Received: 8 February 2022 | Accepted: 12 April 2023

DOI: 10.1111/jora.12858

JOURNAL OF Research on Adolescence SRA Society for Research on Adolescence www.s-r-a.org

EMPIRICAL ARTICLE

Screening for disruptive behavior in adolescents at risk using the strengths and difficulties questionnaire

Pascalle Spaan[1] | **Frank van den Boogert**[1] | **Nina H. Grootendorst-van Mil**[1,2] | **Witte J. G. Hoogendijk**[1] | **Sabine J. Roza**[1,3]

[1]Department of Psychiatry, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[2]Epidemiological and Social Psychiatric Research Institute (ESPRi), Department of Psychiatry, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[3]Netherlands Institute for Forensic Psychiatry and Psychology, The Hague, The Netherlands

Correspondence
Sabine Roza, Department of Psychiatry, Erasmus MC University Medical Center, Dr. Molewaterplein 40, Rotterdam 3015 GE, The Netherlands.
Email: s.roza@erasmusmc.nl

Funding information
GGz Breburg; GGz Delfland; GGz Westelijk Noord-Brabant; Parnassia Psychiatric Institute Antes; Yulius

Abstract

Disruptive behavior in adolescents is burdensome and may continue into adulthood if left unidentified. The strengths and difficulties questionnaire (SDQ) can screen for disruptive behavior, but its psychometric properties in high-risk samples and ability to predict delinquency warrant further investigation. In 1022 adolescents, we investigated the predictive validity (on average 1.9 years after screening) of the self-reported SDQ on disruptive behavior disorders and delinquency, measured with multi-informant questionnaires and structured interviews. We compared three scoring methods: total, subscale, and dysregulation profile scoring. In this high-risk sample, SDQ subscale scores predicted disruptive behavior outcomes best. Predictive values for the specific types of delinquency were small. Concluding, the SDQ can be used in high-risk settings for early identification of youth with disruptive behavior.

KEYWORDS
delinquency, disruptive behavior disorders, psychometrics, screening, SDQ

INTRODUCTION

Disruptive behavior in adolescents is both common and costly, especially considering the risk of further development of problems into adulthood if left unidentified and untreated (Cohen & Piquero, 2009; Foster et al., 2005). In addition to the societal costs, disruptive behavior can be a burden on classrooms, families, and the children performing it themselves. The term disruptive behavior describes a broad range of behaviors and can be defined in terms of psychiatric diagnoses such as attention-deficit/hyperactivity disorder (ADHD) and disruptive behavior disorders (DBDs), like conduct disorder (CD), and oppositional defiant disorder (ODD; American Psychiatric Association, 2013), as well as in terms of externalizing behaviors like drug abuse and illegal behavior like delinquency (Tremblay, 2010). Individuals with ODD show disobedient behavior, deviance against authority figures, and emotion dysregulation whereas individuals with CD show more callous–unemotional traits and antisocial, rule breaking, and aggressive behavior; ODD is often considered a milder variant or precursor of CD (Loeber & Farrington, 2000). ADHD is characterized by hyperactivity, attention problems, and impulsivity. Although ADHD is not considered a DBD, the comorbidity with CD and ODD is incredibly high and indicative of more problematic behavior (Nordström et al., 2013). Research suggests that a majority of those with CD or ODD have comorbid ADHD and reversely up to 50% of those with ADHD have CD or ODD (Kutcher et al., 2004). Various treatments for DBDs and delinquency are available, mostly in the form of school- or family-focused interventions, which show long-term effectiveness in both decreasing disruptive behavior and costs (Dopp et al., 2017; Sawyer et al., 2015; von Sydow et al., 2013). However, early identification of youth at risk for developing disruptive behavior is needed so that they can be referred to such programs and reap the benefits (Arango et al., 2018; Kutcher et al., 2004).

Timely prevention and intervention are particularly important considering the prevalence and development of disruptive behavior across the life span. First, DBDs are among

the most common diagnoses in child and adolescent mental healthcare (Nordström et al., 2013). Second, in terms of delinquency, the adolescent period is of particular interest as well, because the prevalences for many forms of delinquency are at their peak then (Moffitt, 1993). Delinquent behavior has also been linked to CD, ODD, and ADHD, with CD being the most important predictor of offending behavior that develops into adulthood (Byrd et al., 2012). Furthermore, the development of disruptive behaviors is marked by a high amount of continuity (Farrington et al., 2009; Keenan et al., 1998). In terms of clinical disorders, a DBD, and especially CD, is indicative of later antisocial personality disorders (Loeber et al., 2002). Moreover, early onset is often indicative of a worse prognosis. In children with an onset in preadolescence, delinquent behavior has been linked to a two- to threefold increase in the chances of developing chronic serious and violent offending in adulthood, when compared with children with an onset after age 13 (Loeber & Farrington, 2000). Because of the high prevalence of disruptive behavior in this developmental period and potential continuity thereafter, screening in adolescence is particularly important for preventive efforts.

Currently, one of the most widely used screening instruments for mental health in children and adolescents is the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). SDQ scores can be based on self-report, parent report, teacher report, or any combination of these. Advantages of this screener are the short duration, it takes about 5 min to complete, and the combination of positively and negatively formulated items. It results in a total score, indicative of general emotional and behavioral problems, and five subscale scores: emotional symptoms, conduct problems, hyperactivity/inattention, peer problems, prosocial behavior. A wealth of studies support the reliability and validity of the original SDQ as a screener for child psychopathology in the general population (Achenbach et al., 2008; Goodman, 2001; Goodman et al., 2003, 2010; Goodman & Goodman, 2009). In addition to using the SDQ total and subscale scores, research is increasingly investigating the SDQ dysregulation profile (SDQ-DP) scoring method. The profile is derived from the dysregulation profile on the Child Behavior Checklist and measures dysregulation of affect, thoughts, and behavior (Holtmann et al., 2011). So far, studies have found that high dysregulation profile scores are related to a higher likelihood of ODD, CD, and ADHD and are indicative of higher severity and more dysfunction (Caro-Cañizares et al., 2017; Deutz et al., 2018; Holtmann et al., 2011). However, more research is needed to see how this way of scoring performs in comparison to total and subscale scoring methods.

Although the psychometric properties of the SDQ in both universal prevention strategies and clinical settings have been addressed extensively, less is known about the use of this screener for populations at increased risk, that is for selective primary prevention efforts. The idea behind selective primary prevention is that screening in groups that are known to be at increased risk might be more efficient and effective (Arango et al., 2018). An example of this would be to target, screen, and offer treatment to children of parents with severe mental illness (SMI), who are at high increased risk of developing disruptive behavior and other mental health disorders. A recent meta-analysis showed that children of parents with SMI (schizophrenia, bipolar disorder, major depressive disorder) had a 1.77 times higher risk of a developing any mental disorder before age 20 compared to children of parents without SMI, with the prevalence of any mental health disorder before age 20 being 57% (Rasic et al., 2013). Interventions in children of parents with mental illness have been shown to reduce mental health problems, including externalizing symptoms (Siegenthaler et al., 2012; Thanhäuser et al., 2017). However, using the SDQ as a screening tool in a selective primary prevention group like this might lead to different psychometric properties, because a selection of children at risk was already made, and it begs the question if further selection by screening could still be beneficial.

Furthermore, the SDQ is increasingly being used in young populations that have conducted delinquent behavior, like youth who have started fires and youth in correctional facilities, or those at high risk of delinquency, such as children of parents with a history of offending (Colins et al., 2013; Lambie & Krynen, 2017; Whitten et al., 2019). Similar to children of parents with SMI, these samples might also be considered selective primary prevention and intervention targets, because mental health problems are known to be increased in youths in the community that have conducted delinquent behavior, adolescents in correctional facilities, and children of parents with a history of offending (Atilola et al., 2018; Casswell et al., 2012; Whitten et al., 2019). In addition to the lack of knowledge on the psychometric properties of the SDQ in such settings, little is known about the predictive ability of the SDQ with regard to delinquency outcomes. Some studies have been conducted. Bailey and Tarbuck (2006) reviewed screeners for mental health in youth who had committed a crime, including the SDQ, and concluded that predictive values were better than chance, but not ideal. Paalman et al. (2011) studied a sample of 140 boys with a Dutch–Moroccan background and found a distinction could be made between a group defined as one-time offenders, boys with a registration from the police before age 12, and controls, boys without such a police registration, based on a clinical score on the emotional symptoms subscale of the parent-report or self-report SDQ. However, Colins et al. (2013) found that self-report SDQ subscale scoring method did not predict future violent offending and weakly predicted property offending in 444 adolescents in correctional facilities (mean age 16.65 years) from Dutch, Moroccan, and Surinamese origin and the authors conclude that the SDQ should not be used for assessing risk of future violence. Lambie and Krynen (2017) studied 57 adolescents aged 13–17 years and found that adolescents that started fires had a higher risk of clinical scores on the conduct problems and hyperactivity/inattention scales than typically developing secondary school students. Combined, these results suggest that the predictive ability of the SDQ might vary per subscale and for different

JOURNAL OF Research on Adolescence

types of delinquency, such as violent, property, and vandalism or public order delinquency. Considering these mixed findings on the predictive ability of the SDQ with regard to delinquency, studies with larger sample sizes, using multiple informant information, and considering various types of delinquency are needed to strengthen conclusions.

The current study investigated the Dutch self-report version of the SDQ for ages 11–17 (sometimes referred to as "adolescent version" or "youth version" [SDQ-Y]). Regarding this version, recent studies have investigated its psychometric properties in community and clinical adolescent samples, reporting acceptable convergent, discriminative, and concurrent criterion validity, and support for its norm-scores and five-factor structure, but also a six-factor structure with the positively formulated items as separate factor (Vugteveen et al., 2018, 2020, 2021). However, in terms of looking at disruptive behavior outcomes, these studies included either clinician outcome ratings or self-report questionnaires separately and looked at cross-sectional associations (concurrent criterion validity) rather than longer term predictions (predictive criterion validity). Furthermore, clinicians were not blinded to SDQ screening scores which may have influenced the results and DBDs were often treated as one outcome (e.g., ODD/CD). The current study aimed to further investigate the value of the Dutch self-report SDQ as a screener in adolescents to predict disruptive behavior, including various types of delinquency, approximately 2 years after screening. In this paper, we use the term *prediction* to refer to the estimation of events that are measured at a later time point than the screening, that either have not yet occurred or have not yet been observed. The disruptive behaviors studied included CD, ODD, ADHD, comorbid ODD or CD with ADHD, serious delinquency, violent delinquency, property delinquency, and vandalism and public order delinquency. Multi-informant standardized measurements and clinical interviews were used to diagnose CD, ODD, and ADHD, as is generally recommended in child and adolescent psychiatry to obtain a comprehensive assessment of behavior across multiple settings (Van Der Ende & Verhulst, 2005). Interviewers in our study were blind to the SDQ scores, so that knowledge of screening did not influence the outcomes. Therefore, our study addressed some of the limitations of previous studies. Furthermore, we aimed to compare the performance of three scoring methods: total, subscale, and dysregulation profile scoring. A third aim was to investigate the performance of the SDQ in selective primary prevention strategies, by comparing the psychometric properties of the SDQ between adolescents of parents with and without SMI.

METHOD

Study design

The current study used data from the iBerry study, a prospective population-based cohort study on adolescents at high risk of psychopathology; the study design is discussed in detail, including descriptions regarding statistical power, participant exclusion, and representativeness of the sample, by Grootendorst-van Mil et al. (2021). The iBerry study was designed to investigate the transition from subclinical symptoms to psychiatric disorders. In a larger area (including rural and urban regions) in the Netherlands, adolescents in the first year of high school between 2014 and 2016 completed the SDQ as part of a health screening by the Centre for Children and Families. From these 16,736 adolescents aged 12–14 years, a selection was made based on SDQ score. All top 15% scoring adolescents (cut-off: girls >14, boys >13) and a random sample of the lower 85% scoring adolescents were selected, with a 2.5:1 ratio favoring high-scoring adolescents, to create a cohort at high risk of emotional and behavioral problems with sufficient power to study the less common outcomes. Adolescents could not participate if they were already participating in another local cohort study, refused the information leaflet, could not be contacted, or simply declined. At the first cohort measurement (T0) conducted between September 2015 and 2019, 1022 adolescents participated (728 high-scoring and 294 lower scoring adolescents, 54% response rate) usually accompanied by a parent or primary caregiver. Adolescents and (accompanying) parents or caregivers provided informed consent and completed interviews, questionnaires, and biological measurements. Adolescents received a small monetary compensation. The study team was blind to screening status. The current study uses data from all 1022 adolescents in the high-risk cohort who participated in the screening stage and the first cohort measurement, on average 1.9 years after screening ($SD = 0.84$). The Medical Ethical Commission of the Erasmus Medical Center, Rotterdam, approved the study protocol (MEC-2015-007).

Measurements

Demographic characteristics

The demographic characteristics of adolescents and accompanying parents were obtained. Educational level of the adolescent was coded as secondary pre-vocational education (VMBO), secondary general education (HAVO), or pre-university education (VWO). Educational level of the parent was based on the highest obtained diploma as lower (primary school or secondary pre-vocational training; corresponding to up to 12 years of education), intermediate (vocational training, secondary general or pre-university education; corresponding to about 13–15 years of education), or higher (higher or academic education; corresponding to 16 or more years of education). For adolescents and parents, we used the country of birth of the participant and their parents as indicator of cultural and geographic background and proxy of ethnic background (Stronks et al., 2009). Cultural and geographic background was coded into three groups: Dutch, other-Western, and non-Western. A Dutch background was

based on the participant and both parents being born in the Netherlands; a Western background on the participant or their parent being born in Europe (excluding Turkey), North America, or Oceania; and a non-Western background on the participant or their parent being born in Africa, Latin America, or Asia. If both parents were born abroad, the mother's country of birth was given priority.

Strengths and difficulties questionnaire

The Dutch translation of the youth self-report SDQ (Van Widenfelt et al., 2003) was used to screen for emotional and behavioral problems. It consists of 25 statements that measure five subscales, namely emotional symptoms, conduct problems, hyperactivity/inattention, peer problems, and prosocial behavior, with five items per subscale. Answers can be scored as *not true* (0), *somewhat true* (1), or *certainly true* (2); items 7, 11, 14, 21, and 25 are reverse coded. A total score is calculated by adding four subscale scores, excluding prosocial behavior. Higher scores indicate a higher risk of psychopathology. In addition to the total and subscale scoring, five items were summed to obtain the dysregulation profile score (Holtmann et al., 2011): *restlessness* (2), *worrying* (8), *fighting* (12), *sadness* (13), and *stealing* (22). The Dutch self-report SDQ was found to have satisfactory psychometric properties in a community sample of 562 children and adolescents, with support for its five-factor structure, acceptable internal consistency and test–retest stability, and good concurrent validity (Muris et al., 2003).

Psychiatric diagnoses

The diagnoses of CD, ODD, and ADHD were determined based on multi-informant assessment using the clinical interview outcomes of the Mini-International Neuropsychiatric Interview for Children (MINI-KID) and standardized assessment outcomes of the Achenbach System of Empirically Based Assessment (ASEBA) self- and parent-report instruments. By combining assessments from multiple informants, information about a wide range of disruptive behaviors and across different settings was obtained (Van Der Ende & Verhulst, 2005). An adolescent was diagnosed with ODD, CD, or ADHD if one of the three informants indicated that the psychiatric disorder was present. Finally, combining the multi-informant ODD, CD, and ADHD outcomes, a comorbid diagnosis (ODD/CD + ADHD) was set if ODD or CD was present with ADHD.

MINI-KID

The MINI-KID is a relatively short, semistructured clinical interview to determine psychopathology in children and adolescents (Sheehan et al., 2010); the Dutch translation was used by well-trained research assistants with a background in psychology, pedagogical sciences, or medicine (Bauhuis et al., 2013). It consists of 23 modules that correspond to DSM-IV diagnostic categories. The CD, ODD, and ADHD outcomes were scored dichotomously based on symptoms experienced in the past 6 months leading to significant dysfunction. The original MINI-KID showed substantial to excellent convergent validity, substantial sensitivity, substantial to excellent specificity, and good to excellent interrater and test–retest reliability (Sheehan et al., 2010).

ASEBA

The Youth Self Report form (YSR 11–18) and the Child Behavior Checklist (CBCL 6–18) of the ASEBA were used to measure ODD, CD, and ADHD problems over the past 6 months using the DSM-oriented scales (Achenbach et al., 2008). The ODD scale consists of five items in both the YSR and CBCL, the ADHD scale of seven items, and the CD scale has 15 items in the YSR and 17 in the CBCL. A diagnosis on ODD, CD, and ADHD was scored (yes/no) using the clinical cutoff, which differs per scale and by sex and represents scores in the 98–100th percentile of the norm group. The DSM-oriented scales have good internal consistency and good test–retest reliability (Achenbach et al., 2008).

Self-report early delinquency (SRED)

Delinquency in the past 6 months was measured using a Dutch adaptation (van der Laan et al., 2010) of the SRED (Moffitt & Silva, 1988). Adolescents were interviewed on how often they performed a wide range of antisocial and delinquent behaviors, such as stealing, vandalism, and violence: *never, once, two to three times, four to six times, seven times or more*. In accordance with van der Laan et al. (2010), items that did not concern criminal behavior were excluded (e.g., *truancy* and *substance use*). One more item (*hitting someone at home*) was removed because reliability analyses indicated that it did not fit the scale and that seemed theoretically plausible. It contained 23 items; item prevalences are presented in Table S1. A total score was calculated based on item seriousness and frequency, with higher scores indicating more delinquency. Additionally, we categorized different types of delinquency based on the Statistics Netherlands (2011) standard crime classification, resulting in violent, property, and vandalism and public order delinquency subscales. The primary delinquency outcome was binary, with the highest 10% total scorers categorized as serious delinquents and the rest as minor or nondelinquents. van der Laan et al. (2010) found that the SRED showed good internal consistency and construct validity.

MINI-PLUS

The MINI-PLUS, the adult version of the MINI-KID interview, assessed in Dutch, was used to determine SMI in accompanying parents (van Vliet & de Beurs, 2007). It consists of 26 modules that correspond to DSM-IV diagnostic categories. A form of SMI was considered present if there was any lifetime diagnosis in the following modules: depressive disorder, bipolar disorder, or psychotic disorder. The MINI-PLUS has good to very good sensitivity and specificity, and good inter-rater and test–retest reliability (Lecrubier et al., 1997).

Analytic strategy

First, sample characteristics, interrater agreement, and prevalences of outcomes are reported. Interrater agreement based on the dichotomous variables was assessed using Cohen's Kappa and interpreted as: ≤0 poor, 0.01–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect (Landis & Koch, 1977). Next, the SDQ internal consistency was assessed using Cronbach's α and the average inter-item correlation (IIC). Internal consistency was interpreted as Cronbach's α <.5 as unacceptable, ≥.5 and <.6 poor, ≥.6 and <.7 questionable, ≥.7 and <.8 acceptable, ≥.7 and <.8 good, and above .9 excellent (Mallery & George, 2003). IIC was considered adequate between .15 and .5, indicating related items but no redundancy (Clark & Watson, 1995). To study the predictive validity of the SDQ total score, subscale scores, and dysregulation profile, logistic regressions and receiver operating characteristic curves (ROC curves) were examined for psychiatric diagnoses and serious delinquency. Area under the curve (AUC) was interpreted as 0.7–0.8 acceptable, 0.8–0.9 excellent, and ≥0.9 outstanding (Hosmer et al., 2013); unpaired comparisons based on the method of DeLong were conducted (DeLong et al., 1988). For the continuous violent, property, and vandalism or public order delinquency outcome variables, linear regression was used. To investigate the predictive validity in adolescents of parents with and without SMI, main effect and interaction terms were added to the logistic regression models. Age and sex were considered potential covariates (Holling et al., 2008); therefore, sensitivity analyses were conducted in the same way as the SMI analyses. An α of .01 was used as a cutoff to adjust for multiple testing. SPSS V.25 and the R packages "MICE" and "pROC" were used for analyses (IBM Corp, 2017; Robin et al., 2011; Van Buuren & Groothuis-Oudshoorn, 2011).

Multiple imputation

Of all relevant variables, a maximum of 5.8% of values was missing. Adolescents' sex, age, and SDQ variables had no missing values. Measures of interest had the following amounts of missing values: demographics data (0.0%–11.7%), MINI-KID diagnoses (5.2%), MINI-PLUS diagnoses (10.7%), SRED (5.7%), YSR (5.2%–5.5%), and CBCL scores (11.9%). Missing values resulted mainly from declined interviews or unreturned questionnaires. We assumed these values were missing at random; that is missingness was assumed to be related to measured covariates and outcomes. For example, adolescents and parents with more social economical strain might have more unreturned questionnaires. For sum scores, when 75% of the items were valid, the average item score on valid items was multiplied by the number of items in the scale to estimate scores. Missing values after these calculations were handled by multiple imputation. Five imputed datasets were created under fully conditional specification (FCS) with 1000 iterations per chain. Scale variables were imputed based on predictive mean matching (PMM) and binary variables based on logistic regression. Outcome variables could be imputed because auxiliary variables were available: binary indicators of alcohol, smoking, and illicit drug use of the adolescent, brief symptoms inventory (BSI) subscale scores, and standardized assessment of personality—abbreviated scale (SAPAS) total score for parents. Regression coefficients were pooled across imputation sets automatically, based on Rubin's rules which take into account both within and between imputation variance (Rubin, 2004). For the X^2 tests, Nagelkerke's R^2, and AUC values, which were not pooled automatically, median values were reported (Marshall et al., 2009).

RESULTS

Sample characteristics

Data from 1022 adolescents were inspected; see Table 1 for the sample characteristics. Boys and girls participated equally. Adolescents' average age at baseline was 15.0 (SD = 0.9) years old, most were in pre-vocational education and had a Dutch cultural and geographical background. Accompanying parents were most often mothers, intermediate to highly educated, and of Dutch cultural and geographical background. In terms of demographics, the only difference between girls and boys was that girls were more often accompanied by their mother than boys (χ^2 (1) = 15.28, p < .001).

Agreement between raters for psychiatric outcomes

The inter-rater reliabilities between adolescent (YSR) and clinical interviewer (MINI-KID) were slight to fair for all outcomes (κ range: 0.16–0.29), fair for adolescent–parent (respectively YSR and CBCL, κ range: 0.16–0.23), and slight to fair for parent–interviewer (respectively, CBCL and MINI-KID, κ range: 0.09–0.20). Looking at agreement per outcome, ADHD and ODD agreements were fair (respectively, κ range: 0.16–0.29; κ range: 0.10–0.21) and CD agreement was slight to fair (κ range: 0.09–0.28).

TABLE 1 Sample characteristics of adolescents and accompanying parents by adolescent sex[a].

	Total sample $N = 1022$	Girls $n = 522$	Boys $n = 500$
Adolescent			
Age at baseline in years, *M* (± *SD*)	15.0 (± 0.9)	15.0 (± 0.9)	15.0 (± 0.9)
Education level			
Pre-vocational	48.7%	50.0%	47.3%
Secondary general	22.8%	24.4%	21.2%
Pre-university	19.3%	18.6%	20.0%
Undecided	9.2%	7.0%	11.5%
Cultural and geographic background			
Dutch	77.5%	77.8%	77.2%
Other-Western	6.0%	6.2%	5.8%
Non-Western	16.5%	16.0%	17.0%
Parent			
Age at baseline in years, *M* (± *SD*)	46.6 (± 5.7)	46.2 (± 5.8)	46.9 (± 5.5)
Sex, female	83.2%	87.8%	78.3%
Education level			
Lower	19.1%	18.9%	19.3%
Intermediate	37.0%	36.7%	37.4%
Higher	33.4%	33.6%	33.2%
Other	10.5%	10.8%	10.1%
Cultural and geographic background			
Dutch	76.4%	76.3%	76.5%
Other-Western	7.2%	7.8%	6.6%
Non-Western	16.4%	15.9%	17.0%

[a]Results based on non-imputed data.

TABLE 2 Prevalences of disruptive behavior by adolescent sex and parental SMI-status[a].

	Total sample, $n = 1022$ (%)	Girls, $n = 522$ (%)	Boys, $n = 500$ (%)	χ^2 test	Parents without SMI, $n = 680$ (%)	Parents with SMI, $n = 342$ (%)	χ^2 test
ODD	9.5	12.1	6.9	7.74*	7.9	12.9	7.72*
CD	8.2	5.7	10.8	8.81*	6.9	10.7	2.42
ADHD	27.2	27.8	26.5	0.30	24.2	33.0	8.73*
ODD/CD + ADHD	8.4	7.8	9.1	0.82	6.2	13.0	12.66**
Serious delinquency	9.9	5.3	14.6	25.27**	8.9	11.7	2.00

Abbreviation: SMI, severe mental illness.

[a]Results based on imputed data.

*$p < .01$, **$p < .001$.

Prevalences of disruptive behavior and comorbidity

Based on the pooled data, the prevalences of diagnoses in the current sample of high-risk adolescents were 9.5% ODD, 8.2% CD, 27.2% ADHD, 8.4% ODD/CD + ADHD, and 9.9% serious delinquents (resulting from the highest 10% operationalization); see Table 2. Most diagnoses resulted from the diagnostic interviews (8.6% ODD, 7.5% CD, 19.4% ADHD), next from self-report (1.2% ODD, 2.1% CD, 10.0% ADHD), and finally from parent-reported assessment (1.1% ODD, 0.5% CD, 7.9% ADHD). As expected, comorbidity was common. Adolescents with ODD often also had CD (37.9%) and the majority of adolescents diagnosed with ODD and CD had comorbid ADHD (respectively, 68.0% and 56.5%). In terms of sex differences, girls more often received a diagnosis of ODD than boys, while boys were more often classified than girls as having CD and conducting serious delinquency. Delinquent behavior was also more common in adolescents with disruptive behavior diagnoses, with a prevalence of 27.7% in those with ODD, 51.9% in those with CD, 19.8% in those with ADHD, and 35.3% in adolescents with ODD/CD + ADHD.

JOURNAL OF Research on Adolescence

SDQ scores and internal consistency

The average SDQ total score in the overall sample was 14.67 ($SD = 5.46$) and average subscale scores were as follows: emotional symptoms ($M = 3.93$, $SD = 2.47$), conduct problems (M = 2.47, $SD = 1.70$), hyperactivity/inattention ($M = 5.81$, $SD = 2.50$), peer problems ($M = 2.46$, $SD = 1.85$), and prosocial behavior ($M = 7.71$, $SD = 1.76$). Males had lower scores on the prosocial behavior and emotional problems scales (respectively, t (1020) = −8.74, $p < .001$ and t (1020) = −10.83, $p < .001$) and higher scores on the conduct problems scale than girls (t (1020) = 3.88, $p < .001$).

In this high-risk sample of adolescents, the SDQ total score had acceptable internal consistency ($\alpha = .66$). The subscales had poor to acceptable Cronbach's α's and acceptable IIC ($\alpha = .47$–$.76$; IIC = .15–.39), with the hyperactivity/inattention scale performing best and conduct problems performing worst. The dysregulation scoring method had unacceptable internal consistency ($\alpha = .32$; IIC = .08).

SDQ scores and disruptive behavior

Logistic regression was used to investigate associations of the SDQ total score, subscale scores, and dysregulation profile with disruptive behavior; see Table 3. In general, all scoring methods were predictive of all disruptive behavior outcomes. The total score was positively associated with all psychiatric diagnoses and serious delinquency, with the highest ORs for ODD and ADHD (respectively, OR = 1.18, $p < .001$, 95% CI [1.13, 1.24]; OR = 1.18, $p < .001$, 95% CI [1.14, 1.22]). These odd ratios indicate that for each point increase of the SDQ total score, the likelihood of the outcome, respectively, ODD and ADHD, was 1.18 times higher. The SDQ total score explains about 16% of the variance for ADHD and 11% for ODD, as indicated by the Nagelkerke's R^2 values. Next, looking at the SDQ subscales, the conduct problems subscale in particular was positively associated with all disruptive behavior outcomes, with the strongest OR for CD (OR = 1.63, $p < .001$, 95% CI [1.39, 1.90]). The hyperactivity/inattention scale was positively associated with all disruptive behavior outcomes except for CD and had a particularly high OR for ADHD (OR = 1.49, $p < .001$, 95% CI [1.37, 1.63]). The peer problems subscale predicted ADHD only. The emotional problems and prosocial behavior subscales were not associated with disruptive behavior outcomes. Finally, the dysregulation profile score was positively associated with all outcomes. However, the effect sizes of the dysregulation profile were low. For each outcome, the subscale model performed best and explained a considerable amount of variance in the outcomes (Nagelkerke's R^2 range: .13–.23).

Next, ROC curves were estimated for the three methods of scoring on each disruptive behavior outcome; for the subscale scoring method, the most relevant subscale was used; see median ROC curves and AUC statistics in Figure 1. AUCs were compared using the DeLong method. For ODD, there were no significant differences between the scoring methods in terms of AUC. For CD, the conduct problems AUC was higher than the AUCs for the total score ($p = .011$) and dysregulation profile ($p = .002$). For ADHD, the hyperactivity/inattention subscale model had a larger AUC than the dysregulation profile model ($p = .003$), but did not significantly differ from the total score. For ODD/CD + ADHD, the difference in AUC for the three curves did not reach statistical significance. For the serious delinquency outcome, the conduct problems subscale AUC was larger than the total score AUC ($p = .010$) and the dysregulation profile AUC ($p = .031$). In short, the subscale scoring method performed best in terms of AUC as well. AUCs were all acceptable or slightly below acceptable.

TABLE 3 Logistic regression results for the associations of SDQ on disruptive behavior[a].

	ODD	CD	ADHD	ODD/CD + ADHD	Serious delinquency
SDQ score	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
TOT	**1.18 (1.13–1.24)****	**1.11 (1.06–1.16)****	**1.18 (1.14–1.22)****	**1.15 (1.10–1.21)****	**1.10 (1.05–1.15)****
χ^2; R^2	55.4**; .11	19.7**; .04	118.9**; .16	41.0**; .09	21.7**; .04
ES	1.04 (0.94–1.14)	0.91 (0.82–1.02)	1.02 (0.95–1.09)	0.93 (0.83–1.03)	0.92 (0.84–1.02)
CP	**1.33 (1.16–1.53)****	**1.63 (1.39–1.90)****	**1.16 (1.05–1.28)***	**1.39 (1.20–1.60)****	**1.48 (1.29–1.70)****
HA	**1.27 (1.13–1.42)****	1.06 (0.95–1.19)	**1.49 (1.37–1.63)****	**1.32 (1.16–1.49)****	**1.15 (1.04–1.28)***
PP	1.18 (1.04–1.34)	1.03 (0.90–1.19)	**1.17 (1.07–1.28)***	1.18 (1.02–1.35)	0.98 (0.86–1.11)
PB	1.04 (0.91–1.18)	1.03 (0.90–1.18)	1.09 (0.99–1.20)	1.05 (0.91–1.21)	0.99 (0.87–1.13)
χ^2; R^2	70.72**; .14	67.64**; .15	179.65**; .23	71.83**; .16	67.01**; .13
DP	**1.44 (1.27–1.64)****	**1.29 (1.13–1.48)****	**1.45 (1.31–1.59)****	**1.36 (1.19–1.56)****	**1.36 (1.19–1.55)** **
χ^2; R^2	31.91**; .07	13.85**; .03	70.82**; .10	21.11**; .05	23.06**; .05

Note: Significant odds ratio in bold. *$p < .01$; **$p < .001$.

Abbreviations: CI, Confidence Interval; CP, Conduct Problems; DP, Dysregulation Profile; ES, Emotional Symptoms; HA, Hyperactivity/Inattention; OR, Odds Ratio; PB, Prosocial Behavior; PP, Peer Problems; R^2, Nagelkerke's R^2; TOT, Total Score.

[a]Results based on imputed data.

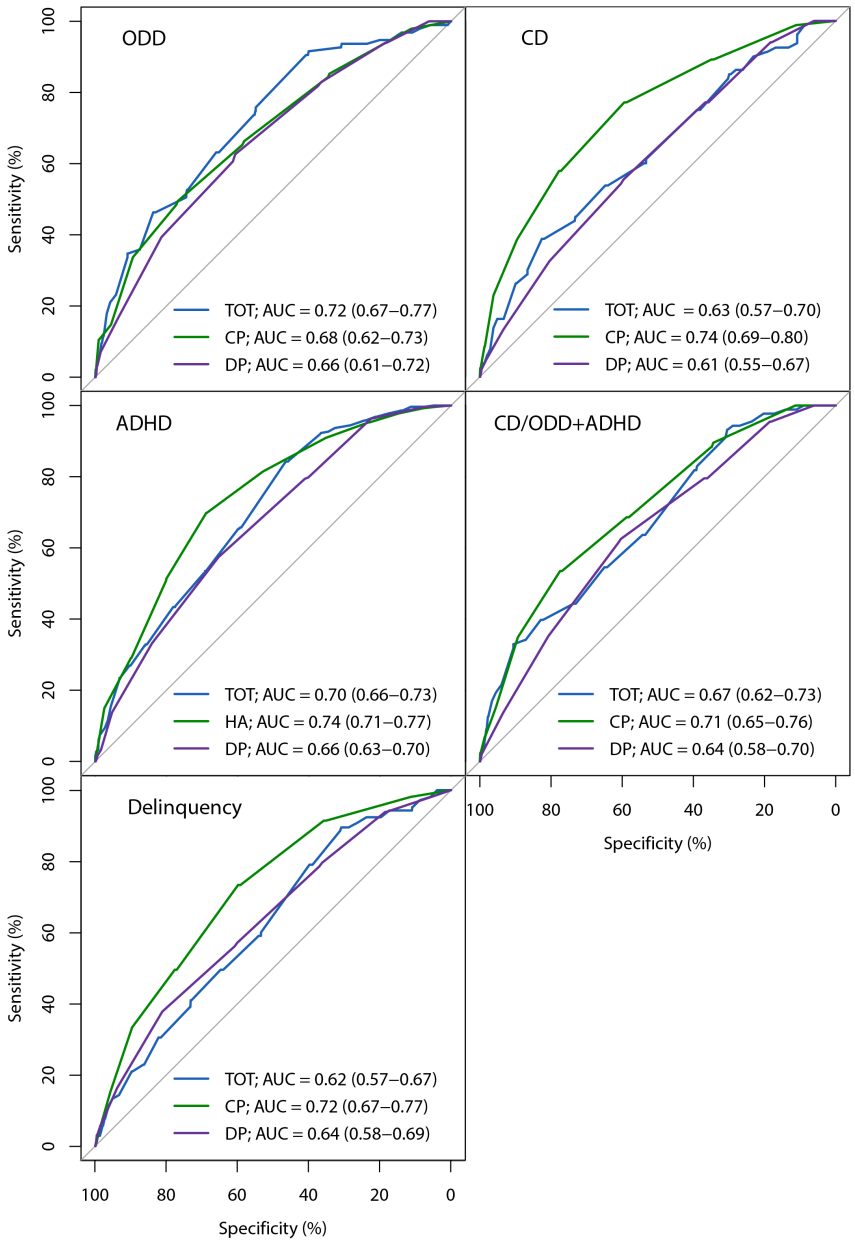

FIGURE 1 Median ROCs and AUCs with 95% confidence intervals for SDQ scoring method on disruptive behavior. Results based on imputed data. CP, Conduct Problems; DP, Dysregulation Profile; HA, Hyperactivity/Inattention; TOT, Total Score.

SDQ scores and types of delinquency

Simple linear regression was used to study associations of SDQ total score, subscale scores, and dysregulation profile with types of delinquency; see Table 4. The total SDQ score was associated with all types of delinquency: violent, property, and vandalism and public order. However, the predictive effects were small, as evidenced by small coefficients and R^2 values. Looking at the subscales, the conduct problems subscale was associated with all types of delinquency and explained more variance than the total score, ranging from 3% to 6%. The dysregulation profile score was associated with property and vandalism and public order delinquency, but not violent delinquency. Similar to the earlier models, the subscale scoring method had the best predictive validity; however, predictive effects were small for specific type of delinquency outcomes.

SDQ scores and disruptive behavior in adolescents of parents with SMI

About a third of all accompanying parents (33.5%) had experienced a form of SMI during their lifetime, with 30.8% reporting a depressive episode or disorder, 4.1% a bipolar disorder, and 4.9% a psychotic episode or disorder. In parents with a lifetime SMI, 15.3% reported multiple SMI diagnoses. The prevalences of ODD, ADHD, and ODD/CD + ADHD were higher in adolescents of parents with SMI than without SMI; see Table 2. The prevalences of CD and serious

JOURNAL OF Research on Adolescence

TABLE 4 Linear regression results of the associations of SDQ on types of delinquency[a].

	Delinquency type		
	Violent	Property	Vandalism/public order
SDQ score	*b* (95% CI)	*b* (95% CI)	*b* (95% CI)
TOT	**0.04 (0.01–0.06)***	**0.07 (0.03–0.10)****	**0.03 (0.01–0.05)****
F; R^2	8.47*; .01	15.88**; .02	12.67**; .01
ES	−0.04 (−0.11–0.04)	−0.04 (−0.12–0.05)	−0.04 (−0.08−−0.01)
CP	**0.21 (0.11–0.31)****	**0.30 (0.17–0.42)****	**0.19 (0.13–0.25)****
HA	−0.01 (−0.08–0.06)	0.06 (−0.02–0.15)	0.03 (−0.02–0.07)
PP	0.04 (−0.05–0.013)	−0.03 (−0.14–0.08)	−0.01 (−0.06–0.04)
PB	−0.06 (−0.16–0.05)	−0.10 (−0.23–0.03)	−0.03 (−0.09–0.03)
F; R^2	7.01 **; .03	10.77**; .05	13.64**; .06
DP	0.12 (0.03–0.21)	**0.22 (0.11–0.33)****	**0.10 (0.04–0.15)***
F; R^2	6.4; .01	15.6**; .02	11.25**; .01

Note: Significant coefficients in bold. *$p < .01$, **$p < .001$.

Abbreviations: CI, Confidence Interval; CP, Conduct Problems; DP, Dysregulation Profile; ES, Emotional Symptoms; HA, Hyperactivity/Inattention; PB, Prosocial Behavior; PP, Peer Problems; TOT, Total Score.

[a]Results based on imputed data.

delinquency were higher in parents with SMI than without SMI in absolute terms, but these differences did not reach statistical significance.

To investigate if predictive validity differed between adolescents of parents with and without SMI, the main effect of SMI and interaction terms with each subscale (e.g., of SMI status*conduct problems score) were added to the multiple logistic regression models presented earlier. Only the SDQ subscales were considered here, because this method performed best in the overall sample. None of the interaction terms were statistically significant, indicating that the ORs for the SDQ subscales did not differ between adolescents of parents with and without SMI on any of the disruptive behavior outcomes.

Sensitivity analyses by sex and age

In a similar manner as for parental SMI, differences in predictive validity based on adolescents' sex and age were investigated by adding interaction terms to the multiple regression models with the SDQ subscales. None of the interaction terms were significant, indicating that the ORs for the SDQ subscales did not differ by sex or age on any of the disruptive behavior outcomes.

DISCUSSION

The Dutch self-reported SDQ scores were related to disruptive behavior and delinquency measured on average 1.9 years after screening ($SD = 0.84$) in this high-risk sample of adolescents. The internal consistency of the total and subscale scores ranged from poor to acceptable. These findings, with the conduct disorder and peer problems' subscales showing the lowest internal consistency, are highly similar to those in Dutch and Belgian community samples (Muris et al., 2003; Van Leeuwen et al., 2006; Van Widenfelt et al., 2003; Vugteveen et al., 2021) and roughly similar to those of the original English self-report SDQ (Goodman, 2001). Poor Cronbach's alpha values could result from a low number of items, poor inter-relatedness between items or heterogeneous constructs (Tavakol & Dennick, 2011). Looking at the conduct problems subscale and considering it consists of only five items that target a wide range of behaviors (*getting angry, doing as one is told, stealing, fighting,* and *lying or cheating*), the lower alpha values make sense and likely underestimate the subscale reliability (Tavakol & Dennick, 2011). The conduct problems scale could also be influenced by social desirability bias, where adolescents may consider some behaviors (e.g., illegal behavior like *stealing*) less desirable to report than others, possibly resulting in less inter-item relatedness. Furthermore, Van Widenfelt et al. (2003) argue that a low level of awareness in adolescents regarding their conduct problems may cause underreporting and the authors suggest multi-informant measurement, using the parent or teacher version in addition to the adolescent version, as a potential solution. The prevalences of ODD, CD, ADHD, and comorbid ODD or CD with ADHD were higher in the current high-risk sample than would be expected in community samples (Canino et al., 2010; Sayal et al., 2018), but the internal consistency and predictive validity of the SDQ were similar. Overall, comparing the predictive validity of the total SDQ total score, subscale scores, and the dysregulation profile scale, the subscale scoring method performed best and showed the highest effect sizes and area under the curve values. As expected, the two most relevant subscales for disruptive behavior were conduct problems and hyperactivity/inattention. The conduct problems subscale predicted all disruptive behavior outcomes, while the hyperactivity/inattention subscale predicted all outcomes

except CD and specific types of delinquency. The AUC values for the conduct problems subscale predicting ODD and CD and the hyperactivity/inattention subscale predicting ADHD in this high-risk sample were similar to those found in a Dutch clinical sample of adolescents (Vugteveen et al., 2018); all subscale AUC values were considered acceptable except ODD which was slightly lower. In terms of delinquency, the SDQ total and subscale scoring methods predicted all types of delinquency. Furthermore, the conduct problems subscale was able to distinguish between serious delinquents and none or minor delinquents with similar predictive validity as for the psychiatric diagnoses. This is in contrast with some previous findings, where violent offending was not predicted by SDQ subscales (Colins et al., 2013) and where the emotional symptoms subscale, but not the behavioral subscale, was related to delinquency (Paalman et al., 2011). However, it is in line with findings that the conduct problems subscale was related to property offending (Colins et al., 2013) and vandalism (Lambie & Krynen, 2017). Although the self-report SDQ subscale scores were associated with delinquency subtypes, we found only weak effect sizes. These small effects may explain mixed and null-findings in previous studies with smaller sample sizes, and beg the question whether meaningful prediction of specific types of delinquency is possible.

The SDQ-DP scale in the current study lacked internal consistency, which makes sense given the wide range of cognitions and behaviors that the items measure. Regardless of this, higher dysregulation profile scores were related to all disruptive behavior outcomes except for violent delinquency. However, effect sizes were small even with severe problems such as the combined diagnosis of ODD or CD with ADHD and serious delinquency. It is possible that affect regulation is more effective in prediciting the severity of disorders and the amount of dysfunction caused by them, rather than predicting the presence or absence of diagnoses (Caro-Cañizares et al., 2017; Deutz et al., 2018). Or it may predict other problems better than disruptive behavior, even though two of five items measure disruptive behaviors, that is, *fighting* and *stealing*. Another consideration is whether the scale items are indeed the best fit, as they were chosen on discriminant analysis based on parent report in reference to the CBCL-dysregulation profile (Holtmann et al., 2011). It could be the case that other items in the self-reported SDQ are more informative of affective dysregulation. However, previously reported findings suggest invariance across the self-report and parent-reported SDQ-DP (Deutz et al., 2018). More research on the predictive validity of the SDQ-DP in predicting severity and dysfunction in adolescents is recommended.

There was no evidence of differences in predictive validity between adolescents of parents with and without SMI or by adolescent's sex or age. In line with expectations, disruptive behavior was more common in adolescents of parents with SMI than without SMI (Rasic et al., 2013; Siegenthaler et al., 2012; Thanhäuser et al., 2017). A limitation of our findings is that SMI of only one parent was assessed, mostly the mother. Taking into account psychopathology in both parents might lead to different findings. For now, results indicate that even in settings with multiple indicators of being at high risk, in this case adolescents at risk of psychopathology with parental SMI, the SDQ subscale scores are predictive of disruptive behavior outcomes. Findings that the SDQ subscales perform equally well across adolescents' sex and age are in line with most of the literature and norms for scoring (Achenbach et al., 2008; Goodman et al., 2003). Regardless, taking these factors into account seems sensible, because differential effects have been reported in some studies (Holling et al., 2008; Muris et al., 2003).

Strengths and limitations

The current study has several important strengths. First, the study provides new insights about the predictive validity of the self-report SDQ for disruptive behavior outcomes using longitudinal data from a large sample of high-risk adolescents. Where a lot of validation studies use cross-sectional data (e.g., Muris et al., 2003; Van Leeuwen et al., 2006; Van Widenfelt et al., 2003), this study provides more information over the predictive validity of the SDQ over a longer term by using follow-up data around 2 years after screening. Furthermore, the ability to predict delinquency was studied, for which literature has been scarce. The use of multi-informant measurement to assess psychiatric diagnoses is an important strength, because raters might complement each other's information and represent multiple settings (Van Der Ende & Verhulst, 2005). Considerable differences between informants in this study also indicated that this had clear added value and that diagnoses would have been missed if a single informant was used. Comorbidity between ODD, CD, and ADHD, which was highly prevalent, was taken into account specifically. Next, the clinical interviewers were blind to SDQ scores and selection status. Finally, missing data were handled with multiple imputation to prevent bias.

There were limitations as well. First, the disruptive behavior diagnoses were based on the positive result of at least one of three informants, which has merits in terms of considering different sources of information, but may have also lead to false positives. However, because most diagnoses resulted from the reliable and validated structured interview conducted by trained interviewers with a relevant educational background, the bias is assumed to be small. Furthermore, the most stringent cutoff scores (clinical) of the ASEBA instruments were used, indicating a score above 97% of the norm group and thus very serious behavior problems. Second, the design may be nonintuitive, because the SDQ was used to create a high-risk cohort in which its psychometric properties were then studied. A sample of adolescents at high risk for psychopathology could have been recruited in many different ways, for example, based on biological, social, or psychological risk factors (Arango et al., 2018). For the iBerry study, the SDQ was used and its psychometric properties were not a primary goal of this

cohort study. However, it is all the more efficient to use these gathered data to answer various research questions and even though this study has used the SDQ both for selection and for prediction, the results are still valid if we assume that the current sample is indeed at high risk for psychopathology. Which seems to be the case, judging from the high prevalences of mental disorders in both the adolescents and their parents (Grootendorst-van Mil et al., 2021). Third, adolescents' disruptive behavior disorders and delinquent behavior at the time of screening were unknown and therefore the predictive validity we discuss regards later measurement of disruptive behavior rather than later development of "new" disruptive behavior. This should be taken into account when interpreting the findings. Fourth, selection bias is a limitation that cannot be ruled out completely and adolescents at high risk of emotional and behavioral problems were slightly more likely to participate in this cohort than those at low risk (Grootendorst-van Mil et al., 2021). Another possible limitation is the use of self-reported delinquency data only, which is inherently different and may yield different conclusions than official police data (Farrington et al., 2003). Note that official data have its limitations as well, in that it only measures delinquency that has been detected and reported.

Future research and implications

Arguably, the most important area for future research and clinical application of this knowledge would be to investigate referral options for adolescents at risk for disruptive behavior. Evidence-based specialized treatments for disruptive behavior are available, in the form of school- or family-focused interventions (Dopp et al., 2017; Sawyer et al., 2015; von Sydow et al., 2013). However, in the Netherlands, where the current study was conducted, specialized child and adolescent mental healthcare settings are dealing with a high demand for treatment as evidenced by long wait-lists that have made national headlines (NOS, 2021). Furthermore, the SDQ is already used country-wide to screen adolescents in high-school settings as part of primary care, presenting a great opportunity for prevention. Targeted intervention efforts for adolescents with subclinical problems may prevent further development and continuity of problems and the need for referral to specialized mental healthcare. This could be a partial structural solution to lower the pressure on specialized child and adolescent mental healthcare. Digital interventions that use cognitive behavioral techniques may be particularly fitting for this age group (Bergin et al., 2020), but other effective and attractive interventions for adolescents should also be considered. More research on referral guidelines and fitting interventions for adolescents with disruptive behavior is therefore recommended. Future research could also look into the predictive ability of the SDQ for official arrests and convictions and thereby provide information which may be more relevant from a criminological and juvenile justice perspective. Finally, future studies could investigate the possible added value of other instruments, biological measurements, or tasks to improve the ability to predict disruptive behaviors. Adolescents at risk may benefit from further developing and implementing knowledge on early identification, selective prevention, and treatment for disruptive behavior.

ACKNOWLEDGEMENTS

The iBerry study is funded by the Erasmus University Medical Center and the following mental healthcare institutes: Parnassia Psychiatric Institute Antes, GGz Breburg, GGz Delfland, GGz Westelijk Noord-Brabant, and Yulius. All funding organizations participate in the Epidemiological and Social Psychiatric Research Institute (ESPRi), a consortium of academic and nonacademic research groups.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

CONSENT TO PARTICIPATE

All participants, and when necessary their parents or legal guardians, provided informed consent.

ORCID

Pascalle Spaan https://orcid.org/0000-0003-1482-8311

REFERENCES

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM* (5th ed.). American Psychiatric Association.

Arango, C., Díaz-Caneja, C. M., McGorry, P. D., Rapoport, J., Sommer, I. E., Vorstman, J. A., McDaid, D., Marín, O., Serrano-Drozdowskyj, E., Freedman, R., & Freedman, R. (2018). Preventive strategies for mental health. *The Lancet Psychiatry*, *5*(7), 591–604. https://doi.org/10.1016/S2215-0366(18)30057-9

Atilola, O., Ola, B., Abiri, G., & Adewuya, A. O. (2018). Correlations between psychopathology and self-reported quality of life among adolescents in youth correctional facilities in Lagos, Nigeria: A short report. *Criminal Behaviour and Mental Health*, *28*(1), 28–35. https://doi.org/10.1002/cbm.2042

Bailey, S., & Tarbuck, P. (2006). Recent advances in the development of screening tools for mental health in young offenders. *Current Opinion in Psychiatry*, *19*(4), 373–377. https://doi.org/10.1097/01.yco.0000228756.72366.de

Bauhuis, O., Jonker, K., Verdellen, C., Reynders, J., & Verbraak, M. (2013). [The introduction of a Dutch instrument to make DSM-IV-Tr-diagnoses in children] De introductie van een Nederlandstalig instrument om DSM-IV-Tr-diagnoses bij kinderen te stellen. *Kind & Adolescent Praktijk*, *12*(1), 20–26. https://doi.org/10.1007/s12454-013-0005-5

Bergin, A. D., Vallejos, E. P., Davies, E. B., Daley, D., Ford, T., Harold, G., Hetrick, S., Kidner, M., Long, Y., Merry, S., Morriss, R., Sayal, K., Sonuga-Barke, E., Robinson, J., Torous, J., & Merry, S. (2020). Preventive digital mental health interventions for children and young people: A review of the design and reporting of research. *NPJ Digital Medicine*, *3*(1), 133. https://doi.org/10.1038/s41746-020-00339-7

Byrd, A. L., Loeber, R., & Pardini, D. A. (2012). Understanding desisting and persisting forms of delinquency: The unique contributions of

disruptive behavior disorders and interpersonal callousness. *Journal of Child Psychology and Psychiatry*, *53*(4), 371–380. https://doi.org/10.1111/j.1469-7610.2011.02504.x

Canino, G., Polanczyk, G., Bauermeister, J. J., Rohde, L. A., & Frick, P. J. (2010). Does the prevalence of CD and ODD vary across cultures? *Social Psychiatry and Psychiatric Epidemiology*, *45*(7), 695–704. https://doi.org/10.1007/s00127-010-0242-y

Caro-Cañizares, I., Serrano-Drozdowskyj, E., Pfang, B., Baca-García, E., & Carballo, J. J. (2017). SDQ dysregulation profile and its relation to the severity of psychopathology and psychosocial functioning in a sample of children and adolescents with ADHD. *Journal of Attention Disorders*, *24*, 1557–1564. https://doi.org/10.1177/1087054717691829

Casswell, M., French, P., & Rogers, A. (2012). Distress, defiance or adaptation? A review paper of at-risk mental health states in young offenders. *Early Intervention in Psychiatry*, *6*(3), 219–228. https://doi.org/10.1111/j.1751-7893.2012.00344.x

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

Cohen, M. A., & Piquero, A. R. (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, *25*(1), 25–49. https://doi.org/10.1007/s10940-008-9057-3

Colins, O. F., Boonmann, C., Veenstra, J., van Domburgh, L., Buffing, F., Doreleijers, T. A., & Vermeiren, R. R. (2013). Mental health problems and recidivism among detained male adolescents from various ethnic origins. *European Child & Adolescent Psychiatry*, *22*(8), 481–490. https://doi.org/10.1007/s00787-013-0384-z

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, *44*(3), 837–845.

Deutz, M. H., Shi, Q., Vossen, H. G., Huijding, J., Prinzie, P., Deković, M., van Baar, A. L., & Woltering, S. (2018). Evaluation of the strengths and difficulties questionnaire-dysregulation profile (SDQ-DP). *Psychological Assessment*, *30*(9), 1174–1185. https://doi.org/10.1037/pas0000564

Dopp, A. R., Borduin, C. M., White, I., Mark, H., & Kuppens, S. (2017). Family-based treatments for serious juvenile offenders: A multilevel meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*(4), 335–354. https://doi.org/10.1037/ccp0000183

Farrington, D. P., Jolliffe, D., Hawkins, J. D., Catalano, R. F., Hill, K. G., & Kosterman, R. (2003). Comparing delinquency careers in court records and self-reports. *Criminology*, *41*(3), 933–958. https://doi.org/10.1111/j.1745-9125.2003.tb01009.x

Farrington, D. P., Ttofi, M. M., & Coid, J. W. (2009). Development of adolescence-limited, late-onset, and persistent offenders from age 8 to age 48. *Aggressive Behavior*, *35*(2), 150–163. https://doi.org/10.1002/ab.20296

Foster, E. M., Jones, D. E., & Conduct Problems Prevention Research Group. (2005). The high costs of aggression: Public expenditures resulting from conduct disorder. *American Journal of Public Health*, *95*(10), 1767–1772. https://doi.org/10.2105/AJPH.2004.061424

Goodman, A., & Goodman, R. (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/chi.0b013e3181985068

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. https://doi.org/10.1007/s10802-010-9434-x

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2003). Using the strengths and difficulties questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *International Review of Psychiatry*, *15*(1–2), 166–172. https://doi.org/10.1192/bjp.177.6.534

Grootendorst-van Mil, N. H., Bouter, D. C., Hoogendijk, W. J. G., van Jaarsveld, S., Tiemeier, H., Mulder, C. L., & Roza, S. J. (2021). The iBerry study: A longitudinal cohort study of adolescents at high risk of psychopathology. *European Journal of Epidemiology*, *36*(4), 453–464. https://doi.org/10.1007/s10654-021-00740-w

Holling, H., Kurth, B. M., Rothenberger, A., Becker, A., & Schlack, R. (2008). Assessing psychopathological problems of children and adolescents from 3 to 17 years in a nationwide representative sample: Results of the German health interview and examination survey for children and adolescents (KiGGS). *European Child & Adolescent Psychiatry*, *17*(Suppl 1), 34–41. https://doi.org/10.1007/s00787-008-1004-1

Holtmann, M., Becker, A., Banaschewski, T., Rothenberger, A., & Roessner, V. (2011). Psychometric validity of the strengths and difficulties questionnaire-dysregulation profile. *Psychopathology*, *44*(1), 53–59. https://doi.org/10.1159/000318164

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. *398*). John Wiley & Sons. https://doi.org/10.1002/9781118548387

IBM Corp. (2017). *IBM SPSS statistics for windows, version 25.0*. IBM Corp.

Keenan, K., Shaw, D., Delliquadri, E., Giovannelli, J., & Walsh, B. (1998). Evidence for the continuity of early problem behaviors: Application of a developmental model. *Journal of Abnormal Child Psychology*, *26*(6), 441–452. https://doi.org/10.1023/a:1022647717926

Kutcher, S., Aman, M., Brooks, S. J., Buitelaar, J., Van Daalen, E., Fegert, J., Findling, R. L., Fisman, S., Greenhill, L. L., Huss, M., Kusumakar, V., Pine, D., Taylor, E., & Huss, M. (2004). International consensus statement on attention-deficit/hyperactivity disorder (ADHD) and disruptive behaviour disorders (DBDs): Clinical implications and treatment practice suggestions. *European Neuropsychopharmacology*, *14*(1), 11–28. https://doi.org/10.1016/s0924-977x(03)00045-2

Lambie, I., & Krynen, A. (2017). The utility of the strengths and difficulties questionnaire as a screening measure among children and adolescents who light fires. *The Journal of Forensic Psychiatry & Psychology*, *28*(3), 313–330. https://doi.org/10.1080/14789949.2016.1275747

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., Janavs, J., & Dunbar, G. C. (1997). The MINI international neuropsychiatric interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, *12*(5), 224–231. https://doi.org/10.1016/S0924-9338(97)83296-8

Loeber, R., Burke, J. D., & Lahey, B. B. (2002). What are adolescent antecedents to antisocial personality disorder? *Criminal Behaviour and Mental Health*, *12*(1), 24–36. https://doi.org/10.1002/cbm.484

Loeber, R., & Farrington, D. P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and Psychopathology*, *12*(4), 737–762. https://doi.org/10.1017/S0954579400004107

Mallery, P., & George, D. (2003). *SPSS for windows step by step: A simple guide and reference* (Vol. *11.0 update*, 4th ed.). Allyn & Bacon.

Marshall, A., Altman, D. G., Holder, R. L., & Royston, P. (2009). Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Medical Research Methodology*, *9*(1), 57. https://doi.org/10.1186/1471-2288-9-57

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Moffitt, T. E., & Silva, P. A. (1988). Self-reported delinquency: Results from an instrument for New Zealand. *Australian and New Zealand Journal of Criminology*, *21*(4), 227–240. https://doi.org/10.1177/000486588802100405

Muris, P., Meesters, C., & van den Berg, F. (2003). The strengths and difficulties questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*(1), 1–8. https://doi.org/10.1007/s00787-003-0298-2

Nordström, T., Ebeling, H., Hurtig, T., Rodriguez, A., Savolainen, J., Moilanen, I., & Taanila, A. (2013). Comorbidity of disruptive behavioral disorders and attention-deficit hyperactivity disorder—Indicator of severity in problematic behavior? *Nordic Journal of Psychiatry*, *67*(4), 240–248. https://doi.org/10.3109/08039488.2012.731431

NOS. (2021). *[Child in need waits on average 10 months for help] Kind in nood wacht gemiddeld tien maanden op hulp*. NOS. https://nos.nl/artikel/2383242-kind-in-nood-wacht-gemiddeld-tien-maanden-op-hulp

Paalman, C. H., van Domburgh, L., Stevens, G. W., & Doreleijers, T. A. (2011). Individual, family and offence characteristics of high risk childhood offenders: Comparing non-offending, one-time offending and re-offending Dutch-Moroccan migrant children in The Netherlands. *Child and Adolescent Psychiatry and Mental Health*, *5*(1), 33. https://doi.org/10.1186/1753-2000-5-33

Rasic, D., Hajek, T., Alda, M., & Uher, R. (2013). Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: A meta-analysis of family high-risk studies. *Schizophrenia Bulletin*, *40*(1), 28–38. https://doi.org/10.1093/schbul/sbt114

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. https://doi.org/10.1186/1471-2105-12-77

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons. https://doi.org/10.1002/9780470316696

Sawyer, A. M., Borduin, C. M., & Dopp, A. R. (2015). Long-term effects of prevention and treatment on youth antisocial behavior: A meta-analysis. *Clinical Psychology Review*, *42*, 130–144. https://doi.org/10.1016/j.cpr.2015.06.009

Sayal, K., Prasad, V., Daley, D., Ford, T., & Coghill, D. (2018). ADHD in children and young people: Prevalence, care pathways, and service provision. *The Lancet Psychiatry*, *5*(2), 175–186. https://doi.org/10.1016/s2215-0366(17)30167-0

Sheehan, D. V., Sheehan, K. H., Shytle, R. D., Janavs, J., Bannon, Y., Rogers, J. E., Milo, K. M., Stock, S. L., & Wilkinson, B. (2010). Reliability and validity of the mini international neuropsychiatric interview for children and adolescents (MINI-KID). *The Journal of Clinical Psychiatry*, *71*, 313–326. https://doi.org/10.4088/jcp.09m05305whi

Siegenthaler, E., Munder, T., & Egger, M. (2012). Effect of preventive interventions in mentally ill parents on the mental health of the offspring: Systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(1), 8–17.e18. https://doi.org/10.1016/j.jaac.2011.10.018

Statistics Netherlands. (2011). *[Standard Classification for Offences 2010] Standaardclassificatie misdrijven 2010*. https://www.cbs.nl/nl-nl/onzediensten/methoden/classificaties/misdrijven/standaardclassificatie-misdrijven-2010/

Stronks, K., Kulu-Glasgow, I., & Agyemang, C. (2009). The utility of 'country of birth' for the classification of ethnic groups in health research: The Dutch experience. *Ethnicity & Health*, *14*(3), 255–269. https://doi.org/10.1080/13557850802509206

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thanhäuser, M., Lemmer, G., de Girolamo, G., & Christiansen, H. (2017). Do preventive interventions for children of mentally ill parents work? Results of a systematic review and meta-analysis. *Current Opinion in Psychiatry*, *30*(4), 283–299. https://doi.org/10.1097/yco.0000000000000342

Tremblay, R. E. (2010). Developmental origins of disruptive behaviour problems: The 'original sin' hypothesis, epigenetics and their consequences for prevention. *Journal of Child Psychology and Psychiatry*, *51*(4), 341–367. https://doi.org/10.1111/j.1469-7610.2010.02211.x

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(1), 1–67. https://doi.org/10.18637/jss.v045.i03

Van Der Ende, J., & Verhulst, F. C. (2005). Informant, gender and age differences in ratings of adolescent problem behaviour. *European Child & Adolescent Psychiatry*, *14*(3), 117–126. https://doi.org/10.1007/s00787-005-0438-y

van der Laan, A. M., Veenstra, R., Bogaerts, S., Verhulst, F. C., & Ormel, J. (2010). Serious, minor, and non-delinquents in early adolescence: The impact of cumulative risk and promotive factors. The TRAILS study. *Journal of Abnormal Child Psychology*, *38*(3), 339–351. https://doi.org/10.1007/s10802-009-9368-3

Van Leeuwen, K., Meerschaert, T., Bosmans, G., De Medts, L., & Braet, C. (2006). The strengths and difficulties questionnaire in a community sample of young children in Flanders. *European Journal of Psychological Assessment*, *22*(3), 189–197. https://doi.org/10.1027/1015-5759.22.3.189

van Vliet, I. M., & de Beurs, E. (2007). [The MINI-international neuropsychiatric interview. A brief structured diagnostic psychiatric interview for DSM-IV en ICD-10 psychiatric disorders] Het MINI Internationaal Neuropsychiatrisch interview (MINI). Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV- en ICD-10-stoornissen. *Tijdschrift voor Psychiatrie*, *49*(6), 393–397.

Van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D., & Goodman, R. (2003). Dutch version of the strengths and difficulties questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*(6), 281–289. https://doi.org/10.1007/s00787-003-0341-3

von Sydow, K., Retzlaff, R., Beher, S., Haun, M. W., & Schweitzer, J. (2013). The efficacy of systemic therapy for childhood and adolescent externalizing disorders: A systematic review of 47 RCT. *Family Process*, *52*(4), 576–618. https://doi.org/10.1111/famp.12047

Vugteveen, J., De Bildt, A., Hartman, C., & Timmerman, M. (2018). Using the Dutch multi-informant strengths and difficulties questionnaire (SDQ) to predict adolescent psychiatric diagnoses. *European Child & Adolescent Psychiatry*, *1-13*, 1347–1359. https://doi.org/10.1007/s00787-018-1127-y

Vugteveen, J., de Bildt, A., Serra, M., de Wolff, M. S., & Timmerman, M. E. (2020). Psychometric properties of the Dutch strengths and difficulties questionnaire (SDQ) in adolescent community and clinical populations. *Assessment*, *27*(7), 1476–1489. https://doi.org/10.1177/1073191118804082

Vugteveen, J., de Bildt, A., Theunissen, M., Reijneveld, S. A., & Timmerman, M. (2021). Validity aspects of the strengths and difficulties questionnaire (SDQ) adolescent self-report and parent-report versions among Dutch adolescents. *Assessment*, *28*(2), 601–616. https://doi.org/10.1177/1073191119858416

Whitten, T., Laurens, K. R., Tzoumakis, S., Kaggodaarachchi, S., Green, M. J., Harris, F., Carr, V. J., & Dean, K. (2019). The influence of parental offending on the continuity and discontinuity of children's internalizing and externalizing difficulties from early to middle childhood. *Social Psychiatry and Psychiatric Epidemiology*, *54*, 965–975. https://doi.org/10.1007/s00127-019-01670-5

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Spaan, P., van den Boogert, F., Grootendorst-van Mil, N. H., Hoogendijk, W. J. G., & Roza, S. J. (2023). Screening for disruptive behavior in adolescents at risk using the strengths and difficulties questionnaire. *Journal of Research on Adolescence*, *00*, 1–13. https://doi.org/10.1111/jora.12858