total ASQ score which includes all three domains. The factor analytic findings are provocative, however, in their indication that the communication items span three factors. A few are included in the first factor, which is largely concerned with social deviance, emphasising that the communication abnormalities are closely connected with problems in social interaction. Many communication items load on the second factor, but those concerned with qualitatively abnormal language features (such as verbal rituals and pronoun reversal) load separately on factor 3. The implication is that it may be useful to differentiate between language deviance and language deficit, although both are a part of autism. However, it may be that the findings simply reflect the currently poor discriminative validity of these items on the ASQ.

A relatively weak diagnostic differentiation was provided by repetitive stereotyped behaviours, most items of which loaded on factor 4. This parallels the ADI and ADOS findings. Although, clinically, this constitutes a key feature of autism (indeed it is one especially emphasised by Kanner (1943), in his original description of the syndrome), it has proved quite difficult to devise an effective means of measurement that works equally well with people with severe retardation and those of normal non-verbal intelligence. Some stereotyped behaviours are quite common in individuals with severe retardation and, although those differ in quality from those most strongly associated with autism, it has not proved easy to reflect this in item construction. Many other repetitive behaviours differentiate autism more clearly but most have quite low base rates even within autistic groups. In many respects, the greatest problem lies with those individuals with autism who are of normal intelligence. Their stereotyped behaviour is more likely to be manifest in circumscribed interest patterns than in the grosser forms of repetitive behaviour. However, the ASQ, like the ADI–R, has only one item on circumscribed interests. We conclude that there is a need to develop better measures for the repetitive features in the behaviour of individuals with mild autism of normal non-verbal intelligence.

Implications

The findings on the value of the ASQ as a screening questionnaire may be summarised as follows. In an unselected general population sample aged at least four years, scarcely any children will have a score of 15 or greater. Accordingly, there will be almost no false positives. Conversely, most individuals with autism have an ASQ score well above 15 (we found a mean of 24.2). However, a significant minority have scores near the 15 mark, and a cut-off greatly above the 15 mark would result in an unacceptable increase in false negatives. The ASQ scores of children with specific or general learning disabilities are substantially above general population norms (reflecting the fact that, as shown by detailed studies, many show some autistic features, albeit well short of the criteria for the diagnosis of autism, although not so far outside those for PDD). At a group level, the ASQ scores of individuals with PDD and of those with other diagnoses are very different (we found means of 22.3 *v.* 11.2), but a substantial minority of individuals with severe non-autistic developmental disorders will have an ASQ score of 15 or greater. The lower the mental age, the greater the chance of a false positive. Even so, the ASQ was found to provide a significant diagnostic differentiation at all IQ levels. In short, the ASQ is a highly effective screening instrument for children aged four years or above but, like any other questionnaire, it cannot be expected to provide individual diagnoses.

ACKNOWLEDGEMENTS

We are most grateful to the families who helped us with this study and to Professor Anthony Cox, Dr Margaret Murphy and Dr Anne Gilchrist, who kindly gave us access to their ADI data and allowed us to approach families under their care to ask them to complete the ASQ. We also thank Dr Reyhan Bilgic for statistical help in relation to the factor analyses.

REFERENCES

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM–IV). Washington, DC: APA.

Baron-Cohen, S., Allen, J. & Gillberg, C. (1992) Can autism be detected at 18 months? The needle, the haystack and the CHAT. *British Journal of Psychiatry*, **161**, 839–843.

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press

Bolton, P., Macdonald, H., Pickles, A., et al (1994) A case–control family history of autism. *Journal of Child Psychology and Psychiatry*, **35**, 877–900.

Dahlgren, S. O. & Gillberg, C. (1989) Symptoms in the first two years of life: a preliminary population study of infantile autism. *European Archives of Psychiatry and Neurological Sciences*, **238**, 169–174.

Di Lavore, P., Lord, C. & Rutter, M. (1995) The Pre-Linguistic Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, **25**, 355–379.

Fombonne, E. (1991) The use of questionnaires in child psychiatry research: measuring their performance and choosing an optimal cut-off. *Journal of Child Psychology and Psychiatry*, **32**, 677–693.

Hanley, J. A. & McNeil, B. J. (1982) The meaning and the use of the area under a receiver operating characteristics (ROC) curve. *Radiology*, **143**, 29–36.

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

Krug, D. A., Arick, J. & Almond, P. (1980) Behaviour checklist for identifying severely handicapped individuals with high levels of autistic behaviour. *Journal of Child Psychology and Psychiatry*, **21**, 221–229.

Le Couteur, A., Rutter, M., Lord, C., et al (1989) Autism Diagnostic Interview: a standardized investigator-based instrument. *Journal of Autism and Developmental Disorders*, **19**, 363–387.

Lord, C., Rutter, M., Goode, S., et al (1989) Autism Diagnostic Observation Schedule: a standardized observation of communicative and social behaviour. *Journal of Autism and Developmental Disorders*, **19**, 185–212.

—, — & Le Couteur, A. (1994) Autism Diagnostic Interview – Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, **24**, 659–685.

Mullen, E. (1989) *Mullen Scales of Early Learning*. Cranston, RI: TOTAL Child.

Parks, S. L. (1988) Psychometric instruments available for the assessment of autistic children. In *Diagnoses and Assessment in Autism* (eds E. Schopler & G. B. Mesibov), pp. 123–136. New York: Plenum.

Raven, J. C., Court, J. H. & Raven, J. (1990) *Coloured Progressive Matrices*. Oxford: Oxford Psychologist Press.

CLINICAL IMPLICATIONS

- The Autism Screening Questionnaire, a 40-item parental questionnaire, provides a useful screening measure for pervasive developmental disorders.
- It may be useful to differentiate between language deviance and language deficit, although both are associated with autism.
- The differentiation between autism and other pervasive developmental disorders is problematic.

LIMITATIONS

- The findings derive from a sample whose parents had already been interviewed using the original or revised version of the Autism Diagnostic Interview. Replication with naïve parents is needed.
- Retest reliability data are needed.
- Further findings with very young children are needed.

SIBEL KAZAK BERUMENT, PhD, MICHAEL RUTTER, FRS, CATHERINE LORD, PhD, ANDREW PICKLES, PhD, ANTHONY BAILEY, MB, BS, MRC Child Psychiatry Unit and Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, London

Correspondence: Dr Michael Rutter, Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, De Crespigny Park, London SE5 8AF

(First received 4 January 1999, final revision 8 April 1999, accepted 20 April 1999)

Rimland, B. (1964) *Infantile Autism: The Syndrome and its Implications for a Neural Theory of Behavior*. New York: Appleton–Century–Crofts.

Rutter, M., Andersen-Wood, L., Beckett, C., et al (1999) Quasi-autistic patterns following severe early global privation. *Journal of Child Psychology and Psychiatry*, **40**, 537–549.

Schopler, E. & Mesibov, G. B. (eds) (1988) *Diagnosis and Assessment in Autism*. New York: Plenum.

Volkmar, F. R., Cicchetti, D. V., Dykens, E., et al (1988) An evaluation of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, **18**, 81–97.

Wechsler, D. (1989) *Wechsler Pre-School and Primary Scales of Intelligence–Revised*. San Antonio, TX: Psychological Corporation.

— (1991) *Manual for the Wechsler Child Intelligence Scale–Revised*. San Antonio, TX: Psychological Corporation.

World Health Organization (1992) *The Tenth Revision of the International Classification of Diseases and Related Health Problems* (ICD–10). Geneva: WHO.

https://doi.org/10.1192/bjp.175.5.444 Published online by Cambridge University Press