



# The Strengths and Difficulties Questionnaire self-report-, parent-, and teacher version in children with intellectual and developmental disabilities

Sabine Kaiser<sup>a,\*</sup>, Marianne Berg Halvorsen<sup>b</sup>

<sup>a</sup> Regional Center for Child and Youth Mental Health - North, UiT The Arctic University of Norway, Tromsø, Norway

<sup>b</sup> Department of Pediatric Rehabilitation, University Hospital of North Norway, Tromsø, Norway

## ARTICLE INFO

No. of reviews completed is 2

### Keywords:

Strengths and Difficulties Questionnaire  
SDQ  
Intellectual and developmental disability  
Psychometric properties  
Factorial validity

## ABSTRACT

**Background:** The Strengths and Difficulties Questionnaire (SDQ) is a frequently used behavioral screening instrument. However, its psychometric properties have been rarely examined among children with intellectual and developmental disabilities (IDD).

**Aims:** The main aims of this study were to examine the internal consistency (i.e., McDonald's Omega), the convergent validity (by correlating the Total difficulties score with the Aberrant Behavior Checklist [ABC]), the divergent validity (by correlating the Total difficulties score with the Vineland Adaptive Behavior Composite; VABS-II Total) and the factorial validity (by the means of confirmatory factor analyses [CFA]) of the SDQ self-report-, parent-, and teacher version in a sample of children with IDD.

**Method:** Participants were 365 children and adolescents (males  $n = 238$ ; 65 %) aged 4–18 years ( $M = 10.11$ ,  $SD = 3.82$ ) referred for a developmental/neurological assessment to the neuro-pediatric outpatient clinics in the specialist health services. The SDQ was filled in by 115 children, 337 parents, and 248 teachers.

**Results:** McDonald's Omega was overall lowest for the self-report version. Correlations of the SDQ Total difficulties score and the ABC subscales were strongest for the parent version. The results of the CFA indicated best model fit for the six-factor model that included a method factor for all three versions of the SDQ, however, model fit was overall not good.

**Conclusions:** Further research that examines the psychometric properties of the SDQ among multiple informants in large samples of children with IDD is needed.

## 1. Introduction

Studies have reported higher emotional and behavioral problems in children with intellectual and developmental disabilities (IDD) compared to typically developing children (Einfeld et al., 2011; Emerson, 2005; Kaptein et al., 2008). Mental health problems result in reduced functioning and increased need for help in everyday life at home and at school, in addition to the difficulties due to IDD, and are associated with reduced wellbeing for the child and the family (Halvorsen, Mathiassen et al., 2019). Early identification of these challenges is important to facilitate the right intervention and to prevent a chronic development of mental health problems. Psychometrically sound measurement instruments are necessary so that children that need help can be identified at an early stage.

\* Corresponding author.

E-mail addresses: [sabine.kaiser@uit.no](mailto:sabine.kaiser@uit.no) (S. Kaiser), [marianne.berg.halvorsen@unn.no](mailto:marianne.berg.halvorsen@unn.no) (M.B. Halvorsen).

<https://doi.org/10.1016/j.ridd.2022.104194>

Received 18 August 2021; Received in revised form 27 January 2022; Accepted 4 February 2022

Available online 8 February 2022

0891-4222/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

There are few instruments that have been designed for children and adolescents with IDD (e.g., the Aberrant Behavior Checklist [ABC; Aman & Singh, 1986] or the Developmental Behavior Checklist; Einfeld & Tonge, 1992), and accordingly instruments not originally designed for this population are commonly used (e.g., the Achenbach System of Empirically Based Assessment [ASEBA] or the Strengths and Difficulties Questionnaire [SDQ; Goodman, 1997]). An advantage with the SDQ and the ASEBA, that were designed for the general child population is, that they offer a youth self-report form compared to the IDD instruments that are proxy based. However, there is a need for more knowledge of the suitability of the non-IDD instruments among children and adolescents with IDD with accompanying communication difficulties, atypical symptom presentation, and symptom overlap between IDD and mental health problems (Flynn et al., 2017; Halvorsen et al., 2022).

The SDQ is a behavioral screening questionnaire that can be filled in by parents and teachers as well as by children aged 11 years or older (Goodman, 1997). The SDQ measures strengths and difficulties on five scales called Emotional problems, Conduct, Hyperactivity/inattention, Peer problems, and Prosocial behaviors. In addition, a Total difficulties score can be generated by adding up the four problem scales Emotional problems, Conduct, Hyperactivity/inattention, and Peer problems.

Several studies have examined the SDQ in typically developing children (e.g., Dickey & Blumberg, 2004; McCrory & Layte, 2012; Palmieri & Smith, 2007). Dickey and Blumberg (2004) identified a three-factor model for the SDQ parent version in the USA. The three factors were an externalizing factor, an internalizing factor, and a positive construal factor consisting of eight of the ten positive worded items.

Palmieri and Smith (2007) examined three models of the SDQ parent version in the USA. The first model was a five-factor higher order model, where the four problem scales loaded on the higher order factor called Total difficulties scale with which the Prosocial factor was correlated. The second model was a five-factor model, where the five scales that comprise the SDQ were correlated. The third model was a six-factor model, including the five scales that comprise the SDQ in addition to a method factor for items that are positive worded. Although all three models fitted the data well, the third model, which included the method factor, fitted the data best. Palmieri and Smith (2007) reported furthermore that Cronbach's alpha was good for four of the five SDQ scales in addition to the Total Difficulties scale ( $\alpha$  ranged from .71 to .88), and except for the Peer Problems scale ( $\alpha = .62$ ).

McCrory and Layte (2012) examined four models of the SDQ parent version in Ireland: A three factor model related to Dickey and Blumberg (2004) findings, where the items from the emotional problems and from the peer problem scale loaded on the internalising factor, an externalising factor with the items from the hyperactivity/inattention and conduct problems scales, and a prosocial factor consisting of the Prosocial behavior items. The remaining three models that were tested by McCrory and Layte (2012) were in accordance to the study from Palmieri and Smith (2007). Also, McCrory and Layte (2012) found that the six-factor model fit the data best.

However, few studies have examined the psychometric properties of the SDQ in samples of children with IDD (Embregts et al., 2010; Emerson, 2005; Haynes et al., 2013; Kaptein et al., 2008). Validating instruments in a specific population is important to be able to make informed and the right decisions where the clinician can also consider the strengths and weaknesses of the instrument.

Embregts et al. (2010) used the SDQ self-report, parent, and teacher version in a small sample of children with mild intellectual disability (ID) to screen for behavioral problems. Cronbach's alpha was the lowest for the self-report version (range for the different scales  $\alpha = .11-.58$ ) followed by the parent version (range for the different scales  $\alpha = .64-.83$ ) and the teacher version of the SDQ (range for the different scales  $\alpha = .72-.85$ ).

Emerson (2005) used the SDQ self-report, parent, and teacher version in a representative sample of children and divided the sample into a small sample of children with and into a large sample without ID. Cronbach's alpha for the self-report version was lowest for children with ID (range for the different scales  $\alpha = .30-.71$ ) compared to children without ID (range  $\alpha = .41-.78$ ). When comparing the correlations for the typically developed children between the SDQ scales from the self-report, parent and teacher versions, there were higher correlations between the self-report and the parent version than between those two versions and the teacher version on most scales, respectively. There was a more diverse picture when comparing the correlations between the three versions for children with ID. The self-report and the parent version of the SDQ were highest correlated on the Emotional problems, the Conduct problems, and the Peer problems scales. Emerson (2005) found furthermore that SDQ self-reported difficulties were related to relevant ICD-10 diagnoses.

Kaptein et al. (2008) found that Cronbach's alpha was slightly lower in children without ID (range  $\alpha$  from .55 – .80) compared to children with ID (range  $\alpha$  from .58 – .84) for the scales of the SDQ parent version.

Only one study has conducted factor analyses in a sample of children with ID (Haynes et al., 2013) and two studies in samples of children with learning difficulties (Bryant et al., 2020; El-Keshky & Emam, 2015). Haynes et al. (2013) conducted first Exploratory Factor Analyses and found that a three-factor model had the best fit for the SDQ self-report. The three factors were named Negative behavior, Positive relationships, and Emotional competence and are not the same as Dickey and Blumberg (2004) three factor model. Confirmatory Factor Analysis (CFA) of their three-factor model found that the chi-squared ( $\chi^2$ ) degrees of freedom (df) ratio was good ( $\chi^2/df = 1.58$ ) and the Root Mean Square Error of Approximation (RMSEA) was acceptable (RMSEA = .068). However, other fit indices were not reported, which makes it difficult to fully evaluate model fit. Haynes et al. (2013) reports a Cronbach's alpha of .68 for the Total difficulties scale. El-Keshky and Emam (2015) did a cross-cultural examination of the SDQ in two Arab countries. They examined the SDQ teacher version and concluded that a three-factor model fit the data best in comparison to a five-factor model. Bryant et al. (2020) conducted a CFA of the five-factor model of the SDQ parent version and concluded that model fit was acceptable ( $\chi^2/df = 3.80$ ; RMSEA = .061; CFI = 0.98). Furthermore, they found that SDQ scores were related to mental health problems and suggested that "the SDQ could be a valuable screening tool for identifying existing mental health difficulties in children recognized as struggling, as it can be in typically developing children and those with specific diagnoses" (Bryant et al., 2020, p. 1).

### 1.1. Study aim

The main aims of this study were to examine the internal consistency, the convergent-, divergent-, and the factorial validity of the SDQ self-report-, parent-, and teacher version in a sample of children with IDD. Internal consistency was examined using McDonald's omega. The convergent validity was examined by correlating the Total difficulties score of the SDQ self-report-, parent-, and teacher version with another behavioral rating scale (i.e., the ABC) that is frequently used among children with IDD (Halvorsen, Aman et al., 2019). The divergent validity was examined by correlating the Total difficulties score of the SDQ self-report-, parent-, and teacher version with the Vineland Adaptive Behavior Scale (i.e., VABS-II Total; the Adaptive behavior Composite; Sparrow et al., 2011). The factorial validity of the SDQ self-report-, parent-, and teacher version was evaluated by conducting CFA of a three-factor, a five-factor, a five-factor second order model, and a model with six factors including a method factor for the positive worded items in accordance to McCrory and Layte (2012). In addition, means and standard deviations for the five scales of the SDQ and the Total difficulties score were reported as well as correlations between the different scales of the SDQ self-report-, parent-, and teacher version.

## 2. Method

### 2.1. Participants and procedure

Participants were 365 children and adolescents (males  $n = 238$ ; 65 % and females  $n = 127$ ; 35 %) aged 4–18 years ( $M = 10.11$ ,  $SD = 3.82$ ) referred for a developmental/neurological assessment to the neuropsychiatric outpatient clinics in the specialist health services at the University Hospital of North Norway and the Finnmark Hospital Trust, in the northern parts of Norway. Assessments in the clinics were interdisciplinary and included specialist such as pediatricians, neuropsychologists, special education therapist, and physiotherapists (for more information see Halvorsen, Mathiassen et al. (2019)). The most frequent IDD in the sample were in descending order: specific developmental disorder (33.2 %), ID (21.1 %), other diseases of the nervous system such as cerebral palsy and epilepsy (17.2 %), Autism Spectrum Disorder (ASD; 15.3 %), Attention Deficit Hyperactivity Disorder (ADHD; 13.7 %), and congenital malformations and chromosomal abnormalities (12.6 %). The diagnoses were not mutually exclusive, so a given participant could have more than one diagnosis.

The SDQ for self-completion was filled in by 115 11–17-year-old, while about 337 parents filled in the SDQ for parents of 4–17-year-old, and about 248 teachers the SDQ for the teachers of 4–17-year-old. The mean age of all children and adolescents was 10.11 ( $SD = 3.82$ ) and the majority were males ( $n = 238$ ; 65 %). The mean of the Full-scale intelligence quotient was 76.22 ( $SD = 17.19$ ). More specifically, the mean age of the 115 children and adolescents that filled in the SDQ-self-report was 13.66 years ( $SD = 2.06$ ) and 69 were male (60 %). The mean of the Full-scale intelligence quotient was 79.74 ( $SD = 16.52$ ). The study was approved by the appropriate ethics committee. Children and parents were informed about the aim of the study. Written informed consent was obtained before the participants were included in the study. For participants who were younger than 12 years, their parents gave consent. For participants who were between 12 and 18 years, written consent was obtained from both the parents and the adolescents. The data protection officer at UNN and Finnmark Hospital Trust approved the use of de-identified data for research purposes.

### 2.2. Instruments

#### 2.2.1. The Strengths and Difficulties Questionnaire

Three versions of the Norwegian SDQ were filled in: The SDQ for self-completion, the SDQ for parents, and the SDQ for teachers. The SDQ for self-completion was filled in by adolescents age 11 years or older and the other two version by parents and teachers of 4–17 years old, respectively. What these three versions of the SDQ have in common is that they consist of 25 items distributed on five scales each; Emotional problems (five items, e.g., for the SDQ self-completion “I worry a lot” and for the SDQ parent version “Many worries or often seems worried”), Conduct problems (five items, e.g., for the SDQ self-completion “I fight a lot. I can make other people do what I want” and for the SDQ parent version “Often fights with other youth or bullies them”), Hyperactivity/inattention (five items, e.g., for the SDQ self-completion “I am restless, I cannot stay still for long” and for the SDQ parent version “Restless, overactive, cannot stay still for long”), Peer problems (five items, e.g., for the SDQ self-completion “I would rather be alone than with people of my age” and for the SDQ parent version “Would rather be alone than with other youth”), and Prosocial behaviors (five items, e.g., for the SDQ self-completion “I try to be nice to other people. I care about their feelings” and for the SDQ parent version “Considerate of other people's feelings”). In addition, there is the Total difficulties scale consisting of the 20 items that comprise the Emotional problems, Conduct problems, Hyperactivity/inattention, and Peer problems scales. The items are rated on a 3-point scale (0 = Not true, 1 = Somewhat true, and 2 = Certainly true).

The psychometric properties of the Norwegian SDQ for self-completion (Kornør & Heyerdahl, 2013), for parents (Kornør & Heyerdahl, 2017), and for teachers (Kornør & Heyerdahl, 2014) have been examined in in several studies as summarized by the beforementioned reviews. Internal consistencies was found to be highest for the teacher version (range of Cronbach's alpha from .64 – .88 for the five scales; Kornør & Heyerdahl, 2014), followed by the parent version (range of Cronbach's alpha from .45 – .82 for the five scales; Kornør & Heyerdahl, 2017), and the self-report version (range of Cronbach's alpha from .44 – .76 for the five scales; Kornør & Heyerdahl, 2013). Construct validity of the three instruments is supported through CFAs and expected group differences (Kornør & Heyerdahl, 2013, 2014, 2017). The Norwegian SDQ for self-completion, parent-, and teacher version have no national norms but regional ones (Kornør & Heyerdahl, 2013, 2014, 2017). For the SDQ self-completion in Northern Norway and the SDQ Total difficulties scale, a score between 0–14 is considered the normal range, between 15–17 borderline, and between 18–40 clinical range for 11- to

16-year-old (Rønning et al., 2004). For the SDQ parent version and the SDQ Total difficulties scale, a score between 0–10 is considered the normal range, between 11–13 borderline, and between 14–40 clinical range for 10- to 13-year-old in Akershus (Van Roy et al., 2010). The SDQ teacher version has only regional norm data for under 9-year-old (Kornør & Heyerdahl, 2014).

### 2.2.2. The Aberrant Behavior Checklist

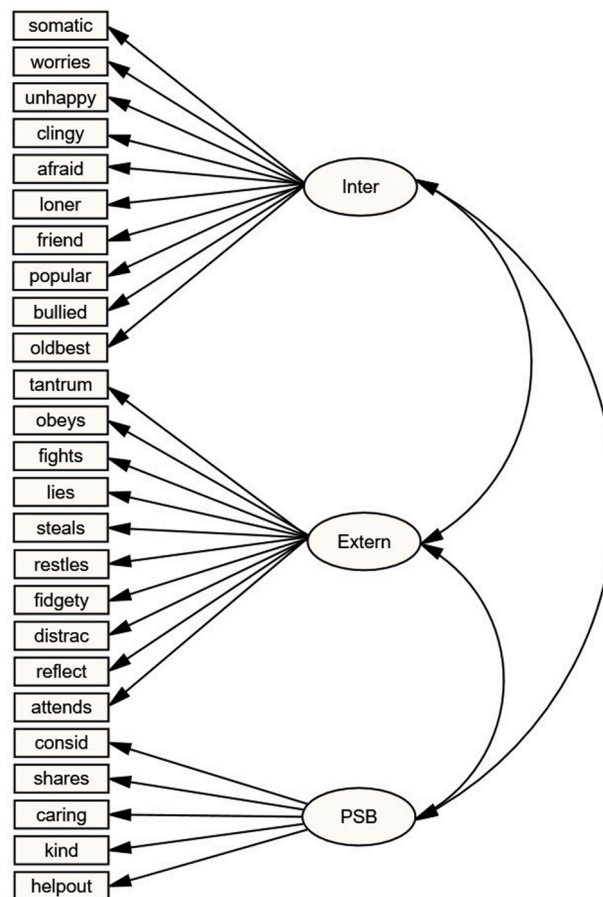
The ABC is a behavior rating scale that assess behavior problems in children and adults with IDD and was filled in by the parents (Aman & Singh, 1986). The questionnaire consists of 58 items and five subscales, which are Irritability (consisting of 15 items), Passivity and Social Withdrawal (consisting of 16 items), Stereotype Behavior (consisting of 7 items), Hyperactivity (consisting of 16 items), and Inappropriate Speech (consisting of 4 items). Each item is rated on a 4-point scale from 0 = “no problem” to 3 = “serious problem”. The Norwegian version of the ABC has been found to have satisfactory psychometric properties (Halvorsen, Aman et al., 2019).

### 2.2.3. The Vineland Adaptive Behavior Scales 2nd edition

The Vineland Adaptive Behavior Scales 2nd edition (VABS-II; Sparrow et al., 2011) is a semi-structured interview and was used to assess the adaptive level of functioning in communication, daily living skills, socialization, and motor skills. The total score was used (VABS-II Total, which is also called the Adaptive Behavior Composite) in the present study. The psychometric properties of the Norwegian version of the VABS-II have been described in the manual using a representative Scandinavian sample of 1673 parents of 2–21 years old indicating good psychometric properties (Sparrow et al., 2011). Construct validity of the instruments is among others supported through CFAs. Split-half reliability was satisfactory for the VABS-II Total (Sparrow et al., 2011).

### 2.2.4. Intellectual function

The standardized Wechsler Intelligence Tests (Wechsler, 2007, 2008a, 2008b, 2009, 2012) appropriate for their ages were used to assess the Full-scale intelligence quotient (FSIQ). For a small number of children ( $n = 14$ ), the Raven’s Colored Progressive Matrices (Raven, 2004) was used. The psychometric properties of the Norwegian versions of Wechsler Intelligence tests have been described in



**Fig. 1.** The three-factor model of the SDQ consisting of the internalizing factor (Inter), the externalising factor (Extern), and the prosocial factor (PSB).

the manuals. The Norwegian version of the WISC-IV was adjusted to Norwegian context and has Norwegian norms. The norms were based on a Norwegian standardization selection of 418 children age 6 years to 16 years and 11 months. The Norwegian version has proven to have good reliability, with coefficients between 0.86 and 0.98 (Wechsler, 2009). The validity of the Norwegian version examined by confirmatory factor analysis showed that the original WISC-IV factor structure was transferable to the Norwegian version (Wechsler, 2009).

### 2.3. Statistical analyses

SPSS 26 and 28 were used to calculate the descriptive statistics like means and standard deviations in addition to correlations, McDonald's Omega, and McDonald's Omega if item deleted (Hayes & Coutts, 2020). Information about participants who scored in the clinical range of the SDQ Total difficulties scale were only calculated for the self-report version, as this was the only version providing appropriate norms for the current sample (Kornør & Heyerdahl, 2013). Internal consistency coefficients of > .70 are considered adequate (European Federation of Psychologists' Association (EFPA), 2013).

A total of twelve CFAs were conducted with MPlus using the robust weighted least square (WLSMV) estimator. A three-factor, a five-factor, a five-factor second order model, and a model with six factors including a method factor for the positive worded items in accordance to McCrory and Layte (2012) were evaluated for the SDQ self-report, parent, and teacher version, respectively. The three-factor model consisted of the internalizing factor, which comprised the five emotional symptoms items and the five peer problem items, an externalising factor, which comprised the five hyperactivity/inattention items and the five conduct problems items, in addition to the prosocial factor with its five items (Fig. 1). The three factors were correlated. The five-factor model comprised the Emotional problems, Conduct problems, Hyperactivity/inattention, Peer problems, and Prosocial behaviors scale, which were correlated (Fig. 2). For the five-factor second order model the four problem scales loaded on another latent variable called SDQ Total difficulties scale, which in turn was correlated with the Prosocial behaviors scale (Fig. 3). In the last model, the five-factor model was extended with a sixth factor on which the positive worded items loaded (Fig. 4). The five SDQ scales were correlated, while the method factor was uncorrelated with the other factors. For all models, except for the six-factor model, the positive worded items were recoded.

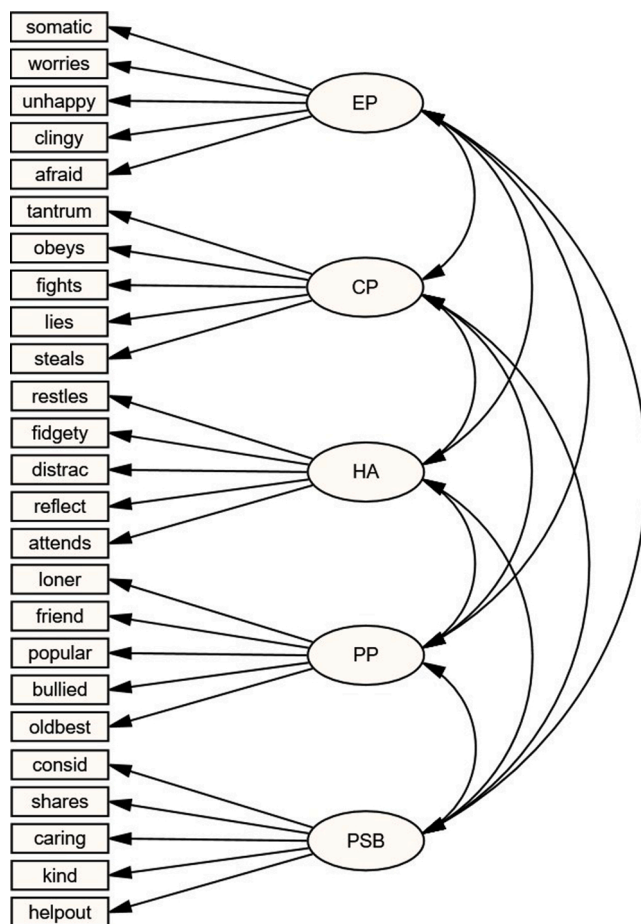


Fig. 2. The five-factor model of the SDQ comprise the Emotional problems (EP), Conduct problems (CP), Hyperactivity. inattention (HA), Peer problems (PP). and the Prosocial behaviors (PSB) scales.

For the data from the self-report version one case was deleted due to missing values on most on the SDQ items. The parent version had no missing values. The teacher version had less than 1 % missing values on 11 variables and less than 2 % on one variable. In the one case where a person had two missing items those were on different scales and the person was therefore not excluded from the analyses. One case had nine missing values and was deleted. Model fit was evaluated by using the  $\chi^2$ /degrees of freedom ratio ( $\chi^2$ /df), the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), and the RMSEA. Acceptable threshold levels are 3.00 or 2.00 for the  $\chi^2$ /df ratio, greater values than .95 for the CFI and TLI, and smaller values than .07 or .06 for the RMSEA (Hooper et al., 2007).

### 3. Results

Table 1 presents the means, standard deviations, correlations, and McDonald’s Omegas for the five scales and for the three versions of the SDQ. Overall, the highest scores on the Total difficulties scale are reported by parents ( $M = 14.72, SD = 6.98$ ) followed by teachers ( $M = 14.31, SD = 6.24$ ), and children ( $M = 13.19, SD = 6.23$ ). Of the 115 children and adolescents who filled in the SDQ-self-report, 69 (60.0 %) scored in the normal range (i.e.,  $< 15$ ), 17 (14.8 %) scored in the borderline range (i.e., between 15–17), and 29 (25.2 %) scored in the clinical range (i.e., 18 or higher) on the SDQ Total difficulties scale.

The correlations between the SDQ self-report and the parent version range from .41 (Hyperactivity/inattention scale) to .62 (Emotional problems scale), from .39 (Conduct problems scale) to .49 (Peer problems scale) between the parent and the teacher version, and from .29 (Hyperactivity/inattention scale) to .49 (Emotional problems scale) between the self-report and the teacher version. McDonald’s Omegas for the five subscales varied from .60 to .81 for the self-report version, from .66 to .80 for the parent version, and from .65 to .82 for the teacher version. Table 2 presents the McDonald’s Omegas if item deleted for the five scales of the SDQ for the self-report-, the parent-, and the teacher versions. For the Hyperactivity/inattention scale McDonald’s Omega could be improved from .64 to .74 by deleting the item “Restless” from the self-report version.

Table 3 presents among others the correlation of the SDQ Total difficulties scale with the five ABC subscales for the self-report, the parent, and the teacher version. These correlations were positive and significant, except for the self-report version of the SDQ and the ABC Social Withdrawal- and Inappropriate Speech subscales, respectively. Correlations between the SDQ Total difficulties scale and the VABS-II Total score were negative, not significant for the self-report version ( $r = -.11, p > .05$ ), and significant for the parent ( $r = -.41, p < .001$ ) and the teacher version ( $r = -.26, p < .001$ ).

Table 4 presents the fit indices from the CFAs for the SDQ self-report-, parent-, and the teacher version for the four different models

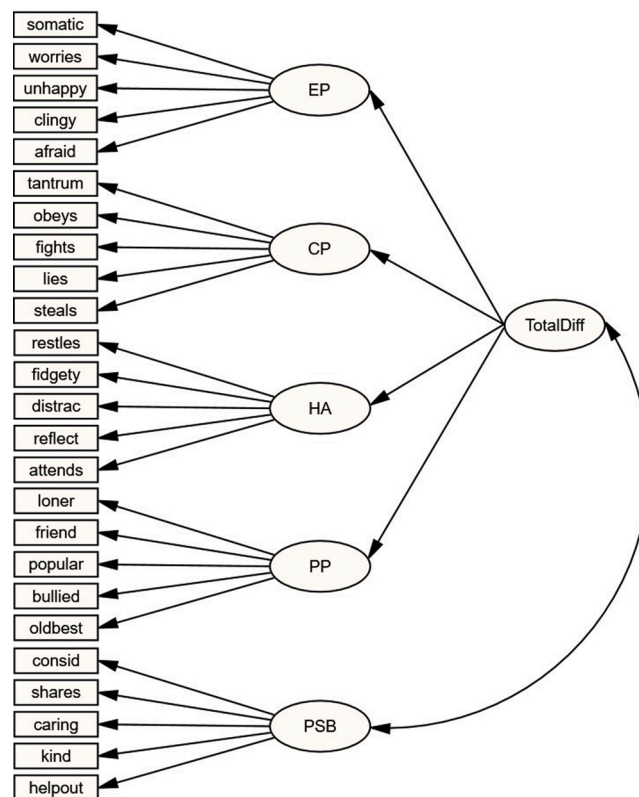


Fig. 3. The five-factor second order model of the SDQ comprises the Emotional problems (EP), Conduct problems(CP), Hyperactivity/inattention (HA), Peer problems (PP) scales, which load on a latent variables called SDQ Total difficulties scales (TotalDiff), which correlated with the prosocial behaviours (PSB) scales.

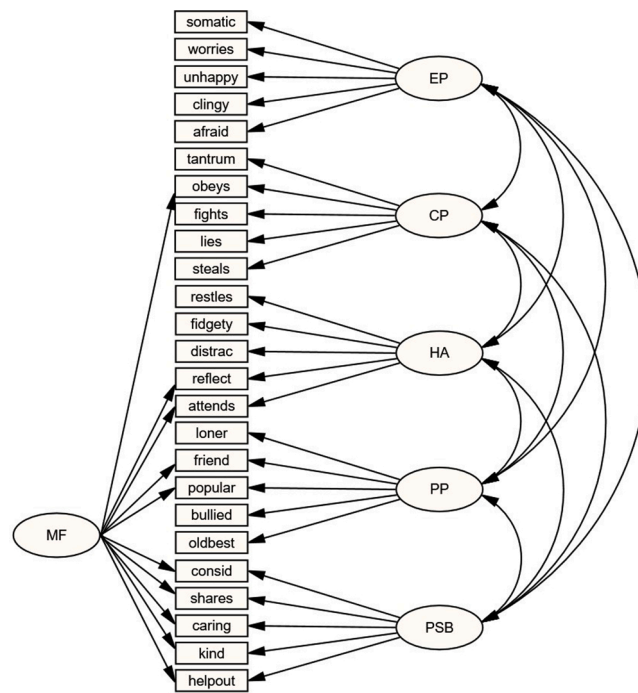


Fig. 4. The five-factor model Emotional problems (EP), Conduct problems (CP), Hyperactivity/inattention (HA), Peer problems (PP), and the Prosocial behaviors (PSB) scales is extended with a sixth factor on which the positive worded items load (MF).

Table 1

Mean (M), Standard Deviation (SD), Correlations (r), and McDonald’s Omega for the SDQ Self-Report (n = 115), Parent (n = 318 – 337), and Teacher Version (n = 233 – 249).

SDQ scale	M(SD)			Correlation			McDonald’s Omega		
	Self-report	Parent	Teacher	S vs. P <sup>a</sup> r	P vs. T <sup>b</sup> r	T vs. S <sup>c</sup> r	Self-report	Parent	Teacher
Total difficulties scale	13.19(6.23)	14.72(6.98)	14.31(6.24)	.52***	.42***	.32**	–	–	–
Emotional problems	3.46(2.70)	3.40(2.62)	2.85(2.29)	.62***	.48***	.49***	.81	.76	.65
Conduct problems	2.16(1.72)	2.13(2.02)	1.81(1.83)	.53***	.39***	.36**	.60	.66	.70
Hyperactivity/inattention	4.20(2.28)	5.32(2.59)	5.71(2.62)	.41***	.46***	.29**	.64	.77	.77
Peer problems	3.37(2.11)	3.87(2.50)	3.94(2.50)	.62***	.49***	.37**	.60	.72	.73
Prosocial behaviors	7.42(1.98)	7.08(2.36)	5.64(2.66)	.52***	.43***	.43***	.70	.80	.82

Note. S = SDQ self-report; P = SDQ parent version; T = SDQ teacher version. Correlation is significant at \*p < .05, \*\*p < .01 and \*\*\*p < .001.

- <sup>a</sup> n = 110.
- <sup>b</sup> n = 238.
- <sup>c</sup> n = 79.

of the SDQ, respectively. Model fit was best for the six-factor model which included a method factor for positively worded items in all three models. The parent version got the best model fit when looking at the  $\chi^2/df$  ratio (2.32), the CFI (.93), TLI (.92), and RMSEA (.062).

The standardized factor loadings for the SDQ six-factor model for the self-report-, the parent-, and the teacher version are presented in Table 5. All items were significant, except for the two items Friend and Attends that loaded on the Method factor ( $p > .05$ ) for the parent- and the teacher version. The correlations between the latent factors for the three different versions of the SDQ six-factor model are presented in Tables 6 and 7. All correlations were in the expected direction.

#### 4. Discussion

The SDQ, originally developed for the general child population, offers a youth self-report form, and is frequently used in the IDD population to assess general mental health problems. Few studies have examined the suitability of the different SDQ forms in general, and particularly the self-report form in this population (Embregts et al., 2010; Emerson, 2005; Haynes et al., 2013; Kaptein et al., 2008).

In the current study, internal consistency of the SDQ subscales was lowest for the self-report version (McDonald’s Omega < .70 on

**Table 2**

McDonald's Omega if Item Deleted for the Five Scales of the SDQ Self-Report- ( $n = 115$ ), the Parent- ( $n = 337$ ), and the Teacher ( $n = 249$ ) Version.

	McDonald's Omega if Item Deleted		
	SDQ Self-Report	SDQ parent version	SDQ teacher version
Emotional problems			
Somatic	.82	.77	.69
Worries	.73	.69	.54
Unhappy	.77	.70	.58
Clingy	.76	.73	.63
Afraid	.77	.71	.63
Conduct problems			
Tantrum	.55	.68	.61
Obeys <sup>a</sup>	.59	.60	.66
Fights	.53	.60	.63
Lies	.55	.63	.68
Steals	.54	.62	.70
Hyperactivity/inattention			
Restless	.74	.71	.70
Fidgety	.53	.69	.70
Distractible	.61	.77	.76
Reflective <sup>a</sup>	.57	.73	.76
Attends <sup>a</sup>	.74	.78	.70
Peer problems			
Loner	.55	.61	.64
Friend <sup>a</sup>	.51	.70	.65
Popular <sup>a</sup>	.59	.70	.70
Bullied	.59	.73	.74
Oldest	.51	.63	.66
Prosocial behaviors			
Considerate <sup>a</sup>	.58	.75	.77
Shares <sup>a</sup>	.71	.79	.80
Caring <sup>a</sup>	.61	.71	.76
Kind <sup>a</sup>	.66	.78	.81
Helpout <sup>a</sup>	.69	.79	.82

Note. <sup>a</sup>Items are positively worded.

**Table 3**

Mean ( $M$ ), Standard Deviations ( $SD$ ), and Correlations between the SDQ Total Difficulties Scale for the Self-Report ( $n = 107$ - $115$ ), Parent ( $n = 301$ - $337$ ), and Teacher ( $n = 222$ - $249$ ) Version with Clinical Variables.

Variable	$M(SD)$	SDQ Total difficulties scale		
		Self-report	Parent	Teacher
VABS-II Total	67.10(15.16)	-.11	-.41***	-.26***
ABC Irritability	6.11(7.29)	.38***	.60***	.34***
ABC Social Withdrawal	5.73(6.22)	.18	.45***	.19**
ABC Stereotypic Behavior	1.53(2.73)	.34***	.30***	.17*
ABC Hyperactivity/Noncompliance	9.70(9.84)	.26**	.60***	.39***
ABC Inappropriate Speech	1.72(2.26)	.17	.48***	.27***

Note. ABC = Aberrant Behavior Checklist; VABS-II Total = Vineland Adaptive Behavior Scale 2<sup>nd</sup> Edition Total Score.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

three subscales) and about similar between the parent and the teacher versions. This finding is comparable to previous studies like Embregts et al. (2010) who found internal consistency, as measured by Cronbach's alpha, was the lowest for the self-report version followed by the parent version and the teacher version of the SDQ. Internal consistency could be improved from inadequate (internal consistency coefficient  $< .70$ ) to adequate (internal consistency coefficient  $> .70$ ; European Federation of Psychologists' Association (EFPA, 2013) for the Hyperactivity/inattention scale by deleting the item "Restless" in the self-report version. The other changes in McDonald's Omegas when item deleted were not as drastic.

In relation to construct validity (i.e., the correlations between instruments assessing similar constructs), there were meaningful correlations between the SDQ Total difficulties score and the ABC subscales, which is an IDD instrument. Furthermore, correlations between the SDQ Total difficulties score and the parent version were medium to large, small to medium for teachers, and from not significant on two scales to medium strong for the self-report version of the SDQ. With regard to divergent validity, correlations between the SDQ total difficulties score and the VABS-II Total score (i.e., the Adaptive behavior Composite score) were negative, not significant for the self-report version but negative and significant for the parent version and the teacher version. In general, lower



**Table 4**Fit Indices from the CFA for the SDQ Testing Different Models for the Self-Report- ( $n = 115$ ), the Parent- ( $n = 337$ ), and the Teacher ( $n = 249$ ) Version.

Version	$\chi^2$	df	$\chi^2/df$	CFI	TLI	RMSEA (90 % CI)
<b>Self-report</b>						
3-factors	507.25***	272	1.86	.71	.68	.087 (.075, .098)
5-factors	412.04***	265	1.55	.81	.78	.072 (.058, .084)
2 <sup>nd</sup> order factor	454.28***	270	1.68	.77	.75	.077 (.065, .089)
6-factor model (1 method factor)	361.75***	255	1.42	.87	.85	.060 (.045, .074)
<b>Parent</b>						
3-factors	1023.84***	272	3.76	.84	.83	.091 (.085, .097)
5-factors	664.72***	265	2.51	.92	.90	.067 (.061, .073)
2 <sup>nd</sup> order factor	704.06***	270	2.61	.91	.90	.069 (.063, .075)
6-factor model (1 method factor)	590.42***	255	2.32	.93	.92	.062 (.056, .069)
<b>Teacher</b>						
3-factors	1082.90***	272	3.98	.81	.79	.109 (.103, .116)
5-factors	715.08 ***	265	2.70	.89	.88	.083 (.075, .090)
2 <sup>nd</sup> order factor	743.05***	270	2.75	.89	.88	.084 (.077, .091)
6-factor model (1 method factor) <sup>a</sup>	613.39***	256	2.40	.92	.90	.075 (.067, .082)

Note. df = Degrees of Freedom; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation; 90 % CI = 90 % Confidence Interval.

<sup>a</sup> Because of model convergence problems, the first factor loading of each factor was freed, the factor variances fixed to 1, and the correlation between Prosocial behaviors and Conduct problems was constrained to  $-0.7$ .

**Table 5**Standardized Factor Loadings from the CFA for the SDQ Six-Factor Model for the Self-Report- ( $n = 115$ ), the Parent- ( $n = 337$ ), and the Teacher ( $n = 249$ ) Version.

	Factor loading		
	Self-report	Parent	Teacher
<b>Emotional problems</b>			
Somatic	.48	.34	.44
Worries	.84	.76	.64
Unhappy	.79	.88	.83
Clingy	.83	.73	.80
Afraid	.78	.73	.68
<b>Conduct problems</b>			
Tantrum	.94	.74	.85
Obeys <sup>a</sup>	-.35 (.54)	-.65 (.37)	-.74 (.46)
Fights	.88	.72	.80
Lies	.54	.76	.70
Steals	.75	.91	.60
<b>Hyperactivity/inattention</b>			
Restless	-.45	-.86	.92
Fidgety	-.44	-.90	.93
Distractible	-.82	-.87	.86
Reflective <sup>a</sup>	.44 (.49)	.60 (.31)	-.50 (.54)
Attends <sup>a</sup>	.45 (.29)	.73 (.13)	-.84 (.08)
<b>Peer problems</b>			
Loner	.62	.78	.60
Friend <sup>a</sup>	-.50 (.44)	-.47 (.12)	-.64 (.15)
Popular <sup>a</sup>	-.35 (.61)	-.69 (.22)	-.84 (.39)
Bullied	.74	.54	.57
Oldbest	.68	.82	.67
<b>Prosocial behaviors</b>			
Considerate <sup>a</sup>	.65 (.70)	.62 (.54)	.74 (.55)
Shares <sup>a</sup>	.35 (.28)	.61 (.38)	.74 (.36)
Caring <sup>a</sup>	.44 (.51)	.48 (.80)	.45 (.76)
Kind <sup>a</sup>	.54 (.51)	.66 (.40)	.73 (.34)
Helpout <sup>a</sup>	.36 (.43)	.36 (.61)	.30 (.63)

Note. The following items were not significant at  $p < .001$ : The items Obeys ( $p < .01$ ), Popular ( $p < .01$ ), Shares ( $p < .01$ ), and three items (Friend ( $p < .01$ ), Attends ( $p < .05$ ), and Shares ( $p < .05$ )) from the Method factor for the self-report version; Three items (Friend ( $p > .05$ ), Popular ( $p < .01$ ), and Attends ( $p > .05$ )) from the Method factor for the parent version. For the teacher version the item Helpout from the Prosocial behaviors factor was significant at  $p < .01$ , and the two items Friend and Attends from the Method factor were not significant ( $p > .05$ ).

<sup>a</sup> Items are positively worded and load additionally on the method factor. The standardized factor loadings for the method factor are presented in the parentheses.

**Table 6**

Correlations between the Latent Factors of the SDQ Six-Factor Model for the Self-Report Version (presented below the diagonal) and the Parent Version (presented above the Diagonal).

	1.	2.	3.	4.	5.
1 Emotional problems	–	.37***	–.36***	.51***	–.30**
2 Conduct problems	.28**	–	–.62***	.55***	–.84***
3 Hyperactivity/inattention	–.50***	–.54***	–	–.44***	.45***
4 Peer problems	.54***	.49***	–.47***	–	–.65***
5 Prosocial behaviors	.09	–.76***	.01	–.41**	–

Note. The Method factor was not correlated with the other five-factors by syntax-command.

\* $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .

**Table 7**

Correlations between the Latent Factors of the SDQ Six-Factor Model for the Teacher Version.

	1.	2.	3.	4.	5.
1 Emotional problems	–				
2 Conduct problems	.27***	–			
3 Hyperactivity/inattention	.17*	.58***	–		
4 Peer problems	.44***	.55***	.27***	–	
5 Prosocial behaviors	–.13	–.70 <sup>a</sup>	–.48***	–.60***	–

Note. The Method factor was not correlated with the other five-factors by syntax-command.

<sup>a</sup> Correlation was fixed to  $-.70$ .

\*  $p < .05$ .

\*\*\*  $p < .001$ .

correlations were expected as the degree of overlap between the SDQ total and the VABS-II Total do not measure identical constructs.

The results of the CFA indicated that none of the models or versions of the SDQ tested had an especially good fit. However, the six-factor model, which included a method factor, had the best fit for all three versions of the SDQ: For the self-report version the  $\chi^2/df$  ratio and the RMSEA were under the recommended cut-off values but the Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) were smaller than  $< .90$ . The six-factor model for the parent version had the highest CFI and TLI but they were under the recommended cut-off value of  $.95$ , while the  $\chi^2/df$  ratio and the Root Mean Square Error of Approximation were slightly higher compared to the self-report version in absolute values.

Also, in the other studies, conducted on samples with typically developed children, the six-factor model fitted the data best (McCrary & Layte, 2012; Palmieri & Smith, 2007). Dickey and Blumberg (2004) stated in relation to the construal factor that the “likelihood that this factor represents a methodological artifact is increasingly strong” (p. 1165). However, they did not recommend deleting the positive worded items from the SDQ because this “could result in decreased acceptability by parents and a decrease in the reporting of problem behaviors” (p. 1165). McCrary and Layte (2012) concluded “that the existence of method effects does not present any great threat to the structural validity of the instrument” (p. 882).

The current study identified a good concurrence between the self-report and the parent versions when correlating the different scales of the SDQ. Correlations were higher between those two versions of the SDQ than compared to the correlations between parent or self-report and teacher version of the SDQ, respectively. Similar findings have been reported previously (Heyerdal, 2003). Heyerdal (2003) emphasized using multiple informants to get a full picture of a child’s mental health. They assess information on a different basis and the correlation between their assessments is an “expression of correspondence between different perceptions of a phenomenon” (p. 130).

#### 4.1. Clinical implications

Non-IDD instruments like the SDQ have generally been more thoroughly researched than tools developed for children and adolescents with IDD, and mental health professionals in general are more likely to be familiar with them. However, findings from the present study indicate that among relatively high-functioning IDD adolescents (mean FSIQ = 79.74,  $SD = 16.52$ ) the reliability (internal consistency) and construct validity of in particularly the self-report version was compromised. In the current absence of broad band self-report instruments developed for the IDD adolescents, the SDQ can be somewhat useful in an incipient assessment as it enables the young person to participate in their own mental health assessment. Participation and co-determination are central in mental health assessment. Little or no participation from patients themselves is in itself a weakness, and may reduce the reliability and validity of a diagnosis.

## 4.2. Strengths and limitations

The current study is one of few that examined the internal consistency, convergent validity, and factorial validity of the SDQ self-report-, parent-, and teacher version in a relatively large sample of children with IDD (Embregts et al., 2010; Emerson, 2005; Haynes et al., 2013). However, for the CFA of the self-report version the sample size was rather small, and it would be desirable to repeat the analyses with a larger sample. Also, the sample recruited among patients undergoing a neurodevelopmental/neurological assessment in the specialist health service was not representative of the general IDD population, limiting the generalizability of the findings. The high correlations between the SDQ Total difficulties scale and the ABC scales can also be related to the fact, that the ABC was filled in by the parents of the children.

As mentioned under Table 3 the correlation between the two scales Prosocial behaviors and Conduct problems was constrained to -0.70 because of model convergence problems, when estimating the six-factor model for the teacher version. The correlation was chosen based on the correlation of the five-factor model between Prosocial behaviors and Conduct problems scales which was -.78 for the teacher version. It has to be taken into account that a lower correlation between the two factors would have worsened model fit. We also had to free the first factor loadings of each factor and to fix the factor variances to 1 for the model to run. However, this should not influence model fit.

## 4.3. Conclusion

This study contributed to the field of mental health assessment among children and adolescents with IDD by examining the psychometric properties of the SDQ self-report-, parent-, and teacher versions. Overall, our results suggested more support for the parent version in terms of evidence of reliability and construct validity. Further development of the usage of self-reporting that meet youths' with IDD's level of communication to enable them to participate in their own mental health examination, will be an important development area for the field.

## Funding

This study was supported by the Northern Norway Regional Health Authority (grant number PFP1178-14). The study was also supported in part by The Norwegian Directorate of Health and the Regional Centre for Child and Youth Mental Health and Child Welfare North at UiT The Arctic University of Norway, Tromsø, Norway. The authors would like to thank the participants who contributed to the dataset.

## Role of the funding source

The funding sources had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the article; and in the decision to submit the article for publication.

## CRediT authorship contribution statement

MBH designed the study, conducted the data acquisition, wrote part of the Methods section, and revised the manuscript critically. SK conducted the analyses and drafted the article. All authors approved the final version of the manuscript.

## Data availability

The data underlying this paper cannot be shared because it contains patient information.

## Declaration of Competing Interest

The authors report no declarations of interest.

## References

- Aman, M. G., & Singh, N. N. (1986). *Aberrant behavior checklist manual*. Educational Publishing.
- Bryant, A., Guy, J., The CALM Team, Holmes, J., Astle, D., Baker, K., ... Zhang, M. (2020). The strengths and difficulties questionnaire predicts concurrent mental health difficulties in a transdiagnostic sample of struggling learners. *Frontiers in Psychology, 11*(3125). <https://doi.org/10.3389/fpsyg.2020.587821>
- Dickey, W. C., & Blumberg, S. J. (2004). Revisiting the factor structure of the strengths and difficulties questionnaire: United States, 2001. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(9), 1159–1167. <https://doi.org/10.1097/01.chi.0000132808.36708.a9>
- Einfeld, S. L., & Tonge, B. J. (1992). *Manual for the developmental behaviour checklist*. Monash University for Developmental Psychiatry and School of Psychiatry, University of New South Wales.
- Einfeld, S. L., Ellis, L. A., & Emerson, E. (2011). Comorbidity of intellectual disability and mental disorder in children and adolescents: A systematic review. *Journal of Intellectual & Developmental Disability, 36*(2), 137–143. <https://doi.org/10.1080/13668250.2011.572548>
- El-Keshky, M., & Emam, M. (2015). Emotional and behavioural difficulties in children referred for learning disabilities from two Arab countries: A cross-cultural examination of the Strengths and Difficulties Questionnaire. *Research in Developmental Disabilities, 36*, 459–469. <https://doi.org/10.1016/j.ridd.2014.10.039>
- Embregts, P. J. C. M., du Bois, M. G., & Graef, N. (2010). Behavior problems in children with mild intellectual disabilities: An initial step towards prevention. *Research in Developmental Disabilities, 31*(6), 1398–1403. <https://doi.org/10.1016/j.ridd.2010.06.020>

- Emerson, E. (2005). Use of the Strengths and Difficulties Questionnaire to assess the mental health needs of children and adolescents with intellectual disabilities. *Journal of Intellectual & Developmental Disability*, 30(1), 14–23. <https://doi.org/10.1080/13668250500033169>
- European Federation of Psychologists' Association (EFPA). (2013). *EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers, v 4.2.6*.
- Flynn, S., Vereenoghe, L., Hastings, R. P., Adams, D., Cooper, S. A., Gore, N., Hatton, C., Hood, K., Jahoda, A., Langdon, P. E., McNamara, R., Oliver, C., Roy, A., Totsika, V., & Waite, J. (2017). Measurement tools for mental health problems and mental well-being in people with severe or profound intellectual disabilities: A systematic review. *Clinical Psychology Review*, 57, 32–44. <https://doi.org/10.1016/j.cpr.2017.08.006>
- 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>
- Halvorsen, M. B., Helverschaou, S. B., Axelsdottir, B., Brøndbo, P. H., & Martinussen, M. (2022). General measurement tools for assessing mental health problems among children and adolescents with an intellectual disability: A systematic review. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-021-05419-5>
- Halvorsen, M., Aman, M. G., Mathiassen, B., Brøndbo, P. H., Steinsvik, O. O., & Martinussen, M. (2019). Psychometric properties of the Norwegian aberrant behavior checklist and diagnostic relationships in a neuro-pediatric sample. *Journal of Mental Health Research in Intellectual Disabilities*, 12(3–4), 234–255. <https://doi.org/10.1080/19315864.2019.1630872>
- Halvorsen, M., Mathiassen, B., Myrbakk, E., Brøndbo, P. H., Sætrum, A., Steinsvik, O. O., & Martinussen, M. (2019). Neurodevelopmental correlates of behavioural and emotional problems in a neuropaediatric sample. *Research in Developmental Disabilities*, 85, 217–228. <https://doi.org/10.1016/j.ridd.2018.11.005>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Haynes, A., Gilmore, L., Shochet, I., Campbell, M., & Roberts, C. (2013). Factor analysis of the self-report version of the strengths and difficulties questionnaire in a sample of children with intellectual disability. *Research in Developmental Disabilities*, 34(2), 847–854. <https://doi.org/10.1016/j.ridd.2012.11.008>
- Heyerdal, M. (2003). SDQ - Strength and Difficulties Questionnaire: En orientering om et nytt spørreskjema for kartlegging av mental helse hos barn og unge [SDQ - Strength and Difficulties Questionnaire: A briefing on a new questionnaire for mapping mental health in children and adolescents]. *Norsk Epidemiologi*, 13(1), 127–135. <https://doi.org/10.5324/nje.v13i1.320>
- Hooper, D., Coughlan, J., & Mullen, M. (2007). Structural equation modeling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6(1), 53–60.
- Kaptein, S., Jansen, D. E. M. C., Vogels, A. G. C., & Reijneveld, S. A. (2008). Mental health problems in children with intellectual disability: Use of the Strengths and Difficulties Questionnaire. *Journal of Intellectual Disability Research*, 52(2), 125–131. <https://doi.org/10.1111/j.1365-2788.2007.00978.x>
- Kornør, H., & Heyerdahl, S. (2013). Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, selvrappport (SDQ-S) [Psychometric properties of the Norwegian version of the Strengths and Difficulties Questionnaire, self-report (SDQ-S)]. *PsykTestBarn*, 2(6), 1–14. <https://doi.org/10.21337/0029>
- Kornør, H., & Heyerdahl, S. (2014). Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, lærerrappport (SDQ-L) [Psychometric properties of the Norwegian version of the Strengths and Difficulties Questionnaire, teacher version (SDQ-T)]. *PsykTestBarn*, 2(5), 1–15. <https://doi.org/10.21337/0036>
- Kornør, H., & Heyerdahl, S. (2017). Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, foreldrerappport (SDQ-P) [Psychometric properties of the Norwegian version of the Strengths and Difficulties Questionnaire, parent version (SDQ-P)]. *PsykTestBarn*, 1(1), 1–30. <https://doi.org/10.21337/0048>
- McCrary, C., & Layte, R. (2012). Testing competing models of the Strengths and Difficulties Questionnaire's (SDQ's) factor structure for the parent-informant instrument. *Personality and Individual Differences*, 52(8), 882–887. <https://doi.org/10.1016/j.paid.2012.02.011>
- Palmieri, P. A., & Smith, G. C. (2007). Examining the structural validity of the Strengths and Difficulties Questionnaire (SDQ) in a U.S. sample of custodial grandmothers. *Psychological Assessment*, 19(2), 189–198. <https://doi.org/10.1037/1040-3590.19.2.189>
- Raven, J. (2004). *The Raven's - Educational coloured progressive matrices and Crichton vocabulary scale. Norwegian manual supplement*. Pearson Assessment.
- Rønning, J. A., Handegaard, B. H., Sourander, A., & Mørch, W. T. (2004). The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *European Child & Adolescent Psychiatry*, 13(2), 73–82. <https://doi.org/10.1007/s00787-004-0356-4>
- Sparrow, S., Cicchetti, D., & Balla, D. (2011). Vineland adaptive behavior scales. In *Norwegian manual supplement* (2nd ed., Vol. 2). Pearson Assessment.
- Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent-child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. *BMC Psychiatry*, 10. <https://doi.org/10.1186/1471-244x-10-56>
- Wechsler, D. (2007). Wechsler abbreviated scale of intelligence (WASI). *Norwegian manual supplement*. Pearson Assessment.
- Wechsler, D. (2008a). Wechsler adult intelligence scale. (WAIS-IV). *Norwegian manual supplement* (4th ed.). Pearson Assessment.
- Wechsler, D. (2008b). Wechsler preschool and primary scale of intelligence. (WPPSI-III). *Norwegian manual supplement* (3rd ed.). Pearson Assessment.
- Wechsler, D. (2009). Wechsler intelligence scale for children. (WISC-IV). *Norwegian manual supplement* (4th ed.). Pearson Assessment.
- Wechsler, D. (2012). Wechsler preschool and primary scale of intelligence. (WPPSI-IV). *Norwegian manual supplement* (4th ed.). Pearson Assessment.