

Screening for Co-occurring Conditions in Adults With Autism Spectrum Disorder Using the Strengths and Difficulties Questionnaire: A Pilot Study

James Findon,* Tim Cadman,* Catherine S. Stewart, Emma Woodhouse, Hanna Eklund, Hannah Hayward, Daniel De Le Harpe Golden, Eddie Chaplin, Karen Glaser, Emily Simonoff, Declan Murphy, Patrick F. Bolton, and Fiona S. McEwen

Adolescents and adults with autism spectrum disorder (ASD) are at elevated risk of co-occurring mental health problems. These are often undiagnosed, can cause significant impairment, and place a very high burden on family and carers. Detecting co-occurring disorders is extremely important. However, there is no validated screening tool for this purpose. The aim of this pilot study is to test the utility of the strengths and difficulties questionnaire (SDQ) to screen for co-occurring emotional disorders and hyperactivity in adolescents and adults with ASD. The SDQ was completed by 126 parents and 98 individuals with ASD (in 79 cases both parent and self-report were available from the same families). Inter-rater reliability, test-retest stability, internal consistency, and construct validity were examined. SDQ subscales were also compared to clinically utilized measures of emotional disorders and hyperactivity to establish the ability to predict risk of disorder. Inter-rater reliability ($r = 0.42$), test-retest stability ($r = 0.64$), internal consistency ($\alpha = 0.52\text{--}0.81$) and construct validity ($r = 0.42\text{--}0.57$) for the SDQ subscales were comparable to general population samples. Parent- and self-report SDQ subscales were significantly associated with measures of anxiety, depression and hyperactivity (62–74% correctly classified). Parent-report performed significantly better than self-report; adults with ASD under-reported difficulties. The SDQ shows promise as a simple and efficient way to screen for emotional disorders and hyperactivity in adolescents and adults with ASD that could help reduce the impact of these disorders on individuals and their families. However, further more systematic attempts at validation are warranted. *Autism Res* 2016, 9: 1353–1363. © 2016 International Society for Autism Research, Wiley Periodicals, Inc.

Keywords: autism spectrum disorder; co-occurring disorder; co-morbid disorder; hyperactivity; attention deficit hyperactivity disorder; anxiety; depression; obsessive–compulsive disorder

Introduction

Autism spectrum disorder (ASD) is a highly heritable neurodevelopmental disorder characterized by impairment in social communication and repetitive, stereotyped behaviors [World Health Organization, 1993]. ASD is usually a life-long disorder with prevalence rates among adults estimated at 1 in 100 [Brugha et al., 2011]. Rates of psychiatric comorbidity are particularly high among individuals with ASD. The most prevalent comorbidities are mood disorders, anxiety disorders, attention-deficit hyperactivity disorder (ADHD) and

obsessive–compulsive disorder (OCD), with prevalence rates of approximately 53, 50, 43, and 24%, respectively [Buck et al., 2014; Hofvander et al., 2009; Vannucchi et al., 2014]. Psychiatric comorbidity in ASD causes significant functional impairment to the individual, resulting in greater contact with services [Leyfer et al., 2006], and a level of burden comparable to that reported by caregivers of persons with acquired brain injury [Cadman et al., 2012]. Co-occurring conditions in ASD are potentially treatable [Russell et al., 2013] and recognizing them should be a priority for community and specialist ASD services. However, comorbidity often

From the King's College London, Institute of Psychiatry, Psychology & Neuroscience, London, UK; (J.F., T.C., C.S.S., E.W., H.E., H.H., D.D.L.H.G., E.C., K.G., E.S., D.M., P.F.B., F.S.M.) and South London & Maudsley NHS Foundation Trust, London, UK (C.S.S., D.D.L.H.G., E.S., P.F.B.), and King's College London, Sackler Institute for Translational Neurodevelopment, Institute of Psychiatry, Psychology & Neuroscience (D.M.).

Received June 01, 2015; accepted for publication February 24, 2016

Address for correspondence and reprints: Fiona S. McEwen, PO80, MRC Social, Genetic and Developmental Psychiatry Centre and Department of Child & Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, 16 De Crespigny Park, London SE5 8AF, UK. E-mail: fiona.s.mcewen@kcl.ac.uk

This article was published online on 28 March 2016. An error was subsequently identified. This notice is included in the online and print versions to indicate that both have been corrected 27 April 2016.

*Denotes joint first authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Published online 28 March 2016 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.1625

Copyright © 2016 The Authors Autism Research published by Wiley Periodicals, Inc. on behalf of International Society for Autism Research

goes unrecognized in ASD, with less than half of those with co-occurring conditions receiving a community-based psychiatric diagnosis [Buck et al., 2014; Hayward et al., Service use and needs among those with an ASD in adolescence and young adulthood: Unrecognised psychiatric conditions, unpublished manuscript]. To assist in identification of psychiatric comorbidity, clinicians need valid and reliable screening tools. However, there are currently no screening tools that screen for a range of psychiatric comorbidity validated for use in adults with ASD. The aim of this pilot study was to establish whether a modified version of the Strengths and Difficulties Questionnaire (SDQ) – one of the most widely used screening tools for psychopathology in children and adolescents in the United Kingdom – could fulfill this purpose.

The SDQ is a brief, 25-item questionnaire with five subscales: emotional problems, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior. It has excellent psychometric properties – mean Cronbach alpha = 0.73 [Goodman, 2001] and mean test-retest reliability = 0.81 [Muris, Meesters, & van den Berg, 2003] – and can be administered as self- or informant-report [Goodman, 2001]. The SDQ has the potential to be a good instrument for use in adolescents and adults with ASD because it: (a) can be used to screen for psychopathologies that are commonly associated with ASD (anxiety, depression, and ADHD); (b) is relatively brief; (c) is likely that cooperation is better when using a single questionnaire rather than several separate questionnaires (for anxiety, depression, ADHD, etc.); and (d) can be administered to individuals with ASD and their parents (or other informants). We investigated the SDQ in adolescents and adults with ASD, testing inter-rater reliability, test-retest stability, internal validity, and external validity (against clinically utilized measures of anxiety, depression, and ADHD). We hypothesized that: (1) the SDQ would demonstrate inter-rater reliability, test-retest stability, internal consistency, and construct validity that is comparable to values from non-ASD populations; (2) the SDQ emotional problems subscale would be significantly correlated with multiple other measures of anxiety disorders and depression, and both self- and informant-report would show good sensitivity and specificity in relation to clinical risk of emotional disorders; and (3) the SDQ hyperactivity subscale would be significantly correlated with multiple measures of ADHD, and both self- and informant-report would show good sensitivity and specificity in relation to clinical risk of ADHD.

Method

Sample

Individuals with a clinical diagnosis of ASD and their parents were recruited from Child and Adolescent Mental Health Services, adult clinics, charities, and

clinical research databases. One-hundred and forty-five families participated: from these families, 98 individuals with ASD (age 14–59 years) and 126 parents (mostly mothers) provided data; in 79 families both self- and parent-report data were available (see supporting information Tables S1 and S2 for details). Of participating families, 71.0% of patients were currently in or had recently sought help from clinical services because of ASD-related difficulties. Diagnosis of ASD was confirmed using the autism diagnostic interview-revised (ADI-R) [Lord, Rutter, M., & Le Couteur, 1994] or the autism diagnostic observation schedule-generic (ADOS-G) [Lord et al., 2000]. The study was approved by The Joint South London and Maudsley and The Institute of Psychiatry NHS Research Ethics Committee.

Measures

Strengths and Difficulties Questionnaire

The SDQ is a brief, 25-item questionnaire that can be administered as self-report, or to parents and teachers of children and adolescents [youthinmind, 2012]. It has been adapted for use in adults with minor alteration of some items (e.g., “I often offer to help others (family members, friends, colleagues)” instead of “I often volunteer to help others (parents, teachers, children)” [youthinmind, 2013]. It covers common areas of emotional and behavioral difficulties, with five subscales – emotional problems, conduct problems, hyperactivity, peer problems, and prosocial behavior – and an impact supplement, which enquires about chronicity, distress, social impairment, and burden to others. Scores for subscales are classified as “normal,” “borderline,” or “abnormal” according to published cut-offs [Goodman, 1997]. A three factor model is produced by summing scores on the emotional and peer problems subscales (internalizing) and hyperactive and conduct problems subscales (externalizing).

Previously validated algorithms [Goodman et al., 2004] were used to predict psychiatric disorder by bringing together information on symptoms and impact from self- and parent-report SDQ for ADHD/hyperkinetic disorders and anxiety-depressive disorders. Each is predicted to be “unlikely,” “possible,” or “probable.” “Probable” hyperactivity-inattention disorder requires evidence of pervasiveness through parent and teacher ratings; because teacher ratings were not available, only “unlikely” or “possible” predictions were produced.

Psychiatric Diagnosis

Psychiatric comorbidity was measured using the development and wellbeing assessment (DAWBA) [Goodman, Ford, Richards, Gatward, & Meltzer, 2000] and clinical notes where available ($n = 72$). The DAWBA is a computer-based psychiatric interview that can be

completed by youths/adults and their parents. The DAWBA was originally developed for children and adolescents but has been adapted for use in adults by modifying the language to make the content appropriate to adults (see *youthinmind* [2015] for sample interviews). Around 95% of the content remains the same as the original version and the while formal validation studies of the adult version have yet to be published, the DAWBA is increasingly used in longitudinal studies of adults and trained clinical raters have reported that the rating process is very similar to the adolescent version [Robert Goodman, personal communication, November 9, 2015]. The DAWBA can be administered by interview or can be completed online by patients and their parents. The online version has been shown to produce broadly similar prevalence estimates for disorders as those produced using interview techniques, in both a general population sample and a sample affected by neurological disorder [Goodman, 2013] and agreement with clinical diagnosis of specific disorders has been shown to range between 74 and 99% in an outpatient sample [Aebi et al., 2012]. The DAWBA uses closed and open-ended questions that cover the operationalized diagnostic criteria of the DSM-IV, DSM-5 and ICD-10 [American Psychiatric Association, 2000]. Screening questions and the SDQ are used to shorten sections of the interview when it is unlikely that the person has difficulties in that area. In instances where the answers to the screening questions do not indicate the presence of disorder, but the SDQ scores do, the section is entered. If the answers to the screening questions do indicate the presence of disorder, the SDQ scores are not considered. An algorithm uses responses from self-report and informants to generate probability bands for the presence of particular psychiatric disorders. SDQ scores do not contribute to the DAWBA probability bands. Clinical raters then review all information and assign a diagnosis for each disorder. In this study, two trained clinical raters assigned diagnoses and final DAWBA diagnosis for each disorder was decided through consensus between the two raters.

Where available, clinical notes were reviewed and comorbid psychiatric disorders were coded as “no,” “possible” (e.g., suspected but the individual had not been further assessed), or “definite.”

Dichotomous variables indicating high risk of particular disorders were also created: high risk was indicated by either a DAWBA diagnosis by a clinical rater or a definite diagnosis recorded in their clinical notes.

In addition to the DAWBA, we assessed the external validity of the SDQ against screening questionnaires that are frequently used in clinical settings.

Screening Questionnaires

Symptoms of ADHD were measured using the Barkley Adult ADHD Rating Scale-IV (BAARS-IV) [Barkley &

Murphy, 2006]. The BAARS-IV is a commonly used, validated 18 item rating scale of DSM-IV criteria for ADHD that has been used in adult ASD populations [Johnston et al., 2013; Wilson et al., 2013]. Parent- and self-report versions were used.

Anxiety and depression were measured using the hospital anxiety and depression scale (HADS) [Zigmond & Snaith, 1983]. The HADS is a commonly used, 14-item self-report screening tool for anxiety and depression. It has demonstrated excellent psychometric properties and is used routinely in clinical and research contexts, including in adult ASD populations [Abell & Hare, 2005; Hare, Wood, Wastell, & Skirrow, 2015; Kanai et al., 2011; Luke, Clare, Ring, Redley, & Watson, 2011].

The obsessive-compulsive inventory-revised (OCI-R) [Foa et al., 2002] was used to measure symptoms of obsessive-compulsive disorder (OCD). The OCI-R is an 18-item self-report measure with excellent psychometric properties [Foa et al., 2002] that performs well in adults with ASD [Cadman et al., 2015].

Procedure

Following informed consent, login details for the SDQ and DAWBA were issued to participants. Each individual had log in details specific to them to enable adolescents/adults and parents to separately log in and complete them. The SDQ was completed as part of the DAWBA package, online (92% of adolescents/adults, 91% of parents) or by telephone or face-to-face interview (8% of adolescents/adults, 9% of parents); in all cases the SDQ was completed before the DAWBA. Paper versions of the BAARS-IV, HADS, and OCI-R were posted to participants to complete and return using a pre-paid envelope. If participants consented to their medical records being reviewed then relevant records were requested from their current mental health service.

Statistical Analysis

Data were analyzed using SPSS for Windows (version 20.0) [IBM Corp., 2011] and Mplus (version 7.3) [Muthén & Muthén, 1998].

The SDQ was validated against other indicators of high risk of disorder by examining the correlation (Spearman's ρ) between SDQ subscale scores and other measures of psychiatric disorder, namely: DAWBA probability bands, clinical diagnosis (no, possible, or definite diagnosis), and standard screening questionnaires. To ensure that associations were not driven by the severity of overall problems (e.g., individuals with more severe problems being rated non-specifically across all SDQ subscales) correlations were repeated controlling for SDQ impact scores (partial ρ). Additionally, correlations

Table 1. Test-Retest Stability: Reliability Coefficients for SDQ Scores

	Time 1 × Time 2 correlations	
	Parent-report	Self-report
SDQ scale	(<i>n</i> = 69)	(<i>n</i> = 35)
Emotional problems	0.55	0.64
Behavioral Problems	0.70	0.64
Hyperactivity	0.70	0.55
Peer problems	0.58	0.72
Prosocial behavior	0.62	0.69
Impact	0.60	0.61
Total difficulties	0.59	0.74

All Spearman's correlation coefficients significant at $P < 0.01$.

with the three-category ordinal clinical diagnosis variables were repeated as polyserial correlations.

The second step was to establish whether the SDQ bands – normal, borderline, abnormal – predict high risk of disorder, using parent-report SDQ, self-report SDQ, and predictive algorithm based on combined parent- and self-report SDQ. SDQ emotional problems and hyperactivity subscales were compared to high risk of disorder (indexed by having a DAWBA clinical rater assigned diagnosis and/or a definite diagnosis recorded in their clinical notes). For parent- and self-report SDQ, borderline and abnormal bands were collapsed into one category. When the predictive algorithm was used, possible and probable categories were collapsed for emotional disorders (probable category not available for hyperactivity; see Measures: SDQ, above). Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for each rater and the predictive algorithm, as well as the percentage of individuals correctly classified and the odds ratio. To establish whether parent- or self-report SDQ was significantly better at identifying cases, the proportion of true positives and negatives was compared between parent- and self-report SDQ using the *Z* test.

Construct validity was tested using multitrait–multimethod (MTMM) [Campbell & Fiske, 1959; Goodman, Lamping, & Ploubidis, 2010] analysis. MTMM is based on producing a correlation matrix of multiple traits (e.g., the subscales of the SDQ) measured by multiple methods (e.g., parent- and self-report). Construct validity is assessed through comparisons across raters and across subscales. For example, the correlation between parent- and self-report emotional subscale scores (a convergent correlation coefficient) would be expected to be higher than between parent-report emotional and self-report hyperactivity subscales (a discriminant correlation coefficient). Where there is no difference between convergent and discriminant correlations, it is inferred that these subscales may not be tapping into the same constructs across raters. MTMM correlations were

performed using Spearman's ρ and compared using Steiger's *Z* tests. We also present Cronbach's alpha, a coefficient of internal consistency, for each subscale in each rater.

Results

Description of Sample

SDQ and DAWBA data were available from 126 parents (offspring gender, 81% male; offspring age, mean [SD] = 23.5 [10.8] years) and 98 adolescents/adults with ASD (75% male; age, mean [SD] = 27.4 [12.3] years); of these, there were 79 dyads with both parent- and self-report data. Of these dyads, subsets of cases had clinical notes relating to co-occurring disorders ($n = 39$) or questionnaire data ($n = 28$). See supporting information Tables S1 and S2 for more detail.

All participants had a previous clinical diagnosis of an autism spectrum disorder. Participants met ADI-R autism genetic resource exchange criteria for “autism,” “not quite autism” or “broad spectrum” [AGRE, 2015]. In all but five cases classified as “broad spectrum,” this resulted from missing data for domain D; they otherwise met criteria for “autism” or “not quite autism” across the three symptom domains. Two participants classified as “broad spectrum” additionally had ADOS data and met criteria for autism. Reviewing clinical notes and diagnoses made by DAWBA clinical rater, 46.5% of participants were at high risk of co-morbid ADHD and 59.9% were at high risk of a co-morbid emotional disorder. Formal IQ data were not available but 8% of the sample was diagnosed with intellectual disability (also known as learning disability in UK health services).

Inter-rater Reliability

Pearson's correlation between parent- and self-report total SDQ scores was $r = 0.42$, $P < 0.05$. This is comparable to the value previously reported in a large scale epidemiological sample ($r = 0.48$) [Goodman, 2001].

Test-Retest Stability

The SDQ was re-administered to parents ($n = 69$) and adolescents/adults ($n = 35$) with ASD after an average interval of 11.8 months (SD = 2.75; range, 6–12). The mean retest stability was $r = 0.64$ and did not differ by informant (parent-report: $r = 0.62$; self-report: $r = 0.66$; see Table 1).

External Validity Against High Risk of Disorder

Parent- and self-report SDQ emotional problems and hyperactivity subscales (normal vs. borderline/abnormal bands) and multi-rater predictive algorithm (unlikely vs. possible/probable bands) were compared to high risk

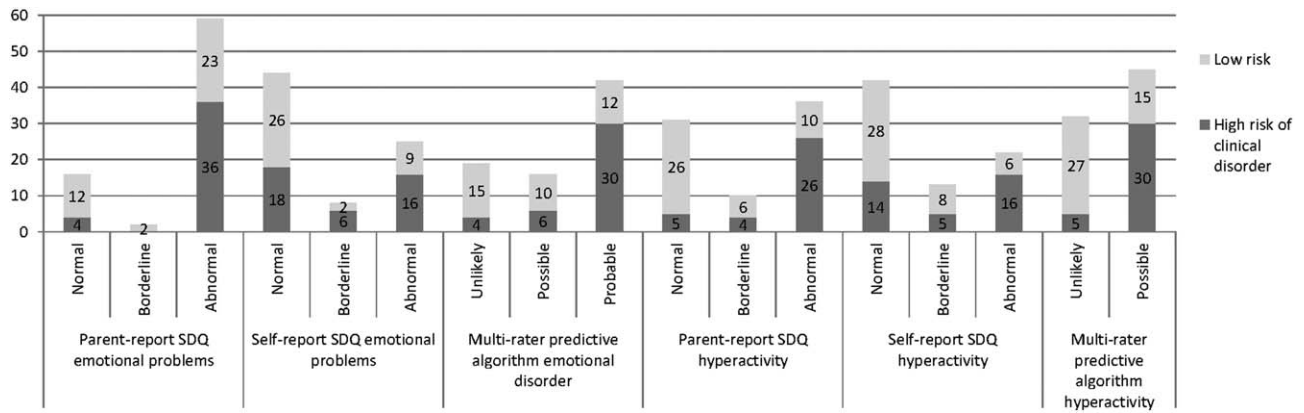


Figure 1. Number of individuals at high or low risk of clinical diagnosis of any type of emotional disorder (for SDQ emotional problems) or ADHD/hyperkineses (for SDQ hyperactivity) in each SDQ band (parent-report, self-report, and multi-informant predictive algorithm). NB Probable band for multi-rater predictive algorithm for hyperactivity requires evidence of pervasiveness across settings from parent and teacher SDQ, which was not available in this sample.

of emotional disorder or ADHD, defined as having a DAWBA clinical rater diagnosis and/or diagnosis recorded in clinical notes.

Emotional Disorder

Figure 1 shows ratings for SDQ emotional problems or disorder in those at high and low risk of clinical diagnosis of any emotional disorder. For both parent- and self-report emotional problems, more than half of those rated as borderline or abnormal were at high risk of emotional disorder and the majority of those rated as normal were at low risk of emotional disorder. Parents were more likely to give abnormal ratings, resulting in more true and false positives, and self-report was more likely to result in normal ratings, resulting in more true and false negatives.

Table 2 shows the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of SDQ emotional problems subscale in predicting high risk of emotional disorders. Parent-reported SDQ emotional problems had high sensitivity (90% of cases detected) whereas only around half (55%) of cases were detected using self-ratings. The PPV showed that 60–70% of cases identified were true positives and 30–40% were false positives. The NPV showed that 60–75% of normal ratings reflected low risk of disorder, whereas 25–40% represented missed cases. The odds of clinical disorder were significantly elevated in those identified using the SDQ, using parent- or self-ratings. Multi-rater predictive algorithm performed similarly to parent-report SDQ, with only slightly improved ability to correctly identify risk of clinical disorder. Restricting analyses to those who were currently in or had recently sought help from clinical services for ASD-related difficulties further reduced performance of self-report SDQ emotional problems; it did not perform significantly

above chance level (see supporting information Table S3).

There was no significant difference in the PPV between parent- and self-rated SDQ emotional problems (Z score = -1.03 , $P = 0.303$); for NPV, parent-report was significantly better than self-report (Z score = -2.11 , $P = 0.034$).

Hyperactivity

A similar pattern was observed for SDQ hyperactivity. Figure 1 and Table 2 show that the majority (60–65%) of those with borderline or abnormal ratings were at high risk of ADHD and the majority (67–84%) of those with normal ratings were at low risk. Odds ratios were significant for parent-ratings, self-ratings, and multi-rater predictive algorithm. Parent-ratings detected more cases (86%) than self-ratings (60%). In two cases, elevated SDQ scores, rather than DAWBA screening questions, had resulted in participants completing the DAWBA hyperactivity module; excluding these two cases did not change estimates of sensitivity, specificity, PPV or NPV. Restricting analyses to those who were in or had recently sought help from clinical services resulted in an increase in sensitivity for both parent- and self-report, though a decrease in specificity for self-report (see supporting information Table S3).

There was no significant difference in the PPV between parent- and self-rated SDQ hyperactivity (Z score = 0.64 , $P = 0.522$); for NPV, parent-report was significantly better than self-report (Z score = 2.45 , $P = 0.014$).

Specificity of SDQ Subscales

Spearman's rank correlations (ρ) and, where appropriate, polyserial correlations between SDQ subscales and other measures of comorbid disorders (DAWBA

Table 2. Classification Statistics for SDQ Emotional Problems and Hyperactivity Subscales

SDQ subscale		Sens	Spec	PPV	NPV	% Correctly classified	Odds ratio	95% CI
Emotional problems	Parent	0.90	0.32	0.59	0.75	62	4.32	1.25–14.95
	Self	0.55	0.70	0.67	0.59	62	2.89	1.13–7.40
Emotional disorder	Multi-rater	0.90	0.41	0.62	0.79	66	6.14	1.81–20.87
Hyperactivity	Parent	0.86	0.62	0.65	0.84	73	9.75	3.14–30.28
	Self	0.60	0.67	0.60	0.67	64	3.00	1.18–7.62
Hyperactivity disorder	Multi-rater	0.86	0.64	0.67	0.84	74	10.80	3.46–33.70

SDQ emotional problems/disorder tested against high risk of any emotional disorder (generalized anxiety disorder, specific phobia, social phobia, panic disorder/agoraphobia, obsessive-compulsive disorder, major depression); SDQ hyperactivity/hyperactivity disorder tested against high risk of any ADHD/hyperkinesia disorder; high risk of disorder was defined as having DAWBA clinical rater diagnosis and/or definite diagnosis recorded in clinical notes. SDQ borderline/abnormal and possible/probable ratings were collapsed and contrasted to normal or unlikely ratings. Sens: sensitivity, Spec: specificity, PPV: positive predictive value, NPV: negative predictive value. $n = 77$. Power achieved for: emotional problems parent-report SDQ = 99%; self-report SDQ = 49%; multi-rater algorithm = 99%; hyperactivity parent-report SDQ = 95%; self-report SDQ = 50%; multi-rater algorithm = 97%.

probability bands, clinical diagnosis, and standard screening questionnaires) were examined (see supporting information Tables S4 and S5). Parent-report SDQ scores showed the expected pattern of associations for emotional problems, correlating significantly with measures of anxiety disorders ($\rho = 0.48\text{--}0.74$), OCD ($\rho = 0.32\text{--}0.58$) and depression ($\rho = 0.58$ for DAWBA probability bands; other measures of depression were not significant) and hyperactivity, which correlated significantly with measures of ADHD ($\rho = 0.49\text{--}0.77$; correlations with self-report BAARS-IV was not significant). This was true even when the effect of SDQ impact scores were controlled (using partial correlations), suggesting that these associations were not being driven by severity of problems regardless of domain. SDQ parent-reported peer problems also showed some association with anxiety disorders and hyperactivity, though only the latter was still significant after controlling for impact scores. Self-report SDQ showed similar correlations between SDQ emotional problems and measures of anxiety disorders, OCD and depression, and between SDQ hyperactivity and measures of ADHD (supporting information Table S5). However, self-report appeared to be less specific than parent-report: hyperactivity and peer problems subscales both correlated significantly with anxiety disorders and OCD, even after controlling for impact scores.

Construct Validity of the SDQ Subscales

Table 3 presents MTMM analysis of the SDQ subscales. Cronbach alpha coefficients ranged between 0.52 and 0.81. The cross-method correlations of the same traits are presented in bold; all were significantly different from zero ($P < 0.001$) and in the moderate range (0.42–0.57).

In the majority of cases, the convergent correlations were significantly larger than the other correlation coefficients in the same row or column (the discriminant correlations), were of moderate size, and were larger

than those previously reported in an epidemiological sample of adolescents [Goodman et al., 2010]. The exceptions to this were: (a) the peer problems subscale, where both parent- and self-rated SDQ showed some evidence of poor discriminant validity relative to emotional problems reported by the other rater; (b) the parent-reported behavioral problems subscale, which demonstrated poor discriminant validity relative to self-reported emotional problems; and (c) self-reported behavioral problems, which demonstrated poor discriminant validity relative to parent-reported hyperactivity. However, the statistical power achieved was low (34–40%) for these discriminant correlations. The three factor (internalizing–externalizing–prosocial) contrast provided much better convergent and discriminant validity (see Table 4).

Discussion

Rates of psychiatric comorbidity are extremely high in adolescents and adults with ASD, causing significant impairment to affected individuals and great burden to their families. Regular screening for common comorbid disorders, such as anxiety, depression, and ADHD, is extremely important, yet there are no screening tools that have been validated in this population. We identified the SDQ as a tool that could potentially be used for this purpose and carried out a pilot study to test its reliability and validity in a sample of adolescents and adults with ASD who were either currently in or had recently sought help from clinical services for ASD-related difficulties. Our findings suggest that the SDQ shows some promise in this population; however, more systematic attempts at validation are warranted.

We assessed the reliability of the SDQ by examining inter-rater agreement, internal consistency, and test-retest stability. Inter-rater agreement between parent- and self-report was comparable to the agreement

Table 3. MTMM Analyses for the Five SDQ Subscales

		Parent-report SDQ					Self-report SDQ				
		Emo	Behav	Hyp	Peer	Pro	Emo	Behav	Hyp	Peer	Pro
Parent SDQ	Emo	$\alpha = 0.78$									
	Behav	0.36	$\alpha = 0.73$								
	Hyp	0.31	0.40	$\alpha = 0.77$							
	Peer	0.44	0.34	0.23	$\alpha = 0.52$						
	Pro	-0.14	-0.45	-0.16	-0.26	$\alpha = 0.81$					
Self-report SDQ	Emo	0.57***^A	0.43 ^{A,1}	0.08	0.29 ²	-0.06	$\alpha = 0.76$				
	Behav	-0.06	0.42***^{B,1}	0.32 ^B	-0.01	-0.07	0.14	$\alpha = 0.61$			
	Hyp	0.18	0.15	0.56***	0.05	-0.07	0.35	0.21	$\alpha = 0.72$		
	Peer	0.25 ^C	-0.06	-0.20	0.45***^{C,2}	-0.12	0.57	0.16	0.09	$\alpha = 0.62$	
	Pro	-0.02	-0.10	-0.02	-0.07	0.47***	-0.15	-0.23	-0.13	-0.18	$\alpha = 0.73$

MTMM = multi-trait multimethod, Emo = SDQ emotional problems subscale, Behav = behavioral problems, Hyp = hyperactivity, Peer = peer problems, Pro = prosocial behavior. $n = 126$ parents; $n = 98$ adolescents/adults with ASD; $n = 79$ for parent-self-report comparison. Values are Spearman's correlation coefficients, except values in the diagonals which are Cronbach's alphas. Cross-method correlations of same traits are presented in bold.

*** $P < 0.001$.

Coefficients in the same row (^{A,B,C}) or the same column (^{1,2}) with the same superscript do not differ from each other using Steiger's Z tests (^A $P = 0.173$; ^B $P = 0.375$; ^C $P = 0.068$, ¹ $P = 0.936$; ² $P = 0.095$).

Table 4. MTMM Analyses for the Internalizing, Externalizing, and Prosocial SDQ Subscales

		Parent-report SDQ			Self-report SDQ		
		Int	Ext	Pro	Int	Ext	Pro
Parent-report SDQ	Int	$\alpha = 0.75$					
	Ext	0.40	$\alpha = 0.78$				
	Pro	-0.22	-0.33	$\alpha = 0.81$			
Self-report SDQ	Int	0.53***	-0.02	-0.10	$\alpha = 0.80$		
	Ext	0.06	0.53***	-0.10	0.27	$\alpha = 0.69$	
	Pro	-0.02	-0.10	0.47***	-0.19	-0.23	$\alpha = 0.73$

MTMM = multi-trait multimethod, Int = internalizing, Ext = externalizing, Pro = prosocial. $n = 126$ parents; $n = 98$ adolescents/adults with ASD; $n = 79$ for parent-self-report comparison. Values are Spearman's correlation coefficients, except values in the diagonals which are Cronbach's alphas. Cross-method correlations of same traits are presented in bold.

*** $P < 0.001$.

reported in a non-ASD population using the SDQ [Goodman, 2001]. Internal consistency was generally good across all subscales and values for Cronbach's alpha were mostly higher than those previously reported in a non-ASD sample [Goodman, 2001]. Test-retest stability was also comparable to that reported in a non-ASD population, despite being carried out after a longer interval [Goodman, 2001]. Therefore, overall the SDQ appears to perform as reliably in adolescents and adults with ASD as it does in non-ASD populations of children and adolescents.

Having demonstrated good reliability, we next explored to what extent the parent- and self-report versions of SDQ showed good external validity when compared to clinically utilized measures of comorbid disorders. The parent-report SDQ emotional problems subscale showed reasonable validity, being specifically associated with measures of emotional disorders (generalized anxiety disorder, specific phobia, social phobia,

panic and agoraphobia, OCD, and depression) and detecting 90% of individuals at high risk of disorder. Similarly, the parent-report SDQ hyperactivity subscale was specifically associated with measures of ADHD and detected 86% of those at high risk of ADHD. Therefore obtaining a parent-report SDQ and taking borderline and abnormal scores on emotional problems or hyperactivity subscales to indicate risk of disorder would be an effective way of detecting the majority of those in need of further assessment and treatment. In practice it might be possible to use just the emotional and hyperactivity subscales as a very brief 10-item screening tool that could be used during routine appointments. However, the specificity was relatively low and the rate of false positives was up to 40% and so intensive work up of SDQ positive cases may not always be warranted; further screening with more in depth tools (such as the DAWBA) in the first instance could provide a more cost-effective approach.

In contrast to the excellent validity shown by the parent-report SDQ, self-report SDQ produced a less specific pattern of correlations with measures of disorders; emotional and hyperactivity subscales identified fewer cases (55–60%); and there were more false negatives. Some adults with ASD might under-report these difficulties or rate them in atypical ways, for example rating anxiety symptoms under hyperactivity items (e.g., “I am easily distracted”). Given this non-specificity, if only self-report SDQ is available then high scores in *any* subscale should be taken to indicate the possibility of co-occurring disorder and further assessment planned accordingly. Ideally, parent-report (or other informant) should be sought; in practice, this might mean that the SDQ has greater utility in adolescent or young adult populations, who are more likely to be accompanied to appointments by their parents.

As well as testing the validity of parent- and self-report SDQ, we tested multi-rater predictive algorithms that incorporate information from the SDQ impact supplement. This did not markedly improve detection of comorbidities over parent-report SDQ. This suggests that in adolescents and adults with ASD, there is little justification for combining parent- and self-report SDQ and impact scores using a statistical package. In practice, it may be sufficient, and indeed easier, to use a paper version of the parent-report SDQ and the published subscale bands.

Having explored the external validity of the SDQ in this population, we set out to test its construct validity in measuring hyperactivity, emotional, behavioral and peer problems, and prosocial behavior using multitrait-multimethod (MTMM) analysis of parent- and self-report SDQ. The convergent validity coefficients were higher than those previously reported in a non-ASD sample [Goodman et al., 2010]. We generally found good discriminant validity between the subscales, providing some evidence that the subscales identify distinct areas of psychopathology in adolescents and adults with ASD. Despite this, there was some evidence of poor discriminant validity between the peer problems subscale and the emotional symptoms subscale, where convergent and divergent correlations did not differ significantly. However, the convergent correlations were around twice the magnitude of the divergent correlations and this difference was close to significance ($P < 0.10$). Furthermore, within each rater emotional and peer problems subscales were highly correlated, suggesting that these difficulties may be closely linked in individuals with ASD (e.g., anxiety may be linked to social situations and peer relationships). There was also evidence for poor discrimination between parent-reported behavior problems and self-reported emotional problems. It is possible that this shows that some problem behavior(s) identified by parents (e.g., temper tan-

trums) are an indicator of emotional difficulties (e.g., fears) but are not clearly articulated as such by the individual with ASD. Finally, there was partial evidence for poor discrimination between parent-reported hyperactivity and self-reported behavior problems. This has been found in other samples [Goodman et al., 2010] and it may be difficult to separate these two aspects of externalizing behavior using a brief questionnaire measure. The low statistical power achieved in this analysis may have, in part, contributed to the apparent poor discrimination between some subscales. The three factor model – internalizing, externalizing, and prosocial behavior subscales – showed much better discriminant validity, indicating that the SDQ can clearly differentiate between internalizing and externalizing problems in adolescents and adults with ASD.

Strengths and Limitations

The promising results of this first attempt to validate the SDQ in adolescents and adults with ASD should be interpreted in the context of the strengths and limitations inherent in the study, including the nature of the sample, the data collection strategy, and the measures used. The sample included a large group of adolescents and adults who had both a clinical diagnosis of ASD and assessment with ‘gold-standard’ diagnostic tools (ADI-R and/or ADOS-G). However, our sample primarily consisted of high-functioning individuals, which is not representative of ASD as a whole. For example, 8% of our sample was diagnosed with an intellectual disability in comparison to an estimated prevalence in children and adolescents of around 40% [Idring et al., 2012; Investigators AaDDMNS, 2012; Pedersen et al., 2012; Saemundsen, Magnússon, Georgsdóttir, Egilsson, & Rafnsson, 2013]. Therefore, we can only generalize our findings to higher functioning individuals with ASD.

The SDQ was administered to parents and their offspring with ASD at two time-points. Administering to parents and individuals with ASD allowed us to establish inter-rater reliability and gain important information about construct validity. Additionally, re-administering the SDQ to both groups allowed us to explore test-retest stability for both parent- and self-report SDQ. Test-retest reliability is usually measured by re-administering the assessment after a brief interval of approximately 1–4 weeks. However, in the current study retest was not carried out until on average 11 months after initial assessment. Therefore, changes in the score may reflect genuine alterations in the respondent’s psychological adjustment as well as poor measurement reliability. Nevertheless, we report a satisfactory lower bound for true test-retest reliability which was similar across raters and comparable to a non-ASD population tested after a shorter interval.

We also compared the SDQ to a range of measures of psychopathology and using a variety of methods, including some of the most widely used screening tools for psychopathology in the UK (BAARS-IV, HADS, and the OCI-R) and clinical diagnoses assigned by practicing clinicians. The pattern of converging results makes it likely that the associations with the SDQ are robust and not entirely attributable to common method variance (i.e., SDQ and other questionnaires being completed by the same rater). For example, parent-report SDQ emotional problems correlated with the OCI-R, which is a self-report questionnaire, and with diagnosis of OCD assigned by clinician.

However, there were limitations to the data. While the sample consisted of adolescents and adults with ASD who were largely in contact with services because of ASD-related difficulties, they were recruited from a variety of sources, making it harder to generalize results to one clear population. Future studies should recruit samples from single sources (e.g., specific healthcare settings) to determine the utility of the SDQ in each setting. It was also clear from clinical notes that despite having evidence of comorbid disorder some individuals had not received further assessment. Nonetheless, we attempted to capture the possibility of disorder in these cases, while not artificially inflating rates of disorder, by rating them as possible rather than definite disorder. Furthermore, we were not able to employ a 'gold-standard' diagnostic approach to measuring co-occurring disorders and the other measures to which we compared the SDQ (DAWBA and screening questionnaires) have not been validated in adults with ASD. Indeed, the lack of validated screening tools is what motivated this pilot study. We therefore compared the SDQ to other indicators of high risk of co-occurring disorders, rather than definitive diagnoses. The converging evidence across multiple methods of assessment lends support to the validity of the SDQ; however, further validation is required using definitive diagnoses. In some individuals information about the time gap between assessment with the SDQ and with other measures (screening questionnaires, clinical assessment) was not available; consequently, disagreement between the SDQ and other measures may have reflected genuine change in mental state rather than poor performance of the SDQ. This would have acted to decrease the strength of association between the SDQ and other measures; despite this, a consistent pattern of significant associations emerged across measures. Finally, whilst disruptive behavior may be a problem in adolescents and adults with ASD, we did not have sufficient clinical data to validate the conduct problems subscale.

In summary, this first attempt to validate the SDQ suggests that it may be a reliable and valid way to screen for co-morbid emotional disorders and ADHD in adolescents and adults in clinical services or seeking help for ASD-

related difficulties. Given the high prevalence of co-occurring mental health conditions, impact on individuals with ASD, and burden to their families, it is extremely important to detect the emergence of these disorders during adolescence and adulthood. The SDQ shows promise as a tool that could be used to regularly screen for co-occurring disorders in people with ASD, helping to reduce the cost to them and their families.

Acknowledgments

We are grateful to all the individuals and families who participated. We thank Nicola Gillan for her assistance in the recruitment of cases in South London, and the Mental Health Research Network (Joanne Almeida, East Midlands and South Yorkshire; Christine Haecker, Heart of England) for their assistance in recruiting cases from Nottinghamshire and North Staffordshire. This study was supported by a National Institute for Health Research program grant (D.M., Grant No. RP-PG-0606-1045). Funding for D.M., E.S., and P.B. was also provided by NIHR Biomedical Research Centre at South London & Maudsley National Health Service Foundation Trust and King's College London; and the Innovative Medicines Initiative Joint Undertaking under Grant Agreement No. 115300, resources of which are composed of financial contribution from the European Union's Seventh Framework Program (FP7/2007-2013) and EFPIA companies' in kind contribution. An NIHR Senior Investigator award supports P.B. The authors alone bear responsibility for collection, analysis, and interpretation of the data. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflict of Interest

None.

Ethical Standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Abell, F., & Hare, D.J. (2005). An experimental investigation of the phenomenology of delusional beliefs in people with Asperger syndrome. *Autism, 9*(5), 515–531.
- Aebi, M., Kuhn, C., Metzke, C.W., Stringaris, A., Goodman, R., & Steinhausen, H.-C. (2012). The use of the development and well-being assessment (DAWBA) in clinical practice: A

- randomized trial. *European Child & Adolescent Psychiatry*, 21(10), 559–567.
- AGRE (2015). Autism genetic resource exchange. Diagnostic information. Retrieved March 24, 2015, from <http://research.agre.org/program/diag.cfm>.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Publishing, Inc.
- Barkley, R.A., & Murphy, K.R. (2006). *Attention-deficit hyperactivity disorder: A clinical workbook*. New York: Guilford Publication.
- Brugha, T.S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, ... Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives General Psychiatry*, 68(5), 459–465.
- Buck, T.R., Viskochil, J., Farley, M., Coon, H., McMahon, W.M., Morgan, J., & Bilder, D.A. (2014). Psychiatric comorbidity and medication use in adults with autism spectrum disorder. *Journal Autism and Developmental Disorders*, 44(12), 3063–3071.
- Cadman, T., Eklund, H., Howley, D., Hayward, H., Clarke, H., Findon, J., ... Glaser, K. (2012). Caregiver burden as people with autism spectrum disorder and attention-deficit/hyperactivity disorder transition into adolescence and adulthood in the United Kingdom. *Journal American Academy of Child & Adolescent Psychiatry*, 51(9), 879–888.
- Cadman, T., Spain, D., Johnston, P., Russell, A., Mataix-Cols, D., Craig, M., ... MRC AIMS Consortium (2015). Obsessive-compulsive disorder in adults with high-functioning autism spectrum disorder: What does self-report with the OCI-R tell us? *Autism Research*, 8(5), 477–485.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105.
- Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P.M. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment*, 14(4), 485.
- Goodman, A., Lamping, D.L., & Ploubidis, G.B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, 38(8), 1179–1191.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581–586.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal American Academy of Child & Adolescent Psychiatry*, 40(11), 1337–1345.
- Goodman, R. (2013). Online child and adolescent mental health surveys can be good enough. *Social Psychiatry and Psychiatric Epidemiology*, 48, 1317–1325.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41(5), 645–655.
- Goodman, R., Ford, T., Corbin, T., & Meltzer, H. (2004). Using the strengths and difficulties questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders. *European Child & Adolescent Psychiatry*, 13(2), ii25–ii31.
- Hare, D.J., Wood, C., Wastell, S., & Skirrow, P. (2015). Anxiety in Asperger's syndrome: Assessment in real time. *Autism*, 19(5), 542–552.
- Hofvander, B., Delorme, R., Chaste, P., Nydén, A., Wentz, E., Ståhlberg, O., ... Leboyer, M. (2009). Psychiatric and psychosocial problems in adults with normal-intelligence autism spectrum disorders. *BMC Psychiatry*, 9, 35.
- IBMCorp. (2011). *IBM SPSS statistics for Windows*. 20.0 Edition. Armonk, NY: IBM Corp.
- Idring, S., Rai, D., Dal, H., Dalman, C., Sturm, H., Zander, E., ... Magnusson, C. (2012). Autism Spectrum disorders in the stockholm youth cohort: Design, prevalence and validity. *PLoS ONE* 7(7), e41280.
- Investigators AaDDMNS (2012). Prevalence of autism spectrum disorders – Autism and developmental disabilities monitoring network, 14 sites, centers for disease control and prevention, United States, 2008. *MMWR Surveillance Summaries*, 61(3), 1–19.
- Johnston, K., Dittner, A., Bramham, J., Murphy, C., Knight, A., & Russell, A. (2013). Attention deficit hyperactivity disorder symptoms in adults with autism spectrum disorders. *Autism Research*, 6, 225–236.
- Kanai, C., Iwanami, A., Hashimoto, R., Ota, H., Tani, M., Yamada, T., & Kato, N. (2011). Clinical characterization of adults with Asperger's syndrome assessed by self-report questionnaires based on depression, anxiety, and personality. *Research in Autism Spectrum Disorders*, 5, 1451–1458.
- Leyfer, O.T., Folstein, S.E., Bacalman, S., Davis, N.O., Dinh, E., Morgan, J., ... Lainhart, J.E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, 36(7), 849–861.
- Lord, C., Risi, S., Lambrecht, L., Cook Jr, H., Leventhal, B.L., DiLavore, P.C., ... Rutter, M. (2000). The autism diagnostic observation schedule – Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685.
- Luke, L., Clare, I.C.H., Ring, H., Redley, M., & Watson, P. (2011). Decision-making difficulties experienced by adults with autism spectrum conditions. *Autism*, 16(6), 612–621.
- Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child & Adolescent Psychiatry*, 12, 1–8.
- Muthén, L.K., & Muthén, B.O. (1998). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Pedersen, A., Pettygrove, S., Meaney, F.J., Mancilla, K., Gotschall, K., Kessler, D.B., ... Cunniff, C. (2012). Prevalence of autism spectrum disorders in hispanic and non-hispanic white children. *Pediatrics*, 129(3), e629–e635.

- Russell, A.J., Jassi, A., Fullana, M.A., Mack, H., Johnston, K., Heyman, I., ... Mataix-Cols, D. (2013). Cognitive behavior therapy for comorbid obsessive-compulsive disorder in high-functioning autism spectrum disorders: A randomized controlled trial. *Depression & Anxiety*, 30(8), 697–708.
- Saemundsen, E., Magnússon, P., Georgsdóttir, I., Egilsson, E., & Rafnsson, V. (2013). Prevalence of autism spectrum disorders in an Icelandic birth cohort. *BMJ Open* 2013(3), e002748.
- Vannucchi, G., Masi, G., Toni, C., Dell’Osso, L., Marazziti, D., & Perugi, G. (2014). Clinical features, developmental course, and psychiatric comorbidity of adult autism spectrum disorders. *CNS Spectrums*, 19(2), 157–164.
- Wilson, C.E., Gillan, N., Spain, D., Robertson, D., Roberts, G., Murphy, ... Murphy, D.G.M. (2013). Comparison of ICD-10R, DSM-IV-TR and DSM-5 in an adult autism spectrum disorder diagnostic clinic. *Journal of Autism & Developmental Disorders*, 43, 2515–2525.
- World Health Organization (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. World Health Organization.
- youth*in*mind (2012). SDQ. Retrieved March 24, 2015, from <http://www.sdqinfo.com/>
- youth*in*mind (2013). Versions of the Strengths and Difficulties Questionnaire (SDQ) for adults. Retrieved March 24, 2015, from <http://www.sdqinfo.org/Adult/>
- youth*in*mind (2015). DAWBA interviews and questionnaires; View online versions (age in years). Retrieved November 17, 2015, from <http://dawba.info/py/dawbainfo/b4.py>
- Zigmond, A.S., & Snaith, R.P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Table S1. Recruitment sources

Table S2. Sample characteristics for overlapping groups used in analyses

Table S3. Classification statistics for SDQ emotional problems and hyperactivity subscales, restricted to those currently in or who had recently sought help from clinical services

Table S4. Correlations between parent SDQ and DAWBA probability bands, clinical diagnosis, and questionnaires

Table S5. Correlations between self-report SDQ and DAWBA probability bands, clinical diagnosis, and questionnaires