Validating the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva): a diagnostic interview for autism spectrum disorders in adults

Kiri Clarke

D.Clin.Psy thesis (Volume 1), 2015

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Kiri Clarke

Date: 19th June 2015

Overview

Department of Health (2010) guidelines highlight the importance of diagnosis for adults with autism spectrum disorders (ASD) who have not previously had their condition recognised. Reliable, valid and user-friendly diagnostic tools must therefore be available.

Part 1: This section critically appraises and systematically reviews 12 studies examining the NICE (2012) recommended adult ASD diagnostic tools. It concludes that there is good evidence to support the use of the Ritvo Asperger Diagnostic Scale-Revised (RAADS-R) and Autism Diagnostic Observation Schedule (ADOS), with some support for the use of the Autism Diagnostic Interview-Revised (ADI-R). However the Adult Asperger Assessment (AAA), Asperger Syndrome Diagnostic Interview (ASDI) and Diagnostic Interview for Social and Communication Disorders (DISCO) appear to have insufficient evidence at present. Further research is indicated for all the instruments.

Part 2: This section presents a study of a new informant report diagnostic tool, the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva). The 3Di-sva interview was completed with an informant for 27 ASD and 27 non-clinical comparison participants. It demonstrated good psychometric properties, including good internal consistency and inter-rater reliability, and strong sensitivity and specificity. The 3Di-sva is a time and cost-efficient tool, which could be suitable for use as part of a multi-dimensional adult ASD assessment. The study was completed as part of a joint project with McKenner (2015), who examined the 3Di-sva when used in a clinical comparison population.

Part 3: This section is a critical appraisal which reflects upon areas relevant to both the literature review and empirical paper. The main focus of the discussion is upon on the idea of ideal versus achievable research and upon my learning process about research within NHS settings.

Table of Contents

List of Tables

Part 1: Literature Review

Part 2: Empirical Paper

List of Figures

Part 1: Literature Review

Part 2: Empirical Paper

Acknowledgements

I would like to thank everyone who helped me to complete this thesis. In particular I would like to thank Dr Will Mandy for his continuing encouragement, inspiration, guidance and excellent supervision throughout. I would also like to thank Michele McKenner, I'm not sure how I would have completed this without her being ever present to plan, troubleshoot, moan, and celebrate with. Thanks also go to Dr Jason Crabtree, for his much needed ongoing support of the project, and to Dr Andrew Greenhill who stepped in with enthusiasm to help us collect data. I am also grateful to Dr Andre Strydom and the clinicians at the ASD clinic for welcoming me into their team and assisting me with my recruitment. We are also indebted to Robert Cobb who obligingly completed our blind scoring exceptionally quickly. Finally I extend my thanks to my family and friends for being there when needed, and in particular to Alex, for putting up with the temporary loss of his wife to thesis work.

PART 1: LITERATURE REVIEW

A systemic literature review of the NICE recommended diagnostic tools for autism spectrum disorders in adults

Abstract

Aims. It is necessary to increase our understanding of the best methods of diagnosing autism spectrum disorders (ASD) in adults, in order to improve adult diagnostic services. This paper systematically identified and reviewed papers examining the psychometric properties of the NICE-recommended adult ASD diagnostic assessment tools – the Adult Asperger Assessment (AAA), Autism Diagnostic Interview-Revised (ADI-R), Autism Diagnostic Observation Schedule (ADOS), Asperger Syndrome Diagnostic Interview (ASDI), Diagnostic Interview for Social and Communication Disorders (DISCO) and the Ritvo Asperger Diagnostic Scale-Revised (RAADS-R).

Method. PsychInfo and Medline were searched to identify relevant papers. A total of 415 papers were found, of which 12 met criteria for inclusion. The quality of each paper was systematically evaluated using an adapted version of the BMJ Clinical Evidence (2014) tool. Measures of reliability and validity for each of the tools were examined and compared.

Results. The RAADS-R and ADOS were found have the best quality evidence available for their use, whilst the ASDI and AAA had the poorest quality evidence. Where reported, measures of reliability were generally good; however overall there was a lack of reporting of reliability information, particularly for the AAA and ADI-R. Overall sensitivity and specificity was high across the measures, with the exception of the specificity of the DISCO.

Conclusions. The review provides support for the use of the RAADS-R and the ADOS, and some support for the use of the ADI-R. It considered the evidence for the AAA, ASDI and DISCO insufficient at present to provide support for their current use in diagnosing adults. Further research is indicated for all the tools, particularly the informant report instruments.

Introduction

Overview

There has been a great deal of research on the diagnosis of Autism Spectrum Disorders (ASD) in children; however the diagnosis of ASD in adults has received less attention. Some individuals with ASD reach adulthood without receiving a diagnosis and it is therefore important that there are reliable, valid and user-friendly tools available for making a diagnosis in adults. Current NICE recommendations advise the use of at least one tool from a list of six when assessing adults for ASD (NICE, 2012). This review will examine the existing literature concerning the NICE recommended tools, in order to progress our understanding of the best methods of diagnosing ASD in adults.

Autism spectrum disorders

Autism spectrum disorders (ASD) are conceptualised as conditions in which two groups of symptoms are observed: social communication and interaction difficulties, and restricted and repetitive behaviours (DSM-5; American Psychiatric Association, 2013). Symptoms must begin in early childhood although they may not be recognised until later in life, and must cause functional impairment. ASD is a new diagnostic entity described in DSM-5 (APA, 2013) that subsumes previous separate categories of diagnosis such as autism, Aspergers, and atypical autism previously described in DSM-IV (APA, 2000). The prevalence of ASD among UK children has been found to be 157 per 10,000 children, with the ratio of known to unknown cases estimated at 3:2 respectively (Baron-Cohen et al., 2009). Therefore a high number of cases of childhood ASD go undetected, yet it has been shown that children with ASD do not grow out of the condition (Howlin & Moss, 2012). The estimated occurrence

of ASD within community adult populations in England is similar to that reported for children (Brugha et al., 2011), with the prevalence of undiagnosed adults with autism reportedly being around 1% (Brugha et al., 2009). Ritvo, Ritvo, Freeman & Mason-Brothers (1994) suggested that individuals with mild or late-appearing symptoms were less likely to come to clinical attention until adolescence or adulthood.

Following the Autism Act 2009, the government set out a Strategy for Adults with Autism in England (Department of Health, 2010) with the aim of improving the lives of adults with ASD. Several key actions were stated, including development of local autism teams, planning and commissioning of autism services, and improving access to diagnosis and post-diagnostic support. This document highlights the importance of diagnosis for adults who have previously not had their condition recognised, and points out that the capacity for diagnosis must be increased. Similarly, recent NICE guidelines on autism in adults (NICE, 2012) note that there is wide variation in rates of identification and diagnostic practice for adults with features of autism, which lead to delays in diagnosis and access to appropriate services. The guidelines highlight the importance of a clear and consistent care pathway to diagnosis and aftercare for adults with ASD.

Diagnostic tools for adults

It is therefore important that good quality psychometric tools to assess and diagnose ASD in adulthood are available. However, many of the instruments currently used for the assessment of ASD are focused on toddlers and children and there is a lack of well validated diagnostic instruments suitable for use with adults. NICE guidelines (2012) recommend that for complex diagnosis and assessment of adults without learning disabilities, the following formal assessment tools are

considered: the Adult Asperger Assessment (AAA; Baron-Cohen et al., 2005), the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1997), the Autism Diagnostic Observation Schedule – Generic (ADOS-G; Lord et al., 2000), the Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI; Gillberg et al., 2001), and the Ritvo Asperger Diagnostic Scale-Revised (RAADS-R; Ritvo et al., 2011). For adults with learning disabilities they suggest the ADOS-G and ADI-R. Additionally, the guidelines suggest that the ADOS-G, ADI-R or Diagnostic Interview for Social and Communication Disorders (DISCO; Wing et al., 2002) are used to organise and structure the process of complex assessment of ASD in adults.

Although some evidence is available suggesting that these tools are suitable for use with adults, the research in this area is limited. Thus far there has been little scrutiny of the existing research or comparison of the available evidence. It is important to increase our understanding of the best methods of diagnosing ASD in adulthood to enable the improvement of diagnostic services for adults with suspected ASD. We need to identify whether new tools need to be developed and where existing tools require updating and improving.

Aim of this review

The aims of this literature review are as follows:

1. To systematically identify and review all published research papers examining the psychometric properties of the NICE-recommended adult ASD diagnostic assessment tools.

2. To carry out a detailed review of the methodological quality of each study examining the NICE-recommended adult ASD diagnostic assessment tools based upon a formal and systemic evaluation using a critical appraisal tool.

The NICE (2012) guidelines are based upon a review of the psychometric properties of each tool, examined in a subsection of a review conducted in 2011 by the National Collaborating Centre for Mental Health (NCCMH; 2012). Some important and good quality research has been published more recently, such as a new paper examining a revised algorithm for Module 4 of the updated ADOS (Hus & Lord, 2014). This review will therefore update that that by NCCMH (2012) by searching for additional research published since this time. It will also include the first detailed review of study quality in the papers examining the recommended tools, which will be formally and systematically evaluated using a critical appraisal tool. This is designed to advance current understanding of the best methods of diagnosing ASD in adults.

As no papers reviewing the DISCO when used in an adult population were available when the NICE (2012) guidelines were established, the DISCO was recommended as a tool to 'organise and structure the diagnostic process' rather than specified as a diagnostic tool. The DISCO will be included here in order to find and review any research that has been published since the guidelines were issued.

Method

Inclusion and exclusion criteria

The inclusion criteria for the review were:

1. The study gives details on the psychometric properties of one of the six tools recommended for ASD diagnostic assessments as listed in the NICE guidelines (AAA, ADI-R, ADOS, ASDI, RAADS-R, or DISCO).
2. The study offers information on the instrument's ability to distinguish ASD from non-ASD.
3. The mean age of the participants included in the study is at least 18.

Exclusion criteria for the review were:

1. The aim of the study is purely to look at the factor structure of the instrument.

Search strategy

Studies were identified from searches up to September 2014 in two electronic databases: PsychInfo and Medline. The first concept for the search was 'Autism' and the following text words were used to identify autism and combined using OR:

i. autis*
ii. ASD
iii. Asperger*
iv. Pervasive developmental disorder
v. PDD

Within PsychInfo the subject heading "Pervasive developmental disorders", which included autism spectrum disorders was selected and combined with the text words using OR. Within Medline, due to the differences in subject heading options, the subject headings "child developmental disorders, pervasive/ or Asperger syndrome/ or autistic disorder" were combined with the text words using OR.

The second concept for the search was 'NICE recommended adult ASD diagnostic tools'. These were searched for using the names of each tool. The following text words were therefore used to identify the tools and combined using OR:

i. Adult Asperger assessment
ii. AAA
iii. Autism diagnostic interview*
iv. ADI*
v. Autism diagnostic observation schedule
vi. ADOS*
vii. Asperger syndrome diagnostic interview
viii. ASDI
ix. Ritvo autism Asperger diagnostic scale*
x. RAADS*
xi. Diagnostic interview for social communication disorders
xii. DISCO

Within PsychInfo, the above search for the diagnostic tool text words was limited to words within the following four areas: 1) title, 2) abstract, 3) key concepts, and 4) subject headings. Within Medline, due to differences in the options, the search for the diagnostic tool text words was limited to words within the following six areas: 1) title, 2) abstract, 3) keyword heading, 4) keyword heading word, 5) MESH subject heading, and 6) subject heading word. The text word search was limited to these areas in order to reduce the irrelevant papers to a manageable number to search through. Without these limits an excess of papers were found listing the diagnostic tools as measures used within the research but not examining the measure itself.

The text words for autism and the NICE recommended measures were combined using AND. The searches were limited to human, English language, peer reviewed journals, and adulthood. In order to identify any additional relevant papers, the reference lists of the final selected papers were examined and the citation tool on Google Scholar was checked.

Identification of relevant papers from search results

The search of PsychInfo yielded a total of 415 papers. The papers were all examined for relevance based on the title and abstract. Papers were initially disregarded on the basis that they had clearly not examined the psychometric properties of any of the included ASD measures. This left 45 papers, of which a further 13 were disregarded as they included only children as participants. Full manuscripts were obtained for the remaining 32 studies. Of these 12 were found to meet the inclusion criteria. No additional papers were found following the search on Medline, or from examining the reference lists or Google Scholar citations. This search revealed no other previous reviews of the tools to diagnose ASD in adults. See Figure 1 for flow diagram of search outcome.

It should be noted that two studies did not include a mean age of the participants and it cannot therefore be confirmed that the mean age was over 18 years. One study on the ADOS module 4 (Brugha et al., 2012) was certain to have a mean age over 18 as it included 618 adults over age 16, and participants were grouped by age with high numbers of participants included in older age groups including up to >75 years. The study on the ASDI (Gillberg et al., 2001) is less certain to have had a sample with a mean age over 18 as it included just 24 individuals aged 6-55 and no indication of the distribution of these ages is given. The

study is included as it is the only study examining the ASDI, however the results of this paper in must be interpreted with caution with regards to its use with adults.

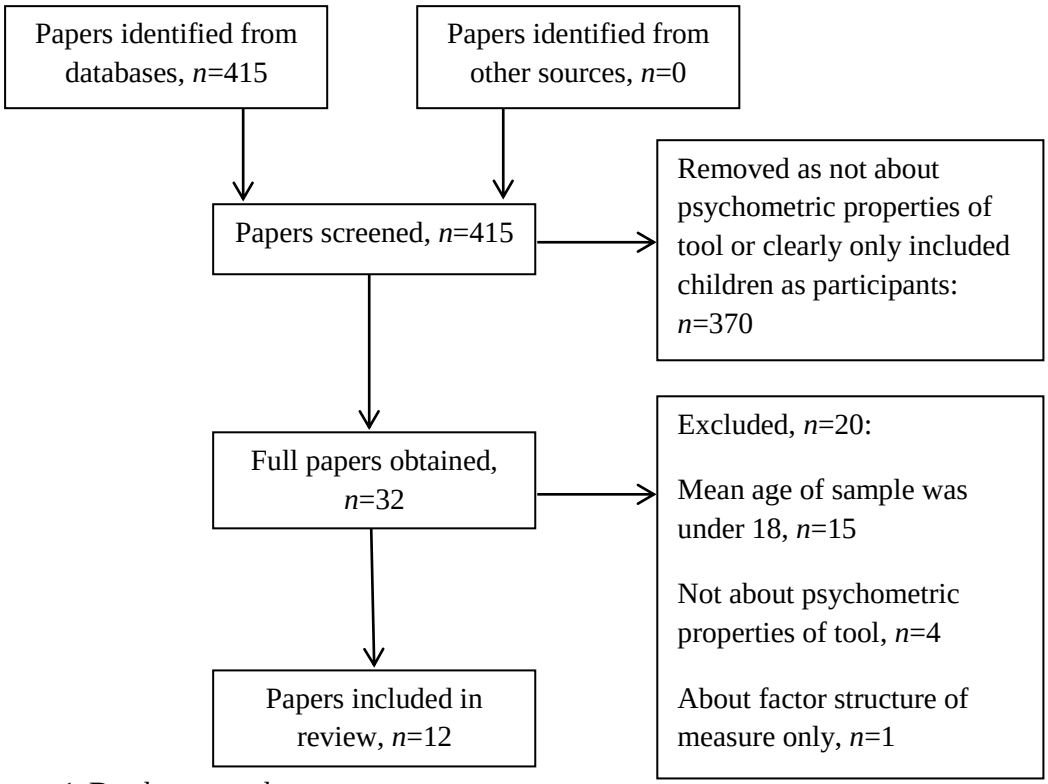

Figure 1. Database search outcome

Analysis of study quality

The quality of each study was systematically analysed using the BMJ Clinical Evidence (2014) tool for critically appraising diagnostic test studies. The tool was adapted slightly to optimise it for use with the ASD diagnostic studies under consideration in this review. See Appendix 1 for a copy of the adapted tool used, including notes about how it was administered for the purpose of this review. The changes made to the tool include:

1. The addition of a question on sample size. This was set as a sample of at least 20 ASD vs 20 non-ASD participants. This sample size was chosen in order to differentiate those studies using particularly small samples from the rest.

2. The addition of a question on whether cognitive ability was examined. This is because differing cognitive ability is a potential confound if not examined and controlled for.
3. Removal of a section of three questions focusing on applying the diagnostic test to a specific patient. The questions removed were: 'Is the patient similar to the people in the study in terms of clinical and demographic characteristics?', 'Is the diagnostic test available and if so does it reflect current practice?', and 'Will the results change the way the patient is managed?'. These questions are not relevant in this case as we are not looking to apply the instruments to a specific patient at this time.
4. The addition of a list of important psychometrics, to identify which of these were reported by the study. Studies were scored according to whether they provided detail on inter-rater reliability, test-retest reliability, internal consistency, convergent validity, and any correlation between score on the tool and participant characteristics such as age.

Each study was examined and scored a 1 for a 'yes' and 0 for a 'no' or 'unclear' answer. Studies were also awarded a score of 0.5 on certain items if they partially fulfilled the criteria. A table showing how each paper was scored is displayed in Appendix 2.

Results

The key characteristics of the 12 papers included are presented in Table 1. Of these studies four included the ADOS module 4, three the RAADS/RAADS-R, two the ADI/ADI-R, and there was one on each of the AAA, the ASDI, and the DISCO.

There was also one paper on both the ADI-R and the ADOS specifically for use with adults with learning disabilities.

The key features of each NICE recommended ASD diagnostic assessment tools are provided in Table 2.

Of the 12 papers retrieved, seven were the same as those included in the NCCMH (2012) review (Baron-Cohen et al., 2005, Brugha et al., 2012, Gillberg et al., 2001, Lord et al., 1997, Lord et al., 2000, Ritvo et al., 2008, Ritvo et al., 2011, and Sappok et al., 2013). Therefore the current literature search retrieved five additional relevant studies not included in those reviewed for the NICE guidelines (NCCMH, 2012).

The additional papers include one on the RAADS-R (Anderson et al., 2011), two on the ADOS module 4 (Bastiaansen et al., 2011, Hus & Lord, 2014), and one on the DISCO and ADI-R (Ngyren et al., 2009). Only one of these has been published since the NICE guidelines were published. On review of the search strategy for the NCCMH (2012) review it is apparent that Anderson et al. (2011) and Nygren et al. (2009) may not have been included as they examine Swedish translation versions of the instruments. It is unclear as to why the Bastiaansen et al. (2011) paper was not retrieved. It was decided that it was helpful to include the two Swedish translation studies as this broadens the evidence base regarding usefulness of the measures. Additionally, there are no papers published on the DISCO for use with adults other than this Swedish translation version.

Table 1

Study characteristics

Study	Tool	Participants	Study design & recruitment	Country	Diagnosis of ASD	Procedure
Anderson et al. (2011)	RAADS-R (Swedish version)	75 ASD aged 26-62 (M= 31; M:F = 36:35) 197 non-ASD aged 19-75 (M= 34; M:F = 80:116)	Case-control Recruited from multiple sites including clinics (ASD) and universities (comparison).	Sweden	Confirmed by either positive ADOS-G or DISCO	All participants completed RAADS-R. Subset of both groups completed AQ.
Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith (2005)	AAA	42 'patients' (age range not reported but M = 34.1; M:F = 36:6)	Cohort Patients attending one diagnostic clinic for adults with suspected ASD.	England	Based on clinician judgement of meeting DSM-IV criteria (including using information gathered from AAA)	All participants completed the AAA. AAA score was compared to diagnosis based on meeting DSM-IV criteria.
Bastiaansen et al. (2011)	ADOS module 4	38 ASD aged 18-66 (M= 31.82) 18 Schizophrenia aged 19-61 (M= 37) 16 Psychopathy aged 23-60 (M= 39) 21 non-clinical controls aged 21-53 (M= 34.23) Males only	Case-control Recruited from multiple sites including mental health organisations and clinics.	Netherlands	According to DSM-IV-TVR criteria based on review of developmental history, current daily functioning and observation.	All participants completed the ADOS module 4, completing all the standard activities and optional daily living items. IQ estimated for all except for nine participants using GIT2 (n= 80) or WAIS (n=4)

Brugha et al. (2012)	ADOS module 4	618 adults over age 16 (M:F = 344: 366). Mean age not stated.	Cohort Community sample of adults randomly selected from general population.	England	Confirmed using either DISCO and ADI-R, or clinical consensus based on vignettes containing record of ADOS, AQ-20, and clinical interview.	Participants completed AQ and ADOS module 4. Subset of sample completed DISCO and ADI-R (n=56).
Gillberg, Gillberg, Rastam & Wentz (2001)	ASDI	24 individuals aged 6-55 (M:F = 18:6). Mean age not stated. 17 neuropsychiatric disorder (10 AS, 3 atypical autism, 2 OCD, 2 ADHD, 1 MPD) 7 non-clinical controls.	Cohort Recruitment site information not stated.	Sweden	Based on consensus between two neuropsychologists, based Gillberg and Gillberg (1989) criteria for Asperger syndrome.	Informants who knew the participants well (including when they were children) completed the ASDI.
Hus & Lord (2014)	ADOS module 4 – new algorithm	177 autism aged 10-55 (M= 20.12) 170 other-ASD aged 9-54 (M= 21.14) 90 non-ASD (84% clinical, 16% non-clinical) aged 13-62 (M= 25.17) Participant sex not stated.	Case control Recruited from multiple sites – research studies and clinic referrals.	USA	Varied according to site but based on best estimate clinical diagnoses based on DSM-IV-TVR criteria.	Varied according to site but all participants completed ADOS module 4. IQ estimated for 91% of participants using various validated IQ measures, most commonly WASI or DAS.

Lord et al. (1997)	ADI/ ADI-R	60 autism aged 12-40 (M= 21.4) 36 non-autism comparison (including clinical cases) aged 7-38 (M= 17.5) M:F distribution not reported. Adult subset of sample reported here.	Cohort Recruited from multiple sites, mainly based at universities.	England, USA, France	Clinical judgement of principal investigator/ senior research associates using Rutter (1978) criteria.	ADI/ADI-R interview conducted with informant for each participant. Non-verbal IQ estimated for majority of participants using variety of validated measures.
Lord et al. (2000)	ADOS module 4	16 Autism (M age= 18.65; M:F= 14:2) 14 PDDNOS (M age= 21.59; M:F= 11:3) 15 Non-ASD comparison (clinical and non-clinical) (M age= 19.11; M:F= 12:3) Overall age range: 10-40	Case control (all matched on verbal IQ) Recruited from referrals to a developmental disorders clinic (ASD) and other clinics/groups (comparison)	USA	Consensus clinical diagnosis based on clinical impression from history, physical examination, and ADI-R score.	All participants completed ADOS-G module 4. Verbal and non-verbal IQ estimated for all participants using variety of validated measures.
Nygren et al. (2009)	DISCO and ADI-R (Swedish version)	27 ‘patients’ aged 15.2-39.7 (M= 24.7; M:F= 14:13). Only adult subset of larger sample reported here.	Cohort Patients with suspected ASD referred to diagnostic clinic.	Sweden	Assigned at case conference. Diagnoses based on DSM-IV criteria, except Asperger diagnosis (based on Gillberg & Gillberg (1989) criteria).	DISCO conducted with informant (close relative) for each participant. ADI-R also completed with subset of sample (n=21).

Ritvo et al. (2011)	RAADS-R	66 Autism age 18+ (M= 30.81; M:F= 52:14) 135 Asperger age 18+ (M= 32.01; M:F= 93:42) 276 no previous DSM-IV-TR diagnosis age 18+ (M=41.52; M:F= 114:162) 302 DSM- IV-TR axis I diagnosis other than ASD age 18+ (M=42.04; M:F= 134:168)	Case-control Recruited from multiple research sites with ongoing clinical and research programmes focussing on autism.	USA, Canada, England, Australia	Diagnoses based on DSM-IV-TR criteria, plus meet criteria on ADOS module 4 or ADI and ADOS.	Varied according to site but included diagnostic and IQ screening, all participants completed RAADS-R in presence of clinician.
Ritvo et al. (2008)	RAADS	37 ASD aged 18-65 (M=35; M:F= 22:15) 57 comparison (inc DSM-IV-TR axis 1 diagnosis and no diagnosis) aged 18-65 (M= 41; M:F= 25:32)	Case-control Recruited from multiple sites including clinics and universities.	USA	Diagnosis by two clinicians using DSM-IV-TR criteria.	All participants completed RAADS-R in presence of clinician.
Sappok et al. (2013)	ADOS / ADI-R	55 ASD aged 18+ (M=36; M:F= 42:13) 24 non-ASD aged 18+ (M=35; M:F= 17:7)	Cohort Referrals to a hospital specialising in learning disabilities.	Germany	Made by team of professionals in case conference, based on ICD-10 criteria.	Where possible participants completed the ADOS (module dependent on participant) and ADI-R. ADOS – 68% feasible to test. ADI-R – 37% feasible to test

Notes: M = mean; M:F = number of male:female participants; AQ = Autism Quotient (Baron-Cohen et al., 2001); AS = Asperger Syndrome; MPD = multiple personality disorder

Table 2

Measure characteristics

Measure	Main charactaristics	How administered	Time to complete	Designed for/as	Availability and training
AAA	4 sections each describing a group of symptoms (social interaction, restricted and repetitive behaviour, verbal/non-verbal communication and imagination) and a final section of 5 prerequisites Uses data from AQ and EQ screening questionnaires	Electronic, data-based, computer scored, clinician administered Completed by person with suspected ASD, and informant (for developmental history)	Takes around 3 hours (including feedback on diagnosis)	Specifically for adults (of normal intelligence) To be used as a complete diagnostic system	Freely available Unclear if training needed
ADI-R	93 items arranged in three functional domains: language/communication, reciprocal social interactions, restricted repetitive and stereotyped behaviours and interests.	Standardized clinician-based interviews Completed with informant who knows person and knew them as a child	90 – 150 minutes including scoring	For children and adults. Most studies examine properties of measure using child samples. Revised version of ADI	Available to buy Requires extensive training and practice in the administration and scoring.
ADOS (module 4)	Semi-structured assessment of social imagination, communication, play, and imaginative use of materials Consists of 10-15 activities with 31 accompanying ratings.	Clinician administered observation tool Completed with person with suspected ASD	Around 45 minutes	For adults with fluent use of language A clinical adjunct diagnostic tool (no developmental history acquired or information about functioning in other contexts)	Available to buy Requires extensive training and practice in the administration, observation and scoring.

ASDI	20 items covering 6 different areas (reciprocal social interaction, narrow interest patterns, routines rituals and interests, speech and language peculiarities, non-verbal communication, motor clumsiness	Clinician administered interview Completed by informant who knows person and knew them as a child	Around 10 minutes	As an adjunct diagnostic tool for Asperger's syndrome and high functioning autism in adults in accordance with Gillberg & Gillberg (1989) criteria	Freely available No training required
DISCO	362 items looking at developmental history and current behaviour. Examines skills, deficits and untypical behaviour.	Semi-structured investigator-based interview Completed with informant who knows person and knew them as a child	Not stated but interview is long (contains the most items of all measures examined)	To understand person's pattern of behaviour and needs as well as diagnosis To cover the broad autism spectrum at all ages (although most studies examine properties of measure using child samples).	Available to buy Requires extensive training and practice in the administration and scoring.
RAADS-R	80-item self-rating scale, four subscales: social interaction, language, circumscribed interests, sensory motor symptoms 64 symptom based questions, 16 non-symptom based responses	Self-report but administered by a clinician in a clinical setting Completed by person with suspected ASD	Around 30 minutes	Specifically for adults (of normal intelligence) A clinical adjunct diagnostic tool Revised version of RAADS	Freely available Unclear if training needed

Notes: AQ = Autism Quotient (Baron-Cohen et al., 2001); EQ = Empathy Quotient (Baron-Cohen & Wheelright, 2004).

Reliability

Reliability refers to the reproducibility of the instrument, in other words how consistent the measure is and how prone it is to measuring random error (Barker, Pistrang & Elliott, 2002). Internal consistency, test-retest reliability and inter-rater reliability for each paper reviewed is displayed in Table 3 and described below. Although there are no absolute criteria against which to judge reliability, Barker et al. (2002) suggest that a figure of 0.8 or above is considered good reliability and 0.5 or below is considered poor reliability. They also propose that 0.7 is acceptable and 0.6 is marginal reliability.

Internal consistency

Internal consistency is a measure of the inter-item reliability of a scale consisting of multiple similar items. It asks whether the different items do indeed relate to the same construct (Barker et al., 2002). Internal consistency was not reported in papers examining the AAA (Baron-Cohen et al., 2005), ASDI (Gillberg et al., 2001), DISCO (Nygren et al., 2009), or the ADI-R as used in non-LD populations (Lord et al., 1997, Nygren et al., 2009). One of the ADOS module 4 papers did not report internal consistency figures (Brugha et al., 2012) and another reported internal consistency figures but not for module 4 (i.e. the adult population) separately from the other groups (Lord et al., 2000) meaning the relevant figures could not be established.

For the RAADS-R, all subscales had acceptable or good internal consistency in the international study (Ritvo et al., 2011) and three out of the four subscales had acceptable or good internal consistency in the Swedish study (Anderson et al., 2011). The internal consistency of the language subscale in the Swedish translation of the

RAADS-R was poor (Anderson et al., 2011). Sappok et al., 2013 showed very good internal consistency for all ADOS modules when used in an LD population; however the internal consistency of the ADI-R when used in an LD population was low.

Both of the papers reporting on the ADOS module 4 (Bastiaansen et al., 2011, Hus & Lord, 2014) examined internal consistency of the social affect scale. This is a new domain suggested by Gotham, Risi, Pickles & Lord (2007) which reorganises and synthesises items from the two previous ADOS domains of social interaction and communication, in order to reflect the new DSM-5 criteria for ASD. Bastiaansen et al. (2011) also looked at the internal consistency of the pre-existing domains of social interaction and communication. Based on their work to revise and calibrate the module 4 algorithms, Hus and Lord (2014) also introduced another domain, the restricted and repetitive behaviour domain. In both studies the social affect scale showed good internal consistency. The restricted and repetitive behaviour subscale was acceptable.

Test-retest reliability

Test-retest reliability examines the consistency of a measure over time (Barker & Pistrang, 2002). Test-retest reliability was not reported in the majority of the papers. It was reported by the two RAADS-R papers (Anderson et al., 2011, Rivto et al., 2011). Anderson et al. (2011) examined test-retest reliability on a subset of 12 participants with a 3-6 month interval between completions of the measure. Ritvo et al. (2011) examined test-retest reliability on a subset of 30 participants, with a mean 1 year interval between completions of the measure. The figures show good test-retest reliability for the total scale in both cases. Only Anderson et al. (2011) reported figures for the subscales, which show acceptable or good test-retest

reliability for three of the four subscales, with the language subscale of this Swedish translation version being noticeably less reliable. The ASDI paper (Gillberg et al., 2001) also examined test-retest reliability, by re-interviewing 20 participants at 10-15 months after the first interview. They found good test-retest reliability.

Inter-rater reliability

Inter-rater reliability measures the degree of agreement between two different raters using the instrument. Reporting of inter-rater reliability was mixed; where applicable most papers made reference to inter-rater reliability, although many failed to report a kappa value, instead reporting only the percentage of agreement between raters, which is problematic as it does not correct for agreement due to chance. Inter-rater reliability was mentioned in three of the ADOS module 4 papers. Bastiaansen et al. (2011) found acceptable inter-rater reliability for overall classification on the ADOS module 4. Lord et al. (2000) found that raters were in agreement on overall classification on the ADOS module 4 in 84% of cases and Hus & Lord (2014) reported that agreement was consistently above 75% throughout the study. Gillberg et al. (2001) reported very good inter-rater reliability on the ASDI. Nygren et al. (2009) did not report on inter-rater reliability for overall classification on the DISCO, however they demonstrated good agreement on the majority of items on the measure. Lord et al. (1997; ADI-R) and Brugha et al. (2009; ADOS module 4) did not examine inter-rater reliability for the cases included in the study but reported raters must have reached an agreement of at least 90% during training. It was not applicable to measure inter-rater reliability on the RAADS-R as this is self-rating measure.

Table 3
Reliability of each instrument

Paper	Internal consistency (Cronbach's alpha, α)	Test re-test reliability	Inter-rater reliability
AAA (Baron-Cohen et al, 2005)	Not reported	Not reported	Not reported
ADI-R (Lord et al, 1997)	Not reported	Not reported	90% item by item agreement criteria before study commenced (no kappa reported)
ADI-R (Nygren et al, 2009)	Not reported	Not reported	Not reported
ADI-R (LD) (Sappok et al, 2013)	0.58	Not reported	Not reported
ADOS module 4 (Bastiaansen et al, 2011)	Social interaction, old algorithm: 0.84 Communication, old algorithm: 0.60 Social affect, revised algorithm: 0.87	Not reported	Overall classification: 89.2% agreement, kappa=0.73
ADOS module 4 (Brugha et al, 2012)	Not reported	Not reported	90% item by item agreement before study commenced (no kappa reported)
ADOS module 4 (Hus & Lord, 2014)	Social affect: 0.84 Restricted and repetitive behaviour: 0.61	Not reported	80% initial item by item agreement. Consistently exceeded 75% item by item agreement (no kappa reported)
ADOS module 4 (Lord et al, 2000)	Figures are mixed for all four modules (i.e. cannot separate adult sample)	Figures are mixed for all four modules (i.e. cannot separate adult sample)	Overall classification: 84% agreement (no kappa reported)

ADOS (LD) (Sappok et al, 2013)	Module 1: 0.81 Module 2: 0.89 Module 3: 0.93 Module 4: 0.92	Not reported	Not reported
ASDI (Gillberg et al, 2001)	Not reported	Reported as kappa= 0.92 (complete agreement on 97% of ratings)	Kappa = 0.91
DISCO (Nygren et al, 2009)	Not reported	Not reported	Reported only for individual items. Kappa's ranged from 0.35 to 0.91. Kappa's for >90% items were over 0.60
RAADS (Ritvo et al, 2008)	Language and communication: 0.65 Sensorimotor and stereotypies :0.73 Social relatedness: 0.86	Not reported	Not applicable
RAADS-R (Anderson et al, 2011)	*ASD/comparison* Total scale: 0.92/0.94 Circumscribed interests: 0.73/0.73 Language: 0.58/0.22 Sensory motor: 0.81/0.77 Social interaction: 0.87/0.89	 Total scale: r=0.80, p=0.002 Circumscribed interests: r=0.73, p=0.002 Language: r=0.43, p=0.161 Sensory motor: r=0.84, p=0.001 Social interaction: r=0.76, p=0.004	 Not applicable
RAADS-R (Ritvo et al, 2011)	Circumscribed interests: 0.90 Language: 0.79 Sensory motor: 0.91 Social relatedness: 0.92	Total scale: r=0.987 (p not reported)	Not applicable

N.B. Green denotes good reliability (r>0.8)
Orange denotes acceptable reliability (r=0.6 – 0.79)
Red denotes poor reliability (r<0.59)

Validity

Validity looks at whether an instrument measures what it is purported to measure. Criterion validity asks how well the instrument correlates with an indicator of the construct it is assessing (Barker et al., 2002). In this case it examines the degree of agreement between the measure and whether or not the person has ASD. Criterion validity for diagnostic measures is measured in terms of sensitivity and specificity. Sensitivity measures the accuracy of the instrument in picking out people who do have the condition (i.e. the number of true positives versus false negatives) and specificity measures how well the instrument avoids picking up people who do not have the condition (i.e. the number of true negatives versus false positives).

Criterion Validity: Sensitivity and specificity

The sensitivity and specificity with respect to ASD diagnosis for each measure according to each paper reviewed is shown in Figure 2. In Figure 3, an overall sensitivity and specificity figure is shown for each measure, based on samples combined across studies. Sensitivity and specificity figures were reported in nine of the twelve papers reviewed. For the three papers in which these figures were not reported (Baron-Cohen et al., 2005, AAA; Gillberg et al., 2001, ASDI; Nygren et al., 2009, DISCO and ADI-R), they were calculated using figures given within the papers regarding numbers of correctly and incorrectly identified cases.

Where figures for sensitivity and specificity were provided for both autism and ASD (Lord et al., 2000 and Sappok et al., 2013) the figures for ASD were used, based on the fact that under new DSM-5 criteria all autism related diagnoses are now classified as ASD. Where papers presented figures for existing algorithms versus revised algorithms (Bastiaansen et al., 2011; Hus & Lord, 2014; Sappok et al., 2013),

the figures for the revised algorithm were used based on the principle that these are the optimal algorithms for the measure.

It can be seen that overall sensitivity and specificity was fairly high across the measures. The RAADS/-R demonstrates the highest overall levels of sensitivity and specificity. The RAADS/-R and the ADOS module 4 demonstrate the smallest confidence intervals around their estimates of sensitivity and specificity, reflecting the large overall sample sizes of these studies. One measure that noticeably underperforms in terms of specificity is the DISCO (Nygren at al., 2009; specificity of 0.50). However it is worth noting that only 6 cases were used to calculate this figure meaning the validity of this figure may be questionable. When used in an LD population of adults, the ADOS also demonstrates lower specificity than many of the others (0.65; Sappok et al., 2013). Finally, the ADOS module 4 as reported by Bastiaansen et al. (2011) appears to show slightly lower sensitivity (0.71) and specificity (0.82) than many of the others. However these figures are still within an acceptable range and this study was one of the few to compare separate groups of adults with clinical diagnoses that may be confused for ASD, effectively making it a more stringent test of the measure than some of the others.

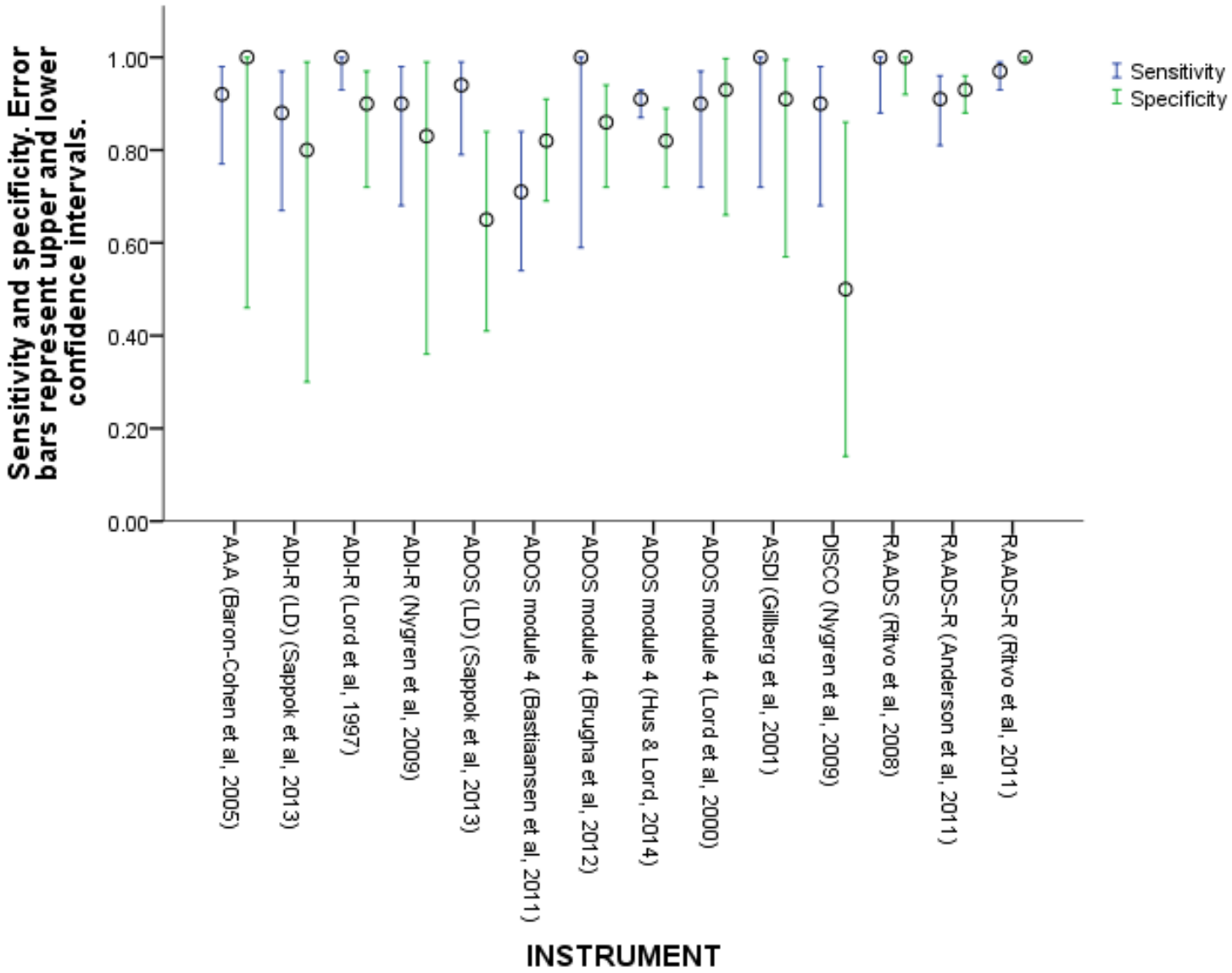

Figure 2. ASD diagnostic tool sensitivity and specificity by paper

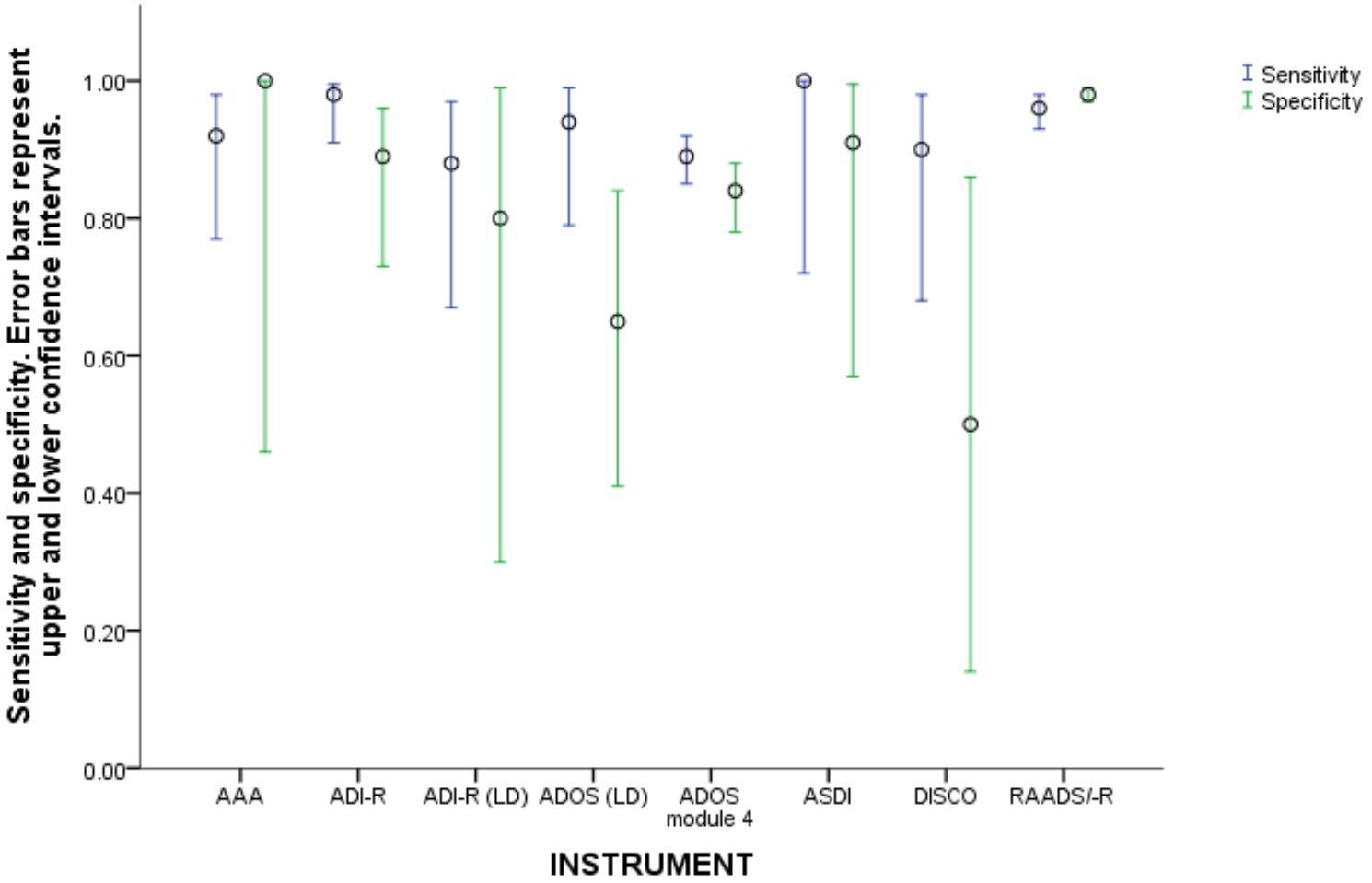

Figure 3. Combined sensitivity and specificity for each ASD diagnostic tool

Study quality and methodological considerations

The quality of each study was systematically analysed using the BMJ Clinical Evidence (2014) tool. Table 4 displays the studies in order of rank according to score using the critical appraisal tool. Table 5 displays an overall score and rank order for each tool, created by calculating the mean score for studies evaluating each tool.

The RAADS/-R and ADOS both score highly, with the RAADS/-R being rated as tool with the best quality evidence for its use. They are followed by the ADI-R and then the DISCO. The ASDI and AAA score lowest in terms of the quality of the evidence available for their use. It is interesting to note that the three lowest scoring measures have only one published paper each examining the tool when used with adults.

Table 4

Study ranking according to the modified BMJ Clinical Evidence score

Study	Score
RAADS-R (Ritvo et al, 2011)	20
ADOS module 4 (Bastiaansen et al, 2011)	19.5
ADOS (LD) (Sappok et al, 2013)	18.5
ADOS module 4 (Hus & Lord, 2014)	18
RAADS-R (Anderson et al, 2011)	18
ADI-R (LD) (Sappok et al, 2013)	17.5
ADOS module 4 (Lord et al, 2000)	17.5
RAADS (Ritvo et al, 2008)	16.5
DISCO/ADI-R (Nygren et al, 2009)	15.5
ADI-R (Lord et al, 1997)	14
ASDI (Gillberg et al, 2001)	13.5
ADOS module 4 (Brugha et al, 2012)	12.5
AAA (Baron-Cohen et al, 2005)	10

Table 5

Overall ranking according to modified BMJ Critical Evidence tool

Study	Score
RAADS/-R	18.17
ADOS[1]	17.1
ADI/-R[1]	15.67
DISCO	15.5
ASDI	13.5
AAA	10

Note. [1] Calculated including the papers examining instrument in both LD and non-LD populations.

The two studies which score lowest according to the critical appraisal are the AAA and the ASDI. Particular problems in the AAA paper (Baron-Cohen et al., 2005) include the lack of an appropriate sample size (only 5 participants did not have ASD according to DSM-IV criteria or 8 participants according to the AAA), no comparison group of participants with other psychopathology that might be mistaken for ASD, lack of blinding of the assessor, no measure of cognitive ability of the participant, and lack of reporting of psychometric properties of the measure. Reasons the ASDI paper (Gillberg et al., 2001) scored low include the limited sample size (13 participants with ASD and 11 participants without), lack of information about the study methodology such as participant details, and lack of reported estimate of diagnostic accuracy. The reason the DISCO (Nygren et al., 2009) scores lower than others include the lack of consideration of important psychometrics and a sample with only six non-ASD participants.

Although the ADOS scores highly on the combined analysis, it is worth noting that Brugha et al. (2012) scored noticeably lower than other studies of this

measure. Reasons for the low scoring of this paper include lack of gold reference standard, no inclusion of a population with psychopathology similar to autism (i.e. a clinical control group), no measure of cognitive ability, inclusion of ADOS module 4 record in vignettes used to make diagnosis and compare to ADOS module 4 results, and lack of some of the desired psychometrics. However, it should be considered that this study was unlike all the other validation studies considered, including the others on the ADOS module 4, in that it sought to validate it as a survey method for identifying cases of ASD among adults in the community. With such a vast sample (*n*=618) it may not have been feasible to measure cognitive ability, and in this case the inclusion of a clinical control population was not relevant. Although this study does not score highly on the BMJ Critical Appraisal tool, it is an important and unique validation of the ADOS module 4 that no other tools have received.

Discussion

Key findings

The main findings of the review are summarised below for each instrument. This review has considered the psychometric properties of each diagnostic assessment tool and the quality of each study reviewed. Although reliability, validity and study quality are the key concerns when comparing different instruments for diagnosing ASD in adults, other factors were also considered, such as how the instrument is administered and the availability and cost of the instrument. All of these aspects are considered when drawing conclusions about the overall value of each tool.

It should be noted that the widely accepted gold standard for an ASD assessment involves information collected from multiple perspectives, such as direct

observation, informant report and self-report (Baird et al., 2006). The only tool reviewed here which collects information from more than one perspective is the AAA, which requires the person themselves to complete two questionnaires, the AQ and the EQ, and then asks questions of an informant. The other measures reviewed collect information from one source and are intended to be used as adjunct tools, as part of a multi-dimensional assessment. Although it may be advantageous to be able to reach a diagnostic conclusion with just one instrument in terms of time, effort and cost of the assessment, due to the complexity of making a diagnosis of ASD, particularly in an adult, triangulation of information from different tools and methods should be encouraged to reach the most reliable conclusion.

Informant-based tools

Using informant-based tools means information can be provided about the person during early childhood, which is key to making a diagnosis according to DSM-IV/DSM-5 criteria, which state that symptoms must begin in early childhood, as well as ICD-10 criteria which state that symptoms must begin before three years of age (World Health Organisation, 1992). However, informant-based tools can be problematic for an adult population, where an informant may not be available or it may be difficult for the informant to recall developmental information. Four informant-based interviews were considered here, the AAA (which also includes self-report information from the patient), the ADI-R, the ASDI and the DISCO.

AAA

The one study evaluating the AAA scored particularly low on the critical appraisal tool. The level of methodological concern about the study, including a lack of an appropriate sample size, blinding and measure of cognitive ability, is

significant and would suggest that its results should be treated as preliminary. As such, the current state of research on the AAA does not appear justify its recommendation for use by NICE (2012). Further more rigorous research is needed before it can confidently be recommended. It should be noted that although it demonstrated good sensitivity and specificity, there were no reliability data (such as internal consistency, test-retest reliability, or inter-rater reliability) reported for the AAA, meaning we have no knowledge about how consistently this instrument performs. Asides from the low critical appraisal score, it is difficult to advocate the use of a tool in which the reliability is currently unknown. The AAA does have potential to be a useful tool if future research was able to demonstrate sound psychometric properties, as it is freely available and does not appear to require extensive training, making it easily accessible for clinicians. However it is a rather time-consuming tool, taking around 3 hours to complete, and was designed to be more stringent than DSM-IV criteria to avoid false positives.

ADI/-R

The ADI-R studies scored best out of the informant-report instruments, and in between the two best and two worst scoring instruments overall, according to the critical appraisal tool. There were some methodological concerns with the available research and therefore it would be important to conduct further, more rigorous investigation of the ADI-R for use with adults; however it seems fair to consider the existing research sufficiently sound in order to contemplate the tool further.

The ADI-R demonstrated good sensitivity and specificity. Reliability data were however somewhat limited for the ADI-R. There was no information available on its test-retest reliability, and data on inter-rater reliability were available from one

study but only as a figure that clinicians had to reach to become reliable rather than a check of the reliability on ratings completed for the purpose of the study. Internal consistency was only reported for the ADI-R when used with adults with learning disabilities, and this fell below acceptable levels (0.58; Sappok et al., 2013). The ADI-R also requires training and practice in its administration and takes over an hour and half to complete. ADI-R kits cost £273 each, with each additional interview booklet costing over £16, and the DVD training package retails at over £1000 (retrieved from www.pearsonclinical.co.uk). It is therefore a somewhat expensive and time-consuming tool to train in and use.

In terms of clinical use of the ADI-R, this review concludes there is evidence, albeit somewhat limited, that the ADI-R is suitable for use when an informant is available in order to gather information about the individual's developmental history. As it is not a stand-alone diagnostic tool, other instruments such as the ADOS module 4 and RAADS-R would be helpful clinical adjunct tools for making a diagnosis when the ADI-R is used.

The ADI-R is the one of the only tools validated for use with adults with learning disabilities, along with the ADOS. However it does still demonstrate some difficulties within this population, as it currently appears to have low internal consistency when used with adults with learning disabilities. This highlights the importance of using it as an adjunct tool alongside the ADOS when assessing adults with learning disabilities.

ASDI

Good test-retest reliability, inter-rater reliability, and sensitivity and specificity were demonstrated for the ASDI; but there was no report on the internal

consistency of the tool. However, like the AAA, the ASDI study scored particularly low according to the critical appraisal tool. The level of methodological concern about the study, including a limited sample size and lack of description of participants, is again significant and would suggest that any evidence for its value should be considered preliminary. Although recommended for use by NICE (2012), this review suggests that further more rigorous research of the ASDI is needed before it can confidently be recommended for use. Furthermore, the ASDI is developed to make diagnoses according to Gillberg & Gillberg (1989) criteria rather than DSM-IV/DSMI-5 or ICD-10 criteria, again limiting the usefulness of this tool for diagnostic services in the UK. The ASDI is however freely available and does not require training, meaning it is an easily accessible tool for clinicians to use.

DISCO

Like the ADI-R, the DISCO studies fell in between the two best and two worst scoring instruments overall according to the critical appraisal tool, but rated second best out of the four informant report tools. There were methodological concerns, namely a very small sample of participants without ASD (*n*=6), and a lack of important psychometric data included in the DISCO research. Although the tool scores sufficiently well to warrant considering it further, it is important that further, more rigorous investigation of the DISCO when used with adults is conducted.

The DISCO was the only measure considered in this review in which specificity was notably low at only 50 percent, meaning there is a high risk of over diagnosis. However, as highlighted earlier, this may be a very imprecise measure of the specificity due to methodological flaws in that not enough non-ASD participants were included in the study. Data on the inter-rater reliability of the DISCO were

mixed; no information is given for the overall scale, however most (>90%) individual items scored within an acceptable range. No data were available on internal consistency or test-retest reliability. It is acknowledged that NICE (2012) recommend the DISCO to 'organise and structure a diagnostic assessment' rather than as a diagnostic tool due to the lack of investigation of its diagnostic utility in an adult population. Although this review uncovered a study of the DISCO with adults, it confirms that DISCO cannot be recommended for diagnostic use at this time, on the basis that important reliability information is unknown and currently unacceptable levels of specificity have been demonstrated when used with adults.

Self-report tools

Self-report tools overcome the difficulties described above with informant based interviews as they require only the person in question to answer questions. However it should be considered that there can be problems with self-report, as it has been suggested that individuals with ASD have poor self-referential cognition (Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007) and limited insight (Bishop & Seltzer, 2012) which may make self-report more difficult. Indeed informant score yielded higher sensitivity and specificity than self-report on the Broad Autism Phenotype Questionnaire (BAPQ; Hurley, Losh, Parlier, Reznick & Piven, 2007) and it has been found that people with ASD underscore their symptoms when using the AQ (a self-report ASD screening tool; Bishop & Seltzer, 2012). The only fully self-report tool considered here is the RAADS-R.

RAADS/-R

The RAADS-R studies were the most methodologically sound according to the critical appraisal tool. It can therefore be considered to have good quality

evidence for its use. From the three papers on the RAADS-R there is clear evidence of good internal consistency and test-retest reliability, and good sensitivity and specificity. A measure of inter-rater reliability of this instrument is not applicable due to it being a self-rating measure. The RAADS-R is a freely available tool, which would require little training due to its self-report nature. It is also relatively quick (around 30 minutes) to administer. This makes the RAADS-R a good option for clinicians with limited resources.

This review suggests that the RAADS-R would be a suitable option for making a diagnosis in an adult for whom there was no informant available. As it is not a stand-alone instrument, it would be helpful in this case to combine the RAADS-R with a tool such as the ADOS. The RAADS-R could also be used alongside an informant-based interview such as the ADI-R, as a means of triangulating information from various sources.

There is also evidence that the RAADS-R can distinguish between those with ASD and those with other DSM-IV-TVR axis 1 diagnoses, as Ritvo et al. (2011) included a large clinical comparison group in their study. Although this area needs more investigation, this implies that RAADS-R may also be a useful tool for cases in which someone has or is suspected of having another clinical diagnosis.

Observational tools

Observational tools overcome the difficulties described above experienced by informant and self-report. However these tools involve one-off observations and therefore cannot provide information on everyday functioning and developmental history. The only observation tool considered here was the ADOS.

ADOS

The ADOS studies scored highly on the critical appraisal tool, suggesting that overall the studies are methodologically sound and the ADOS can be considered to have good quality evidence for its use. It showed acceptable to good internal consistency and inter-rater reliability, however no study has yet reported on the test-retest reliability of this tool when used with adults. Sensitivity and specificity were good, although specificity was slightly lower in the ADOS specifically when used with adults with learning disabilities (0.65; Sappok et al., 2013).

It should be noted that the ADOS is an expensive tool to use which requires a significant amount of training. The cost of the ADOS-2 kit is cited as £2078, with each additional scoring booklet costing just under £6, and the DVD training packages retailing at over £1000 each (retrieved from www.pearsonclinical.co.uk). The ADOS also requires ongoing reliability monitoring, making this a demanding tool in terms of time and effort required from clinicians. Despite its cost, current evidence suggests that the ADOS module 4 would be useful clinical adjunct tool for making a diagnosis in an adult, both when an informant is available and when one is not.

Although some other instruments included mixed clinical comparison groups, e.g. any axis 1 DSM-IV-TR diagnosis (Ritvo et al. 2008, 2011), the ADOS module 4 was the only instrument be investigated when used with comparison groups with other specified clinical diagnoses, namely groups with psychopathy and schizophrenia. The ADOS module 4 can be used to distinguish between males with ASD and psychopathy, using the revised algorithms suggested by Bastiaansen et al. (2011), making it a particularly useful clinical adjunct tool for adult males with psychopathy. To some degree it can also help distinguish between adult males with

ASD and schizophrenia, although this distinction is more difficult to make and is better based on individual ADOS module 4 items rather than overall score (Bastiaansen et al., 2011).

The ADOS is one of the only tools, along with the ADI-R, with evidence supporting its use in a population of adults with learning disabilities. Although it is noted above that the ADOS demonstrates lower specificity in learning disability populations, this is only when using ASD cut offs (as used in the current review for standardisation purposes). Higher specificity was found for the ADOS in this population when autism cut offs were used (specificity = 0.8; Sappok et al., 2013). The ADOS should therefore be the tool of choice, along with the ADI-R, when assessing a person with learning disabilities, as these are the only tools in which clear attempts have been made to validate them for use with adults with learning disabilities.

Suggestions for future research

There is clearly the need for further research into tools for the diagnosis of adults with ASD. Of the six tools recommended by NICE, half had had only one paper published in relation to their use with adults, and all of these had significant limitations as discussed above. The other tools had a least three papers each covering their use with adults, but all would still benefit from further investigation to fully understand their strengths and weaknesses with different populations of adults with suspected ASD.

It is notable that the evidence for the existing informant report tools is weak or mixed. This review has concluded that at present the AAA and ASDI do not currently have adequate quality research to recommend their use, and the DISCO

does not demonstrate satisfactory psychometric properties in order to recommend its use with adults. Although there was some evidence for use of the ADI-R this was mixed. There is clearly a need for further research into these tools, or for the development of new informant based tools. As the AAA, DISCO and ADI-R are particularly time consuming interviews, and the DISCO and ADI-R require extensive training, the development of a new informant based interview which was shorter and demanded less training would be ideal.

Based on the critical appraisal tool used for this literature review, which was highly relevant to the validation of tools to diagnose ASD, future research on new or existing tools would benefit from the inclusion of as many as possible of the suggestions below. It should be noted that there are many practical challenges to carrying out such research, for example it can be difficult to recruit large samples where research budgets are small and to collect psychometrics such as test-retest reliability where time is limited. It is therefore recognised that it may not be possible to carry out research which fulfils all the suggestions; however it would be best practice to adhere to as many as practically possible.

An ideal study would be one in which a gold reference standard were used for diagnosis of ASD, namely clinician consensus diagnosis made according to DSM-5 or ICD-10 criteria. The sample of participants would be well described (including age range, mean and standard deviation as well as gender ratio of participants), recruited from different locations or clinics, and formed of separate groups of at least 20 participants both with and without ASD (or ideally more). An identifiable group of participants with other clinical conditions that may have features that could be confused with ASD, such as symptoms of psychosis or social anxiety, would be important – as it is likely to be easier to distinguish between ASD and people with no

clinical symptoms compared to distinguishing between ASD and those with symptoms that might overlap with ASD. It is also helpful for participants to have completed a measure of cognitive ability in order to understand the cognitive profile of the group with which the measure is validated, as well as ensuring similar cognitive abilities between comparison groups. Assessors would be blind and the reference standard diagnosis would be made without the influence of the test under investigation (i.e. the index tool would not be used when making the diagnosis to which this tool was compared). The paper would provide a detailed description of the method and how data was analysed, including procedures for scoring and dealing with missing data. Ideally the paper would report sensitivity and specificity figures, giving the raw numbers from which this was derived, as well as other psychometric data looking at different types of reliability and validity.

Conclusions

This review examined the body of literature available on the psychometric properties of the NICE recommended tools for the diagnosis of ASD in adults. The review provides support for the use of the RAADS-R and the ADOS with adults, and some support for the use of the ADI-R. The best tools for use in different circumstances are discussed. The review considered the evidence for the AAA, ASDI and DISCO too weak at present to provide good evidence for their use in diagnosing adults. This does not mean however that they will not present themselves to be highly useful tools for use with adults in the future, but further research is necessary. Indeed for all the tools, further research is indicated, especially to validate the tools for use within specific populations of adults with other clinical diagnoses. The development of new informant report tools for adults may also be beneficial.

References

Andersen, L. M., Näswall, K., Manouilenko, I., Nylander, L., Edgar, J., Ritvo, R. A., Ritvo, E., & Bejerot, S. (2011). The Swedish version of the Ritvo autism and Asperger diagnostic scale: Revised (RAADS-R). A validation study of a rating scale for adults. *Journal of Autism and Developmental Disorders, 41*(12), 1635-1645.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Autism Act (2009). Retrieved from http://www.legislation.gov.uk/ukpga/2009/15.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet, 368*(9531), 210-215.

Barker, C., Pistrang, N., & Elliott, R. (2002). *Research Methods in Clinical Psychology*. Chichester: Wiley.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry, 194,* 500-509.

Baron-Cohen, S., & Wheelright, S. (2004). The empathy quotient (EQ). An investigation of adults with Asperger Syndrome or high functioning autism,

and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163-175.

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The adult Asperger assessment (AAA): a diagnostic method. *Journal of Autism and Developmental Disorders*, *35*(6), 807-819.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keysers, C., & de Bildt, A. (2011). Diagnosing Autism spectrum disorders in adults: the use of Autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders*, *41*(9), 1256-1266.

Bishop, S. L., & Seltzer, M. M. (2012). Self-reported autism symptoms in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2354-2363.

BMJ Clinical Evidence (2014). *Diagnostic test studies: assessment and critical appraisal.* Retrieved from: http://clinicalevidence.bmj.com.libproxy.ucl.ac.uk/x/set/static/ebm/toolbox/665061.html

Brugha, T. S., McManus, S., Bankart, J., Scott. F., Purdon, S., Smith, J., Bebbington, P., Jenkins, R., & Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry*, *68*(5), 458-466.

Brugha, T., McManus, S., Meltzer, H., Smith, J., Scott, F. J., Purdon, S., Harris, J., & Bankart, J. (2009). Autism spectrum disorders in adults living in households throughout England: Report from the adult psychiatric morbidity survey 2007.*Leeds: The NHS Information Centre for Health and Social Care*.

Brugha, T. S., McManus, S., Smith, J., Scott, F. J., Meltzer, H., Purdon, S., Berney, T., Tantam, D., Robinson, J., Radley, J., & Bankart, J. (2012). Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. *Psychological Medicine*, *42*(03), 647-656.

Department of Health. (2010). Fulfilling and rewarding lives: The strategy for adults with autism in England. UK: Author.

Gillberg, I. C., & Gillberg, C. (1989). Asperger syndrome—some epidemiological considerations: a research note. *Journal of Child Psychology and Psychiatry*, *30*(4), 631-638.

Gillberg, C., Gillberg, C., Råstam, M., & Wentz, E. (2001). The Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI): a preliminary study of a new structured clinical interview. *Autism*, *5*(1), 57-66.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613-627.

Howlin, P., & Moss, P. (2012). In review: Adults with autism spectrum disorders. *Canadian Journal of Psychiatry*, *57*(5), 275.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679-1690.

Hus, V., & Lord, C. (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, 1-17.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One, 2*(9), e883.

Lord, C., Pickles, A., McLennan, J., Rutter, M., Bregman, J., Folstein, S., Fombomme, E., Leboyer, M., & Minshew, N. (1997). Diagnosing autism: analyses of data from the Autism Diagnostic Interview. *Journal of Autism and Developmental Disorders, 27*(5), 501-517.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

National Institute for Health and Care Excellence (2012). *Autism: recognition, referral, diagnosis and management of adults on the autism spectrum. CG142*. London: Author.

National Collaborating Centre for Mental Health (2012). *Autism: the NICE guideline on recognition, referral, diagnosis and management of adults on the autism spectrum.* UK: The British Psychological Society and The Royal College of Psychiatrists.

Nygren, G., Hagberg, B., Billstedt, E., Skoglund, Å., Gillberg, C., & Johansson, M. (2009). The Swedish version of the diagnostic interview for social and communication disorders (DISCO-10). Psychometric properties. *Journal of Autism and Developmental Disorders, 39*(5), 730-741.

Ritvo, E. R., Ritvo, R., Freeman, B. J., & Mason-Brothers, A. (1994). Clinical characteristics of mild autism in adults. *Comprehensive Psychiatry*, *35*(2), 149-156.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., Yuwiler, A., Ritvo, M. J., & Weisbender, L. (2008). A scale to assist the diagnosis of autism and Asperger's disorder in adults (RAADS): A pilot study. *Journal of Autism and Developmental Disorders*,*38*(2), 213-223.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., Ritvo, M. J., Hufnagel, D. H., McMahon, W., Tonge, B., Mataix-Cols, D., Jassi, A., Attwood, T., & Eloff, J. (2011). The Ritvo Autism Asperger diagnostic scale-revised (RAADS-R): A scale to assist the diagnosis of autism spectrum disorder in adults: An international validation study. *Journal of Autism and Developmental Disorders*, *41*(8), 1076-1089.

Sappok, T., Diefenbacher, A., Budczies, J., Schade, C., Grubich, C., Bergmann, T., Bölte, S., & Dziobek, I. (2013). Diagnosing autism in a clinical sample of adults with intellectual disabilities: How useful are the ADOS and the ADI-R? *Research in Developmental Disabilities*, *34*(5), 1642-1655.

Wing, L., Leekam, S. R., Libby, S. J., Gould, J., & Larcombe, M. (2002). The diagnostic interview for social and communication disorders: Background, inter-rater reliability and clinical use. *Journal of Child Psychology and Psychiatry*, *43*(3), 307-325.

World Health Organization. (1992). *International Classification of Diseases: Diagnostic criteria for research* (10th edition). Geneva, Switzerland: Author.

PART 2: EMPIRICAL PAPER

Validating the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva): a diagnostic interview for autism spectrum disorders in adults

Abstract

Aims. There is a lack of validated diagnostic tools for adults with autism spectrum disorders (ASD). This study aims to evaluate the reliability and validity of a new, 71-question informant-report tool, the Dimensional, Developmental and Diagnostic Interview – short version for adults (3Di-sva). The 3Di-sva generates scores for subscales reflecting the DSM-5 ASD diagnostic criteria.

Methods. The 3Di-sva was administered to a parent (or an alternative informant) of 27 adults with ASD and 27 non-clinical comparison adults. A subset (ASD *n*=17, comparison *n*=24) of participants completed an estimate of IQ, and where possible interviews were audio-recorded and independently coded to evaluate inter-rater reliability (ASD *n*=10, comparison *n*=15). Participants with ASD also completed the Autism Diagnostic Observation Schedule (ADOS).

Results. The 3Di-sva demonstrated good reliability as measured by internal consistency and inter-rater reliability. Criterion validity was strong: ASD participants scored significantly higher than comparison participants on all subscales, and sensitivity (93%) and specificity (100%) were high. In the ASD group, there was however low correlation between 3Di-sva scores and ADOS scores. Construct validity was partially demonstrated: as expected there was strong correlation between scores on the two main scales of the 3Di-sva. However scores on the 3Di-sva subscales were not significantly associated with IQ score or gender, but were associated with age for the A-scale (Social Interaction and Communication) in the comparison group.

Conclusions. The 3Di-sva demonstrates good psychometric properties and is a time and cost-efficient tool that could be suitable for use a part of a multi-dimensional ASD assessment. Future research should examine the test re-rest reliability of the 3Di-sva and its reliability and validity when used with a clinical control population.

Introduction

Autism spectrum disorders

Autism spectrum disorders (ASD) are conceptualised by The Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM-5; American Psychiatric Association [APA], 2013) as conditions in which deficits are seen in two functional domains, often referred to as a dyad of impairments. The two domains are social communication and interaction and restricted and repetitive behaviours. For a diagnosis to be made, symptoms must begin in early childhood, although they may not be recognised until later in life, and must cause functional impairment. The concept of a dyad of impairments is a change from the previously accepted notion of a triad of impairments, in which symptoms were clustered into three domains: language and communication, reciprocal social interaction, and restricted, repetitive and stereotyped behaviours and interests (DSM-IV-TR; APA, 2000). The DSM-5 diagnostic entity of an autism spectrum disorder subsumes the DSM-IV-TR separate categories of diagnosis such as autism, Aspergers, and atypical autism.

The prevalence of ASD among UK children is around 1.5%, with the ratio of known to unknown cases estimated to be 3:2 respectively (Baron-Cohen et al., 2009). As many cases go undetected in childhood, and a recent review of outcome studies has shown that children with ASD do not grow out of the condition (Howlin & Moss, 2012), a number of individuals with ASD are likely to progress to adulthood without receiving a diagnosis. In particular, individuals with mild or late-appearing symptoms may be less likely to come to clinical attention until they reach adolescence or adulthood (Ritvo, Ritvo, Freeman & Mason-Brothers, 1994). The occurrence of ASD within community adult populations in England is similar to that

reported for children, and there is no significant decrease in prevalence of ASD across adult age groups (Brugha et al., 2011).

It is well known that there is a higher rate of ASD amongst males, with male to female ratios found to be between 3:1 and 4:1 (e.g. Baird et al., 2006, Chakrabarti & Fombonne, 2001). Within the typically developing general population, males have demonstrated significantly higher levels of ASD traits than females in both child (Constantino & Todd, 2003) and adult populations (Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001). Evidence also shows that females are less likely than males to receive a diagnosis of ASD at equivalent levels of autistic traits (Dworzynski, Ronald, Bolton & Happé, 2012). This may be because females have developed better compensation or adaptation strategies, or due to gender stereotypes in the diagnostic process (Dworzynski et al., 2012). Therefore, although within the general population males may demonstrate higher ASD traits than females, within a clinical population this is less likely, as females will often need to display more severe traits of ASD in order to receive a diagnosis.

According to a US-based community survey of children with ASD, 31% had an IQ in the intellectual disability range (IQ below 70) and 69% did not (IQ above 70) (Centres for Disease Control and Prevention, 2014), which suggests that majority of people with ASD do not have an intellectual disability. There has however been found to be a relationship between verbal IQ score and prevalence of ASD within the normal range of intelligence (i.e. over 70), with higher levels of ASD found in adults who score lower on a verbal IQ test (Brugha et al., 2009).

After the publication of the Autism Act 2009, the government set out the Strategy for Adults with Autism in England (Department of Health, 2010). Key

actions stated within the strategy included the development of local autism teams, planning and commissioning of autism services, and improving access to diagnosis and post-diagnostic support. The importance of diagnosis for adults who have previously not had their condition recognised is highlighted, and it is stated that capacity for adult diagnosis of ASD must be increased. Recent National Institute of Clinical Excellence guidelines on autism in adults (NICE, 2012) also note that there is wide variation in diagnostic practice for adults with features of autism, which lead to delays in diagnosis and access to appropriate services. The guidelines also highlight the importance of a clear and consistent care pathway to diagnosis for adults with ASD.

Diagnostic tools for adults

It is therefore important that good quality tools to assess and diagnose ASD in adulthood are available. However, many of the instruments currently used for the assessment of ASD are focused on toddlers and children and there is a lack of well validated diagnostic instruments suitable for use with adults. There are currently three types of standardised tools used for adults: direct observation, self-report, and informant report.

The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) is the only NICE (2012) recommended observational tool. It is suitable for use with both children and adults, with the ADOS module 4 being designed specifically for use with adolescents and adults with fluent speech (Lord et al., 2000). It is a well validated instrument with good psychometric properties, for example good inter-rater reliability (Bastiaansen et al., 2011) and sensitivity and specificity exceeding 80% for Module 4 (Hus & Lord, 2014). The ADOS therefore has value when assessing ASD

in adults, although there are some drawbacks in that extensive training is required to administer it and it is expensive to procure and use (Charman & Gotham, 2013).

The Ritvo Autism and Asperger Diagnostic Scale: Revised (RAADS-R; Ritvo et al., 2011) is the only NICE (2012) recommended self-report tool. It has demonstrated good psychometric properties, for example very high test-retest reliability, good internal consistency, and sensitivity and specificity of over 90%, (Ritvo et al., 2011; Anderson et al., 2011). This tool is therefore also valuable for assessing ASD in adults, however it should be noted that some individuals with ASD have poor self-referential cognition (Lombardo, Barnes, Wheelwright & Baron-Cohen, 2007) and limited insight (Bishop & Seltzer, 2012) which may make self-report difficult in such cases. Indeed it has been found that informant score yielded higher sensitivity and specificity than self-report on the Broad Autism Phenotype Questionnaire (BAPQ; Hurley, Losh, Parlier, Reznick & Piven, 2007) and that people with ASD underscore their symptoms when using a self-report screening tool, the Autism-Spectrum Quotient (AQ; Bishop & Seltzer, 2012).

Informant report instruments also provide valuable information for an adult ASD assessment. The NICE (2012) recommended informant report tools are the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter & Le Couteur, 1994), the Diagnostic Interview for Social and Communication Disorders (DISCO; Wing, Leekham, Libby, Gould, & Larcombe, 2002), the Adult Asperger Assessment (AAA; Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005), and the Asperger Syndrome Diagnostic Interview (ASDI; Gillberg, Gillberg, Rastam & Wentz, 2001).

The AAA and the ASDI were designed specifically for adults, with the AAA being a semi-structured and the ASDI being a more highly structured interview.

There is currently however only limited evidence available for both of these, as only one paper published by the developer of the instrument is available for each (Baron-Cohen et al., 2005, Gillberg et al.2001) and the sample sizes are notably limited in both cases. There currently exists no published reliability data, such as inter-rater reliability, internal consistency or test-retest reliability, for the AAA, which is also a time consuming instrument to administer, taking around 3 hours to complete (Baron-Cohen et al., 2005). Additionally, the AAA was designed to be more stringent than DSM-IV criteria to avoid false positives and the ASDI is developed to make diagnoses according to Gillberg & Gillberg (1989) criteria rather than DSM-IV-TR, DSM-5 or ICD-10 (World Health Organisation, 1992) criteria, therefore limiting their usefulness for diagnostic services in the UK.

The ADI-R and the DISCO are semi-structured interviews initially designed for use with the parents of children but also recommended as suitable for the assessment of adults (NICE, 2012). However, the only published paper examining the use of the DISCO with an adult population includes a sample of only six non-ASD participants and demonstrates poor specificity (Nygren et al., 2009), meaning that currently there is limited evidence for its value when used with adults. The DISCO is also a time consuming instrument to complete, taking between two and four hours (Charman & Gotham, 2013). The ADI-R presents as a more promising tool, for example it has been shown to demonstrate high sensitivity and specificity (Lord et al., 1997). However, there is still a lack of good quality data published regarding the reliability of the instrument when used with adults. Additionally, there are drawbacks to the ADI-R; it takes around two hours to complete (Constantino et al., 2003), requires extensive training in its administration, and is costly to procure and use (Charman & Gotham, 2013).

Therefore, although promising observational and self-report tools exist, the NICE (2012) recommended informant-report tools demonstrate several problems. There is a lack of sufficient psychometric data available on the instruments, several of the tools are costly to train in and use, and, with the exception of the ASDI, they take several hours to complete. An informant-report instrument is needed which demonstrates good psychometric properties, is easy to train in and non-expensive to use, which is also short enough to be a realistic and useful part of an assessment in health services which are often stretched in terms of time due to financial restraints.

Further development and research of informant-report tools is therefore worthwhile, in order to improve our diagnostic capacity for adults with ASD. There are several reasons why it is important that good quality informant-report instruments are available. Although it is recognised that ASD symptoms may not become fully manifest until adolescence or adulthood, the DSM-5 criteria for a diagnosis of ASD state that symptoms must have been present during early childhood, meaning it is very important to explore an individual's early symptoms with someone who knew them well as a child. Informant report can also be particularly helpful when considering alternative diagnoses. ASD is associated with extensive comorbidity (Mukaddes, Hergüner, & Tanidir, 2010), and it can present with similar features to other disorders (for example schizophrenia (Baastiansen et al., 2011), and some forms of anxiety (Zandt, Prior, & Kyrios, 2009)). Informant report regarding developmental history can assist a clinician to differentiate between such diagnoses and ASD. Also, as discussed earlier, individuals with ASD can experience difficulties with self-report instruments meaning that availability of third party information is likely to be a useful adjunct in any assessment of adult ASD.

The 3Di-sv

A new informant-report tool which could be useful for the diagnosis of ASD in adults is the Dimensional, Developmental and Diagnostic Interview (3Di; Skuse et al., 2004). The 3Di is a standardized parent interview designed to measure autistic features dimensionally, which can be administered to unselected clinical and general populations. The original version contained a diagnostic algorithm with 113 items, however a new shorter version was later developed (3Di-sv; Santosh et al., 2009) comprising a subset of just 53 items. The 3Di is well validated and reliable for use with child populations (Skuse et al., 2004, Santosh et al., 2009); however there is no evidence yet that it is suitable for use in an adult population.

The measure was however recently adapted into a specific adult version (Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version; 3Di-sva). The adult version was developed by analysing which questions in the 3Di-sv were best able to discriminate between those with and without ASD in older adolescent populations. Following this analysis some items were modified to make them more relevant to adults, and some new questions were added based upon knowledge of the ASD phenotype in adults and upon new DSM-5 criteria. The adult version includes 71 questions, which ask both about the individual's development as a child and their functioning in the present day. It is intended for use with adults with intellectual ability within the normal range, reflecting the fact that the original 3Di was designed for higher-functioning individuals.

Small-scale pilot research has suggested that the measure is able to discriminate effectively between adults with and without ASD, which indicates that the 3Di-sva is worth investigating further. The pilot research explored the validity of

3Di-sva subscales using a prototype of the instrument designed to measure the DSM-IV-TR construct of the 'triad of impairments'. As this has now been updated to a 'dyad of impairment' in the DSM-5, the 3Di-sva has since been revised to generate subscale scores according to the criteria for these two dimensions as opposed to the three originally developed.

It takes around one hour to train individuals how to score responses and 45 minutes to conduct the 3Di-sva. Scoring can be done easily using a computer algorithm. The interview is highly structured and consequently suitable for administration both in person and by telephone. The 3Di-sva is therefore potentially a time-efficient and practical informant-report instrument compared to the ADI-R, DISCO and AAA. Further benefits of the tool are that it is designed to make both dimensional and categorical assessments, giving an overall diagnosis as well as scores on subscales linked to the DSM-5 criteria, and it allows for the identification of specific areas of ability or impairment in both clinical and non-clinical populations. The 3Di-sva is therefore potentially a useful diagnostic tool for adult ASD which could form part of a standardised assessment.

Research aims

In order to evaluate the 3Di-sva for potential use within adult ASD diagnostic services, we need to investigate the reliability, criterion validity and construct validity of the instrument. To assess criterion validity, it is essential to investigate the 3Di-sva's ability to discriminate between ASD and non-ASD populations. This report focuses on its ability to discriminate ASD and a non-clinical comparison population, however the research was completed as part of a joint project in which

ability to discriminate ASD and a clinical comparison population was also examined and is reported elsewhere (McKenner, 2015).

To further investigate criterion validity it will be useful to explore the 3Di-sva's relationship to another adult ASD diagnostic tool, the ADOS module 4. To assess construct validity, correlation between the subscales of the 3Di-sva will be examined, as well as correlation between score and IQ (as there is usually a correlation between ASD traits and IQ) and score and gender in the comparison population (as traits of ASD are generally higher within males in the typically developing but not the clinical population). It would also be interesting to check for any correlation between age and 3Di-sva score; none would be expected as there is no known relationship between age and ASD traits.

The current study therefore addresses the following questions:

1) Does the 3Di-sva demonstrate good reliability, in terms of having high:
    a. inter-rater concordance, and
    b. internal consistency?
2) Does the 3Di-sva have criterion validity, as demonstrated by:
    a. effective discrimination between ASD and a non-clinical comparison population, and
    b. a significant correlation between 3Di-sva scores and scores on the ADOS module 4, in people diagnosed with ASD?
3) Does the 3Di-sva have construct validity, as demonstrated by:
    a. a significant correlation between scores on the social communication and interaction scale and the restricted and repetitive behaviours scale,
    b. a significant correlation between 3Di-sva score and estimated IQ,

c. significantly higher 3Di-sva scores in males in the comparison population, and

d. no significant correlation between 3Di-sva score and age?

Method

Design

This study used a cross sectional, between-subjects design to assess the psychometric properties of the 3Di-sva when used with people with and without ASD, using quantitative methods. The study was conducted as part of a joint research project with another UCL DClinPsy trainee, Michele McKenner (McKenner, 2015). See Appendix 3 for breakdown of individual contributions to project.

Participants

Inclusion and exclusion criteria

Participants in both the ASD and comparison groups were required to be aged 18 or over for inclusion. Exclusion criteria for both groups included 1) participant learning disability (as indicated by an estimated IQ under 70), as the interview is designed to assess people with an IQ within the normal range, and 2) no informant available to complete the 3Di-sva.

Additional inclusion criteria for the ASD group included, 1) meet threshold for a diagnosis of ASD on the ADOS module 4 (i.e. score at least 7 or above on the combined communication and social interaction scale), and 2) ASD diagnosis confirmed by clinician consensus opinion (based upon DSM-IV-TR criteria for autistic or Asperger's disorder or DSM-5 criteria for autism spectrum disorder). It was considered essential that all participants first met the threshold for ASD on the ADOS in order to avoid circularity, as although the 3Di-sva algorithm was not used when making diagnoses, information gathered from informants during the 3Di-sva

did contribute to the diagnostic decisions made by clinicians. Clinician consensus opinion was then used to control against the inclusion of false positive cases from the ADOS module 4.

Additional exclusion criteria for the comparison group were: 1) current mental health difficulties (a separate group with mental health difficulties were recruited and are reported on by McKenner (2015)), 2) any current or previous concerns around having an ASD, and 3) participant or informant unable to speak fluent English (due to lack of resources to provide an interpreter).

Sample

The overall sample collected as part of the joint work with McKenner (2015) included 74 participants aged 18 – 59. Three separate groups were recruited, 27 participants with ASD, 27 typically developing comparison participants, and 20 comparison participants with mental health difficulties. McKenner (2015) conducted analyses examining the use of the 3Di-sva when used with adults with mental health difficulties. The current analysis looks at the 3Di-sva when used with adults with ASD (ASD group) and non-clinical control adults (comparison group). The demographics of the sample are displayed in Table 1.

The number of participants recruited for the study was not based upon a power analysis as we were not concerned with the 3Di-sva's capacity to detect small, subtle between-group differences. The number of participants recruited was instead determined by the maximum number it was feasible to recruit within given time and financial restrictions. Post hoc power analyses show that sufficient power (0.80) was achieved to detect a large effect size in all of the t-test, Mann Whitney U, and correlational analyses.

Table 1

Participant demographics.

Group	*N*	% Male	Mean age (*SD*), min-max	Mean est. IQ (*SD*), mix-max
ASD	27	67	35.63 (13.32), 18 – 59	109.47 (16.89), 72 – 138
Comparison	27	56	29.54 (8.87), 18 – 52	115.7 (10.30), 89 – 133
Overall sample	54	61	32.64 (11.66), 18 – 59	113.12 (13.59), 72 – 138

Note. Age is unknown in one comparison group case. Mean estimated IQ is based on scores for 17 ASD group cases and 24 comparison group cases. IQ scores were based on estimates provided by the TOPF in all cases except two ASD group cases which completed the WASI and six ASD group cases who completed the WAIS-IV. Where scores were not obtained, IQ was assumed to be in the normal range.

The ASD group were recruited from two adult ASD diagnostic clinics in London. The standard diagnostic process within both clinics included the completion a clinical interview, ADOS module 4, and 3Di-sva interview (where an informant was available). Diagnosis was based upon the consensus decision of the team, consisting of clinical psychologists and consultant psychiatrists. Of the ASD group, 15 cases had completed their assessment within the past two years and consented for their anonymised data to be included in research. Each of these cases were contacted by the researchers and asked to complete the Test of Premorbid Functioning (TOPF; Wechsler, 2009) for the purposes of the current research. A further 12 cases were recruited at the time that they attended the clinic for their assessment. See Figure 1 for a flowchart of recruitment to the ASD group.

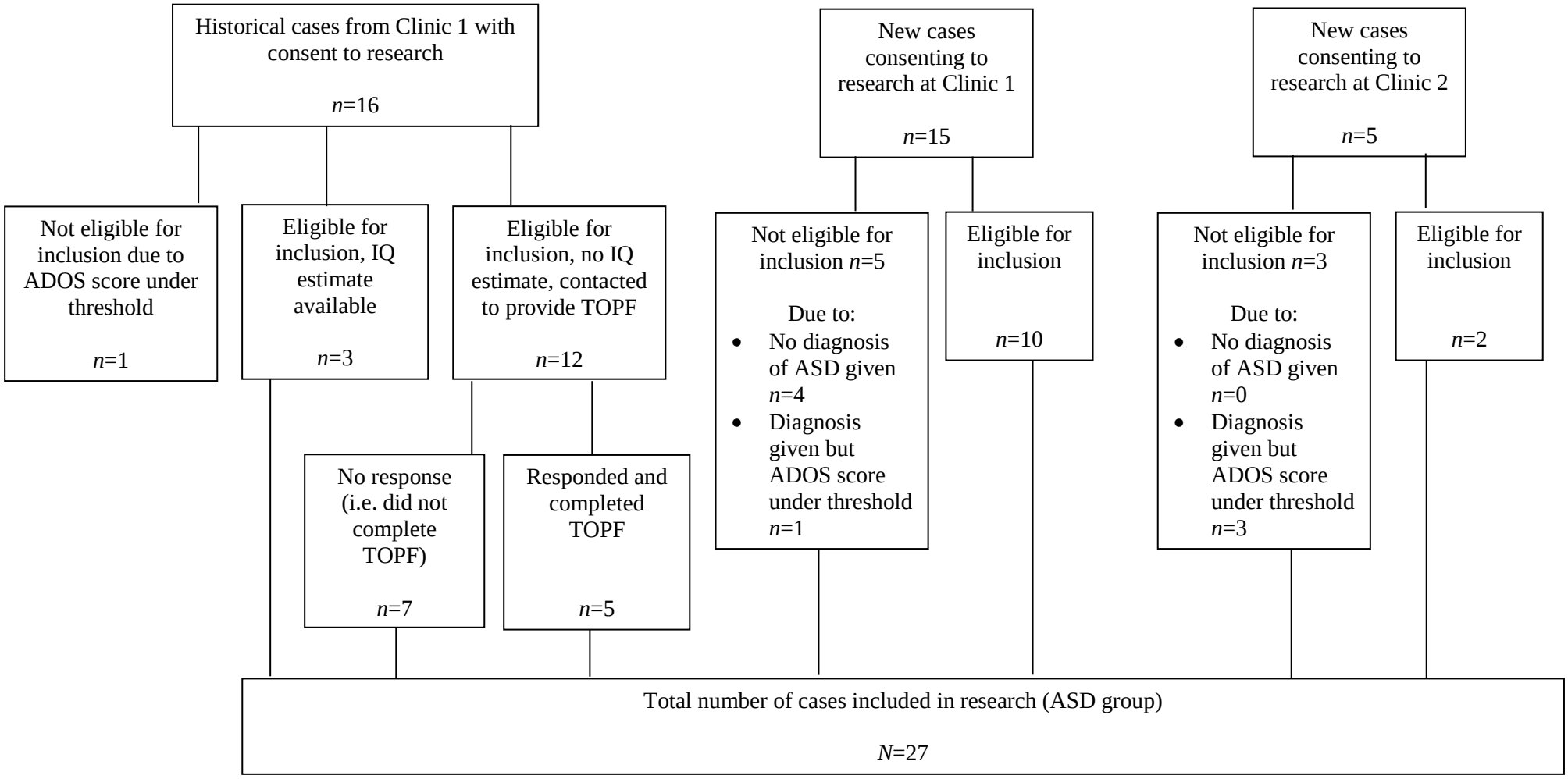

Figure 1. ASD group recruitment flowchart.

The comparison group were recruited using convenience sampling methods, from adverts placed around a university campus and sent to friends and colleagues of the researchers. Potential participants were asked to contact the researchers to express their interest in taking part.

Ethics

All aspects of the study were approved by the Bloomsbury NRES Committee and by relevant local Research & Development departments. All participants recruited provided informed consent before taking part in the research, and historical ASD group cases provided consent at the time of their assessment for their anonymised data to be included in research. All research data was collected and stored according to the Data Protection Act 1998. See Appendix 4 for letter giving ethical approval, Appendix 5 for information sheets, Appendix 6 for consent forms, and Appendix 7 for invitation letter sent to the historical ASD cases.

Measures

Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva)

The 3Di-sva is an informant report, semi-structured interview designed to assess and diagnose autism spectrum disorders in adults. It was adapted from the short form 3Di (Santosh et al., 2009) used in child and adolescent populations. It provides an assessment of the areas of autistic impairment highlighted by the DSM-5. The interview asks questions about both developmental history and current behaviour. It is carried out with an informant who knew the person both as a child and currently. In most cases this is a parent, but it could also be other family members or friends.

The 3Di-sva consists of 71 interview questions, 67 of which are included in the scoring algorithm. The remaining four questions measure developmental milestones. Questions included in the algorithm are arranged into two main scales, the 'A-scale' which reflects the DSM-5 Social Interaction and Communication dimension, and the 'B-scale' which reflects the DSM-5 Restricted, Repetitive Patterns of Behaviour, Activities or Interests dimension. The A-scale and B-scale are comprised of separate subscales reflecting the DSM-5 diagnostic criteria, forming a total of seven subscales. The arrangement of questions within the subscales is displayed in Figure 2.

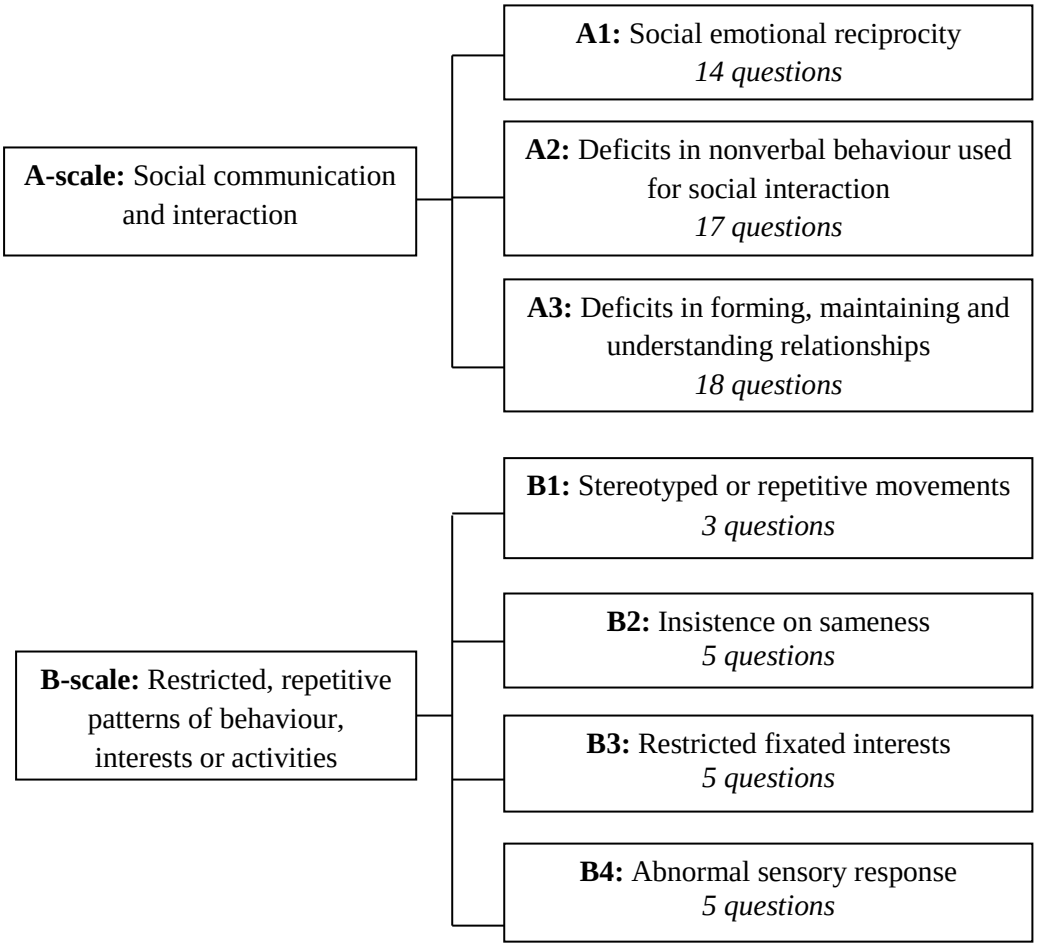

Figure 2. 3Di-sva scoring algorithm arrangement.

Questions are scored on either a three point (either 0= Often, 1= Sometimes, 2= Never, or vice versa) or four point Likert scale (0= No, 1= Yes, minimal, 2= Yes, persistent, 3= Yes, persistent with functional impairment). The only exception is the four questions about developmental milestones, in which two scoring options are available (0= within the normal range, 1= outside of the normal range). All questions receiving a score of 3 are recoded to 2 when calculating algorithm scores in order to ensure that all items within a scale carry equivalent weight. Scores for each of the seven subscales are generated by totalling the responses to each of the relevant questions and an overall score for the A-scale (Social Communication) and the B-scale (Restricted repetitive behaviour, interests or activities) are generated. As there are many more items making up the subscales of the A-scale than the B-scale, all scores are scaled in order to give each subscale an equal weighting.

Test of Premorbid Functioning – UK Version (TOPF; Wechsler, 2009)

The TOPF is a brief measure used to predict full scale IQ for individuals aged 16 to 89 years. The test involves reading out a list of up to 70 words that have atypical grapheme to phoneme translations. Individuals are asked to read the words out loud in order and stopped if they pronounce more than 5 words incorrectly in a row. It takes around five minutes to complete. Full scale IQ score is predicted based on number of words correctly pronounced, number of years of education and age. The TOPF has been shown to demonstrate good internal reliability (0.95), good test-retest reliability (0.89-0.95), and high correlation (0.81) with full scale IQ score as predicted by the Wechsler Adult Intelligence Scale- fourth edition (WAIS-IV; Wechsler, 2008). It has been validated for use in various populations including individuals with ASD.

Autism Diagnostic Observation Schedule (ADOS-G) Module 4 (Lord et al., 2000) and ADOS Diagnostic Observation Schedule 2 (ADOS-2) Module 4 (Lord et al., 2012)

The ADOS is a standardized semi-structured observational assessment consisting of tasks and questions which an examiner carries out with the individual with suspected ASD. It assesses communication, reciprocal social interaction, imagination/creativity, and stereotyped behaviours and restricted interests. Module 4 of the ADOS is used for adolescents and adults with fluent speech. It places a greater emphasis on conversation, as opposed to play as in the other ADOS modules, to gather information about social-communication. An ADOS module 4 usually takes between 45 minutes to one hour to complete. The ADOS-2 is the recently updated version of the original ADOS-G, however for Module 4 both versions are very similar and there are no differences in the scoring algorithm used. The ADOS-G was used in 11 cases and the ADOS-2 in 17 cases in the current study.

For Module 4 of both the ADOS-G and ADOS-2, various observed behaviours are coded using a three or four point Likert scale, which ranges from 0 indicating no abnormalities, to 2 or 3 indicating definite difference or abnormality. A scoring algorithm is given in which a Communication and a Social Interaction score are generated. These are then combined to give a total score, which must reach a cut-off score of 7 or above to indicate a diagnosis of ASD. Imagination/Creativity and Stereotyped Behaviours and Restricted Interests scores are also generated but not included in the diagnostic algorithm. The ADOS is often used as part of a gold-standard ASD assessment and has demonstrated good psychometric qualities (e.g. Lord et al., 2000, Bastiaansen et al., 2011, Hus & Lord, 2014).

Procedure

Participants were recruited between August 2014 and May 2015. The 3Di-sva interview was carried out with an informant for all participants in both groups. For the comparison group, the informant was the mother in the majority of cases (*n*=25), however a father (*n*=1) and older sister (*n*=1) also acted as an informant. For the newly recruited participants in the ASD group, the informants were mothers in all cases. Data is unavailable on who the informant was for the historical cases included in the ASD group, however clinician opinion was that the vast majority of historical interviews were also done with mothers. For the comparison group, interviews were carried out by the researchers over the telephone (*n*=24) or in person (*n*=4). For the newly recruited ASD group cases, the interviews were either carried out in person at the clinic (*n*=7) or over the telephone (n=5), by either the study researchers (*n*=10) or clinicians at the ASD clinics (*n*=2). All interviews for historical cases included in the ASD group were conducted by clinic clinicians. The method by which the information was collected for the historical cases is unknown, however clinician impression was that the majority included were conducted in person.

All researchers and clinicians conducting 3Di-svas included in the study had been trained in its use. Interviews were audio recorded where possible. Recordings were gathered for 10 ASD participants and 15 comparison participants. Audio recordings were listened to and scored by one psychology undergraduate trained in using the 3Di-sva who was blind to participant group.

Participants in both groups were asked to complete a TOPF as an estimate of IQ where possible (*n*=9 ASD group, *n*=24 comparison group). ASD group cases were not asked to complete a TOPF when a more comprehensive IQ test had been

completed as part of their clinic assessment; in these cases the IQ estimate generated from the alternative test was used. This applied to eight ASD group cases (*n*= 2 Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), *n*=6 WAIS-IV). All ASD group participants also completed an ADOS module 4 with a clinician at the ASD clinic.

For the comparison group, participants and the informants who completed the 3Di-sva interview were given a £10 voucher to thank them for their time. No payment was given to the ASD group as data was collected as part of the routine clinic assessment and it was deemed unethical to pay participants from this group when other individuals attending the clinic would not be paid if they were not suitable for inclusion in the research.

Analysis

Data analyses were conducted in SPSS, version 22. Missing data was dealt with by prorating subscale scores using the mean item score if less than 50% of the data was missing.

Preliminary analyses were conducted to assess normal distribution of the variables and differences between groups on IQ and age were analysed. Inter-rater reliability of the 3Di-sva interviews was assessed using intra-class correlation coefficients. Cronbach's alpha was used to test for internal consistency of the subscales. Independent samples t-tests or Mann Whitney U tests (depending on distribution of variables) were used to look for statistically significant differences between the two groups (ASD vs comparison) on scores for the 3Di-sva A-scale (Social communication and interaction**)** and B-scale (Restricted repetitive patterns of behaviour, interests or activities). Gender differences in comparison group for the

two 3Di-sva scale scores were also examined using independent sample t-tests or Mann Whitney U tests as appropriate. A Receiver Operating Characteristics (ROC) curve was generated to examine ability of the overall 3Di-sva score to discriminate between the ASD and comparison group. This was used to set optimal thresholds for indicating an ASD diagnosis which maximised sensitivity and specificity of the measure. Pearson or Spearman correlations (depending on distribution of data) were examined between the 3Di-sva scale scores and the ADOS module 4 scores. Pearson or Spearman correlations were also examined between 3Di-sva scale scores and age and estimated IQ.

Results

Preliminary analyses

Normal distribution

The distribution of each of the subscales, age, and estimated IQ were examined visually and using the Kolmogorov-Smirnoff test. Results of the Kolmogorov-Smirnoff test are displayed in Table 2. Distribution of the data for the subscales was found to differ significantly from normal in all cases for the combined group, and all cases apart from the overall A-scale for the comparison group. In the ASD group the subscale data was normally distributed, with the exception of scale B2. The distribution of age differed significantly from normal in the combined and ASD groups, but not in the comparison group. Estimated IQ was normally distributed in all cases. All analyses involving variables that were not normally distributed were carried out using non-parametric statistics.

Distribution of ADOS subscale data for the ASD group were also examined and found to differ significantly from normal in all cases (ADOS Communication:

D(27)=.18, *p*=.03; ADOS Social Interaction: *D*(27)=.29, *p*<.001; ADOS Imagination/Creativity: *D*(25)=.32, *p*<.001; ADOS Stereotyped Behaviours and Restricted Interests: *D*(25)=.32, *p*<.001; ADOS Combined Communication and Social Interaction: *D*(27)=.22, *p*=.002.). Analyses using ADOS data were therefore conducted using non-parametric statistics.

Table 2

Kolmogorov-Smirnoff test for normal distribution of variables

	Kolmogorov-Smirnoff test		
	Comparison group	ASD group	Combined group
A-scale	*D*(27)=.14, *p*=.16	*D*(27)=.10, *p*=.20	*D*(54)=.24, *p*<.001***
A1	*D*(27)=.22, *p*=.002**	*D*(27)=.10, *p*=.20	*D*(54)=.20, *p*<.001***
A2	*D*(27)=.28, *p*<.001***	*D*(27)=.15, *p*=.11	*D*(54)=.17, *p*<.001***
A3	*D*(27)=.18, *p*=.03*	*D*(27)=.13, *p*=.20	*D*(54)=.20, *p*<.001***
B-scale	*D*(27)=.17, *p*=.04*	*D*(27)=.08, *p*=.20	*D*(54)=.25, *p*<.001***
B1	*D*(27)=.53, *p*<.001***	*D*(26)=.13, *p*=.20	*D*(53)=.30, *p*<.001***
B2	*D*(27)=.41, *p*<.001***	*D*(27)=.21, *p*=.003**	*D*(54)=.25, *p*<.001***
B3	*D*(27)=.25, *p*<.001***	*D*(27)=.14, *p*=.19	*D*(54)=.21, *p*<.001***
B4	*D*(27)=.51, *p*<.001***	*D*(27)=.13, *p*=.20	*D*(54)=.30, *p*<.001***
Age	*D*(26)=.14, *p*=.20	*D*(27)=.17, *p*=.04*	*D*(53)=.15, *p*=.005**
IQ	*D*(24)=.08, *p*=.20	*D*(17)=.11, *p*=.20	*D*(41)=.11, *p*=.20

* = distribution of variable differs significantly from normal at p<.05; ** = distribution of variable differs significantly from normal at p<.01, *** = distribution of variable differs significantly from normal at p<.001.

Between group differences

Using a Mann Whitney U analysis, the difference in age between the ASD and comparison groups did not reach significance, U=269.00, z= -1.47, p=.14, r=-.20. Using independent sample t-tests, the difference between groups on estimated IQ was also not significant, t(39)=-1.47, p=.15, d=.47, 95% CI [-14.82, 2.36].

Missing data

Examination of the percentages of missing data showed that no particular question had an excessive amount of missing data. The overall maximum amount of missing data for any one question was 13% (for one question: ED5 'Once [name] started talking did they have conversations just to be sociable?'). The majority of the missing data for this question was missing from the ASD group (86%). There were 19 questions with no missing data at all. Subscales scores were successfully generated for all cases apart from one ASD group case for scale B1. This person was missing data for two out of the three questions on this scale and therefore it was not possible to compute an overall scale score.

Reliability

Internal consistency

Cronbach's alphas are displayed in Table 3. Internal consistency was good for all subscales. Although internal consistency for B1: 'Repetitive motor movements or speech' was slightly lower than the other subscales this still falls within the acceptable range. The overall internal consistency of the two scales was very high.

Table 3

Subscale Cronbach's alphas (α)

A: Social Communication and Interaction	**0.97**
A1: Social Emotional Reciprocity	0.87
A2: Deficits in nonverbal behaviour used for social interaction	0.93
A3: Deficits in forming, maintaining and understanding relationships	0.95
B: Restricted repetitive patterns of behaviour, interests or activities	**0.92**
B1: Stereotyped or repetitive movements	0.71
B2: Insistence on sameness	0.87
B3: Restricted fixated interests	0.79
B4: Abnormal sensory response	0.82

Cronbach's alpha were examined for all scale items and for no scale was the Cronbach's alpha significantly improved by the deletion of any items. Item-total correlations were also examined, with a correlation of less than .4 being considered particularly low. When examining all the items within the overall A-scale, some items were found to have low item-total correlations. These were L29 'Do conversations with [name] tend to go off in unexpected directions?' (r=.30), NVC45 'Can [name] look disgusted?' (r=.38), NVC47 'Do his/her expressions ever appear to be exaggerated or put on?' (r=.29), and SE57 'How about sharing his/her excitement with others?' (r=.23). Similarly some items within the overall B-scale had low item-total correlations. These were I64 'Has [name] ever seemed unusually interested in,

and absorbed by, things that spin?' (*r*=.38), I66 'Has [name] ever seemed unusually sensitive to sensations like touch or smell?' (*r*=.36), and I68 'Has [name] ever shown any hand or finger mannerisms when excited or distressed?' (*r*=.35). It was deemed unhelpful to remove any of these items as it did not improve the Cronbach's alpha and removal would negatively affect the content validity of the interview by reducing its coverage of the DSM-5 criteria.

Item-total correlations were also examined for each scale item within the individual subscales. Within the subscales, only one question within scale A1 (SE55 'Comes to show you something that interests him/her') was found to have a low item-total correlation at *r*=.27. As it did not improve the Cronbach's alpha to delete this item it was not removed from the interview.

Inter-rater reliability

Inter-rater reliability was very good for all subscales, both within the ASD and comparison groups separately and for the two groups combined. Intra-class correlations using one-way random single measures are displayed in Table 4.

Table 4

Intraclass correlation coefficients

	ASD	Comparison	Combined
A: Social Communication and Interaction	**.92**	**.94**	**.99**
A1: Social Emotional Reciprocity	.96	.87	.99
A2: Deficits in nonverbal behaviour used for social interaction	.94	.98	.99
A3: Deficits in forming, maintaining and understanding relationships	.87	.99	.99
B: Restricted repetitive patterns of behaviour, interests or activities	**.98**	**.90**	**.99**
B1: Stereotyped or repetitive movements	.86	1	.92
B2: Insistence on sameness	.94	.94	.98
B3: Restricted fixated interests	.96	.81	.97
B4: Abnormal sensory response	.93	.94	.95

Criterion Validity

Discrimination between ASD and comparison population

Mean scores for the 3Di-sva subscales for the ASD group and comparison group are displayed in Table 5. The difference between the scores of the two groups was highly significant for all subscales, with very large effect sizes in all cases.

Examination of histograms showing the distribution of scores for all subscales showed that for the majority of cases data was normally distributed for the ASD group, with the exception of scale B2 which was negatively skewed. For the comparison group data was positively skewed in all cases. Histograms showing the distribution of scores for the A-scale and B-scale are shown in Figures 3 and 4 respectively. There was no overlap in scores between the groups on the A-scale, with the majority of comparison group cases scoring very low and a distribution of higher scores attained by the ASD group. For the B-scale, the comparison group again score extremely low in all cases. There is a slight overlap between ASD and comparison cases on this scale, with ASD group scores being distributed more evenly across the range of possible scores.

Table 5

Difference in 3Di-sva scores by group

	ASD		**Comparison**		**Significance of difference**	**Effect size**
	Mean (SD) Range[1]	Median	Mean (SD) Range[1]	Median		
A: Social Communication and Interaction	3.13 (.74) 1.65-4.48	3.09	.32 (.24) .00-1.00	.29	*t*(31.15)=18.73, *p*<.001, 95% CI [2.50, 3.11] ***	*d*=6.71
A1: Social Emotional Reciprocity	1.07 (.27) .36-1.67	1.08	.13 (.14) .00-.64	.14	*U*=1.50, *z*= -6.30, p<.001 ***	*r*=-.86
A2: Deficits in nonverbal behaviour used for social interaction	.91 (.40) .29-1.80	.80	.07 (.09) .00-.29	.00	*U*=1.00, *z*= -6.35, *p*<.001 ***	*r*=-.86
A3: Deficits in forming, maintaining and understanding relationships	1.15 (.31) .65-1.94	1.06	.12 (.13) .00-.47	.12	*U*<.001, *z*= -6.33, *p*<.001 ***	*r*=-.86

B: Restricted repetitive patterns of behaviour, interests or activities	4.65 (2.36) .25-8.43	4.67	.41 (.42) .00-1.32	.25	*U*=27.50, *z*= -5.84, *p*<.001 ***	*r*=-.80
B1: Stereotyped or repetitive movements	.68 (.47) .00-1.50	.75	.03 (.08) .00-.25	.00	*U*=66.00, *z*= -5.50, *p*<.001 ***	*r*=-.76
B2: Insistence on sameness	2.13 (.98) .00-3.33	2.33	.15 (.23) .00-.67	.00	*U*=43.50, *z*= -5.71, *p*<.001 ***	*r*=-.79
B3: Restricted fixated interests	1.06 (.64) .00-2.00	1.00	.18 (.21) .00-.80	.20	*U*=64.00, *z*= -5.27, *p*<.001 ***	*r*=-.72
B4: Abnormal sensory response	.79 (.65) .00-2.00	.80	.05 (.13) .00-.40	.00	*U*=106.00, *z*= -4.87, *p*<.001 ***	*r*=-.66

Note. [1] Possible range for the A-scale is 0-6 and for the B-scale is 0-8. Possible range for subscales A1, A2, A3, B1, B2, B3 and B4 is 0-2. *** = significant between group difference at p<.001.

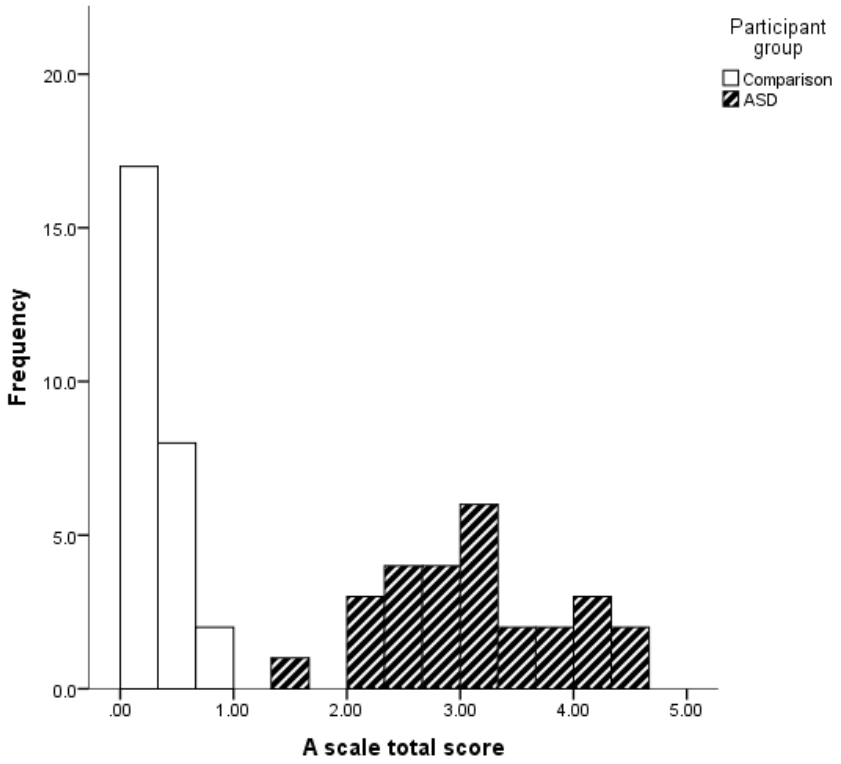

Figure 3. ASD group and comparison group total scores on the 3Di-sva A-scale.

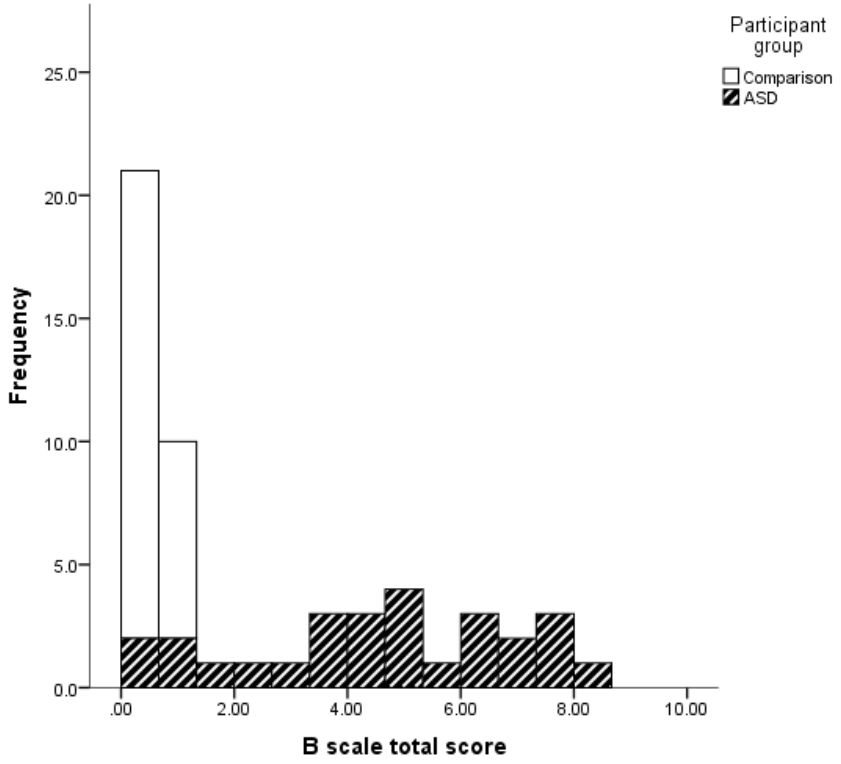

Figure 4. ASD group and comparison group total scores on the 3Di-sva B-scale.

Receiver Operating Characteristic (ROC)

A ROC curve was generated to analyse the ability of the 3Di-sva A-scale and B-scale to discriminate between ASD and comparison participants. The ROC curve is displayed in Figure 5. Discriminatory ability of the two scales was assessed using the area under the curve (AUC), which indicated high overall accuracy of both scales. For the A-scale, AUC=1 (*SE*<.001*)*, *p*<.001, 95% CI [1,1]. For the B-scale, AUC=.96 (*SE*=.02*)*, *p*<.001, 95% CI [.95, 1].

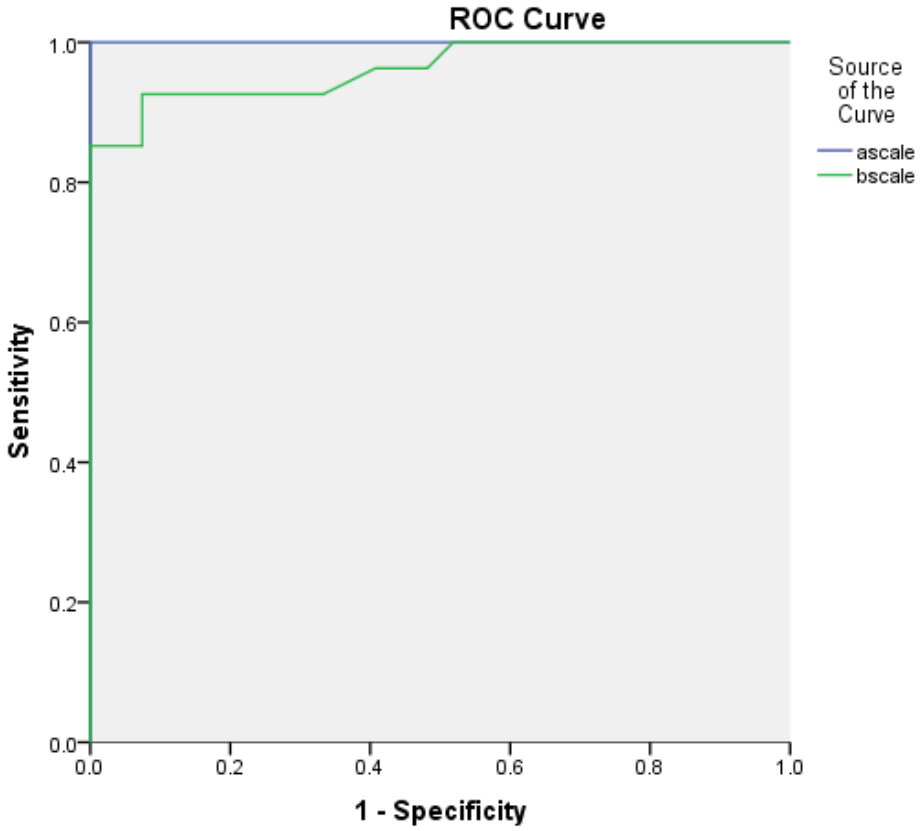

Figure 5. ROC curve of the 3Di-sva A-scale and B-scale

Sensitivity and specificity

Cut points which maximised both sensitivity and specificity for both the A-scale and B-scale were identified. For the A-scale (range 0 to 6) this was a scaled

score of 1.4 and for the B-scale (range 0 to 8) this was a scaled score of 1. In order to be categorised as having ASD by the 3Di-sva a person must score above the cut-off on both scales (in line with DSM-5 criteria). The number of cases correctly categorised by the 3Di-sva is displayed in Table 6.

Table 6

Diagnosis according to 3Di-sva

3Di-sva diagnosis	Participant group	
	ASD	Comparison
Non-ASD	2	27
ASD	25	0

Using the above figures, sensitivity and specificity figures were calculated. Sensitivity (the probability that the 3Di-sva algorithm result is positive when ASD is present) was .93, 95% CI [.74, .99] and specificity (the probability that the result is negative when ASD is not present) was 1, 95% CI [0.85 – 1]. Similarly, in this sample the positive predictive value (the probability that ASD is present when the test is positive) was 1, 95% CI [.83, 1] and the negative predictive value (the probability that ASD is not present when the test is negative) is .93, 95% CI [.76 - .99].

Correlation between 3Di-sva scores and scores on the ADOS module 4

Correlations between each of the 3Di-sva subscales and the ADOS module 4 subscales are displayed in Table 7. The majority of the correlations between scores on the 3Di-sva subscales and the ADOS subscales were found not to be significant, with the exception of the correlation between the 3Di-sva B2 subscale (Insistence on

sameness) and the ADOS Social Interaction scale, and between the 3Di-sva B4 subscale (Abnormal sensory response) and both the ADOS Social Interaction scale and the ADOS Imagination and Creativity scale. However, these correlations were significant at the p=.02 - .04 level and due to the large number of correlations carried out, may be attributable to Type 1 error.

Table 7

Correlations between 3Di-sva subscales and ADOS module 4 subscales for ASD group

	ADOS C	ADOS SI	ADOS I/C	ADOS SBRI	ADOS C&SI
A-scale	$r_s(27)=.15$, $p=.45$	$r_s(27)=.19$, $p=.35$	$r_s(25)=.13$, $p=.53$	$r_s(25)=-.06$, $p=.79$	$r_s(27)=.18$, $p=.38$
A1	$r_s(27)=.15$, $p=.45$	$r_s(27)=.22$, $p=.26$	$r_s(25)=.14$, $p=.50$	$r_s(25)=.02$, $p=.92$	$r_s(27)=.21$, $p=.30$
A2	$r_s(27)=.10$, $p=.62$	$r_s(27)=.27$, $p=.17$	$r_s(25)=.04$, $p=.83$	$r_s(25)=.01$, $p=.96$	$r_s(27)=.19$, $p=.34$
A3	$r_s(27)=.22$, $p=.28$	$r_s(27)=-.13$, $p=.52$	$r_s(25)=.11$ $p=.62$	$r_s(25)=-.07$, $p=.76$	$r_s(27)=-.01$, $p=.95$
B-scale	$r_s(27)=-.02$, $p=.93$	$r_s(27)=.35$, $p=.08$	$r_s(25)=.28$, $p=.17$	$r_s(25)=.05$, $p=.82$	$r_s(27)=.17$, $p=.40$
B1	$r_s(26)=-.03$, $p=.90$	$r_s(26)=.11$, $p=.59$	$r_s(24)=.01$, $p=.98$	$r_s(24)=.06$, $p=.78$	$r_s(26)=.07$, $p=.74$
B2	$r_s(27)=.05$, $p=.81$	$r_s(27)=.44$, $p=.02$*	$r_s(25)=.38$, $p=.06$	$r_s(25)=-.05$, $p=.82$	$r_s(27)=.26$, $p=.20$
B3	$r_s(27)=-.01$, $p=.98$	$r_s(27)=.13$, $p=.53$	$r_s(25)=-.05$, $p=.80$	$r_s(25)=.07$, $p=.75$	$r_s(27)=.002$, $p=.99$
B4	$r_s(27)=-.17$, $p=.41$	$r_s(27)=.40$, $p=.04$*	$r_s(25)=.46$, $p=.02$*	$r_s(25)=-.04$, $p=.84$	$r_s(27)=.12$, $p=.57$

Note. ADOS C = ADOS communication total; ADOS SI = ADOS social interaction total; ADOS I/C= ADOS imagination creativity total; ADOS SBRI = ADOS stereotyped behaviours and restricted interests total; ADOS C&SI= ADOS combined communication and social interaction total; * = significant at p<.05.

Construct validity

Correlation between scores on the 3Di-sva A-scale and B-scale

There was a significant correlation between scores on the 3Di-sva A-scale and B-scale for the ASD group, $r(27)=.55$, $p=.003$, and the overall sample, $r_s(54)=.79$, $p<.001$. The correlation between scores on the two scales was not significant in the comparison group, $r_s(27)=.07$, $p=.72$.

Correlation between 3Di-sva score and estimated IQ

There was not a significant correlation between scores on the 3Di-sva A-scale and estimated IQ for either the ASD group, $r(17)=.06$, $p=.81$, the comparison group, $r(24)=-.08$, $p=.72$, or the overall sample, $r_s(41)=-.19$, $p=.25$. There was also not a significant correlation between scores on the 3Di-sva B-scales and estimated IQ for either the ASD group, $r(17)=-.01$, $p=.97$, the comparison group, $r_s(24)=-.13$, $p=.54$, or the overall sample, $r_s(41)=-.21$, $p=.20$.

Gender differences in the comparison population

There was not a significant difference between scores for males and females in the comparison population on the A-scale, (males: *M*=.37, *SD*=.25, *Mdn*=.*32*; females: *M*=.26, *SD*=.21, *Mdn*=2.7), $t(25)=1.19$, $p=.25$, $d=.48$, 95% CI [-.08, .29]. There was also not a significant difference in scores for males and females in the comparison population on the B-scale, (males: *M*=.41, *SD*=.37, *Mdn*=.33; females: *M*=.40, *SD*=.49, *Mdn*=.20), *U*=78.5, z=-.57, p=.58, *r*=-.11.

Correlation between 3Di-sva score and age

The correlation between score on the A-scale and age was not significant for the ASD group, $r_s(27)=-.23$, p=.26, or the combined groups, $r_s(53)=.02$, p=.87, however it did reach significance in the comparison group, $r(26)= -.42$, $p=.03$. The correlation between score in the B-scale and age did not reach significance in any group (ASD group: $r_s(27)=.17$, p=.41, comparison group: $r_s(26)= .14$, $p=.49$, combined groups: $r_s(53)=.26$, p=.07).

Discussion

The current study examined the psychometric properties of a new informant report tool for facilitating the diagnosis of ASD in adults, the 3Di-sva. The findings show that the 3Di-sva is a reliable instrument. The internal consistency of the subscales ranged from acceptable to excellent, suggesting that items within each subscale are sufficiently reflective of the same underlying concept. There was also a high level of agreement between raters for all the subscales, demonstrating that the 3Di-sva can be consistently scored by raters, one of whom is blind to participant group.

The 3Di-sva also demonstrated strong criterion related validity. Participants with ASD had significantly and substantially higher scores than comparison participants across all subscales. The high Area under the Curve, and the high sensitivity (93%) and specificity (100%) provide evidence that the 3Di-sva is able to correctly classify individuals as having ASD or not in the vast majority of cases.

Another aspect of criterion validity examined was the correlation between 3Di-sva subscale scores and ADOS module 4 scores for the ASD group. Three significant correlations were found; however the significant correlations are between

scales that appear to have little relationship conceptually, whereas correlations which may be more anticipated (e.g. between the 3Di-sva A scale and the ADOS combined Communication and Social Interaction scale) were not found. As noted previously, the significance of the correlations found could be attributable to Type 1 error due to the large number of correlations carried out. It is also notable that low correlations between the ADOS and parent report have occasionally been found elsewhere, for example between diagnosis on the ADOS and ADI-R (Bishop & Norbury, 2002; De Bildt et al., 2004).

There may also have been a lack of power to detect other significant correlations between scores on the ADOS and 3Di-sva. The power achieved was enough to detect significance for large effects, however due to differences in the type of measures being compared (an informant report looking at developmental history and current functioning versus an observation assessing functioning at one moment in time), it is conceivable that smaller effects would be expected. Some of the correlations (e.g. between 3Di-sva A2 score 'Deficits in nonverbal behaviour used for social interaction' and ADOS Social Interaction score) did show a medium effect size, despite being statistically insignificant.

In terms of construct validity, as expected there was a strong positive correlation between scores on the A-scale and the B-scale for the ASD group, suggesting that the higher symptoms are in one area, the higher they are in the other. This finding did not extend to the comparison group. Although it might be expected that some degree of traits on one scale would be related to traits on the other (albeit below thresholds), it appears that people without ASD scored so low on both scales that there was not sufficient variability to identify any correlation between them. A possible interpretation of this is that although the 3Di-sva provides a clear categorical

assessment of ASD, it is less successful in this sample at providing a dimensional assessment of ASD traits.

Although a correlation between score and IQ was hypothesised, on the basis that a relationship between higher ASD traits and lower IQ has previously been demonstrated (Brugha et al., 2009), no correlation was found here for either group. This is a positive finding in that it suggests score on the 3Di-sva is not influenced by IQ. Similarly, although a relationship between score and gender in the comparison population was hypothesised, on the basis that traits of ASD are generally higher within males in the typically developing but not the clinical population (Constantino, & Todd, 2003; Baron-Cohen et al., 2001), no significant relationship was found. This may suggest that 3Di-sva score is not influenced by gender; however it should be noted that despite an insignificant result, a medium effect size was detected for the relationship between gender and A-scale score, with males generally scoring higher than females on this scale, again suggesting that the insignificant result could be attributed to lack of power to detect the effect.

Interestingly, an unexpected significant correlation was found between age and 3Di-sva A-scale score in the comparison group. One possible explanation for the finding is that some of the 'current' items on the A-scale may pick up behaviours towards parents that may be considered fairly normal in some young adults, such as 'Does [name] have conversations with you just to be sociable, for instance, does s/he made small talk?'. It was found that parents of younger participants occasionally noted that their child did not, but they felt it was an expected phase of behaviour. This could explain why older comparison participants scored lower on this scale overall. This is not however a problem for the validity of the 3Di-sva, as despite

slightly higher A-scale scores for younger comparison participants, all comparison participants scored lower on this scale than ASD participants.

As well as the criterion and construct validity investigated here, it appeared that participants in the ASD group subjectively felt the 3Di-sva had good face validity. Although this was not formally measured, informants often commented when asked that they felt the interview had covered all their areas of concern in relation to a possible diagnosis of ASD. The 3Di-sva also has good content validity, as it was designed to represent all the DSM-5 ASD diagnostic criteria. It is notable however that the content of the interview is weighted towards to the A-scale, with the B-scale subscales having many fewer items than those in the A-scale. This was controlled for by creating weighted scale scores, however as the B-scale subscales also demonstrated slightly lower (although still acceptable) internal consistency, it may be helpful to investigate whether this can be improved by the addition of further items within the B-scale. This is an issue which does not affect the 3Di-sva alone; other ASD diagnostic tools used with adults such as the ADOS module 4 and ADI-R also include more social and communication items than repetitive and stereotyped behaviour items. There a general need to improve the measurement of repetitive and stereotyped behaviour in adult ASD diagnostic tools.

It is notable that the 3Di-sva would have had perfect sensitivity and specificity if using A-scale score alone to indicate diagnosis, as opposed to meeting the threshold on both the A-scale and B-scale. The two cases that were not correctly classified by the 3Di-sva were missed on the basis of a low score on the B-scale. It was considered important that individuals met the threshold on both scales when considering sensitivity and specificity, as DSM-5 criteria requires individuals to have experienced symptoms in both areas. There is known to be group of individuals who

display only the social reciprocity and communication deficits seen in ASD without the significant repetitive and stereotyped behaviours (Mandy, Charman, Gilmour & Skuse, 2011), and the DSM-5 introduced a new diagnosis for such cases: Social Communication Disorder (SCD). Although the two cases in question here did receive a diagnosis of ASD, it is possible that SCD could be a more suitable diagnosis. Nonetheless, one could argue that when used in clinical practice, a score above cut-off on both the A-scale and B-scale could be taken as a strong indication of a positive diagnosis of ASD, with cases which score above cut-off on the A-scale but not B-scale also indicating possible diagnosis of ASD. Cases that do not score above cut-off on the A-scale (even if they do on the B-scale) should be considered highly unlikely to indicate a positive diagnosis of ASD.

The fact that B-scale score was less sensitive to diagnosis of ASD than A-scale score may be related to the idea that in an adult population, the DSM-5 criteria of Restricted, Repetitive Patterns of Behaviour, Interest or Activities are less persistent. It has been found that adults are likely to show less restricted and repetitive behaviours and interests symptoms whilst displaying the same communication symptoms as younger cohorts (Seltzer et al., 2003), and that prevalence of symptoms related to social reciprocity and non-verbal communication are generally higher than symptoms of repetitive behaviours and stereotyped interests in adults (Shattuck et al., 2007). For this reason it may be helpful to focus any new B-scale questions on behaviours that may have been present during childhood rather than current behaviours; or to get a better understanding of how this aspect of ASD manifests in adulthood.

The current research also confirms that the 3Di-sva is a time and resource efficient tool suitable for use within ASD diagnostic clinics. Within the ASD group,

length of the interview varied between 23 and 75 minutes, with the mean length of interview being 50 minutes. Interviews were successfully conducted both in person and on the telephone, meaning that parents who were unable to attend clinic appointments were still able to be interviewed. All researchers and clinicians who took part in the study, including the psychology undergraduate who carried out scoring for inter-rater reliability purposes, received around one hour of instruction in its use, showing that training on the 3Di-sva is quick and straightforward.

Limitations and future directions

One limitation of the current research is that it does not include a comparison group with other clinical diagnoses. As we know that ASD is associated with extensive comorbidity (Mukaddes et al., 2010) and people with ASD can display similar features to other disorders (Baastiansen et al., 2011, Zandt et al., 2009), a more difficult and ecologically relevant test of the 3Di-sva would be to examine its ability to discriminate adults with ASD from those in other clinical populations. It is necessary to test the 3Di-sva in populations with psychosis, anxiety, and depression, whose presentations may include features and symptoms that could potentially be picked up by the 3Di-sva, for example difficulties with social interactions and restricted behaviours. Such research has already begun, as the current research was conducted as part of a joint project and McKenner (2015) reports on the 3Di-sva when used with a clinical control group. However this research contains participants with a mixture of different disorders such as anxiety, depression, psychosis and personality disorder. This could be extended further by comparing clinical control groups with separate diagnoses in order to provide information on the diagnostic utility of the 3Di-sva in specific clinical areas.

The current research was restricted by the fact that the time was not available to investigate test retest reliability. In order to ensure that the 3Di-sva provides a reliable measure of ASD symptoms across time, there is a need for the future analysis of test retest reliability. This could be done by approaching the current sample and conducting the 3Di-sva again with the same informant. Another additional factor to investigate in future research would be the correlation between the 3Di-sva and other tools purporting to measure ASD symptoms. Although limited correlation was found with the ADOS module 4 in the current research, it would also be interesting to further investigate the validity in respect to the 3Di-svas relationship to other types of instruments such as self-report (e.g. RAADS-R) and other informant report (e.g. ADI-R) tools.

As previously alluded to, the relatively small sample size has been somewhat of a limitation in this research. The sample size was enough to detect large effect sizes, and we were able to clearly demonstrate the criterion validity of the 3Di-sva in terms of its ability to accurately distinguish between groups. However, as discussed above, it is possible that some of the other hypotheses were rejected as a result of not enough power to detect smaller effects. Future research with a larger sample size may help reach clearer conclusions about these effects. Additionally, the current sample did not allow for investigation of internal consistency of the scales for separate groups, due to the very low variance within groups amongst some of the subscales, particularly in the comparison group. A large sample size would likely lead to more variance, allowing a valid demonstration of internal consistency in the separate populations. Furthermore, a larger sample size would be of benefit in terms of further analysing the construct validity of the 3Di-sva using factor analysis. It is also worth noting that only 10 ASD and 15 comparison cases were audio-recorded

and therefore included in the inter-rater reliability analyses; this is a small number and so results should be treated with some degree of caution. Similarly, the relatively small sample size could affect the precision of the estimates of sensitivity and specificity. Further investigation of these factors in a larger sample would be of value.

It is acknowledged that there was a lack of formal screening for symptoms of ASD or mental health difficulties in the control group, due to the limited time and financial resources available. It would be of benefit for future research of the 3Di-sva to include tools such as the AQ, Generalised Anxiety Disorder Assessment (GAD-7; Spitzer, Kroenke, Williams & Löwe, 2006) and Patient Health Questionnaire (PHQ-9; Spitzer, Kroenke & Williams, 1999) to measure self-reported symptoms of ASD, anxiety and depression respectively. Participants were asked whether they experienced any symptoms of ASD or mental health difficulties when screening for eligibility, but it is possible that participants could be unaware of or unwilling to directly report any such symptoms. I would argue however that any presence of unreported ASD or mental health symptoms would make it more difficult to demonstrate the between-group differences found here, suggesting the presence of any such symptoms would not affect the validity of the conclusions reached.

Another potential confound to acknowledge is that we did not explore whether the method of administration (i.e. in person versus telephone) or who acted as the informant (i.e. mother versus other informant) influenced the results. The vast majority of informants in both groups were mothers, meaning there was not sufficient variability to investigate differences in outcomes depending on informant. The majority of comparison group interviews were done over the phone, whereas the number of interviews completed in person versus on the phone in the ASD group

was more equal across the cases for which method of administration was known. As we cannot be certain of the method of administration in the historical ASD group cases, which make up half of the ASD group, there were not sufficient numbers with which to analyse any differences related to method of administration. Due to the highly structured nature of the 3Di-sva, it is unlikely that method of administration would influence outcome, and subjectively there appeared to be no difference in interviews conducted in person and on the phone. However, it would be worth investigating whether these variables affect the results in the future, using a larger sample.

Implications and Conclusions

This research has shown that 3Di-sva is a potentially useful tool as part of an ASD assessment for adults. It has proved to be reliable, in terms of good internal consistency and high inter-rater reliability, as well as highly accurate at discriminating between individuals with and without ASD. It is also a time and cost efficient tool, which is easy to administer and score. It provides an indication of diagnosis according to DSM-5 criteria, allowing for assessment of symptoms across the range of the DSM-5 criteria. The 3Di-sva could be used as part of a multi-dimensional assessment, providing valuable informant report information alongside other sources of information such as observation using the ADOS module 4.

As previously discussed, informant report is an important source of information, which should be included whenever possible when carrying out an adult ASD assessment. The 3Di-sva has potential to be an improvement on other currently available NICE (2012) recommended informant report tools, namely the ADI-R, AAA, ASDI and DISCO. This research has provided good initial evidence for the

psychometric properties of the 3Di-sva; evidence which is lacking for the other tools. Additionally the time and cost efficiency makes the 3Di-sva more suitable for use in clinics with limited resources than the ADI-R and DISCO, and it is the only known informant report tool currently available which assesses adult ASD in line with DSM-5 criteria.

This research is the first step in validating the instrument and further work to continue to demonstrate its usefulness is still required. It would be helpful to complete further research using a larger sample size, and it is essential that test-retest reliability and ability to discriminate between different clinical control groups is investigated.

References

Andersen, L. M., Näswall, K., Manouilenko, I., Nylander, L., Edgar, J., Ritvo, R. A., Ritvo, E., & Bejerot, S. (2011). The Swedish version of the Ritvo autism and Asperger diagnostic scale: Revised (RAADS-R). A validation study of a rating scale for adults. *Journal of Autism and Developmental Disorders*, *41*(12), 1635-1645.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Autism Act 2009. Retrieved from http://www.legislation.gov.uk/ukpga/2009/15.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a

population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210-215.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry*, *194*(6), 500-509.

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The adult Asperger assessment (AAA): a diagnostic method. *Journal of Autism and Developmental Disorders*, *35*(6), 807-819.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keysers, C., & de Bildt, A. (2011). Diagnosing Autism spectrum disorders in adults: the use of Autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders*, *41*(9), 1256-1266.

Bishop, D. V., & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. *Journal of Child Psychology and Psychiatry*, *43*(7), 917-929.

Bishop, S. L., & Seltzer, M. M. (2012). Self-reported autism symptoms in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2354-2363.

Brugha, T. S., McManus, S., Bankart, J., Scott. F., Purdon, S., Smith, J., Bebbington, P., Jenkins, R., & Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry*, *68*(5), 458-466.

Brugha, T., McManus, S., Meltzer, H., Smith, J., Scott, F. J., Purdon, S., Harris, J., & Bankart, J. (2009). Autism spectrum disorders in adults living in households throughout England: Report from the adult psychiatric morbidity survey 2007.*Leeds: The NHS Information Centre for Health and Social Care*.

Centres for Disease Control and Prevention (2014). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years – Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. MMWR 2014; 63(SS02), 1-21.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Jama*, *285*(2

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders–lessons from research and practise. *Child and Adolescent Mental Health*, *18*(1), 52-63.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of General Psychiatry*, *60*(5), 524-530.

De Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., & Minderaa, R. (2004). Interrelationship between autism diagnostic observation schedule-generic (ADOS-G), autism diagnostic interview-revised (ADI-R), and the diagnostic and statistical manual of mental disorders (DSM-IV-TR)

classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, *34*(2), 129-137.

Department of Health. (2010). Fulfilling and rewarding lives: The strategy for adults with autism in England. UK: Author.

Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders?. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(8), 788-797.

Gillberg, I. C., & Gillberg, C. (1989). Asperger syndrome—some epidemiological considerations: a research note. *Journal of Child Psychology and Psychiatry*, *30*(4), 631-638.

Gillberg, C., Gillberg, C., Råstam, M., & Wentz, E. (2001). The Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI): a preliminary study of a new structured clinical interview. *Autism*, *5*(1), 57-66.

Howlin, P., & Moss, P. (2012). In review: Adults with autism spectrum disorders. *Canadian Journal of Psychiatry*, *57*(5), 275.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679-1690.

Hus, V., & Lord, C. (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, 1-17.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One*, *2*(9), e883.

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*(5), 659-685.

Lord, C., Pickles, A., McLennan, J., Rutter, M., Bregman, J., Folstein, S., Fombomme, E., Leboyer, M., & Minshew, N. (1997). Diagnosing autism: analyses of data from the Autism Diagnostic Interview. *Journal of Autism and Developmental Disorders, 27*(5), 501-517.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part 1): Modules 1-4.* Torrance, CA: Western Psychological Services.

Mandy, W., Charman, T., Gilmour, J., & Skuse, D. (2011). Toward specifying pervasive developmental disorder—not otherwise specified. *Autism Research,4*(2), 121-131.

McKenner, M. (2015). *Validating the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva) for use in a clinical population* (Unpublished doctoral thesis). University College London, London.

Mukaddes, N. M., Hergüner, S., & Tanidir, C. (2010). Psychiatric disorders in individuals with high-functioning autism and Asperger's disorder: similarities and differences. *World Journal of Biological Psychiatry*, *11*(8), 964-971.

National Institute for Health and Care Excellence (2012). *Autism: recognition, referral, diagnosis and management of adults on the autism spectrum. CG142*. London: National Institute for Health and Care Excellence.

Nygren, G., Hagberg, B., Billstedt, E., Skoglund, Å., Gillberg, C., & Johansson, M. (2009). The Swedish version of the diagnostic interview for social and communication disorders (DISCO-10). Psychometric properties. *Journal of Autism and Developmental Disorders*, *39*(5), 730-741.

Ritvo, E. R., Ritvo, R., Freeman, B. J., & Mason-Brothers, A. (1994). Clinical characteristics of mild autism in adults. *Comprehensive Psychiatry*, *35*(2), 149-156.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., Ritvo, M. J., Hufnagel, D. H., McMahon, W., Tonge, B., Mataix-Cols, D., Jassi, A., Attwood, T., & Eloff, J. (2011). The Ritvo Autism Asperger diagnostic scale-revised (RAADS-R): A scale to assist the diagnosis of autism spectrum disorder in adults: An international validation study. *Journal of Autism and Developmental disorders*, *41*(8), 1076-1089.

Santosh, P. J., Mandy, W. P., Puura, K., Kaartinen, M., Warrington, R., & Skuse, D. H. (2009). The construction and validation of a short form of the developmental, diagnostic and dimensional interview. *European Child & Adolescent Psychiatry*, *18*(8), 521-524.

Seltzer, M. M., Krauss, M. W., Shattuck, P. T., Orsmond, G., Swe, A., & Lord, C. (2003). The symptoms of autism spectrum disorders in adolescence and adulthood. *Journal of Autism and Developmental Disorders*, *33*(6), 565-581.

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., Lounds, J., & Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1735-1747.

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 & Adolescent Psychiatry*, *43*(5), 548-558.

Spitzer, R., Kroenke, K., Williams, J. (1999). Validation and utility of a self-report Version of PRIME-MD: the PHQ Primary Care Study. *Journal of the American Medical Association, 282*, 1737-1744.

Spitzer, R. L., Kroenke, K, Williams, J. B.W. & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine,166,* 1092-1097.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale–Fourth Edition (WAIS–IV).* San Antonio, TX: Pearson.

Wechsler, D. (2009). *Test of Premorbid Functioning – UK Edition.* London: Pearson Assessment.

Wing, L., Leekam, S. R., Libby, S. J., Gould, J., & Larcombe, M. (2002). The diagnostic interview for social and communication disorders: Background, inter-rater reliability and clinical use. *Journal of Child Psychology and Psychiatry*, *43*(3), 307-325.

World Health Organization. (1992). *International classification of diseases: Diagnostic criteria for research* (10th edition). Geneva, Switzerland: Author.

Zandt, F., Prior, M., & Kyrios, M. (2009). Similarities and differences between children and adolescents with autism spectrum disorder and those with obsessive compulsive disorder executive functioning and repetitive behaviour. *Autism*, *13*(1), 43-57.

PART 3: CRITICAL APPRAISAL

Introduction

This critical appraisal extends the discussion of both my literature review and empirical paper, reflecting on two main areas of the research process which have been challenging and thought-provoking. The first is that of ideal versus achievable research, in which I consider how my own study led me to recognise the difficulties faced in achieving the ideals I set up within my literature review for autism spectrum disorder (ASD) diagnostic tool research. The second is a reflection on my learning process about research within NHS settings, including how my expectations for others' roles within my research did not fit with reality, but how this ultimately benefitted me in terms of my own appreciation for and understanding of the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva).

Ideal versus achievable research

Following my literature review, it was of interest to me that no study I reviewed fulfilled all the quality criteria outlined in the critical appraisal tool used (the BMJ Critical Evidence (2014) tool), and several studies fulfilled troublingly few. I was surprised at the poor quality of some of the research, and concluded the review with remarks about what an ideal study into an ASD diagnostic tool would consist of. I feel it is important for me to reflect upon how I experienced the reality of completing my own research into an ASD diagnostic tool and how this links to the literature I reviewed.

The literature review was beneficial in aiding my thinking about the design of my own study, highlighting some of the shortfalls present in existing research and helping me to hold in mind the important and desirable factors for mine. My study

met the majority of the criteria within the critical appraisal tool used for the literature review; however it did fall short of the 'ideal' study described in two main ways. The first was the lack of a clinical control group and the second was a measure of test-retest reliability. Completing my own research allowed me to better comprehend some of the barriers that appear to be limiting the standard of research often produced in this area, particularly in relation to research on informant report tools.

Recruiting participants

It became apparent that, whilst recruiting ASD participants for the research was fairly straightforward, raising the interest of people without ASD to take part was considerably more challenging. Not only were we asking individuals to participate in research on a topic to which they necessarily had no association (as it was a requirement that comparison participants did not suspect they had traits of ASD), we were also asking them to recruit someone else to take part in it too – as research on informant reports inevitably requires both the individual and an informant to agree to take part. In an attempt to gain the interest of non-ASD participants we began by offering a £5 voucher for their time. We quickly became aware however that this was not enough of an incentive and raised this to £10. The higher amount did result in more non-ASD participants coming forward, however this then constrained the total number of people we could recruit due to financial limitations.

Recruitment of non-clinical comparison participants was a challenge, yet I have posited in both my literature review and my empirical paper that it is also essential that research of ASD diagnostic tools demonstrates the validity of the instrument when used with different clinical control populations. Within clinical

control populations it is likely that recruitment would be even more difficult, as not only is ASD again a condition unrelated to the person, but the individuals in question are also likely to have their own set of significant difficulties and challenges to focus on. Potential participants could also be put off taking part if they do not want their parent to know that they are receiving mental health treatment. Recent or current illness is cited as a factor which adversely affects recruitment of research participants (Patel, Doku &Tennakoon, 2003), with the level of additional demand put upon patient participants influencing their decision to take part in research (Ross et al., 1999). It seems probable that the more impaired someone is by their mental health difficulties the less likely they are to be interested in participating in something that offers very little in return to them, especially given the additional burden of needing to seek participation from another person as well as themselves.

My insight into the reality of recruiting control populations helps me comprehend why many of studies reviewed presented data in which the clinical control population consisted of individuals who were assessed for ASD but found to have a mental health rather than ASD diagnosis, instead of being recruiting from specific mental health populations. Only one study (on the ADOS module 4, Bastiaansen et al., 2011) investigated the diagnostic tool when used with separate clinical control populations (schizophrenia and personality disorder), despite that fact that it is important to understand the validity of ASD diagnostic tools within separate diagnostic populations. Without such research we cannot uncover potential variations in the validity of the diagnostic tools in different clinical populations, which will be masked when a comparison group consists of such a diverse mix of diagnoses. However whilst we can identify that this is ideal, and indeed necessary, the reality is

that with the difficulties in recruiting for such studies, research is published using whatever type of control group has been achievable.

A lack of the funding needed to compensate participants from control groups is not confined to ASD diagnostic tool research completed as a part of a doctoral thesis. In the current economic climate money for research is something that can often be hard to secure. The Autism Alliance (2015) note that finance for ASD research is often assigned to specific areas such as genetics and early intervention, plus only 30% of all ASD research funding goes to work focussed on adults. Patel et al., (2003) suggest that participants conduct a personal cost-benefit analysis when deciding when to take part in research. Although some benefit may be gained, for example enjoyment of the contact with researchers, it is likely that costs involved in taking part in this type of research exceed the benefits for control group participants, unless they receive suitable financial compensation. It is perhaps simplistic to think that more funding is the answer to the difficulties in recruiting control groups, but my experience did indicate that individuals are somewhat more willing to give up their time if they do get something in return.

As having comparison groups is essential to proving the validity of any ASD diagnostic tool, I envisage that the difficulty of recruiting these groups has been a significant barrier to more studies being published on the existing tools. Of the National Institute of Clinical Excellence (NICE; 2012) recommended informant report tools, only the ADI-R has more than one paper published examining its use in adults. This highlights the fact that more interest and funding is needed for this area, especially if we are to improve diagnostic services for adults with ASD, as recommended in the Strategy for Adults with Autism (Department of Health, 2010).

Reference standards

A dilemma that I came across during my research is the difficulty of having a suitable reference standard against which to judge the sensitivity and specificity of an ASD tool. The ideal reference standard would be an expert clinician consensus diagnosis of ASD made according to the DSM-IV (American Psychiatric Association (APA), 2000), DSM-5 (APA, 2013) or ICD-10 (World Health Organisation, 1992) criteria. The majority of papers reviewed in the literature met this criterion, as did all cases included in my research. However, even when this ideal reference standard is in place, an issue of circularity arises if information gleaned from the instrument under investigation is used when making such diagnoses. This issue was present in some of the papers included in the literature review (Baron-Cohen, Wheelwright, Robinson & Woodbury-Smith, 2005; Hus & Lord, 2014). In my own research, although 3Di-svas completed for the study were not scored at the time of the assessment, we could still have been criticised for including information obtained during the 3Di-sva interview when making diagnostic decisions. We therefore set the criterion that research participants first had to meet criteria for a diagnosis according to the ADOS module 4. Following this, to rule out false positives on the ADOS and ensure diagnoses were made according to DSM-5 criteria, the diagnosis was also confirmed by clinician consensus opinion. The ADOS criterion protected us from the issue of circularity, but brought with it its own complications.

Having the ADOS score criteria meant that we lost cases from the ASD group who did receive a diagnosis of ASD but did not score up on the ADOS. The ADOS itself, although often considered to be 'gold standard', is not a perfect measure. As noted in the literature review, the overall sensitivity and specificity demonstrated across the available ADOS module 4 papers (Bastiaansen et al., 2011; Brugha et al.,

2012; Hus & Lord, 2014; Lord et al., 2000) suggest a sensitivity and specificity of around 89% and 84% respectively. Therefore cases who did have ASD, but whose symptoms were more subtle or better masked during a one-off observation, were excluded from the analysis. This raised concerns for me that we were only allowing the cases with a clearer diagnosis of ASD to be included in our analysis, which makes it easier to demonstrate high sensitivity and specificity. For my own peace of mind I did score the cases for which I had data but were excluded on this basis and found that they would have scored above the 3Di-sva thresholds according to those set in the empirical paper; however it still seems somewhat unsatisfactory to need to exclude these cases.

As an ideal, to overcome these issues, the 3Di-sva would have been used on a group of participants who had received a diagnosis of ASD according to DSM-5 criteria, without the 3Di-sva having been part of the diagnostic process. However this again raises issue of what is ideal versus what is achievable. In this situation we did not have the capacity to complete 3Di-svas, in addition to another interview or instrument that would have been necessary to include in its place, to reach diagnostic conclusions. It was therefore felt that the inclusion criteria set were the optimal way of controlling for the issues around reference standard and circularity that arose.

Measuring cognitive ability

In line with one of the criteria in the appraisal tool used for the literature review, it was considered essential to include an estimate of participant IQ in my research, in order to understand and control for any influence of IQ on 3Di-sva score. It was clear that it would not be practicable to complete a full WAIS-IV (Wechsler, 2008) with each participant, and after careful consideration it was also felt that

completing the WASI (Wechsler, 1999), which takes between 15 and 30 minutes depending on the number of subtests used, would also not be realistic. The Test of Premorbid Functioning (TOPF, Wechsler, 2009), an updated version of the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) which takes around five minutes to complete and requires few materials, was eventually chosen as the most feasible instrument given the time and financial restrictions. The TOPF score, along with demographic information, is used to predict full scale IQ on the WAIS-IV.

However, the TOPF is a tool designed to predict premorbid intellectual function, when there is a suspected loss of cognitive function. This was not the case in the majority our sample, although it was believed that this could be helpful in the clinical control cases (reported on by McKenner, 2015) as it is known that cognitive function can be affected by mental health difficulties (McDermott & Ebmeier, 2009; Michel et al., 2013). It should however be acknowledged that our method of estimating IQ was not perhaps the ideal choice to use across all participants and should we have had unlimited time, finances and enthusiasm from participants, the WASI or WAIS-IV would be used. However, this is where a compromise had to be made and a feasible method of estimating of IQ was certainly better than none at all.

In some cases it proved challenging to even complete the TOPF (a very quick measure which simply involves reading out a list of words) with all participants. This was often the case in the ASD group, where I was frequently given cases to complete a 3Di-sva over the phone, meaning my first chance to complete a TOPF with the participant was when they attended their diagnostic feedback session. Due to the often emotive nature of such an appointment, it was important that I had a test that was quick and undemanding. I consequently think that we would have been less

successful in gaining the data we needed had we chosen the WASI, and therefore under the circumstances made the correct decision.

Unacceptable compromise

My own research enabled me to sympathise with the difficulty of achieving some of the study features required for a high score according the critical appraisal tool, and highlighted that sometimes compromises have to be made. However it has also strengthened my assertion that some of studies examined in the literature review are not of suitable quality, based on the currently available evidence, to be recommended for use. It seems reasonable to expect studies to publish data on the majority of the psychometric properties added to the critical appraisal tool if they are to be used in clinical practice. My research failed to measure just one from the list, test retest reliability, which required more time to investigate than was available. However it seems inexcusable that straightforward data such as internal consistency and association with participant characteristics (e.g. correlation between scores and age or IQ) is not reported in some studies on ASD diagnostic tools. Such statistics are vital to investigate before we can go on to conclude on the validity of an instrument.

During the course of my research I have become more aware of the diagnostic process used within different NHS diagnostic clinics. Although both clinics I was involved with used clinical interview, ADOS module 4 and the 3Di-sva, other services base their diagnostic decisions on the AAA (Baron-Cohen et al., 2005), on the basis that the AAA describes itself to be a complete diagnostic system and is recommended by NICE (2012). However when one considers that there is no data published regarding any form of reliability for the AAA, this does begin to seem

less than ideal. It troubles me that some instruments recommended by NICE (2012) and used by NHS diagnostic services are ones which we can't yet be sure are reliably measuring the concept they claim to measure. I am aware that research into adult ASD diagnostic tools is a relatively new and under researched area, and I hope that over time research in this area develops further, giving us a clearer evidence base upon which to select our diagnostic tools.

Research within NHS diagnostic clinics

Another area I would like to reflect upon is that of the realities of completing research within NHS settings, particularly when you are not working within the particular setting you are recruiting from. As someone who would like to continue clinical research post qualification, the experience of completing my empirical paper has been an important lesson in managing expectations when completing research within the NHS. I believe this is important to consider, as the expectations I had upon others at the start of my research proved to be too high and this led to significant delays in the data collection that could have otherwise been avoided. Not only was my recruitment far more successful once I changed my expectations, being made to change my approach was also beneficial for my own understanding of the 3Di-sva.

High expectations

We began with just one ASD recruitment site, and it was initially conceptualised that my research partner and I would be required to spend little time at the ASD clinic and instead would focus our efforts on recruiting those from the comparison groups. The rationale was that the ASD group data would, with the agreement of the clinic lead and the assistance of clinicians at the service, 'collect itself'. The 3Di-sva had already become incorporated into the standard clinic

assessment process, and therefore the only additional elements needed on top of the normal procedure was for clinicians to get signed consent, complete the TOPF, and audio-record their 3Di-sva assessment. Each of these tasks seemed, to me, to be relatively quick and straightforward. However, over a period of many months', data was collected for just two people (without audio-recordings), despite the fact that many more assessments had been completed. It became clear that expecting the clinic to collect what we needed on our behalf was not going to work, and so I began attending the service on a weekly basis to collect data myself. This approach was far more successful, as across the course of 12 weeks I was able to complete and record 13 3Di-sva interviews (of which eight were eligible for inclusion in the ASD group). Had I taken this approach from the start, the number of participants we had been able to recruit overall would have been significantly higher.

This situation highlighted to me that we had expected too much of the ASD clinic. The expectations were formed following discussions with and agreement from the head of the clinic, and as outsiders to the service it had seemed like relatively little to ask clinicians to include a few additional steps in their routine in order to collect our data. However, for clinicians in a busy service that runs only one day a week, expecting them to hold our research in mind and complete any additional tasks was not realistic. My expectation was that clinicians would be on board with the data collection as they would benefit from the 3Di-sva being validated and having diagnostic thresholds. However I came to understand that, despite the fact that I saw my research as important and interesting, this did not mean clinicians would feel the same way when their main task is to get through a waiting list of patients to meet targets set by NHS service commissioners. Furthermore, the clinic itself went through a long period of change and upheaval, with the physical location of the clinic

moving and a great deal of staff change, making our research even more of a burden for them and something even less at the forefront of their minds.

Due to the difficulties of recruiting from the first site, we added a second ASD site, which did successfully collect data itself. This clinic was headed by one of the co-supervisors of the project, who had worked at the first clinic at the point that the project was begun. This clinic was smaller and was not experiencing any upheaval. In this case, the clinic was able to successfully collect data for us without me needing to be physically present at the service. Unfortunately we only began using this site many months into the research, and the assessment process was slower as it was smaller clinic, meaning we recruited few participants from the site.

Overall I have taken from this experience that recruiting within NHS services is likely to be most successful if you are collecting data within the service yourself, and if not there ideally needs someone else within the clinic with a vested interested in the project. NHS clinics are busy environments with many targets to meet, and expecting others to take on extra tasks for the benefit of an outsider researcher is unlikely to be a feasible approach.

Benefits of greater involvement

Although my significantly increased involvement in collection of the data for the ASD group came at the price of a reduced ability to be involved in collection of the data for the clinical control group (reported on by McKenner, 2015), I appreciate now that I would have missed out on the important experience of using the 3Di-sva as a clinician within an ASD diagnostic clinic, had I not had to re-think my recruitment strategy. Completing a number of 3Di-svas with the ASD group, as well as participating in team meetings and diagnostic discussions, enabled me to take a

scientist-practitioner role in my research, gaining an important first hand understanding of using the instrument as a clinician within the population for which it is intended. Using the tool with the ASD population proved to me that for a clinician in a diagnostic clinic, it is a suitable and user-friendly tool.

I also came to realise that had I only conducted the interviews with the non-clinical and clinical control groups I would have been less aware of questions which are occasionally misinterpreted or hard to score within the ASD population. Many of the questions are less relevant and therefore more easily scored as 'behaviour not present' in the control populations, so collecting data solely from these would have meant I was less aware of areas needing scoring clarification. Gaining this understanding was essential in terms of writing a scoring manual with which to train others, which in turn is vital to ensure the high inter-rater reliability of the instrument.

Interviewing the parents of the ASD group also enabled me to form conclusions about the face validity of the instrument. After completing the 3Di-sva I had a chance to informally discuss with informants whether they felt the relevant and important areas had been covered. I was reassured to hear consistently they had been. Had I not repeatedly used the instrument with the ASD population, I could not have reached such a conclusion.

I also felt very positive from using the 3Di-sva myself within the service, and participating in team discussions around diagnoses, that the 3Di-sva was indeed a useful adjunct tool, that was generally easy to complete, and provided necessary and helpful information needed when making diagnostic decisions. This in turn has further increased my passion for continued investigation of the tool, as I feel

convinced that the clinical implications of this research genuinely are that this tool could be of great benefit to ASD diagnostic services.

Concluding remarks

The process of completing my literature review and empirical paper has been both challenging and rewarding. It has been an interesting process to begin by reviewing literature and forming conclusions about how ASD diagnostic tool research should be done, to then personally experience the difficult reality of completing such a study. I have learnt that a lot of determination and flexibility are needed to complete studies within clinical settings, but that ultimately such research is worthwhile. I have been able to experience first-hand the benefits of taking a scientist-practitioner role, combining clinical work and research. Seeing the value of the tool I was researching first-hand preserved my determination and enabled me to remain motivated in the face of the inherent challenges.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Autism Alliance (2015). Research introduction: the challenges of research. Retrieved from http://www.autism-alliance.org.uk/research/research-introduction

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The adult Asperger assessment (AAA): a diagnostic method. *Journal of Autism and Developmental Disorders, 35*(6), 807-819.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keysers, C., & de Bildt, A. (2011). Diagnosing Autism spectrum disorders in adults: the use of Autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders*, *41*(9), 1256-1266.

BMJ Clinical Evidence (2014). *Diagnostic test studies: assessment and critical appraisal.* Retrieved from: http://clinicalevidence.bmj.com.libproxy.ucl.ac.uk/x/set/static/ebm/toolbox/665061.html

Brugha, T. S., McManus, S., Smith, J., Scott, F. J., Meltzer, H., Purdon, S., Berney, T., Tantam, D., Robinson, J., Radley, J., & Bankart, J. (2012). Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. *Psychological Medicine*, *42*(03), 647-656.

Department of Health. (2010). Fulfilling and rewarding lives: The strategy for adults with autism in England. UK: Author.

Hus, V., & Lord, C. (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders,* 1-17.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

McDermott, L. M., & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of affective disorders*, *119*(1), 1-8.

McKenner, M. (2015). *Validating the Developmental, Diagnostic and Dimensional Interview - Short Form Adult Version (3Di-sva) for use in a clinical population* (Unpublished doctoral thesis). University College London, London.

Michel, N. M., Goldberg, J. O., Heinrichs, R. W., Miles, A. A., Ammari, N., & Vaz, S. M. (2013). WAIS-IV profile of cognition in schizophrenia. *Assessment, 20*(4), 462-473.

National Institute for Health and Care Excellence (2012). *Autism: recognition, referral, diagnosis and management of adults on the autism spectrum. CG142*. London: National Institute for Health and Care Excellence.

Patel, M. X., Doku, V., & Tennakoon, L. (2003). Challenges in recruitment of research participants. *Advances in Psychiatric Treatment, 9*(3), 229-238.

Ross, S., Grant, A., Counsell, C., Gillespie, W., Russell, I., & Prescott, R. (1999). Barriers to participation in randomised controlled trials: a systematic review. *Journal of clinical epidemiology, 52*(12), 1143-1156.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale–Fourth Edition (WAIS–IV).* San Antonio, TX: Pearson.

Wechsler, D. (2009). *Test of Premorbid Functioning – UK Edition.* London: Pearson Assessment.

World Health Organization. (1992). *International classification of diseases: Diagnostic criteria for research* (10th edition). Geneva, Switzerland: Author.

APPENDIX 1

BMJ Clinical Evidence (2014) tool for critically appraising diagnostic test studies.

REFERENCE STANDARD

Was there a clear question for the study to address?

[Is all the following information included in the paper?]

	YES	NO	UNCLEAR
Population			
Test			
Setting			
Outcome			

Is there comparison with an appropriate (gold) reference standard for diagnosing ASD? I.e. DSM or ICD.

YES	NO	UNCLEAR

SAMPLE [additional question]

Does the study include an adequate sample size for analysis?

[Arbitrary but defined here at least 20 ASD and 20 non-ASD]

YES	NO	UNCLEAR

POPULATION

Did the study include people with other disorders that are commonly confused with ASD?

YES [separate clinical comparison group]	NO	UNCLEAR [combined in one comparison group within non-clinical controls – score 0.5]

BLINDING

Were the people assessing the results of the index diagnostic test blinded to the results of the reference standard?

YES	NO	UNCLEAR/PARTIAL [Score 0.5 if partial blinding]

TESTING

Was the reference standard applied regardless of the index test result?

[This question is used to discriminate studies which used the index test to inform the reference standard]

YES	NO	UNCLEAR

CONFOUNDS [additional question]

Was a measure of cognitive ability used with at least part of the each group?

YES	NO	UNCLEAR

Was the diagnostic test validated in a second independent group of patients?

[Was the sample collected from more than one site?]

YES	NO	UNCLEAR

METHODS

Were the methods of the diagnostic test described in enough detail?

Rationale for the reference standard?
[Automatically given a 1 if gold standard (DSM/ICD) but scored for those with alternative reference standard that therefore needs explaining]

YES	NO	UNCLEAR

Technical specifications or references for running the index test and reference standard?
[Sufficient detail to replicate study]

YES	NO	UNCLEAR

Methods for calculating or comparing measures of diagnostic accuracy?
[Question used to show if the method of statistical analysis is described]

YES	NO	UNCLEAR

Results – what should ideally be included?

POPULATION

Are there sufficient clinical and demographic characteristics of the people in the study?
[Need to include age (mean, standard deviation and range) and gender. If one of these is missing then marked as partial and score as 0.5]

YES	NO	UNCLEAR/PARTIAL

Do the results include how indeterminate results, missing results and outliers of the index test were handled?

YES	NO	UNCLEAR

Do results include criteria for defining severity of the target disorder?

YES	NO	UNCLEAR

Do the results include cross-tabulation of the index test results by the reference standard results? Or is there enough information to generate this?

YES	NO	UNCLEAR

Do the results include estimates of diagnostic test accuracy?

YES	NO	UNCLEAR

Do the results include important psychometrics? [additional question]

Inter-rater reliability	YES	NO	UNCLEAR
Test-retest reliability	YES	NO	UNCLEAR
Internal consistency	YES	NO	UNCLEAR
Convergent validity	YES	NO	UNCLEAR
Correlation with participant characteristics	YES	NO	UNCLEAR

TOTAL SCORE:

Note. Questions in blue represent additional questions added to original tool. Remarks in grey clarify the question and how to score it where necessary.

APPENDIX 2

Scoring of each paper using the modified BMJ Critical Appraisal Tool

Modified BMJ Critical Appraisal Tool Question	Paper			
	RAADS/-R			AAA
	Anderson et al. (2011)	Ritvo et al. (2008)	Ritvo et al. (2011)	Baron-Cohen et al. (2005)
Was there a clear question for the study to address?				
Population	1	1	1	1
Test	1	1	1	1
Setting	1	1	1	1
Outcome	1	1	1	1
Is there comparison with an appropriate (gold) reference standard for diagnosing ASD? I.e. DSM or ICD?	0	1	1	1
Does the study include an adequate sample size for anaylsis?	1	1	1	0
Did the study include people with other disorders that are commonly confused with ASD?	0.5	1	1	0
Were the people assessing the results of the index diagnostic test blinded to the results of the reference standard?	0	0	0	0
Was the reference standard applied regardless of the index test result?	1	1	1	1
Was a measure of cognitive ability used with at least part of the each group?	1	0	1	0
Was the diagnostic test validated in a second independent group of patients?	1	1	1	0

Rationale for the reference standard?	0	1	1	1
Technical specifications or references for running the index test and reference standard?	1	1	1	1
Methods for calculating or comparing measures of diagnostic accuracy?	1	1	1	0
Are there sufficient clinical and demographic characteristics of the people in the study?	1	0	0.5	0.5
Do the results include how indeterminate results, missing results and outliers of the index test were handled?	1	0	0	0.5
Do results include criteria for defining severity of the target disorder?	0	0	0	0
Do the results include cross-tabulation of the index test results by the reference standard results? Or is there enough information to generate this?	0	1	1	1
Do the results include estimates of diagnostic test accuracy?	1	1	1	0
Do the results include important psychometrics?				
Inter-rater reliability	0.5	0.5	0.5	0
Test-retest reliability	1	0	1	0
Internal consistency	1	1	1	0
Convergent validity	1	0	1	0
Correlation with participant characteristics	1	1	1	0
TOTAL	18	16.5	20	10

Modified BMJ Critical Appraisal Tool Question	PAPER			
	ADOS			
	Bastiaansen et al. (2011)	**Brugha et al (2012)**	**Hus & Lord (2014)**	**Lord et al. (2000)**
Was there a clear question for the study to address?				
Population	1	1	1	1
Test	1	1	1	1
Setting	1	1	1	1
Outcome	1	1	1	1
Is there comparison with an appropriate (gold) reference standard for diagnosing ASD? I.e. DSM or ICD?	1	0	1	0
Does the study include an adequate sample size for analysis?	1	0.5*	1	0
Did the study include people with other disorders that are commonly confused with ASD?	1	0	1	0.5
Were the people assessing the results of the index diagnostic test blinded to the results of the reference standard?	0.5	0	0	1
Was the reference standard applied regardless of the index test result?	1	0	0	1
Was a measure of cognitive ability used with at least part of the each group?	1	0	1	1
Was the diagnostic test validated in a second independent group of patients?	1	0	1	1

Rationale for the reference standard?	1	1	1	0
Technical specifications or references for running the index test and reference standard?	1	1	1	1
Methods for calculating or comparing measures of diagnostic accuracy?	1	1	1	1
Are there sufficient clinical and demographic characteristics of the people in the study?	1	0	1	1
Do the results include how indeterminate results, missing results and outliers of the index test were handled?	1	0	1	1
Do results include criteria for defining severity of the target disorder?	0	1	0	1
Do the results include cross-tabulation of the index test results by the reference standard results? Or is there enough information to generate this?	0	1	0	1
Do the results include estimates of diagnostic test accuracy?	1	1	1	1
Do the results include important psychometrics?				
Inter-rater reliability	1	0.5	1	1
Test-retest reliability	0	0	0	0
Internal consistency	1	0	1	0
Convergent validity	0	1	0	0
Correlation with participant characteristics	1	0	1	1
TOTAL	19.5	12	18	17.5

* N.B. 618 participants completed ADOS module 4. However sensitivity and specificity analyses were conducted on subset of n=56 and >20 received an ASD diagnosis.

Modified BMJ Critical Appraisal Tool Question	PAPER			
	ASDI	**ADI-R**	**ADOS/ADI-R (LD)**	**DISCO/ADI-R**
	Gillberg et al. (2001)	**Lord et al. (1997)**	**Sappok et al. (2013)**	**Nygren et al. (2009)**
Was there a clear question for the study to address?				
Population	1	1	1	1
Test	1	1	1	1
Setting	1	1	1	1
Outcome	1	1	1	1
Is there comparison with an appropriate (gold) reference standard for diagnosing ASD? I.e. DSM or ICD?	1	0	1	1
Does the study include an adequate sample size for analysis?	0	1	1 (ADOS) 0 (ADI-R)	0
Did the study include people with other disorders that are commonly confused with ASD?	0.5	0.5	1	0.5
Were the people assessing the results of the index diagnostic test blinded to the results of the reference standard?	1	0.5	1	1
Was the reference standard applied regardless of the index test result?	1	1	1	1
Was a measure of cognitive ability used with at least part of the each group?	1	1	0	1
Was the diagnostic test validated in a second independent group of patients?	0	1	0	0

Rationale for the reference standard?	1	0	1	1
Technical specifications or references for running the index test and reference standard?	1	1	1	1
Methods for calculating or comparing measures of diagnostic accuracy?	0	1	1	1
Are there sufficient clinical and demographic characteristics of the people in the study?	0	0.5	0.5	1
Do the results include how indeterminate results, missing results and outliers of the index test were handled?	0	1	1	0
Do results include criteria for defining severity of the target disorder?	0	0	1	1
Do the results include cross-tabulation of the index test results by the reference standard results? Or is there enough information to generate this?	1	0	0	1
Do the results include estimates of diagnostic test accuracy?	0	1	1	0
Do the results include important psychometrics?				
Inter-rater reliability	1	0.5	0	0
Test-retest reliability	1	0	0	0
Internal consistency	0	0	1	0
Convergent validity	0	0	1	1
Correlation with participant characteristics	0	0	1	0
TOTAL	13.5	14	18.5 (ADOS) 17.5 (ADI-R)	15.5

APPENDIX 3

Contributions to joint project

Aspects of study completed jointly

- All planning of study methodology.
- Liaison with the ASD clinics from which recruitment took place.
- Writing the project protocol.
- Writing the NHS REC application.
- Attendance at the NHS REC panel interview.
- Recruitment of the non-clinical control group (*n=22* completed by McKenner, *n*=5 completed by Clarke).

Aspects of study completed by Kiri Clarke

- All liaison with the IAPT services from which participants were recruited to the clinical control group.
- Writing of NHS R&D applications relevant to the IAPT service and ASD services.
- Recruitment and interviewing of IAPT participants for the clinical control group (*n*=8 recruited, *n*=7 interviewed).
- Recruitment and interviewing of cases from ASD clinic (*n*=13).
- Analysis and write-up of data for ASD group versus non-clinical control group.

Aspects of study completed by Michele McKenner

- All liaisons with the psychosis services and other non-IAPT services from which participants were recruited to the clinical control group.
- Writing of the NHS R&D application relevant to the above services
- Writing of a substantial amendment relevant to recruitment of psychosis participants.
- Recruitment and interviewing of non-IAPT participants included in clinical control group (*n*=8), plus interviewing of one IAPT participant.
- Analysis and write-up of data for ASD group versus clinical control group.

APPENDIX 4

Letter of approval from National Research Ethics Service Committee

NHS
Health Research Authority

National Research Ethics Service

NRES Committee London - Bloomsbury
HRA NRES Centre Manchester
Barlow House 3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7815
Fax: 0161 625 7299

16 July 2014

Dr Will Mandy
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London
WC1E 7HB

Dear Dr Mandy

Study title: **Validating the 3DI-sva: a short form diagnostic interview for autism spectrum disorders in adults**
REC reference: **14/LO/1134**
IRAS project ID: **145776**

Thank you for your email of 16 July 2014. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 09 July 2014.

Documents received

The documents received were as follows:

Document	Version	Date
Other [Schedule of Information Sheets, Consent Forms and Advertising Documents]		
Participant consent form [Parent - Non-clinical Controls]	V6.2	16 July 2014
Participant consent form [Parent - Affective Control Group]	V7.2	16 July 2014
Participant consent form [Parent - Psychosis Control Group]	V8.2	16 July 2014
Participant consent form [Parent - Pre 3DI]	V9.2	16 July 2014
Participant consent form [Parent - Post 3DI]	V10.2	16 July 2014
Participant consent form [Child - Affective Control Group]	V2.2	16 July 2014
Participant consent form [Child Psychosis Control Group]	V3.2	16 July 2014
Participant consent form [Child - Pre 3DI]	V4.2	16 July 2014
Participant consent form [Child - Post 3DI]	V5.2	16 July 2014
Participant consent form [Child - Non-clinical Controls]	V1.2	16 July 2014
Participant information sheet (PIS) [Parent - Non-clinical Controls]	V6.2	16 July 2014

Participant Information sheet (PIS) [Parent - Affective Control Group]	V7.2	16 July 2014
Participant Information sheet (PIS) [Parent - Psychosis Control Group]	V8.2	16 July 2014
Participant Information sheet (PIS) [Parent - Pre 3DI]	V9.2	16 July 2014
Participant Information sheet (PIS) [Parent - Post 3DI]	V10.2	16 July 2014
Participant Information sheet (PIS) [Child - Non-clinical Controls]	V1.2	16 July 2014
Participant Information sheet (PIS) [Child - Affective Control Group]	V2.2	16 July 2014
Participant Information sheet (PIS) [Child Psychosis Control Group]	V3.2	16 July 2014
Participant Information sheet (PIS) [Child - Pre 3DI]	V4.2	16 July 2014
Participant Information sheet (PIS) [Child - Post 3DI]	V5.2	16 July 2014

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Affective control group flyer]	V1.1	12 March 2014
Copies of advertisement materials for research participants [Psychosis control group flyer]	V2.1	12 March 2014
Copies of advertisement materials for research participants [Psychosis control group poster]	V2.1	12 March 2014
Copies of advertisement materials for research participants [Affective control group poster]	V1.1	12 March 2014
Copies of advertisement materials for research participants [Non-clinical control group poster]	V3.1	12 March 2014
Evidence of Sponsor Insurance or Indemnity (non NHS Sponsors only)	Arthur J Gallagher	26 July 2013
Letter from funder [Trainee Indemnity & employment arrangements letter]		06 November 2009
Letter from sponsor	UCL	23 May 2014
Letters of invitation to participant [Invitation letter]	V1.1	26 March 2014
Non-validated questionnaire [3DI-sva]	V1	23 May 2014
Other	Michele McKenner CV	23 May 2014
Other [Schedule of Information Sheets, Consent Forms and Advertising Documents]		
Participant consent form [Child Psychosis Control Group]	V3.2	16 July 2014
Participant consent form [Parent - Post 3DI]	V10.2	16 July 2014
Participant consent form [Child - Pre 3DI]	V4.2	16 July 2014
Participant consent form [Child - Non-clinical Controls]	V1.2	16 July 2014
Participant consent form [Parent - Affective Control Group]	V7.2	16 July 2014
Participant consent form [Child - Post 3DI]	V5.2	16 July 2014
Participant consent form [Parent - Non-clinical Controls]	V6.2	16 July 2014
Participant consent form [Parent - Psychosis Control Group]	V8.2	16 July 2014
Participant consent form [Child - Affective Control Group]	V2.2	16 July 2014
Participant consent form [Parent - Pre 3DI]	V9.2	16 July 2014
Participant Information sheet (PIS) [Child - Non-clinical Controls]	V1.2	16 July 2014
Participant Information sheet (PIS) [Parent - Psychosis Control Group]	V8.2	16 July 2014
Participant Information sheet (PIS) [Child - Affective Control Group]	V2.2	16 July 2014

Participant Information sheet (PIS) [Parent - Pre 3DI]	V9.2	16 July 2014
Participant Information sheet (PIS) [Parent - Post 3DI]	V10.2	16 July 2014
Participant Information sheet (PIS) [Child - Pre 3DI]	V4.2	16 July 2014
Participant Information sheet (PIS) [Child Psychosis Control Group]	V3.2	16 July 2014
Participant Information sheet (PIS) [Child - Post 3DI]	V5.2	16 July 2014
Participant Information sheet (PIS) [Parent - Non-clinical Controls]	V6.2	16 July 2014
Participant Information sheet (PIS) [Parent - Affective Control Group]	V7.2	16 July 2014
REC Application Form	3.5	22 May 2014
Referee's report or other scientific critique report [Oliver Mason peer review]		03 October 2013
Research protocol or project proposal	V1	26 March 2014
Summary CV for Chief Investigator (CI)	Will Mandy	
Summary CV for student	Kirl Clarke	
Summary, synopsis or diagram (flowchart) of protocol in non technical language [Flow diagram]	V1	13 March 2014

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

14/LO/1134	Please quote this number on all correspondence

Yours sincerely

[redacted]

Dr Ashley Totenhofer
REC Manager

E-mail: nrescommittee.london-bloomsbury@nhs.net

Copy to: Dr Clara Kalu - University College London

Mrs Angela Williams – Noclor

Ms Michele McKenner - University College London

Ms Kirl Clarke - University College London

APPENDIX 5

Sample information sheets

Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

Participant Information Sheet

This sheet is for you to keep and tells you more about this study. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and do not hesitate to ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Autism Spectrum Disorder (ASD) is a lifelong developmental disability which affects how a person communicates with and relates to other people. It also affects how they make sense of the world around them. We are researching ways of telling who has ASD and who does not have ASD.

We are not asking you to take part because we believe you might have ASD.

This study is investigating whether an interview called the 3DI can tell the difference between people with ASD and people who do not have ASD. The 3DI is a new tool for assessing ASD in adults. It involves interviewing the parents of the person with suspected ASD. The questions are about the person's childhood development, as well as any current difficulties with social interaction, communication and flexibility. There is already evidence that the 3DI is effective, this study aims to investigate in greater detail how useful and effective it is.

Why is this important?

Some people with ASD reach adulthood without being diagnosed. Getting a diagnosis enables individuals to access support from health and social services, and helps others understand the individual's needs and behaviours. Currently there is a lack of effective, valid tools for diagnosing ASD in adults. Your participation in this research helps us validate a new diagnostic interview to assess adults for ASD.

Why am I being asked to participate?

We are recruiting a group of people who do not have ASD or any mental health difficulties to see if the 3DI can distinguish this group from a group of people with ASD. This is important to ensure the 3DI is an effective diagnostic tool. We are **not** recruiting you because we think you have ASD.

What would I need to do?

We will ask you to complete one brief task which will involve reading out a list of words. This should take around 5 minutes and can be done in person or over the phone.

We would also like your parent to take part in the research. We would like you to get their permission for us to contact them. If it is not possible to speak to your parent we could talk to

someone else who knew you well as a child, like a guardian or another relative. If they agree to take part, we would interview them over the phone, asking questions about your development as a child. All information will be kept strictly confidential, as explained in more detail below.

The reason we need to interview someone who has known you since you were a child is that the symptoms of ASD first appear in early childhood and so we need to find out about your childhood development from someone who knew you when you were growing up.

Is there anything that would stop me taking part?

We will not be able to include you in this study if you have ever had any concerns that you yourself may have ASD. This is because we are trying to see whether the 3DI can tell the difference between people who are known to have ASD and people who do not. If you are worried you might have ASD and wish to find out more then please see your GP. We will also not be able to include you if you do not have a parent or relative who can also take part because, as explained above, the 3DI relies on asking someone who knows you well questions about your early childhood. Unfortunately, we will not be able to include you if have difficulty understanding or speaking English, or if your parent has difficulties with English.

If we are worried that you do not understand what the study involves, we will not include you in the study. However if, after you have taken part, something happens which means you no longer understand what the research involves, we may still use the information we have collected about you. All personal information will have been removed from the data and it will not be possible to identify you.

Do I have to take part?

It is entirely up to you and your parent to decide whether or not to take part in this study. If you decide to take part, you are free to withdraw at any time, without giving a reason.

Are there incentives for taking part?

To thank you for your time, and for introducing us to your parent, once we have received all the relevant information you will receive a £10 Amazon voucher. You will also be entered into a prize draw to win a £50 Amazon voucher when the study is complete - you will have approximately a 1 in 60 chance of winning this voucher. Your parent will also be entitled to a £10 Amazon voucher once they have taken part, and will be entered into a separate prize draw to win a £50 Amazon voucher.

What are the possible disadvantages and risks of taking part?

We do not think there will be any disadvantages for you if you take part in the study. If you want to stop your participation for any reason, you are free to do so immediately.

What happens if the interview suggests I might have ASD?

We will ask you, before you take part, whether you want us to let you know if the interview suggests you might have ASD. Even if the interview does suggest this, there would need to be further investigation to confirm any diagnosis. If the interview does suggest you may have ASD and you have asked us to let you know, we will contact you to discuss this and any next steps that you may wish to take.

Will my taking part in this study be kept confidential?

All data will be collected and stored in accordance with the Data Protection Act 1998. Only the researchers involved in the study will have access to data you or your parent provide(s) as part of this study. This information will be kept either in locked cabinets or on a secure and encrypted computer account, and will be stored anonymously at University College London (UCL). Any information you or your parent has provided will be kept strictly confidential. The only exception to this is if you mention something during the study which indicates that you may be a risk to yourself or others. In this case we would need to let someone relevant know, eg your GP. If this happens we will tell you.

What will happen to the results of the research?

The information will be analysed and compared with similar information provided by groups of people who have ASD, or who have other mental health difficulties, or who don't have ASD or mental health difficulties. These findings may be written up as reports, initially as part of a doctoral thesis at UCL. However, names and other identifying information will be removed. The results of the study may be presented at national and international conferences and published in academic journals. You will not be personally identified in any of these reports or publications.

Who is organising and funding this research?

The project is funded by UCL. It is being conducted by the Research Department of Clinical, Educational and Health Psychology at UCL.

Contact for further information or assistance

If you have any further questions or would like assistance at any point during the study, please contact **Michele McKenner** or **Kiri Clarke** at UCL on 07989 085 846 or email adult.autism.study@gmail.com. In the case of a complaint, please contact Dr Will Mandy via w.mandy@ucl.ac.uk.

Thank you for reading this information

Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

Participant Parent Information Sheet

This sheet is for you to keep and tells you more about this study. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and do not hesitate to ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Autism Spectrum Disorder (ASD) is a lifelong developmental disability which affects how a person communicates with and relates to other people. . It also affects how they make sense of the world around them. We are researching ways of telling who has ASD and who does not have ASD.

We are looking for parents of people WITHOUT ASD to take part.

This study is investigating whether an interview called the 3DI can tell the difference between people with ASD and people who do not have ASD. The 3DI is a new tool for assessing ASD in adults. It involves interviewing the parents of the person with suspected ASD. The questions are about the person's childhood development, as well as any current difficulties with social interaction, communication and flexibility. There is already evidence that the 3DI is effective, this study aims to investigate in greater detail how useful and effective it is.

Why is this important?

Some people with ASD reach adulthood without being diagnosed. Getting a diagnosis enables individuals to access support from health and social services, and helps others understand the individual's needs and behaviours. Currently there is a lack of effective, valid tools for diagnosing ASD in adults. Your participation in this research helps us validate a new diagnostic interview to assess adults for ASD.

Why am I being asked to participate?

We are recruiting the parents of a group of people who do not have ASD or any mental health difficulties to see if the 3DI can distinguish this group from a group of people with ASD. This is important to ensure the 3DI is an effective diagnostic tool. We are **not** asking you to take part because we think your child might have an ASD.

What would I need to do?

Once we have your consent and your child's consent, we will contact you and carry out the 3DI interview with you over the phone, asking questions about your child's development. We

would like to audio record the interview. It should take no more than 45-60 minutes. We will also meet with your child on one occasion for around 15 minutes. All information will be kept strictly confidential, as explained in more detail below.

Is there anything that would stop me taking part?

We will not be able to include you / your child in this study if there has ever been any concern that they may have ASD. This is because we are trying to see whether the 3DI can tell the difference between a group of people who are known to have an ASD and groups of people who do not. If you do have concerns about your child having an ASD and wish to find out more then please discuss this with your GP. In addition, unfortunately, we will not be able to include you if either you or your child has difficulty understanding or speaking English.

If we are worried that you or your child does not understand what the study involves, we will not include you in the study. However if, after you have taken part, something happens which means you no longer understand what the research involves, we may still use the information we have collected about you. All personal information will have been removed from the data and it will not be possible to identify you.

Do I have to take part?

It is entirely up to you and your child to decide whether or not to take part in this study. If you decide to take part, you are free to withdraw at any time, without giving a reason.

Are there incentives for taking part?

To thank you for your time, once we have received all the relevant information you will receive a £10 Amazon voucher. You will also be entered into a prize draw to win a £50 Amazon voucher when the study is complete - you will have approximately a 1 in 60 chance of winning this voucher. Your child will also be entitled to a £10 Amazon voucher once they have taken part, and will be entered into a separate prize draw to win a £50 Amazon voucher.

What are the possible disadvantages and risks of taking part?

We do not think there will be any disadvantages for you if you take part in the study. If you want to stop your participation for any reason, you are free to do so immediately.

Will my taking part in this study be kept confidential?

All data will be collected and stored in accordance with the Data Protection Act 1998. Only the researchers involved in the study will have access to data you or your child provide(s) as part of this study. This information will be kept either in locked cabinets or on a secure and encrypted computer account, and will be stored anonymously at University College London (UCL). Any information you or your child has provided will be kept strictly confidential. The only exception to this is if you mention something during the study which indicates that you may be a risk to yourself or others. In this case we would need to let someone relevant know, eg your GP. If this happens we will tell you.

What will happen to the results of the research?

The information will be analysed and compared with similar information provided by groups of people who have ASD, or who have other mental health difficulties, or who don't have ASD or mental health difficulties. These findings may be written up as reports, initially as part of a

doctoral thesis at UCL. The results of the study may be presented at national and international conferences and published in academic journals. You will not be personally identified in any of these reports or publications.

Who is organising and funding this research?

The project is funded by UCL. It is being conducted by the Research Department of Clinical, Educational and Health Psychology at UCL.

Contact for further information or assistance

If you have any further questions or would like assistance at any point during the study, please contact **Michele McKenner** or **Kiri Clarke** at UCL on 07989 085 846 or email adult.autism.study@gmail.com. In the case of a complaint, please contact Dr Will Mandy via w.mandy@ucl.ac.uk.

Thank you for reading this information

Camden and Islington NHS
NHS Foundation Trust

UCL

Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

Participant Information Sheet

This sheet is for you to keep and tells you more about this study. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and do not hesitate to ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Some people with Autism Spectrum Disorders (ASD) reach adulthood without ever having had their condition recognised. This can prevent them from getting much needed support. We need more effective tools to help diagnose adults with suspected ASD, including tools which look at early childhood development and difficulties.

A relatively new tool for assessing ASD in adults is being used in the Asperger's Syndrome Diagnostic and Consultation Service in Camden which you are attending (the Service). This tool is called the Developmental, Diagnostic and Dimensional Interview (or 3DI).

The 3DI involves interviewing a parent or relative of the person with suspected ASD to find out about their childhood development, as well as any current difficulties with social interaction and communication. There is already evidence that the interview is effective, but this study aims to investigate in greater detail how useful and effective it is.

What would I need to do?

The 3DI interview will be carried out with your parent(s) or relative as part of the routine assessment process with the Service. We would like your permission to use the information provided by your parent / relative in order to assess how effective the 3DI is. For example, we want to check that the 3DI is able to tell the difference between people who have ASD and those who do not.

Do I have to take part?

It is entirely up to you and your parent to decide whether or not to take part in this study. If you decide to take part, you are free to withdraw at any time, without giving a reason. **Please note that the clinical care you receive will not be affected in any way – whether or not you decide to take part in the study.**

What does the study involve?

The study will involve us accessing and using information provided by you and your parent as part of your routine assessment by the Service. This will include information given by your parent in the 3DI interview and information from some tasks and measures that you complete. All information will be kept strictly confidential, as explained in more detail below.

Is there anything that would stop me taking part?

If we are worried that you do not understand what the study involves, we will not include you in the study. However if, after you have taken part, something happens which means you no longer understand what the research involves, we may still use the information we have collected about you. All personal information will have been removed from the data and it will not be possible to identify you.

What are the possible disadvantages and risks of taking part?

We do not think there will be any disadvantages for you if you take part in the study. However, if you want to stop your participation for any reason, you are free to do so immediately.

Will my taking part in this study be kept confidential?

All data will be collected and stored in accordance with the Data Protection Act 1998. Only the researchers involved in the study will have access to data you or your parent provide(s) as part of this study. This information will be kept either in locked cabinets or on a secure and encrypted computer account, and will be stored anonymously at University College London (UCL). Any information you or your parent has provided will be kept strictly confidential. The only exception to this is if you mention something during the study which indicates that you may be a risk to yourself or others. In this case we would need to let someone relevant know, eg your GP. If this happens we will tell you.

What will happen to the results of the research?

The information will be analysed and compared with similar information provided by different groups of people who don't have ASD. These findings may be written up as reports, initially as part of a doctoral thesis at UCL. The results of the study may be presented at national and international conferences and published in academic journals. You will not be personally identified in any reports or publications of the research.

Who is organising and funding this research?

The project is funded by UCL. It is being conducted by the Research Department of Clinical, Educational and Health Psychology at UCL.

Contact for further information or assistance

If you have any further questions or would like assistance at any point during the study, please feel free to contact **Michele McKenner** or **Kiri Clarke** at UCL on 07989 085 846 or email adult.autism.study@gmail.com. In the case of a complaint, please contact Dr Will Mandy via w.mandy@ucl.ac.uk.

Thank you for reading this information

Camden and Islington NHS
NHS Foundation Trust

UCL

Validating the 3Di-sva: diagnostic interview for autism spectrum disorders in adults

Participant Parent Information Sheet

This sheet is for you to keep and tells you more about this study. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and do not hesitate to ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Some people with Autism Spectrum Disorders (ASD) reach adulthood without ever having had their condition recognised. This can prevent them from getting much needed support. We need more effective tools to help diagnose adults with suspected ASD, including tools which look at early childhood development and difficulties.

A relatively new tool for assessing ASD in adults is being used in the Asperger's Syndrome Diagnostic and Consultation Service in Camden which your child is attending (the Service). This tool is called the Developmental, Diagnostic and Dimensional interview (or 3DI).

The 3DI involves interviewing the parents of the person with suspected ASD to find out about their childhood development, as well as any current difficulties with social interaction and communication. There is already evidence that this tool is effective, this study aims to investigate in greater detail how useful and effective it is.

What would I need to do?

The 3DI interview will be carried out with you as part of the routine assessment process with the Service. We would like your permission to audio record this interview and to use the information provided by you in order to assess how effective the 3DI is. For example, we want to check that the 3DI is able to discriminate between people who have ASD and those who do not.

Do I have to take part?

It is entirely up to you and your child to decide whether or not to take part in this study. If you decide to take part, you are free to withdraw at any time, without giving a reason. **Please note that the clinical care your child receives will not be affected in any way – whether or not you decide to take part in the study.**

What does the study involve?

The study will involve us accessing and using information provided by you and your child as part of your child's routine assessment by the Service. This will include information given by you in the 3DI interview and information from some tasks and measures that your child completes. All information will be kept strictly confidential, as explained in more detail below.

Is there anything that would stop me taking part?

If we are worried that you or your child does not understand what the study involves, we will not include you in the study. However if, after you have taken part, something happens which means you no longer understand what the research involves, we may still use the information we have collected about you. All personal information will have been removed from the data and it will not be possible to identify you.

What are the possible disadvantages and risks of taking part?

We do not think there will be any disadvantages for you if you take part in the study. However, if you want to stop your participation for any reason, you are free to do so immediately.

Will my taking part in this study be kept confidential?

All data will be collected and stored in accordance with the Data Protection Act 1998. Only the researchers involved in the study will have access to data you or your child provide(s) as part of this study. This information will be kept either in locked cabinets or on a secure and encrypted computer account, and will be stored anonymously at University College London (UCL). Any information you or your child has provided will be kept strictly confidential. The only exception to this is if you mention something during the study which indicates that you may be a risk to yourself or others. In this case we would need to let someone relevant know, eg your GP. If this happens we will tell you.

What will happen to the results of the research?

The information will be analysed and compared with similar information provided by different groups of people who don't have ASD. These findings may be written up as reports, initially as part of a doctoral thesis at UCL. The results of the study may be presented at national and international conferences and published in academic journals. Neither you nor your child will personally identified in any reports or publications of the research.

Who is organising and funding this research?

The project is funded by UCL. It is being conducted by the Research Department of Clinical, Educational and Health Psychology at UCL.

Contact for further information or assistance

If you have any further questions or would like assistance at any point during the study, please feel free to contact **Michele McKenner** or **Kiri Clarke** at UCL on 07989 085 846 or email adult.autism.study@gmail.com. In the case of a complaint, please contact Dr Will Mandy via w.mandy@ucl.ac.uk.

Thank you for reading this information

APPENDIX 6

Sample consent forms

CONSENT FORM (Version 1.2)

Title of Project: Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

	Yes/No*
I confirm that I have read and understand the information sheet dated 16/07/2014 (version 1.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	YES / NO
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected.	YES / NO
I understand that all personal data relating to participants is anonymised and held and processed in the strictest confidence.	YES / NO
I understand that the researchers will contact my parents using the contact details provided by me and will interview them about me and my development.	YES / NO
I understand that relevant sections of my data collected during the study may be looked at by individuals from University College London, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	YES / NO
I agree to take part in this study.	YES / NO

* delete as appropriate

I **would / would not** * like the researchers to contact me if the 3Di interview suggests that I might have ASD

Participant

....

First name Last name Date Signature

Person taking consent

....

Name Date Signature

(1 copy for participant; 1 copy for researcher)

Names of researchers:
Dr Will Mandy, University College London
Michèle McKenner, University College London
Kiri Clarke, University College London

CONSENT FORM (Version 6.2)

Title of Project: Validating the 3Di-eva: a diagnostic interview for autism spectrum disorders in adults

	Yes/No*
I confirm that I have read and understand the information sheet dated 16/07/2014 (version 6.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	YES / NO
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected.	YES / NO
I understand that all personal information I give regarding my child is anonymised and held and processed in the strictest confidence.	YES / NO
I understand that relevant sections of my data collected during the study may be looked at by individuals from University College London, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	YES / NO
I agree for my interview to audio recorded.	YES / NO
I agree to take part in this study.	YES / NO

* delete as appropriate

Participant

....

First name Last name Date Signature

Person taking consent

....

Name Date Signature

(1 copy for participant; 1 copy for researcher)

Names of researchers:
Dr Will Mandy, University College London
Michele McKenner, University College London
Kiri Clarke, University College London

Camden and Islington NHS
NHS Foundation Trust

UCL

CONSENT FORM (Version 4.2)

Title of Project: Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

	Yes/No*
I confirm that I have read and understand the information sheet dated 16/07/2014 (version 4.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	YES / NO
I understand that agreeing for my data to be used for this research is voluntary and that I am free to withdraw it at any time, without giving any reason, and without my health care or legal rights being affected.	YES / NO
I understand that all personal data relating to participants is anonymised and held and processed in the strictest confidence when used for the purposes of this study.	YES / NO
I understand that researchers will not have access to any of my NHS records other than the data from my clinic assessment required for this study.	YES / NO
I understand that relevant sections of my data collected during the study may be looked at by individuals from University College London, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	YES / NO
I agree for data from my clinic assessment to be included in this study.	YES / NO

* delete as appropriate

Participant

....

First name Last name Date Signature

Person taking consent

....

Name Date Signature

(1 copy for participant; 1 copy for researcher)

Names of researchers:
Dr Will Mandy, University College London
Dr Jason Crabtree, University College London
Michèle McKenner, University College London
Kiri Clarke, University College London

Camden and Islington NHS
NHS Foundation Trust

UCL

CONSENT FORM (Version 9.2)

Title of Project: Validating the 3Di-sva: a diagnostic interview for autism spectrum disorders in adults

	Yes/No*
I confirm that I have read and understand the information sheet dated 16/02/2014 (version 9.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	YES / NO
I understand that agreeing for the data I give about my child to be used for this research is voluntary and that I am free to withdraw it at any time, without giving any reason, without my legal rights being affected or my child's health care being affected.	YES / NO
I understand that all personal information I give regarding my child is anonymised and held and processed in the strictest confidence when used for the purposes of this study.	YES / NO
I understand that researchers will not have access to any of my child's NHS records other than the data from my clinic assessment required for this study.	YES / NO
I understand that relevant sections of my data collected during the study may be looked at by individuals from University College London, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	YES / NO
I agree for my interview to audio recorded.	YES / NO

* delete as appropriate

Participant

....

First name Last name Date Signature

Person taking consent

....

Name Date Signature

(1 copy for participant; 1 copy for researcher)

Names of researchers:
Dr Will Mandy, University College London
Dr Jason Crabtree, University College London
Michèle McKenner, University College London
Kiri Clarke, University College London

APPENDIX 7

Invitation letter for historical ASD group cases

Camden and Islington NHS
NHS Foundation Trust

UCL

Dear xxxxxx

Re: Research project into diagnostic interview for Autism Spectrum Disorders (ASD)

When you visited the Asperger's Syndrome Diagnostic and Consultation Service (ASDCS), you indicated that you were happy for anonymised information about you to be stored on the confidential ASDCS database, solely for the purpose of audits and research. You also agreed to be added to our Autism and Asperger Research Register, a register of individuals who are willing to be contacted about opportunities to participate in new research projects into Asperger's Syndrome and other Autism Spectrum Conditions.

New research project
We are contacting you about a new research project which is investigating an interview used for diagnosing ASD in adults. The interview is called the 3DI (the Developmental, Diagnostic and Dimensional Interview). The 3DI helps us gather information about a person's childhood development as well as current difficulties with social interaction and communication.

When you visited the Clinic, a family member, or someone else who knows you well, completed the 3DI as part of your assessment. This data has been stored, in an anonymised form, on the ASDCS database.

There is already evidence that the 3DI is effective. Our research aims to investigate in greater detail how useful and effective it is. We want to show how well the interview can distinguish between people who have ASD and people who don't. This is really important for improving diagnostic services for other adults with suspected ASD.

You already have my 3DI, so what do you want me to do now?
In order to be able to use your data in our research project we would need to collect a bit more information from you. We are writing to ask if you would be willing to complete an additional brief task – which simply involves reading out a list of words – to help us with our research.

This information will help us ensure that the abilities of the people we use from the Asperger's Clinic are similar to those of other people taking part who do not have ASD.

The task should take no more than 5-10 minutes. It could be done by meeting in person on one occasion or, if you have access to an email account on a computer, it could be done over the phone.

Names of researchers:
Dr Will Mandy & Dr Jason Crabtree, University College London
Michèle McKenner & Kirt Clarke, University College London

Why should I help?

Some people with Autism Spectrum Disorders reach adulthood without ever having had their condition recognised. This can prevent them from getting much needed support.

Whilst there is a lot of research looking at how to diagnose children with suspected ASD, we need more research showing us the best ways to diagnose adults.

By taking 10 minutes to provide us with this extra bit of information you will help us to develop user-friendly and effective ways of diagnosing ASD in adults.

What should I do now?

- Please contact us on 07969 085 846 (you may call or text) or email us at adult.autism.study@gmail.com to let us know if you are interested in taking part.
- *Please also let us know if you are not interested in participating so that we can remove you from our database.*

Registering your interest does not mean you have agreed to provide the additional data requested. You will be given further, more detailed information about the study before consenting to take part.

Please note that if we do not hear from you regarding whether you would like to take part or not within two weeks of sending this letter, we will attempt to contact you by phone to follow up this invite. If we cannot get through to you after two attempts at calling we will not try again and will assume you do not wish to take part.

Yours sincerely,

[redacted]

Kiri Clarke & Michele McKenner
Trainee Clinical Psychologists
University College London

In collaboration with Dr Andre Strydom and Dr Bano Hassan, Asperger's Syndrome Diagnostic and Consultation Service.