
Diagnostic instruments for autism spectrum disorder

A brief review

April 2011



April 2011

Prepared for the
Ministries of Health and Education
by the New Zealand Guidelines Group.

Authors

Joanna M McClintock PhD, Clinical Psychologist
John Fraser, Manager Implementation Services, NZGG



Contents

Introduction	1
1. Instruments reviewed	2
1.1 Instruments for autism	2
1.2 Instruments for Asperger's disorder	2
2. Review of instruments	4
2.1 Autism Diagnostic Interview-Revised (ADI-R)	5
2.2 Autism Diagnostic Observation Schedule (ADOS)	6
2.3 Childhood Autism Rating Scale (CARS)	7
2.4 Gilliam Autism Rating Scale – Second Edition (GARS-2)	9
2.5 Social Communication Questionnaire	10
2.6 Social Responsiveness Scale	11
2.7 Developmental, Dimensional and Diagnostic Interview (3di)	12
2.8 Diagnostic Interview for Social and Communicative Disorders	13
2.9 Gilliam Asperger's Disorder Scale – 2003 Update	14
2.10 Asperger Syndrome Diagnostic Scale	15
2.11 Krug Asperger's Disorder Index	16
2.12 Autism Spectrum Screening Questionnaire	17
3. Potentially preferable combinations of instruments: criteria for preferences	20
4. Potentially preferable combinations: autism	21
4.1 Autism: screening	21
4.2 Autism: diagnosis	21
5. Potentially preferable combinations: Asperger's disorder	23
5.1 Asperger's disorder: screening	23
5.2 Asperger's disorder: diagnosis	23
6. Conclusion	24
References	25

List of Tables

1	Summary of the instruments reviewed	19
2	Potentially preferable instrument combinations for the screening and diagnosis of autism	22

Introduction

In April 2008, New Zealand was the first country in the world to publish a clinical guideline for the diagnosis and management of autism spectrum disorders (ASD). The *New Zealand Autism Spectrum Disorder Guideline* was launched on World Autism Awareness Day, 1 April 2008. The guideline recommends a comprehensive assessment for the diagnosis of ASD, including the use of validated diagnostic instruments for use in combination with expert clinical judgment.

This document:

- a. contains a brief review of certain of the available instruments, describing their basic characteristics including appropriate use and setting, statistical properties, requirements in terms of user qualifications and training, and licensing arrangements. It then,
- b. sets out some potentially preferable combinations of instruments for screening and diagnosis of autism, and for the screening for Asperger's disorder.¹

It is intended as a reference resource for practitioners in the health, disability and education sectors. It is important not to infer that diagnostic instruments alone are adequate for the recognition and diagnosis of autism. Diagnosis can be made only by experienced clinicians, integrating information from their expertise and training, from clinical findings, and from information collected from the person being assessed, their family/whānau and the person's referrer. Instruments are an important aid to diagnosis, not a substitute for clinical expertise.

The following organisations endorse the content of this review and draw their members' attention to it as relevant for practice:

- New Zealand Association of Occupational Therapists
- New Zealand College of Clinical Psychologists
- New Zealand Psychological Society
- New Zealand Speech Language Therapists Association.

¹ All of the instruments for assessing the likelihood of Asperger's disorder stop short of making a definitive diagnosis (see section 2).

1. Instruments reviewed

1.1 Instruments for autism

Instruments reviewed here for screening and diagnosis of autism were selected on the basis that they are all:

- listed as appropriate in the New Zealand Autism Spectrum Disorder Guideline (Ministry of Health, 2008); and,
- listed in the 'Practice parameters for the screening and diagnosis of autism' by the American Academy of Neurology and the Child Neurology Society (Filipek et al, 2000), current as of July 2006 (National Guidelines Clearinghouse, nd).

The instruments meeting these criteria are:

- Autism Diagnostic Interview-Revised (Lord, Rutter and Le Couteur, 1994)
- Autism Diagnostic Observation Schedule (Lord, et al, 2000)
- Childhood Autism Rating Scale (Schopler, Reichler, DeVellis, Daly, 1980)
- Gilliam Autism Rating Scale – Second Edition (Gilliam, 2005)
- Social Communication Questionnaire (Rutter, Bailey and Lord, 2003)
- Social Responsiveness Questionnaire (Constantino and Gruber, 2005)
- Developmental, Dimensional and Diagnostic Interview (Skuse, Warrington, Bishop, Chowdhury, Lau, Mandy, Place, 2004)
- Diagnostic Interview for Social and Communicative Disorders (Wing, Leekam, Libby, Gould and Larcombe, 2002).

Many of the instruments assessed here are used in New Zealand. However, none of them has been formally validated in New Zealand for use with the New Zealand population. This report does not address this issue.

1.2 Instruments for Asperger's disorder

Instruments reviewed for assessment of Asperger's disorder were selected on the basis that they are:

- listed as appropriate in the New Zealand Autism Spectrum Disorder Guideline (Ministry of Health, 2008); and,
- reviewed in the Mental Measurement Yearbook Tests Online, from the Buros Institute at the University of Nebraska; the Yearbook provides independent, expert testing and review of instruments.

The instruments meeting these criteria are:

- Gilliam Asperger's Disorder Scale (2003 Update; Gilliam, 2003)
- Asperger Syndrome Diagnostic Scale (Myles, Bock and Simpson, 2001)
- Krug Asperger's Disorder Index (Krug and Arick, 2003)
- Autism Spectrum Screening Questionnaire (Ehlers, Gillberg and Wing, 1999).

2. Review of instruments

The review of each instrument canvasses:

- administrative issues such as ease of administration and scoring, required experience or qualifications for assessors, and the duration of assessments
- The 'comparison sample' population or populations of people with or without ASD in which each instrument was trialled and developed
- a summary of performance statistics for each instrument, wherever possible addressing each of:
 - instrument sensitivity (power to detect a person who does have ASD) (Greenhalgh, 1997)
 - instrument specificity (power to exclude a person who does not have ASD) (Greenhalgh, 1997)
 - instrument reliability, that is:
 - inter-rater reliability (degree to which to different assessors get consistent results)
 - temporal stability (consistency across time of the instrument)
 - internal consistency (consistency of results across different items within the test)
 - instrument validity, that is:
 - content validity (how appropriately the items in the instrument measure ASD)
 - construct validity (assessment of how the instrument reflects theories of ASD)
 - discriminant validity (whether an instrument can discriminate between two distinct phenomena)
- licensing arrangements.

Regarding user training, most of the instruments reviewed are available from more than one distributor in Australasia. Commonly, it is distributors who define criteria for levels of training or expertise required for competent use of an instrument. These criteria can vary between distributors for the same instrument, and can vary also from recommendations made by the publisher of the instrument. On the basis that flexibility as to use is desirable, the 'licensing arrangements' sections summarise only those licensing arrangements, current as at October 2008, which allow for the widest use of each instrument.

2.1 Autism Diagnostic Interview-Revised (ADI-R)

The Autism Diagnostic Interview-Revised (ADI-R) (Lord et al, 1994) is a semi-structured interview designed to assess the three core aspects of ASD; social, communication, and restricted behaviours or interests. The ADI-R is designed for individuals aged 18 months and older (Lord et al, 1994), and can be used for treatment and educational planning, regardless of whether a diagnosis of ASD is obtained.

2.1.1 Administration issues

The ADI-R is designed to be administered by an experienced clinician and is for a parent or caregiver who is very knowledgeable about the individual being assessed. The ADI-R consists of 93 items covering areas of family background, developmental history, language, communication, social development, interests, and general behaviour (Le Couteur, Lord, Rutter, 2003). As the procedure is standardised it needs to be followed carefully, and the interviewer records and codes responses based on descriptions of behaviours by caregiver. Definitions of the behaviours being assessed are provided in the ADI-R which allows for more accurate coding. An algorithm is used to code the interview items and summary scores are provided for the four domains required for diagnosis: reciprocal social interaction; communication; restricted, stereotypic behaviour; and age of presentation. Cut-off scores are then used to determine the presence of ASD; there are not separate cut-offs for autism and ASD (Naglieri and Chambers, 2009).

The interview can take between two and three hours to administer. The authors advocate for training and clinical experience when using the ADI-R. Training programmes with guidebooks and exercises are available from the publishers.

2.1.2 Comparison sample

Information about the comparison sample for the ADI-R is limited (Naglieri and Chambers, 2009). The ADI-R was administered to several hundred caregivers of individuals with and without autism, aged from preschool to early adulthood. No further information about the standardised sample, such as ethnicity or spoken language is provided.

2.1.3 Statistics

For each domain, a range of sensitivity (.86–1.0) and specificity (.75–.96) values are reported for various combinations of score (either total score or cut-off score) and language ability of the individual assessed. Lord et al (2004) report internal consistency for each assessment of each domain as follows: social (.95); restricted and repetitive behaviours (.69), verbal (.85) and communication (.84). Agreement over time on a sample of six mothers by different interviewers was 91% over a two to three month period. Naglieri and Chambers (2009) also report that

the manual indicates test-retest reliability for a sample of 94 preschool children over a 2 to 5 month period. Rutter, Le Couteur, and Lord (nd) state that the ADI-R has proven effective in differentiating autism from other developmental disorders, as well as assessing syndrome boundaries, and quantifying autistic symptomatology.

2.1.4 Licensing

The ADI-R is published by the Western Psychological Services, and is currently available through the Australia Council of Educational Research (ACER) and is available to professionals with accredited training in psychology, health sciences, counselling, education, medicine and other specialist areas.

2.2 Autism Diagnostic Observation Schedule

The Autism Diagnostic Observation Schedule (ADOS) is an observation measure designed to assess reciprocal social interaction and communication, play, and use of imagination (Lord et al, 2000). The ADOS attempts to set a 'social world' in which behaviours associated with ASD can be observed via play, tasks, and conversation. The ADOS can be used to assist with educational planning (Lord et al, 2000). The ADOS was originally developed to be used in conjunction with the Autism Diagnostic Interview (ADI). This combination of instruments has been deemed the 'gold standard' for the assessment of ASD (Filipek, 2000).

2.2.1 Administration issues

Within the ADOS there are four 'modules', one of which is administered depending on the individual's verbal ability. Module 1 is used for children who are preverbal or have single-word language. Module 2 is appropriate for individuals with phrase speech abilities. Module 3 is used for children and adolescents who are verbally fluent. Verbally fluent adolescents and adults would be assessed with Module 4 (Lord et al, 2000; Naglieri and Chambers, 2008). A module takes approximately half an hour to complete and more than one module can be administered if the examiner judges that a more or less advanced module is appropriate. Each item is typically scored on a three-point scale, from no evidence of abnormality (0) to definite evidence (2) and abnormalities so severe that interfere with observation (3). There is standardised information about procedures, rating scales and age focus, and operationalised criteria are provided, as well as detailed descriptions of the behaviour being investigated (Lord et al, 2000).

An ADOS user requires a high level of knowledge and clinical experience of ASD, as well as experience in working with children. Training in the use of the ADOS is also required, because the behaviours of the examiner must be standardised and clinical judgment is required as to when a child can be 'pressed' (Naglieri and Chambers, 2009; Nebel-Schwalm and Matson, 2008). If using the ADOS for research purposes then personal attendance at training workshops is required. For solely clinical purposes a DVD training package is available.

The ADOS classification system does not assign a diagnosis. The ADOS has thresholds for social interaction, communication and communication-social interaction total; a person may reach the threshold on all three scales but not receive

a diagnosis of ASD, because of late presentation of difficulties or no problems in the area of restricted behaviours. The authors stress the importance of using the ADOS in conjunction with a developmental history, corroborating information from other sources, and the use of clinical judgment (Lord et al, 2000).

2.2.2 Comparison sample

The comparison sample included individuals with diagnoses of autism, pervasive developmental disorder, not otherwise specified (PDD-NOS) and individuals not on the spectrum. This sample is from America, and includes European American, African American, Hispanic, Asian American and mixed ethnic groups. All members of this sample had English as their first language (Lord et al, 2000).

2.2.3 Statistics

A range of sensitivity and specificity data is provided depending on cut-off scores. Typically the instrument has sensitivity in the upper 90% range and specificity in the upper 80% to lower 90% range (Naglieri and Chambers, 2009). Agreement between raters for diagnostic classification when assessing individuals with autistic disorder, PDD-NOS, and non-ASD ranged from 81% to 93% for the four modules.

Lord et al calculated internal consistency for all domains and modules, which ranged from .47 to .94. The lower results were found for stereotyped behaviours and restricted interests in module 3. Excellent inter-rater reliability within each module (.65–.78) and test-retest reliability (.59–.82) over an average period of nine months is also reported.

2.2.4 Licensing

The ADOS is published by the Western Psychological Services, and is currently available from the Australia Council of Educational Research (ACER) and is available to professionals with accredited training in psychology, health sciences, counselling, education, medicine and other specialist areas.

2.3 Childhood Autism Rating Scale

The Childhood Autism Rating Scale (CARS) is an observation instrument and was developed to identify children with autism compared to children with other developmental disabilities and determine symptoms severity (Schopler, Reichler and Rothen Renner, nd). The CARS was developed for children over the age of two years. As no upper age limit is provided it is not clear at what age the CARS would be inappropriate.

The items for the CARS were derived from five different theoretical perspectives of autism, including Kanner's description of the syndrome to the DSM-III-R. Due to this, the CARS does not include some constructs considered important to the diagnosis of autism (Ozonoff, Goodlin-Jones and Solomon, 2005). The CARS is also unable to

clearly differentiate Asperger's disorder or pervasive developmental disorders, however, as Rellini et al (2004) point out the CARS was developed prior to the concept of an autism spectrum.

2.3.1 Administration issues

The CARS consists of 15 four-point scales (or seven-point scale if half points values are used) where a child's behaviour is rated for chronological age. Specific descriptive examples are provided for each of the behaviours being assessed. The scores are then summated to categorise a child on a continuum from 'non-autistic' to 'mild to moderate' to 'severe autism' (Prizant, nd). No specific training is required in the use of the CARS.

The CARS was designed as a screening tool that requires minimal training and can be used by a range of professionals. It is important to note that the CARS does assume the assessor has knowledge of chronological age-appropriate functioning across the domains assessed (Prizant, nd).

2.3.2 Comparison sample

The comparison group to develop the scores on the CARS was children with autism participating in a North Carolina programme, *Treatment and Education of Autistic and Related Communication-Handicapped Children (TEACCH)* (Prizant, nd).

2.3.3 Statistics

Nebel-Scwalm and Matson (2008) summarise reliability and validity research for the CARS. Internal consistency reliability (.94), test-retest for CARS diagnoses (.64) and for CARS scores (.81), and inter-rater reliability (.71), are all acceptable. Validity studies have shown no difference between psychoeducational testing and other techniques [interview (.75), observation (.86) and case history review (.63)]. Validity tests have demonstrated the use of the CARS across professionals from different disciplines (.81). Rellini et al (2004) compared the CARS with DSM-IV criteria and found 100% sensitivity for children with autism. Scores on the CARS can be skewed by an individual's cognitive function. If an individual had limited cognitive abilities they would get a higher score on the CARS regardless of whether they demonstrated difficulties with social interaction (Prizant, nd).

2.3.4 Licensing

The CARS is published by the Western Psychological Services, and is currently available through the Pearson Clinical and Talent Assessment Australia and New Zealand. Assessors need to be registration level B or M, which is an allied health or special education professional, or a medical practitioner. Demonstration and practice videos are also available for assessors.

2.4 Gilliam Autism Rating Scale – Second Edition

The Gilliam Autism Rating Scale – Second Edition (GARS-2) was developed for use in screening and diagnosis of individuals with autism and to assist in educational planning and research (Gilliam, nd). The second edition of the GARS provides specific information on instructional objectives and has revised subscales (Ward-Fairbank, nd). The GARS-2 is for use with people aged three to 22 years of age.

2.4.1 Administration issues

The GARS-2 contains three subscales: stereotyped behaviours, communication and social interaction, and is based on definitions of autism from the Autism Society of America and the diagnostic criteria of the DSM-IV-TR. The GARS-2 is administered through three different techniques, a parent interview, observation, and the assessor answering key questions and interpretations. The three subscales are summed to provide an autism index. A higher scores and autism indices are suggestive of problematic behaviour. There is also a probability of autism classification (very likely, possibly, unlikely). These scales are reportedly vague (Garro, nd). The measure takes approximately 10 to 15 minutes to complete. No extra training is required.

2.4.2 Comparison sample

The validation sample for the GAR-2 consisted of 1107 individuals between three and 22 who had been diagnosed with autism (as rated by staff from schools and centres specialising in autism, and other disabilities, as well as data from the Asperger Syndrome Information and Support website) (Gilliam, nd; Ward-Fairbank, nd). The sample is similar to the 2001 US Census with reference to geographic region, and ethnicity. The gender ratio was reportedly similar to that of individuals with ASD (81% male). However, the age was skewed to younger individuals, with more than half of the sample being aged three to eight years (Ward-Fairbank, nd).

2.4.3 Statistics

Sensitivity and specificity data are not available for the GAR-2. Good internal consistency reliability (.94 total test) is reported, and although test-retest reliability was good (.84 for Autism Index), the sample size for assessing stability over time was small (n=37 subjects) (Ward-Fairbank, nd).

There is reportedly evidence for construct and content validity, however, as noted by Garro (nd) there needs to be more research on the differentiation of the subscales. The GAR-2 may have a tendency to 'over-diagnose' because it departs from theoretical frameworks of ASD. Preliminary data supports the use of the Autism Index as a general diagnostic indicator. More research is required on the psychometric properties of the GAR-2 with reference to inter-rater reliability and construct and criterion-related validity. Garro (nd) recommends GAR-2 as a screening tool for pervasive developmental disorders, including ASD.

2.4.4 Licensing

The GAR-2 is published by the Western Psychological Services, and is currently available from Psychological Assessments Australia and requires a user level C, which means users must be 'graduates in a field related to the area of test usage, and have had some experience and/or additional training in test administration and interpretation'. Non-graduates with extensive relevant experience and the completion of appropriate training programmes will also be considered for registration at this level.

2.5 Social Communication Questionnaire

The Social Communication Questionnaire (SCQ) is a rating scale developed to assess symptoms associated with ASD (Rutter, Bailey and Lord, nd). An older version was called the Autism Screening Questionnaire. The SCQ is based on the DSM-IV and the content of the ADI-R and the items have identical words. The SCQ was developed as a screening tool for ASD, as well as for research, and progress of symptoms.

2.5.1 Administration issues

Rather than as a structured interview, the SCQ involves yes/no questions and takes approximately 10 minutes to administer. The raw scores are then added for a total score which is interpreted based on appropriate cut-off scores. There are two formats: the lifetime and current behaviour. The lifetime form takes into account developmental history compared to the current behaviour form that looks at behaviours in the last three months. Administration takes approximately 10 minutes (Rutter et al, nd).

2.5.2 Comparison sample

The group from which cut-off scores were determined consisted of 200 children who had a range of developmental disabilities and clinical diagnoses. The same sample had been involved in studies using the ADI-R (Naglieri and Chambers, 2009).

2.5.3 Statistics

Sensitivity and specificity data are provided for a range of cut-off scores. Using a total score of 15 or higher for differentiating ASD from other diagnoses, sensitivity of .85 and specificity of .75 are reported. Using the same cut-off for differentiating autism from other diagnoses (excluding intellectual disability), sensitivity of .96 and specificity of .80 are reported. Using a cut-off score of 22 for differentiating autism from other ASD, sensitivity of .75 and specificity of .60 are reported (Naglieri and Chambers, 2009).

Naglieri and Chambers (2009) report on the reliability and validity of the SCQ and comment that internal consistency has been demonstrated by the authors (.81–.93 from two different studies). The SCQ has also been shown to differentiate children with ASD from children with other disabilities. There is also a high correlation

between the SCQ and ADI-R with an average agreement between the items on the two tests of 70.8% (Naglieri and Chambers, 2009).

2.5.4 Licensing

The SCQ is published by Western Psychological Services, and is currently available through Australia Council of Educational Research (ACER) and is available to professionals with accredited training in psychology, health sciences, counselling, education, medicine and other specialist areas.

2.6 Social Responsiveness Scale

The Social Responsiveness Scale (SRS) was developed to identify ASD in children, and to screen and support clinical diagnoses. The domains of the questionnaire are social behaviour, communication and repetitive behaviours associated with ASD. The SRS was designed for children between four and 18 years, and can be used to assist with diagnosis, for research, and also to measure progress of symptoms (Constantino, nd).

2.6.1 Administration issues

The questionnaire consists of 65 items which are rated on a scale of 1 (not true) to 4 (true). There are forms for both parents and teachers to complete. The items are combined for a overall score. There are also five treatment subscale scores which can aid with treatment planning. Administration and scoring can be completed in about 20 minutes (Constantino, nd). Case examples are provided in the manual to guide how the test can be applied to different diagnostic categories in ASD.

2.6.2 Comparison sample

To determine psychometric properties and cut-off scores, a comparison sample was developed from five different large research projects, involving over 1600 children from America. Results from the five different groups were compared to ensure that they could be combined. The scoring is standardised and there are data for parents and teacher reports, as well as for both male and females. Reviewers of the SRS (Venn, nd; Conway, nd) caution about how the items were developed for the instrument and the way in which the normative information was developed as the sample may not represent the general population, which needs to be taken into consideration by users.

2.6.3 Statistics

The SRS has been found to discriminate children with ASD from children with other psychiatric diagnoses, and sensitivity (.85) and specificity (.75) for any ASD as rated by expert clinicians (Naglieri and Chambers, 2009). Good internal consistency (.93–.97 for parent, teacher and clinical ratings), over 17 months test-retest reliability was .85 for males and .77 for females. The agreement between raters has all been

reported (parent-parent = .91; mother-teacher = .92; father-teacher = .75) (Conway, nd).

2.6.4 Licensing

The SRS is published by Western Psychological Services and is currently available through Australian Council for Educational Research (ACER). There are currently no restrictions on qualifications of the user.

2.7 Developmental, Dimensional and Diagnostic Interview

The Developmental, Dimensional and Diagnostic Interview (3di) (3di; Skuse, Warrington, Bishop, Chowdhury, Lau, Mandy, Place, 2004) is a computerised interview designed to assess symptoms of ASDs from a dimensional perspective, so taking into account dimensions of impairment. The 3di is a parental interview which can be used in both clinical and non-clinical samples.

2.7.1 Administration issues

The interview consists of 183 items, and covers areas of demographics, family, developmental history and motor skills. There are 266 questions that are concerned with disorders on the autism spectrum, and 291 questions that relate to mental state relevant to other diagnoses. The questions need to be asked in the way that they are written and the authors have attempted to make them sound as natural as possible. It is possible to abbreviate the questionnaire if a specific diagnosis is suspected or if certain modules are not relevant (eg, verbal questions for a non-verbal individual) (Skuse et al, 2004).

The full interview takes 90 minutes to administer. The authors have developed a 'pre-interview package' for parents to complete in advance and this information can be entered into the computer. The child can then be assessed using an abbreviated interview which takes 45 minutes to complete (Skuse et al, 2004). The 3di also establishes comorbidity across a full range of child psychiatric disorders.

2.7.2 Statistics

Sensitivity (1.0) and specificity (0.97) were both near perfect for the 3di, for discriminating between the autism spectrum and non-autism spectrum. Test retest and inter-rater reliability were excellent, with most correlation coefficients being greater than 0.9. There was high agreement (0.92) for autism (including Asperger syndrome), atypical autism, and PDD-NOS (Skuse et al, 2004).

Content validity was demonstrated by the items being selected based on literature, and the DSM and ICD classification systems. In a clinical study of 60 children, of whom 27 had an ASD diagnosis by clinician, the 3di diagnosed 29 children as having a significant degree of autism disorder based on ICD-10 criteria for childhood autism, atypical autism, or PDD-NOS (Skuse et al, 2004).

2.7.3 Licensing

The 3di is available from the publishers and training is required in the use of the 3di as part of purchase. The authors have developed semi-automated training using DVD technology and a DVD/internet-based system to ensure that users are using the 3di appropriately. The developers provide ongoing support for one year after purchase. The course is suitable for medical, health or education professionals who have some experience of working with children with autism (Warrington, personal communication, 26 June 2009).

2.8 Diagnostic Interview for Social and Communicative Disorders

The Diagnostic Interview for Social and Communicative Disorders (DISCO) is a semi-structured interview designed to assess impairments in the areas of social interaction, social communication and social imagination, and repetitive behaviours associated with ASD. The DISCO is for children and adults (Wing, Leekam, Libby, Gould and Larcombe, 2002). The DISCO has been revised 10 times and was developed independently of classification systems such as the ICD and DSM (Leekam, Libby, Wing, Gould and Taylor, 2002).

2.8.1 Administration issues

The DISCO takes a dimensional approach and obtains a profile of development and behaviour, as well as identifying specific features associated with ASD. Information about developmental history and current functioning is obtained. The interview takes between two and three hours to complete. The questions are flexible so can be adapted depending on the individual's level of functioning, prior information, and cultural background. That the DISCO is not associated with diagnostic systems ensures that it will still be relevant as changes get made to the DSM and ICD systems (Leekam et al, 2002; Wing et al, 2002). A range of algorithms are provided to assist with a diagnosis of ASD, pervasive developmental disorders and/or psychiatric disorders, and are based on the DSM-III-R, DSM-IV and ICD-10 (Wing et al, 2002). Training is available on the use of the DISCO.

2.8.2 Statistics

The authors report high inter-rater reliability (>.75). Wing et al (2002) and Leekam et al (2002) have also demonstrated that the DISCO is valid in differentiating ASD from other developmental and psychiatric disorders, using the ASD algorithm, rather than the ICD-10 algorithm, as this failed to accurately discriminate between some children with ASD and some without ASD (Nebel-Scwalm and Matson, 2008).

2.8.3 Licensing

The DISCO is available from the authors. It is unclear what level of education is required for administration.

2.9 Gilliam Asperger's Disorder Scale – 2003 Update

The 2003 update of the Gilliam Asperger's Disorder Scale (GADS) is based on the 2001 edition, and is a 32-item behaviour checklist developed to identify people with Asperger's disorder, assess unique behaviour difficulties, develop education plans, document progress, and measures Asperger's disorder in research. The GADS was designed for use with people aged three to 22 years (England, nd).

2.9.1 Administration issues

The items are specific and observable behaviours, and are rated using frequency-based ratings within a six-hour period, ranging from never observed, seldom observed (1–2 times), sometimes observed (3–4 times), frequently observed (5 or more times). The 32 items are divided into four subscales: social interaction, restricted patterns, cognitive patterns and pragmatic skills, as well as a parent interview form (the interview form replaces the early development subscale from the 2001 edition). Subtest scaled scores are summed to get an Asperger's disorder quotient, where scores indicate probability of the presence of Asperger's disorder, ranging from low/not probable, borderline and high score/probable. Administration time is approximately five to 10 minutes (England, nd; McGregor, nd).

2.9.2 Comparison sample

The GADS was standardised on a sample of 371 people aged between three and 22 years who had a previous diagnosis of Asperger's disorder. An international sample was used, including United States of America, Britain, Mexico, and Australia. The majority of the sample was between the age of seven and 13 years. Independent diagnoses were not obtained for any of the individuals in the standardisation sample (McGregor, nd).

2.9.3 Statistics

No sensitivity or specificity data are reported for the GADS (Campbell, 2005). Internal consistency is reported at 0.87 for the Asperger's disorder quotient, which is lower than recommended for a reliable test (Bracken, 1987). Test-retest reliability (.93) and inter-rater reliability (.77) are reported to be adequate. However, the sample sizes are very small (10 individuals for test-retest and 16 individuals for inter-rater), so need to be considered cautiously.

Content, construct and discriminant validity are reported by the authors of the GADS. However, England (nd) cautions about assuming validity because of the subjective nature of rating the behaviours being rated, and sample sizes in the validity studies were very small, and therefore, may not accurately discriminate people with Asperger's disorder from other behavioural disorders.

2.9.4 Licensing

The GADS is published by the Western Psychological Services, and is currently available from Psychological Assessments Australia and requires a user level C, which means users must be 'graduates in a field related to the area of test usage, and have had some experience and/or additional training in test administration and interpretation'. Non-graduates with extensive relevant experience and the completion of appropriate training programmes can be considered for registration at this level.

2.10 Asperger Syndrome Diagnostic Scale

The Asperger Syndrome Diagnostic Scale (ASDS) is a 50-item rating scale developed to aid in the identification of people with Asperger syndrome, document progress, assist with educational programmes, and for research purposes. The ASDS is for use with people aged five to 18 years.

2.10.1 Administration issues

The ASDS consists of 50 items and can be completed by someone who knows the individual well and has had close contact with them in the previous two weeks. The respondent rates the presence or absence of each of the behaviours. The ASDS has five subscales: language, social, maladaptive, cognitive and sensorimotor. These subscales are not recommended for diagnostic purposes, but rather for clinical interest. The subscale scores are summed to yield an 'Asperger syndrome quotient', which is used to determine the probability of diagnosis of Asperger syndrome, with probability rated as being very likely, likely, possibly, unlikely and very unlikely. Administration time is approximately 10 to 15 minutes.

2.10.2 Comparison sample

The ASDS was standardised using a sample of 115 individuals aged five to 18 years from American who had a previous diagnosis of Asperger's disorder. The biggest criticism of the ASDS is that an independent diagnosis of Asperger's disorder was not obtained for the standardisation sample. In addition, no details are given as to the cognitive functioning of the standardisation sample, which impacts upon the cognitive and language subscales.

2.10.3 Statistics

Sensitivity and specificity data are not reported for the ASDS (Campbell, 2005). Internal consistency is reported at 0.83 for the Asperger syndrome quotient which is below the standard recommended for acceptable reliability (Bracken, 1987). Inter-rater reliability is reported 0.93, however, only 14 pairs of teacher-parent raters were compared. Temporal stability data is not available.

Content and construct validity are demonstrated by the items for the test being derived from the DSM and ICD classification systems, as well as seminal research on Asperger's disorder and current literature. Discriminant validity was reported by comparing the standardised sample with a sample of 177 people who had a range of

disorders including autism, behavioural difficulties, attention deficit hyperactivity disorder, and learning disabilities. Eighty-five percent of the participants were correctly identified using the Asperger syndrome quotient. However, the diagnoses of the comparison sample were also not independently verified.

2.10.4 Licensing

The ASDS is published by the Western Psychological Services, and is currently available through the Psychological Assessments Australia and requires a user level B, which means users must be 'registered psychologists or four-year psychology graduates undergoing supervision'.

2.11 Krug Asperger's Disorder Index

The Krug Asperger's Disorder Index was developed to 'distinguish individuals with Asperger's disorder from those who have other forms of high-functioning autism' (Krug and Arick, nd), as well as assist with education planning and for research purposes. The KADI has forms for two different age brackets, children aged six to 11 years and people aged 12 to 21 years.

2.11.1 Administration issues

The KADI consists of 32 items, with the assessor indicating the presence or absence of behaviours that are indicative of Asperger's disorder. The raw scores are weighted and summed to yield a total standard score which indicates the likelihood of a diagnosis of Asperger's disorder. The higher the score the higher is the probability that the individual has Asperger's disorder. There are two groups of items, the first is a 'prescreening' scale which 'immediately identifies normal individuals' (Krug and Arick, nd); if the score on these 11 items does not exceed 18 then the assessor does not continue with the assessment. A parent, teacher or caregiver can complete the KADI at school or home and takes approximately 15 to 20 minutes.

2.11.2 Comparison sample

The authors report that the KADI was standardised on a sample of 486 individuals, with 130 having diagnoses of Asperger's disorder 162 diagnosed with autism, and 194 individuals described as 'normal' from 30 states in America and 10 countries. In the mental measurements review Nellis (nd) reports that 'careful reading revealed' the standardisation sample was only the 130 individuals with Asperger's disorder. The diagnoses were not independently confirmed.

2.11.3 Statistics

Sensitivity for the KADI is reported at .78, and specificity as .94 (Campbell, 2005). Internal consistency coefficient is .93, and over a two-week period temporal stability is .98, in a sample of 25 individuals with Asperger's disorder. Inter-rater reliability is reported as a 90% agreement between 10 pairs of raters. All raters in the sample were parents or relatives, rather than teachers, which could skew the results.

Item development was based on case histories, DSM-IV criteria and items from the autism screening instrument for educational planning by Krug, Arick and Almond (1993; as cited in Campbell, 2005). The final items were based on those that discriminated between the 'normal' sample and the Asperger's disorder sample, and then between those with Asperger's disorder and high-functioning autism disorder (Nellis, nd).

The KADI is reported to differentiate between Asperger's disorder, high-functioning autism and typical peers. Nellis (nd) cautions about the use of the KADI as a tool for differential diagnosis because the two clinical samples (Asperger's Disorder and high-functioning autism) were combined in some analyses. Nellis recommends the need for further research comparing the KADI with other rating scales and interviews.

2.11.4 Licensing

The KADI is published by the Western Psychological Services and is currently only available through them. No specific details are provided about user level.

2.12 Autism Spectrum Screening Questionnaire

The Autism Spectrum Screening Questionnaire (ASSQ) is a 27-item scale that was developed as a screening tool to identify children who may require a further more comprehensive assessment to diagnose Asperger's Disorder or high-functioning autism (Ehlers et al, 1999).

2.12.1 Administration issues

Twenty-seven behavioural descriptions are provided and rated on a three-point scale (not present, somewhat present and definitely present). The scores are then summed to yield a total score. There are both teacher and parent versions, which have different cut-off scores (22 and 19) to determine whether further referral is warranted. Domains that are assessed include: social interaction, communication problems, restricted and repetitive behaviours, motor clumsiness and associated behaviours. The ASSQ takes approximately 10 minutes to administer and score.

2.12.2 Statistics

Reports for sensitivity are variable for correctly identifying Asperger's disorder for both the parent and teacher forms. The ASSQ shows good specificity in correctly identifying non-Asperger's disorder (Ehlers et al, 1999). A very recent study on the validation of the ASSQ as a population screen for seven to nine year olds from the Bergen Child Study, combining parent and teacher formats and provided sensitivity of 0.91 and specificity of 0.86 (Posserud, Lundervold, Gillberg, 2009).

Internal consistency data for the ASSQ is not reported by the authors (Campbell, 2005). Interpreter reliability for parent-teacher agreement (n=20) is .77. For a larger sample (n=105), the parent-teacher agreement was .66. Temporal stability over a two-week period is .94 (n=65) and .96 (n=86) for teachers and parents, respectively (Ehlers et al, 1999).

Item selection for the ASSQ was based on the author's clinical experience, as well as relevant literature on the autism spectrum.(Ehler et al, 1999).

Content validity is claimed on the basis of clinical experience and literature review for item selection. The ASSQ appropriately discriminates between individuals with ASDs, attention deficit hyperactivity disorder, and learning difficulties for both the parent and teacher reports. Diagnoses for each group were established by a clinical case conference (Campbell, 2005).

2.12.3 Licensing

The ASSQ is available from the original publication in the journal, *Autism and Developmental Disorders*.(Ehler et al, 1999) No specific details about user levels are provided.

Table 1: Summary of the instruments reviewed

Tool	Type	Disorder	Age range	User level	Training required	Administration time	Diagnosis or screen	Sens	Specif	I.C.	I.R	T.S	Validity
ADI-R	Interview	Autism	2–Adult	Specialist	Yes	90 minutes	Diagnosis	.86–1.0	.75–.96	.69–.95		91%	Demonstrated
ADOS	Obsev	Autism	18 month +	Specialist	Yes	30–45 minutes	AIDS diagnosis	.90	.80	.47–.94	.65–.82	.59–.82	Demonstrated
CARS	Rating	Autism	2–?	Specialist	No	10–15 minutes	Screening	1.0		.94	.71	.81	Demonstrated
GAR-2	Rating and interview	Autism	3–22 years	Specialist	No	15 minutes	Screening			.94		.84	More research required
SCQ	Rating	Autism	4+	Specialist	No	15 minutes	Screening	>.75	>.60	.81–.93			Demonstrated
SRS	Rating	Autism	4–18 years	Open	No	15–20 minutes	Screening	.85	.75	.93–.97	.75–.91	.77–.85	Demonstrated
3di	Interview	Autism	Children	Trained	Yes	90 minutes	Diagnosis	1	.97		.9	.9	Demonstrated
DISCO	Interview	Autism	3 years +	Unknown	Yes	120–180 minutes	Diagnosis				>.75		More research required
GADS	Interview	Asp D	3–22 years	Specialist	No	5–10 minutes	AIDS diagnosis	.75–.96	.60–.80	.81–.93			More research required
ASDS	Rating	Asp D	5–18 years	Psych	No	10–15 minutes	AIDS diagnosis			.83	.93		Demonstrated
KADI	Rating	Asp D	6–21 years	Unknown	No	15–20 minutes	Screening	.78	.94	.93	90%	.98	Demonstrated
ASSQ	Rating	Asp D	6–17 years	Unknown	No	10 minutes	Screening	.62–.91	>.9		.77	.94	Demonstrated

Type – type of instrument; Age – age range applicable; User level – qualification required to administer and score instrument; Administration time – length of time required to administer; Diagnosis or screen – whether the instrument is intended to diagnose or screen for ASD; Specif – specificity; Sens – sensitivity; I.C – internal consistency; I.R – inter-rater reliability; T.S – temporal stability.

ADI-R Autism Diagnostic Interview-Revised; ADOS – Autism diagnostic Observation Schedule; CARS – Childhood Autism Rating Scale; GARS – Gilliam Autism Rating Scale – Second Edition; GADS – Gilliam Asperger’s Disorder Scale (2003 update); ASDS – Asperger’s Syndrome Diagnostic Scale; 3di – Developmental, Dimensional and Diagnostic Interview (3di); ASSQ – Autism Spectrum Screening Questionnaire; DISCO – Diagnostic Interview for Social and Communicative Disorders; SCQ – Social Communication Questionnaire; SRQ – Social Responsiveness Questionnaire; KADI – Krug Asperger’s Disorder Index.

Spec – specialist – health, educational, medical; Psych – psychologist.

3. Potentially preferable combinations of instruments: criteria for preferences

The following section posits potentially useful combinations of instruments that could be used to improve the formality and consistency of screening and diagnosis for ASD across age groups, and for both autism and Asperger's syndrome in New Zealand. Potentially preferred instrument combinations have been identified on the basis that they:

- (wherever possible) include those instruments which have a larger research base, which have been subjected to independent academic review, and which perform well in terms of reliability, validity and practicality
- are facilitative of a stepwise approach to screening and diagnosis, that is:
 - *Stage 1*: screening assessments – using a rating scale or interview that can be utilised by a range of professionals to assess whether a referral should be made for a diagnostic assessment
 - *Stage 2*: diagnostic assessments – using both interview and observation methods by experienced practitioners
- are flexible as to use by the widest range of practitioner groups/disciplines; this helps avoid 'bottlenecks' in the diagnosis pathway would otherwise be caused by shortages in particular workforces, either locally or nationally
- allow both interview and observation-based assessment, based on international best practice guidelines for ASDs (California Department of Developmental Services, 2003; Filipek et al, 2000; Ozonoff et al, 2005).

Detailed cost analysis is not within the scope of this report, but brief comments on relative cost are made in the following sections.

4. Potentially preferable combinations: autism

4.1 Autism: screening

The SCQ and CARS may be preferred screening instruments for autism. They are both brief and easy to administer and the research into the reliability and validity is strong. Neither instrument requires training, and both instruments can be used by a wide range of health and educational professionals. CARS claims validity for children as young as two years of age, and SCQ for children aged four years or more. The SCQ may be slightly preferable on the basis that it is a newer tool and reflects more recent DSM classifications.

4.2 Autism: diagnosis

The ADI-R and the 3di may be preferred interview-based instruments for diagnosis of autism. Research indicates sound reliability and validity for both. The ADI-R may be slightly preferable because of its larger research base. Training is required for both instruments, but they can be used by a wide range of health and educational professionals. For both instruments, interviews are time-consuming for both assessor and family. However, the 3di can be abbreviated.

On the criterion that instrument-based assessment should include clinician observation data, the ADOS, as the only clinician observation-based instrument, is preferred (as long as used in combination with information from other sources – see section 2.2.1). Available research findings suggest that it has sound reliability and validity. However, the ADOS can be difficult to administer and training of users is required.

The practice parameters for the screening and diagnosis of autism issued by the American Academy of Neurology and the Child Neurology Society (Filipek et al, 2000) and also Ozonoff et al (2005), recommend the CARS as an observational instrument for the diagnosis of autism, even though it was developed as a screening tool. This is controversial to some writers because items in the CARS were developed from outdated diagnostic criteria (DSM-III-R) and do not include some constructs considered important to the diagnosis of autism (see section 2.3).

Table 2 lists potentially preferable combinations of instruments. The combinations are presented in broad order of most to least preferred, based on the criteria defined earlier. In the final analysis these judgments are finely balanced; all combinations are defensible as long as their strengths and weaknesses are understood.

This review does not contain in-depth cost analysis, but partly because the relative merits of different instrument combinations are finely balanced, brief comment is provided in the table regarding the relative cost of the combinations.

Table 2: Potentially preferable instrument combinations for the screening and diagnosis of autism

Combination No.	Screening instrument	Diagnosis instruments		Notes
		Interview	Observation	
Combination 1	SCQ	ADI-R	ADOS	SCQ, ADI-R - strong research base. ADOS - strong research, clinician-observation based
Combination 2	CARS	ADI-R	ADOS	Same as Comb. 1, save use of CARS for screening – potentially slightly lower long term cost
Combination 3	SCQ	3di	ADOS	Same as Comb. 1, save use of 3di – potentially lower long term cost
Combination 4	CARS	3di	ADOS	Same as Comb. 3, save use of CARS for screening – potentially slightly lower long term cost
Combination 5	SCQ	ADI-R	CARS	Same as Comb. 2, save use of CARS for diagnosis (see qualifiers above)
Combination 6	SCQ	3di	CARS	Likely lowest cost option

5. Potentially preferable combinations: Asperger's disorder

5.1 Asperger's disorder: screening

The KADI appears to warrant consideration as a preferred instrument for screening for Asperger's disorder, although initial data only is available. It is brief and relatively easy to administer, and its publishers do not specify requirements as regards user qualifications. In common with all reviewed measures for screening for Asperger's disorder, further research is warranted.

5.2 Asperger's disorder: diagnosis

No instruments in this review are validated for making a definitive diagnosis of Asperger's disorder (see section 2).

6. Conclusion

This document has reviewed a subset of instruments available for the screening and diagnosis of ASDs, including Asperger's disorder, and set out some potentially preferable combinations of instruments.

Many of the instruments reviewed demonstrate acceptable evidence of reliability and validity. A good body of research surrounds several of the instruments, and internationally there are well developed professional and academic fora and peer review processes which serve to test instruments and challenge designers to improve instrument accuracy and utility.

It is recommended that:

- professional bodies whose members screen for, or diagnose, autism and Asperger's disorder should direct their membership to the review
- the Ministries of Health and Education should monitor developments in this field, and periodically update the review so that information available to practitioners about diagnostic instruments remains up to date.

References

- Bracken BA. 1987. Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment* 4: 313–326.
- California Department of Developmental Services. 2003. *Autistic Spectrum Disorders: Best practice guidelines for screening, diagnosis and assessment*. Author: California.
- Campbell JM. 2005. Diagnostic Assessment of Asperger's Disorder: A review of five third-party rating scales. *Journal of Autism and Developmental Disorders* 35: 25–35.
- Constantino J (nd). Retrieved 10 June 2009 from http://portal.wpspublish.com/portal/page?_pageid=53,70492&_dad=portal&_schema=PORTAL.
- Conway F (nd). [Review of the Test Social Responsiveness Scale.] In *The 17th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Ehlers S, Gillberg C, Wing L. 1999. A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders* 29: 129–141.
- England CT (nd). [Review of the Test Gilliam Asperger's Disorder Scale.] In *The 16th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH Jr, Dawson G, et al. 2000. Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology* 55: 468–479.
- Garro A (nd). [Review of the Test Gilliam Autism Rating Scale – 2nd edition.] In *The 17th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Gilliam J (nd). Retrieved 10 June 2009 from http://portal.wpspublish.com/portal/page?_pageid=53,70079&_dad=portal&_schema=PORTAL.
- Krug DA, Arick JR (nd). Retrieved 17 June 2009, from http://portal.wpspublish.com/portal/page?_pageid=53,69545&_dad=portal&_schema=PORTAL.
- Greenhalgh T. 1997. How to read a paper: Papers that report diagnostic or screening tests. *British Medical Journal* 315: 540–543.
- Leekam SR, Libby SJ, Wing L, Gould J, Taylor C. 2002. The Diagnostic Interview for Social and Communication Disorders: Algorithms for ICD-10 childhood autism and Wing and Gould autistic spectrum disorder. *Journal of Child Psychology and Psychiatry* 43: 327–42.
- Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, et al. 2000. Autism Diagnostic Observation Schedule – Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders* 30: 205–223.
- Lord C, Rutter M, 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: 659–85.

- Le Couteur A, Lord C, Rutter M. 2003. Autism Diagnostic Interview-Revised (ADI-R). Retrieved 10 June 2009 from Autism Genetic Resource Exchange, Los Angeles, CA. Website: <http://www.agre.org/program/aboutadi.cfm>.
- McGregor CM (nd). [Review of the Test Gilliam Asperger's Disorder Scale.] In *The 16th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Ministry of Health. 2008. *New Zealand Autism Spectrum Guideline*. Wellington: Ministry of Health.
- Naglieri J, Chambers. 2009. Psychometric issues and current scales for assessing autism spectrum disorders. In Goldstein S, Naglieri JA, Ozonoff S. *Assessment of Autism Spectrum Disorders*. New York: Guilford Press.
- National Guidelines Clearinghouse (nd). Retrieved 10 June 2009, from <http://www.guideline.gov>.
- Nebel-Scwalm MS, Matson JL. 2008. Differential Diagnosis. In JL Matson (ed) *Clinical Assessment and Intervention for Autism Spectrum Disorder* (pp.91–129). Burlington, MA: Elsevier.
- Nellis LM (nd). [Review of the Test Krug Asperger's Disorder Index.] In *The 16th Mental Measurements Yearbook*. Retrieved 17 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Ozonoff S, Goodlin-Jones B, Solomon M. 2005. Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology* 34: 523–540.
- Posserud MB, Lundervold AJ, Gillberg C. 2009. Validation of the autism spectrum screening questionnaire in a total population sample. *Journal of Autism and Developmental Disorders* 39: 126–134.
- Prizant BM (nd). [Review of the Test The Childhood Autism Rating Scale]. In *The 11th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Rellini E, Tortolani D, Trillo S, Carbone S, Montecchi F. 2004. Childhood autism rating scale (CARS) and autism behavior checklist (ABC) correspondence and conflicts with DSM-IV criteria in diagnosis of autism. *Journal of Autism and Developmental Disorders* 34: 703–708.
- Rutter M, Bailey A, Lord C (nd). Retrieved 10 June 2009, from http://portal.wpspublish.com/portal/page?_pageid=53,70432&_dad=portal&_schema=PORTAL.
- Schopler E, Reichler RJ, Rothen Renner B (nd). Retrieved 10 June 2009 from http://portal.wpspublish.com/portal/page?_pageid=53,69417&_dad=portal&_schema=PORTAL.
- Skuse D, Warrington R, Bishop D, Chowdhury U, Lau J, Mandy W, Place M. 2004. The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 43: 548–558.
- Statistics New Zealand. 2007. *National Population Projections: 2006 (base)–2061*. Retrieved 24 June 2009, from <http://www.stats.govt.nz/NR/rdonlyres/152C242E-E6F8-4A7A-87C78C0A3E48D790/0/nationalpopulationprojections2006basehotp.pdf>.

- Statistics New Zealand (nd). Population and Change and Structure. Retrieved 24 June 2009. <http://www.stats.govt.nz/analytical-reports/dem-trends-07/downloadable-excel-tables.htm>.
- Tadevosyan-Leyfer O, Dowd M, et al. 2003. A principal components analysis of the autism diagnostic interview-revised. *Journal of the American Academy of Child and Adolescent Psychiatry* 42: 864–872.
- Ward-Fairbank D (nd). [Review of the Test Gilliam Autism Rating Scale – 2nd edition]. In *The 17th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.
- Wing L, Leekam SR, Libby SJ, Gould J, Locombe M. 2002. The diagnostic interview for social and communication disorders: background, inter-rater reliability and clinical use. *Journal of Child Psychology and Psychiatry* 43: 307–25.
- Venn JJ (nd). [Review of the Test Social Responsiveness Scale]. In *The 17th Mental Measurements Yearbook*. Retrieved 10 June 2009 from EBSCOHost Mental Measurements Yearbook database.