**MEASURE NAME:** Strengths and Difficulties Questionnaire-Child Report  
**Acronym:** SDQ-Child

<table>
<thead>
<tr>
<th>Basic Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s):</strong></td>
</tr>
</tbody>
</table>
| **Author Contact:** | Robert Goodman, Ph.D  
Department of Child and Adolescent Psychiatry  
Institute of Psychiatry  
DeCrespigny Park  
London SE5 8AF, United Kingdom |
| **Author Email:** | N/A |
| **To Obtain:** | Extensive information regarding the SDQ, as well as copies of the measure in multiple languages, may be obtained from www.sdqinfo.com.  
The website states that the measure is copyrighted, but “paper versions may be downloaded and subsequently photocopied without charge by individuals or non-profit organizations provided they are not making any charge to families.” |
| **E-mail:** | sdq@youthinmind.net |
| **Website:** | **www.sdqinfo.com** |
| **Cost per copy (in US $):** | $0.00 |
| **Copyright:** | Yes |
| **Description:** | The SDQ is a widely and internationally used brief behavioral-screening instrument assessing child positive and negative attributes across 5 scales: 1) Emotional Symptoms, 2) Conduct Problems, 3) Hyperactivity-Inattention, 4) Peer Problems, 5) Prosocial Behavior. The measure also yields a Total Difficulties score. The SDQ was designed to be administered to parents or teachers in parallel versions, a child self-report version is also available (each version is reviewed separately in this database).  
The SDQ has been extensively researched with various populations and has been translated into over 40 languages. An extended version is available and includes an impact supplement that asks if the respondent thinks the young person has a problem, and if so, inquires about Chronicity, Distress, Social Impairment, and Burden for Others. |
| **Theoretical Orientation Summary:** | The measure is based on the concepts that underpin the Diagnostic and Statistical Manual of Mental Disorders (4th Ed.) |
and the ICD-10. For example, items in the SDQ Inattention-Hyperactivity scale were selected because they reflect key symptoms for a DSM-IV diagnosis of ADHD or ICD-10 diagnosis of hyperkinesis (Goodman & Scott, 1999).

### Domains Assessed:

| 1. | General symptomatology (child) |
| 2. | ADHD symptoms/disorders (child) |
| 3. | Externalizing Symptoms (child) |
| 4. | Mood and Anxiety Symptoms (child) |
| 5. | Peer group relationships (child) |
| 6. | Psychosocial functioning (child) |

### Languages Available:

Afrikaans, Amharic, Arabic, Bangla, Basque, Bengali, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Farsi, Finnish, French, Gaelic, Gallego, German, Greek, Gujarati, Hindi, Hungarian, Icelandic, Irish, Italian, Japanese, Khmer, Korean, Lithuanian, Macedonian, Malay, Norwegian, Polish, Portuguese, Punjabi, Romanian, Russian, Serbian, Sinhalese, Slovene, Spanish, Swedish, Tamil, Thai, Turkish, Ukranian, Urdu, Welsh

### Age Range: 11.0 - 16.0

### # of Items: 25

### Measure Type: Screening

### Measure Format: Questionnaire

### Time to Complete (min): 5

### Reporter: Self

### Time to Score (min): 5

### Education Level: 99.00

### Periodicity: Standard SDQ is last 6 months. Follow-up (for intervention) is last month.

### Response Format: Symptom Scales: 3-point rating: 0 = Not True, 1 = Somewhat True, 2 = Certainly True

### Materials Needed:

<table>
<thead>
<tr>
<th>Material</th>
<th>Required</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper and pencil</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Computer</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Video equipment</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

### Material Notes:

While there is no manual in the typical form, all information (including administration procedures and scoring instructions) is on the website (www.sdqinfo.com). All of this information is provided free of charge and is downloadable from the website.

All versions of the SDQ (in all translations) can be downloaded from the website including:

1. Materials for scoring including instructions for hand scoring, black-and-white transparent overlays for hand-scoring, and access to online scoring (www.sdqscore.net).

2. An optional computerized scoring and report-writing program, which runs using Microsoft Access, is available free of charge by contacting sdq@youthinmind.net. Instructions for obtaining the program are on the website.

3. Scoring syntax in different languages (e.g., SPSS, SAS, STATA).
6. Computerized algorithms for predicting psychiatric disorder using information from different versions of the SDQ (Self, Parent, Teacher). The algorithm makes separate predictions for conduct-oppositional disorders, hyperactivity-inattention disorders, and anxiety-depressive disorders as Unlikely, Possible, or Probable.

Sample Items:

<table>
<thead>
<tr>
<th>Domains</th>
<th>Scale</th>
<th>Sample Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Difficulties</td>
<td>Conduct Problems</td>
<td>I am often accused of lying or cheating.</td>
</tr>
<tr>
<td></td>
<td>Inattentio-Hyperactivity</td>
<td>I am restless, I cannot stay still for long.</td>
</tr>
<tr>
<td>Emotional Symptoms</td>
<td></td>
<td>I worry a lot.</td>
</tr>
<tr>
<td>Peer Problems</td>
<td></td>
<td>Other children or young people pick on me or bully me.</td>
</tr>
<tr>
<td>Prosocial</td>
<td></td>
<td>I try to be nice to other people. I care about their feelings.</td>
</tr>
</tbody>
</table>

Notes (additional scales and domains):
If the impact supplement is given, it yields the following scores: 1) Impact score, 2) Chronicity (1 item), 3) Burden (1 item).

Information Provided: (check all that apply)

| No/Diagnostic information DSM-III | Yes/Standard Scores          | No/Diagnostic information DSM-IV | Yes/Percentile |
| Yes/Strengths                   | No/Graph (e.g., of elevated scale) | Yes/Areas of concerns/risks       | No/Dichotomous assessment |
| No/Program evaluation information | Yes/Clinical friendly output | Yes/Continuous assessment       | No/Written feedback |
| Yes/Raw Scores                  | No/Other                      |                                |                  |

Training Notes: The SDQ is designed to be administered by researchers, clinicians, and educators. Specific data on training needed to administer and interpret is not provided.
### Parallel or Alternate Forms

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel Forms?</td>
<td>No</td>
</tr>
<tr>
<td>Alternate Forms:</td>
<td>No</td>
</tr>
<tr>
<td>Forms for Different Ages:</td>
<td>Yes</td>
</tr>
<tr>
<td>If so, are forms comparable:</td>
<td>Yes</td>
</tr>
<tr>
<td>Any Altered Versions of Measure:</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Describe:

There are multiple versions of the SDQ to meet the needs of researchers and clinicians. All versions have the 25 items that comprise the scale.

Details are from the website (unless otherwise cited):

1. The teacher version is the same as the parent version but is reviewed separately in this database in order to present data specifically by reporter.

2. There is a parent/teacher version for 3-4 year olds. It contains 22 identical items. One item on reflectiveness is slightly reworded (original=“thinks things out before acting”; 3-4 yr olds: “can stop and think things out before acting.”)

2 items on antisocial behavior are replaced by items on oppositionality (original: “often lies or cheats”); 3-4 year olds: “often argumentative with adults” and original: steals from home, school or elsewhere”; 3-4 year olds: “can be spiteful to others.”

3. Multiple versions exist for different language groups.

4. An impact supplement is available, which first asks whether the respondent thinks the youth has a problem or not, and, if so, gathers data regarding chronicity, distress, social impairment, and burden to others.

5. There are follow-up questionnaires for use at post-test, following an intervention. This version has the 25 basic items, the impact question, and 2 follow-up questions regarding change due to intervention. The timeframe for this measure is also changed from “last six months or this school year” to “last month.”

6. There is an Adolescent Self-Report version, which is also reviewed in this database. The wording on this version is slightly different.

7. There is a computerized version developed for the Child Self-Report version. This version was examined with a group of children aged 8-15. No differences were found between means when the measure was completed on the computer versus on paper in a clinic sample.
The computerized version was more highly correlated with parent report and had a better test-retest reliability ($r=.83$, 40 children tested after 6 weeks) than the paper report did, although the difference between the computerized report and paper report was not statistically significant. Children who used the computerized form were more likely to report that the questionnaire was easy to complete. The computerized version appeared to discriminate between clinic and community samples (Truman et al., 2003).

### Population Used to Develop Measure

Psychometrics for the SDQ were originally examined in a sample of 346 parent respondents and 185 teacher respondents. Children, aged 4-16, were recruited from two London child psychiatric clinics or a children’s dental hospital in London.

Psychiatric Sample: $M=9.8$ years; 63% male, 37% female
Dental Sample: $M=10.8$ years; 53% male, 47% female.
No other demographic information was available (Goodman, 1997).

### Psychometrics

**Global Rating (scale based on Hudall Stamm, 1996):**

- Psychometrically matured, used in multiple peer reviewed articles by different people
- Norms:
  - For separate age groups: Yes
  - For clinical populations: Yes
  - Separate for men and women: Yes
  - For other demographic groups: Yes

**Notes:**

- Normative data has been obtained in several countries, using several translations of the SDQ (see website). Two of the largest scale normative studies have been conducted in the United Kingdom and in the United States.

1. **UK**
   - Normative data was obtained on a total of 10,438 children aged 5 to 15.
   - Information was obtained from:
     - 10,298 parents (99% of sample)
     - 8,208 teachers (79% of sample)
     - 4,228 children, aged 11-15 (93% of this age band)
   - Samples of children aged 5-10 and 11-15: 50% male and 50% female; from urban, semi-rural, and rural areas.
   - Note: For children aged 5-10, there are parent and teacher norms, by child gender, but there are no norms for self-report. For children aged 11-15, there are norms for Parent, Teacher, and Self-Report by gender.

2. **UNITED STATES**
   - The SDQ was included in the 2001 National Health Interview Survey (NHIS)
Information on the sample child was obtained from a knowledgeable adult residing in the household.

Of the 10,367 children in the survey who were aged 4-17, there was complete data for 9,878 children on all sections of the SDQ, and normative data is available for this sample.

The sample included children aged 4-7, 8-10, and 11-14; and had equal representation from both genders. Respondents included parents (biological, adoptive, or step: 92%) and grandparents (4.4%).

Norms are available on the website and in Bourdon, Goodman, Rae, Simpson, & Koretz (2005).

Note: Normative data are available only for the Parent Report, but not for Child or Teacher report. They are available by gender and age (4-7, 8-10, 11-14, 15-17).

<table>
<thead>
<tr>
<th>Clinical Cutoffs:</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specify Cutoffs:</td>
<td>While there are no cutoffs, scores at or above the 90th percentile are used to predict psychiatric disorder.</td>
</tr>
<tr>
<td>Used in Major Studies:</td>
<td>Yes</td>
</tr>
<tr>
<td>Specify Studies:</td>
<td>Multiple across different countries e.g., Bhui, et al. (2005); Calam, Gregg, &amp; Goodman (2005); Goodman, Gledhill, &amp; Ford (2003); Lynch, Mills, Daly, &amp; Fitzpatrick (2004); Ronning et al. (2004).</td>
</tr>
</tbody>
</table>
Reliability:

<table>
<thead>
<tr>
<th>Type:</th>
<th>Rating</th>
<th>Statistics</th>
<th>Min</th>
<th>Max</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test-Retest-# days:</td>
<td>24</td>
<td>Acceptable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal Consistency:</td>
<td></td>
<td>Pearson correlation</td>
<td>0.66</td>
<td>0.82</td>
<td>0.71</td>
</tr>
<tr>
<td>Inter-Rater:</td>
<td>Acceptable</td>
<td>Cronbach’s alpha</td>
<td>0.41</td>
<td>0.81</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Acceptable</td>
<td>Pearson correlation (parentxchild)</td>
<td>0.3</td>
<td>0.48</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Parallel/Alternate Forms:

Notes:

RELIABILITY
The internal consistency and inter-rater reliability (Parent x Child) data presented in the table above are from Goodman (2001) because in this database we typically report reliabilities presented by the measure’s author. Test-retest data are from Mellor (2004) because they are reported separately by scale, and the timeframe is more appropriate and comparable to other studies. Additional psychometric data from studies conducted in other countries are presented when they were available in the Notes in the “Content Validity” section (under “USE IN OTHER COUNTRIES”).


TEST-RETEST RELIABILITY
Stability of scores over a 4-6 month interval: mean test-retest stability was .62.

INTERNAL CONSISTENCY (alpha)
Goodman (2001) reported the following data:
Total (.80), Emotional Symptoms (.66), Conduct Problems (.60), Hyperactivity/Inattention (.67), Peer Problems (.41), Prosocial Behavior (.66), Impact (.81).

INTER-RATER RELIABILITY
Correlations between raters using Pearson product moment correlation: as reported by Goodman, nearly all correlations were greater than those reported in a meta-analysis of cross-informant correlations (Achenbach et al., 1987).

Parent x Child:
Total Difficulties (.48), Emotional Symptoms (.37), Conduct Problems (.44), Hyperactivity/Inattention (.41), Peer Problems (.40), Prosocial Behavior (.30), Impact (.30)

Teacher x Child:
Total Difficulties (.33), Emotional Symptoms (.21), Conduct Problems (.30), Hyperactivity/Inattention (.32), Peer Problems (.29), Prosocial Behavior (.23), Impact (.23)


TEST-RETEST RELIABILITY: subset of 120 families over a 2-week period.
Good test-retest reliability was reported for children aged 11-17 and children aged 7-11. There were no differences between older and younger children except for the peer problems scale, with younger children being less consistent in their responses.
Total Difficulties (.72), Emotional Symptoms (.68), Conduct Problems (.66), Hyperactivity/Inattention (.82), Peer Problems (.67), Prosocial (.69)

INTERNAL CONSISTENCY (alpha)
Similar internal consistencies were reported for children aged 11-17 and children aged 7-
### CHILDREN AGED 11-17
Total Difficulties (.70), Emotional Symptoms (.65), Conduct Problems (.64), Hyperactivity/Inattention (.75), Peer Problems (.59), Prosocial (.66)

### CHILDREN AGED 7-11
Total Difficulties (.72), Emotional Symptoms (.68), Conduct Problems (.66), Hyperactivity/Inattention (.73), Peer Problems (.55), Prosocial (.62)

### ALL CHILDREN
Total Difficulties (.71), Emotional Symptoms (.67), Conduct Problems (.65), Hyperactivity/Inattention (.74), Peer Problems (.58), Prosocial (.64)

### Interrater Reliability
Correlations among reporters: All correlations (e.g., Parent and Teacher, Parent and Child, Teacher and Child) were significant at p<.01 and ranged from .18-.50 (average correlation=.37). Analyses examining differences between older (aged 11-17) and younger children (aged 7-11) showed that older children’s reports are more consistent with parents’ reports, but there are no differences between older and younger children’s reports with teachers.

### Parent-Child
Total Difficulties (.38), Emotional Symptoms (.32), Conduct Problems (.37), Hyperactivity/Inattention (.46), Peer Problems (.34), Prosocial (.34)

### Teacher-Child
Total Difficulties (.36), Emotional Symptoms (.24), Conduct Problems (.39), Hyperactivity/Inattention (.44), Peer Problems (.35), Prosocial (.29)

### Content Validity:
As described by Goodman (1997) the SDQ was designed with the following specifications:

1. Applicable to children aged 4-16.
2. One version for parents and teachers and a similar version for child self-report.
3. Both strengths and difficulties well represented.
4. 5 items each on five relevant dimensions (Conduct Problems, Emotional Symptoms, Hyperactivity/Inattention, Peer Relationships, and Prosocial Behavior).

The dimensions selected were based on factors identified in an analysis of an expanded version of the Rutter parent questionnaire (Goodman, 1994).

Items were also based on nosological concepts and on concepts that underpin the Diagnostic and Statistical Manual of Mental Disorders (4th Ed.) and the ICD-10 (Goodman & Scott, 1999). For example, items in the SDQ Hyperactivity/Inattention scale were selected because they reflect key symptoms for a DSM-IV diagnosis of ADHD or ICD-10 diagnosis of hyperkinesis.
Validity Type | Not known | Not found | Nonclinical Samples | Clinical Samples | Diverse Samples
---|---|---|---|---|---
Convergent/Concurrent | Yes | Yes | Yes
Discriminant | Yes | Yes | Yes
Sensitive to Change | Yes | Yes
Longitudinal/Maturation Effects | Yes | Yes
Sensitive to Theoretically Distinct Groups | Yes | Yes | Yes

Notes: Only studies that administered the SDQ Child Self-Report version were included in the summary below. Given the large number of studies, not all were reviewed. We focused the review on the use of the measure with trauma exposed and diverse populations. Although studies are grouped by headings (e.g., “USE WITH TRAUMA EXPOSED POPULATIONS” and “USE IN OTHER COUNTRIES”), there is a lot of overlap among categories.

The SDQ Child version, along with other SDQ versions have been used in many studies of conduct disorder and behavior problems (e.g., Taha et al., 2005) including twin studies examining genetic and environmental influences (e.g., Scourfield, Van den Bree, Martin, & McGuffin, 2004).

CORRELATIONS WITH OTHER MEASURES
1. The SDQ correlates with the Rutter (Goodman, 1997) and relates well with Teacher and Parent SDQ scores (Goodman, Meltzer, & Bailey, 2003).

2. In a study of bullying behavior (Woods & White, 2005), behavior problems in the clinical range on all SDQ scores were associated with clinical levels of over arousal, as assessed using the Arousal Predisposition Scale. Additional data are also summarized below under "USE IN OTHER COUNTRIES."

DISCRIMINANT VALIDITY
1. The measure differentiates between different groups of children. A number of studies show that the measure discriminated between clinic and community samples (e.g., Goodman et al., 2003).

2. Anxious children have been found to score higher than do nonclinical and externalizing groups on the Total and Emotional Symptoms Subscale, and externalizing children score higher than do anxious and nonclinical groups on Hyperactivity/Inattention. (Lyneham & Rapee, 2005).

FACTOR ANALYSIS
As described below, under “USE IN OTHER COUNTRIES” a number of factor analytic studies have been conducted with somewhat divergent results.

1. Goodman (2001) reported that, as hypothesized, a 5-factor solution was found for Parent, Teacher, and Child Self-Report using eigen values>1 to determine the number of factors.
2. Studies in Germany (Becker, Hagenberg, Roessner, Woerner, & Rothenberger, 2004) and the Netherlands (Muris, Meesters, & van den Berg, 2003) found a similar factor structure as that reported by Goodman using British samples.

3. However, factor analysis with Norwegian (Ronning, Handegaard, Sourander, & Morch, 2004), Finnish (Koskelainen, Sourander, & Vauras, 2001), and Arab (Gaza) children (Thabet, Stretch & Vostanis, 2004) suggest that the factor structure proposed by Goodman has a somewhat variable and questionable fit.

4. Koskelainen et al. (2001) proposed an alternate 3-factor structure (externalizing, internalizing, and prosocial).

TREATMENT OUTCOME
A number of studies have shown that at least some SDQ scales are sensitive to treatment effects using pre- to post-test comparison designs (e.g., inpatient: Gavida-Payne, Littlefield, Halgren, Jenkins, & Coventry, 2003; outpatient: Callaghan, Young, Pace, & Vostanis, 2004).

1. For example, Callaghan et al. (2004) found that following 5 months of treatment the peer relationship problems score showed significant improvement.

2. In a small sample of child reports (n=8) Mathai, Anderson, & Bourne (2003) reported significant declines in total SDQ scores following treatment.

USE WITH YOUNGER CHILDREN
1. Muris, Meesters, Vincken, & Eijkelenboom (2004) examined the psychometrics of the SDQ self-report in a non-clinical sample of children aged 8-13 years old. These data are reviewed separately because the measure was not developed for children under age 11. Overall, the results provide some support for the use of the Self-Report with younger children.

The authors reported questionable internal reliability for younger children aged 8-10: Conduct (.45); Peer Problems (.36); Emotional Symptoms (.56), and Prosocial behavior (.57). Children's SDQ scores were, however, correlated with teacher SDQ scores, mean correlation=.28, and teacher scores on the externalizing scales of the Teacher Report Form.

Factor analysis suggested the presence of four factors: Emotional Symptoms, Prosocial Behavior, Hyperactivity/Inattention, and a mixed factor of Peer and Conduct Problems. Children identified by teachers as displaying behavior problems at school had higher Difficulty and lower Prosocial scores. SDQ scores appeared to be better at predicting behavior problems (teacher rating) than Youth Self-Report Scores (YSR). SDQ scores were found to correlate with YSR scores in the expected direction.

USE WITH DIVERSE SAMPLES
1. All versions of the SDQ (Child, Parent, and Teacher) were found to have acceptable internal consistency and validity and to be considered a robust measure for children and adolescents with intellectual disabilities (Emerson, 2005).

2. The measure has been used with many different ethnic groups within England, including immigrant and refugee children (e.g., Leavey, 2004). For
example, in a multiethnic group of 26,23 adolescents in East London (included 690 Bangladeshi, 250 Indian, 184 Pakistani, 166 black Carribean, 279 Black African, 121 Black British 191 mixed ethnicity, 581 White UK, and 161 White other youth), Bangladeshi youth who were more culturally integrated were found to have fewer mental health problems as assessed by the SDQ (Bhui, Stansfeld, Head, Haines, Hillier, Taylor, et al., 2005).

3. The measure has been used in samples of Indian and White British children, with results showing different connections between symptoms and father’s involvement based on ethnicity (Flouri, 2005).

4. The SDQ was used in a large (n=2,790) study of ethnically diverse children aged 11-14 to examine whether ethnic differences in the prevalence of psychological distress are associated with deprivation (Stansfeld, Haines, Head, Bhui, Viner, Taylor, et al., 2004).

USE IN OTHER COUNTRIES (EXCEPT BRITAIN)
The SDQ has been used in many countries to examine rates of psychopathology (e.g., Ireland: Lynch, Mills, Daly, & Fitzpatrick, 2004). Studies have also examined the psychometrics of the measure in different countries. Many of these studies are detailed below.

1. Woerner et al. (2004) report on the use of the SDQ overseas (beyond Europe) in Brazil, Canada, the Middle East, Asia, and Australia. They suggest that the data provides support for the psychometric properties of the measure.

BRAZIL
The SDQ (Parent, Teacher, and Child versions) were used in a study that examined child mental health problems in a rural African-Brazilian community (Goodman, dos Santos, Nunes, de Miranda, Fleitlich-Bilyk, & Filho, 2005). The authors report significant agreement between the SDQ and the Development and Well-Being Assessment (DAWBA).

AUSTRALIA
The SDQ is part of the health department in Queensland Australia’s standardized assessment procedure, which is routinely administered at intake and discharge at child and youth mental health facilities throughout the state (Harnett, Loxton, Sadler, Hides, & Baldwin, 2005).

BANGLADESH
Mullick & Goodman (2001) examined the psychometrics of a Bangla version (translated and backtranslated) with a sample of 99 clinic and 162 community Bangladeshi children aged 4-16. They found that SDQ scores distinguish between community and clinic samples and between children with different psychiatric diagnoses. Using ROC curves for each SDQ scale, AUC (Area Under Curve) = >.80 were found for Total Impact, Emotional Problems, and Hyperactivity/Inattention. For Parent, Teacher, and Child reports, Emotional Symptoms were able to distinguish between clinic cases with and without an emotional disorder, Conduct Problems were able to distinguish between clinic cases with and without conduct disorder, and Hyperactivity was able to distinguish between those with and without a hyperactivity disorder.

GAZA
A study of Arab children living in the Gaza Strip suggests that the standard
factor structure may not be appropriate for these children and that certain items appeared to have different meaning for these participants compared to Western participants (Thabet, Stretch, & Vostanis, 2000).

SOUTHERN EUROPE
Marzocchi et al. (2004) described the use of the SDQ in southern European countries (Italy, Spain, Portugal, Croatia, France).

Spanish
The Spanish version of the SDQ has been used in a number of studies. García, Goodman, Mazaira, Torres, Rodríguez-Sacristán, Hervas & Fuentes (2000) reported on the initial psychometrics comparing the SDQ with the CBCL and Child Behavior Questionnaire. The original article cannot be obtained, so it is unclear as to which version of the SDQ they used.

NORDIC COUNTRIES
A review article on the use of the SDQ in Nordic countries (Obel, Heiervang, Odriguez, Heyerdahl, Smedje, Sourander, et al., 2004) suggested that the distributions of the SDQ are similar across countries and suggested collaboration in developing norms for Nordic countries. The authors described the use of the SDQ in Sweden, Finland, Norway, Denmark, and Iceland, detailing studies in each of these countries that had used the SDQ.

Norway
1. Oppedal, Roysamb, & Heyerdahl (2005) analyzed data from 1,295 10th-grade immigrants in Norway. The sample included children who had immigrated from recent conflict zones. In this sample, they reported the following internal consistencies for the SDQ-Self-Report version: Total Score (.72), Emotional Problems (.69), Conduct Problems (.46), Hyperactivity Problems (.53), Peer problems (.44). They reported differences on SDQ scores by generation, ethnicity, and gender.

2. Ronning, Handegaard, Sourander, & Morch (2004) examined the psychometrics of the SDQ self-report version in a sample of 4,167 Norwegian children aged 11-16. They reported internal consistencies (58-.67). Confirmatory factor analysis suggested that the factor structure proposed by Goodman had a somewhat variable and questionable fit. The authors suggested that it “might be worth reformulating some items” at least in the Norwegian version and making modifications to improve the measure's psychometric strength.

3. In a sample of 4,130 Norwegian 6th-10th graders self-perceived harassment was found to be related to SDQ symptoms (Ronning et al. 2004).

Finland
1. Koskelainen, Sourander, & Kaljonen (2000) reported on the psychometrics of the Parent, Teacher, and Child SDQ in a sample of Finnish children aged 7-15 (n=735). They reported on the internal consistency for all three reporters as alpha=.63-.86. Teachers had the best internal consistency (M=.79) compared to parents and children (.67 and .65, respectively).

Inter-rater reliability (correlations): .28-.40 for children and parents, .28-.38 for
children and teachers, and .29-.45 for parents and teachers. The validity was supported through correlations with strong correlations with the CBCL and Youth Self Report.

For example the total CBCL and Parent SDQ were correlated at r=.75, and the total child self-report SDQ and YSR Total were correlated at .71.

2. Koskelainen, Sourander, & Vauras (2001) conducted a factor analysis using data from 1,458 Finnish youth. They forced a 5-factor solution and reported that the first 3 factors were structured in accordance with the original SDQ. The remaining factors were somewhat problematic. A second factor analysis with the number of factors unspecified, suggested the presence of 3 factors. They reported internal consistency for the 5 scales as .53-.71. Mean scores and cutoffs are provided.

GERMANY
The psychometrics of the German version of the SDQ have been detailed in a number of published studies.

1. Klasen, Woerner, Rothenberger, & Goodman (2003) described the psychometric properties of the German SDQ Parent, Teacher, and Self-Report. These data are summarized from the abstract as the article is in German. They reported that factor analysis replicated the original scale structure. The SDQ was correlated with the German version of the CBCL as expected. They suggested that the German version is as useful and valid as the English version.

2. Becker, Hagenberg, Roessner, Woerner, & Rothenberger (2004) reported the following internal consistencies for the SDQ German version: Total Difficulties (.78), Emotional (.77), Conduct Problems (.58), Hyperactivity/Inattentiveness (.65), Peer Problems (.65), Prosocial Behavior (.78). Principal components analysis with a German sample suggested a 5-factor solution explaining 51.4% of the variance, with high concordance with the original SDQ scales (Becker et al., 2004).

The authors also reported correlations between Child and Parent SDQ scores for each scale and for Child and Teacher SDQ scores. Correlations were all significant and moderate for parents and children (.30-.57), and teachers and children (.27-.50). The authors compared the predictive power of the SDQ with the CBCL in terms of detecting psychiatric diagnosis. The child self report was reportedly as effective as the Youth Self-Report or CBCL.

Netherlands
The psychometrics of the Dutch version of the SDQ have been reported on in at least two studies (Muris, Meesters, & van den Berg, 2003; van Widenfelt, Goedhart, Treffers, & Goodman, 2003).

1. Muris, et al. (2003) examined the psychometrics of the SDQ (Parent and Child versions) in a sample of Dutch 562 children aged 9 to 15 (M=12.3). Factor analysis of the Child Self-Report data suggested a 5-factor solution accounting for 43.9% of the variance, with the majority of items (all but 4) loading on their hypothesized factors. They also reported the following data:

**TEST-RETEST RELIABILITY (average 2-month intraclass correlation):**
Total Difficulties (.87), Emotional Symptoms (.76), Conduct Problems (.66), Hyperactivity/Inattentiveness (.88), Peer Problems (.83), Prosocial Behavior (.59)
INTERNAL CONSISTENCY (alpha):
- Total Difficulties (.78)
- Emotional Symptoms (.71)
- Conduct Problems (.45)
- Hyperactivity/Inattention (.72)
- Peer Problems (.54)
- Prosocial Behavior (.62)

INTERRATER RELIABILITY (correlations between Parent and Child reports):
- Total Difficulties (.46)
- Emotional Symptoms (.43)
- Conduct Problems (.31)
- Hyperactivity/Inattention (.42)
- Peer Problems (.43)
- Prosocial Behavior (.21)

VALIDITY
The Child Self-Report SDQ correlated significantly with the Youth Self-Report (YSR), Children’s Depression Inventory (CDI), Revised Children’s Manifest Anxiety Scale (RCMAS), and ADHD Questionnaire (ADHDQ), as expected. All correlations are reported in the article.

USE WITH TRAUMA-EXPOSED POPULATIONS

General Trauma
Correlations between Parent and Child reports on the SDQ were examined in both an At-Risk Norwegian sample and a community sample (Waaktaar, Borge, Christie, & Torgersen, 2005). The At-Risk sample included children, many of whom had experienced traumas including death of a parent or sibling, war exposure, refugee experiences.

Correlations for the At-Risk sample:
- Total Problems (.43)
- Disruptive Difficulties (.52)
- Emotional Difficulties (.22)
- Prosocial Behavior (.19)

Correlations for the community sample:
- Youth-Mother: Total Problems (.30)
- Disruptive Difficulties (.34)
- Emotional Difficulties (.30)
- Prosocial Behavior (.16)

- Youth-Father: Total Problems (.33)
- Disruptive Difficulties (.37)
- Emotional Difficulties (.21)
- Prosocial Behavior (.16)

MALTREATED CHILDREN
1. In a study of maltreatment among Palestinian youth in the Gaza Strip, youth who were maltreated scored higher on many SDQ scales (Self and Teacher reports) than did non-maltreated youth.

2. Emotional Problems scores were predicted by a coping strategy of “trying to feel better by eating, drinking, smoking, using drugs, or medication (Thabet, Tischler, & Vostanis, 2004).

CHRONIC ILLNESS
1. Parent and Child SDQs were collected on a sample of 43 children with end stage renal disease. Means and SD are presented and compared to a normative sample (Madden, Hastings, & Hoff, 2002).
Predictive Validity:

Postdictive Validity:

Sensitivity Rate(s):

Specificity Rate(s):

Positive Predictive Power:

Negative Predictive Power:

Goodman (2001) reported on the predictive validity of the SDQ in predicting independently diagnosed DSM-IV diagnoses. Statistics are reported separately by scale and by diagnosis. For total SDQ scales (Youth Report only) and any DSM-IV diagnosis: Specificity (95%), Sensitivity (43%), Negative Predictive Value (94%), Positive Predictive Value 44%. These data are reported in the above table.

A computerized algorithm has also been developed to predict child psychiatric diagnosis using SDQ Symptom and Impact scores from multiple informants (Parents, Teachers, and Children). The algorithm yields scores of unlikely, possible, or probable for 4 categories of disorder: 1) Conduct Disorder, 2) Emotional Disorders, 3) Hyperactivity Disorders, and 4) Any Psychiatric Disorders.

A number of studies have examined the predictive validity of this algorithm in terms of its ability to screen for children with psychiatric disorders (e.g., Goodman, Ford, Simmons, Gatward, & Meltzer, 2000; Goodman, Ford, Simmons, Gatward, & Meltzer, 2003).

1. Using this algorithm, Goodman, Renfrew, & Mullick (2000) found that agreement between SDQ prediction and independent clinical diagnosis was highly significant (Kendall’s tau-b ranging from .49-.73).

When the scores were dichotomized (only “probably” counted as positive), across disorders (Conduct, Emotional, and Hyperactivity) and samples (London & Dhaka), they reported Sensitivity (.81%-90%), Specificity (47%-84%), Positive Predictive Power (35%-86%), and Negative Predictive Power (.83%-98%). They reported that “the algorithm is good at detecting disorder . . . but at the expense of being over-inclusive.”

2. Goodman, Ford, Corbin, & Meltzer (2004) present Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value using the algorithm to predict psychiatric status in foster children. Using multiple informants they reported the following data: Sensitivity=84.8%, Specificity=80.1%, Positive Predictive Value=74.2%, Negative Predictive Value=88.7%.

For a private household sample: Sensitivity=63.3%, Specificity=94.6%, Positive Predictive Value=52.7%, Negative Predictive Value=96.4%.

The authors suggested that the SDQ predictive algorithm works best when
Data are completed by caregivers and teachers. Caregivers and teachers provide data of similar predictive value. When data from an adult informant are already being used, self-report data appears to contribute little additional information.

Limitations of Psychometrics and Other Comments Regarding Psychometrics:
The SDQ has been extensively researched with different age groups, different informants, diverse cultural groups, and with various translations. Research indicates strong psychometric properties as well as research and clinical utility.

As noted by Goodman, Renfrew, & Mullick (2000) the SDQ algorithm for predicting child psychiatric diagnosis is good at detecting disorder but is overly inclusive.

Consumer Satisfaction
No formal data are available; however, given that the SDQ has been translated into over 40 languages and is widely used in research, clinical, and educational practice, it appears that consumer satisfaction is high.

Languages Other than English

<table>
<thead>
<tr>
<th>Language</th>
<th>Translation Quality (check all that apply)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 = Has been translated</td>
</tr>
<tr>
<td></td>
<td>2 = Has been translated and back translated - translation appears good and valid.</td>
</tr>
<tr>
<td></td>
<td>3 = Measure has been found to be reliable with this language group.</td>
</tr>
<tr>
<td></td>
<td>4 = Psychometric properties overall appear to be good for this language group.</td>
</tr>
<tr>
<td></td>
<td>5 = Factor structure is similar for this language group as it is for the development group.</td>
</tr>
<tr>
<td></td>
<td>6 = Norms are available for this language group.</td>
</tr>
<tr>
<td></td>
<td>7 = Measure was developed for this language group.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finnish</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portuguese (Brazil)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bangla/Bangladesh</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finnish</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabic (Gaza)</td>
<td>Yes</td>
<td>Yes</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>Unk</td>
<td>No</td>
</tr>
</tbody>
</table>

Use with Trauma Populations

Populations for which measure has demonstrated evidence of reliability and validity:

| Yes | 1. Physical abuse | Unk | 2. Natural disaster |
| Unk | 5. Neglect | Unk | 6. Imprisonment |
| Unk | 7. Domestic Violence | Unk | 8. Witness death |
| Unk | 9. Community violence | Unk | 10. Assault |
| Yes | 11. Medical trauma | Yes | 12. War/combat |
|       |             |   |                  |

Strengths and Difficulties Questionnaire-Child Report
NCTSN Measure Review Database
www.NCTSN.org
### Use with Diverse Populations

#### USE WITH DIVERSE POPULATIONS RATING SCALE
1. Measure is known (personal communication, conference presentation) to have been used with members of this group.
2. Studies in peer-reviewed journals have included members of this group who have completed the measure.
3. Measures have been found to be reliable with this group.
4. Psychometric properties well established with this group.
5. Norms are available for this group (or norms include a significant proportion of individuals from this group)
6. Measure was developed specifically for this group.

<table>
<thead>
<tr>
<th>Population Type:</th>
<th>Degree of Usage: (check all that apply)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1. Developmental disability</td>
<td>Yes</td>
</tr>
<tr>
<td>2. Disabilities</td>
<td>Unk</td>
</tr>
<tr>
<td>3. Lower socio-economic status</td>
<td>Yes</td>
</tr>
<tr>
<td>4. Rural populations</td>
<td>Yes</td>
</tr>
<tr>
<td>5. Below average IQ</td>
<td>Yes</td>
</tr>
</tbody>
</table>

#### Notes (including other diverse populations):

The measure has been used with a variety of different ethnic groups (e.g., African Americans, Black British, British Indians, British Pakisatanis, Black Carribeans, Black Africans, British Bangladeshis).

Details regarding other translations are available on the website (www.sdqinfo.com).

**Translations/Countries**
- English (Australia, UK, USA)
- Portuguese (Brazil, Portugal)
- Spanish (inc. Río de la Plata)

#### Pros and Cons/Qualitative Impression

**Pros:**

1. The SDQ has been extensively researched in a wide variety of settings.
2. The SDQ appears to be a very useful tool for screening of mental health problems.
3. Multiple comparable informant versions are available (Parent, Child Self-Report, Teacher).
4. It is brief (much shorter than comparable measures).
5. It is easy to administer and score.
6. The SDQ Parent Report has been associated with service utilization outcomes.
7. The subscales and items correspond to major categories and criteria of current classification systems (Rothenberger & Woerner, 2004).
8. The measure is easily available in more than 40 languages at www.sdqinfo.com.
9. May be good for cross-cultural studies because it is short and available in multiple languages (Rothenberger & Woerner, 2004).
Cons:

1. While no cons are indicated for use of the SDQ as a screening tool, further research is necessary to examine it as a tool for guiding treatment and for examining outcome resulting from treatment.

2. The SDQ emotional subscale may have some weaknesses in terms of its ability to detect specific disorders that are not the focus of the measure’s attention such as specific phobias, panic disorders, separation anxiety, and eating disorders (Goodman et al., 2000; Quinton & Murray, 2002).

3. Neither naturalistic nor interventional longitudinal studies have repeatedly administered the SDQ (Rothenberger & Woerner, 2004).

4. For trauma-exposed children, it should be noted that there are no specific scales focusing on trauma symptomatology. The correlation found between the SDQ and the PTSD Reaction Index is of small magnitude (r=.22), and the authors (McDermott et al., 2005) suggested that the SDQ should not be used on its own to screen for trauma-related symptoms.

5. Studies on the factor structure of the SDQ that have used Child Self-Reports have conflicting findings, with some studies replicating the original 5-factor structure and others suggesting other structures (e.g., 3 factors). However, it should be noted that few measures have undergone such rigorous testing of their factors’ structure, with numerous factor analyses conducted in different countries and cultural groups.
A PsychInfo search (8/05) of “Strengths and Difficulties Questionnaire" or SDQ" anywhere revealed that the measure has been referenced in 329 peer-reviewed journal articles.

Note: Because it was not possible to conduct a search that identified which specific version of the SDQ (Parent, Teacher, Child Self-Report) was used, this number represents the total for all SDQ versions.

However, the articles cited below (for the most part) included the Child version. The number is most likely an underestimate, given that the SDQ is internationally used and citations in foreign journals may not all be included in PsychInfo.


A PsychInfo search 8/05 of “Strengths and Difficulties Questionnaire” or "SDQ" anywhere revealed the measure has been referenced in 6 conferences and 0 dissertations.

Note: Because it was not possible to conduct a search that identified which specific version of the SDQ (Parent, Teacher, Child Self-Report) was used, this number represents the total for all SDQ versions.

The number is most likely an underestimate, given that the SDQ is internationally used and citations in foreign journals may not all be included in PsychInfo.

<table>
<thead>
<tr>
<th>Number of Published References:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(based on author provided information and a PsychInfo search, not including dissertations)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Unpublished References:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(based on a PsychInfo search of unpublished doctoral dissertations)</td>
</tr>
</tbody>
</table>

The author read the review and indicated he was pleased with it. His feedback was integrated into the review.

Citation for Review: Chandra Ghosh Ippen, Ph.D., Amie Alley, Ph.D.

Editor of Review: Chandra Ghosh Ippen, Ph.D., Madhur Kulkarni, M.S.

Last Updated: 8/20/2005

PDF Available: yes

This project was funded by the Substance Abuse and Mental Health Services Administration (SAMSHA), U.S. Department of Health and Human Services (HHS). The views, policies and opinions expressed are those of the authors and do not necessarily reflect those of SAMSHA or HHS.