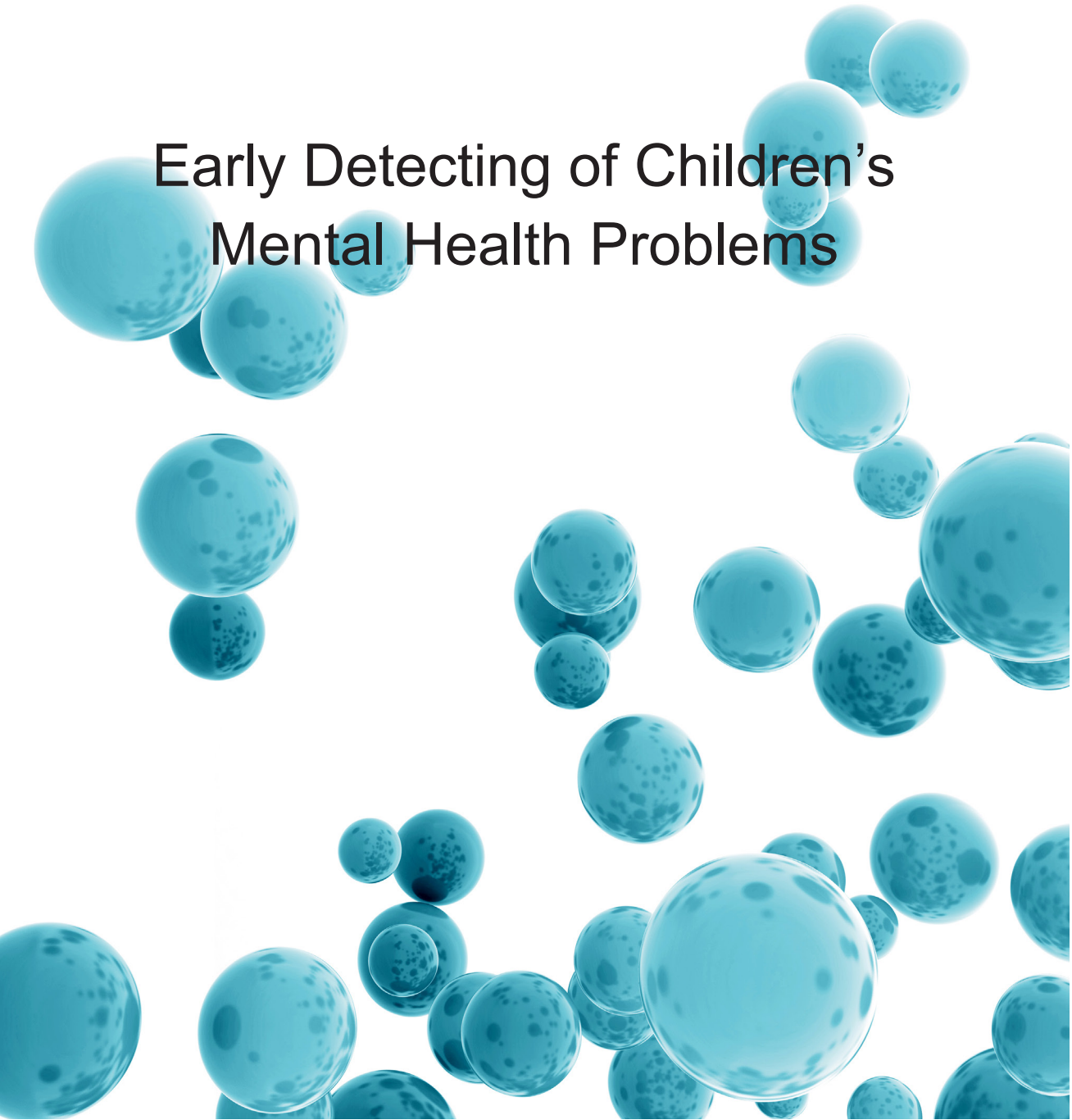
Acta Universitatis Tamperensis 2076

ANNE-MARI BORG

Early Detecting of Children's Mental Health Problems



ANNE-MARI BORG

Early Detecting of Children's Mental Health Problems

Acta Universitatis Tamperensis 2076
Tampere University Press
Tampere 2015

UNIVERSITY OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere, School of Medicine
Finland

Supervised by
Professor Tuula Tamminen
University of Tampere
Finland
Emeritusprofessor Matti Joukamaa
University of Tampere
Finland

Reviewed by
Docent Linnea Karlsson
University of Turku
Finland
Docent Päivi Santalahti
University of Turku
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Cover design by
Mikko Reinikka

Distributor:
verkkokauppa@juvenesprint.fi
https://verkkokauppa.juvenes.fi

Acta Universitatis Tamperensis 2076
ISBN 978-951-44-9863-3 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1570
ISBN 978-951-44-9864-0 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2015

To my dear family

Contents

List of original communications

This review is based on original communications referred to in the text by their Roman numerals I–IV.

I Borg, A-M., Kaukonen, P., Salmelin, R., Joukamaa, M. and Tamminen, T. (2012). Reliability of the Strengths and Difficulties Questionnaire among Finnish 4-9-year-old children. Nordic Journal of Psychiatry, 66(6), 403-413.

II Borg, A-M., Kaukonen, P., Joukamaa, M. and Tamminen, T. (2014). Finnish norms for young children on the Strengths and Difficulties Questionnaire. Nordic Journal of Psychiatry, 68(7), 433-442.

III Borg, A-M., Salmelin, R., Kaukonen, P., Joukamaa, M. and Tamminen, T. (2014). Feasibility of the Strengths and Difficulties Questionnaire in assessing children's mental health in primary care: Finnish parents', teachers' and public health nurses' experiences with the SDQ. Journal of Child and Adolescent Mental Health, 26(3), 229-238.

IV Borg, A-M., Salmelin, R., Joukamaa, M. and Tamminen, T. (2014). Cutting a long story short? The clinical relevance of asking parents, nurses, and young children themselves to identify children's mental health problems by one or two questions. The Scientific World Journal, vol. 2014, Article ID 286939. doi: http://dx.doi.org/10.1155/2014/286939

Abbreviations

ASEBA	Achenbach System of Empirically Based Assessment
AUC	Area under curve
BITSEA	Brief Infant-Toddler Social-Emotional Assessment
DAWBA	Development and Well-Being Assessment
NPV	Negative predictive value
OR	Odds ratio
PPV	Positive predictive value
PSYBOBA	Psychosocial Problems in Primary Education
ROC	Receiver operating characteristics
SDQ	Strengths and Difficulties Questionnaire
SDQ-Fin	Finnish version of the SDQ

Abstract

The overall aim of the dissertation was to assess suitable methods for detecting young children's mental health problems in primary health care in a multi-informant context consisting of the children, their parents, public health nurses and preschool and school teachers. More precisely, the study focused on exploring the psychometric properties of the Strengths and Difficulties Questionnaire (SDQ) among young Finnish children (I, II). The adjusted Finnish cut-offs of the SDQ were defined, and their capacity to identify the children suffering from psychiatric symptoms and disorders was explored (II). The focus of interest was also assessing, as briefly as possible, a simple and easy to use one-question screen for the child, the parent and the public health nurse (IV). In addition, the feasibility aspects of the Finnish version of the SDQ (SDQ-Fin) and of the child self-evaluation enquiry were evaluated (III, IV).

The target population of the study comprised 4–9-year-old children (n = 2,682) receiving regular health check-ups in child health clinics and school health care clinics from March 2008 to March 2009. The study was conducted as part of a project entitled "Developing children's mental health work, 2007–2009" in the Pirkanmaa and South Karelia hospital districts. In the first phase of the study, multi-informant questionnaire assessments were conducted in the context of health check-ups: the SDQs were completed by parents and by preschool and school teachers; the one-question screen was filled in by parents and public health nurses, and children filled in the self-evaluation enquiry. In the second phase, a stratified subgroup of the participating children (n = 646) were invited to the diagnostic interview of the Development and Well-Being Assessment (DAWBA) after the check-up visit. Thirdly, feedback questionnaires on the feasibility of the SDQ-Fin and the child's self-evaluation enquiry were collected.

The SDQ-Fin had accurate reliability properties of internal consistency and inter-rater and test-retest reliability. Significant and clinically important differences were found in the distributions of the SDQ-Fin scores between parent and teacher reports and between genders and age groups of the children. The adjusted lower cut-off was 9/10 and the higher cut-off 11/12 for the parent- and teacher-rated SDQ-Fin total scores. The sensitivity of the adjusted higher cut-off of the SDQ-

Fin total score was 90% in parent reports and 70% in teacher reports; the respective specificities were 74% and 66%. The SDQ-Fin had a good capacity for discriminating between the children with low risk and high risk for a psychiatric disorder.

The one-question screen had fairly good inter-rater reliability between the parents' and public health nurses' perceptions. The sensitivities of the one-question screen were 65% for the parents', 68% for the public health nurses' and 79% for their combined reports; the respective specificities were high. Difficulties identified by parents and nurses were strongly related to child psychiatric disorders. Of the young children, 2–5% reported a low mood and negative expectations, which was related to a twofold risk for any psychiatric disorder and a threefold risk for an emotional disorder and negative situational family factors. The SDQ-Fin was found to be a feasible method, and it had positive effects on cooperation between the parents and professionals in assessing children's mental health. The child's self-evaluation enquiry was evaluated to *be* an appropriate method and not burdensome in assessing the psychosocial well-being of the children.

The parent- and teacher-rated SDQ-Fin was found to be a reliable, valid and feasible method in detecting children's mental health problems among 4–9-year-olds visiting for regular health check-ups. As an important clinical implication, the adjusted cut-offs on the SDQ-Fin for young children were defined, and they had a high sensitivity in identifying the children at high risk for a psychiatric disorder. The SDQ-Fin can thus be recommended for routine clinical use in the context of children's regular health check-ups when it is ensured that adequate treatment and help are offered for those children identified with mental health problems. The one-question screen for parents and public health nurses showed good reliability and validity properties, and it can thus be suggested as a first-stage screening method for professionals evaluating the need for a more comprehensive assessment of the mental status and functioning of the child. The children's self-evaluation of emotional well-being brought clinically relevant information complementary to adult reports on the risk of mental health problems and especially emotional problems. These findings emphasise the necessity of the multi-informant approach in detecting children's mental health problems using standardised and culturally valid methods.

Key words: children, mental health, child psychiatry, screening, detecting, the Strengths and Difficulties Questionnaire, questionnaire, psychometric properties, reliability, validity, feasibility, self-evaluation, Finnish

Tiivistelmä

Väitöstutkimuksen tavoitteena oli tutkia ja arvioida lasten mielenterveysongelmien tunnistamiseen soveltuvia menetelmiä perusterveydenhuollossa yhteistyössä lapsen, vanhempien, terveydenhoitajien ja päivähoidon sekä koulun opettajien kanssa. Tutkimus kohdentui Vahvuudet ja vaikeudet -kyselyn (Strengths and Difficulties Questionniare, SDQ) psykometristen ominaisuuksien arviointiin suomalaisilla lapsilla (I, II). Tutkimuksessa määritettiin SDQ -kyselyn katkaisupistemäärien raja-arvot suomalaisaineistossa ja arvioitiin kyselyn kapasiteettia tunnistaa psyykkisesti oireilevat ja psykiatrisista häiriöistä kärsivät lapset (II). Kiinnostuksen kohteena oli lisäksi kehittää ja arvioida mahdollisimman lyhyt, yksinkertainen ja helppokäyttöinen yhden tai kahden kysymyksen seula lapselle, vanhemmalle ja terveydenhoitajalle (IV). Tutkimuksessa arvioitiin myös suomenkielisen SDQ -kyselyn (SDQ-Fin) ja Lapsen oma-arvio hyvinvoinnistaan -kyselyn käyttökelpoisuutta (III, IV).

Tutkimusotos koostui 4─9-vuotiaista lapsista (n = 2682), jotka kävivät lastenneuvolan tai kouluterveydenhuollon terveystarkastuksessa maaliskuun 2008 ja maaliskuun 2009 välisenä aikana. Tutkimusaineisto koottiin ”Lasten mielenterveystyön kehittäminen 2007─2008” hankkeen yhteydessä Pirkanmaan ja Etelä-Karjalan sairaanhoitopiireissä. Tutkimuksen ensimmäisessä vaiheessa koottiin kyselylomakkeita terveystarkastusten yhteydessä: vanhempien ja päivähoidon sekä koulun opettajien täyttämät SDQ -lomakkeet, yhden kysymyksen seula vanhempien ja terveydenhoitajan täyttämänä sekä Lapsen oma-arvio hyvinvoinnistaan -kysely. Tutkimuksen toisessa vaiheessa määritetty osaotos (n = 646) tutkimukseen osallistuneista kutsuttiin terveystarkastuksen jälkeen lastenpsykiatriseen diagnostiseen arvioon, Kehityksen ja hyvinvoinnin arviointi -haastatteluun (Development and Well-Being Assessment, DAWBA). Kolmannessa vaiheessa koottiin kyselylomakkeilla palautetta SDQ-Fin kyselyn ja Lapsen oma-arvio hyvinvoinnistaan -kyselyn käyttökelpoisuudesta.

Reliabiliteetin eli toistettavuuden osa-alueet toimivat SDQ-Fin -kyselyllä hyvin mittarin sisäisen yhdenmukaisuuden (internal consistency) ja mittaajien välisten arvioiden yhdenmukaisuuden (inter-rater reliability) osalta sekä testi-uusinta tutkimusasetelmassa (test-retest reliability). SDQ-Fin pistemäärien jakaumissa oli

merkittäviä ja tilastollisesti merkitseviä eroja vanhempien ja opettajien arvioissa sekä lasten sukupuolten ja ikäluokkien välillä. Määritetyt raja-arvot vanhemman ja opettajan SDQ-kyselyn kokonaispistemäärille olivat alemman katkaisupisteen osalta 9/10 ja ylemmän katkaisupisteen osalta 11/12. Ylemmän SDQ-Fin katkaisupisteen sensitiivisyys vanhemman kyselylle oli 90 % ja opettajan kyselylle 70 %, vastaavat spesifisyys arvot olivat 74 % ja 66 %. SDQ-Fin erotteli hyvin toisistaan ne lapsiryhmät, joilla oli matala ja korkea psykiatrisen häiriön riski.

Vanhempien ja terveydenhoitajien vastaukset yhden kysymyksen seulassa olivat melko yhdenmukaisia. Yhden kysymyksen seulan sensitiivisyys oli vanhempien arviossa 65 %, terveydenhoitajien arviossa 68 % ja molempien vastaajien yhdistetyssä arviossa 79 %, vastaavat spesifisyysarvot olivat korkeita. Vanhempien ja terveydenhoitajien tunnistamat lasten vaikeudet olivat voimakkaasti yhteydessä lapsen psyykkiseen häiriöön. Lapsista 2─5 % raportoi alhaista mielialaa ja negatiivisia tulevaisuuden odotuksia ja nämä asiat olivat yhteydessä kaksinkertaiseen riskiin psykiatriselle häiriölle sekä kolminkertaiseen riskiin lapsen tunne-elämän häiriölle ja kielteisille perhetekijöille. Lisäksi SDQ-Fin kyselyn arvioitiin olevan käyttökelpoinen menetelmä ja lisäävän vanhempien ja työntekijöiden yhteistyötä arvioitaessa lapsen psykososiaalista hyvinvointia. Lapsen oma-arvio hyvinvoinnistaan arvioitiin käyttökelpoiseksi ja ei kuormittavaksi menetelmäksi.

Tämä väitöstutkimus osoitti vanhemman ja opettajan vastaaman Vahvuudet ja vaikeudet -kyselyn (SDQ-Fin) olevan luotettava, pätevä ja käyttökelpoinen menetelmä 4─9-vuotiaiden lasten mielenterveysongelmien tunnistamisessa terveystarkastuksissa. SDQ-kyselyn pistemäärien katkaisupisteiden määrittämistä suomalaisille lapsille voi pitää tutkimuksen merkittävänä kliinisenä sovelluksena. Näillä katkaisupisteillä SDQ-Fin kyselyllä oli korkea sensitiivisyys kohonneen psykiatrisen häiriön riskin tunnistamisessa. SDQ-Fin kyselyä voi suositella käytettäväksi lasten terveysseurannassa silloin kun huolehditaan, että oireileville lapsille tarjotaan apua ja asianmukaista hoitoa. Vanhemman ja terveydenhoitajan yhden kysymyksen seulan hyvät reliabiliteetti ja validiteetti ominaisuudet puoltavat sen käyttöä ensiarviona ohjaamassa terveydenhuollon työntekijän tarkempaa harkintaa lapsen psyykkisen voinnin ja toimintakyvyn tutkimisesta. Kysymällä lapselta hänen omaa arviotaan hyvinvoinnistaan saadaan kliinisesti merkittävää ja aikuisten arvioita täydentävää tietoa lapsen riskistä psykiatriseen sairastavuuteen ja erityisesti riskistä tunne-elämän ongelmiin. Tutkimustulosten perusteella on erityisen tärkeää, että lasten mielenterveysongelmien tunnistamisessa huomioidaan usean tahon arviot lapsen tilanteesta käyttämällä standardisoituja menetelmiä, joiden soveltuvuus kyseisessä kulttuurissa on arvioitu.

1 Introduction

Several factors argue for an early detection of children's mental health problems: children have high prevalence rates of psychiatric symptoms and disorders; children's mental health problems are known to have high continuity to adolescence and adulthood. Early referral and care most likely improves children's mental health prognosis, and the prevention of mental health disorders has been considered lucrative. Before anything, the early identification of children's need for psychosocial help should be premised based on their human rights. These aspects will be briefly introduced in the review.

In Finland, public child health care and school health care are established parts of municipal primary services, and are responsible for monitoring and supporting the development and health of children and the well-being of their families. In Finland, virtually entire age groups of children participate in the regularly administered health check-ups in child health clinics from birth until six years of age and after that in school health care clinics. There was lack of national guidelines and norms on monitoring children's mental health before the 2011 government decree (Finlex, 338/2011). These national recommendations emphasise the comprehensive evaluation of child and family well-being in extensive health assessment check-ups at least five times between infancy and the end of primary school. The aspects of multi-informant approach in identifying children with psychosocial difficulties and providing early care and support are highlighted in the decree.

Screening and health examinations are not distinguished clearly in public discussion (Sauni et al., 2014). The regularly administered children's health check-ups include elements of screening. However, it is evident that the context of assessing screening tests and developing screening programmes involves a complex of issues (Hakama and Malila, 2008). The health care screening programmes are steered nationally by the Ministry of Social Affairs and Health (Mäkelä et al., 2014) and by government decree (Finlex, 339/2011). Only some of the ten principles of screening for disease suggested by the World Health Organisation (Wilson and Jungner, 1968) were assessed and discussed in the present study. This dissertation

focused on assessing reliable, valid and feasible methods for detecting and monitoring children's mental health problems.

Standardised methods are, however, not yet established practice in assessing Finnish children's mental health in primary health care. Standardised rating scales, including questionnaires, are acknowledged to help detect children's mental health problems. Standardised questionnaires can ensure systematic assessments of symptoms and provide quantifiable information on the presence, frequency and severity of symptoms (Myers and Winters, 2002). In addition, using standardised rating scales allows comparison with repeated measurements, comparison with peers, comparison with overall population, and cross-cultural comparison (Myers and Winters, 2002). Standardised methods make it possible to monitor population health. In the primary health care system, however, the regular and comprehensive use of standardised questionnaires is rare (Batty et al., 2013; Gold et al., 2009). Thus, there is a current need for research into suitable methods for detecting children's mental health problems.

When assessing children's mental health in front-line services, the questionnaires need to be short and easy to use and interpret, in addition to having accurate psychometric properties. This kind of standardised method is in short supply, however. With the lack of suitable methods, professionals seem to have their own practices of asking children and parents ordinary questions such as "How are you?" and "Have you perceived any difficulties or do you have any concerns about your child?" Only a few studies have examined how valid and relevant such questions are in detecting children's mental health problems.

The Strengths and Difficulties Questionnaire (SDQ) is an internationally used and studied brief questionnaire for assessing children's and adolescents' mental health in community and clinical settings (R. Goodman, 1999; R. Goodman, 2001). The present study was based on the need to study the reliability and validity of the SDQ in young, under-ten-year-old Finnish children because the psychometric properties of the method had only been studied among older school-aged children and adolescents (Koskelainen, Sourander and Kaljonen, 2000; Koskelainen, Sourander and Vauras, 2001).

The overall aim of the present study was to assess suitable methods for detecting 4–9-year-old children's mental health problems in primary health care in a multi-informant context consisting of the children, their parents, public health nurses and preschool and school teachers. More precisely, the study focused on exploring the psychometric properties and feasibility of the Strengths and Difficulties Questionnaire in assessing the mental health of young Finnish children

in a community sample. In addition, the focus of interest was on assessing a brief, simple and easy-to-use one-question screen for children, parents and public health nurses.

2 Review of the literature

2.1 Rationale for detecting early mental health problems in children

2.1.1 Prevalence of children's mental health problems

Mental health problems occur commonly among children of all ages. Of 5–17-year-old children and adolescents, 3–18% have been found to suffer from a psychiatric disorder causing significant functional impairment (Costello, Egger and Angold, 2005; Ford, Goodman and Meltzer, 2003; Merikangas, He, Brody et al., 2010). The reported prevalence rates have varied widely in differing study samples and depend on the measures used in assessing psychopathology, the severity of the scoring criteria, and whether functional impairment is included or ignored (Costello et al., 2005). In addition, cross-cultural differences in the prevalence rates of child psychiatric disorders assessed by the same diagnostic measure have ranged from 2% to 17% (A. Goodman et al., 2011).

The prevalence rates of child psychiatric disorders and patterns of comorbidity among under-school-aged children have corresponded to the prevalence rates among older children (Egger and Angold, 2006). According to present knowledge, child psychiatric disorders can already be diagnosed from the age of one and a half or two years on (Egger and Angold, 2006; Skovgaard, Houmann, Landorph and Christiansen, 2004; Skovgaard, Houmann, Christiansen and Andreasen, 2005). Epidemiological studies examining the prevalence of child psychiatric disorders according to structured diagnostic interviews have already been conducted among very young children between the ages of 18 months and five years. In a Norwegian community sample of four-year-old children, the prevalence rate for any child psychiatric disorder was 7% and comorbidity was common (Wichstrom et al., 2012). In a Romanian sample of children aged 18–60 months, the prevalence of disorders was 9% (Gleason et al., 2011), and the prevalence of psychopathology in 18-month-old children was 16–18% in a Danish cohort study (Skovgaard et al., 2007). Most frequent diagnoses were relationship disorders (9%) and regulatory

disorders (7%) according to DC 0–3 (Zero To Three, 1994), and neurodevelopmental disorders (7%), emotional and behavioural diagnoses (4%) and eating disorders (3%) according to ICD-10 (World Health Organisation, 1994) (Skovgaard et al., 2007).

In the Nordic countries, the prevalence rates of child psychiatric symptoms and disorders has generally been lower than in the United Kingdom, in the United States and in many other countries (Achenbach et al., 2008; Elberling, Linneberg, Olsen, Goodman and Skovgaard, 2010; A. Goodman et al., 2011; Heiervang et al., 2007; Heiervang, Goodman and Goodman, 2008; Jozefiak, Larsson, Wichstrom and Rimehaug, 2012; Koskelainen et al., 2000; Kristensen, Henriksen and Bilenberg, 2010; Obel et al., 2004; Rescorla et al., 2007; Wichstrom et al., 2012). This might reflect genuine cross-cultural differences in the mental health of children but also informants' different reporting styles across cultures. The findings of prevalence rates in the Nordic countries have been relatively consistent.

In Finland, 24% of children were evaluated to have psychiatric symptoms according to the Rutter questionnaires, and 9% were in need of psychiatric treatment based on a diagnostic interview in an epidemiological sample of 8–9-year-old children (n = 5813) in 1989 (Almqvist, Kumpulainen et al., 1999; Almqvist, Puura et al., 1999). In three cross-sectional representative samples of eight-year-old children (in 1989, [n = 986]; in 1999, [n = 831]; and in 2005, [n = 870]), 16–24% of the boys and 10–12% of the girls were reported by parents or teachers using Rutter questionnaires to have emotional or behavioural symptoms (Sourander, Niemelä, Santalahti, Helenius and Piha, 2008). Among 12-year-old children, 6% have been reported to be suffering from behavioural or emotional problems according to parent-rated Child Behaviour Checklist questionnaires (n = 908) (Pihlakoski et al., 2004). In the Child Health Monitoring Development Pilot Study (2007–2008), public health nurses reported at least minor concerns on the psychosocial development and health of 15% of five-year-old children (n = 217) and 12–13% of primary-school-aged children (n = 444) (Mäki et al., 2010). In this above-mentioned sample, parents reported symptoms of deviant behaviour in 12–17% of boys and 5–8% of girls and low mood in 2–3% of boys and 3–5% of girls. Among three-year-old children (n = 374), the prevalence of parent-rated behavioural and emotional difficulties was 8% (Sourander, 2001). General practitioners have evaluated 3% of the 4–18-month-old infants (n = 363) as showing signs of social withdrawal (Puura et al., 2010).

2.1.2 Continuity of mental health problems

Mental health disorders in adults commonly have their onset already in childhood (Costello et al., 2005; Kessler et al., 2005; Merikangas, He, Burstein et al., 2010). The continuity of psychopathology has been found moderate to strong in prospective studies among pre-schoolers (Kerr, Lunkenheimer and Olson, 2007; Klein, Otto, Fuchs, Reibiger and von Klitzing, 2014). The pathways of symptoms and global functioning among pre-schoolers and from early childhood to adolescence have, however, been complex (Kerr et al., 2007; Klein et al., 2014; Pihlakoski et al., 2006). Still, the high continuity of externalising symptoms has been replicated in longitudinal studies (Kerr et al., 2007; Pihlakoski et al., 2006).

Of the Finnish three-year-old children having parent-rated emotional or behavioural difficulties, almost 30% were perceived as still having difficulties at the age of 12 (Pihlakoski et al., 2006). In this longitudinal study sample, aggressive behaviour had the strongest stability among boys and girls from age three to 15 years (Pihlakoski et al., 2006; Sourander et al., 2006), and it predicted a poor sense of coherence at 18 years of age (Honkinen et al., 2009).

The childhood predictors of later psychopathology and other adverse outcomes have been assessed in the longitudinal samples of the Finnish 1981 Birth Cohort Study and in the Northern Finland Birth Cohort 1986 study. Psychopathology at the age of eight years has been found to be a long-lasting risk factor for severe psychiatric disorders requiring hospitalisation and antidepressant medication (Gyllenberg et al., 2010; Gyllenberg et al., 2011). In the "From a Boy to a Man Study", boys with combined conduct and internalising problems at age eight had the highest longitudinal risk of psychiatric disorders, criminal offenses and self-reported problems (Sourander et al., 2007). In addition, childhood psychopathology among boys has been found to be a risk factor for drug offences at age 18, adult smoking and a poor sense of coherence (Niemela et al., 2008; Niemela et al., 2009; Ristkari et al., 2009). The externalisation of problems during childhood has preceded adolescent substance use in both genders; among boys, substance use was also associated with criminal offences (Miettunen et al., 2014). Girls having externalising difficulties at the age of eight had an increased risk of becoming teenage mothers (Lehti et al., 2012).

In addition, childhood externalising and internalising psychopathologies were found to be associated with adverse health behaviours and health outcomes in midlife, as well as with increased long-term mortality (Jokela et al., 2009; Stumm et al., 2011).

2.1.3 Referral to care and use of services

The majority of children with mental health symptoms and disorders have not received mental health services (Ikäheimo, 1999; Pihlakoski et al., 2004; Santalahti, Sourander and Piha, 2009; Sayal and Ford, 2010; Sourander et al., 2008; Wichstrom, Belsky, Jozefiak, Sourander and Berg-Nielsen, 2014). Of the children with emotional or behavioural problems, 11% at the age of four and 25% at the age of seven had received mental health services (Wichstrom et al., 2014). Among 12-year-old Finnish children, 7% had received some health and social services because of behavioural or emotional difficulties, and half of them had received mental health services (Pihlakoski et al., 2004). However, referral to outpatient child psychiatric treatment has continuously increased in Finland over the last decade (SOTKAnet; Santalahti et al., 2009; Sourander et al., 2008). In addition, children with perceived emotional and behavioural problems have often received some support at school (Heiervang et al., 2007; Sourander et al., 2008).

The nature of a child's psychopathology and functional impairment affects help-seeking and referral to care. Mental health service use among school-aged children has been found to be most common in cases of hyperactivity (75%) and conduct disorders (41%) but rare in the case of emotional disorders (13%) (Heiervang et al., 2007). The behavioural but not emotional difficulties of the child have been associated with and predicted the use of services at different ages (Pihlakoski et al., 2004; Puura et al., 1998; Wichstrom et al., 2014). In addition, a child's functional impairment causing parental distress has predicted help-seeking (Pihlakoski et al., 2004; Wichstrom et al., 2014). Girls have been referred to mental health services less frequently than boys (Sourander et al., 2008; Wichstrom et al., 2014).

The process of referral has been proposed to consist of several stages, from recognition, help-seeking and decisions, to referral (Zwaanswijk et al., 2003). The advance along these stages is influenced by the numerous characteristics of the child, parents, family, environment, availability of services and professionals (Ikäheimo, 1999; Zwaanswijk et al., 2003). For example, progress at the referral process stages has been found to be associated with the child's physical illness and factors connected to the parents' psychiatric, marital and family problems (Ikäheimo, 1999).

2.1.4 Ethical and economic aspects

Children have the right to well-being in the present and to healthy development, including psychosocial development and health. The Convention on the Rights of the Child (CRC) was ratified in Finland in 1991 (UNICEF, http://www.unicef.org/crc/). According to the CRC (Article 24), "States Parties recognize the right of the child to the enjoyment of the highest attainable standard of health and to facilities for the treatment of illness and rehabilitation of health." The necessity of developing early detection and care of children's mental health disorders is thus ethically justified.

Mental health problems have been identified as the most significant health problem of children (World Health Organization, 2004). There are preventive mental health interventions for children and their parents that have been documented to be both effective and cost-effective in improving the outcomes of children and families (Karoly, Kilburn and Cannon, 2005; National Collaborating Centre for Mental Health, 2010; World Health Organization, 2004). Early childhood interventions from birth to five years of age targeted at families with risk factors for healthy child development have shown convincing evidence of favourable outcomes in the lives of participating children compared with control groups in longitudinal data (Karoly et al., 2005; Olds et al., 1997; World Health Organization, 2004). In addition, the effectiveness of behavioural and cognitive-behavioural parenting interventions has been demonstrated in the prevention and treatment of early onset conduct problems in children aged three to twelve years (Furlong et al., 2012; National Collaborating Centre for Mental Health, 2010). In Finland, there has also been widespread interest in assessing the effectiveness of children's mental health interventions (Aronen and Arajärvi, 2000; Björklund et al., 2014; Laajasalo and Pirkola, 2012; Punamäki et al., 2013; Solantaus et al. 2010; Williford et al., 2012).

From the perspective of national health, the World Health Organization (WHO) has named the prevention of mental disorders as the most central challenge (World Health Organization, 2004). Early-onset mental health disorders are known to be associated with substantial societal costs in terms of long-lasting risks for mental and physical disorders, adverse life course outcomes and reduced achievements in education and financial status (Kessler et al., 2009). Furthermore, early intervention in children's mental health is evaluated to show significant lifetime economic returns by an average benefits-costs ratio of six to one

(Campion, Bhui, Bhugra and European Psychiatric Association, 2012; Karoly et al., 2005).

Other ethical aspects of the effects of detecting children's mental health problems also require consideration. Detection of early mental health problems should not cause harm, such as unnecessary concerns to the child or the family, for example. The detection should not classify and stigmatise the children, but benefit the child's healthy development as an individual. The child and the family have the right to know the aim of the detection and the results of the assessment. For the children identified with high risk for psychiatric disorder, there should be facilities for the more comprehensive assessment of their mental health status, the evaluation of risk and protective factors for their development (World Health Organization, 2004) and adequate treatment (Wilson and Jungner, 1968).

2.2 Methods and clinical aspects in detecting children's mental health

2.2.1 Principles and challenges in primary health care

There are many challenges in ensuring early detection of children's mental health problems in primary health care. Firstly, few parents express their concerns about the mental health of their child to professionals (Dulcan et al., 1990; Sayal and Ford, 2010). When parents reported having these concerns, the concerns were often not confirmed by professionals (Reijneveld, de Meer, Wiefferink and Crone, 2008). Thirdly, parents have reported several barriers to seeking help: an insufficient length of visits in the primary care system, a discontinuity of care and contact with professionals, and psychological aspects such as embarrassment, the stigma of mental health problems, or concerns about being labelled with or receiving a diagnosis (Sayal and Ford, 2010).

From the professionals' point of view, assessing children's mental health is a complex issue. It might be difficult to identify psychopathology from the typical course of a child's psychosocial development (Angold and Egger, 2007). Again, if the child has socio-emotional or behavioural problems, they need to be considered in the context of the child's developmental level (Carter, Briggs-Gowan and Davis, 2004). In addition, a child's symptoms and level of functioning must first be evaluated in the context of the child's family and second in the context of other

significant social environments (Ederer, 2004). It is also necessary to assess children's psychosocial functioning in multi-axial terms instead of trying to capture a "present versus absent" assessment of the problems (Achenbach, McConaughy and Howell, 1987).

It is generally acknowledged that a multi-informant approach is a crucial principle in assessing children's mental health. However, discrepancies are common between different informants' evaluations of the child's psychopathology (De Los Reyes and Kazdin, 2005). A symptom is regarded as being present if any of the informants (child, parent or teacher) report it present (Angold, Egger 2007). In addition, the levels of agreement between informants vary with the broad spectrum of children's psychic symptoms (Ederer, 2004). Thus, the integration and interpretation of multi-informant data are challenging tasks. The findings of a meta-analysis of agreement levels between different informants by Achenbach et al. (1987) have been suggested as benchmarking the levels of cross-informant agreement in this context (R. Goodman, 2001; Stone, Otten, Engels, Vermulst and Janssens, 2010). In the analysed reports on children's behavioural and emotional problems, the mean correlation between similar types of informants (e.g. mothers and fathers) was $r = 0.60$ (Pearson correlation coefficient), between different kinds of informants (e.g. parent and teacher) $r = 0.28$, and between the subjects and other informants (e.g. child/adolescent and parent/teacher) $r = 0.22$ (Achenbach et al., 1987). All these correlations were statistically significant. The modest to moderate cross-informant agreement levels have since been replicated, and the findings have been interpreted as reflecting the perceived variations in the child's functioning in different surroundings (R. Goodman, 2001; Stone, Otten, Engels, Vermulst and Janssens, 2010).

The role of young children as informants in assessing their mental health has been unclear and vague. Parents' and teachers' reports are often administered, but young children are seldom asked to self-evaluate their well-being. Standardised self-reports are usually available for children over 11 years old (Achenbach TM, 2001; R. Goodman, 2001; Kovacs, 1992). However, especially children's emotional problems seem to be underestimated without self-reports and relying only on parental reports (Michels et al., 2013). Children report having emotional symptoms more commonly than perceived by parents and teachers (Ederer, 2004; Michels et al., 2013; Seiffge-Krenke and Kollmar, 1998; Van Roy, Groholt, Heyerdahl and Clench-Aas, 2010). In general, a higher level of agreement between children's and adults' reports has been found for externalising than for internalising problems

(Ederer, 2004). There seems to be a need for further developing appropriate self-reporting methods, also for young children.

2.2.2 Questionnaires and queries

Questionnaires are used in epidemiologic studies, screening and monitoring in normative settings, identifying children with psychosocial symptoms or at greater risk of a psychiatric disorder, and evaluating treatment outcome (Myers and Winters, 2002). It is important to select the test with the best psychometric properties and feasibility for the population and purpose in question (Myers and Winters 2002). As mentioned before, the comprehensive use of standardised questionnaires is rare in primary health care (Batty et al., 2013; Gold et al., 2009).

There are only a few short multi-dimensional questionnaires that have been widely reported on and have gained acceptance in children's front-line mental health services. The Achenbach System of Empirically Based Assessment (ASEBA) questionnaires, comprising the Child Behaviour Checklist (CBCL), the Teacher's Report Form (TRF) and the Youth Self-Report (YSR), have often been used as criteria in assessing the psychometric properties of other questionnaires (Achenbach and Rescorla, 2001). The ASEBA questionnaires, however, consist of so many items that they seem too burdensome for widespread use in primary practice. The Rutter questionnaires were long-established behavioural screening questionnaires for parents and teachers but are no longer commonly used (Elander and Rutter, 1996). Developed on the basis of the Rutter questionnaires, the Strengths and Difficulties Questionnaire (SDQ) incorporates additional items on psychopathology, a distinct dimension of prosocial behaviour and an assessment of global functioning (R. Goodman, 1997; R. Goodman, 1999; R. Goodman, 2001). The SDQ is a widely reported method for measuring children's mental health in community and in clinical settings, both for research and clinical purposes (Achenbach et al., 2008; Bourdon, Goodman, Rae, Simpson and Koretz, 2005; Du, Kou and Coghill, 2008; R. Goodman, Slobodskaya and Knyazev, 2005; Hawes and Dadds, 2004; Marzocchi et al., 2004; Obel et al., 2004; Rothenberger et al., 2008; Woerner et al., 2004). The Pediatric Symptom Checklist (PSC) is also a well-documented brief multi-dimensional questionnaire but not, to the author's knowledge, used in Finland (Jellinek, Murphy and Burns, 1986; Jellinek et al., 1988). Screening tools for very young children have also been developed and documented (Carter et al., 2004). The Brief Infant-Toddler Social-Emotional

Assessment (BITSEA) is currently undergoing a validation process in Finland (Briggs-Gowan, Carter, Irwin, Wachtel and Cicchetti, 2004; Haapsamo et al., 2009).

Asking parents briefly, in only one or a few questions, about their concerns or perceptions about their child's behaviour and emotions has been found to be useful in identifying high-risk children (Ford, Sayal, Meltzer and Goodman, 2005; A. Goodman and R. Goodman, 2011; R. Goodman, 1999). The first question of the SDQ impact supplement, which asks parents and teachers what their general perceptions are of the child's difficulties (in emotions, concentration, behaviour or social competence), identifies this almost as well as the entire SDQ assesses high-risk children (A. Goodman and R. Goodman, 2011; R. Goodman, 1999).

2.2.3 Methods in Finnish primary health care

Also in Finland, standardised questionnaires for assessing children's mental health have been used in child health clinics and school health care only rarely for children under 12 years old (Hakulinen-Viitanen, Pelkonen, Saaristo, Hastrup and Rimpelä, 2008). One-third of primary health care units have reported using some method or a questionnaire designed for assessing children's psychosocial health and need for support, but the methods were usually locally designed and not standardised (Rimpelä, Rigoff, Wiss and Hakulinen-Viitanen, 2006). The SDQ was introduced as a standardised questionnaire in evaluating children's psychosocial health in the Finnish handbook of child health assessment methods (Mäki, Wikström, Hakulinen-Viitanen and Laatikainen, 2011). The SDQ has also been evaluated by a Finnish network of experts and found suitable for identifying children's psychiatric symptoms in front-line clinical practice (TOIMIA, 2013, http://www.thl.fi/toimia/tietokanta/). The national criteria for specialised non-urgent child psychiatric care were validated in 2005 (Kaukonen et al., 2010).

Finnish general practitioners (GPs) have found their competence and skills to be inadequate in assessing children's mental health status and need for psychiatric treatment (Heikkinen, Puura, Ala-Laurila, Niskanen and Mattila, 2002). Of the GPs participating in the study, 40% reported being short of time at health check-ups (Heikkinen et al., 2002). GPs and public health nurses have very important roles in identifying children with mental health problems and referring them to care. More education is needed for these front-line professionals in the methods and clinical aspects of detecting children's mental health.

2.3 Measurement properties of assessment methods

There is broad variation in the terminology and definitions of specific measurement properties of assessment methods in child psychiatry and, generally, in the medical sciences. This makes it difficult for a clinician or a researcher to study the literature on assessing the measurement properties of different instruments and make comparisons between them. In particular, there seems to be a confusing variety of terms and definitions for the validity properties in the literature.

In the COSMIN study (Mokkink et al. 2010), consensus-based standards were searched to select the most important measurement properties and their adequate terms and definitions in the medical and health sciences. In addition, guidelines were drawn up on how the measurement properties should be evaluated. The taxonomy of measurement properties according to the COSMIN terminology (De Vet, Terwee, Mokkink and Knol, 2011; Mokkink et al., 2010) is represented in Table 2.3. The definitions of these measurement properties according to the COSMIN panel are briefly represented in the next chapters, i.e. 2.3.1, 2.3.2 and 2.3.3. In addition, the concept of feasibility is reviewed in Chapter 2.3.4.

Table 2.3. Taxonomy of measurement properties according to the COSMIN terminology (Mollink et al., 2010; De Vet et al., 2011).

Measurement properties	**Aspects of measurement properties**		
Reliability	Internal consistency		
	Inter-rater reliability		
	Test-retest reliability		
	Measurement error		
Validity	Content validity		
	Criterion validity	Concurrent validity	
		Predictive validity	
	Construct validity	Structural validity	
		Hypotheses testing	Convergent validity
			Discriminative validity
			Known groups validity
		Cross-cultural validity	
	Responsiveness		
Interpretability			

2.3.1 Definitions of reliability properties

Reliability has been defined as "the degree to which the measurement is free from measurement error" (Mokkink et al., 2010). Aspects of reliability can be assessed by repeated measurements using the same instrument in different circumstances. The concept of reliability represents variation in measurements from many sources: the measurement instrument, the respondents or observers, different surroundings and the time-points of the measurement.

In a multi-item instrument, internal consistency measures the inter-relatedness among the items. The correlations between items indicate whether the item is a part of the scale and to which extent the items assess the same construct. The best known parameter for assessing the internal consistency of a scale is Cronbach's alpha.

Inter-rater reliability is defined as repeated measurements with the same instrument on the same occasion by different individuals and, respectively, the intra-rater reliability on different occasions by the same individuals. The more commonly used term for intra-rater reliability is the test-retest. Different parameters are obtained for calculating these correlations, depending on whether the variables are continuous or categorical.

Test-retest reliability assesses the variation over time in repeated measurements by the same respondents. There is no rule for the time interval between the initial test and the re-test. In questionnaire studies, however, a time interval of two weeks has been suggested in order to find a balance between assessing the stability of the measurement and the stability of the assessed phenomena (De Vet et al., 2011).

The magnitude of measurement error is necessary information in measuring changes in health status but rarely reported in studies assessing psychometric properties of instruments. Thus, this reliability property is not reviewed here.

2.3.2 Definitions of validity properties

The COSMIN panel has defined the concept of validity as "the degree to which an instrument truly measures the construct(s) it purports to measure" (Mokkink et al., 2010). Three different kinds of the main aspects of the validity can be distinguished and, further, several subtypes for each of them; see Table 2.3.

In the textbook *Measurement in Medicine* (De Vet et al., 2011), many ideas and principles are incorporated into the concept of validity: the construct intended to be measured should be clearly described; knowledge about the construct drives formulation of testing hypotheses; the validity of a measurement instrument is population- and context-dependent (e.g. language and culture or form of administration); validation focuses not on the instrument itself but on the scores it produces in specific situations. In addition, the validation is defined as a continuous process of assessing the degree of validation of the measurement with the combination of various aspects of validity (De Vet et al., 2011).

The validation process for a measurement instrument starts with content validation, which means assessing whether the content has corresponds in a relevant and comprehensive way with the construct it is intended to measure (Mollink et al., 2010; De Vet et al., 2011). Content validation is based on a subjective judgment of how well the instrument reflects the construct (face validity), and no statistical testing is involved. An expert panel or the users of the method are asked to evaluate how adequately the instrument seems to reflect the assessed construct, to study the relevance and comprehensiveness of the questions or items of the instrument, and often also to compare the content with other measurement instruments assessing the same construct.

Criterion validity assesses how well the scores of the instrument agree with the scores on the gold standard (Mollink et al., 2010; De Vet et al., 2011). The gold standard is assumed to represent the true state of the construct of interest. In reality, a perfectly valid instrument for a gold standard does not exist. In order to be considered an appropriate instrument for the gold standard, information about the validity and reliability of the instrument must be provided. Concurrent validity considers the scores of the measurement instrument and the gold standard at the same time. Predictive validity assesses the extent to which the scores of the instrument predict the scores of the gold standard in the future.

Assessing criterion validity is often utilised for evaluative and diagnostic purposes. A hypothesis is needed to specify the extent of agreement between the scores of the instrument and the gold standard in order to study whether the instrument is sufficiently valid for its clinical purpose (De Vet et al., 2011). Statistical parameters often used in assessing the diagnostic accuracy include sensitivity, specificity, positive and negative predictive values (scales of dichotomous outcome; see Figure 2.3.2) and receiver operating characteristics curves (ROC) (dichotomous or continuous scales).

Table 2.3.2. Definitions of validity, sensitivity, and positive and negative predictive values (Altman, 1991; Santalahti, 1998; Uhari and Nieminen, 2001).

		Reference (gold standard, disease status) positive	negative	Total
Test result (screening test)	positive	a (true positive)	b (false positive)	a + b
	negative	c (false negative)	d (true negative)	c + d
	Total	a + c	b + d	n

Sensitivity = a / (a + c) = Proportion of patients with disease who have positive test result

Specificity = d / (b + d) = Proportion of those without the disease who have negative test result

Positive predictive value = a / (a + b) = Proportion of correctly diagnosed patients with disease in subjects with positive test results

Negative predictive value = d / (c + d) = Proportion of those without the disease in subjects with negative test results

Construct validity is provided when there is no gold standard. Construct validity is subdivided into three aspects: structural validity, hypothesis testing and cross-cultural validity (Mollink et al., 2010; De Vet et al., 2011). Structural validity uses factor analysis to assess how adequately the scores of the instrument reflect the dimensionality of the construct. In hypothesis testing, the relationships of the scores of the instrument under study are compared with the scores of other instruments or the differences in the scores of the instrument are assessed in the subgroups of patients. In convergent validity, a hypothesis formulates that the instrument measures constructs similar to those measured by another comparable instrument. In discriminant validity, it is hypothesised that the instrument measures constructs that are different from the comparison method. Known groups or discriminative validity assesses expected differences in the scores of the measurement instrument between the subgroups of patients.

Cross-cultural validity assesses differences between items or questions in the translated or culturally adapted instrument compared with the original version of the instrument (De Vet et al., 2011). The validation starts with an accurate translation process. Guidelines have been laid down that define the essential steps of the recommended stages of cross-cultural adaptation during the translation process of a questionnaire (Beaton, Bombardier, Guillemin and Ferraz, 2000).

Differences in the items may be induced by the translations or by differences in the cultural meanings of the language. In evaluating the construct validity of a cross-culturally adapted instrument, measurement invariances are assessed in order to find whether the items after translation have retained the same meanings as in the original version (De Vet et al., 2011).

Responsiveness is considered as an aspect of validity in a longitudinal context. Responsiveness refers to the ability of an instrument to detect change in the construct over time (Mokkink et al., 2010). The concept is not reviewed here more precisely because it lies outside the focus of the dissertation.

2.3.3 Definition of interpretability

Interpretability is not a measurement property, but the concept is included in the COSMIN taxonomy because of its importance in the well-considered use of an instrument in clinical practise and in research. It is defined as "the degree to which one can assign qualitative meaning to an instrument's quantitative scores or change in scores" (Mokkink et al., 2010). Interpretability refers to what the scores of an instrument mean in general. It is important to examine the distribution of the scores in order to know in what kind of a population the scores are to be interpreted. Also, interpreting the reliability and validity properties of an instrument necessitates information about the distributions of the scores in the population in question (De Vet et al., 2011). The interpretability of changes in scores in a longitudinal context can be evaluated using a number of specific methods and statistical parameters not represented here.

2.3.4 The concept of feasibility

A commonly shared view is that, in addition to possessing adequate psychometric properties, a measurement instrument has to be suitable for routine use before it is accepted by the users and respondents in everyday clinical practice. No consensus on the concept has yet been found, but several important aspects and elements have been claimed as necessary for an instrument to be feasible (Fitzpatrick, Davey, Buxton and Jones, 1998; Myers and Winters, 2002; Slade, Thornicroft and Glover, 1999; Slade et al., 2001). It has been stated that feasibility should be systematically investigated before a measurement instrument can be recommended for routine clinical use (Slade et al., 2001).

The feasibility of an instrument has been defined as "the extent to which an assessment is suitable for use in a routine, sustainable and meaningful basis in typical clinical settings, when used in a specific manner and for a specific purpose" (Slade et al., 1999). A feasible measurement instrument is suggested by Slade et al. (1999, 2001) to incorporate six properties: 1) brief (looks short, easy to use), 2) simple to use (no training required) and to complete (meaning of ratings is explicit), 3) relevant to clinical judgement and to respondents, 4) acceptable to the profession (what is measured, how the instrument is administrated and what is the purpose of the measurement), 5) available and 6) valuable (the benefits of the measurement outweigh the costs and using the measurement results in a more comprehensive or detailed assessment than without it).

According to Fitzpatrick et al. (1998), the low response rates of a measurement may reflect the low acceptability of the measurement method among the patients. In addition, the important feasibility properties of an instrument from the clinical point of view should be easy to administer, process and interpret; a translated or culturally adapted version should be available; and norms or cut-offs for the scores should be available.

2.4 Review of the measurement properties of the SDQ

There is great deal of published information available on the SDQ, but only the most essential studies on the psychometric properties and feasibility of the method are reviewed here. The review focuses on the following studies: representations of the original development and testing of the psychometric properties of the SDQ, studies on children under 12 years old, studies conducted on community samples, and earlier studies on the SDQ parent and teacher reports in Finland and other Nordic countries.

2.4.1 Review of the reliability aspects of the SDQ

2.4.1.1 Internal consistency of the SDQ

The reliability aspect of the internal consistency of the SDQ total score has, in general, reached well-accepted values of Cronbach's alpha (R. Goodman, 2001; Koskelainen et al., 2000; Stone et al., 2010), with some exceptions (Dave, Nazareth,

Senior and Sherr, 2008; Du et al., 2008); see also Table 2.4.1. A guideline for a well-accepted value of the Cronbach's alpha is between 0.70 and 0.90 (De Vet et al., 2011). In the British epidemiologic study, the values of the internal consistency were $\alpha = 0.82$ for the parent-reported SDQ, $\alpha = 0.87$ for the teacher-reported SDQ and $\alpha = 0.80$ for the self-reported SDQ (R. Goodman, 2001). In a meta-analysis of 26 studies involving children under 12 years of age (Stone et al., 2010), the weighted mean of internal consistencies for the parent-reported SDQ was $\alpha = 0.81$ (0.53–0.84) and for the teacher-reported SDQ $\alpha = 0.82$ (0.62–0.85). In the Finnish study involving 7–12 year-old children, the alpha was 0.71 for all informants (parent, teacher, adolescent) (Koskelainen et al., 2000). Among Finnish adolescents (13–19 years old), the internal consistency for the self-reported total scores was $\alpha = 0.64$ (Koskelainen et al., 2001).

The internal consistencies of the SDQ subscales have varied a great deal with the study populations. The hyperactivity subscale has most commonly had the highest alphas, and there has been more variation in which subscale has the lowest alpha. In the earlier Nordic studies, the lowest internal consistencies in the parent- and teacher-reported SDQs have been in the conduct subscale (Koskelainen et al., 2000; Malmberg, Rydell and Smedje, 2003; Niclasen et al., 2012; Sanne, Torsheim, Heiervang and Stormark, 2009) and the highest in the hyperactivity subscale. In Danish cohorts and in a German sample, the internal consistencies were higher for boys than for girls (Niclasen et al., 2012; Rothenberger et al., 2008). In addition, teacher-reported SDQ scales have usually had higher internal consistencies than parent-reported ones (Niclasen et al., 2012; Stone et al., 2010).

2.4.1.2 Inter-rater reliability of the SDQ

The results of inter-rater reliability of the SDQ total scores between different pairs of informants have shown moderate correlations (R. Goodman, 1997; R. Goodman, Meltzer and Bailey, 1998; R. Goodman, 2001; Koskelainen et al., 2000; Stone et al., 2010; see also Table 2.4.1). The correlation r value (Pearson's and Spearman's correlation coefficients) may take on a range of values from -1 to 0 to +1, where ±1 indicates perfect linear positive or negative association between the two variables and $r = 0$ indicates no association between the measured variables (Mukaka, 2012; Taylor, 1990). Guidelines for roughly interpreting the size of correlation have been suggested: $r \leq 0.30$ or 0.35 represents low or weak correlation; $r > 0.30$ or 0.36 to $r = 0.67$ or 0.70 indicates moderate correlation; and $r \geq 0.68$ or 0.70 represents strong or high correlation (Mukaka, 2012; Taylor, 1990).

In the original reliability studies of the SDQ, the correlations of the inter-rater agreement between parents and teachers have varied between 0.43 and 0.62 (R. Goodman, 1997; R. Goodman et al., 1998). In an earlier Finnish study on school-aged children (Koskelainen et al., 2000), parent-teacher agreement was r = 0.44 (Pearson correlation coefficient), and in the Danish cohort studies the agreement varied between r = 0.45 and r = 0.53 (Niclasen et al., 2012).

Of the subscales, hyperactivity has reached the highest inter-rater reliability values (R. Goodman, 2001; Koskelainen et al., 2000; Niclasen et al., 2012; Sanne et al., 2009; Stone et al., 2010; Van Leeuwen K. and Bosmans G., De Medts L:, Braet C., 2006; van Widenfelt, Goedhart, Treffers and Goodman, 2003). In a review of under-12-year-old children the lowest weighted correlations of the parent and teacher inter-rater agreement were 0.26 (0.22–0.30) in the prosocial subscale and 0.28 (0.23–0.41) in the emotional symptoms (Stone et al., 2010).

The values for inter-rater agreement between mothers and fathers have seldom been reported. In a British study (Dave et al., 2008), the interparental agreement for the SDQ total score was poor: 0.27 (kappa coefficient). The respective agreement between mothers' and fathers' ratings was considered moderate (r = 0.53–0.61) in a Chinese study (Mellor, Wong and Xu, 2011). The highest agreement ratio was found in externalising problems in both studies. In addition, inter-rater reliability between mother and father was generally higher for boys than for girls.

The results for inter-rater agreement between adolescent and parent and between adolescent and teacher are not reviewed here.

2.4.1.3 Test-retest reliability of the SDQ

The test-retest reliability values, the correlations of the SDQ total scores in repeated measurements, have varied between moderate and strong (Du et al., 2008; R. Goodman, 2001; Hawes and Dadds, 2004; Muris, Meesters and van den Berg, 2003; Stone et al., 2010); see also Table 2.4.1. In a British epidemiologic study, the stability of the SDQ total scores according to the parent-ratings was 0.72 (Pearson correlation) and according to the teacher-ratings 0.80 after four to six months (R. Goodman, 2001). For the extended version of the SDQ, the test-retest reliability of the parent-rated SDQ total scores was 0.85 (intraclass correlation) and of the impact scores 0.54 in the time interval of three to four weeks (R. Goodman, 1999). The time interval between the initial and second measurement has varied from some weeks to one year in the reviewed studies (Table 2.4.1). The teacher-reported SDQ scores have had higher test-retest reliability correlations than the parent-rated

scores (Stone et al., 2010). In Finland, the test-retest reliability of the SDQ had not been tested before the present study.

Table 2.4.1. Summary of the SDQ reliability studies included in the review.

Study/ author	Country	N	Child's age	Informant[1]	Internal consistency (α)[2]	Inter-rater reliability (correlation)[3]	Test-retest (correlation)[3,4]
Goodman (1997)	United Kingdom	403	4–16	P, T		Total score: 0.62 (r) Subscores: 0.37–0.65	
Goodman (1998)	United Kingdom	199	11–16	P, T, (S)		Total score: 0.43 (EQS) Subscores: 0.14–0.38	
Goodman 1999	United Kingdom	34	5–15	P			Total score:0.85 (ICC) Impact: 0.54 (3–4 weeks)
Smedje et al. (1999)	Sweden	900	6–10	P	Total score: 0.76 Subscores: 0.51–0.75		
Koskelainen et al. (2000)	Finland	735	7–15	P, T, (S)	Total score: 0.71 Subscores:0.59–0.86	Total score: 0.44 (r) Subscores: 0.29-0.45	
Goodman (2001)	United Kingdom	10,438	5–15	P, T, (S)	Total score:0.80–0.87 Subscores:0.57–0.88 Impact: 0.85	Total score: 0.46 (r) Subscores: 0.27–0.48 Impact: 0.37	Total score: 0.72–0.80 Subscores:0.57–0.82 Impact: 0.57–0.68 (4–6 months)
Hawes & Dadds (2003)	Australia	1359	4–9	P, (T)			Total score: 0.77 (r) Subscores: 0.61–0.77 Impact: 0.63 (12 months)

(cont.)

Table 2.4.1. (continued)

Study/ author	Country	N	Child's age	Informant[1]	Internal consistency (α)[2]	Inter-rater reliability (correlation)[3]	Test-retest (correlation)[3,4]
Muris et al. (2003)	Netherlands	562	9–15	P, (S)	Total score: 0.80 Subscores: 0.55–0.78		Total score:0.88 (ICC) Subscores: 0.75–0.91 (2 months)
van Widenfelt et al. (2003)	Netherlands	300	8–16	P, T	Total score:0.81–0.88 Subscores: 0.57–0.89	Total score: 0.52 (r) Subscores: 0.23–0.54	
Bordon et al. (2005)	United States	10,367	4–17	P	Total score: 0.83 Subscores:0.46–0.77 Impact: 0.80		
Van Leeuwen et al. (2006)	Netherlands	523 + 1086	4–8	P, T	Subscores: 0.48–0.84	Subscores: 0.22–0.50	
Davé et al. (2008)	United Kingdom	248	4–6	M, F	Total score:0.61–0.62 subscores:0.36–0.74	Total score:0.27 (kappa) Subscores:0.025–0.36	
Du et al. (2008)	China	1965	3–17	P, T	Total score:0.59–0.60 Subscores:0.30–0.83	Total score: 0.46 (r) subscores:0-25–0.44	Total score: 0.55–0.72 Subscores: 0.40–0.79 (12 months)
Rothenberger et al.(2008)	Germany	2,406	7–16	P	Total score: 0.82 Subscores: 0.58–0.79		
Sanne et al. (2009)	Norway	6,430 (P) 8,999 (T)	7–9	P, T	Subscores: 0.55–0.82		

(cont.)

Table 2.4.1. (continued)

Study/ author	**Country**	**N**	**Child's age**	**Informant**[1]	**Internal consistency** (α)[2]	**Inter-rater reliability** (correlation)[3]	**Test-retest** (correlation)[3,4]
Stone et.al (2010) review		131,223	4–12	P, T	Weighted means: Total score:0.80–0.82 Subscores:0.53–0.85 Impact:0.81–0.83	Weighted means: Total score: 0.44 Subscores:0.26–0.47	Weighted means: Total score:0.76–0.84 Subscores:0.65–0.85 Impact:0.57–0.68
Mellor & Wong (2011)	China	700	mean 8.7	M, F	Total score:0.70–0.71 Subscores: 0.25–0.69	Total score: 0.53–0.61 Subscores:0.38–0.61	
Niclasen et al. (2012)	Denmark	71,840 (cohort study)	5, 7, 10–12	P, T	Total score:0.75–0.88 Subscores: 0.44–0.86	Total score: 0.45–0.53 Subscores:0.29–0.50 Impact: 0.41–0.50	

[1] P = parent, T = teacher, (S) = self (the results are not represented here), M = mother, F = father.
[2] α = Cronbach's alpha.
[3] Parameters in assessing correlations: r = Pearson correlation coefficient, EQS = Structural equation modelling.
[4] ICC = Intraclass correlation.

2.4.2 Review of the validity aspects of the SDQ

2.4.2.1 Content validity of the SDQ

The SDQ was developed for a brief behavioural screening questionnaire measuring children's prosocial behaviour (strengths) and symptoms in four distinct dimensions (difficulties): emotional symptoms, conduct problems, hyperactivity and peer problems (R. Goodman, 1997; R. Goodman and Scott, 1999; R. Goodman, 2001). In addition, the extended version of the questionnaire was intended to measure the chronicity, overall distress, social impairment and burden on others caused by the difficulties (R. Goodman, 1999). The developer of the SDQ has stated that the selection of items measuring the difficulties and of questions for the impact supplement was based on the concepts and classifications of childhood psychopathology (American Psychiatric Association, 1994; World Health Organisation, 1994) and factor analysis (Elander and Rutter, 1996; R. Goodman, 1997; R. Goodman, 1999). The content of the SDQ has been compared with other instruments considered as assessing the same construct, Rutter and ASEBA questionnaires (Achenbach and Rescorla, 2001; Elander and Rutter, 1996). These results are represented below, in Chapter 2.4.2.3.

2.4.2.2 Criterion validity of the SDQ

There are relatively few studies assessing the criterion validity against a diagnostic gold standard when we consider the large number of studies on the psychometric properties of the SDQ. In the review of the psychometric properties of the parent- and teacher-rated SDQ for 4–12-year-olds, only nine of the 48 studies included assessed the criterion validity of the method (Stone et al., 2010). This dissertation outlines the first Finnish criterion validity study against a diagnostic method.

In the British epidemiologic study (R. Goodman, 2001), SDQ total scores above the 90th percentile (high-risk group) were strongly associated with independently diagnosed psychiatric disorders, according to the diagnostic assessment by the Development and Well-Being Assessment (DAWBA; R. Goodman, Ford, Richards, Gatward and Meltzer, 2000). The mean odds ratio for a

psychiatric disorder in this high-risk group compared with the low-risk group was 15.7 for parent-rated SDQ scales, 15.2 for teacher-rated scales and 6.2 for youth self-rated scales (R. Goodman, 2001). It has been replicated that higher SDQ scores are associated with a greater probability of a child psychiatric disorder (Hawes and Dadds, 2004).

On the basis of the parent-, teacher- and self-rated SDQ symptom scores and the impact scores, a computerised algorithm was developed to predict child psychiatric diagnoses (R. Goodman, Renfrew and Mullick, 2000). When the SDQ is completed online, the predictive algorithm opens up as a part of the computer-assisted analysis (Youthinmind, www.sdqscore.org). The algorithm was assessed against independent diagnoses assigned by clinicians (R. Goodman et al., 2000). The predictive algorithm generates three classes of probability (unlikely, possible or probable) for four categories of diagnoses (conduct disorders, emotional disorders, hyperactivity disorders or any disorders). Among patients in Britain (n = 101) and in Bangladesh (n = 89), the "probable" prediction identified 81–91% of the children with diagnoses (R. Goodman et al., 2000). Correlations between the SDQ predictive algorithms and diagnoses assigned by clinicians were lower in an Australian sample than in the above-mentioned study (Mathai, Anderson and Bourne, 2004).

Sensitivity of the SDQ in identifying children suffering from psychiatric disorders has been found to be higher according to multi-informant ratings (parent, teacher and adolescent) than according to single-informant ratings in a large community sample (R. Goodman, Ford, Simmons, Gatward and Meltzer, 2000). The sensitivity of the multi-informant SDQ predictive algorithms was 63.3% in identifying children with a psychiatric disorder by the DAWBA; the specificity was 94.6%. The multi-informant predictive algorithm of the SDQ had a more than 70% sensitivity in identifying conduct, hyperactivity, depressive disorders and some anxiety disorders but only under half of the specific phobias, separation anxiety disorders and eating disorders were identified (R. Goodman et al., 2000). Again, the sensitivity of the parent-rated SDQ total scores in detecting psychiatric disorders has been found to be accurate (Hysing, Elgen, Gillberg, Lie and Lundervold, 2007).

2.4.2.3 Construct validity of the SDQ

2.4.2.3.1 Structural validity of the SDQ

The structural validity of the SDQ has been re-examined by item-level confirmatory factor analysis, and the results have mostly supported the five-factor structure (R. Goodman, 2001; Stone et al., 2010); see also Table 2.4.2. The original five-factor model was confirmed to fit in the British epidemiological study (R. Goodman, 2001). In the meta-analysis of the studies assessing the structural validity of the SDQ, support was found for the five-factor structure in ten of 13 studies; two studies found it a poor model with respect to fit; and the results in one study supported a three-factor model (Stone et al., 2010). In Nordic studies, the five-factor structure for parent and teacher versions of the SDQ has shown acceptable goodness of fit (Niclasen et al., 2012; Niclasen, Skovgaard, Andersen, Somhovd and Obel, 2013; Sanne et al., 2009; Smedje, Broman, Hetta and von Knorring, 1999); see also Table 2.4.2. In the Finnish sample of adolescents (13–17 years old), the five-factor structure was confirmed for the self-report (Koskelainen et al., 2001).

2.4.2.3.2 Hypothesis testing of the SDQ

The correlations between the SDQ scores and the scores of other questionnaires assessing theoretically similar constructs have been found to be considerable and strong. The scores of the SDQ and Rutter questionnaires rated by parents and teachers were highly correlated, and both methods discriminated well between psychiatric and non-psychiatric samples (R. Goodman, 1997).

The results of the SDQ and ASEBA questionnaires have been found to be convergent in several studies (R. Goodman and Scott, 1999; Koskelainen, Sourander and Kaljonen, 2000; Stone et al., 2010). When comparing the scores between the parent-rated SDQ and the parent-rated ASEBA questionnaire (Child Behaviour Checklist, CBCL), the correlation between the total scores was $r = 0.87$ ($p < 0.001$) and the correlations between the symptom subscores were $r = 0.71–0.84$ ($p < 0.001$) (R. Goodman and Scott, 1999). In addition, these methods discriminated well between the high-risk and low-risk children: the area under curve (AUC) was 0.95 for both methods (R. Goodman and Scott, 1999). Stone et al. (2010) reviewed the concurrent validity of the SDQ against the ASEBA parent- and teacher-rated questionnaires in nine studies. In the review, the weighted correlations between the total scores of the SDQ and the ASEBA questionnaires were 0.76 for both parent and teacher ratings. At

the subscale level, the highest weighted correlations were found for the conduct problems and hyperactivity (0.69–0.79) (Stone et al., 2010). Among Finnish school-aged children, the concurrent validity between the total scores of the SDQ and the CBCL was 0.75 and for the youth reports 0.71 (Koskelainen et al., 2000).

The SDQ was included in the comparison study on the convergent validity of three short questionnaires assessing psychosocial dysfunction among 7–12-year-old Dutch children (Vogels, Crone, Hoekstra and Reijneveld, 2009). In this study, the CBCL was used as a criterion, and the psychometric performance of all the three questionnaires (the SDQ, the Pediatric Symptom Checklist [PSC] and a Dutch-origin questionnaire [Psychosocial Problems in Primary Education, PSYBOBA]) supported their validity in detecting children's psychosocial dysfunction.

The total scores of the SDQ have adequately discriminated between the subgroups of high- and low-risk children (R. Goodman, 1997; R. Goodman, Meltzer and Bailey, 1998; R. Goodman and Scott, 1999; R. Goodman, 1999; Koskelainen et al., 2000; Stone et al., 2010). In the meta-analysis of eight studies assessing the discriminative validity of the SDQ, the weighted AUC value for the parent-rated total scores (0.87) was slightly higher than for the teacher ratings (0.83) (Stone et al., 2010). The lowest AUC values were found for the teacher-rated peer problems (0.57) and prosocial behaviour (0.65). Among Finnish 7–12-year-old children, parent- and teacher-rated high total scores (at or above the 90 percentile) were strongly associated with the parent-reported variables of problematic behaviour of the child and help-seeking for the child (Koskelainen et al., 2000). It is also notable that the impact scores of the SDQ have discriminated better than the total difficulties scores between the community and clinical samples (R. Goodman, 1999).

2.4.2.3.3 Cross-cultural validity of the SDQ

The cross-cultural validity of the SDQ has been widely assessed (Achenbach et al., 2008; Bourdon, Goodman, Rae, Simpson and Koretz, 2005; Du et al., 2008; R. Goodman et al., 2005; Hawes and Dadds, 2004; Marzocchi et al., 2004; Obel et al., 2004; Rothenberger et al., 2008; Woerner et al., 2004); see also Table 2.4.2. The questionnaire has been translated into more than 70 languages, but the national norms, assessed against a diagnostic assessment, have seldom been represented. The cross-cultural comparisons of the SDQ data and diagnostic assessments by the DAWBA showed considerable differences in the prevalence of children´s mental health disorders and in parents' reporting styles in seven countries (A. Goodman et al., 2011). It was concluded that population-specific norms on the SDQ are needed

before it is possible to estimate the prevalence of child mental disorders according to this questionnaire. In addition, it was suggested that prevalence levels based on short questionnaire assessments should be cautiously interpreted between different countries and cultures (A. Goodman et al., 2011).

The importance of easy accessibility of the computerised multicultural norms of the assessment method for practitioners and researchers has been emphasised (Achenbach et al., 2008). For the SDQ, computerised norms are available for six countries at the websites of the method for the moment (Youthinmind, b). Normative and computerised SDQ data are represented for British children (4 – 15-year-olds), American children (USA; 4–17-year-olds), Australian children (7–17-year-olds), Danish children (5–12-year-olds), Italian children (3–17-year-olds) and, most recently, Japanese children (4–15-year-olds). In addition, means and standard deviations and frequency distributions of the SDQ scores in many other countries have been published.

In the Nordic countries, the distributions of the SDQ scores have been quite similar, and the Nordic score distributions have been lower than their British counterparts (Heiervang, Goodman and Goodman, 2008; Koskelainen et al., 2000; Malmberg, Rydell and Smedje, 2003; Niclasen et al., 2012; Obel et al., 2004; Smedje et al., 1999). In Finland, the means and standard deviations of the SDQ total scores rated by parents, teachers and adolescents have been represented (Koskelainen et al., 2000). In this Finnish study, boys were scored higher than girls and adolescents higher than younger children. In addition, the means and standard deviations of all the SDQ scales have been assessed for the Finnish adolescents (Koskelainen et al., 2001). The present study focused on younger children than in the earlier studies, which focused on older school-aged children and adolescents in Finland. In addition, the present study was the first criterion validity study on the Finnish version of the SDQ.

Most recently, there has been a growing interest in assessing the psychometric properties of the SDQ among young children (Ezpeleta, Granero, de la Osa, Penelo and Domenech, 2013; Sim et al., 2013; Sveen, Berg-Nielsen, Lydersen and Wichstrom, 2013; Theunissen, Vogels, de Wolff and Reijneveld, 2013); see also Table 2.4.2. In Norwegian four-year-old pre-schoolers, the screening efficiency of the SDQ for emotional and behavioural disorders was found to be similar to those of older children in a criterion validity study (Sveen et al., 2013). Following the encouraging evidence of adequate reliability and validity properties in several studies, the preschool version of the SDQ for three- to four-year-old children was re-labelled to the early-years SDQ for two- to four-year-olds in June 2014

(www.sdqinfo.org). For the moment, normative and computerised SDQ data on children under five years old are represented for the following samples: Scottish 2–4-year-old children, Spanish 2–4-year-olds, Italian 3–5-year-olds and Japanese 4–5-year-old children (Youthinmind, www.sdqinfo.org).

Table 2.4.2. Summary of the SDQ validity studies included in the review.

Study/Author	Country	N	Child's age	Informant[1]	Main findings on criterion and construct validity
Goodman (1997)	United Kingdom	403	4–16	P, T	The correlation between the SDQ and Rutter total scores was 0.88 for parent and 0.92 for teacher reports. Both methods discriminated well between psychiatric and non-psychiatric samples: AUC values 0.84–0.87.
Goodman (1999)	United Kingdom	467+ 232	5–15	P, T, S	The score distributions on the extended version of the SDQ were markedly and significantly higher in the psychiatric clinic sample than in the community sample. Impact scores discriminated better than symptom scores between these groups.
Goodman & Scott (1999)	United Kingdom	132	4–7	P	The correlations between the SDQ and the CBCL[2] were for total scores r = 0.87 and for subscores r = 0.71–0.84.
Smedje et al. (1999)	Sweden	900	6–10	P	The five-factor structure was confirmed.
Goodman et al. (2000)	Britain Bangladesh	101+ 89	11–16	P, T, S	The "probable" SDQ predictive algorithm identified 81-91% of the children with diagnoses assigned by clinicians.
Goodman et al. (2000)	United Kingdom	7984	5–15	P, T, S	The sensitivity of the multi-informant SDQ predictive algorithms was 63.3% in identifying children with a psychiatric disorder according to the DAWBA[3]; the specificity was 94.6%.
Koskelainen et al. (2000)	Finland	735	7–15	P, T, (S)	Correlations between the SDQ parent reports and the CBCL[2] were 0.75 for total scores and 0.41–0.70 for subscores. Scoring at or above the 90th percentile on the SDQ by teachers or parents was strongly associated (OR 3.2 [1.5–7.1] – 10.0 [5.3–18.6]) with parent-reported variables of problematic behaviour and help-seeking for the child. (cont.)

Table 2.4.2. (continued)

Study/Author	Country	N	Child's age	Informant[1]	Main findings on criterion and construct validity
Goodman (2001)	United Kingdom	10,438	5–15	P,T,(S)	SDQ total scores above the 90th percentile were strongly associated with independently diagnosed psychiatric disorders by the DAWBA[3]. The predicted five-factor structure was confirmed.
Hawes & Dadds (2003)	Australia	1359	4–9	P,(T)	SDQ scores above the 90th percentile were strongly associated with a child psychiatric diagnosis assigned by DISCAP[4]: the odds ratio for total scores was 11.7.
Mathai et al. (2004)	Australia	130	4–15	P, T, S	The level of agreement between the SDQ predictive algorithm and clinical team diagnoses was 0.39–0.56. The sensitivities were 93% for conduct, 44% for hyperactivity and 36% for emotional disorder.
Hysing et al. (2007)	Norway	7007	7–9	P, T	The sensitivity of the parent-rated SDQ total scores for any disorder was 77.3% and specificity 88.5%, calculated against DAWBA[3] assessment.
Sanne et al. (2009)	Norway	6,430 (P) 8,999 (T)	7–9	P, T	The five-factor structure (slightly modified) was confirmed.
Vogels et al. (2009)	Netherlands	2,066	7–12	P	The SDQ, PSC[5] and PSYBOBA[5] questionnaires were compared and evaluated against the CBCL[2]. The internal consistencies were α = 0.80–0.89, sensitivities between 0.78–0.86% and AUCs 0.93–0.96. The added values of the questionnaires to offer new information were between 29.3 and 68.5 (odds ratios).

(cont.)

Table 2.4.2. (continued)

Study/Author	Country	N	Child's age	Informant[1]	Main findings on criterion and construct validity
Stone et.al (2010)	(meta-analysis)	131,223	4–12	P, T	Weighted correlations between the SDQ and the ASEBA (k^6 = 9) were 0.76 for total scores and 0.46–0.76 for subscores. The weighted AUC values (k^6 = 8) were 0.87 (0.64–0.91) for the parent-rated total scores and 0.83 (0.65–0.91) for teacher ratings.
Niclasen et al. (2012) and Niclasen et al. (2013)	Denmark	71,840 (cohort study)	5, 7, 10–12	P, T	The five-factor structure was confirmed.
Ezpeleta et al. (2013)	Spain	1341	3	P, T	Criterion validity against DICA-PPY[7] showed a good to very good ability to differentiate between the children with and without DSM-IV diagnoses (sensitivity for total scores over 70%). The five-factor structure was confirmed. Internal consistency for parent reports was 0.87 and for teacher reports 0.91. Convergent validities against the ASEBA questionnaire and C-GAS[8] were moderate.
Sim et al. (2013)	United Kingdom	543	30 months	P	Of the children, 8.8% scored in the abnormal range of the SDQ total scores. Children assigned at higher developmental risk had more often (13.5%) high SDQ total scores than children at no developmental risk (4.0%).
Sveen et al. (2013)	Norway	845	4	P, T	Of the parent-rated total score: the sensitivity was 54%, specificity 89% and the AUC 0.76 for any disorder. For emotional and behavioural problems, the screening efficiency was better: sensitivity 65%, specificity 89% and AUC 0.83. Teacher information did not improve the prediction accuracy.

(cont.)

Table 2.4.2. (continued)

Study/Author	Country	N	Child's age	Informant[1]	Main findings on criterion and construct validity
Theunissen et al. (2013)	Netherlands	839	3–4	P	Validity properties assessed against CBCL[2] as criterion were AUC 0.94, sensitivity 0.79, specificity 0.93 and odds ratio 36.5 for added information. Internal consistency for total score was α = 0.78 and for subscores 0.50–74.

[1] P = parent, T = teacher, (S) = self; the results are not represented here; M = mother, F = father.
[2] CBCL = Child Behaviour Checklist.
[3] DAWBA = the Development and Well-Being Assessment.
[4] DISCAP = Diagnostic Interview Schedule for Children, Adolescents and Parents.
[5] PSC = Pediatric Symptom Checklist, PSYBOBA = Dutch-origin Questionnaire (Psychosocial Problems in Primary Education).
[6] k = number of studies included in the meta-analyses.
[7] DICA-PPY = Diagnostic Interview for Children and Adolescents for Parents of Preschool and Young Children.
[8] C-GAS = Children's Global Assessment Scale.

2.4.3 Review of the feasibility aspects of the SDQ

The feasibility aspects of brevity and simplicity have been taken into consideration when developing the SDQ: the questionnaire fits at one or two pages; it is applicable to a wide age range of children, multi-informant versions of the questionnaire are similar in structure and contents (R. Goodman, 1997; Youthinmind, b). In addition, the SDQ is available free of charge online for non-commercial use and has been translated into more than 70 languages (Youthinmind, b). Regardless of the widespread use of the SDQ, little interest was shown in the feasibility aspects of the method in earlier studies.

In earlier qualitative studies, the SDQ attracted wide acceptance in focus groups evaluating respondents' views of routine outcome measures in child and adolescent mental health services (Moran, Kelesidi, Guglani, Davidson and Ford, 2012; Stasiak et al., 2013; Williamson et al., 2010). In order to improve the cultural appropriateness of the SDQ among Australian aboriginals, some practical key recommendations were given to clinicians about the use of the method: the questionnaire was to be administered orally and the purpose of the measurement and interpretation of the results should be explained to parents (Williamson et al., 2010).

The author has found only one earlier study evaluating the feasibility aspects of the SDQ in the context of assessing children's mental health in routine health check-ups in primary care. In that comparison study of three short questionnaires (the SDQ, the PSC and the PSYBOBA), the usability of the methods was rated by parents and professionals (Vogels et al., 2009). The SDQ was preferred by the professionals, but they argued that calculating the scores of the SDQ subscales was difficult. However, calculating the total scores of the SDQ was considered sufficient in the first distinction of the child's mental health problems. The parents favoured both the SDQ and PSYBOBA.

3 Aims of the study

The aims of the present study were:

1) To explore the psychometric properties and reliability of the Finnish version of the Strengths and Difficulties Questionnaire (SDQ-Fin) in 4–9-year-old children visiting for regular health check-ups at child health clinics and school health care clinics (I).

2) To define and select the adjusted Finnish cut-offs of the SDQ in young children against the diagnostic Development and Well-Being Assessment (DAWBA) and to explore the capacity of the SDQ-Fin to identify children suffering from mental health problems (II).

3) To assess the reliability and validity of one-question screens presented to children, their parents and public health nurses in identifying children at elevated risk for mental health problems (IV).

4) To evaluate the feasibility of using the SDQ-Fin as a screening questionnaire in assessing young children's mental health and to evaluate the feasibility of directly asking a young child to evaluate his/her emotional well-being in the context of regular health check-ups (III, IV).

The detailed research questions of the study are presented in Table 3 below.

Table 3. An overview of the detailed research questions.

		Detailed research questions	Focus of measurements	Data	Articles
1.	a)	Are there differences in the distributions of the SDQ-Fin parent and teacher scores and between genders and age groups of the children?	Exploring the distributions of the SDQ-Fin parent and teacher reports.	SDQ-Fin for the mother, father and preschool and school teacher; SDQ parent report for the second time within 12 weeks [1)]	I
	b)	What is the extent of agreement between informants?	Assessing the inter-rater reliability between the informant pairs (mother-father, average parent-teacher, teacher-teacher), the internal consistency and the test-retest of the SDQ-Fin.		
	c)	How consistent is the SDQ-Fin?			
	d)	How stable is the SDQ-Fin?			
2.	a)	What are the cut-offs (norms) on the SDQ in a Finnish community sample of 4–9-year-old children?	Searching the cut-offs by defining the 80th and 90th percentiles of the SDQ scores and by ROC analysis against a diagnostic assessment.	SDQ parent and teacher report, Development and Well-Being Assessment (DAWBA) for parents and teachers	II
	b)	How accurately does the SDQ-Fin with the suggested cut-offs identify the children suffering from psychiatric symptoms and disorders?	Assessing sensitivity and specificity for the suggested cut-offs, and assessing the diagnostic discriminative capacity of the SDQ.		
3.	a)	Are the one-question screens reliable and valid methods?	Assessing the inter-rater reliability, examining the discriminative validity of the one-question screens against a diagnostic assessment and assessing the associations between the one-question screens and child psychiatric disorders.	One-question screen for parent and nurse, child's self-evaluation enquiry on emotional well-being, SDQ-Fin for parent and teacher, DAWBA for parents and teachers	IV
	b)	Is it clinically relevant to directly ask a young child to evaluate his/her emotional well-being?			
4.	a)	What kind of experiences and opinions do parents, public health nurses and teachers in preschool education and at school have about the SDQ-Fin and its use?	Analysing the collected quantitative and qualitative feedback data.	SDQ feasibility questionnaire for parents, preschool and school teachers and public health nurses, feedback questionnaire for the child's self-evaluation enquiry for parents and public health nurses	III, IV
	b)	How appropriate and burdensome do the parents and public health nurses find the child's self-evaluation enquiry?			

[1)] In addition, data were collected on participation information, socio-demographic characteristics of the participants and the nurse's one-question screen for participants and non-participants.

4 Material and methods

4.1 Study design

4.1.1 The pilot study

A small-scale pilot study was conducted in the city of Tampere from November 2006 to February 2007. The pilot study was a part of a project entitled "Developing indicators for children´s psychosocial health 2006-2007" which was funded by the Ministry of Social Affairs and Health of Finland and the Pirkanmaa Hospital District.

The aims of the pilot study were 1) to test the suitability of the methods designed to be employed in the main study, 2) to test how the study procedure worked in everyday clinical practice and 3) to choose a suitable comparison method, the gold standard for the questionnaires to be assessed in the main study.

Eleven child health clinics and five school health care clinics were willing to participate in the pilot study. Of the eligible five-year-old children, 29 out of 43 participated in the study and, of the eligible seven-year-olds, eight out of 45.

As a result of the pilot study, some minor changes were made to the design of the main study and to the contents of study information letters and feedback questionnaires. The ASEBA questionnaires were excluded from the study procedure based on parents' and teachers' personal communications about the ASEBA being too burdensome to be used in the context of regular health check-ups. The Development and Well-Being Assessment (DAWBA) interview was chosen as the diagnostic assessment method. Furthermore, the parents gave positive feedback on conducting the DAWBA interview by phone.

4.1.2 Study procedure

The main study was conducted as a part of a project entitled "Developing children's mental health work, 2007-2009". The data was collected from March 2008 to March 2009. The project was funded by Finland's Ministry of Social Affairs and Health and the participating hospital districts.

The target population of the study was 4−9-year-old children visiting for regular health check-ups at child health clinics and school health care clinics.

Training sessions were held in the municipalities participating in the project in order to instruct in and implement both the project and the study procedures. Brief training sessions on the assessment methods used in the study were offered to medical professionals in the primary health care system, to preschool teachers and to teachers at schools. These training sessions consisted of an overview of the contents and use of the assessment methods.

The study consisted of three phases: 1) multi-informant questionnaire assessments in the context of regular health check-ups, 2) a diagnostic assessment for a stratified subgroup of participating children after the check-up visit and 3) collecting feedback from the informants on the questionnaires used in the study. An overview of the study design is presented in the Figure 4.1.2 below.

4.1.2.1 The first phase of the study: Questionnaire assessments (I, IV)

The study was introduced by public health nurses to parents who were booking an appointment for their 4−9-year-old child's regular health check-up during the appointed time of data collection. Prior to their visit to the clinic, the public health nurse sent study information and questionnaires to the homes of interested parents: the informed consent form, a socio-demographic questionnaire including a parent's one-question screen and the SDQs for both parents and the child's teacher. The participating parents filled in the questionnaires and asked the child's teacher to complete the SDQ. More precisely, two preschool teachers and one teacher at school were asked to complete the SDQ. The parents then returned all the completed forms to the public health nurse at the health check-up visit.

During this visit, the public health nurses asked the accompanying parent or parents whether the child and the family had decided to participate in the study or not. In addition, the public health nurses evaluated, by asking one question, whether the child had any emotional or behavioural difficulties. This one-question screen was completed for every child at the health check-up and completed

anonymously for the non-participants. The SDQs could be used for discussion with the parents at the check-up, but the sum scores or reports produced by the method were not made available to the public health nurses. Each child completed a self-evaluation enquiry about his/her emotional well-being with the help of the public health nurse.

The public health nurses collected the above-mentioned completed forms and questionnaires, and these papers were then submitted to the researchers to be processed.

4.1.2.2 The second phase of the study: The diagnostic assessment (II, IV)

After the check-up visit, the parents' and teachers' SDQ reports were used to stratify the participating children into two subgroups: 1) the children who had scored at or above the British 80th percentile cut-off, i.e. screen-positives, according to any informant, and 2) the children who had scored below the above-mentioned percentile cut-off, i.e. screen-negatives, according to every informant. Every parent of the screen-positive subgroup was invited to the diagnostic assessment of the Development and Well-Being Assessment (DAWBA) interview. For every two such screen-positive cases (at the beginning of the study for every such case), a parent of a screen-negative child, matched for child's age group and gender, was invited to the DAWBA interview. One parent per child was interviewed by phone within an interval of 2–17 weeks after the health check-up visit. With the parents' permission, the child's preschool or school teacher was also asked to complete the DAWBA in the form of a paper questionnaire.

The SDQ is included in the DAWBA assessment, and thus the parents who participated in the second phase of the study completed the SDQ for the second time by phone. The parent's initial and second SDQ reports were utilised for assessing the test-retest reliability of the SDQ (I).

The parent phone interviews were conducted by five interviewers who were trained in the use of the DAWBA method and experienced in mental health work. The author held meetings with the interviewers at regular intervals in order to standardise the practices of interviewing and registering the data.

The author reviewed all the DAWBA assessments blind to the initial SDQ screening status of the children and assigned the diagnoses according to the ICD-10 (World Health Organization, 1994) and DSM-IV (American Psychiatric Association, 1994; World Health Organization, 1994).

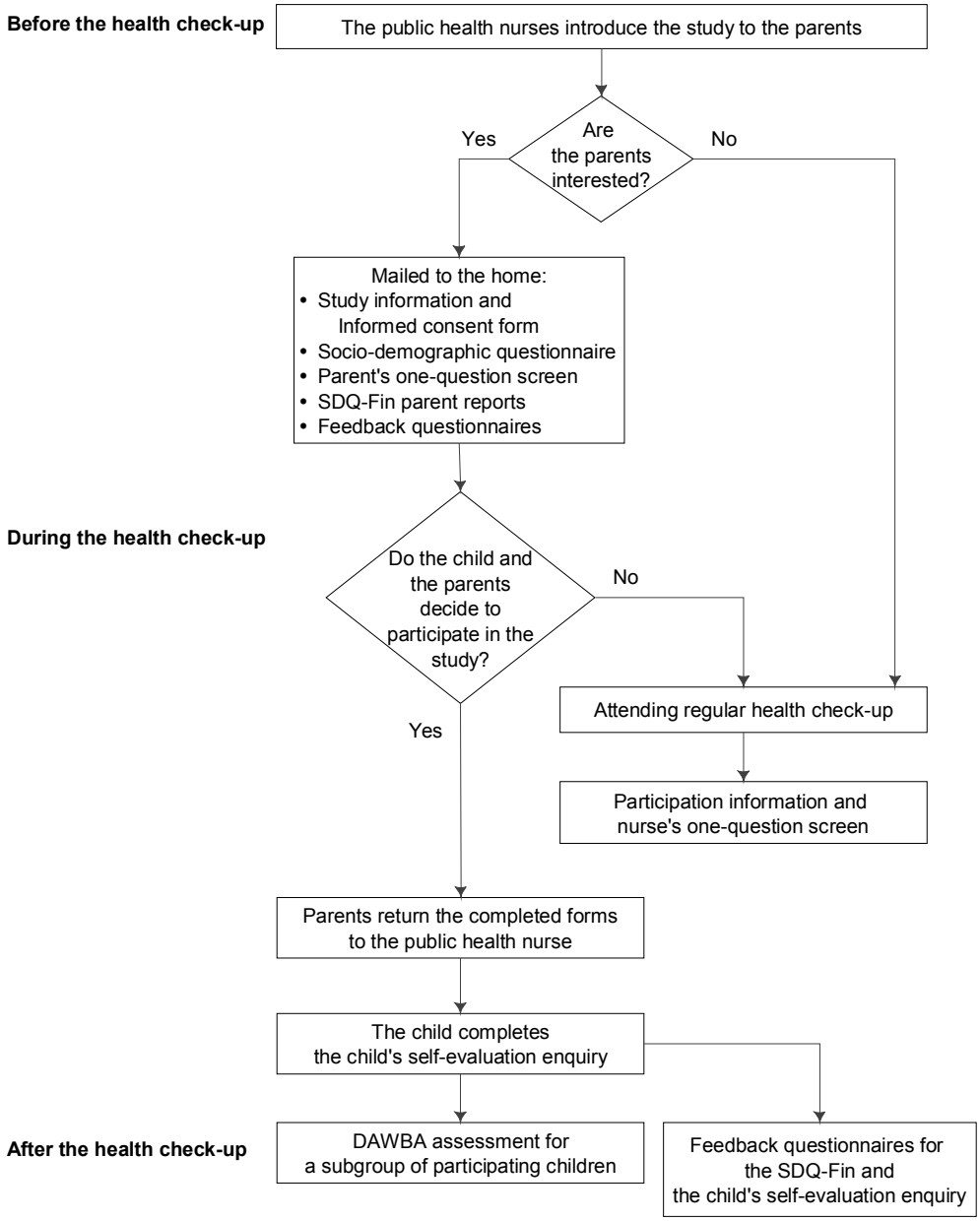


Figure 4.1.2. An overview of the study design (I-IV).

4.1.2.3 The third phase of the study: Collecting feedback on questionnaires (III, IV)

Feedback questionnaires on the feasibility of the SDQ and the child's self-evaluation enquiry were collected after the check-up visit. The SDQ feasibility questionnaire was completed anonymously by the participating parents and only once by each teacher and public health nurse involved in the process. The feedback questionnaire for the child's self-evaluation enquiry was completed anonymously by parents and once by each public health nurse involved in the study process.

The completed feedback questionnaires were returned to the researchers via the public health nurses or by the parents themselves in return envelopes.

4.1.3 Sample

The study was conducted in two of the 21 hospital districts in Finland, covering a total population of 634,526 in 2007 (Statistics Finland, 2014). In the Pirkanmaa Hospital District (central Finland), 18 out of 25 municipalities and, in the South Karelian Hospital District (eastern Finland), seven out of ten municipalities participated in the study as their willingness and resources allowed. There were urban (n = 11), semi-urban (n = 7) and rural (n = 17) municipalities among those enrolled and non-enrolled (Statistics Finland, 2015). A total of 154 child health clinics and school health care clinics took part in the study.

Children four to nine years old who were attending regular health check-ups during the study period (1 March 2008 to 31 March 2009) were included in the study. The sample consisted of 4−6-year-old preschool-aged children in child health clinics and 7-9-year-old children in school health care clinics; some of the latter had just passed their tenth birthday. Families not speaking Finnish were excluded from the study (n = 40).

4.1.3.1 Sample in the first phase of the study (I, IV)

A subject flow chart of the participation rates in the first and second phases of the study (I, II) is presented in Figure 4.1.3.

A total of 4,178 eligible children (49.5% girls) and their parents were invited to participate in the study. Three-fifths of the eligible children were preschool-aged (n = 2,596) and the rest school-aged (n = 1,582); see Figure 4.1.3.

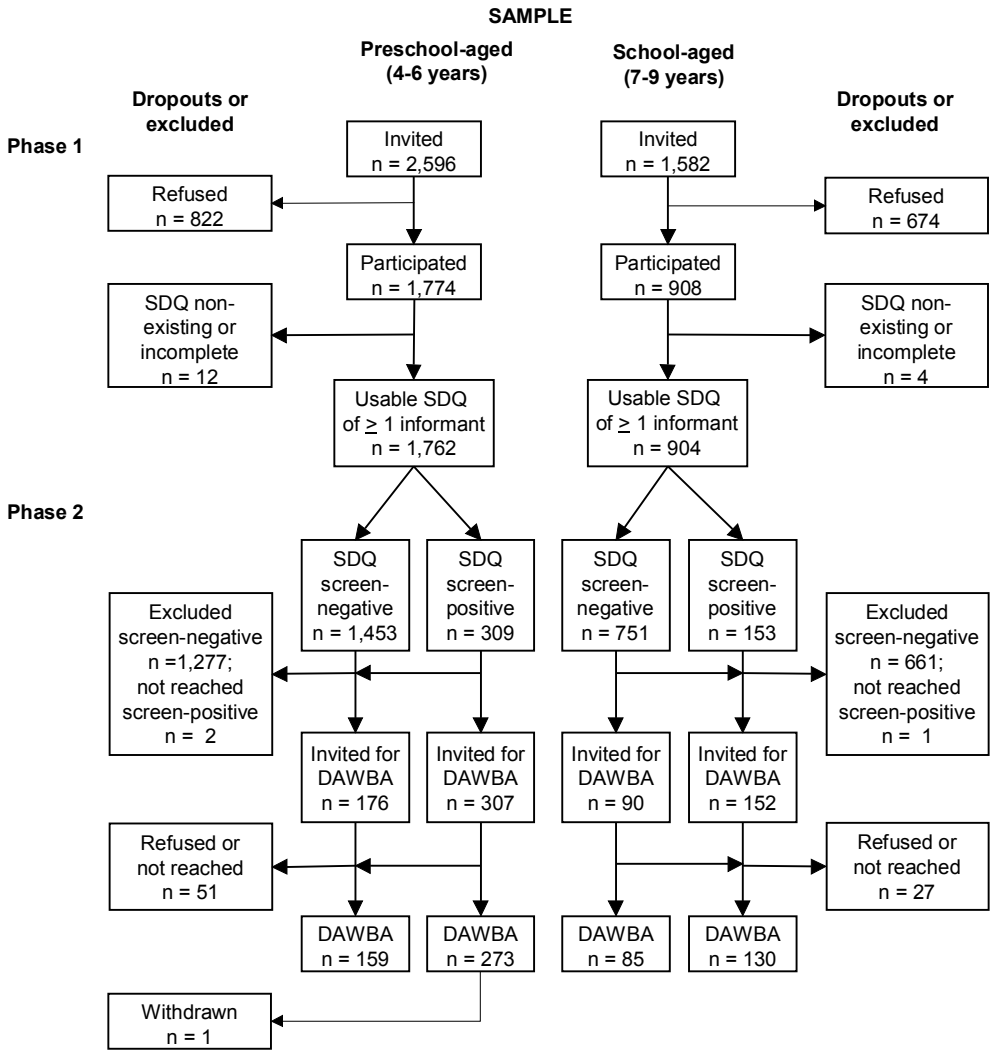


Figure 4.1.3. Subject flow chart of the reliability and validity study of the Finnish version of the Strengths and Difficulties Questionnaire in 4-9-year old children (I, II).

The participation rate in the total sample was 64.2% (n = 2,682/4,178); see Figure 4.1.3. The participation rate among preschool-aged children was 68.3% and among school-aged children 57.4%. The genders of the participating children were

evenly distributed. The participants in the Pirkanmaa Hospital District accounted for 67.0% of the total sample. A total of 71.9% of the preschoolers and 57.5% of the school-aged children were collected from Pirkanmaa. The socio-demographic data of the participants are presented in more details in the Study I (Table 1). No statistically significant differences between the hospital districts were found regarding the gender of the child, the family structure, the number of siblings or family income. The parents of the preschool-aged children in Pirkanmaa had a higher level of education ($p < 0.001$) than those in South Karelia. This difference was considered to enrich the data; consequently, the two hospital districts were pooled in the analysis.

The socio-demographic variable means of the family structure, the number of children in families, day-care arrangements and family income among participants did not substantially differ from the Finnish demographic statistics (2005–2012) on families with children (Hiltunen, 2007; Olkkonen, 2014 a; Olkkonen, 2014 b; Helminen and Pietiläinen, 2014). In addition, the level of education among the population aged 15 or over has been found to be higher in the Pirkanmaa region than in South-Karelia (Suomen virallinen tilasto, 2013), corresponding the findings of educational differences among the participating parents in the present study.

Almost all the participating children (2,666/2,682) had at least one informant's SDQ that was usable in the analysis (Figure 4.1.3). There were 2,582 usable mother-completed SDQs and 1,935 usable father-completed SDQs, and both parents had completed the SDQs in 70.2% of the cases (n = 1,882). Of the participating preschool-aged children, 76.3% (1,354/1774) had a teacher-rated SDQ available. In addition, another teacher had completed the SDQ in 53.3% (945/1774) of the preschool-aged participants. In the group of school-aged participants, 97.5% (n = 885) had a teacher's SDQ report available.

A nurse's one-question screen was available for 99.3% of the participants and for all 1,496 non-participants. The parents' one-question screen was completed by 98.9% of the participants. Of the 2,682 participating children, 97.8% completed the self-evaluation enquiry. Both the parents' one-question screen and the child's self-evaluation enquiry were available for 96.8% (n = 2,595) of the participating children.

4.1.3.2 Sample in the second phase of the study (II, IV)

In the second phase of the study, the SDQ was used to stratify the children into screen-positive and screen-negative subgroups in both age groups (Figure 4.1.3). In the

total sample, 17.3% of the children had scored above the British 80th percentile cut-off in the parent or teacher report or both. Of the screen-positives, 66% were boys (305/462). The respective proportion of the screen-positives was 17.5% in preschool-aged and 16.9% in school-aged children. DAWBA information was obtained from 93.3% (431/462) of the screen-positives in the second-phase sample. One family in the screen-positive subgroup decided to withdraw from the second phase of the study after the DAWBA interview. In the subgroup of those screen-negatives who were invited to the DAWBA, the participation rate was 80.8% (215/266). Of the participants in the DAWBA assessment, 67% were preschool-aged children and 66% were boys. Of the participating parents in the DAWBA interview, 91% were mothers, 7% fathers and 2% some other person. A teacher's report was obtained from 75% (n = 486/646) of the participants in the DAWBA.

4.1.3.3 Sample in the third phase of the study (III, IV)

For the SDQ feasibility questionnaire, the response rate among parents was 58% (1,546/2,682). Of the parental respondents, 65% (n = 1,009) had preschool-aged children and the remainder school-aged children. A total of 225 public health nurses were involved in the study process, and 70% (n = 156) of them completed the SDQ feasibility questionnaire (only once each, not once per child). Of the nurse respondents, 107 worked at child health clinics and 49 at school health care clinics. Preschool teachers returned 1,286 SDQ feasibility questionnaires, but the exact number of participating preschool teachers could not be determined and thus participation rates cannot be calculated. There were 123 school teachers involved in the study process: four of them did not return the questionnaire, and thus the response rate was 97%.

The feedback questionnaire on the feasibility of the child's self-evaluation enquiry was completed by 55% of the participating parents (preschoolers n = 964, school-aged n = 523) and 99% of the public health nurses (child health care n = 105, school health care n = 49).

4.1.4 Attrition

The participation rates among girls and boys in both age groups in the first and second phases of the study were regularly monitored during the data collection

process. The response rates in the first phase of the study increased when reminders were sent to parents and public health nurses and when the public health nurses received support in their administrative tasks during the study process. In addition, participating professionals were encouraged to contact the project coordinator or researcher if they had any further questions. In the diagnostic phase of the study, the interviewers tried to reach the parents on the phone several times, and if they did not succeed, an invitation letter to the DAWBA interview was sent to the home. In addition, parent wishes were taken into account when arranging times for the interviews.

The nurse's one-question screen (nurse's concern enquiry) was completed for all invited children (see Figure 4.1.3), including both participating and non-participating children, in order to assess the possible difference between these groups; see Table 4.1.4. According to the public health nurses, the non-participating children had difficulties more often than those participating in both gender and age groups. The frequency of boys was higher than that of girls in the group of non-participants. The attrition analysis was not conducted on the sixteen participating children who were dropouts or excluded because of incomplete forms (see Figure 4.1.3).

Table 4.1.4. Comparison of the frequencies of children's difficulties identified by the public health nurses between participants and non-participants.

	Total Sample			Boys[1]		Girls[1]		Preschool- aged children		School-aged children	
	Part.[2] %	Non-part.[3] %	All %	Part.[2] %	Non-part.[3] %	Part.[2] %	Non-part.[3] %	Part.[2] %	Non-part.[3] %	Part.[2] %	Non-part.[3] %
	(n = 2,682)	(n = 1,496)	(n = 4,178)	(n = 1,315)	(n = 759)	(n = 1,367)	(n = 732)	(n = 1,774)	(n = 822)	(n = 908)	(n = 674)
Gender of the child											
Boy	49.1	50.9	49.6	49.0	50.7			49.5	54.9	47.0	49.3
Nurse's one-question screen											
no difficulties	73.3	66.4	70.8	65.5	62.2	80.8	70.9	73.0	65.5	73.8	67.7
yes - minor difficulties	17.2	20.4	18.3	21.7	21.6	12.8	19.1	17.5	20.6	16.5	20.2
yes - definite/severe difficulties	6.6	12.0	8.5	9.4	15.3	4.0	8.6	6.6	12.8	6.6	11.1
cannot say/do not know	3.0	1.1	2.3	3.5	0.9	2.5	1.4	2.9	1.2	3.1	1.0

[1]Information on child gender was missing in five cases.

[2] Part. = participants, [3] Non-part. = non-participants. The differences between participating and non-participating subsamples were all statistically significant ($p < 0.05$)

4.2 Measures

4.2.1 Questionnaires

4.2.1.1 The Strengths and Difficulties Questionnaire (SDQ)

The SDQ is a screening questionnaire for 3–16-year-olds to be completed by parents, teachers and 11–16-year-old children themselves (R. Goodman, 1997; R. Goodman, 1999; R. Goodman, 2001). The method includes three components: 1) items on the psychological attributes of the child's behavioural and emotional problems and social skills, concerning both positive and negative behavioural traits, 2) impact supplement and 3) follow-up questions. The extended version of the SDQ, including the items on psychological attributes (on the front of the sheet) and the impact supplement (on the back), was used in this study. The parents' and teachers' SDQ reports were utilised in the study as well (I-IV).

The Finnish version of the SDQ and instructions for scoring by hand have been translated from English into Finnish, back-translated, and approved by the copyright owner of the method. The current version of the SDQ-Fin (Appendix 5) was the same as used in the earlier study of Koskelainen et al. (2000). The scoring instructions in Finnish are found on the websites for the method (Youthinmind, b); see Appendix 6.

The 25 items on psychological attributes are divided into five subscales, each including five items. The subscales collate emotional symptoms, conduct problems, hyperactivity/inattention difficulties, peer relationship problems and prosocial behaviour. Items are originally scored as 1 for "somewhat true" and, depending on the item, as 0 or 2 for "not true" or for "certainly true", and for analysis recoded as 0 to 2 with increasing severity. The scores from all the scales except the prosocial scale are summed to generate a total difficulties score ranging from 0 to 40. On the prosocial scale, higher scores stand for desirable, positive behavioural traits, whereas on the other four scales and on the total score, higher scores reflect increasingly negative behavioural traits and symptoms (Youthinmind, b). The scoring instructions for the Finnish versions completed by parents or teachers are attached in Appendix 6.

In the SDQ analysis, the scores can be used as continuous variables or as categorised bandings. The 80th and 90th percentiles of the SDQ scores have been presented as provisional cut-off points for the groups of "borderline" and "abnormal" (R. Goodman, 1997). This original three-band categorisation of the SDQ scores into the groups of normal, borderline and abnormal was used in the study.

In the impact supplement, the first question on the parent and teacher version asks, "Overall, do you think that your child/this child has difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people?" The response alternatives are no, yes/minor difficulties, yes/definite difficulties, yes/severe difficulties. If difficulties are reported, the rest of the impact supplement questions enquire about the duration or chronicity of the difficulties, overall distress, social impairment and burden to others. The items on overall distress and social impairment are scored from 0 to 2 and can be summed to generate an impact score which is a continuous variable ranging from 0 to 10 for the parent-rated version and from 0 to 6 for the teacher-rated version. Teachers, unlike parents, are not asked to report about the home life and leisure activities of the child. The items on chronicity and burden to others are not included in the impact score. The total impact score can be classified as normal for a score of 0, borderline for a score of 1 and abnormal for a score of 2 or more. This classification was also used in the present reliability study of the SDQ-Fin (I).

4.2.1.2 The nurse's one-question screen

The nurse's one-question screen was consistent with the first question in the parent's and teacher's SDQ impact supplement. Public health nurses assessed, based on their clinical evaluation, whether, overall, the child had difficulties in one or more of the following areas: emotions, behaviour, concentration, or being able to get on with other people. The response alternatives were no, yes/minor difficulties, yes/definite difficulties, yes/severe difficulties, can't say/do not know. The last answering option was added to the original alternatives of the above-mentioned question on the SDQ.

The enquiry was called a nurse's concern enquiry in Study I and a nurse's one-question screen in Study IV.

4.2.1.3 The parent's one-question screen

Parents were asked to assess by answering one question about whether their child had any emotional problems or difficulties in behaviour, concentration or social skills. The enquiry was answered on a four-step scale (no difficulties/not many difficulties/ quite many difficulties/very many difficulties). The enquiry was a slightly modified version of the first question on the parent's SDQ impact supplement.

4.2.1.4 Child's self-evaluation enquiry of emotional well-being

The child's self-evaluation enquiry of his/her emotional well-being was developed for this study (IV), and it consisted of two questions; see Appendix 7. The first question assessed the child's self-evaluation of his/her mood and emotions. The second question explored the child's expectations regarding the future. The written response alternatives (five-step scale) had visual analogues in the form of facial expressions. The public health nurse read the questions and the response alternatives aloud for the child even if he/she could read. The child chose and marked with an "X" the answer best describing his/her feelings and expectations.

4.2.1.5 The SDQ feasibility questionnaire

Experiences and opinions on the SDQ and its use were collated from parents, public health nurses and teachers in preschool education and at school using the feasibility questionnaire developed for this study (III); see Appendix 8 and 9. The impact of using the SDQ in cooperation between parents and public health nurses or pre-school teachers was also assessed. The questionnaire was developed on the basis of both earlier research on feasibility aspects of measures (Myers and Winters, 2002; Slade et al., 1999; Slade et al., 2001) and clinical considerations.

All respondents were asked to report how long it took to complete the SDQ (in minutes) and how burdensome they found it (not at all/not much/rather a lot/very much). They were also asked to estimate how age-appropriate (to the developmental stage) this method was in assessing the psychosocial well-being of the child (very good/fairly good/not good, not poor/rather poor/very poor). Schoolteachers answered only these first three questions.

In the next five items, answered on a five-step scale (totally agree/somewhat agree/somewhat disagree/totally disagree/can't say), parents, public health nurses and preschool and school teachers were asked to assess the impact the use of the SDQ had on cooperation between parents and professionals. These items collated how useful the SDQ was in gathering information and reaching a common understanding about the child's mental health and possible need for support. All these respondents were also asked how parents felt about the method and how burdensome a tool they found it as part of the medical check-up of the child and in the conversations between parents and preschool teachers.

In open-ended questions, public health nurses and preschool teachers were asked to list the methods they had previously used when evaluating children's psychosocial well-being, social skills and possible need for support. They were also asked to compare their observations on the SDQ and these earlier methods. In addition, all respondents were offered a chance to write feedback in their own words on their experiences with the method.

4.2.1.6 The feedback questionnaire for the child's self-evaluation enquiry

Parents and public health nurses evaluated the feasibility aspects of the child's self-evaluation enquiry in the feedback questionnaire developed for this study (IV); see Appendix 8. They were asked how appropriate this method was in assessing the psychosocial well-being of the child (very good/fairly good/not good, not poor/rather poor/very poor). The public health nurses were also asked to report how long (on average, with five minutes' precision) it took to complete the child's self-evaluation enquiry and how burdensome they found it (not at all/not very/ rather/very burdensome). An open-ended question about feedback on the informants' experiences with the method was not utilised in this study.

4.2.2 Measures used in assessing sample characteristics

Characteristics of the participating and non-participating study samples were assessed by participation information completed by the public health nurses and the nurse's one-question screen; see Appendix 3. Data on the age and gender of each participating and non-participating child were recorded by the public health nurses. In addition, the participating parents completed the socio-demographic form developed for this study

(I); see Appendix 4. The socio-demographic characteristics utilised in the study are presented in more detail in Study I.

4.2.3 Diagnostic assessment

The diagnostic assessment was conducted using the parent and teacher versions of the DAWBA method (R. Goodman et al., 2000). The Finnish version of the DAWBA consists of a semi-structured interview of the parents of children aged 5–17 years and of children over 11 years old, and a briefer questionnaire version for teachers (Youthinmind, a). The DAWBA method can be administered by an interviewer working with a computerised or paper version, or the respondent himself/herself can complete the interview online. In the present study, the interviewers used the computerised online version of the DAWBA.

The DAWBA begins with the SDQ. The interview consists of sections of structured questions about psychiatric symptoms and their impact. If definite symptoms are identified, parents are asked to describe the problems in more detail in open-ended responses. On the other hand, skip rules can reduce the length of the interview when screening questions do not indicate any of the identified symptoms covered by the section. The skip rules will not activate when the start-of-section screening questions indicate problems or the relevant corresponding SDQ scale is above the 80th percentile.

The structured questions in various sections follow closely the diagnostic criteria of the ICD-10 (World Health Organization, 1994) and DSM-IV (American Psychiatric Association, 1994), and most child psychiatric disorders are covered. Information from the structured questions answered by different informants is collated by a computer program that assigns each child to a level of probability bands that represent the prevalence of any diagnosis and several specific diagnosis categories in epidemiological samples (A. Goodman, Heiervang, Collishaw and Goodman, 2011). The categorisation of the predictive measure (the prevalence of any disorder) offered to the clinical rater and used in the study is < 1% (very low), < 5% (low), 20% (moderate) and 75% (high). This categorisation offers the clinical rater a starting point for reviewing the data comprehensively. To decide on definitive diagnoses, the clinical rater reviews all relevant information: the structured, closed and the open accounts of all available informants and the prevalence level produced by computer algorithms.

The DAWBA method has been translated into 28 languages in paper versions and into 19 languages in online versions (Youthinmind, a). The method has shown good inter-rater reliability (Aebi et al., 2012; Ford, Goodman and Meltzer, 2003; Frigerio et al., 2009; Heiervang et al., 2007). The method discriminates well between community and clinical samples (R. Goodman et al., 2000). Agreement between clinical diagnoses and diagnoses assigned by the DAWBA rater has been fair to moderate in concordance with earlier findings between diagnoses assigned by standardised diagnostic interviews and clinical diagnoses (Aebi et al., 2012; Frigerio et al., 2009; R. Goodman et al., 2000; Rettew, Lynch, Achenbach, Dumenci and Ivanova, 2009). In a comparative study of three diagnostic interviews, lower prevalence rates for diagnoses were generated with the DAWBA than with the other interviews, which were the Diagnostic Interview Schedule for Children (DISC) and the Child and Adolescent Psychiatric Assessment (CAPA) (Angold et al., 2012). In this comparative study, the DAWBA was considered a good choice for clinical trials.

In the present study, the researcher assigned the definite diagnoses according both to the ICD-10 and DSM-IV classifications, but only the ICD-10 diagnoses were used in Studies II and IV. The assigned specific diagnoses were placed in five categories: emotional, conduct, hyperactivity and other diagnosis (Tic/Tourette's, pervasive developmental disorders, not otherwise specified mental disorders) and situational factors (Z61 - Problems related to negative life events in childhood, Z62 - Other problems related to upbringing and Z63 - Other problems related to primary support group, including family circumstances). Diagnoses of developmental disorders (enuresis, encopresis, stuttering, developmental disorders of speech and language and of scholastic skills) were excluded from analysis.

The competence of the clinical rater was ensured by several arrangements because it was not possible to arrange inter-rater reliability testing. The clinical rater was trained in an intensive two-day training course and practised the cases in the training manual at the DAWBA website. The rater also consulted the author of the method on practical and methodological issues. A consultation group of four experienced child psychiatrists was invited in order to obtain consensus diagnoses in cases where the clinical rater found it difficult to assign diagnoses. The frequencies of the diagnoses assigned by the clinical rater were compared with the computer-predicted level of prevalence of any disorder, and the associations were statistically significant ($p < 0.001$): in the group of very low prevalence (< 5%), 11% of children had been assigned diagnoses, in the group of moderate prevalence (≥ 20%, < 75%), 45% and in the group of high prevalence (≥ 75%), 93%.

4.3 Statistical analyses

In the questionnaire assessments (I, III, IV), the frequencies and distributions of categorised variables were described as percentages. As continuous variables were non-normally distributed, they were described by medians and quartiles, except SDQ scores for which the 80th and 90th percentiles were used instead of quartiles (I). The significance of the SDQ score differences between groups defined by age and gender were examined using the nonparametric Mann-Whitney U test (I).

Most of the analysis in the study was conducted by stratifying the children by gender and age group. However, the results are presented without stratification if the stratified analysis did not alter the findings and when the number of cases in some of the stratified categories was small.

Due to the lack of Finnish SDQ norms for young children, the 80th British percentile was used to suggest a screen-positive case (I, II, IV) when the categorised bandings of the SDQ scores were utilised. This borderline cut-off was selected because, firstly, the particular aim was to cover possible cases extensively and, secondly, the prevalence and quality of psychiatric symptoms in this age group of young children were insufficiently studied in Finland.

In assessing the distributions of the SDQ scores, in the test-retest of the method (I) and in further validity testing of the SDQ-Fin (II), the scores of the mother- and father-reported SDQs were combined into a single parent score and those of two teachers in preschool education into a single teacher score. More precisely, when the values of both informants were available, the combined score was the mean of the scores; otherwise the score of the only informant. This could be done because of the good inter-rater reliability between both the parents and the teachers in preschool education (I).

In the Study I, the internal consistency of the SDQ scores was examined using Cronbach's alpha. Based on continuous SDQ scores, the agreements between informants (between mother and father; between the combined parent report and teacher in preschool education or school; between two professionals in preschool education) and the parent-reported test-retest were examined using Spearman's correlation coefficient. The multivariate methods of Cox regression and two-way analysis of variance (ANOVA) were utilised in examining the possible interaction of the child's age and gender on the parent- and teacher-reported SDQ total scores. In Cox regression, all cases were considered as having had the "event", and the inverse of the SDQ score represented the "survival time".

In Study IV, agreement between the informants in the one-question screens was examined using the ɣ coefficient. The response alternatives in the one-question screen for the parent and public health nurse as well as the child's self-evaluation enquiry were dichotomised for further analysis when examining the validity properties of these enquiries and when conducting logistic regression and stepwise logistic regression. The categorisations are represented in more detail in Study IV.

In the second phase of the study, the DAWBA method was used as the gold standard when examining the validity properties of both the SDQ (II) and of the one-question screen for parents and public health nurses and of the child's self-evaluation enquiry (IV). Two measures based on the DAWBA were utilised: the DAWBA computer-generated predictive measure (the prevalence of any disorder) and the DAWBA clinical rater assigned diagnoses.

The computer-generated predictive measure offered to the DAWBA rater was re-categorised from four to three levels in the analysis: < 5% prevalence (low probability of disorder), ≥ 20% and < 75% (moderate), and ≥ 75% (high probability of disorder). In Study II, this three-level computer-generated predictive measure was utilised in 1) examining the association between the DAWBA predictive measure and the SDQ total score by curve estimation, 2) assessing the discriminating capacity of the SDQ total scores by using receiver operating characteristics (ROC) analysis, 3) exploring the lower and upper cut-offs of the SDQ-Fin using ROC analysis and 4) evaluating the sensitivity and specificity values resulting from the various adjusted cut-off candidates. Furthermore, the computer-generated predictive measure was dichotomised (< 75% vs. ≥ 75%) in Study IV and utilised in 1) assessing the validity properties (sensitivity, specificity, positive and negative predictive values [PPV, NPV]) of the one-question screen for parents and nurses and of the child's self-evaluation enquiry and 2) conducting logistic regression analysis in order to assess how substantial risk factors the above-mentioned one-question screens, entered simultaneously as explanatory variables, were for the high probability of a disorder (outcome variable).

The diagnoses assigned by the DAWBA clinical rater, the second gold standard DAWBA measure, were exploited as dichotomous variables in both Study II and Study IV. In Study II, this measure was used 1) in assessing the frequencies of psychiatric diagnoses in the study sample, 2) as an extra control test in adjusting the Finnish upper cut-offs of the SDQ and 3) in estimating the diagnostic discriminative capacity of the method in the study sample. In Study IV, the existence of any or selected specific diagnoses assigned by the DAWBA rater were used, one at a time, as outcome variables in logistic regression analysis which was

conducted to determine how substantial risk factors the one-question screens were for the assigned diagnoses types.

In Study II, the discriminating capacity of the SDQ scores was examined using area under the curve (AUC) in ROC analysis. This analysis was performed separately for the SDQ lower cut-off (normal vs. borderline or abnormal score) against the DAWBA computer-predicted lower cut-off (low vs. moderate or high probability of disorder) and for the SDQ higher cut-off (normal or borderline vs. abnormal score) against the DAWBA computer-predicted higher cut-off (low or moderate vs. high probability of disorder). The AUC values of 0.90–1 were considered excellent, 0.80–0.90 good, 0.70–0.80 fair, 0.60–0.70 poor and < 0.60 fail (Tape, http://gim.unmc.edu).

The Finnish adjusted lower and higher cut-offs for the parent- and teacher-reported SDQs were sought by defining the 80th and 90th percentiles of the SDQ scores and by ROC analysis (II). The sensitivities and specificities of the suggested SDQ-Fin cut-offs were then examined against the DAWBA measures (computer-generated disorder probabilities, rater-assigned diagnoses), as described above. When choosing the cut-offs for the SDQ-Fin in the community sample of young children, the aim was to balance between recognising psychiatric symptoms sensitively enough and avoiding false negatives. Therefore, values at least 60% were required for the sensitivity of the lower cut-off and for the specificity of the higher cut-off.

In Study III, the written responses were analysed by inductive and qualitative data dissection by the researcher. The data were analysed manually without preconceived hypotheses. Firstly, the responses were explored thoroughly and then encoded according to the main themes emerging from different respondents. Secondly, the incidences of these codes were tabulated as percentage proportions of both respondents, and all responses and parallel themes were combined. The choices made by the researcher were discussed in the steering group.

The statistical analysis was made using SPSS for Windows statistical software. Version 15 was used in Study I and version 19 in Studies II, III and IV. In the present study, p-values < 0.05 were considered statistically significant.

4.4 Ethics

The study was conducted in accordance with the Declaration of Helsinki. The ethics committees of the participating hospital districts approved the study. Written

informed consent was obtained from all participating parents. Participation in the study was voluntary for the families, and the public health nurses and GPs were responsible for the process of the children's health examination and the process of referral.

5 Summary of the results

5.1 The psychometric properties and reliability of the SDQ-Fin in 4–9-year-old children (I)

5.1.1 Distributions of the scores of the SDQ-Fin

The frequency distributions of the SDQ-Fin scores reported by parents and teachers in the total sample and in the subsamples of genders and age groups (preschool-aged, n = 1,762, and school-aged, n = 904) are represented in Appendix 10 (Tables 1-14) and in Figure 2 of Study I. In addition, the means and standard deviations of the SDQ-Fin scores were calculated, despite the skewedness of the distributions, in order to enable comparison with earlier studies; see Appendix 11.

Boys were rated higher than girls in total scores as well as on the conduct and hyperactivity subscales in both preschool- and school-aged groups reported by both parents and teachers. Also on the peer problems subscale, the preschool-aged boys had higher scores than the girls did. On the emotional symptoms subscale, there were no major differences in the distributions of the scores between boys and girls in either age group or according to either informant. On the prosocial skills subscale, both informants assigned higher scores to girls than to boys in both age groups. The results on the impact scale were reversed.

Parents assigned higher total scores for children than did teachers. The differences in the parent- and teacher-rated SDQ scores between boys and girls in both age groups were all significant except on the emotional symptoms subscale. Comparing the SDQ scores between age groups revealed significant differences in all the parent-rated and teacher-rated scales except in peer relationships and prosocial skills.

Both child gender and child age had an independent effect on the parent- and teacher-rated total scores (Study I, Table 4). Boys and preschool-aged children had significantly higher scores than other groups. The gender and age of a child had no interaction on the parent-rated total score, but their interaction was significant on

the teacher-rated total score. Boys had the highest scores regardless of age group, and school-aged girls had the lowest scores.

5.1.2 The inter-rater reliability of the SDQ-Fin

The extent of agreement between mothers' and fathers' ratings for the SDQ-Fin total score was 0.65 (Spearman's correlation), for the subscales 0.50–0.65 and for the impact scale 0.54 in the total sample; see Table 3 in Study I. The highest agreement between parents was found in the hyperactivity subscale and the lowest agreements in the peer problems and emotional symptoms of the child. Parent evaluations were quite congruent between preschool- and school-aged children, with the exception that a higher level of agreement was found in school-aged children than in preschool-aged children in conduct problems and in the impact score.

The inter-rater reliability between parents and teachers for the total score was 0.43 (Spearman's correlation) for the subscales 0.26–0.47 and 0.30 for the impact scale in the total first-phase sample. The highest level of agreement between parents and teachers was found in hyperactivity and the lowest in prosocial skills. The agreement between parents' and teachers' SDQ total scores was greater in preschool-aged than in school-aged children.

On all the scales, the level of agreement was higher between mothers' and fathers' average evaluations than between teachers' and parents'. In addition, the level of agreement between mothers' and fathers' average evaluations and between parents' and teachers' average evaluations was higher for boys than for girls. The agreement level for the total score was higher between mothers and teachers ($r = 0.43–0.54$) than between fathers and teachers ($r = 0.36–0.42$).

The correlation between the total scores rated by two informants in day care was 0.81, and the correlations for the subscales were 0.59–0.71 and for the impact score 0.77. The highest level of agreement was found in hyperactivity and the lowest in emotional symptoms.

5.1.3 The internal consistency of the SDQ-Fin

In the first phase total sample, the internal consistency of the total score rated by mothers was 0.78, by fathers 0.77 and by teachers 0.86 (Cronbach's alpha). The internal consistencies of all the scales were higher in teachers' than in parents'

reports, except in the impact scale. Of the subscales, the highest internal consistency was in hyperactivity (α = 0.76–0.87) and the lowest consistencies were in peer problems (α = 0.46–0.52) and emotional symptoms (α = 0.53–0.56). The alphas were higher for boys than for girls in all the SDQ scales except in the father- and teacher-reported emotional symptoms. In the parents' reports, younger children had higher internal consistencies than older ones except in peer problems and prosocial skills. The internal consistencies of the SDQ-Fin scales rated by parents and teachers in the total sample and split by gender and age group are represented in more detail in Table 2 of Study I.

5.1.4 The test-retest reliability of the SDQ-Fin

In the test-retest study, the SDQ was administered the second time by phone to those parents who participated in the diagnostic DAWBA interview. The parent (mother or father) who completed the initial SDQ was interviewed in 92% of cases (592/646). The test-retest correlation for the parent-rated total score was 0.76 (Spearman correlation) within an interval of 12 weeks, the cumulative percent of all the available reports being 64.8% within the time interval. The stability was highest for the hyperactivity subscore 0.79; for the other subscores the correlations were between 0.60 and 0.68 and for the impact score 0.45. The test-retest correlations of the SDQ scores generally showed no remarkable differences when the ratings over the time interval of two weeks were compared with those of over 17 weeks. Only the impact scores correlation fell from 0.75 to 0.57.

5.2 The adjusted Finnish cut-offs for the SDQ in young children and the validity of the method (II)

5.2.1 The cut-offs for the SDQ-Fin and their sensitivity and specificity

In the total first phase sample, both defining the 80th and 90th percentiles of the SDQ scores and ROC analysis suggested the same cut-offs for the SDQ-Fin total scores rated by parents and teachers; see Study II, Table 2. The cut-off candidates were also defined for the SDQ total scores in the subsamples formed by the gender and age group of the children. The final selection of the optimal cut-offs was then

made by evaluating the sensitivity and specificity values against the DAWBA computer-generated disorder probabilities for all these above-mentioned cut-off candidates; see Study II, Table 3. In addition, to adjust the upper cut-offs, the DAWBA-rater-assigned diagnoses were used as an extra reference method (Study II, Table 3). This analysis or the analysis stratified by gender and age group of the children did not alter the selection of the optimal cut-offs. The selected cut-offs for the SDQ total score, their sensitivity and specificity values, and the proportion of children selected by them among the participants are represented in Table 5.1.

The nationally adjusted cut-offs considered were lower than the British ones for 4–17-year-olds (original three-band categorisation; Youthinmind, b). A comparison between the sensitivity and specificity values of the adjusted Finnish and British cut-offs in this study sample is represented in the Table 5.1.

The frequencies of children in the study sample who were scored at or above the selected Finnish or British cut-offs by parents and by teachers are shown in Table 5.1. The proportion of children selected by either the parents or the teachers according to the Finnish lower cut-off was 24% and according to the higher cut-off 17%. The respective proportions of children selected by both the parents' and the teachers' SDQ-Fin were 6% and 3%.

Table 5.1. The suggested Finnish and the British cut-offs for the parent- and teacher-rated total scores of the Strengths and Difficulties Questionnaire (SDQ), their sensitivities and specificities and the frequency of children identified by these cut-offs among Finnish 4–9-year-old children. The sensitivity and specificity values were also calculated for the British cut-offs in this Finnish study sample and compared with the nationally adjusted ones.

SDQ total score	Finnish				British			
	Adjusted cut-offs[1]	Sens[2] %	Spec[2] %	Identified children %	British cut-offs[3]	Sens[2] %	Spec[2] %	Identified children %
Parent-reported								
Lower cut-off	9/10	76	69	18	14/15	46	89	9
Higher cut-off	11/12	90	74	11	17/18	36	97	-
Teacher-reported								
Lower cut-off	9/10	66	63	15	12/13	59	71	13
Higher cut-off	11/12	70	66	11	16/17	49	86	-

[1]Suggested by the 80th and 90th percentiles and by ROC analysis.
[2]Sens = sensitivity. Spec = specificity. Calculated against the DAWBA computer-generated disorder probabilities.
[3] www.sdginfo.com, original three-band categorisation.

5.2.2 The concurrent validity of the SDQ-Fin

The discriminative capacity of the SDQ-Fin to identify psychiatric symptoms in children was assessed by ROC analysis. The AUC value for the lower cut-off of the parent-rated total score was fair (0.79) and for the higher cut-off good (0.87). For the higher cut-offs, the AUC values for the parent-rated subscores were between 0.68–0.83 and for the impact score 0.82. The AUC values were fair for the lower (0.71) and higher (0.76) cut-offs of the teacher-rated total score. For the higher cut-offs of the teacher-rated subscores, the AUC values were 0.54–0.78 and those for the impact score 0.71. The lowest AUC values were found for teacher-rated emotional symptoms (0.54) and parent-rated peer problems (0.68). The ROC curves for the total scores of both informants are represented in Appendix 12 and the detailed AUC values in Table 1 of Study II.

In curve estimation, the relationships between SDQ parent- and teacher-rated total scores and the DAWBA computer-generated disorder probabilities were positive and linear ($p < 0.001$) (Study II, Figure 2).

The frequencies of the clinician-assigned diagnoses according to the DAWBA assessment in the second phase study sample were 20% for any diagnosis (128/646), 9% for any emotional disorder, 7% for any conduct disorder, 5% for hyperactivity and 4% for other psychiatric diagnoses. Developmental diagnoses were removed from the group of other diagnoses. There were 37 children whose only diagnosis was one of the removed ones. An estimate of 8.5% for the overall frequency of any diagnosis in the entire first phase sample was extrapolated on the basis of the frequencies of assigned diagnoses in the second phase sample; the extrapolation is described in more detailed in Study II.

There were clinically important and significant differences between the frequencies of assigned diagnoses in the groups of normal, borderline and abnormal defined by the Finnish cut-offs of the parent and teacher-rated SDQ scores, represented in detail in Table 4 of Study II. The frequencies of any diagnoses in the normal and abnormal groups defined by parental total scores were 7% and 41%. The respective proportions regarding teachers' reports were 10% and 34%.

5.3 The psychometric properties of the one-question screens (IV)

5.3.1 The reliability and validity of the one-question screens

The inter-rater reliability was fairly good (ɣ = 0.73) between the parents' and public health nurses' one-question screens. The agreement level was low between a child's self-evaluation and the one-question screen for parent (ɣ = 0.10) and nurse (ɣ = 0.15). The agreement between the parent-rated one-question screen and the first question on the SDQ impact supplement was ɣ = 0.92.

The frequencies of parents', nurses' and children's own perceptions of definite or severe difficulties according to the one-question screen are represented in Table 5.2. The frequencies of all the response alternatives in the screens stratified by gender and age group are represented in more detail in Table 1 of Study IV.

The parent's and nurses' one-question screen had fairly good sensitivity (68-65%) and high specificity (87−88%) in identifying children suffering from mental health problems (Table 5.2). Of those children who were evaluated to have definite or severe difficulties, 41% had a high (≥ 75%) DAWBA computer-predicted prevalence level of any psychiatric disorder (PPV). When the adults identified no difficulties, 95% of the children had a low or moderate prevalence level (< 75%) of any psychiatric disorder (NPV). The child's self-evaluation questions had very low sensitivities and PPVs and high specificities and NPVs. The results of the above-mentioned analysis stratified by gender and age group are represented in more detail in Table 2 of Study IV.

The combination of the adults' reports and of all three informants' reports produced a sensitivity of 79%, and the respective specificities were 80% and 75%.

Logistic regression analysis was performed in order to assess how substantial risk factors the one-question screens were for the high probability of a disorder and for the assigned diagnosis types. Difficulties identified by parents and public health nurses were strongly associated (OR 14) with a DAWBA computer-predicted high prevalence level (≥ 75%) of a child psychiatric diagnosis (Table 5.2). The odds ratios for the rater-assigned psychiatric diagnoses were between 4.0 and 34.4 ($p < 0.001$) (Study IV, Table 3). The strongest association was found between the nurse's identification of difficulties and a hyperactivity diagnosis (OR 34.4). Difficulties identified by parents and nurses were least strongly associated with any emotional diagnoses (OR 4.0-4.5).

Table 5.2. Frequencies of definite or severe difficulties or concerns in parents' and nurses' one-question screens and in children's self-evaluations, and the validity of these screens in identifying children suffering from mental health problems.

	Frequency %	**Sens**[1,6] %	**Spec**[2,6] %	**PPV**[3,6] %	**NPV**[4,6] %	**OR**[5,6] (95% CI)
Parent's one-question screen quite many/very many difficulties	(n = 2652) 6.4	(n = 637) 68	(n = 637) 87	(n = 637) 41	(n = 637) 95	14.4** (8.4–24.9)
Nurse's one-question screen yes – definite/severe difficulties	(n = 2602) 6.8	(n = 622) 65	(n = 622) 88	(n = 622) 41	(n = 622) 95	13.6** (7.8–23.5)
Combined parent's and nurse's one-question screen		(n = 613) 79	(n = 613) 80	(n = 613) 34	(n = 613) 97	
Child's self-evaluation						
How are you? often / almost always sad	(n = 2623) 2.1	(n = 629) 7	(n = 629) 98	(n = 629) 26	(n = 629) 89	
What do you expect for your near future? some/many bad things are going to happen	(n = 2620) 4.8	(n = 628) 9	(n = 628) 94	(n = 628) 16	(n = 628) 89	
Combined child's self-evaluation		(n = 629) 14	(n = 629) 93	(n = 629) 20	(n = 629) 90	2.2* (1.1 – 4.7)
Combined parent's, nurse's and child's one-question screen		(n = 598) 79	(n = 598) 75	(n = 598) 29	(n = 598) 97	

[1]Sens = sensitivity, [2]Spec = specificity, [3]PPV= positive predictive value, [4]NPV = negative predictive value, [5]OR = odds ratio.
[6]Assessed against the computer-generated predictive measure of the DAWBA (< 75% vs. ≥ 75%).
*p < 0.05.
**p < 0.001.

In addition, the variables of the one-question screens, gender, age group and the interaction between age and gender were examined together in order to determine the strongest risk factors for psychopathology (Study IV, Table 4). The difficulties identified by both parents and nurses remained the strongest risk factors for all

child psychiatric outcomes, except for a diagnosis of hyperactivity, for which only the nurse's concern remained statistically significant. The school-aged children had a twofold risk for any assigned diagnoses compared with preschool-aged children, and girls had a higher risk (OR 2.3) for emotional diagnoses than the boys did. No significant interaction was found between the gender and age of the children as a risk factor for psychiatric disorders.

5.3.2 The relevance of directly asking a young child to evaluate his/her emotional well-being

Of the children in the present study, 2.1% evaluated themselves as feeling sad or miserable often or almost always, and 4.8% suspected that some or many bad things were going to happen to them. The combined child's self-evaluation enquiry had high specificity (93%) and NPV (90%). Although the sensitivity and PPV values for the negative ratings of the two questions and for the combined child's self-evaluation were low (7−26%), the combined child's self-evaluation was related to a twofold risk for a high DAWBA computer-generated prevalence level of any diagnosis (Table 5.2). Significant associations were also found with any rater-assigned diagnosis (OR 2.4), with any emotional diagnosis (OR 3.0) and with negative situational factors (OR 3.2) (Study IV, Table 3). The combined self-evaluation negative ratings remained a significant risk factor for any emotional diagnosis (OR 2.7) and for negative situational factors (OR 2.9) in the logistic regression analysis (Study IV, Table 4).

5.4 The feasibility of the SDQ-Fin and the child's self-evaluation enquiry (III, IV)

5.4.1 Feedback on the use of the SDQ-Fin (III)

The time spent on completing or going through the SDQ varied according to the informant: it took a maximum of ten minutes for most of the parents (over 88%) and a maximum of 15 minutes for most of the teachers (76%) and public health nurses (89%) (Study III, Table 2). The SDQ was found not at all or not very burdensome by 97% of the parents, by 92% of the teachers and by 62% of the

public health nurses. At least four-fifths of the respondents evaluated the SDQ to be very or fairly age-appropriate for assessing the psychosocial well-being of children.

Most informants felt that the method increased cooperation between parents and professionals (49–71%), augmented knowledge (73–88%) and raised the level of agreement (70–79%) about the children's mental health and need for support (Study III, Figure 1). In addition, parents reported that using the SDQ was a positive experience for them (88%), which was also noticed by the professionals (53–81%).

In the open-ended feedback, the professionals reported that, compared with the previously used methods, the SDQ increased focus on the mental health of the child. Some difficulties in completing and interpreting the SDQ were reported by 21—34% of the respondent groups. The parents reflected strongly on the importance of dialogue with the professionals even when a questionnaire was used in evaluating the well-being of the child. They also gave positive feedback on including the assessment of the child's mental health in regular health check-ups.

5.4.2 Feedback on the child's self-evaluation enquiry (IV)

The child's self-evaluation enquiry was evaluated to be very or fairly appropriate for assessing the psychosocial well-being of the child by 63% of the parents and 71% of the public health nurses in the total first phase sample (Table 5.4.). The method was found to be more appropriate among school-aged than among preschool-aged children. The children completed the enquiry with the public health nurses in ten minutes on average, and the younger children spent less time with it than did the older ones. The public health nurses (94%) did not find these two questions burdensome.

Table 5.4. The feasibility aspects of the child's self-evaluation enquiry reported by the parents and public health nurses.

	Parents			Public health nurses		
	Total sample %	Preschool-aged %	School-aged %	Total sample %	Preschool-aged %	School-aged %
	(n=1487)	(n=1487)	(n=523)	(n=154)	(n=105)	(n=48)
Appropriateness						
Very /fairly good	63	61	67	71	60	96
Not good, not poor	14	15	13	20	27	4
Rather/very poor	8	11	5	9	13	-
Cannot say	14	14	16	-	-	-
Burdensome						
Not at all/not much				96	96	96
Rather/very much				4	4	4
Cannot say				-	-	-
Time (min.)						
< 5				58	70	33
5-10				40	29	65
> 10				1	1	2

5.5 Summary of findings

An overview of the main findings is given in Table 5.5.

Table 5.5. An overview of the main findings.

Research aims	Main findings
1. To explore the psychometric properties and reliability of the Finnish version of the Strengths and Difficulties Questionnaire (SDQ-Fin) in 4–9-year-old children visiting for regular health check-ups in child health clinics and in school health care clinics.	Significant and clinically important differences were found in the distributions of the SDQ-Fin scores between informants and between genders and age groups of the children. Parents rated higher total scores than did teachers. Both informants rated higher total scores for boys than for girls. The inter-rater reliability of the SDQ-Fin total scores between parents was 0.65, between parents and teachers 0.43, and between two preschool teachers 0.81. The highest levels of agreement were found when evaluating the boys and the hyperactivity of the children. The internal consistencies of the SDQ-Fin total score were acceptable (α = 0.78–0.86) in all the informants' reports. The general trend was that the method worked more consistently for evaluating boys than girls and for evaluating younger than older children. The test-retest reliability of the total score in the parent-rated SDQ-Fin was high (r = 0.76) within an interval of 12 weeks.
2. To define and select the adjusted Finnish cut-offs of the SDQ-Fin in young children and to explore the capacity of the SDQ to identify the children suffering from mental health problems.	For the parent- and teacher-rated SDQ-Fin total scores, the adjusted lower cut-off was 9/10 and higher cut-off 11/12. The adjusted cut-offs were remarkably lower than the British ones. With the higher suggested total score cut-off, the sensitivity of the SDQ-Fin was 90% in parent reports and 70% in teacher reports, the respective specificity values being 74% and 66%. The higher cut-off had a good capacity to discriminate psychiatric disorders according to parents' reports (AUC 0.87) and a fair capacity according to teachers´ reports (AUC 0.76). The frequency of the assigned diagnoses was sixfold in the parent-rated and threefold in the teacher-rated abnormal group compared with the normal group defined by the adjusted cut-offs.
3. To assess the reliability and validity of one-question screens presented to children, parents and public health nurses in identifying children at elevated risk for mental health problems.	The inter-rater reliability between the parents' and nurses' one-question screens was fairly good (ɣ = 0.73), but it was low between the adults' reports and the children's own evaluation. The sensitivities of the adults' one-question screens were fairly good (65%–68%) and the specificities high. Combining the parent and nurse reports or all three informants' reports led to 79% sensitivity. Difficulties identified by parents and nurses were strongly related to any child psychiatric diagnosis (OR 14) and specific diagnoses (OR 4.0–34.4). Of the young children, 2–5% reported low mood and negative expectations, which was related to a twofold risk of any psychiatric disorder and a threefold risk of an emotional disorder and negative situational family factors.
4. To evaluate the feasibility of using the SDQ-Fin as a screening questionnaire in assessing young children's mental health and to evaluate the feasibility of directly asking young children to evaluate their emotional well-being in the context of regular health check-ups.	The parents, teachers and public health nurses found the SDQ-Fin an appropriate method in medical check-ups and the use of it had positive effects: increased cooperation and a common understanding and agreement about the child's mental health and need for support. Parents called for more dialogue with the professionals. Almost all parents and teachers and 62% of the public health nurses did not find the SDQ burdensome. Some difficulties in completing and interpreting the SDQ were reported by a minority of the respondents. The parents and public health nurses found the child's self-evaluation enquiry a fairly or very appropriate method and not burdensome in assessing the psychosocial well-being of children. The enquiry was found to be more often suitable for school-aged than for preschool-aged children.

6 Discussion

6.1 Strengths and limitations of the study

6.1.1 Study design

The strength of the study was that it was conducted in a multifaceted design in the routine clinical practice of children's regular health check-ups in child health care and school health care clinics. In addition, the most important informants in the children's everyday life were invited to participate in assessing the mental health of the child: the children themselves, both parents, the school or preschool teacher and the public health nurse. The mental well-being of the children was thus evaluated from many perspectives and using different methods, increasing the usability of the results.

Arising from the context of the study design, several limitations must be borne in mind, and the results may be biased in many ways. In the first place, the interest and resources of the enrolling municipalities and clinics may have differed from the ones that did not enrol. This is influenced by the fact that the study was conducted as a part of the project entitled "Developing children's mental health work, 2007-2009" aimed at improving the collaboration between parents, day care or school personnel, and primary health care professionals when assessing children's mental health in regular check-ups. In addition, the differing attitudes of public health nurses towards study projects in general and towards children's mental health work in particular may have affected the way they introduced the study to parents. This might have further affected the participation rates in the first phase of the study.

The study was designed not to disrupt the course of a health check-up. In a regular health check-up, it is important to share information about the well-being of the child: for this reason, the SDQ reports were shared between informants and could be used for discussion. It is possible that the preschool and school teachers were more cautious and well-considered in their SDQ answers, knowing that the parents would have the opportunity to see the answers as well. In addition, the

parent- and teacher-reported SDQs may have influenced on the results of the one-question screen reported by the public health nurses.

6.1.2 Study sample

The large sample size and the fairly homogenous age distribution of children for the study design strengthen the representativeness of the results. The study design and the sample was not, however, an attempt to reach the level of criteria of an epidemiological prevalence study. This must be kept in mind when interpreting the present results on the frequencies of children's mental health problems according to questionnaire screening in medical health check-ups and according to the diagnostic interview assessment.

The participation rates were modest in the first phase (64%), and this introduces a clear limitation of the study. Information on the non-participating children was collected in order to control for bias. The public health nurses reported that the non-participating children had difficulties twice as commonly as the participating ones (12% versus 7%). This has most likely biased the results of the SDQ score distributions and caused the frequencies of mental health problems to be underestimated. There have also been results in multicultural research indicating that parents' lower response rates on children's behavioural and emotional problems have contributed to lower problem scores than in societies with higher response rates (Rescorla et al., 2007; Rescorla et al., 2011).

In an epidemiological sample of Finnish children in 1989, slight differences in the frequencies of psychiatric symptoms were observed between urban, suburban and rural districts, and no significant differences were found between the five regional study areas (Almqvist et al., 1999). The regional representativeness of the present study is supported by these earlier findings. Languages other than Finnish were spoken by 5.4% of families with children in 2005 (Kartovaara, 2007). Thus, the exclusion of families who do not speak Finnish from the study should not have introduced an important problem for the representativeness of the study sample.

Among the SDQ respondents, however, the response activity of both parents was higher than expected (73%) and very high among schoolteachers (98%). In addition, all the public health nurses and almost all participating parents completed the one-question screen. There was also no notable attrition when the participating children filled in the self-evaluation enquiry during the health check-up visit. These

findings encourage the use of a multi-informant approach when collecting information about the mental health of children in regular health check-ups.

The representativeness of the sample at the second phase must be discussed. During the stratification process, the aim was to reach enough large numeric groups of distinct psychiatric symptoms and disorders, both among the SDQ screen-positives and according to the DAWBA computer-generated disorder probabilities, in the subsamples of preschool-aged and school-aged children and among genders. After informed, preparatory calculations of the study sample size, the data collection was continued until sufficient numbers of children with psychiatric disorders stratified by both age-groups and gender were reached for the study analyses. During the stratification, the borderline cut-off (the 80th British percentile) instead of the abnormal cut-off (the 90th British percentile) was selected to suggest a screen-positive in order to cover possible cases extensively. The sample from the second phase of the study was intentionally enriched with the screen-positive children.

In the second phase, participation activity was high (89%) in the diagnostic interview. This could indicate that the parents who had already decided to participate in the first phase of the study had a firmly positive attitude towards further assessment of the well-being of their children.

In the third phase of the study, feedback questionnaires were available from almost all teachers about the SDQ and from almost all the public health nurses about the child's self-evaluation enquiry. Other informants' activity in responding was lower: between 55-70%. This could have biased the results towards positive or negative feedback evaluations on the SDQ and on the child's self-evaluation enquiry, depending on whether the respondents were more or less favourably disposed than the non-participants. It is not possible to assess this further for two reasons: the feedback was collected anonymously, and no information was collected on non-participants in this phase of the study.

6.1.3 Methods

The ultimate methodological strength of the study was the opportunity offered to conduct criterion validity studies (II, IV). The diagnostic assessment method was utilised as the gold standard in defining the norms on the SDQ-Fin for the first time in Finland, as well as determining the validity of the one-question screens. In addition, the reliability and validity of the SDQ-Fin was evaluated in many aspects.

As the author knows, this was also the first study to carefully examine the multiple feasibility elements of the SDQ widely used in different cultures.

Many methodological limitations have to be considered, however, when interpreting the results of the present SDQ-Fin reliability (I) and validity studies (II). An informant-based bias may have been introduced by using the means of the parents' and teachers' SDQ scores, when two informants within each class of these respondents were available. The other alternative would have been to choose the higher value of two informants' ratings in each informant class. The first alternative was opted for due to the good agreement values between mother's and father's SDQ reports (Spearman's rho 0.65) and between the two day-care teachers (0.81). The mean scores could have reduced the informant-based bias by better depicting the different views of the well-being of the child compared with the other alternative. The maximum values might have produced higher adjusted cut-offs or higher sensitivity than those that were found on the basis of the combined mean scores.

The DAWBA was selected as the gold standard for many reasons. It had already been validated as a particularly suitable method for use in a large population study (R. Goodman, 1999; R. Goodman, 2001). The diagnostic assessment method was expected to adhere to the multi-informant approach, to obtain respondent-based information after a relatively short time spent training interviewers, and to collect the data in an optimally practical way and at low cost. Under these conditions, and because the method had already been translated into Finnish, the DAWBA was the best and most realistic option open to the present study. Further studies on the validity and feasibility of the DAWBA have been published after the data was collected for the present study (Aebi et al., 2012; Angold et al., 2012; Frigerio et al., 2009; A. Goodman, Heiervang, Collishaw and Goodman, 2010; Rettew, Lynch, Achenbach, Dumenci and Ivanova, 2009). These results indicate that the DAWBA was a good choice for the purposes of the present study.

It could be considered a strength or a limitation of the study that the researcher performed all the DAWBA analysis. The single rater ruled out possible disagreements between the assigned specific diagnoses with other raters but may have caused a systematic error. In spite of many efforts, we did not succeed in arranging an inter-rater reliability study. The DAWBA method was not commonly used in Finland at the time the study was conducted and a willing colleague could not be found to participate in the inter-rater process. A sample size of at least 50 patients has been suggested for inter-rater reliability studies (De Vet et al., 2011). Translating at least this number of DAWBA interviews into another language

would have been a precondition for arranging the inter-rater process with a non-Finnish-speaking clinician. Therefore, the competence of the DAWBA-rater was supported by other arrangements, as presented earlier in the Methods section.

The results of the SDQ-Fin validity properties may have biased by the fact that the SDQ and the DAWBA were not quite independent of each other in the present study. The SDQ is included in the DAWBA assessment and the skipping rules of the DAWBA interview were influenced by the SDQ. In addition, some circularity may have biased the validity results of the SDQ-Fin because the clinical DAWBA rater was only partly blind to the SDQ. The DAWBA rater was blind to the initial SDQ but, as part of the DAWBA assessment, the second SDQ was available to the rater. These possible tendencies to circularity might have increased the calculated sensitivity of the SDQ-Fin. However, in a community sample it was only reasonable to employ the DAWBA skipping rules in order to keep the interview minimally burdensome for the parents.

Different practices of administrating the methods could have influenced the ways the parents answered the SDQ questions and their participation activities. The SDQ data was collected in the form of paper questionnaires from all informants in the first phase and phone interviews with the parents in the second phase of the study. There might have been differences between individual informants in understanding the structured questions of the SDQ and choosing the response alternatives depending on whether the questionnaire was completed on paper or in a phone interview. Familiarity with the questionnaire after the first completion might have made it easier for the parents to answer the second time by phone.

It was decided to hold the diagnostic interview by phone instead of interviewing face-to-face or asking respondents to complete a form online. The aspects of simplicity and feasibility were emphasised in carrying out the data collection. The population study was conducted in two hospital districts with long distances within and between them. It was considered to be more time- and money-saving for both parents and interviewers to conduct interviews over the phone rather than face to face. The participation rates might have been lower and the rates of dropouts higher in the online completion of the DAWBA compared with phone interviews. Because the sample was not a clinical one, no technical or comprehensive support to parents completing the DAWBA themselves online would have been available in any child mental health or child psychiatric clinic: this could have caused selective participation.

The high participation rate in the interview phase could indicate that the phone interview arrangements were suitable for parents. As a point of comparison, under half of the invited parents in the Bergen Child Study participated in the DAWBA face-to-face interview (Heiervang et al., 2007). In the present study, the parents informally provided positive feedback on their experience of being interviewed over the phone.

The effects of different diagnostic interview methods have mostly been studied among adults and are not reviewed here. Not many earlier studies can be found in which the child psychiatric diagnostic assessment was administered by phone (Hawes and Dadds, 2004). Preliminary results on high level of agreement between the face-to-face and telephone administration of the parent's version of the Children's Interview for Psychiatric Syndromes (P-ChIPS) have been published (Paing, Weller, Dixon and Weller, 2010). The surveys in which DAWBA was completed online may have been faster and cheaper to administer than personal interviews, but they may also have introduced bias related to selective participation and partial or low response rates (R. Goodman, 2013; Heiervang and Goodman, 2011).

In the third phase of the study, both quantitative and qualitative data were collected in the feedback questionnaires of the SDQ. Qualitative data afforded an opportunity to receive information without preconceived hypotheses influencing what questions were asked and how they were framed.

6.2 The psychometric properties and the adjusted cut-offs of the SDQ-Fin

6.2.1 Reliability and distributions of the SDQ-Fin

The internal consistency of the total difficulties score of the SDQ-Fin rated by parents and teachers had well-accepted values. In addition, the present results replicated the findings of higher internal consistency for the teacher-reported than for the parent-reported SDQ scales and the higher values for boys than for girls. The highest values of internal consistency were in the hyperactivity subscale in concordance with earlier studies (Koskelainen et al., 2000; Malmberg et al., 2003; Niclasen et al., 2012; Stone et al., 2010). The peer problems and the emotional symptoms subscale had the lowest alphas, in concordance with findings among 3–

4-year-old Dutch children as rated by parents (Theunissen et al., 2013). This finding, however, contradicted earlier Nordic results showing the lowest alphas in the conduct problems (Koskelainen et al., 2000; Malmberg et al., 2003; Niclasen et al., 2012). The present low values of the peer problems subscale were in line with the earlier results in the review of 26 studies (Stone et al., 2010) suggesting low inter-relatedness among the items in the subscale. The present alpha values for the emotional symptoms subscale were lower than in the above-mentioned review. The items in the emotional symptoms subscale may thus assess somewhat differing constructs in young preschoolers compared with school-aged children and a different construct among Finnish parents and teachers than in other cultures, surprisingly, including the other Nordic countries as well.

The inter-rater reliability was assessed between three pairs of informants: between the parent and teacher ratings, between the mother and father ratings and between two preschool teachers. The agreement on the SDQ-Fin total score between parents and teachers (ϱ = 0.43) was moderate and close to the earlier studies (R. Goodman, 2001; Koskelainen et al., 2000; Niclasen et al., 2012; Stone et al., 2010). In concordance with earlier studies, the highest agreement between parents and teachers was found in the hyperactivity subscale and the lowest agreements in the peer problems and emotional symptoms of the child (Stone et al., 2010).

The inter-rater reliability between mother and father was considerable (ϱ = 0.65) and higher than in earlier studies assessing the differences in the reliability properties of the SDQ according to the mother's and father's ratings (Dave et al., 2008; Mellor et al., 2011). In addition, an expected result was that the inter-parental agreement was higher than the agreement between the parent and teacher ratings (Achenbach et al., 1987). Parents observe their child in the home surroundings, and teachers and parents see the child in relationships and surroundings of different kinds.

Further information was gained from the inter-rater reliability between the two preschool teachers. The inter-rater agreement between the ratings of these informants was strong (ϱ = 0.81). Also taking into account the high internal consistency (α = 0.86) for the teacher ratings of preschool-aged children, the present results indicate that the SDQ-Fin reports from day-care are to be considered reliable.

There are a few earlier studies comparing the results of the inter-rater reliability on the SDQ between boys and girls and between the age groups of children. The present finding was that all pairs of informants had higher levels of agreement

when evaluating boys than girls. The same trend has been found in inter-parental agreements (Dave et al., 2008; Mellor et al., 2011). In the meta-analysis of different measures than the SDQ, the gender of the child did, however, not affect the results on the inter-rater reliability (Achenbach et al., 1987). In addition, further information was gained on how the age of the child affected the inter-rater reliability of the SDQ. The earlier results on the issue are contradictory. In a Chinese sample (Du et al., 2008), the inter-rater agreement between the parent- and teacher-rated total scores was higher in younger children (under ten years old) than in older children, but, again, in the Danish cohorts reverse results were found (Niclasen et al., 2012). In the present study, the inter-rater reliability between parents' and teachers' SDQ total scores was higher in preschool-aged than in school-aged children, but the agreement between parents was rather congruent in these age groups. A possible interpretation for this finding is that the collaboration between parents and teachers is more regular and active in preschool-aged than in school-aged children, thus confirming shared views on the well-being of the child.

Within an interval of 12 weeks, the test-retest reliability of the parent-rated SDQ-Fin was strong for the total score ($r = 0.76$) and moderate-to-strong for the subscores ($r = 0.60$–0.79); for the impact score it was moderate ($r = 0.45$). The present test-retest result for the SDQ total score is in line with earlier studies with corresponding time intervals between the repeated measurements (R. Goodman, 2001; Muris et al., 2003). When the time interval has been clearly longer (12 months), both corresponding and lower test-retest correlations on the parent-rated total scores were found (Du et al., 2008; Hawes and Dadds, 2004). In the present study, the highest test-retest reliability was found for the hyperactivity subscore and the lowest for the impact score in concordance with the mean test-retest correlations found on six earlier studies in the review of Stone et al. (2010). It should also be noticed that there were no remarkable differences in the parent-rated SDQ-Fin scores, except in the impact scores, when the ratings between the time-interval over two weeks and over 17 weeks were compared. This finding suggests that the present test-retest results represent more the well-functioning reliability property of the SDQ-Fin than the changes in the mental health status of the children.

The distributions, 80th and 90th percentiles, of the parent- and teacher-rated SDQ total scores in the present study sample were lower than in the represented normative and computerised SDQ data (www.sdq.org) for the British (4–15-year old) and American (USA; 4–17-year-old) children (Youthinmind, b). The present frequencies of mental health problems may be underestimated because of the low

participation rates in the first phase of the study, as discussed earlier. However, the Finnish distributions of the total scores split by gender were in line with the Danish normative frequency data for 5–7-year-old girls and boys. Compared with Japanese distributions (4–15-year-olds), the Finnish parents' ratings of total scores were slightly lower but the teachers' ratings were quite comparable. Finnish parents rated higher total scores than did teachers in accordance with above-mentioned British, American, Danish and Japanese normative data (Youthinmind, b). In addition, the present study's findings confirmed the earlier results of parents and teachers rating higher total scores for boys than for girls (Bourdon et al., 2005; Du et al., 2008; Koskelainen et al., 2000; Niclasen et al., 2012; Rothenberger et al., 2008).

Finnish preschool-aged children had significantly higher scores than 7–9-year-old children according to both parents and teachers. The result is to be considered preliminary because earlier findings on the subject have been somewhat conflicting. The present result was in concordance with distribution comparisons between Italian 3–5- and 6–10-year-old children and between British 5–10- and 11–15-year-olds (Youthinmind, b). No such trends, however, have been found in Japanese and American normative distributions (Youthinmind, b). In the present study, the participation rate was higher among preschool-aged than among school-aged children, and this may also have influenced the differing frequency distributions.

6.2.2 Validity and the adjusted cut-offs of the SDQ-Fin

The necessity of nationally adjusted cut-offs for the SDQ was confirmed when the very poor screening properties of sensitivity and specificity assessed by the British cut-offs were found in the present study sample. The best options of the adjusted cut-offs of the SDQ total score were carefully explored using two methods and then confirmed by assessing their validity properties. The chosen adjusted SDQ total score cut-offs were 9/10 for the lower cut-off (normal-borderline) on the parent and teacher reports and 11/12 for the higher cut-off (borderline-abnormal). These nationally adjusted cut-offs were remarkably lower than the British and Chinese cut-offs for parent and teacher reports (Youthinmind, b). The present cut-offs for the parent-reported SDQ were also lower than reported earlier in the USA (4–17-year-olds) (Bourdon et al., 2005) and Germany (7–16-year-olds) (Rothenberger et al., 2008). However, the Finnish cut-offs for parent reports were in line with Danish cut-offs for 5–7 and 10–12-year-old boys and girls

(Youthinmind, b) as well as with Norwegian 4-year-olds (Sveen et al., 2013). In addition, the higher cut-off of the Swedish SDQ parent report (10/11) defined by AUC was very near the present finding (Malmberg et al., 2003).

Unexpectedly, no differences between genders were found in the selected adjusted cut-offs. Although the 80th and 90th percentiles of the SDQ-Fin total scores were higher for boys than for girls, the gender differences were not evident when defining the cut-offs by ROC analysis. Nor did the sensitivity and specificity analysis of the cut-off options stratified by gender provide separate cut-offs for girls and boys. Common cut-off bandings have also been recommended for girls and boys in the USA, Britain, China and Germany (Bourdon et al., 2005a; Du et al., 2008; Rothenberger et al., 2008; Youthinmind, b). In Danish norms, however, separate cut-off bandings defined by percentiles were recommended for genders (Youthinmind, b). These recommended cut-off scores for boys were 1-2 points higher than for girls in both age groups of 5-7- and 10-12-year-old children. In the present study, the assessments against the diagnostic method reinforced the validity of the adjusted cut-offs considerably.

With the adjusted cut-offs, the SDQ-Fin identified the children suffering from psychiatric symptoms and disorders accurately. Scoring at or over the higher SDQ cut-off can be used to identify likely "cases" of mental health disorders (Youthinmind, b). The sensitivity of the higher cut-off of the SDQ-Fin parent report was 90% and of the teacher report 70%; the respective specificities were 74% and 66%. Thus, only one of ten children with a high risk of psychiatric disorder according to the DAWBA was not recognised by parents on the SDQ and three of ten were unidentified by teachers, respectively. However, one of three or four of the identified children had a false positive result. The present sensitivity values of the higher cut-off were higher and specificity values were lower than in earlier concurrent validity studies in community samples (R. Goodman, 1999; R. Goodman, 2001; Sveen et al., 2013). Most screening questionnaires have been shown to exhibit poorer sensitivity than specificity values, which has been considered a problematic issue (Costello, Egger 2005). However, when detecting children's mental health problems on the front line, it is important to recognise the children with problems and offer an accurate clinical consideration of the overall situation to each identified child, and thus also reduce the rate of false positives.

The adjusted SDQ-Fin total scores discriminated well between the low- and high-risk children according to the ROC analyses and frequencies of the clinician-assigned diagnoses in the groups of normal and abnormal defined by the total sores. The AUC values of the SDQ-Fin total score higher cut-off for the parent

reports (0.87) and for the teacher reports (0.76) were higher than in the Norwegian study (Sveen et al., 2013) and fit within the ranges of weighted AUC values specified by informants represented in the review of Stone et al. (2010). Notably, the AUC value of the higher cut-off was fail for the teacher-rated emotional symptoms (0.54) and poor for the parent-rated peer problems (0.68), but for other subscales the AUCs were at least fair (≥70). The impact scores rated by parents (AUC 0.82) and by teachers (AUC 0.71) identified the high-risk children accurately according to the DAWBA, in concordance with earlier results (R. Goodman, 1999). The impact supplement thus introduced information that was clinically relevant in evaluating the symptoms of the children together with the impairment such symptoms caused.

Of the Finnish 4–9-year-old children, 17% were scored at above the higher cut-off and 24% at or above the lower cut-off by either parents or teachers in accordance with earlier Finnish questionnaire-based studies (Almqvist et al., 1999; Koskelainen et al., 2000). The present proportion of the parent-rated difficulties on the SDQ-Fin was, however, higher than earlier findings for 12-year-old children (Pihlakoski et al., 2004). Only cautious comparisons can be made between the present 9% estimate for an overall frequency of child psychiatric diagnoses for the study population and earlier prevalence rates of 9% in Finnish and 7% in Norwegian epidemiological samples (Almqvist et al., 1999; Wichstrom et al., 2012). In the present study, the conflicting result was that emotional disorders were the most frequently assigned diagnoses, although the frequencies were low for the emotional symptoms according to SDQ-Fin parent and teacher reports. Finnish parents and teachers seem to underestimate the prevalence of children's emotional symptoms, in concordance with earlier findings (Heiervang et al., 2008; Michels et al., 2013; Rothenberger et al., 2008; Sveen et al., 2013).

The SDQ-Fin total scores showed the adequate properties of inter-rater reliability, internal consistencies and test-retest reliability. In addition, the Finnish adjusted total scores had high sensitivity in identifying the children at high risk of a psychiatric disorder and discriminated well between the low- and high-risk children. However, remarkable differences on the reliability properties and criterion validity were found between the SDQ subscales. The hyperactivity subscale had the best reliability properties and good diagnostic discriminative validity. The emotional and peer problems subscales had the most problems with reliability and validity properties. The inter-parental agreement and the internal consistencies were low for both subscales. In addition, the emotional subscale had an inadequate capacity to distinguish disorders according to teachers and the peer

problems subscale according to parents. In summary, the present findings support the earlier discussion about the SDQ total scores being more reliable and valid in detecting children's mental health problems than the subscales (Mieloo et al., 2012; Stone et al., 2010; Theunissen et al., 2013).

6.3 The psychometric properties and clinical relevance of the one-question screens

The inter-rater reliability was fairly good ($\gamma = 0.73$) between the parents' and public health nurses' one-question screens. Public health nurses do not necessarily meet the 4–9-year-olds often; the regular health check-up visits are usually held once a year. As for their clinical knowledge, public health nurses have, however, an important contribution to make in evaluating a child's psychosocial health in comparison with the child's own developmental level and with peers. Notably, the public health nurses could not say or know their perception on the child only in 3% of all the eligible children and in 1% of the non-participants. The one-question screen thus offered the public health nurses an opportunity to show their perception of the well-being of the child and to act as an active informant.

The frequencies of the parents', nurses' and children's own perceptions of definite or severe difficulties according to the one-question screens are to be considered preliminary because no comparable earlier findings on very short queries in Finland are available. In the present study sample, the frequency of parent-reported definite or severe difficulties was lower according to the one-question screen (6%) than according to scoring above the SDQ-Fin higher cut-off (11%). Parents might have been more cautious in stating their perception in the form of a single question than in the multiple items of the SDQ. Parents, however, perceived some minor difficulties in almost half of the children, according to the one-question screen. The frequency of British parental concerns according to a short enquiry of four questions was 10% (Ford et al., 2005). In addition, the frequency of the self-reported low mood was notably lower in the present study (2%) than in a Belgian short self-report questionnaire on the emotional problems of 5–10-year-olds (12–16%) (Michels et al., 2013), and than in earlier studies with validated self-report measurements (Ringoot et al., 2013; Sourander et al., 2005).

The parents' and public health nurses' perceptions of children having definite or severe difficulties were fairly sensitive (65-68%) in recognising the children at high risk for a psychiatric disorder, and combining their reports produced an even

higher sensitivity (79%). As the adults reported no difficulties, the specificity, the proportion of true negatives, was high (87-88%), and only 5% of the children of no concern had a high risk for a psychiatric disorder according to the diagnostic assessment. The results were in line with earlier studies on validity of the first question in the SDQ to discriminate between low- and high-risk children and the short four-question enquiry for parents (Ford et al., 2005; R. Goodman, 1999). The combined parents' and nurses' one-question screen had lower sensitivity and higher specificity than the parent-rated SDQ higher cut-off in the present study sample. The differences between the validity properties of the above-mentioned methods could be used in everyday practise. The adults' perception of children having no difficulties seems accurate according to the one-question screen, and any reported difficulties should be further examined using a validated method. The one-question screen is thus suggested as an efficient first-stage screening method for professionals evaluating the need for a more comprehensive assessment of the mental health status and functioning of a child.

The difficulties identified on the parents' and nurses' one-question screen were both substantial and significant risk factors for a child's psychopathology (OR 4–34). The parents' perception of a child having any emotional problems or any difficulties in behaviour, concentration or social skills had a strong association with the high risk of a disorder (OR 14) and with any psychiatric diagnosis assigned by the clinician (OR 10). The strong associations found were in concordance with the earlier validity results on the SDQ parent report in Britain (R. Goodman, 2001). In addition, both the parents' and public health nurses' one-question screen remained significant risk factors for a child's psychopathology. The one-question screen for parents and public health nurses thus produces two sets of valid information complementary to each other on a child's risk of mental health problems.

As expected, the inter-rater reliability values between the children and the parents and public health nurses were low, in concordance with earlier results (Ederer, 2004; Michels et al., 2013). In addition, the differing contents of the questions in the child's self-evaluation and in the adults' one-question screen partly explained the poor agreement. The results discussed above on the low values of internal consistency and inter-rater reliability of the SDQ-Fin emotional subscale support the assumption that it is necessary to ask children directly about their emotional well-being. It is also valuable to provide children with an opportunity to play an active role in evaluating their own health and well-being in an age-appropriate way.

It was found to be clinically relevant to ask young children to self-evaluate their emotional well-being. The children's reports of low mood and negative future expectations were not sensitive in identifying psychiatric disorders, but they were significantly associated with a two- to threefold risk of any disorder, of an emotional disorder and of negative situational family factors. A child's self-evaluation of no concerns was highly specific, and false negative results occurred only in one of ten of the cases. Children's concerns must thus be taken seriously. In addition, the negative ratings on the child's self-evaluation, on the parent's and on the nurse's one-question screen remained all for threefold significant risks for an emotional disorder. Being a girl produced a twofold risk of emotional disorders, but no interaction between a child's age and gender was found for psychopathology. Earlier Norwegian results on the age and gender correlates for psychopathology have been conflicting (Heiervang et al., 2007; Wichstrom et al., 2012). The child's self-evaluation of well-being thus brought a desirable, validated and useful adjunct to the multi-informant evaluation of children's mental health and especially to the detection of emotional problems.

6.4 The feasibility of the SDQ-Fin and the child's self-evaluation

Feedback on the practical feasibility aspects of the brevity and simplicity of the SDQ-Fin in the regular health check-ups was mainly positive according to the parents, teachers and public health nurses. However, some participants among each of the respondent groups commented on having some difficulty in filling in and interpreting the questionnaire, in accordance with earlier findings (Vogels et al., 2009). Some of the public health nurses also criticised the use of the SDQ-Fin as rather burdensome. Notably, the use of any standardised questionnaire in evaluating children's mental health had not been prevailing practice in regular health check-ups, and the public health nurses were not familiar with the use of the SDQ. In addition, it might have been difficult for the public health nurses to distinguish how particularly burdensome the use of the SDQ was in the middle of busy check-up visits which were under the extra strain of the various procedures and administrative tasks of the present study and developing project. After the study process, some of the participated public health nurses provided information off the record that the entire study design was quite burdensome and time-consuming and that they had been forced to learn many new things in addition to the use of the SDQ.

The SDQ-Fin fulfilled the feasibility criteria (Slade et al., 1999; Slade et al., 2001) of being acceptable and relevant to respondents and to clinical judgement and being valuable in providing a more comprehensive assessment than would have been possible without it. The parents encouraged the inclusion of the SDQ-Fin in children´s health monitoring, and the public health nurses indicated that, in their experience, the method increased the focus on the mental health of children.

The use of the SDQ had several effects on the cooperation between the parents and front-line professionals. The parents' experience of the method was positive, and all the informants reported that it increased knowledge and agreement between the parents and professionals about the children's mental health. The use of a standardised questionnaire in assessing children's mental health might thus be interpreted as an intervention having a positive influence on the interaction between parents and professionals. However, the parents highlighted the importance of interactive conversations about their child's situation with the professionals. The collaboration between parents and clinicians has already previously been considered primary in the choice of an assessment method in child mental health services (Stasiak et al., 2013).

The children's self-evaluation enquiry of their emotional well-being was evaluated as being quite appropriate and not burdensome for its purpose. The public health nurses reported a clearly higher appropriateness among the school-aged children than among the preschool-aged children. The public health nurses' impressions were considered important because the children filled in the enquiry with the help of these nurses. The study sample consisted of 4–9-year-old children, and it should be noted that children develop their cognitive and linguistic skills substantially during these years. The older children could probably understand the questions and discuss the subject more easily than the younger children. It is a special challenge to find appropriate self-report methods that are valid and useful during the years of rapid childhood development.

7 Conclusions

1. The parent- and teacher-rated SDQ-Fin was found to be a reliable method in detecting children's mental health problems among 4–9-year-olds visiting child health clinics and school health care clinics for regular health check-ups.

- The properties of inter-rater reliability, internal consistency and test-retest reliability were accurate for the parent- and teacher-rated total scores. The SDQ-Fin total scores were more reliable than the subscales in detecting mental health problems in young Finnish children.
- The study documented that the gender and age of the target population, type of informants and cultural differences are important to assess and should be taken into account when evaluating the psychometric properties and reliability of the SDQ.

2. The suggested adjusted cut-offs for the SDQ-Fin in young children were defined and selected in the criterion validity study, and they were found to be considerably lower than the British norms. The result is likely to be affected by the low participation rates in the present study. The necessity of verifying the nationally adjusted cut-offs was confirmed in order to accurately identify the children suffering from mental health problems.

- The adjusted Finnish SDQ total scores had a high sensitivity in identifying the children with a high risk for a psychiatric disorder and discriminated well between low- and high risk children.
- The frequencies of psychiatric symptoms and disorders among the study sample of young Finnish children were comparable with earlier Finnish findings on school-aged children.

3. The one-question screen for parents and public health nurses showed the good properties of inter-rater reliability and validity to identify the children at elevated risk for mental health problems. The child's self-evaluation enquiry of emotional well-being generated clinically relevant information that was complementary to that in adults' reports on the risk for mental health problems and especially on emotional problems.

- Combining parents' and public health nurses' perceptions from the one-question screen improved the accuracy of the method in identifying the children with an elevated risk for mental health problems.
- The one-question screen is suggested as an efficient first-stage screening method for professionals to evaluate the need for a more comprehensive assessment of the mental status and functioning of a child with standardised methods.
- The child's self-evaluation enquiry of emotional well-being is suggested as a useful way to ask young children for their own perspective on their mental well-being.

4. The SDQ-Fin was found to be a feasible method for assessing young children's mental health according to the feedback from the parents and front-line professionals. Parents and public health nurses found the child's self-evaluation enquiry of emotional well-being to be appropriate and not a burdensome method for 4–9-year-old children's regular health check-ups.

- SDQ-Fin parent and teacher versions were accepted as short, mostly simple to use, appropriate, relevant and valuable for assessing young children's mental health, and the method can thus be recommended for routine clinical use in the context of regular health check-ups.

8 Implications for clinical practice and future research

The detection of children's mental health is one of the basic tasks of health monitoring in primary health care. However, standardised methods are seldom used systematically in children's regular health check-ups. The present results on the reliability, validity and feasibility of the parent- and teacher-rated SDQ-Fin support the use of the method in detecting mental health problems in primary health care in 4–9-year-old children under the conditions outlined below.

Firstly, further evaluations are needed on the principles and practice of the screening process before the SDQ-Fin may be considered an official large-scale screening method. Aspects to be considered include, for example, the ethical and psychological meanings of the screening for children and their families, the facilities for and the costs of the screening procedure, and resources of further assessments and the treatment of the children's mental health problems (Mäkelä and Autti-Rämö, 2014; Wilson and Jungner, 1968).

The SDQ-Fin total scores showed adequate reliability and validity properties, but not all the subscales worked consistently. Therefore, the use and interpretation of the SDQ-Fin total scores is recommended in identifying those children with mental health problems in everyday front-line clinical practice.

As an important clinical implication, the suggested Finnish cut-offs for 4–9-year-old children were represented in the study. With the suggested higher cut-off of 11/12 for the parent- and teacher-rated SDQ-Fin, the method was found to be sensitive in identifying young children with a high risk for psychiatric disorders but not so specific in avoiding unnecessary concerns about the mental health of the child. In interpreting the SDQ results, professionals in primary health care should keep in mind that the results are suggestive and only a starting point for a more comprehensive clinical consideration of a child's overall situation.

In the future, it will hopefully be possible to further validate the SDQ-Fin in representative national cohorts and define the adjusted cut-offs in epidemiologic samples of different age groups of children and adolescents.

A specific interest in the present study was to assess how quickly and easily it would be possible to identify the children at elevated risk for mental health problems. Instead of the systematic administration of standardised assessment

methods, ordinary questions on the well-being of children seem to be prevalent practice in primary health care. The tested one-question screen for parents and nurses showed good reliability and validity properties and could thus be suggested for use in everyday clinical practice in regular health check-ups. Parents' and public health nurses' perceptions of children having no difficulties was found to be accurate, but it is vital that any reported difficulties are investigated further using a validated method. The parents' and public health nurses' one-question screen is suggested as a short first-stage screen to evaluate the need for a more comprehensive assessment of the mental health status and functioning of a child using standardised methods.

Young children at the age of 4–9 were found to be capable providing useful and valid information on their mental health. Detecting children's emotional problems meets special challenges because of the lack of suitable assessment methods for young children and the low agreement between adults' and children's evaluations on internalising symptoms. With its two pictorial questions, the child's self-evaluation brought complementary information to the adults' reports, especially on emotional problems. The result supported earlier reports on young children's competence in evaluating their emotional well-being (Ialongo, Edelsohn and Kellam, 2001; Luby, Belden, Sullivan and Spitznagel, 2007; Michels et al., 2013). A child's self-evaluation of no concerns was highly specific, and thus children's concerns must always be taken seriously.

In future, more research is needed on the reliability and validity of young children's short self-reports. The method was developed for the present study, and further studies on and replication of the psychometric properties of the assessment methods of the same kind are needed to gain an overview of the clinical relevance of such an approach among young children.

The necessity of the multi-informant approach in detecting and evaluating children's mental health was substantiated in the study. Both parents were active in participating in the study, and the agreement between mothers and fathers was high in evaluating their children's mental health. These results encourage the professionals to engage the fathers in more active participation than exists at present in children's mental health services. In order to gain a general view of the child's symptoms and functioning, perceptions must be collected both from the home surroundings and from school or day-care. The SDQ-Fin was found to be useful for this purpose. The public health nurses' contribution proved to be significant and complementary to that from other informants on evaluating children's mental health. Public health nurses have an important role in recognising

children with mental health problems in health monitoring. The present results demonstrated the relevance of placing the public health nurses' own clinical judgement side by side with collecting and interpreting multi-informant information on the child's symptoms and functioning.

In order to be accepted and used in everyday clinical practice, the feasibility of an assessment method is crucial. The SDQ-Fin was found to be feasible according to parents, public health nurses and preschool and school teachers, which supports its implementation in health monitoring. The child's self-evaluation enquiry was also judged appropriate and not burdensome to use for young children at regular health check-ups. The professionals must, however, keep in mind how sensitive a matter it is for parents evaluating their child's mental health. A respectful and conversational collaboration between professionals and parents is important, regardless of the kind of measure that is used.

Further studies are needed on the feasibility of scoring and interpreting the results of the SDQ in health monitoring. The analysis of the SDQ reports was not made available to the public health nurses during health check-up visits, so it was not possible to collect feedback on how the nurses experienced SDQ scoring and analysis, nor was feedback collected on how the use of the SDQ would affect referral to care.

Introducing standardised questionnaires for detecting children's mental health problems at regular health check-ups would be an intervention that could have several advantages and possibly negative effects as well. Negative effects could be the result if the purpose of measurement is unclear and there is incompetence in interpreting the results in the context of the overall situation of the child. Adequate treatment and support must be arranged for the identified children with mental health problems. The comprehensive administration of standardised measures would ensure efficient and homogenous monitoring of children's mental health problems and enable the collection of quantitative data on their prevalence. It was documented in the present study that a standardised method helped the public health nurses focus on mental health issues. In addition, a standardised method might make it easier for the professionals to bring up the sensitive subject of concern about the mental health of a child. The parents approved using the SDQ in detecting issues in children's mental health. The early detection of children's mental health problems calls for developing child mental health networks, treatment plans and opportunities for the front-line professionals to consult specialised mental health professionals.

9 Acknowledgements

I owe my deepest gratitude and sincere respect to my supervisors Professor Tuula Tamminen, M.D., PhD, and professor Matti Joukamaa, M.D., PhD, for guiding me into scientific thinking and encouraging me through all the stages of this study. I was given a chance to enjoy and learn new perspectives from the productive collaboration of these brilliant supervisors. I am grateful to Tuula Tamminen for inspiring and inviting me into the discipline of child psychiatry. Under her wise, understanding and dynamic guidance, I have had the privilege to specialise in child psychiatry and prepare my dissertation. I warmly thank Matti Joukamaa for sharing his profound scientific knowledge from the very start of designing the study. He has guided me gentlemanly by the means of constructive criticism in balance with encouragement and warm support.

I wish to sincerely thank the official reviewers of this dissertation, Docent Päivi Santalahti, M.D., PhD, and Docent Linnea Karlsson, M.D., PhD, for their careful and expert review, constructive advice and wise comments, which helped me to improve the final form of the manuscript.

I am also very grateful to Pälvi Kaukonen, M.D., for her solid support and encouraging feedback as a co-author and as a member of the study steering group. I express my sincere thanks to Pälvi for providing me with the opportunity to carry out this study as a part of the project "Developing children's mental health work, 2007─2009" (LAMIKE).

I am indebted to Raili Salmelin, MSc, PhD, for her patience in teaching me statistical analyses, for her guidance analysing the data, and for her valuable help in preparing the manuscripts. Under stressful conditions, I learned that Raili's timetables were flexible but her work remained precise. Thank you, Raili for your faithful help in every turn!

I wish to acknowledge Professor Kari Mattila, M.D., PhD, for being a member of the study steering group and giving his scientific expertise of general practice at disposal. I appreciate his friendly guidance and trustworthy backup in case of significant turning points in the study process.

I owe my warm-hearted thanks to Virginia M. Mattila, MA, for her competent, quick and friendly responses to my requests for help with the English language in

all the original communications. I am also grateful to the Language Services staff in the University of Tampere for their invaluable assistance in editing this dissertation.

I express my deep gratitude to the participating families, the public health nurses, the preschool teachers and schoolteachers for making the study possible.

I am grateful to all the team members in the LAMIKE project for their contribution to collect and process the data. Particularly, I owe my warm thanks to Minna Sorsa, MSc, for coordinating the pilot study; to Irmeli Nenonen, M.D., and Ulla-Pia Kokko, RN, for conducting the data collection in South Karelia; to the DAWBA interviewers and to Pirkko Koskelainen, office worker. My special thanks belong to Sari Miettinen, RN, for coordinating the data collecting carefully, systematically and tenderly like it was her own child. I also thank her for being such a good and encouraging friend during the ups and downs of the study process, and for our memorable and joyful moments together in conferences.

I warmly thank all of my research colleagues in the University of Tampere for being such inspiring examples as researchers and for their interest in my study. I am particularly grateful to the experienced child psychiatrists, Ilona Luoma, Mirjami Mäntymaa, Reija Latva and Merja Mäki, who friendly gave their time to help me with the difficult cases in the diagnostic phase of the study.

This dissertation was financially supported by grants from the Finnish Child Psychiatric Research Foundation, the Emil Aaltonen Foundation, the Medical Research Foundation of Tampere University Hospital, the Finnish Foundation for Pediatric Research, the Finnish Brain Foundation and the Research Foundation of Tampere, which are all gratefully acknowledged.

I thank my mother Marju and late father Matti for all love and care and my sisters Ilse and Taru, my stepfather Pepe and stepbrother Jussi for being loved just the way I am. I am grateful to my mother- and father-in-law, Anja and Olavi Borg, for their loving encouragement. I am happy to be blessed by friends who have been praying and standing beside me all these years with many happy and sad events in life. My beloved husband Sami has been my best support and emotional anchorage. My life is a fabulous adventure with him and with our treasures: Sirkku, Sointu, Topi, Maisa and Elmo. My children have known wisely when mummy needed extra encouragement and when she had better to rest. Elmo was grown up together with this dissertation and has ensured me the armfuls of joy every day.

Tampere, 14th June 2015

Anne-Mari Borg

10 References

Achenbach, T.M., Rescorla, L.A. (2001). Manual for the ASEBA School-age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.

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

Achenbach, T. M., McConaughy, S. H. and Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232.

Aebi, M., Kuhn, C., Metzke, C. W., Stringaris, A., Goodman, R. and Steinhausen, H. C. (2012). The use of the Development and Well-Being Assessment (DAWBA) in clinical practice: A randomized trial. *European Child & Adolescent Psychiatry, 21*(10), 559-567. doi:10.1007/s00787-012-0293-6; 10.1007/s00787-012-0293-6

Almqvist, F., Kumpulainen, K., Ikaheimo, K., Linna, S. L., Henttonen, I., Huikko, E., . . . Moilanen, I. (1999). Behavioural and emotional symptoms in 8-9-year-old children. *European Child & Adolescent Psychiatry, 8 Suppl 4*, 7-16.

Almqvist, F., Puura, K., Kumpulainen, K., Tuompo-Johansson, E., Henttonen, I., Huikko, E., . . . Tamminen, T. (1999). Psychiatric disorders in 8-9-year-old children based on a diagnostic interview with the parents. *European Child & Adolescent Psychiatry, 8 Suppl 4*, 17-28.

Altman, D. (1991). Diagnostic tests. *Practical statistics for medical research* (pp. 409-419). United States of America: Chapman & Hall/CRC.

American Psychiatric Association (Ed.). (1994). *Diagnostic and statistical manual of mental disorders, 4th edition (DSM-IV).* Washington, DC: American Psychiatric Association.

Angold, A. and Egger, H. L. (2007). Preschool psychopathology: Lessons for the lifespan. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 48*(10), 961-966. doi:10.1111/j.1469-7610.2007.01832.x

Angold, A., Erkanli, A., Copeland, W., Goodman, R., Fisher, P. W. and Costello, E. J. (2012). Psychiatric diagnostic interviews for children and adolescents: A comparative study. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(5), 506-517. doi:10.1016/j.jaac.2012.02.020

Aronen, E. T. and Arajarvi, T. (2000). Effects of early intervention on psychiatric symptoms of young adults in low-risk and high-risk families. *The American Journal of Orthopsychiatry*, 70(2), 223-232.

Batty, M. J., Moldavsky, M., Foroushani, P. S., Pass, S., Marriott, M., Sayal, K. and Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: From present to future practice. *Child and Adolescent Mental Health, 18*(2), 82-87. doi:10.1111/j.1475-3588.2012.00658.x

Beaton, D. E., Bombardier, C., Guillemin, F. and Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186-3191.

Björklund, K., Liski, A., Samposalo, H., Lindblom, J., Hella, J., Huhtinen, H., . . . Santalahti, P. (2014). "Together at school" - a school-based intervention program to promote socio-emotional skills and mental health in children: Study protocol for a cluster randomized controlled trial. *BMC Public Health*, 14, 1042-2458-14-1042. doi:10.1186/1471-2458-14-1042

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G. and Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(6), 557-564. doi:10.1097/01.chi.0000159157.57075.c8

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K. and Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: Screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology, 29*(2), 143-155.

Campion, J., Bhui, K., Bhugra, D. and European Psychiatric Association. (2012). European Psychiatric Association (EPA) guidance on prevention of mental disorders. *European Psychiatry : The Journal of the Association of European Psychiatrists, 27*(2), 68-80. doi:10.1016/j.eurpsy.2011.10.004; 10.1016/j.eurpsy.2011.10.004

Carter, A. S., Briggs-Gowan, M. J. and Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 45*(1), 109-134.

Costello, E. J., Egger, H. and Angold, A. (2005). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(10), 972-986. doi:10.1097/01.chi.0000172552.41596.6f

Dave, S., Nazareth, I., Senior, R. and Sherr, L. (2008). A comparison of father and mother report of child behaviour on the Strengths and Difficulties Questionnaire. *Child Psychiatry and Human Development, 39*(4), 399-413. doi:10.1007/s10578-008-0097-6

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B. and Knol, D. L. (2011). *Measurement in medicine. A practical guide.* New York, United States: Cambridge University Press.

Du, Y., Kou, J. and Coghill, D. (2008). The validity, reliability and normative scores of the parent, teacher and self -report versions of the Strengths and Difficulties Questionnaire in China. *Child and Adolescent Psychiatry and Mental Health, 2*(1), 8. doi:10.1186/1753-2000-2-8

Dulcan, M. K., Costello, E. J., Costello, A. J., Edelbrock, C., Brent, D. and Janiszewski, S. (1990). The pediatrician as gatekeeper to mental health care for children: Do parents' concerns open the gate? *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(3), 453-458. doi:S0890-8567(09)65154-6 [pii]

Ederer, E. M. (2004). Mental health problems in young children: Self-reports and significant others as informants. *Psychology Science, 46*(Suppl1), 123-140.

Egger, H. L. and Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47*(3-4), 313-337. doi:10.1111/j.1469-7610.2006.01618.x

Elander, J. and Rutter, M. (1996). Use and development of the Rutter parent's and teacher's scales. *International Journal of Methods in Psychiatric Research, 6*, 63-78.

Elberling, H., Linneberg, A., Olsen, E. M., Goodman, R. and Skovgaard, A. M. (2010). The prevalence of SDQ-measured mental health problems at age 5-7 years and identification of predictors from birth to preschool age in a Danish birth cohort: The Copenhagen Child Cohort 2000. *European Child & Adolescent Psychiatry, 19*(9), 725-735. doi:10.1007/s00787-010-0110-z

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E. and Domenech, J. M. (2013). Psychometric properties of the Strengths and Difficulties Questionnaire (3-4) in 3-year-old preschoolers. *Comprehensive Psychiatry, 54*(3), 282-291. doi:10.1016/j.comppsych.2012.07.009

Fitzpatrick, R., Davey, C., Buxton, M. J. and Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment (Winchester, England), 2*(14), I-IV, 1-74.

Ford, T., Goodman, R. and Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(10), 1203-1211. doi:10.1097/01.chi.0000081820.25107.ae

Ford, T., Sayal, K., Meltzer, H. and Goodman, R. (2005). Parental concerns about their child's emotions and behaviour and referral to specialist services: General population survey. *BMJ (Clinical Research Ed.), 331*(7530), 1435-1436. doi:331/7530/1435 [pii]

Frigerio, A., Rucci, P., Goodman, R., Ammaniti, M., Carlet, O., Cavolina, P., . . . Molteni, M. (2009). Prevalence and correlates of mental disorders among

adolescents in Italy: The PrISMA study. *European Child & Adolescent Psychiatry, 18*(4), 217-226. doi:10.1007/s00787-008-0720-x

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M. and Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews* (Online), 2, CD008225. doi:10.1002/14651858.CD008225.pub2

Gleason, M. M., Zamfirescu, A., Egger, H. L., Nelson, C. A.,3rd, Fox, N. A. and Zeanah, C. H. (2011). Epidemiology of psychiatric disorders in very young children in a Romanian pediatric setting. *European Child & Adolescent Psychiatry, 20*(10), 527-535. doi:10.1007/s00787-011-0214-0; 10.1007/s00787-011-0214-0

Gold, J., Buonopane, R. J., Caggiano, R. A., Picciotto, M., Vogeli, C., Kanner, N. T., . . . Murphy, J. M. (2009). Assessing outcomes in child psychiatry. *The American Journal of Managed Care, 15*(4), 210-216.

Goodman, A. and Goodman, R. (2011). Population mean scores predict child mental disorder rates: Validating SDQ prevalence estimators in Britain. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 52*(1), 100-108. doi:10.1111/j.1469-7610.2010.02278.x; 10.1111/j.1469-7610.2010.02278.x

Goodman, A., Heiervang, E., Collishaw, S. and Goodman, R. (2011). The 'DAWBA bands' as an ordered-categorical measure of child mental health: Description and validation in British and Norwegian samples. *Social Psychiatry and Psychiatric Epidemiology, 46*(6), 521-532. doi:10.1007/s00127-010-0219-x

Goodman, A., Heiervang, E., Fleitlich-Bilyk, B., Alyahri, A., Patel, V., Mullick, M. S., . . . Goodman, R. (2011). Cross-national differences in questionnaires do not necessarily reflect comparable differences in disorder prevalence. *Social Psychiatry and Psychiatric Epidemiology,* doi:10.1007/s00127-011-0440-2

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 38*(5), 581-586.

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 40*(5), 791-799.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337-1345. doi:10.1097/00004583-200111000-00015

Goodman, R. (2013). Online child and adolescent mental health surveys can be good enough. *Social Psychiatry and Psychiatric Epidemiology,* doi:10.1007/s00127-013-0658-2

Goodman, R., Ford, T., Richards, H., Gatward, R. and Meltzer, H. (2000). The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 41*(5), 645-655.

Goodman, R., Ford, T., Simmons, H., Gatward, R. and Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry : The Journal of Mental Science, 177*, 534-539.

Goodman, R., Meltzer, H. and Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry, 7*(3), 125-130.

Goodman, R., Renfrew, D. and Mullick, M. (2000). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child & Adolescent Psychiatry, 9*(2), 129-134.

Goodman, R. and Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Child Psychology, 27*(1), 17-24.

Goodman, R., Slobodskaya, H. and Knyazev, G. (2005). Russian child mental health-a cross-sectional study of prevalence and risk factors. *European Child & Adolescent Psychiatry, 14*(1), 28-33. doi:10.1007/s00787-005-0420-8

Gyllenberg, D., Sourander, A., Niemelä, S., Helenius, H., Sillanmäki, L., Piha, J., . . . Almqvist, F. (2010). Childhood predictors of later psychiatric hospital treatment: Findings from the Finnish 1981 Birth Cohort Study. *European Child & Adolescent Psychiatry, 19*(11), 823-833. doi:10.1007/s00787-010-0129-1

Gyllenberg, D., Sourander, A., Niemelä, S., Helenius, H., Sillanmäki, L., Ristkari, T., . . . Almqvist, F. (2011). Childhood predictors of use and costs of antidepressant medication by age 24 years: Findings from the Finnish Nationwide 1981 Birth Cohort Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(4), 406-15, 415.e1. doi:10.1016/j.jaac.2010.12.016

Haapsamo, H., Ebeling, H., Soini, H., Joskitt, L., Larinen, K., Penninkilampi-Kerola, V., . . . Moilanen, I. (2009). Screening infants with social and emotional problems: A pilot study of the Brief Infant Toddler Social and Emotional Assessment (BITSEA) in northern Finland. *International Journal of Circumpolar Health, 68*(4), 386-393. doi:804 [pii]

Hakama, M. and Malila, N. (2008). Millainen on hyvä seulonta? Duodecim, Lääketieteellinen Aikakauskirja, 123(2), 193-199.

Hawes, D. J. and Dadds, M. R. (2004). Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *The Australian and New Zealand Journal of Psychiatry, 38*(8), 644-651. doi:10.1111/j.1440-1614.2004.01427.x

Heiervang, E., Goodman, A. and Goodman, R. (2008). The Nordic advantage in child mental health: Separating health differences from reporting style in a cross-cultural comparison of psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 49*(6), 678-685. doi:10.1111/j.1469-7610.2008.01882.x

Heiervang, E. and Goodman, R. (2011). Advantages and limitations of web-based surveys: Evidence from a child mental health survey. *Social Psychiatry and Psychiatric Epidemiology, 46*(1), 69-76. doi:10.1007/s00127-009-0171-9; 10.1007/s00127-009-0171-9

Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M. B., . . . Gillberg, C. (2007). Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service

use. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(4), 438-447. doi:10.1097/chi.0b013e31803062bf

Heikkinen, A., Puura, K., Ala-Laurila, E. L., Niskanen, T. and Mattila, K. (2002). Child psychiatric skills in primary healthcare - self-evaluation of Finnish health centre doctors. *Child: Care, Health and Development, 28*(2), 131-137. doi:252 [pii]

Helminen, M-L. and Pietiläinen, M.(2014). Maahanmuutto moninaistaa lasten perheitä. Hyvinvointikatsaus 1/2014. *Hyvinvointitilastollinen aikakausilehti* 25:22-28. Tilastokeskus. Retrieved from: http://tilastokeskus.fi/artikkelit/2014/art_2014-02-26_002.html

Honkinen, P. L., Aromaa, M., Suominen, S., Rautava, P., Sourander, A., Helenius, H. and Sillanpaa, M. (2009). Early childhood psychological problems predict a poor sense of coherence in adolescents: A 15-year follow-up study. *Journal of Health Psychology, 14*(4), 587-600. doi:10.1177/1359105309103578

Hysing, M., Elgen, I., Gillberg, C., Lie, S. A. and Lundervold, A. J. (2007). Chronic physical illness and mental health in children. Results from a large-scale population study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 48*(8), 785-792. doi:10.1111/j.1469-7610.2007.01755.x

Ikäheimo, K. (1999). Lastenpsykiatrinen hoitoonohjautumisprosessi. Turku: *Turun yliopiston julkaisuja* ser. C, tom. 151.

Jellinek, M. S., Murphy, J. M. and Burns, B. J. (1986). Brief psychosocial screening in outpatient pediatric practice. *The Journal of Pediatrics, 109*(2), 371-378.

Jellinek, M. S., Murphy, J. M., Robinson, J., Feins, A., Lamb, S. and Fenton, T. (1988). Pediatric Symptom Checklist: Screening school-age children for psychosocial dysfunction. *The Journal of Pediatrics, 112*(2), 201-209. doi:http://dx.doi.org/10.1016/S0022-3476(88)80056-8

Jokela, M., Ferrie, J. and Kivimäki, M. (2009). Childhood problem behaviors and death by midlife: The British National Child Development Study. Journal of the American Academy of Child and Adolescent Psychiatry, 48(1), 19-24. doi:10.1097/CHI.0b013e31818b1c76

Jozefiak, T., Larsson, B., Wichstrom, L. and Rimehaug, T. (2012). Competence and emotional/behavioural problems in 7-16-year-old Norwegian school children as reported by parents. *Nordic Journal of Psychiatry, 66*(5), 311-319. doi:10.3109/08039488.2011.638934 [doi]

Karoly, L., Kilburn, R. and Cannon, J. (2005). *Early childhood interventions: Proven results, future promise*. Prepared for the PNC Financial Services Group, Inc. CA: RAND Corporation, 2005, Santa Monica. Retrieved from: http://www.rand.org/pubs/monographs/MG341

Kartovaara, L. (2007). Lasten asuinpaikka, kieli, kansalaisuus ja syntymämaa. In the publication: Suomalainen lapsi 2007, pages 9-30. Tilastokeskus & Stakes. Helsinki: Tilastokeskus 2007.

Kerr, D. C., Lunkenheimer, E. S. and Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 48*(10), 967-975. doi:JCPP1776 [pii]

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., . . . Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) surveys. *Epidemiologia E Psichiatria Sociale, 18*(1), 23-33.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R. and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 593-602. doi:62/6/593 [pii]

Klein, A. M., Otto, Y., Fuchs, S., Reibiger, I. and von Klitzing, K. (2014). A prospective study of behavioral and emotional symptoms in preschoolers. *European Child & Adolescent Psychiatry,* doi:10.1007/s00787-014-0575-2

Koskelainen, M., Sourander, A. and Kaljonen, A. (2000). The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *European Child & Adolescent Psychiatry, 9*(4), 277-284.

Koskelainen, M., Sourander, A. and Vauras, M. (2001). Self-reported Strengths and Difficulties in a community sample of Finnish adolescents. *European Child & Adolescent Psychiatry, 10*(3), 180-185.

Kovacs, M. (1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica,* 46 (5-6), 305-315.

Kristensen, S., Henriksen, T. B. and Bilenberg, N. (2010). The Child Behavior Checklist for ages 1.5-5 (CBCL/1(1/2)-5): Assessment and analysis of parent- and caregiver-reported problems in a population-based sample of Danish preschool children. *Nordic Journal of Psychiatry, 64*(3), 203-209. doi:10.3109/08039480903456595; 10.3109/08039480903456595

Laajasalo, T. and Pirkola, S.(edit.) . (2012). Ennen kuin on liian myöhäistä. Ehkäisevän mielenterveystyön toimivia käytäntöjä palvelujärjestelmän kehittäjille. *Terveyden ja hyvinvoinnin laitos, Raportti 47/2012*. Juvenes Print - Tampereen yliopistopaino Oy. Retrieved from: http://urn.fi/URN:ISBN:978-952-245-686-1

Lehti, V., Niemelä, S., Heinze, M., Sillanmäki, L., Helenius, H., Piha, J., . . . Sourander, A. (2012). Childhood predictors of becoming a teenage mother among Finnish girls. *Acta Obstetricia Et Gynecologica Scandinavica, 91*(11), 1319-1325. doi:10.1111/j.1600-0412.2012.01517.x

Mäki, P., Hakulinen-Viitanen, T., Kaikkonen, R., Koponen, P., Ovaskainen, M., Sippola, R., . . . LATE-työryhmä. (2010). Lasten terveys. LATE-tutkimuksen perustulokset lasten kasvusta, kehityksestä, terveydestä, terveystottumuksista ja kasvuympäristöstä. *Raportti 2/2010. Terveyden ja Hyvinvoinnin laitos*, Helsinki.

Malmberg, M., Rydell, A. M. and Smedje, H. (2003). Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-swe). *Nordic Journal of Psychiatry, 57*(5), 357-363. doi:10.1080/08039480310002697

Marzocchi, G. M., Capron, C., Di Pietro, M., Duran Tauleria, E., Duyme, M., Frigerio, A., . . . Therond, C. (2004). The use of the Strengths and Difficulties Questionnaire (SDQ) in southern European countries. *European Child & Adolescent Psychiatry, 13 Suppl 2*, II40-6. doi:10.1007/s00787-004-2007-1

Mathai, J., Anderson, P. and Bourne, A. (2004). Comparing psychiatric diagnoses generated by the Strengths and Difficulties Questionnaire with diagnoses made by clinicians. *The Australian and New Zealand Journal of Psychiatry, 38*(8), 639-643. doi:10.1111/j.1440-1614.2004.01428.x

Mellor, D., Wong, J. and Xu, X. (2011). Interparent agreement on the Strengths and Difficulties Questionnaire: a Chinese study. *Journal of Clinical Child and Adolescent Psychology : The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, 40*(6), 890-896. doi:10.1080/15374416.2011.614580; 10.1080/15374416.2011.614580

Merikangas, K. R., He, J. P., Brody, D., Fisher, P. W., Bourdon, K. and Koretz, D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001-2004 NHANES. *Pediatrics, 125*(1), 75-81. doi:10.1542/peds.2008-2598

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., . . . Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication - Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980-989. doi:10.1016/j.jaac.2010.05.017

Michels, N., Vanaelst, B., Stoppie, E., Huybrechts, I., Bammann, K., De Henauw, S. and Sioen, I. (2013). Parental and children's report of emotional problems: Agreement, explanatory factors and event-emotion correlation. *Child and Adolescent Mental Health, 18*(3), 180-186. doi:http://helios.uta.fi:2136/10.1111/j.1475-3588.2012.00672.x

Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M. and Jansen, W. (2012). Validity and reliability of the Strengths and Difficulties Questionnaire in 5-6 year olds: Differences by gender or by parental education? *PloS One, 7*(5), e36805. doi:10.1371/journal.pone.0036805

Miettunen, J., Murray, G. K., Jones, P. B., Mäki, P., Ebeling, H., Taanila, A., . . . Moilanen, I. (2014). Longitudinal associations between childhood and adulthood externalizing and internalizing psychopathology and adolescent substance use. Psychological Medicine, 44(8), 1727-1738. doi:10.1017/S0033291713002328

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737-745. doi:10.1016/j.jclinepi.2010.02.006

Moran, P., Kelesidi, K., Guglani, S., Davidson, S. and Ford, T. (2012). What do parents and carers think about routine outcome measures and their use? A focus group study of CAMHS attenders. *Clinical Child Psychology and Psychiatry, 17*(1), 65-79. doi:10.1177/1359104510391859; 10.1177/1359104510391859

Mukaka, MM. (2012). A guide to appropriate use of Correlation coefficient in medical research. Malawi Medical Journal, 24(3): 69–71. Retrieved from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/

Muris, P., Meesters, C. and van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ)--further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child & Adolescent Psychiatry, 12*(1), 1-8. doi:10.1007/s00787-003-0298-2

Mäkelä, M. and Autti-Rämö, I. (2014). Keuhkotaudeista geeniseulontoihin: Seulontojen historia ja periaatteet Suomessa. In the publication of: Seulonnat Suomessa. Terveydenhuollon seulontojen nykytila ja tulevaisuuden näkymät. *Sosiaali- ja terveysministeriön julkaisuja 2014:17*. Helsinki: Sosiaali- ja terveysministeriö. Retrieved from: http://urn.fi/URN:ISBN:978-952-00-3534-1

Myers, K. and Winters, N. C. (2002). Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(2), 114-122. doi:10.1097/00004583-200202000-00004

National Collaborating Centre for Mental Health (UK). (2010). Antisocial Personality Disorder: Treatment, Management and Prevention. The NICE guideline on treatment, management and prevention. *National Clinical Practice Guideline*, Number 77. Commissioned by the National Institute for Health & Clinical Excellence, published by The British Psychological Society and The Royal

College of Psychiatrists. Retrieved from: http://www.nice.org.uk/nicemedia/live/11765/43046/43046.pdf

Niclasen, J., Skovgaard, A. M., Andersen, A. M., Somhovd, M. J. and Obel, C. (2013). A confirmatory approach to examining the factor structure of the Strengths and Difficulties Questionnaire (SDQ): A large scale cohort study. *Journal of Abnormal Child Psychology, 41*(3), 355-365. doi:10.1007/s10802-012-9683-y; 10.1007/s10802-012-9683-y

Niclasen, J., Teasdale, T. W., Andersen, A. M., Skovgaard, A. M., Elberling, H. and Obel, C. (2012). Psychometric properties of the Danish Strength and Difficulties Questionnaire: The SDQ assessed for more than 70,000 raters in four different cohorts. *PloS One, 7*(2), e32025. doi:10.1371/journal.pone.0032025

Niemelä, S., Sourander, A., Elonheimo, H., Poikolainen, K., Wu, P., Helenius, H., . . . Almqvist, F. (2008). What predicts illicit drug use versus police-registered drug offending? Findings from the Finnish "From a Boy to a Man" birth cohort study. *Social Psychiatry and Psychiatric Epidemiology, 43*(9), 697-704. doi:10.1007/s00127-008-0361-x

Niemelä, S., Sourander, A., Pilowsky, D. J., Susser, E., Helenius, H., Piha, J., . . . Almqvist, F. (2009). Childhood antecedents of being a cigarette smoker in early adulthood. The Finnish 'From a Boy to a Man' Study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 50*(3), 343-351. doi:10.1111/j.1469-7610.2008.01968.x

Obel, C., Heiervang, E., Rodriguez, A., Heyerdahl, S., Smedje, H., Sourander, A., . . . Olsen, J. (2004). The Strengths and Difficulties Questionnaire in the Nordic countries. *European Child & Adolescent Psychiatry, 13 Suppl 2*, II32-9. doi:10.1007/s00787-004-2006-2

Okkonen, K-M. (2014). Suomessa lasten toimeentulo on eurooppalaisittain turvattua. Hyvinvointikatsaus 1/2014. *Hyvinvointitilastollinen aikakausilehti* 25:8-15. Tilastokeskus. Retrieved from: http://tilastokeskus.fi/artikkelit/2014/art_2014-02-26_001.html?s=0

Okkonen, K-M. (2014). Suomessa toimeentulovaikeudet koettelevat erityisesti yksinhuoltajaperheitä. Hyvinvointikatsaus 1/2014. *Hyvinvointitilastollinen aikakausilehti* 25:16-21. Tilastokeskus. Retrieved from: http://tilastokeskus.fi/artikkelit/2014/art_2014-02-26_003.html

Olds, D. L., Eckenrode, J., Henderson, C. R.,Jr, Kitzman, H., Powers, J., Cole, R., . . . Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect. fifteen-year follow-up of a randomized trial. *The Journal of the American Medical Association*, 278(8), 637-643.

Paing, W. W., Weller, R. A., Dixon, T. A. and Weller, E. B. (2010). Face-to-face versus telephone administration of the parent's version of the Children's Interview for Psychiatric Syndromes (P-ChIPS). *Current Psychiatry Reports, 12*(2), 122-126. doi:10.1007/s11920-010-0101-2; 10.1007/s11920-010-0101-2

Pihlakoski, L., Aromaa, M., Sourander, A., Rautava, P., Helenius, H. and Sillanpää, M. (2004). Use of and need for professional help for emotional and behavioral problems among preadolescents: A prospective cohort study of 3- to 12-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(8), 974-983. doi:S0890-8567(09)61361-7 [pii]

Pihlakoski, L., Sourander, A., Aromaa, M., Rautava, P., Helenius, H. and Sillanpää, M. (2006). The continuity of psychopathology from early childhood to preadolescence: A prospective cohort study of 3-12-year-old children. *European Child & Adolescent Psychiatry, 15*(7), 409-417. doi:10.1007/s00787-006-0548-1

Punamäki, R. L., Paavonen, J., Toikka, S. and Solantaus, T. (2013). Effectiveness of preventive family intervention in improving cognitive attributions among children of depressed parents: A randomized study. Journal of Family Psychology : *Journal of the Division of Family Psychology of the American Psychological Association* (Division 43), 27(4), 683-690. doi:10.1037/a0033466; 10.1037/a0033466

Puura, K., Almqvist, F., Piha, J., Moilanen, I., Tamminen, T., Kumpulainen, K., . . . Koivisto, A. (1998). Children wih symptoms of depression - what do the adults see? *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 39(4), 577.

Puura, K., Mäntymaa, M., Luoma, I., Kaukonen, P., Guedeney, A., Salmelin, R. and Tamminen, T. (2010). Infants' social withdrawal symptoms assessed with a direct infant observation method in primary health care. *Infant Behavior & Development, 33*(4), 579-588. doi:10.1016/j.infbeh.2010.07.009; 10.1016/j.infbeh.2010.07.009

Reijneveld, S. A., de Meer, G., Wiefferink, C. H. and Crone, M. R. (2008). Parents' concerns about children are highly prevalent but often not confirmed by child doctors and nurses. *BMC Public Health, 8*, 124-2458-8-124. doi:10.1186/1471-2458-8-124

Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., . . . Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders, 15*(3), 130-142. doi:http://helios.uta.fi:2125/10.1177/10634266070150030101

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., . . . Verhulst, F. C. (2011). International comparisons of behavioral and emotional problems in preschool children: Parents' reports from 24 societies. *Journal of Clinical Child and Adolescent Psychology*, 40(3), 456-467. doi:10.1080/15374416.2011.563472

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L. and Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18*(3), 169-184. doi:http://helios.uta.fi:2172/10.1002/mpr.289

Ringoot, A. P., Jansen, P. W., Steenweg-de Graaff, J., Measelle, J. R., van der Ende, J., Raat, H., . . . Tiemeier, H. (2013). Young children's self-reported emotional, behavioral, and peer problems: The Berkeley Puppet Interview. *Psychological Assessment, 25*(4), 1273-1285. doi:10.1037/a0033976; 10.1037/a0033976

Ristkari, T., Sourander, A., Ronning, J. A., Piha, J., Kumpulainen, K., Tamminen, T., . . . Almqvist, F. (2009). Childhood psychopathology and sense of coherence at age 18: Findings from the Finnish From a Boy to a Man Study. *Social Psychiatry and Psychiatric Epidemiology, 44*(12), 1097-1105. doi:10.1007/s00127-009-0032-6

Rothenberger, A., Becker, A., Erhart, M., Wille, N., Ravens-Sieberer, U. and BELLA study group. (2008). Psychometric properties of the parent Strengths and Difficulties Questionnaire in the general population of German children and adolescents: Results of the BELLA study. *European Child & Adolescent Psychiatry, 17 Suppl 1*, 99-105. doi:10.1007/s00787-008-1011-2

Sanne, B., Torsheim, T., Heiervang, E. and Stormark, K. M. (2009). The Strengths and Difficulties Questionnaire in the Bergen Child Study: A conceptually and methodically motivated structural analysis. *Psychological Assessment, 21*(3), 352-364. doi:10.1037/a0016317

Santalahti, P. (1998). Prenatal screening in Finland: Availability and womens´ decision-making and experiences. Research report 94. Helsinki: *Stakes National Research and Development Centre for Welfare and Health*. University of Turku.

Santalahti, P., Sourander, A. and Piha, J. (2009). Use of children's mental health services. [Lasten mielenterveyspalveluiden kaytto] *Duodecim; Laaketieteellinen Aikakauskirja, 125*(9), 959-964.

Sauni, R., Hakulinen-Viitanen, T., Pelkonen, M., Mäkelä, M., Vartiainen, E., Leino, T. and Puumalainen, T. (2014). Terveystarkastustukset. In the publication of: Seulonnat Suomessa. Terveydenhuollon seulontojen nykytila ja tulevaisuuden näkymät. *Sosiaali- ja terveysministeriön julkaisuja 2014:17*. Helsinki: Sosiaali- ja terveysministeriö. Retrieved from: http://urn.fi/URN:ISBN:978-952-00-3534-1

Sayal, K. and Ford, T. (2010). From epidemiology of child and adolescent mental health problems to service utilization. In M. E. Garralda, J. Raynaud, M. Lanham & A. Jason (Eds.), *Increasing awareness of child and adolescent mental health* (pp. 3-25). Retrieved from http://helios.uta.fi/docview/742983062?accountid=14242:

Seiffge-Krenke, I. and Kollmar, F. (1998). Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behaviour: A longitudinal analysis of parent-adolescent agreement on internalising and externalising problem behaviour. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 39*(5), 687-697.

Sim, F., O'Dowd, J., Thompson, L., Law, J., Macmillan, S., Affleck, M., . . . Wilson, P. (2013). Language and social/emotional problems identified at a universal developmental assessment at 30 months. *BMC Pediatrics, 13*, 206-2431-13-206. doi:10.1186/1471-2431-13-206

Skovgaard, A. M., Houmann, T., Christiansen, E. and Andreasen, A. H. (2005). The reliability of the ICD-10 and the DC 0–3 in an epidemiological sample of children 1½ years of age. *Infant Mental Health Journal,* 26 (5), 470-480. doi:10.1002/imhj.20065

Skovgaard, A. M., Houmann, T., Christiansen, E., Landorph, S., Jorgensen, T., CCC 2000 Study Team, . . . Lichtenberg, A. (2007). The prevalence of mental health problems in children 1(1/2) years of age - the Copenhagen Child Cohort 2000. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 48*(1), 62-70. doi:JCPP1659 [pii]

Skovgaard, A. M., Houmann, T., Landorph, S. L. and Christiansen, E. (2004). Assessment and classification of psychopathology in epidemiological research of children 0-3 years of age: A review of the literature. *European Child & Adolescent Psychiatry, 13*(6), 337-346. doi:10.1007/s00787-004-0393-z

Slade, M., Cahill, S., Kelsey, W., Powell, R., Strathdee, G. and Valiakalayil, A. (2001). Threshold 3: The feasibility of the Threshold Assessment Grid (TAG) for routine assessment of the severity of mental health problems. *Social Psychiatry and Psychiatric Epidemiology, 36*(10), 516-521.

Slade, M., Thornicroft, G. and Glover, G. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology, 34*(5), 243-249.

Smedje, H., Broman, J. E., Hetta, J. and von Knorring, A. L. (1999). Psychometric properties of a Swedish Version of the "Strengths and Difficulties Questionnaire". *European Child & Adolescent Psychiatry, 8*(2), 63-70.

Solantaus, T., Paavonen, E. J., Toikka, S. and Punamäki, R. L. (2010). Preventive interventions in families with parental depression: Children's psychosocial

symptoms and prosocial behaviour. *European Child & Adolescent Psychiatry*, 19(12), 883-892. doi:10.1007/s00787-010-0135-3

SOTKAnet. Statistics and Indicator Bank. © 2005 – 2013 Terveyden ja hyvinvoinnin laitos. Retrieved from: http://uusi.sotkanet.fi/portal/page/portal/etusivu/hakusivu. Accessed at: 26 April 2015.

Sourander, A. (2001). Emotional and behavioural problems in a sample of Finnish three-year-olds. *European Child & Adolescent Psychiatry, 10*(2), 98-104.

Sourander, A., Haavisto, A., Ronning, J. A., Multimäki, P., Parkkola, K., Santalahti, P., . . . Almqvist, F. (2005). Recognition of psychiatric disorders, and self-perceived problems. A follow-up study from age 8 to age 18. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 46*(10), 1124-1134.

Sourander, A., Jensen, P., Davies, M., Niemelä, S., Elonheimo, H., Ristkari, T., . . . Almqvist, F. (2007). Who is at greatest risk of adverse long-term outcomes? the Finnish From a Boy to a Man Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(9), 1148-1161. doi:10.1097/chi.0b013e31809861e9

Sourander, A., Niemelä, S., Santalahti, P., Helenius, H. and Piha, J. (2008). Changes in psychiatric problems and service use among 8-year-old children: A 16-year population-based time-trend study. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(3), 317-327. doi:10.1097/CHI.0b013e318160b98f

Sourander, A., Pihlakoski, L., Aromaa, M., Rautava, P., Helenius, H. and Sillanpää, M. (2006). Early predictors of parent- and self-reported perceived global psychological difficulties among adolescents: A prospective cohort study from age 3 to age 15. *Social Psychiatry and Psychiatric Epidemiology, 41*(3), 173-182. doi:10.1007/s00127-005-0013-3

Stasiak, K., Parkin, A., Seymour, F., Lambie, I., Crengle, S., Pasene-Mizziebo, E. and Merry, S. (2013). Measuring outcome in child and adolescent mental health services: Consumers' views of measures. *Clinical Child Psychology and Psychiatry, 18*(4), 519-535. doi:10.1177/1359104512460860

Statistics Finland. [Suomen tilastokeskus] *Statistical databases*: PX-web statfin in English: Population: Population structure: Population according to age (1-year), sex, marital status and language by area 1990 - 2013. Retrieved from: http://193.166.171.75/database/StatFin/vrm/vaerak/vaerak_fi.asp. Last updated: 18 November 2014. Accessed at: 6 February 2015.

Statistics Finland. [Suomen tilastokeskus] Metadata: Statistical grouping of municipalities 2015. Retrieved from: http://www.stat.fi/meta/luokitukset/kuntaryhmitys/001-2015/index_en.html and personal communication with Information service of the Statistics Finland (miia.huomo@stat.fi at 24 April 2015).

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A. and Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: A review. *Clinical Child and Family Psychology Review, 13*(3), 254-274. doi:10.1007/s10567-010-0071-2

Suomen virallinen tilasto (SVT) (2013). Väestön koulutusrakenne [verkkojulkaisu]. ISSN=1799-4586. Helsinki: Tilastokeskus. Retrieved from: http://www.stat.fi/til/vkour/2013/vkour_2013_2014-11-06_tie_001_fi.html. Accessed at: 12 April 2015.

Sveen, T. H., Berg-Nielsen, T. S., Lydersen, S. and Wichstrom, L. (2013). Detecting psychiatric disorders in preschoolers: Screening with the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(7), 728-736. doi:10.1016/j.jaac.2013.04.010

Tape, T. *Interpreting diagnostic tests. Roc Curves*. University of Nebraska Medical Center. Retrieved from: http://gim.unmc.edu/dxtests/roc3.htm. Accessed at: 30 January 2015.

Taylor, R. (1990). Interpretation of the Correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography January*, 6, 35-39. doi:10.1177/875647939000600106

Theunissen, M. H., Vogels, A. G., de Wolff, M. S. and Reijneveld, S. A. (2013). Characteristics of the Strengths and Difficulties Questionnaire in preschool children. *Pediatrics, 131*(2), e446-54. doi:10.1542/peds.2012-0089

TOIMIA. (2013). *Toimintakyvyn mittaamisen ja arvioinnin kansallinen asiantuntijaverkosto. TOIMIA-tietokanta.* © 2011 - 2014 Terveyden ja Hyvinvoinnin laitos. Retrieved from: http://www.thl.fi/toimia/tietokanta/. Accessed at: 30 January 2015.

Uhari, M., & Nieminen, P. (2001). Kliinisen päätöksenteon työkaluja. *Epidemiologia ja biostatistiikka* (pp. 40-43). Jyväskylä: Kustannus Oy Duodecim.

UNICEF. *Convention on the rights of the child.* Retrieved from: http://www.unicef.org/crc/. Accessed at: 30 January 2015.

Valtioneuvoston asetus neuvolatoiminnasta, koulu- ja opiskeluterveydenhuollosta sekä lasten ja nuorten ehkäisevästä suun terveydenhuollosta 338/2011. Retrieved from: http://www.finlex.fi/fi/laki/alkup/2011/20110338

Valtioneuvoston asetus seulonnoista 339/2011. Retrieved from: http://www.finlex.fi/fi/laki/alkup/2011/20110339

Van Leeuwen K. and Bosmans G., De Medts L:, Braet C. (2006). The Strengths and Difficulties Questionnaire in a community sample of young children in Flanders.*22*(3), 189-197.

Van Roy, B., Groholt, B., Heyerdahl, S. and 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*, 56-244X-10-56. doi:10.1186/1471-244X-10-56; 10.1186/1471-244X-10-56

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. and Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry, 12*(6), 281-289. doi:10.1007/s00787-003-0341-3

Vogels, A. G., Crone, M. R., Hoekstra, F. and Reijneveld, S. A. (2009). Comparing three short questionnaires to detect psychosocial dysfunction among primary school children: A randomized method. *BMC Public Health, 9*, 489-2458-9-489. doi:10.1186/1471-2458-9-489; 10.1186/1471-2458-9-489

Wichstrom, L., Belsky, J., Jozefiak, T., Sourander, A. and Berg-Nielsen, T. S. (2014). Predicting service use for mental health problems among young children. *Pediatrics, 133*(6), 1054-1060. doi:10.1542/peds.2013-3184

Wichstrom, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E. and Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 53*(6), 695-705. doi:10.1111/j.1469-7610.2011.02514.x; 10.1111/j.1469-7610.2011.02514.x

Williamson, A., Redman, S., Dadds, M., Daniels, J., D'Este, C., Raphael, B., . . . Skinner, T. (2010). Acceptability of an emotional and behavioural screening tool for children in Aboriginal community controlled health services in urban NSW. *The Australian and New Zealand Journal of Psychiatry, 44*(10), 894-900. doi:10.3109/00048674.2010.489505

Williford, A., Boulton, A., Noland, B., Little, T. D., Karna, A., and Salmivalli, C. (2012). Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *Journal of Abnormal Child Psychology*, 40(2), 289-300. doi:10.1007/s10802-011-9551-1

Wilson, J. and Jungner, G. (1968). Principles and practice of screening disease. Public Health Papers, 34. World Health Organisation, Geneva.

Woerner, W., Fleitlich-Bilyk, B., Martinussen, R., Fletcher, J., Cucchiaro, G., Dalgalarrondo, P., . . . Tannock, R. (2004). The Strengths and Difficulties Questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. *European Child & Adolescent Psychiatry, 13 Suppl 2*, II47-54. doi:10.1007/s00787-004-2008-0

von Stumm, S., Deary, I. J., Kivimäki, M., Jokela, M., Clark, H. and Batty, G. D. (2011). Childhood behavior problems and health at midlife: 35-year follow-up of a Scottish Birth Cohort. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 52(9), 992-1001. doi:10.1111/j.1469-7610.2011.02373.x

World Health Organisation (Ed.). (1994). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva, Switzerland.

World Health Organization. (2004). *Prevention of mental disorders: Effective interventions and policy option : Summary report / a report of the World Health Organization Department of Mental Health and Substance Abuse*. In collaboration with the Prevention Research Centre of the Universities of Nijmegen and Maastricht. Editors: Clemens

Hosman, C., Jané-Llopis, E., Saxena, S. Retrieved from: http://www.who.int/mental_health/media/en/545.pdf

Youthinmind a. *DAWBA, information for researchers and clinicians about the Development and Well-Being Assessment.* Retrieved from: http://www.dawba.info/. Last updated: 23 November 2014. Accessed at: 8 February 2015.

Youthinmind b. *SDQ, information for researchers and professionals about the Strengths and Difficulties Questionnaires.* Retrieved from: http://www.sdqinfo.com/. Last updated: 28 May 2014. Accessed at: 8 February 2015.

Youthinmind c. *SDQ scoring site* Retrieved from: http://www.sdqscore.org. Accessed at: 8 February 2015.

Zero to Three. (1994). *Diagnostic classification of mental health and developmental disorders of infancy and early childhood* (revised edition 2005 ed.). Washington, DC: National Center for Infants, Toddlers and FamiliesPrograms.

Zwaanswijk, M., Verhaak, P. F., Bensing, J. M., van der Ende, J. and Verhulst, F. C. (2003). Help seeking for emotional and behavioural problems in children and adolescents: A review of recent literature. *European Child & Adolescent Psychiatry*, 12(4), 153-161. doi:10.1007/s00787-003-0322-6

Errata

Publications missing from the reference list

De Los Reyes, A. and Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131(4), 483-509. doi: http://helios.uta.fi:2136/10.1037/0033-2909.131.4.483

Hakulinen-Viitanen, T., Pelkonen, M., Saaristo, V., Hastrup, A. and Rimpelä, M. (2008*). Äitiys- ja lastenneuvolatoiminta 2007. Tulokset ja seurannan kehittäminen.* (Stakes 2008, raportti 21.). Valopaino Oy, Helsinki: Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus. Retrieved from: https://www.julkari.fi/bitstream/handle/10024/77329/R21-2008-VERKKO.pdf?sequence=1

Hiltunen, H. (edit.) (2007). *Suomalainen lapsi 2007*. Helsinki: Tilastokeskus.

Ialongo, N. S., Edelsohn, G. and Kellam, S. G. (2001). A further look at the prognostic power of young children's reports of depressed mood and feelings. *Child Development*, 72(3), 736-747.

Kaukonen, P., Salmelin, R. K., Luoma, I., Puura, K., Rutanen, M., Pukuri, T. and Tamminen, T. (2010). Child psychiatry in the Finnish health care reform: National criteria for treatment access. *Health Policy* (Amsterdam, Netherlands), 96(1), 20-27. doi:10.1016/j.healthpol.2009.12.009

Luby, J.L., Belden, A., Sullivan, J. and Spitznagel, E. (2007). Preschoolers' contribution to their diagnosis of depression and anxiety: Uses and limitations of young child self-report of symptoms. *Child Psychiatry & Human Development*, 38(4), 321-338.

Mäki, P., Wikström, K., Hakulinen-Viitanen, T. and Laatikainen, T. (2011). *Terveystarkastukset lastenneuvolassa ja kouluterveydenhuollossa.* Menetelmäkäsikirja (Opas 14 ed.). Juvenes Print, Tampere: Terveyden ja hyvinvoinnin laitos.

Rimpelä, M., Rigoff, A., Wiss, K. and Hakulinen-Viitanen, T. (2006). Seulontatutkimukset 3-7-vuotiaiden lasten terveystarkastuksissa. Stakesin työpapereita 17/2006. Retrieved from: https://www.julkari.fi/bitstream/handle/10024/76176/T17-2006-VERKKO.pdf?sequence=1

11 Appendices

1. Cover letters for parents (A,B)/Study information for parents
2. Informed consent
3. Participation information and the nurse's one-question screen
4. Socio-demographic form, including the parent's one-question screen
5. The Finnish version of the Strength and Difficulties Questionnaire (SDQ-Fin), parent and teacher report
6. Instructions for scoring the SDQ-Fin by hand, the Finnish version (Vahvuudet ja Vaikeudet -kyselyn (SDQ) pisteytys vanhemman ja opettajan vastaamana)
7. The child's self-evaluation enquiry of emotional well-being
8. The feedback questionnaires on the SDQ-Fin and on the child's self-evaluation enquiry of emotional well-being for parents and public health nurses
9. The feedback questionnaire on the SDQ-Fin for preschool and school teachers
10. Frequency distributions of the Strengths and Difficulties Questionnaire (SDQ) for Finnish 4–9 year olds (Tables 1-14)
11. The means and standard deviations of the SDQ-Fin scores
12. The ROC curves for the higher cut-offs of the SDQ-Fin total score

Appendix 1A.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

LASTENNEUVOLAN ASIAKASTIEDOTE

Lastenpsykiatrian klinikka

11.1.2015

Hyvät vanhemmat,

Pyydämme Teitä ja lastanne osallistumaan Lasten mielenterveystyön kehittäminen -hankkeen yhteydessä toteutettavaan tutkimukseen. Tutkimme ja kehitämme neuvolatyöhön sopivia menetelmiä lapsen henkisen hyvinvoinnin, ihmissuhdetaitojen ja näihin liittyvän tuen tarpeen arvioimiseksi. Teemme yhteistyötä vanhempien ja päivähoidon kanssa. Pirkanmaan sairaanhoitopiirin eettinen toimikunta on antanut tutkimuksesta myönteisen lausunnon.

Neuvolan 5- 6 -vuotistarkastuksen yhteydessä Teidän, lapsenne, päivähoidon ja neuvolan työntekijöiden vastattavaksi tulee lapsenne hyvinvointiin ja kehitykseen liittyviä kyselylomakkeita. Ennen neuvolakäyntiä pyydämme Teitä täyttämään oheisen suostumuskaavakkeen, taustatietolomakkeen, Vahvuudet ja vaikeudet -kyselyn (molemmille vanhemmille omakysely) sekä palautelomakkeen. Arvioimme näiden lomakkeiden täyttöön kuluvan yhteensä noin 20 minuuttia. Lisäksi pyydämme Teitä toimittamaan päivähoitoon oheisessa kirjekuoressa olevat kyselylomakkeet.

Pyydämme Teitä ystävällisesti palauttamaan neuvolakäynnillä terveydenhoitajalle kaikki kotona ja päivähoidossa täytetyt lomakkeet. Terveydenhoitajan käynnillä lapselta kysytään hänen arviotaan omasta hyvinvoinnistaan kuva-astekkojen avulla. Lääkäri täyttää neuvolatarkastuksessa kyselyn, jonka avulla kootaan laajemmin tietoa lapsen kehityksestä ja hyvinvoinnista, toimintakyvystä ja mahdollisesta tuen tarpeesta. Käynnin jälkeen Teiltä pyydetään palautetta tästä lääkärin käyttämästä kyselystä. Haluamme varmentaa neuvolatutkimuksessa testattavien menetelmien luotettavuuden haastattelemalla Teitä puhelimitse lapsenne kehitykseen ja hyvinvointiin liittyen. Haastattelu vie aikaanne puolesta tunnista tuntiin ja toteutetaan neuvolakäynnin jälkeen Teidän kanssanne erikseen sovittavana ajankohtana.

Tutkimukseen osallistuminen on vapaaehtoista ja Teillä on oikeus missä tahansa vaiheessa kieltäytyä osallistumisesta, syytä siihen ilmoittamatta. Kieltäytymisenne ei vaikuta Teidän tai lapsenne tarvitsemiin terveyspalveluihin nyt tai myöhemmin. Tutkimuksen tuloksia käsitellään luottamuksellisesti ja nimettöminä, siten ettei yksittäistä vastaajaa voida tunnistaa.

Mikäli Teillä on kysyttävää tai haluatte lisätietoja, vastaamme mielellämme.

Ystävällisesti

Sari Miettinen
hankesairaanhoitaja
sari.miettinen@pshp.fi
puh. 050- 537 6543

Anne-Mari Borg
tutkiva lääkäri, lastenpsykiatri
Tampereen yliopisto, Lääketieteen laitos
anne-mari.borg@uta.fi
puh. 050-372 8362

Pälvi Kaukonen
ylilääkäri
TAYS lastenpsykiatrian klinikka
palvi.kaukonen@pshp.fi

Appendix 1B.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

KOULUTERVEYDEN ASIAKASTIEDOTE

Lastenpsykiatrian klinikka

11.1.2015

Hyvät vanhemmat,

Pyydämme Teitä ja lastanne osallistumaan Lasten mielenterveystyön kehittäminen -hankkeen yhteydessä toteutettavaan tutkimukseen. Tutkimme ja kehitämme kouluterveydenhuoltoon sopivia menetelmiä lapsen psykososiaalisen hyvinvoinnin ja tuen tarpeen arvioimiseksi yhteistyössä vanhempien ja opettajan kanssa. (Pirkanmaan sairaanhoitopiirin eettinen toimikunta on antanut tutkimuksesta myönteisen lausunnon.)

Kouluterveydenhuollon 7-vuotiaiden, koulunsa aloittaneiden lasten määräaikaistarkastuksen yhteydessä Teidän, lapsenne, opettajan, terveydenhoitajan ja lääkärin vastattavaksi tulee lapsenne hyvinvointiin ja kehitykseen liittyviä kyselylomakkeita. Pyydämme Teitä täyttämään oheisen suostumuskaavakkeen, taustatietolomakkeen, Vahvuudet ja vaikeudet -kyselyn (molemmille vanhemmille oma kysely ?) sekä palautelomakkeen. Yhteensä arvioimme näiden lomakkeiden täyttöön kuluvan Teiltä noin 20 minuuttia. Osallistuessanne hankkeeseen ja tutkimukseen myös luokanvalvoja täyttää lapsestanne Vahvuudet ja vaikeudet -kyselyn sekä palautekyselyn.

Lapsenne terveystarkastuksessa lääkäri täyttää kyselyn, jonka avulla kootaan laajemmin tietoa lapsen kehityksestä ja hyvinvoinnista, toimintakyvystä ja mahdollisesta tuen tarpeesta. Käynnin jälkeen Teiltä pyydetään palautetta tästä lääkärin käyttämästä kyselystä. (Käynnillä lapselta itseltään kysytään hänen arviotaan hyvinvoinnistaan piirtämistä avuksi käyttäen?). Haluamme varmentaa tutkimuksessa testattavien menetelmien luotettavuuden haastattelemalla Teitä puhelimitse lapsenne kehitykseen ja hyvinvointiin liittyen. Haastattelu vie aikaanne puolesta tunnista tuntiin ja toteutetaan Teidän kanssanne erikseen sovittavana ajankohtana.

Pyydämme ystävällisesti palauttamaan kouluterveydenhuollon käynnillä kaikki teille kotiin lähetetyt lomakkeet. Tutkimukseen osallistuminen on vapaaehtoista ja Teillä on oikeus missä vaiheessa tahansa kieltäytyä osallistumisesta, syytä siihen ilmoittamatta. Kieltäytymisenne ei vaikuta Teidän tai lapsenne tarvitsemien terveyspalveluiden saatavuuteen nyt tai myöhemmin. Tutkimuksen tuloksia käsitellään luottamuksellisesti ja nimettöminä.

Mikäli Teillä on kysyttävää tai haluatte lisätietoja, vastaamme mielellämme.

Ystävällisesti

Sari Miettinen
hankesairaanhoitaja
sari.miettinen@pshp.fi
puh.

Anne-Mari Borg
tutkiva lääkäri, lastenpsykiatri
Tampereen yliopisto, Lääketieteen laitos
anne-mari.borg@uta.fi
puh

Pälvi Kaukonen
TAYS lastenpsykiatrian ylilääkäri
palvi.kaukonen@pshp.fi

Appendix 2.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

SUOSTUMUS

LASTEN MIELENTERVEYSTYÖN KEHITTÄMINEN -hanke ja sen yhteydessä toteutuva tutkimus

Olen saanut sekä kirjallista että halutessani myös suullista tietoa lapsen henkisen hyvinvoinnin, ihmissuhdetaitojen ja näihin liittyvän mahdollisen tuen tarpeen arviointimenetelmien tutkimuksesta perusterveydenhuollossa ja mahdollisuuden esittää siitä tutkijoille kysymyksiä. Lapselleni annetaan tutkimuksesta suullista tietoa hänen vastatessaan Lapsen oma arvio -lomakkeen kysymyksiin neuvolakäynnillä ja hänellä on halutessaan mahdollisuus esittää tutkijoille kysymyksiä. Täyttämällä tutkimuslomakkeiston annan samalla suostumukseni neuvolalle ja päivähoidolle antaa tietoja lapsestani.

Ymmärrän, että tutkimukseen osallistuminen on vapaaehtoista ja että minulla ja lapsellani on oikeus kieltäytyä siitä milloin tahansa syytä ilmoittamatta. Ymmärrän myös, että tiedot käsitellään luottamuksellisesti.

paikka	aika	paikka	aika

Suostun osallistumaan tutkimukseen:

vanhemman/ huoltajan allekirjoitus

nimen selvennys

vanhemman/ huoltajan syntymäaika

lapsen allekirjoitus

lapsen nimi

lapsen syntymäaika

vanhemman/ huoltajan osoite

vanhemman/ huoltajan puhelinnumero

Suostumuksen vastaanottaja:

tutkijan allekirjoitus

nimen selvennys

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

SUOSTUMUS

LASTEN MIELENTERVEYSTYÖN KEHITTÄMINEN -hanke ja sen yhteydessä toteutuva tutkimus

Olen saanut sekä kirjallista että halutessani myös suullista tietoa lapsen henkisen hyvinvoinnin, ihmissuhdetaitojen ja näihin liittyvän mahdollisen tuen tarpeen arviointimenetelmien tutkimuksesta perusterveydenhuollossa ja mahdollisuuden esittää siitä tutkijoille kysymyksiä. Lapselleni annetaan tutkimuksesta suullista tietoa hänen vastatessaan Lapsen oma arvio -lomakkeen kysymyksiin kouluterveydenhuollon käynnillä ja hänellä on halutessaan mahdollisuus esittää tutkijoille kysymyksiä. Täyttämällä tutkimuslomakkeiston annan samalla suostumukseni kouluterveydenhuollolle ja opettajalle antaa tietoja lapsestani.

Ymmärrän, että tutkimukseen osallistuminen on vapaaehtoista ja että minulla ja lapsellani on oikeus kieltäytyä siitä milloin tahansa syytä ilmoittamatta. Ymmärrän myös, että tiedot käsitellään luottamuksellisesti.

paikka	aika	paikka	aika

Suostun osallistumaan tutkimukseen:

vanhemman/ huoltajan allekirjoitus

nimen selvennys

vanhemman/ huoltajan syntymäaika

Suostumuksen vastaanottaja:

tutkijan allekirjoitus

nimen selvennys

lapsen allekirjoitus

lapsen nimi

lapsen syntymäaika

vanhemman/ huoltajan osoite

vanhemman/ huoltajan puhelinnumero

Appendix 3.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä
Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

TERVEYDENHOITAJAN SEURANTAKAAVAKE (5-6-VUOTIAAT)

EI OSALLISTUJAT (Täytetään kaikista ikäryhmään kuuluvista lapsista, jotka eivät osallistu hankkeeseen)

Kunta	
Lapsen ikä (vuosina)	**Lapsen sukupuoli** 1. Poika 2. Tyttö
Syy, miksi ei osallistu, mikäli tiedossa sitä erikseen kysymättä Perheellä on oikeus kieltäytyä hankkeeseen osallistumisesta syytä siihen ilmoittamatta. Tähän kirjataan kieltäytymisen syy vain, mikäli perhe on sen oma-aloitteisesti kertonut.	**Terveydenhoitajan huolikysymys** Yleisesti ottaen, onko tällä lapsella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden kanssa toimeen tuleminen? (ympyröikää sopivin vaihtoehto) 1. Ei 2. Kyllä – vähäisiä vaikeuksia 3. Kyllä – selviä vaikeuksia 4. Kyllä – huomattavia vaikeuksia 5. En osaa sanoa

OSALLISTUJAT (Täytetään kaikkien hankkeeseen osallistujien kohdalta)

Kunta ja lastenneuvolan toimipiste	
Lapsen nimi ja syntymäaika (ppkkvvvv esim. 01012002)	
Terveystarkastuksen päivämäärä (ppkkvvvv)	**Lääkärintarkastuksen päivämäärä** (ppkkvvvv)
Vanhempien yhteystiedot (nimi, osoite ja puh)	**Terveydenhoitajan huolikysymys** Yleisesti ottaen, onko tällä lapsella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden kanssa toimeen tuleminen? (ympyröikää sopivin vaihtoehto) 1. Ei 2. Kyllä – vähäisiä vaikeuksia 3. Kyllä – selviä vaikeuksia 4. Kyllä – huomattavia vaikeuksia 5. En osaa sanoa
Päivähoidon yhteystiedot (nimi ja toimipisteen osoite)	

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä
Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

TERVEYDENHOITAJAN SEURANTAKAAVAKE (7-8-VUOTIAAT)

EI OSALLISTUJAT (Täytetään kaikista ikäryhmään kuuluvista lapsista, jotka eivät osallistu hankkeeseen)

Kunta	
Lapsen ikä (vuosina)	**Lapsen sukupuoli** 1. Poika 2. Tyttö
Syy, miksi ei osallistu, mikäli tiedossa sitä erikseen kysymättä Perheellä on oikeus kieltäytyä hankkeeseen osallistumisesta syytä siihen ilmoittamatta. Tähän kirjataan kieltäytymisen syy vain, mikäli perhe on sen oma-aloitteisesti kertonut.	**Terveydenhoitajan huolikysymys** Yleisesti ottaen, onko tällä lapsella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden kanssa toimeen tuleminen? (ympyröikää sopivin vaihtoehto) 1. Ei 2. Kyllä – vähäisiä vaikeuksia 3. Kyllä – selviä vaikeuksia 4. Kyllä – huomattavia vaikeuksia 5. En osaa sanoa

OSALLISTUJAT (Täytetään kaikkien hankkeeseen osallistujien kohdalta)

Kunta ja kouluterveydenhuollon toimipiste	
Lapsen nimi ja syntymäaika (ppkkvvvv esim. 01012002)	
Terveystarkastuksen päivämäärä (ppkkvvvv)	**Lääkärintarkastuksen päivämäärä** (ppkkvvvv)
Vanhempien yhteystiedot (nimi, osoite ja puh)	**Terveydenhoitajan huolikysymys** Yleisesti ottaen, onko tällä lapsella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden kanssa toimeen tuleminen? (ympyröikää sopivin vaihtoehto) 1. Ei 2. Kyllä – vähäisiä vaikeuksia 3. Kyllä – selviä vaikeuksia 4. Kyllä – huomattavia vaikeuksia 5. En osaa sanoa
Opettajan yhteystiedot (nimi ja toimipisteen osoite)	

Appendix 4.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Esitietokysymykset vanhemmille

Pyydämme jompaa kumpaa vanhemmista täyttämään tämän esitietolomakkeen. Vastatkaa kysymyksiin rengastamalla oikeaa vaihtoehtoa vastaava numero tai kirjoittamalla vastauksenne sille varattuun tilaan.

1. Lapsen syntymäaika _____ (ppkkvvvv, esim. 03062001)

2. Lapsen sukupuoli
   1 poika
   2 tyttö

3. Alle kouluikäisen lapsen hoitomuoto
   1 kotihoidossa
   2 perhepäivähoidossa
   3 päiväkodissa
   4 jokin muu, mikä? _____

4. Lapsen kanssa asuvat vanhemmat ovat
   1 biologiset vanhemmat
   2 biologinen äiti ja kasvatusisä (esim. avopuoliso)
   3 biologinen isä ja kasvatusäiti (esim. avopuoliso)
   4 biologinen äiti yksin
   5 biologinen isä yksin
   6 adoptiovanhemmat
   7 sijaisvanhemmat
   8 muu, kuka _____

5. Montako lasta perheessänne asuu kaikkiaan? _____

6. Vanhempien koulutustaso? Valitkaa molemmilta sarakkeilta soveltuva, peruskoulutusta ja ammatillista koulutusta vastaava vaihtoehto.

	Äiti tai äidin asemassa oleva	Isä tai isän asemassa oleva
Vanhempien peruskoulutus		
kansakoulu	1	1
perus- tai keskikoulu	2	2
ylioppilas	3	3
koulu keskeytynyt	4	4
Vanhempien ammatillinen koulutus		
ei ammattikoulutusta	1	1
ammattikurssi tai –kursseja	2	2
koulutasoinen ammatillinen koulutus (ammatti- tai kauppakoulu)	3	3
opistoaste tai ammattikorkeakoulu	4	4
yliopistokoulutus	5	5
muu amm. koulutus, mikä?	6	6

_____

7. Mikä seuraavista vaihtoehdoista kuvaa mielestänne parhaiten perheenne nykyistä taloudellista tilannetta?

1 Täysin riittävä toimeentulo, ei käytännössä toimeentulohuolia tai -ongelmia
2 Kutakuinkin riittävä toimeentulo, vain harvoin toimeentulohuolia tai -ongelmia
3 Jokseenkin riittämätön toimeentulo, melko usein toimeentulohuolia tai -ongelmia
4 Ei riittävää toimeentuloa, hyvin usein toimeentulohuolia tai -ongelmia
5 En halua sanoa

8. Onko lapsenne ruumiillisessa terveydentilassa erityistä mainittavaa? (esim. pitkäaikaissairaus, toistuvia sairaalahoitoja)

1 Ei
2 Kyllä, millaisia? _____

_____

_____

9. Onko lapsellanne ruumiillisia oireita, joihin ei ole löytynyt lääketieteellistä selitystä?

1 Ei
2 Kyllä, millaisia? _____

_____

_____

10. Onko lapsellanne mielestänne vaikeuksia tunnetilojen ja käyttäytymisen säätelyssä, keskittymisessä tai sosiaalisissa taidoissa?

1 Ei lainkaan vaikeuksia
2 Ei kovinkaan paljon vaikeuksia
3 Melko paljon vaikeuksia
4 Hyvin paljon vaikeuksia

11. Mikä seuraavista vaihtoehdoista kuvaa parhaiten omaa käsitystänne lapsenne hyvinvoinnista ja pärjäämisestä, kun ajatellette hänen kykyjään tunnetilojen ja käyttäytymisen säätelyssä, keskittymisessä tai sosiaalisissa taidoissa?

1 En ole lainkaan huolissani asiasta
2 En ole kovinkaan huolissani asiasta
3 Olen melko huolissani asiasta
4 Olen hyvin huolissani asiasta

12. Tarvitseeko lapsenne mielestänne erityistä tukea perheenne ulkopuolelta psyykkisen hyvinvointinsa ja pärjäämisensä tueksi?

1 Ei tarvitse
2 Luultavasti ei tarvitse
3 Luultavasti tarvitsee
4 Tarvitsee tai on jo hakeutumassa tai päässyt jonkin tuen piiriin

Appendix5.

Vahvuuksien ja Vaikeuksien Kyselylomake (SDQ-Fin) V 4-16

Pyytäisimme teitä ystävällisesti täyttämään tämän kyselylomakkeen koskien mainitun lapsenne käyttäytymistä viimeisen 6 kk:n (tai kuluvan kouluvuoden) aikana merkitsemällä rasti yhteen kolmesta annetusta vaihtoehdosta: "Ei Päde", "Pätee Jonkinverran", "Pätee Varmasti". On hyvin tärkeää, että vastaatte jokaiseen kohtaan parhaan kykynne mukaan siitäkin huolimatta, että aina ette tunne olevanne asiasta täysin varma - tai, että kysymys kuulostaa älyttömältä.

Lapsen / Nuoren Nimi ..... Tyttö / Poika

Syntymäaika .....

	Ei Päde	Pätee Jonkinverran	Pätee Varmasti
Ottaa muiden tunteet huomioon	☐	☐	☐
Levoton, yliaktiivinen, ei pysty olemaan kauan hiljaa paikoillaan	☐	☐	☐
Valittaa usein päänsärkyä, vatsakipua tai pahoinvointia	☐	☐	☐
Jakaa auliisti tavaroitaan (karkkeja, leluja, värikyniä jne) muiden lasten kanssa	☐	☐	☐
Hänellä on usein kiukunpuuskia, tai hän kiivastuu helposti	☐	☐	☐
Ei näytä kaipaavan seuraa, leikkii usein itsekseen	☐	☐	☐
On yleensä tottelevainen, tavallisesti tekee niinkuin aikuinen käskee	☐	☐	☐
Hänellä on monia huolia, näyttää usein huolestuneelta	☐	☐	☐
Tarjoutuu auttamaan, jos joku loukkaa itsensä, on pahoilla mielin tai huonovointinen	☐	☐	☐
Jatkuvasti hypistelemässä jotakin tai kiemurtelee paikoillaan	☐	☐	☐
Hänellä on ainakin yksi hyvä ystävä	☐	☐	☐
Usein tappelee toisten lasten kanssa tai kiusaa muita	☐	☐	☐
Usein onneton, mieli maassa tai itkuinen	☐	☐	☐
Yleensä muiden lasten suosiossa	☐	☐	☐
Helposti häiriintyvä, mielenkiinto harhailee	☐	☐	☐
Uusissa tilanteissa pelokas tai aikuiseen takertuva, vailla itseluottamusta	☐	☐	☐
Kiltti nuorempiaan kohtaan	☐	☐	☐
Valehtelee tai petkuttaa usein	☐	☐	☐
Muiden lasten silmätikku tai kiusaamisen kohde	☐	☐	☐
Tarjoutuu usein auttamaan muita (vanhempiaan, opettajia, muita lapsia)	☐	☐	☐
Harkitsee tilanteen ennen kuin toimii	☐	☐	☐
Varastaa kotoa, koulusta tai muualta	☐	☐	☐
Tulee paremmin toimeen aikuisten kuin toisten lasten kanssa	☐	☐	☐
Kärsii monista peloista, usein peloissaan	☐	☐	☐
Saattaa tehtävät loppuun, hyvin pitkäjänteinen	☐	☐	☐

Onko teillä hänestä muita kommentteja tai huomautuksia ?

Kääntöpuolella muutama lisäkysymys - Olkaa Hyvä !

Yleisesti ottaen, onko lapsellanne/nuorella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden ihmisten kanssa toimeentuleminen?

Ei	Kyllä - vähäisiä vaikeuksia	Kyllä - selviä vaikeuksia	Kyllä - huomattavia vaikeuksia
☐	☐	☐	☐

Jos vastasitte "Kyllä", olkaa hyvä ja vastatkaa seuraaviin näitä vaikeuksia koskeviin kysymyksiin:

- Vaikeuksien kesto:

Alle 1kk	1 - 5 kk	6 - 12 kk	Yli 1 v.
☐	☐	☐	☐

- Huolestuttavatko nämä vaikeudet lasta/nuorta itseään ?

Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
☐	☐	☐	☐

- Häiritsevätkö nämä vaikeudet lapsen/nuoren elämää seuraavilla alueilla?

	Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
KOTIELÄMÄ	☐	☐	☐	☐
TOVERISUHTEET	☐	☐	☐	☐
KOULUOPPIMINEN	☐	☐	☐	☐
HARRRASTUKSET	☐	☐	☐	☐

- Rasittavatko nämä vaikeudet teitä tai kenties koko perhettä ?

Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
☐	☐	☐	☐

Lomakkeen täyttäjän nimikirjoitus ..... Päivämäärä .....

Äiti / Isä / Joku muu (Olkaa hyvä ja selventäkää:)

Parhaat kiitokset avustanne !

Vahvuuksien ja Vaikeuksien Kyselylomake (SDQ-Fin) O 4-16

Pyytäisimme teitä ystävällisesti täyttämään tämän kyselylomakkeen koskien mainitun lapsen/nuoren käyttäytymistä viimeisen 6 kk:n (tai kuluvan kouluvuoden) aikana merkitsemällä rasti yhteen kolmesta annetusta vaihtoehdosta: "Ei Päde", "Pätee Jonkinverran", "Pätee Varmasti". On hyvin tärkeää, että vastaatte jokaiseen kohtaan parhaan kykynne mukaan siitäkin huolimatta, että aina ette tunne olevanne asiasta täysin varma - tai, että kysymys kuulostaa älyttömältä.

Lapsen / Nuoren Nimi ..... Tyttö / Poika

Syntymäaika .....

	Ei Päde	Pätee Jonkinverran	Pätee Varmasti
Ottaa muiden tunteet huomioon	☐	☐	☐
Levoton, yliaktiivinen, ei pysty olemaan kauan hiljaa paikoillaan	☐	☐	☐
Valittaa usein päänsärkyä, vatsakipua tai pahoinvointia	☐	☐	☐
Jakaa auliisti tavaroitaan (karkkeja, leluja, värikyniä jne) muiden lasten kanssa	☐	☐	☐
Hänellä on usein kiukunpuuskia, tai hän kiivastuu helposti	☐	☐	☐
Ei näytä kaipaavan seuraa, leikkii usein itsekseen	☐	☐	☐
On yleensä tottelevainen, tavallisesti tekee niinkuin aikuinen käskee	☐	☐	☐
Hänellä on monia huolia, näyttää usein huolestuneelta	☐	☐	☐
Tarjoutuu auttamaan, jos joku loukkaa itsensä, on pahoilla mielin tai huonovointinen	☐	☐	☐
Jatkuvasti hypistelemässä jotakin tai kiemurtelee paikoillaan	☐	☐	☐
Hänellä on ainakin yksi hyvä ystävä	☐	☐	☐
Usein tappelee toisten lasten kanssa tai kiusaa muita	☐	☐	☐
Usein onneton, mieli maassa tai itkuinen	☐	☐	☐
Yleensä muiden lasten suosiossa	☐	☐	☐
Helposti häiriintyvä, mielenkiinto harhailee	☐	☐	☐
Uusissa tilanteissa pelokas tai aikuiseen takertuva, vailla itseluottamusta	☐	☐	☐
Kiltti nuorempiaan kohtaan	☐	☐	☐
Valehtelee tai petkuttaa usein	☐	☐	☐
Muiden lasten silmätikku tai kiusaamisen kohde	☐	☐	☐
Tarjoutuu usein auttamaan muita (vanhempiaan, opettajia, muita lapsia)	☐	☐	☐
Harkitsee tilanteen ennen kuin toimii	☐	☐	☐
Varastaa kotoa, koulusta tai muualta	☐	☐	☐
Tulee paremmin toimeen aikuisten kuin toisten lasten kanssa	☐	☐	☐
Kärsii monista peloista, usein peloissaan	☐	☐	☐
Saattaa tehtävät loppuun, hyvin pitkäjänteinen	☐	☐	☐

Onko teillä hänestä muita kommentteja tai huomautuksia ?

Kääntöpuolella muutama lisäkysymys - Olkaa Hyvä !

Yleisesti ottaen, onko tällä lapsella/nuorella mielestänne vaikeuksia yhdellä tai useammalla seuraavista alueista: tunnetilat, keskittyminen, käyttäytyminen tai muiden ihmisten kanssa toimeentuleminen ?

Ei	Kyllä - vähäisiä vaikeuksia	Kyllä - selviä vaikeuksia	Kyllä - huomattavia vaikeuksia
☐	☐	☐	☐

Jos vastasitte "Kyllä", olkaa hyvä ja vastatkaa seuraaviin näitä vaikeuksia koskeviin kysymyksiin:

- Vaikeuksien kesto:

Alle 1kk	1 - 5 kk	6 - 12 kk	Yli 1 v.
☐	☐	☐	☐

- Huolestuttavatko nämä vaikeudet lasta/nuorta itseään ?

Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
☐	☐	☐	☐

- Häiritsevätkö nämä vaikeudet lapsen/nuoren elämää seuraavilla alueilla ?

	Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
TOVERISUHTEET	☐	☐	☐	☐
KOULUOPPIMINEN	☐	☐	☐	☐

- Rasittavatko nämä vaikeudet teitä tai kenties koko luokkaa ?

Eivät ollenkaan	Vain vähän	Aika paljon	Hyvin paljon
☐	☐	☐	☐

Lomakkeen täyttäjän nimikirjoitus ..... Päivämäärä .....

Luokanopettaja / Luokanvalvoja / Joku muu (Olkaa hyvä ja selventäkää:)

Parhaat kiitokset avustanne !

Appendix 6.

Vahvuudet ja Vaikeudet kyselyn (SDQ) pisteytys vanhemman ja opettajan vastaamana

SDQ sisältää 25 väittämää, jotka muodostavat 5 kyselyosiota viiden väitteen sarjoina. Yleensä on helpointa laskea ensin erikseen kaikkien viiden kyselyosion pisteet ennen kokonaispistemäärän laskemista. "Pätee jonkin verran" vastaus antaa aina yhden pisteen, mutta "Ei päde" ja "Pätee varmasti" vastausten pisteytys vaihtelee väittämästä riippuen, kuten alla kussakin kyselyosiossa on nähtävissä. Kunkin kyselyosion pistemäärä voi vaihdella 0 ja 10 pisteen välillä, jos kaikkiin viiteen väittämään on vastattu. Kyselyosion pistemäärä on laskettavissa, jos vähintään kolmeen väitteeseen viidestä on vastattu.

Tunneoireiden kyselyosio	**Ei päde**	**Pätee jonkin verran**	**Pätee varmasti**
Valittaa usein päänsärkyä, vatsakipua tai pahoinvointia	0	1	2
Hänellä on monia huolia, näyttää usein huolestuneelta	0	1	2
Usein onneton, mieli maassa tai itkuinen	0	1	2
Uusissa tilanteissa pelokas tai aikuiseen takertuva vailla…	0	1	2
Karsii monista peloista, usein peloissaan	0	1	2
Käytösoireiden kyselyosio/skaala	**Ei päde**	**Pätee jonkin verran**	**Pätee varmasti**
Hänellä on usein kiukunpuuskia, tai hän kiivastuu helposti	0	1	2
On yleensä tottelevainen, tavallisesti tekee niin kuin…	2	1	0
Usein tappelee toisten kanssa tai kiusaa muita	0	1	2
Valehtelee tai petkuttaa usein	0	1	2
Varastelee kotoa, koulusta tai muualta	0	1	2
Yliaktiivisuuden kyselyosio	**Ei päde**	**Pätee jonkin verran**	**Pätee varmasti**
Levoton, yliaktiivinen, ei pysty olemaan kauan hiljaa…	0	1	2
Jatkuvasti hypistelemässä jotakin tai kiemurtelee paikoillaan	0	1	2
Helposti häiriintyvä, mielenkiinto harhailee	0	1	2
Harkitsee ennen kuin toimii	2	1	0
Saattaa tehtävät loppuun, hyvin pitkäjänteinen	2	1	0
Kaverisuhteiden ongelmien kyselyosio	**Ei päde**	**Pätee jonkin verran**	**Pätee varmasti**
Ei näytä kaipaavan seuraa, leikkii usein itsekseen	0	1	2
Hänellä on ainakin yksi hyvä ystävä	2	1	0
Yleensä muiden lasten suosiossa	2	1	0
Muiden lasten silmätikku tai kiusaamisen kohde	0	1	2
Tulee paremmin toimeen aikuisten kuin toisten lasten kanssa	0	1	2
Prososiaalisen käytöksen kyselyosio	**Ei päde**	**Pätee jonkin verran**	**Pätee varmasti**
Ottaa muiden tunteet huomioon	0	1	2
Jakaa auliisti tavaroitaan (karkkeja, leluja, värikyniä jne)…	0	1	2
Tarjoutuu auttamaan, jos joku loukkaa itsensä, on …	0	1	2
Kiltti nuorempiaan kohtaan	0	1	2
Tarjoutuu usein auttamaan muita	0	1	2

Vaikeuksien kokonaispistemäärän laskeminen:

Saadaan laskemalla yhteen muiden paitsi prososiaalisen käytöksen kyselyosion pisteet. Kokonaispistemäärä voi vaihdella 0 ja 40 välillä (kokonaispistemäärää ei voi laskea, jos jonkun kyselyosion pistemäärä puuttuu).

Oirepisteiden tulkinta ja käyttö oireilevien lasten ("tapausten") tunnistamiseen

Vaikka SDQ pisteitä voidaan käyttää jatkuvina muuttujina, saattaa olla kätevää luokitella pistemäärät seuraavasti: normaali, raja-arvo ja poikkeava (korkea pistemäärä). Käytettäessä alla esitettäviä oirepisteiden luokitusrajoja vaikeuksien kokonaispistemäärän "poikkeava" luokkaa (korkea pistemäärä) voidaan käyttää mielenterveyshäiriön, "tapausten" tunnistamiseen. Tämä on selvästi vain karkea tapa tunnistaa häiriöitä - yhdistämällä useaan vastaajan antamaa tietoa oire- ja vaikutuspisteistä päästään parempaan joskaan ei lähellekään täydelliseen tulokseen. Keskimäärin 10 % osuus väestöotoksesta asettuu pisteytyksessä luokkaan "poikkeava" (korkea pistemäärä) ja edelleen 10 % osuus "raja-arvo" luokkaan. Tarkat osuudet vaihtelevat maan, iän ja sukupuolen mukaan - normatiivista dataa on saatavilla mäliit www- sivuilta. Voi olla tarpeen säätää luokitusrajoja näiden mainittujen tekijöiden mukaan; asettamalla luokitusrajaa korkeammalle, silloin kun on tärkeintä välttää väärä positiivisia löydöksiä, ja asettamalla luokitusrajaa matalammalle, kun on tärkeää välttää väärä negatiivisia löydöksiä..

Vanhemman vastaama	**Normaali**	**Raja-arvo**	**Poikkeava (korkea pistemäärä)**
Vaikeuksien kokonaispistemäärä	0 – 13	14 – 16	17 - 40
Tunnesoireiden pisteet	0-3	4	5 - 10
Käytösoireiden pisteet	0 – 2	3	4 - 10
Yliaktiivisuuden pisteet	0 – 5	6	7 - 10
Kaverisuhteiden ongelmien pisteet	0 – 2	3	4 - 10
Prososiaalisen käytöksen pisteet	6 – 10	5	0 - 4
Opettajan vastaama			
Vaikeuksien kokonaispistemäärä	0 – 11	12 – 15	16 - 40
Tunnesoireiden pisteet	0 – 4	5	6 - 10
Käytösoireiden pisteet	0 – 2	3	4 - 10
Yliaktiivisuuden pisteet	0 – 5	6	7 - 10
Kaverisuhteiden ongelmien pisteet	0 – 3	4	5 - 10
Prososiaalisen käytöksen pisteet	6 – 10	5	0 - 4

Vaikutuspisteiden laskeminen ja tulkinta

Käytettäessä vaikutusosion sisältävää SDQ -kyselyä, voidaan oireiden aiheuttamaa yleistä huolta, vaikeuksia ja sosiaalista toimintakykyhaittaa arvioivista väittämistä laskea yhteen vaikutuspisteet, jotka voivat vaihdella vanhempien kyselyssä 0-10 välillä ja opettajien kyselyssä 0-6 välillä.

Vanhemman vastaama	**Ei ollenkaan**	**Vain vähän**	**Aika paljon**	**Hyvin paljon**
Vaikeudet huolestuttavat lasta/muorta itseään	0	0	1	2
Vaikuttavat elämään ja arkeen KOTONA	0	0	1	2
Vaikuttavat lapsen KAVERISUHTEISIIN	0	0	1	2
Vaikuttavat OPPIMISEEN RYHMÄSSÄ / LUOKASSA	0	0	1	2
Vaikuttavat VAPAA-AIKAAN JA HARRASTUKSIIN	0	0	1	2
Opettajan vastaama				
Vaikeudet huolestuttavat lasta itseään	0	0	1	2
Vaikuttavat IKÄTOVERISUHTEISIIN	0	0	1	2
Vaikuttavat OPPIMISEEN RYHMÄSSÄ / LUOKASSA	0	0	1	2

Vastauksia kysymyksiin vaikeuksien kestosta ja rasittavuudesta ei lasketa mukaan vaikutuspisteisiin. Mikäli vastaus vaikutusosion ensimmäiseen kysymykseen on "ei" (vastaajan arvion mukaan lapsella ei ole vaikeuksia tunne-elämässään tai käytöksessään), vastaajaa ei pyydetä jatkamaan vastaamista muihin kysymyksiin oireiden aiheuttamasta yleisestä huolesta, vaikeuksista ja toimintakykyhaitasta. Tässä tilanteessa kertyy automaattisesti nolla vaikutuspistettä.

Vaikka vaikutuspisteitä voidaan käyttää jatkuvina muuttujina, saattaa olla kätevää luokitella pistemäärät seuraavasti: normaali, raja-arvo ja poikkeava. Pistemäärän ollessa 2 tai yli tulos on poikkeava (korkea pistemäärä), yhdellä pisteellä raja-arvo ja normaali tulos pistemäärän ollessa 0.

Appendix 7.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Lapsen oma arvio hyvinvoinnistaan

Lapsen nimi _____

Lapsen syntymäaika _____ (ppkkvvvv esim. 01012002)

Lomakkeen täyttöpvm _____ (ppkkvvvv)

MITÄ SINULLE KUULUU?

Valitse kuva, joka parhaiten kuvaa elämääsi ja vointiasi. Piirrä rasti (X) kuvan päällä olevaan laatikkoon.

				
1. Minulla on lähes aina surullinen tai kurja olo.	2. Minulla on usein surullinen tai kurja olo.	3. Minulla on yhtä paljon iloisia ja kurjia hetkiä.	4. Olen aika usein iloinen ja hyvällä tuulella.	5. Olen tosi usein iloinen ja hyvällä tuulella.

MITÄ ODOTAT TULEVILTA PÄIVILTÄSI? MILLAISTA ELÄMÄSI TULEE OLEMAAN?

Valitse kuva, joka parhaiten kuvaa, miltä sinusta tuntuu. Piirrä rasti (X) kuvan päällä olevaan laatikkoon.

				
1. Luotan siihen, että tulevat päivät ovat tosi kivoja.	2. Luotan siihen, että tulevat päivät ovat aika mukavia.	3. En juurikaan pelkää tulevia asioita.	4. Pelkään, että minulle tapahtuu melko paljon kurjia asioita.	5. Pelkään, että minulle tapahtuu paljon kurjia asioita.

KIITOS VASTAUKSESTASI !

OHJE:
Lapsi täyttää kaavakkeen yhdessä terveydenhoitajan kanssa. Terveydenhoitaja lukee kysymykset, ohjeet ja vaihtoehdot ääneen lapselle, vaikka lapsi osaisikin lukea.

Appendix 8.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset vanhemmille

Vahvuudet ja Vaikeudet -kyselyn käyttökelpoisuus

Täytitte kotiin postitetun, oikealla esitetyn kaltaisen **Vahvuudet ja Vaikeudet -kyselyn (Strengths and Difficulties Questionnaire, SDQ)**. Tätä kyselylomaketta käytettiin mahdollisesti myös keskustelun tukena terveydenhoitajan vastaanotolla. Pyydämme Teitä vanhemman näkökulmasta lyhyesti arvioimaan kyselyn ymmärrettävyyttä ja sopivuutta 5-6-vuotiaiden lasten hyvinvoinnin arviointivälineenä.

1. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 5-6-vuotiaan lapsen hyvinvoinnin arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

2. Kuinka kuormittavaksi koitte **Vahvuudet ja Vaikeudet -kyselyyn** vastaamisen?

   1 Ei lainkaan kuormittava
   2 Ei kovinkaan kuormittava
   3 Melko kuormittava
   4 Hyvin kuormittava

Palautekysely jatkuu seuraavalla sivulla

3. Mitä mieltä olette seuraavista väitteistä?

	Täysin samaa mieltä	Jokseenkin samaa mieltä	Jokseenkin eri mieltä	Täysin eri mieltä	En osaa sanoa
Vahvuudet ja Vaikeudet -kysely:					
lisää merkittävästi yhteistyötä terveydenhoitajan kanssa	1	2	3	4	5
lisää merkittävästi tietoa lapsen henkisestä hyvinvoinnista ja ihmissuhdetaidoista	1	2	3	4	5
auttaa löytämään yhteisen näkemyksen neuvolan kanssa lapsen tilanteesta ja tuen tarpeesta	1	2	3	4	5
on nykymuodossaan liian raskas väline neuvolaan	1	2	3	4	5
käyttö oli meistä vanhempina yleensä ottaen myönteinen kokemus	1	2	3	4	5

Lapsen oma arvio hyvinvoinnistaan -kyselyn käyttökelpoisuus

Terveydenhoitajan tarkastuksen yhteydessä lapsenne täytti terveydenhoitajan kanssa **Lapsen oma arvio –kyselyn**.

4. Kuinka hyvin tai huonosti **Lapsen oma arvio -kysely** sopii mielestänne lapsen hyvinvoinnin arviointiin?

1 Erittäin hyvin
2 Melko hyvin
3 Ei hyvin eikä huonosti
4 Melko huonosti
5 Erittäin huonosti
6 En osaa sanoa

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset vanhemmille

Vahvuudet ja Vaikeudet -kyselyn käyttökelpoisuus

Täytitte kotiin postitetun, oikealla esitetyn kaltaisen **Vahvuudet ja Vaikeudet -kyselyn (Strengths and Difficulties Questionnaire, SDQ).** Tätä kyselylomaketta käytettiin mahdollisesti myös keskustelun tukena terveydenhoitajan vastaanotolla. Pyydämme Teitä vanhemman näkökulmasta lyhyesti arvioimaan kyselyn ymmärrettävyyttä ja sopivuutta 7-8-vuotiaiden lasten hyvinvoinnin arviointivälineenä.

Vahvuuksien ja Vaikeuksien Kyselylomake (SDQ-Fin) V4-16

Pyytäisimme teitä ystävällisesti täyttämään tämän kyselylomakkeen koskien mainitun lapsenne käyttäytymistä viimeisen 6 kk:n (tai kuluvan kouluvuoden) aikana merkitsemällä rasti yhteen kolmesta annetusta vaihtoehdosta: "Ei Päde", "Pätee Jonkinverran", "Pätee Varmasti". On hyvin tärkeää, että vastaatte jokaiseen kohtaan parhaan kykynne mukaan siitäkin huolimatta, että aina ette tunne olevanne asiasta täysin varma - tai, että kysymys kuulostaa älyttömältä.

Lapsen / Nuoren Nimi .... Tyttö / Poika

Syntymäaika ....

	Ei Päde	Pätee Jonkinverran	Pätee Varmasti
Ottaa muiden tunteet huomioon	☐	☐	☐
Levoton, yliaktiivinen, ei pysty olemaan kauan hiljaa paikoillaan	☐	☐	☐
Valittaa usein päänsärkyä, vatsakipua tai pahoinvointia	☐	☐	☐
Jakaa auliisti tavaroitaan (karkkeja, leluja, värikyniä jne) muiden lasten kanssa	☐	☐	☐
Hänellä on usein kiukunpuuskia, tai hän kiivastuu helposti	☐	☐	☐
Ei näytä kaipaavan seuraa, leikkii usein itsekseen	☐	☐	☐
On yleensä tottelevainen, tavallisesti tekee niinkuin aikuinen käskee	☐	☐	☐
Hänellä on monia huolia, näyttää usein huolestuneelta	☐	☐	☐
Tarjoutuu auttamaan, jos joku loukkaa itsensä, on pahoilla mielin tai huonovointinen	☐	☐	☐
Jatkuvasti hypistelemässä jotakin tai kiemurtelee paikoillaan	☐	☐	☐
Hänellä on ainakin yksi hyvä ystävä	☐	☐	☐
Usein tappelee toisten lasten kanssa tai kiusaa muita	☐	☐	☐
Usein onneton, mieli maassa tai itkuinen	☐	☐	☐
Yleensä muiden lasten suosiossa	☐	☐	☐
Helposti häiriintyvä, mielenkiinto harhailee	☐	☐	☐
Uusissa tilanteissa pelokas tai aikuiseen takertuva, vailla itseluottamusta	☐	☐	☐
Kiltti nuorempiaan kohtaan	☐	☐	☐
Valehtelee tai petkuttaa usein	☐	☐	☐
Muiden lasten silmätikku tai kiusaamisen kohde	☐	☐	☐
Tarjoutuu usein auttamaan muita (vanhempiaan, opettajia, muita lapsia)	☐	☐	☐
Harkitsee tilanteen ennen kuin toimii	☐	☐	☐
Varastaa kotoa, koulusta tai muualta	☐	☐	☐
Tulee paremmin toimeen aikuisten kuin toisten lasten kanssa	☐	☐	☐
Kärsii monista peloista, usein peloissaan	☐	☐	☐
Saattaa tehtävät loppuun, hyvin pitkäjänteinen	☐	☐	☐

Onko teillä hänestä muita kommentteja tai huomautuksia ?

Kääntöpuolella muutama lisäkysymys - Olkaa Hyvä !

1. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 7-8-vuotiaan lapsen hyvinvoinnin arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

2. Kuinka kuormittavaksi koitte **Vahvuudet ja Vaikeudet -kyselyyn** vastaamisen?

   1 Ei lainkaan kuormittava
   2 Ei kovinkaan kuormittava
   3 Melko kuormittava
   4 Hyvin kuormittava

Palautekysely jatkuu seuraavalla sivulla

3. Mitä mieltä olette seuraavista väitteistä?

	Täysin samaa mieltä	Jokseenkin samaa mieltä	Jokseenkin eri mieltä	Täysin eri mieltä	En osaa sanoa
Vahvuudet ja Vaikeudet -kysely:					
lisää merkittävästi yhteistyötä terveydenhoitajan kanssa	1	2	3	4	5
lisää merkittävästi tietoa lapsen henkisestä hyvinvoinnista ja ihmissuhdetaidoista	1	2	3	4	5
auttaa löytämään yhteisen näkemyksen kouluterveydenhuollon kanssa lapsen tilanteesta ja tuen tarpeesta	1	2	3	4	5
on nykymuodossaan liian raskas väline kouluterveydenhuoltoon	1	2	3	4	5
käyttö oli meistä vanhempina yleensä ottaen myönteinen kokemus	1	2	3	4	5

Lapsen oma arvio hyvinvoinnistaan -kyselyn käyttökelpoisuus

Terveydenhoitajan tarkastuksen yhteydessä lapsenne täytti terveydenhoitajan kanssa **Lapsen oma arvio -kyselyn**.

4. Kuinka hyvin tai huonosti **Lapsen oma arvio -kysely** sopii mielestänne lapsen hyvinvoinnin arviointiin?

1 Erittäin hyvin
2 Melko hyvin
3 Ei hyvin eikä huonosti
4 Melko huonosti
5 Erittäin huonosti
6 En osaa sanoa

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset terveydenhoitajalle
Lapsen oma arvio -kyselyn, Vahvuudet ja Vaikeudet -kyselyn sekä LAPS-lomakkeen käyttökelpoisuudesta

OSA A

"Lasten mielenterveystyön kehittäminen" -hankkeeseen liittyen olette täyttänyt 5-6-vuotiaiden määräaikaistarkastuksissa lasten kanssa **Lapsen oma arvio -kyselyn**. Pyydämme Teitä terveydenhoitajan näkökulmasta lyhyesti arvioimaan kyselyn käytettävyyttä, ymmärrettävyyttä ja sopivuutta.

1. Kuinka kauan **Lapsen oma arvio -kysely** vei keskimäärin aikaanne *yhden* lapsen kohdalla? (arvio vähintään 5 min. tarkkuudella)

   _____ minuuttia

2. Kuinka hyvin tai huonosti **Lapsen oma arvio -kysely** sopii mielestänne 5-6-vuotiaan lapsen henkisen hyvinvoinnin arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

3. Kuinka kuormittavaksi koitte **Lapsen oma arvio -kyselyn**?

   1 Ei lainkaan kuormittava
   2 Ei kovinkaan kuormittava
   3 Melko kuormittava
   4 Hyvin kuormittava

OSA B

Kehittämishankkeeseen liittyen 5-6-vuotiaiden määräaikaistarkastuksissa vanhemmat palauttivat kauttanne esitäyttämiään **Vahvuudet ja Vaikeudet -kyselyitä** (Strengths and Difficulties Questionnaire, SDQ). Valitettavasti tämän hankkeen aikana kyselyiden tulokset pistemäärinä eivät olleet vielä terveydenhoitajien käytössä. Mikäli kuitenkin hyödynsitte **Vahvuudet ja Vaikeudet -kyselyä** keskustelun tukena, pyydämme Teitä vastaamaan seuraaviin kysymyksiin.

4. **Vahvuudet ja Vaikeudet -kysely** on menetelmä, jolla kootaan vanhemmilta ja päivähoidosta tietoa lapsen henkisestä hyvinvoinnista, ihmissuhdetaidoista ja näihin liittyvästä mahdollisesta tuen tarpeesta. Mitä menetelmiä olette tähän mennessä käyttänyt samaan käyttötarkoitukseen? Luetelkaa korkeintaan viisi tärkeintä.

_____

_____

5. Kuinka kauan **Vahvuudet ja Vaikeudet -kyselyn** läpikäyminen vei keskimäärin aikaanne *yhden* lapsen/perheen kohdalla? (arvio vähintään 5 min. tarkkuudella)

_____ minuuttia

6. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 5-6-vuotiaan lapsen henkisen hyvinvoinnin, ihmissuhdetaitojen ja näihin liittyvän mahdollisen tuen tarpeen arviointiin?

1 Erittäin hyvin
2 Melko hyvin
3 Ei hyvin eikä huonosti
4 Melko huonosti
5 Erittäin huonosti

7. Kuinka kuormittavaksi koitte keskimäärin omalta kannaltanne **Vahvuudet ja Vaikeudet -kyselyn** käytön neuvolatyössä

1 Ei lainkaan kuormittava
2 Ei kovinkaan kuormittava
3 Melko kuormittava
4 Hyvin kuormittava

8. Kun vertaatte **Vahvuudet ja Vaikeudet -kyselyä** tähän asti neuvolatyössänne käytössä olleisiin arviointimenetelmiin (kysymys 4), miten Vahvuudet ja Vaikeudet -kysely mielestänne poikkeaa aiemmin käyttämistänne menetelmistä?

_____

_____

_____

9. Mitä mieltä olette seuraavista väitteistä liittyen **Vaikeudet ja Vahvuudet -kyselyyn**?

	Täysin samaa mieltä	Jokseenkin samaa mieltä	Jokseenkin eri mieltä	Täysin eri mieltä	En osaa sanoa
Lisää merkittävästi yhteistyötä perheen kanssa	1	2	3	4	5
Lisää merkittävästi tietoa lapsen henkisestä hyvinvoinnista ja ihmissuhdetaidoista	1	2	3	4	5
Auttaa löytämään yhteisen näkemyksen vanhempien kanssa lapsen tilanteesta ja tuen tarpeesta	1	2	3	4	5
On nykymuodossaan liian raskas väline neuvolakäyttöön	1	2	3	4	5
Perheiden vanhemmat kokivat sen käytön yleensä ottaen myönteisesti	1	2	3	4	5

10. Voitte lopuksi antaa vapaamuotoista palautetta **Lapsen oma arvio -kyselyn** tai **Vaikeudet ja Vahvuudet -kyselyn** soveltuvuudesta neuvolatyöhön.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset terveydenhoitajalle
Lapsen oma arvio -kyselyn, Vahvuudet ja Vaikeudet -kyselyn sekä LAPS-lomakkeen käyttökelpoisuudesta

OSA A

"Lasten mielenterveystyön kehittäminen" -hankkeeseen liittyen olette täyttänyt 7-8-vuotiaiden määräaikaistarkastuksissa lasten kanssa **Lapsen oma arvio -kyselyn**. Pyydämme Teitä terveydenhoitajan näkökulmasta lyhyesti arvioimaan kyselyn käytettävyyttä, ymmärrettävyyttä ja sopivuutta.

1. Kuinka kauan **Lapsen oma arvio -kysely** vei keskimäärin aikaanne *yhden* lapsen kohdalla? (arvio vähintään 5 min. tarkkuudella)

   _____ minuuttia

2. Kuinka hyvin tai huonosti **Lapsen oma arvio -kysely** sopii mielestänne 7-8-vuotiaan lapsen henkisen hyvinvoinnin arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

3. Kuinka suoritettavaksi koitte **Lapsen oma arvio -kyselyn**?

   1 Ei lainkaan suoritettava
   2 Ei kovinkaan suoritettava
   3 Melko suoritettava
   4 Hyvin suoritettava

OSA B

Kehittämishankkeeseen liittyen 7-8-vuotiaiden määräaikaistarkastuksissa vanhemmat palauttivat kanttanne esitäyttämiään **Vahvuudet ja Vaikeudet -kyselyitä** (Strengths and Difficulties Questionnaire, SDQ). Valitettavasti tämän hankkeen aikana kyselyiden tulokset pistemäärinä eivät olleet vielä terveydenhoitajien käytössä. Mikäli kuitenkin hyödynsitte **Vahvuudet ja Vaikeudet -kyselyä** keskustelun tukena, pyydämme Teitä vastaamaan seuraaviin kysymyksiin.

4. **Vahvuudet ja Vaikeudet -kysely** on menetelmä, jolla kootaan vanhemmilta ja opettajalta tietoa lapsen henkisestä hyvinvoinnista, ihmissuhdetaidoista ja näihin liittyvästä mahdollisesta tuen tarpeesta. Mitä menetelmiä olette tähän mennessä käyttänyt samaan käyttötarkoitukseen? Luetelkaa korkeintaan viisi tärkeintä.

_____

_____

5. Kuinka kauan **Vahvuudet ja Vaikeudet -kyselyn** läpikäyminen vei keskimäärin aikaanne *yhden* lapsen/perheen kohdalla? (arvio vähintään 5 min. tarkkuudella)

_____ minuuttia

6. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 7-8-vuotiaan lapsen henkisen hyvinvoinnin, ihmissuhdetaitojen ja näihin liittyvän mahdollisen tuen tarpeen arviointiin?

1 Erittäin hyvin
2 Melko hyvin
3 Ei hyvin eikä huonosti
4 Melko huonosti
5 Erittäin huonosti

7. Kuinka suorittavaksi koitte keskimäärin omalta kannaltanne **Vahvuudet ja Vaikeudet -kyselyn** käytön kouluterveydenhuollossa?

1 Ei lainkaan suorittava
2 Ei kovinkaan suorittava
3 Melko suorittava
4 Hyvin suorittava

8. Kun vertaatte **Vahvuudet ja Vaikeudet -kyselyä** tähän asti kouluterveydenhuollossa käytössä olleisiin arviointimenetelmiin (kysymys 4), miten Vahvuudet ja Vaikeudet -kysely mielestänne poikkeaa aiemmin käyttämistänne menetelmistä?

_____

_____

_____

9. Mitä mieltä olette seuraavista väitteistä liittyen **Vaikeudet ja Vahvuudet -kyselyyn**?

	Täysin samaa mieltä	Jokseenkin samaa mieltä	Jokseenkin eri mieltä	Täysin eri mieltä	En osaa sanoa
Lisää merkittävästi yhteistyötä perheen kanssa	1	2	3	4	5
Lisää merkittävästi tietoa lapsen henkisestä hyvinvoinnista ja ihmissuhdetaidoista	1	2	3	4	5
Auttaa löytämään yhteisen näkemyksen vanhempien kanssa lapsen tilanteesta ja tuen tarpeesta	1	2	3	4	5
On nykymuodossaan liian raskas väline kouluterveydenhuoltoon	1	2	3	4	5
Perheiden vanhemmat kokivat sen käytön yleensä ottaen myönteisesti	1	2	3	4	5

10. Voitte lopuksi antaa vapaamuotoista palautetta **Lapsen oma arvio -kyselyn** tai **Vaikeudet ja Vahvuudet -kyselyn** soveltuvuudesta kouluterveydenhuoltoon.

_____

_____

_____

_____

_____

_____

OSA C

11. Lääkäri täytti 7-8-vuotiaiden lasten määräaikaistarkastuksissa **Lastenpsykiatrian arviointimenetelmä -lomakkeen (LAPS)**. Nykyinen lomakeluonnos on tarkoitettu kokoamaan laajemmin informaatiota ja arvioimaan lasten henkisen hyvinvoinnin ja ihmissuhdetaitojen kehitystä, terveyttä ja tuen tarvetta. Voitte halutessanne antaa vapaamuotoista palautetta terveydenhoitajan näkökulmasta LAPS-lomakkeen soveltuvuudesta lääkärin työvälineeksi kouluterveydenhuollossa.

_____

_____

_____

Kiitos palautteestanne!

Appendix 9.

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset päivähoidolle
Vahvuudet ja vaikeudet -kyselyn käyttökelpoisuudesta

"Lasten mielenterveystyön kehittäminen" -hankkeeseen liittyen olette täyttänyt 5-6-vuotiaiden havainnoinnin yhteydessä **Vahvuudet ja Vaikeudet -kyselyn** (Strengths and Difficulties Questionnaire, SDQ) yhdestä tai useammasta lapsesta. Pyydämme Teitä päivähoidon näkökulmasta lyhyesti arvioimaan kyselyn käytettävyyttä, ymmärrettävyyttä ja sopivuutta.

1. Kuinka kauan **Vahvuudet ja Vaikeudet -kyselyn** täyttäminen vei keskimäärin aikaanne yhden lapsen/perheen kohdalla? (arvio vähintään 5 min. tarkkuudella)

   _____ minuuttia

2. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 5-6-vuotiaan lapsen henkisen hyvinvoinnin ja mahdollisen tuen tarpeen arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

3. Kuinka kuormittavaksi koitte **Vahvuudet ja Vaikeudet -kyselyyn** vastaamisen?

   1 Ei lainkaan kuormittava
   2 Ei kovinkaan kuormittava
   3 Melko kuormittava
   4 Hyvin kuormittava

4. **Vahvuudet ja Vaikeudet -kysely** on menetelmä, jolla kootaan vanhemmilta ja päivähoidolta tietoa lapsen henkisestä hyvinvoinnista, ihmissuhdetaidoista ja näihin liittyvästä mahdollisesta tuen tarpeesta. Mitä menetelmiä olette tähän mennessä käyttänyt samaan käyttötarkoitukseen? Luetelkaa korkeintaan viisi tärkeintä.

   _____

   _____

   _____

   _____

5. Kun vertaatte **Vahvuudet ja Vaikeudet -kyselyä** tähän asti päivähoidossa käytössä olleisiin arviointimenetelmiin (kysymys 4), miten Vahvuudet ja Vaikeudet -kysely mielestänne poikkeaa aiemmin käyttämistänne menetelmistä?

_____

_____

6. Mitä mieltä olette seuraavista väitteistä liittyen **Vahvuudet ja Vaikeudet -kyselyyn**?

	Täysin samaa mieltä	Jokseenkin samaa mieltä	Jokseenkin eri mieltä	Täysin eri mieltä	En osaa sanoa
Lisää merkittävästi yhteistyötä perheen kanssa	1	2	3	4	5
Lisää merkittävästi tietoa lapsen henkisestä hyvinvoinnista ja ihmissuhdetaidoista	1	2	3	4	5
Auttaa löytämään yhteisen näkemyksen vanhempien kanssa lapsen tilanteesta ja tuen tarpeesta	1	2	3	4	5
On nykymuodossaan liian raskas väline päivähoitoon	1	2	3	4	5
Perheiden vanhemmat kokivat sen käytön yleensä ottaen myönteisesti	1	2	3	4	5

7. Voitte lopuksi antaa vapaamuotoista palautetta päivähoidon näkökulmasta **Vahvuudet ja Vaikeudet -kyselyn** soveltuvuudesta päivähoitoon. Yllätyittekö esimerkiksi jostakin asiasta myönteisesti tai kielteisesti?

_____

_____

_____

_____

_____

_____

_____

Kiitos palautteestanne!

PIRKANMAAN SAIRAANHOITOPIIRI
Yhdessä terveyttä

ETELÄ-KARJALAN SAIRAANHOITOPIIRI

Lastenpsykiatrian vastuualue
Lasten mielenterveystyön kehittämishanke

Palautekysymykset opettajalle
Vahvuudet ja Vaikeudet -kyselyn käyttökelpoisuudesta

"Lasten mielenterveystyön kehittäminen" -hankkeeseen liittyen olette täyttänyt 7-8-vuotiaiden havainnoinnin yhteydessä **Vahvuudet ja Vaikeudet -kyselyn** (Strengths and Difficulties Questionnaire, SDQ) yhdestä tai useammasta lapsesta. Pyydämme Teitä opettajan näkökulmasta lyhyesti arvioimaan kyselyn käytettävyyttä, ymmärrettävyyttä ja sopivuutta.

1. Kuinka kauan **Vahvuudet ja Vaikeudet –kyselyn** täyttäminen vei keskimäärin aikaanne yhden oppilaan kohdalla? (arvio vähintään 5 min. tarkkuudella)

   _____ minuuttia

2. Kuinka hyvin tai huonosti **Vahvuudet ja Vaikeudet -kysely** sopii mielestänne 7-8-vuotiaan lapsen henkisen hyvinvoinnin ja mahdollisen tuen tarpeen arviointiin?

   1 Erittäin hyvin
   2 Melko hyvin
   3 Ei hyvin eikä huonosti
   4 Melko huonosti
   5 Erittäin huonosti

3. Kuinka kuormittavaksi koitte **Vahvuudet ja Vaikeudet –kyselyyn** vastaamisen?

   1 Ei lainkaan kuormittava
   2 Ei kovinkaan kuormittava
   3 Melko kuormittava
   4 Hyvin kuormittava

4. Voitte lopuksi antaa vapaamuotoista palautetta tai kehittämisideoita opettajan näkökulmasta **Vahvuudet ja Vaikeudet -kyselyn** soveltuvuudesta opettajan käyttöön. Yllätyittekö esimerkiksi jostakin asiasta myönteisesti tai kielteisesti?

   _____

   _____

   _____

   _____

Kiitos palautteestanne!

Appendix 10.

Frequency distributions of the Strengths and Difficulties Questionnaire (SDQ) for Finnish 4—9 year-olds

Table 1. Frequency distributions of the SDQ total scores rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Total difficulties score	Mothers (n = 2582)		Fathers (n = 1935)		Parents' average (n = 2635)		Teachers' average (n = 2242)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	1.7	1.7	2.4	2.4	0.9	0.9	9.9	9.9
1	4.7	6.5	4.1	6.6	2.5	4.2	10.9	22.6
2	6.7	13.2	7.2	13.7	4.2	10.0	10.0	34.8
3	10.7	23.9	10.8	24.5	7.1	19.8	8.9	45.9
4	12.0	35.9	10.5	35.0	7.5	31.3	7.2	54.5
5	11.2	47.1	10.9	45.9	8.3	43.5	6.2	62.6
6	10.8	57.9	11.5	57.4	7.1	55.0	4.8	69.0
7	8.8	66.7	8.9	66.4	6.1	65.1	3.7	73.9
8	7.8	74.5	7.9	74.2	5.5	73.9	4.0	79.5
9	5.8	80.4	6.8	81.0	3.8	79.8	3.4	83.6
10	4.6	84.9	4.4	85.5	2.9	84.9	1.7	86.1
11	3.6	88.5	2.8	88.3	2.4	88.5	1.9	88.4
12	2.7	91.2	2.9	91.3	1.7	91.0	1.7	90.7
13	1.9	93.2	2.0	93.3	1.5	93.5	1.2	92.2
14	1.7	94.9	2.0	95.2	1.1	95.3	0.8	93.2
15	0.9	95.8	1.2	96.5	0.6	96.4	0.8	94.3
16	0.9	96.7	1.0	97.5	0.8	97.5	0.8	95.3
17	0.7	97.4	0.8	98.2	0.3	98.1	0.8	96.3
18	0.7	98.1	0.6	98.8	0.3	98.5	0.7	97.2
19	0.5	98.6	0.3	99.1	0.3	98.9	0.5	97.9
20	0.4	99.0	0.3	99.3	0.3	99.2	0.4	98.4
21	0.3	99.3	0.2	99.5	0.2	99.4	0.6	99.1
22	0.1	99.4	0.2	99.7	0.1	99.5	0.1	99.2
23	0.2	99.6	0.1	99.8	0.1	99.7	0.3	99.5
24	0.2	99.8			0.1	99.7	0.1	99.6
25	0.1	99.9	0.1	99.9	0.1	99.9	0.1	99.7
26	0.0	99.9	0.1	100.0	0.0	100.0	0.2	100.0
27	0.0	100.0						
28							0.0	100.0
29-35								
36	0.0	100.0			0.0	100.0		
37-40								

Table 2. Frequency distributions of the scores of the SDQ emotional symptoms subscale rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Emotional symptoms score	Mothers (n = 2582)		Fathers (n = 1936)		Parents' average (n=2635)		Teachers' average (n=2242)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	41.1	41.1	42.5	42.5	32.4	32.4	51.8	51.8
1	33.0	74.1	34.7	77.2	24.0	70.7	19.5	79.0
2	14.6	88.7	13.1	90.3	9.4	88.0	7.9	89.5
3	5.9	94.6	5.3	95.6	4.1	95.2	3.6	94.3
4	3.1	97.7	2.4	98.0	1.6	97.9	2.5	97.4
5	1.2	98.9	1.1	99.1	0.7	99.0	1.1	98.9
6	0.5	99.4	0.6	99.7	0.3	99.5	0.6	99.5
7	0.2	99.6	0.1	99.8	0.2	99.8	0.3	99.8
8	0.3	99.9	0.2	100.0	0.2	100.0	0.1	99.9
9	0.0	100.0			0.0	100.0	0.0	100.0
10	0.0	100.0					0.0	100.0

Table 3. Frequency distributions of the scores of the SDQ conduct problems subscale rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Conduct problems score	Mothers (n = 2582)		Fathers (n = 1936)		Parents' average (n=2635)		Teachers' average (n=2242)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	26.9	26.9	27.2	27.2	19.6	19.6	58.9	58.9
1	28.2	55.2	29.6	56.9	20.5	50.1	13.2	76.8
2	21.8	76.9	22.2	79.1	16.4	75.9	7.0	86.0
3	13.0	89.9	11.6	90.7	8.2	89.9	5.0	91.6
4	6.2	96.1	6.1	96.8	3.3	96.2	3.2	95.6
5	2.0	98.1	2.0	98.8	1.6	98.6	1.8	97.9
6	1.4	99.4	0.9	99.7	0.3	99.4	0.9	99.2
7	0.3	99.7	0.2	99.8	0.2	99.8	0.3	99.7
8	0.1	99.8	0.1	99.9	0.1	99.9	0.3	100.0
9	0.1	99.9	0.1	100.0	0.0	100.0		
10	0.1	100.0			0.0	100.0		

Table 4. Frequency distributions of the scores of the SDQ hyperactivity and inattention problems subscale rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Hyperactivity score	Mothers (n = 2581)		Fathers (n = 1935)		Parents' average (n=2634)		Teachers' average (n=2242)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	13.7	13.7	12.8	12.8	8.1	8.1	26.9	26.9
1	21.6	35.3	19.9	32.7	15.1	29.0	15.8	46.5
2	26.8	62.1	26.4	59.1	19.1	56.5	13.1	62.7
3	12.9	75.1	14.3	73.4	9.7	72.9	7.8	72.8
4	9.1	84.2	10.9	84.3	6.5	83.0	5.7	80.2
5	7.4	91.6	7.5	91.8	5.6	91.3	5.8	87.6
6	3.7	95.3	3.4	95.1	2.7	95.5	2.9	91.2
7	1.7	97.0	2.3	97.4	1.1	97.3	2.2	94.3
8	1.1	98.1	1.3	98.8	0.8	98.5	1.2	95.9
9	1.2	99.3	0.8	99.5	0.7	99.4	1.7	97.9
10	0.7	100.0	0.5	100.0	0.5	100.0	1.6	100.0

Table 5. Frequency distributions of the scores of the SDQ peer problems subscale rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Peer problems score	Mothers (n = 2582)		Fathers (n = 1934)		Parents' average (n=2635)		Teachers' average (n=2241)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	26.8	26.8	24.0	24.0	16.8	16.8	39.0	39.0
1	32.7	59.5	34.2	58.2	25.8	52.8	22.8	68.8
2	21.5	81.1	23.3	81.4	16.1	80.1	11.2	83.4
3	11.7	92.8	11.6	93.0	7.3	92.7	6.6	91.6
4	4.3	97.1	4.6	97.6	2.8	97.3	3.0	95.2
5	1.9	99.0	1.4	99.0	1.1	99.1	1.9	97.8
6	0.5	99.4	0.5	99.5	0.2	99.5	0.8	98.7
7	0.4	99.8	0.5	100.0	0.2	99.8	0.6	99.5
8	0.2	100.0			0.1	100.0	0.3	99.9
9	0.0	100.0					0.0	100.0
10								

Table 6. Frequency distributions of the scores of the SDQ prosocial behaviour subscale rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Prosocial behaviour score	Mothers (n = 2582)		Fathers (n = 1935)		Parents' average (n=2635)		Teachers' average (n=2235)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	0.0	0.0	0.1	0.1	0.0	0.0	0.4	0.4
1	0.2	0.2	0.1	0.2	0.1	0.1	1.4	1.9
2	0.4	0.6	0.4	0.5	0.2	0.5	2.2	4.3
3	0.9	1.5	1.1	1.6	0.5	1.0	3.0	7.6
4	2.3	3.8	3.4	5.0	1.5	2.8	4.3	12.7
5	8.9	12.7	10.2	15.2	5.0	9.3	9.8	24.1
6	15.3	28.0	15.3	30.5	10.6	23.9	11.5	38.4
7	17.2	45.2	17.2	47.8	12.6	42.2	10.4	51.5
8	19.1	64.2	19.7	67.5	13.7	63.8	14.3	68.8
9	18.9	83.2	17.0	84.5	14.0	84.6	11.1	83.1
10	16.8	100.0	15.5	100.0	9.9	100.0	14.4	100.0

Table 7. Frequency distributions of the impact scores of the SDQ rated by parents and teachers in a Finnish community sample of 4—9 year olds.

Impact score	Mothers (n = 2576)		Fathers (n = 1932)		Parents' average (n=2634)		Teachers' average (n=2230)	
	%	Cumul. %	%	Cumul. %	%	Cumul. %	%	Cumul. %
0	90.2	90.2	92.8	92.8	88.3	88.3	87.9	87.9
1	5.8	96.0	4.1	96.9	3.7	96.0	5.2	95.0
2	2.4	98.4	1.8	98.7	1.6	98.4	2.3	98.0
3	0.7	99.1	0.6	99.3	0.5	99.3	1.1	99.4
4	0.4	99.5	0.4	99.7	0.2	99.5	0.4	99.9
5	0.2	99.7	0.1	99.8	0.2	99.7	0.1	100.0
6	0.1	99.8	0.1	99.8	0.2	99.9	100.0	
7	0.1	99.9	0.1	99.9	0.1	100.0		
8	0.1	100.0	0.1	100.0				
9								
10								

Frequency distributions of the Strengths and Difficulties Questionnaire (SDQ) for Finnish 4—9 year-old boys and girls

Table 8. Frequency distributions of the SDQ total scores rated by parents and teachers in a Finnish community sample of 4—9-year-old boys and girls.

Total difficulties score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1344)	4-6 years (n = 881)	7-9 years (n = 463)	All (n = 1099)	4-6 years (n = 670)	7-9 years (n = 429)	All (n = 1143)	4-6 years (n = 687)	7-9 years (n = 456)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	0.7	0.2	1.6	1.1	0.9	1.5	5.9	4.9	7.5	13.8	13.8	21.5
1	3.3	2.5	4.9	5.0	3.9	7.1	15.5	13.9	17.9	29.4	29.4	41.9
2	8.2	7.2	10.3	11.8	9.9	15.3	26.7	24.3	30.3	42.7	42.7	54.4
3	16.1	14.0	20.4	23.4	20.4	29.2	37.4	35.1	41.0	54.0	54.0	64.7
4	26.6	24.1	31.6	35.9	32.2	43.0	44.9	42.4	49.0	63.8	63.8	73.2
5	37.3	35.1	41.7	49.6	46.5	55.3	53.6	51.9	56.2	71.3	71.3	79.2
6	48.9	47.0	52.7	60.9	58.7	65.0	60.5	59.4	62.2	77.2	77.2	82.9
7	58.6	55.9	64.2	71.3	69.8	74.1	66.1	65.8	66.4	81.5	81.5	86.4
8	68.6	67.0	71.9	78.9	77.9	81.0	72.3	71.8	73.2	86.4	86.4	90.1
9	74.7	73.7	76.6	84.7	84.3	85.5	77.7	77.5	78.1	89.3	89.3	91.4
10	80.3	79.7	81.5	89.2	89.3	89.0	80.6	80.4	80.9	91.3	91.3	93.4
11	84.7	84.4	85.5	92.2	92.3	92.0	83.4	83.3	83.7	93.1	93.1	94.7
12	87.6	87.6	87.6	94.3	94.3	94.2	86.9	87.2	86.5	94.3	94.3	96.1
13	90.4	90.5	90.2	96.4	97.0	95.2	89.3	89.0	89.7	95.0	95.0	96.3
14	93.0	93.3	92.3	97.6	98.4	96.1	90.5	90.4	90.7	95.8	95.8	96.7
15	94.5	94.7	94.1	98.1	99.0	96.5	92.1	92.5	91.4	96.4	96.4	97.1
16	96.2	96.8	95.1	98.7	99.4	97.4	93.4	93.6	93.0	97.2	97.2	97.8

17	97.2	97.7	96.3	99.0	99.5	98.1	94.7	94.6	94.9	97.8	97.8	98.5
18	97.8	98.1	97.2	99.1	99.7		95.7	95.8	95.6	98.6	98.6	98.7
19	98.5	98.7	97.9	99.3		98.5	96.8	96.9	96.7	99.0	99.0	
20	98.8	99.0	98.6	99.5		99.1	97.6	97.3	98.1	99.1	99.1	
21	99.1	99.3		99.6	99.9		98.6	98.5	98.8	99.5	99.5	99.3
22	99.4	99.5	99.1							99.7	99.7	
23	99.5		99.3	99.8	100.0	99.4	99.1	99.1	99.1	99.8	99.8	99.6
24	99.6		99.5	99.9		99.6	99.4		99.5			
25	100.0	100.0	100.0				99.6	99.7				
26				99.9		99.8	99.9	100.0	99.8	100.0	100.0	100.0
27												
28							100.0		100.0			
29-35												
36				100.0		100.0						
37-40												

*The scores of two informants were combined into a single parent or a single teacher score.

Table 9. Frequency distributions of the scores of the SDQ emotional symptoms subscale rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Emotional symptoms score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1344)	4-6 years (n = 881)	7-9 years (n = 463)	All (n = 1099)	4-6 years (n = 670)	7-9 years (n = 429)	All (n = 1143)	4-6 years (n = 687)	7-9 years (n = 456)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	34.1	35.4	31.4	30.7	31.7	28.9	52.5	46.6	61.8	51.2	44.4	61.4
1	71.1	72.6	68.1	70.4	73.1	65.2	79.5	79.3	80.0	78.6	77.3	80.5
2	87.7	89.1	84.8	88.2	89.9	85.1	90.7	92.1	88.6	88.4	87.6	89.5
3	95.0	96.8	91.6	95.4	96.9	92.4	95.5	96.9	93.2	93.2	92.7	93.9
4	97.8	98.6	96.3	98.0	98.8	96.5	97.9	98.7	96.7	96.9	96.7	97.4
5	98.9	99.2	98.4	99.1	99.7	98.1	99.5	99.9	98.8	98.4	98.5	98.2
6	99.3	99.5	98.8	99.6	99.9	99.1	99.5		99.1	99.5	99.6	99.3
7	99.9	100.0	99.8				99.7	100.0	99.3	99.8	99.7	100.0
8	100.0		100.0	99.9	100.0	99.8	100.0		100.0			
9				100.0		100.0				99.9	99.9	
10										100.0	100.0	

*The scores of two informants were combined into a single parent or a single teacher score.

Table 10. Frequency distributions of the scores of the SDQ conduct problems subscale rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Conduct problems score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1344)	4-6 years (n = 881)	7-9 years (n = 463)	All (n = 1099)	4-6 years (n = 670)	7-9 years (n = 429)	All (n = 1143)	4-6 years (n = 687)	7-9 years (n = 456)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	13.3	13.3	23.9	22.3	19.2	28.3	50.7	45.1	59.4	66.8	57.2	81.4
1	41.1	41.1	57.8	53.4	49.8	60.3	69.8	66.4	75.1	83.5	78.9	90.4
2	67.4	67.4	76.6	81.3	79.7	84.2	80.4	79.4	82.1	91.3	89.7	93.9
3	86.9	86.9	88.8	92.2	92.3	92.0	88.7	86.9	91.6	94.4	92.9	96.7
4	94.7	94.7	94.6	97.7	97.8	97.4	94.2	93.1	95.8	97.0	95.8	98.9
5	98.1	98.1	97.4	99.3	99.4	98.9	97.1	96.3	98.4	98.8	98.3	99.6
6	99.2	99.2	98.8	99.6	99.9	99.1	98.5	98.4	98.8	99.9	99.9	100.0
7	99.8	99.8	99.8	99.8	100.0	99.4	99.5	99.6	99.3			
8	100.0	100.0		99.9		99.6	100.0	100.0	100.0	100.0	100.0	
9			100.0									
10				100.0		100.0						

*The scores of two informants were combined into a single parent or a single teacher score.

Table 11. Frequency distributions of the scores of the SDQ hyperactivity and inattention problems subscale rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Hyperactivity score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1343)	4-6 years (n = 881)	7-9 years (n = 462)	All (n = 1099)	4-6 years (n = 670)	7-9 years (n = 429)	All (n = 1143)	4-6 years (n = 687)	7-9 years (n = 456)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	6.0	4.9	8.4	10.1	8.1	13.9	17.3	14.8	21.2	36.2	26.5	50.9
1	22.6	20.5	26.9	35.2	31.7	42.0	35.0	32.2	39.4	57.5	50.5	68.0
2	47.5	45.4	51.8	65.2	62.9	69.7	50.9	50.4	51.5	74.1	70.3	79.8
3	65.9	65.2	67.4	79.7	78.2	82.5	62.1	61.9	62.5	83.1	79.5	88.6
4	77.7	77.1	78.9	88.0	86.8	90.3	71.9	72.4	71.1	88.2	85.2	92.8
5	87.8	87.8	87.8	94.6	94.8	94.4	81.4	82.1	80.4	93.5	92.0	95.8
6	93.7	93.8	93.7	97.2	97.6	96.3	86.3	86.6	85.8	96.0	94.9	97.6
7	96.0	95.9	96.0	98.5	99.0	97.6	91.0	91.3	90.4	97.5	97.1	98.0
8	97.8	98.1	97.2	99.2	99.3	98.9	93.4	94.0	92.5	98.3	98.1	98.5
9	99.0	99.3	98.4	99.9	99.9	99.8	96.6	96.9	96.3	99.2	99.3	99.1
10	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

*The scores of two informants were combined into a single parent or a single teacher score.

Table 12. Frequency distributions of the scores of the SDQ peer problems subscale rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Peer problems score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1344)	4-6 years (n = 881)	7-9 years (n = 463)	All (n = 1098)	4-6 years (n = 670)	7-9 years (n = 428)	All (n = 1143)	4-6 years (n = 687)	7-9 years (n = 456)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	15.7	14.2	18.7	17.9	15.1	23.1	36.1	33.9	39.5	41.9	38.6	46.9
1	49.8	47.3	54.8	55.7	53.3	60.3	65.1	63.9	67.1	72.3	70.0	75.7
2	77.3	75.8	80.3	82.8	81.4	85.5	80.6	79.7	82.0	86.1	84.4	88.6
3	90.3	89.8	91.3	94.9	95.3	94.2	89.4	88.8	90.4	93.6	93.4	93.9
4	95.8	96.3	94.8	98.7	98.8	98.5	93.7	92.7	95.3	96.6	96.8	96.3
5	98.5	98.6	98.1	99.7	99.8	99.6	96.9	96.9	97.0	98.6	98.7	98.5
6	99.2	99.2	99.3	99.9	99.9	99.8	97.9	97.8	98.1	99.5	99.4	99.6
7	99.7	99.7	99.8		100.0		99.1	99.1	99.1	99.9	99.9	100.0
8	100.0	100.0	100.0	100.0		100.0	99.8	99.9	99.8			
9							100.0	100.0	100.0	100.0	100.0	
10												

*The scores of two informants were combined into a single parent or a single teacher score.

Table 13. Frequency distributions of the scores of the SDQ prosocial behaviour subscale rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Prosocial behaviour score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 864)	7-9 years (n = 427)	All (n =1344)	4-6 years (n = 881)	7-9 years (n = 463)	All (n = 1097)	4-6 years (n = 670)	7-9 years (n = 429)	All (n = 1138)	4-6 years (n = 685)	7-9 years (n = 453)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0				0.1		0.2	0.5	0.4	0.7	0.4	0.1	0.7
1	0.1	0.1	0.2	0.1		0.4	2.4	0.9	4.7	1.5	0.9	2.4
2	0.7	0.7	0.7	0.2		0.6	5.5	2.2	10.5	3.2	2.6	4.0
3	1.5	1.6	1.2	0.5	0.2	1.1	10.6	7.0	16.2	4.7	3.4	6.8
4	4.3	4.1	4.7	1.5	1.5	1.5	18.0	14.8	23.2	7.6	6.0	9.9
5	13.1	13.4	12.4	5.7	6.0	5.2	32.3	27.2	40.3	16.3	14.3	19.2
6	30.0	30.6	28.8	18.0	18.8	16.4	48.5	43.9	55.7	28.7	26.3	32.5
7	50.2	50.6	49.4	34.6	36.4	31.1	61.7	59.0	66.0	41.6	40.6	43.0
8	71.3	70.7	72.4	56.6	57.1	55.7	77.1	75.4	79.9	60.8	61.6	59.6
9	88.8	88.9	88.8	80.6	82.0	78.0	88.4	87.9	89.2	77.9	81.5	72.6
10	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

*The scores of two informants were combined into a single parent or a single teacher score.

Table 14. Frequency distributions of the impact scores of the SDQ rated by parents and teachers in Finnish 4—9-year-old boys and girls.

Impact score	Parents*						Teachers*					
	Boys			Girls			Boys			Girls		
	All (n = 1291)	4-6 years (n = 860)	7-9 years (n = 427)	All (n =1343)	4-6 years (n = 881)	7-9 years (n = 462)	All (n = 1093)	4-6 years (n = 667)	7-9 years (n = 426)	All (n = 1137)	4-6 years (n = 683)	7-9 years (n = 454)
	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %	Cumul. %
0	84.6	86.2	81.3	92.0	93.1	89.8	82.6	83.5	81.2	93.1	93.0	93.2
1	94.1	96.4	89.5	97.8	98.8	95.9	91.9	91.9	92.0	97.9	98.4	97.1
2	97.6	98.3	96.3	99.1	99.5	98.3	96.8	97.2	96.2	99.2	99.4	98.9
3	98.9	99.0	98.8	99.6	99.9	99.1	99.0	99.3	98.6	99.7		99.8
4	99.4	99.4	99.3	99.7		99.4	99.8	99.9	99.8	100.0	100.0	100.0
5	99.6	99.5	99.8	99.9		99.8	100.0	100.0	100.0			
6	99.8	99.8	100.0	99.9		100.0						
7	100.0	100.0		100.0	100.0				100.0			
8												
9												
10												

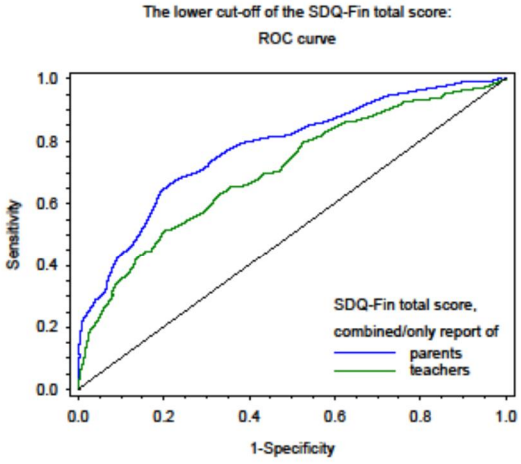
*The scores of two informants were combined into a single parent or a single teacher score.

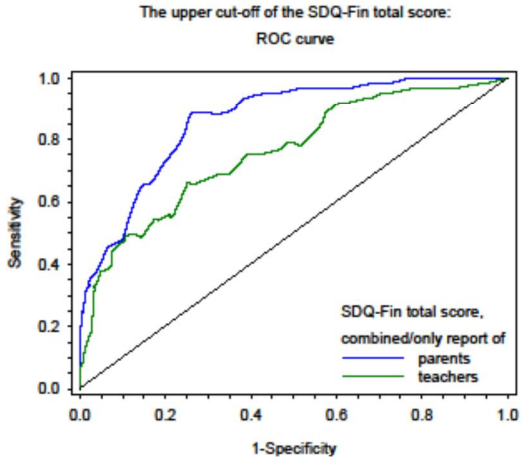
Appendix 11. Means and standard deviations of the SDQ-Fin scores rated by parents and teachers in 4-9-year-old children.

	4-6-year-olds		7-9-year-olds		Boys		Girls		Total sample	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parents' evaluation(*	(n = 1745)		(n = 890-889)		(n = 1291)		(n = 1344-1343)		(n = 2635-2634)	
Total score	6.72	3.80	6.36	4.39	7.23	4.28	5.99	3.63	6.60	4.01
Emotional symptoms	0.96	1.06	1.21	1.38	1.02	1.20	1.06	1.17	1.04	1.19
Conduct problems	1.64	1.30	1.39	1.43	1.72	1.43	1.40	1.25	1.56	1.35
Hyperactivity	2.61	1.88	2.39	2.01	2.91	2.06	2.17	1.72	2.54	1.93
Peer problems	1.52	1.19	1.38	1.26	1.58	1.33	1.36	1.10	1.47	1.22
Prosocial skills	7.52	1.63	7.64	1.66	7.24	1.68	7.86	1.54	7.56	1.64
Impact score	0.13	0.53	0.24	0.73	0.23	0.72	0.10	0.47	0.16	0.61
Teachers' evaluation(*	(n = 1357-1350)		(n = 885-880)		(n = 1099-1093)		(n = 1143-1137)		(n = 2242-2230)	
Total score	5.57	4.89	4.82	5.04	6.33	5.35	4.26	4.32	5.28	4.96
Emotional symptoms	0.85	1.21	0.81	1.38	0.79	1.19	0.88	1.35	0.84	1.28
Conduct problems	1.03	1.54	0.68	1.34	1.16	1.64	0.65	1.24	0.90	1.48
Hyperactivity	2.49	2.45	2.17	2.58	3.05	2.72	1.70	2.07	2.36	2.50
Peer problems	1.22	1.50	1.16	1.49	1.34	1.63	1.05	1.34	1.20	1.50
Prosocial skills	7.08	2.09	6.83	2.52	6.46	2.31	7.50	2.12	6.99	2.28
Impact score	0.16	0.55	0.21	0.65	0.29	0.73	0.10	0.40	0.18	0.59

*) for 4-6-year-olds the mean of the evaluations of two teachers, if available, otherwise and for 7-9-year-olds the only teacher's evaluation.

Appendix 12.


The lower cut-off of the SDQ-Fin total score:
ROC curve
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
SDQ-Fin total score,
combined/only report of
parents
teachers


The upper cut-off of the SDQ-Fin total score:
ROC curve
Sensitivity
1-Specificity
0.0
0.2
0.4
0.6
0.8
1.0
SDQ-Fin total score,
combined/only report of
parents
teachers

12 Original communications

informa
healthcare

Reliability of the Strengths and Difficulties Questionnaire among Finnish 4–9-year-old children

ANNE-MARI BORG, PÄLVI KAUKONEN, RAILI SALMELIN, MATTI JOUKAMAA, TUULA TAMMINEN

Borg A-M, Kaukonen P, Salmelin R, Joukamaa M, Tamminen T. Reliability of the Strengths and Difficulties Questionnaire among Finnish 4–9-year-old children. Nord J Psychiatry 2012; 66:403–413.

Background: Early recognition of children's mental health problems calls for structured methods in front line services. The Strengths and Difficulties Questionnaire (SDQ) is a commonly used short questionnaire in screening child's mental difficulties. *Aim:* To test the reliability and descriptive properties of the SDQ in a community sample of Finnish 4–9-year-old children ($n = 4178$). *Methods:* Both parents, two teachers in day-care or a teacher at school completed the SDQ. To control for possible bias, public health nurses rated their concern about every child's mental health, including non-participants. *Results:* The internal consistencies of the SDQ total score in all informants' reports were satisfactory to good. Agreement (Spearman rho) in total scores between parents was 0.65, between parent and teacher 0.43 and between two teachers in day-care 0.81. The stability in parent's reports over 12 weeks was good. The distributions of the informant-rated scores indicated significant and clinically important gender differences, and the 80th and 90th percentiles were generally below the international cut-off points. Public health nurses reported emotional or behavioural difficulties more commonly in non-participants (12%) than in participants (7%; $p < 0.001$). *Conclusions:* The results supported earlier findings of good internal consistency, inter-rater and cross-informant agreements and test–retest of the method. However, the gender and age of the child, the number of informants and cultural differences in reporting styles affected the results and thus confirmed the need to re-evaluate the SDQ in the culture and population in question.

• *Child psychiatry, Finnish, Reliability, Screening questionnaire, Strengths and Difficulties Questionnaire.*

Anne-Mari Borg, M.D., University of Tampere, School of Medicine, Department of Child Psychiatry, Lääketieteen yksikkö, 33014 Tampereen yliopisto, Finland, E-mail: anne-mari.borg@uta.fi
Accepted 20 January 2012.

Assessing a child's mental health in front line services calls for reliable, validated and feasible methods to be used in collaboration with the family, healthcare workers and the child's day-care or school professionals. Multiple informants increase opportunities to identify the children in need of support. Early recognition and referral to care are crucial because of the high prevalence of child psychiatric disorders (1–3), high continuity to adolescence and adulthood (4, 5), and insufficient referrals to child psychiatric care (2, 5). In primary healthcare, structured methods would ease and improve the assessment of the child's mental health in routine, medical check-ups.

The Strengths and Difficulties Questionnaire (SDQ) is a commonly used child's behavioural screening instrument, designed to both researchers and clinicians. It can be completed by parents and teachers, which adds the comparability of cross-informant agreement (6–8). Better capacity for discriminating caseness is achieved with the extended version, including additional information on the distress, social impairment and burden (7). The psychometric properties of the SDQ are confirmed in a large population sample of children aged 5–15 years (9). When comparing the SDQ with other similar measures, Achenbach's and Rutter's questionnaires, no differences in the predictive validity in discriminating low risk and high risk samples were found (6, 10). Extensive research in different cultures and language versions (2, 3, 11, 12–19) has made the SDQ the top-rated questionnaire in assessing child's mental health.

DOI: 10.3109/08039488.2012.660706

The prevalences of mental problems in children assessed by the SDQ as well as the reliability and validity of the method vary by gender, age and developmental stage of the child, across informants and across cultural contexts (12, 14, 20). Thus re-evaluating the psychometric properties of it in national and cultural contexts is appropriate.

The psychometric properties of the Finnish translation of the SDQ (SDQ-Fin) have been assessed in 7–15-year-old schoolchildren (16, 21). Because both the previous Finnish and international studies have focused on upper primary school-aged children and adolescents, it is important, for early recognition, to study the suitability of the method for younger children, through observing the effects of the gender of the child and informants' roles. The aim of this study was to test the psychometric properties and reliability of the SDQ-Fin in 4–9-year-old children visiting child health clinics and school healthcare clinics.

Material and Methods

Study design

The study was a part of a project "Developing children's mental health work, 2007–2009" in two hospital districts (Pirkanmaa and South Karelia, covering a population of 604,000) in Finland. Twenty-five municipalities (out of 35) participated in the study according to their willingness and resources. There were both urban and rural areas among the enrolled and non-enrolled municipalities. Altogether 154 child health clinics and school healthcare clinics participated. The local ethics committees approved the study. Informed consent was obtained for all participants.

The public health nurse sent study information, informed consent form and questionnaires to the parents prior to the child's regular check-up. Participating parents filled them in and asked two of the child's day-care personnel or teacher at school to complete the SDQ. The parents then returned all these papers when attending the medical check-up. The public health nurses also rated their concern for every child visiting the medical check-up, including, anonymously, those who refused to participate in the study. Within 2–17 weeks, parents of a subgroup of children were contacted again and one of them was asked to complete the SDQ for the second time by phone. Children selected to the subgroup were: 1) those who in the first round had scored at or above the international cut-off according to any informant, and 2) one child scoring below the cut-off with each informant for every two cases scoring above it (at the beginning of the study per every such case) being in the same age group and of the same gender (Fig. 1).

Sample

The sample consisted of 4–6-year-old preschoolers in child health clinics and 7–9-year-old children in school healthcare. Families not speaking Finnish were excluded from the study.

Altogether 4178 eligible children and their parents were invited to participate in the study; three-fifths were preschoolers, the rest school-aged (Fig. 1). The respective participation rates were 68.3% and 57.4%. The proportion of girls was 49.5%. The participants in Pirkanmaa accounted for 67.0%. For other socio-demographic data, see Table 1. No significant differences between the hospital districts were found regarding the gender of the child, the constitution of families, the number of siblings or family income. In Pirkanmaa, the parents of the preschool children had higher education level ($p < 0.001$) than in South Karelia. These differences were considered to enrich the data and consequently the districts were pooled in the analyses.

Almost all the participating children had at least one parent-rated and one teacher-rated SDQ usable in analyses (Fig. 1). The questionnaire from both parents was returned by three-quarters. In day-care, one teacher completed the SDQ for 76.3% and two teachers for 53.3% of the children. The public health nurses returned the Nurse's concern enquiry for 99.3% of participants and for all 1496 non-participants.

Altogether 592 parents correctly completed the SDQ twice, the median test–retest interval being 11.2 weeks.

Measures

The *SDQ* is a screening questionnaire for 3–16-year-olds to be completed by parents, teachers and 11–16-year-old children themselves (9, 13). It consists of 25 items on psychological attributes of the child's symptoms and social skills. The same form is suitable for parents and teachers. In this study the extended version of the method, including the impact supplement, i.e. global assessments of child's mental difficulties, was used (7). The SDQ-Fin has been carefully translated, back-translated and approved by the copyright owner (22).

The SDQ items form five subscales: emotional symptoms, conduct problems, hyperactivity/inattention difficulties, peer relationship problems and prosocial behaviour. Items are originally scored as 1 for "somewhat true" and, depending on the item, as 0 or 2 for "not true" or for "certainly true", and for analysis recoded as 0 to 2 with increasing severity. The scores from all the scales except the prosocial scale are summed to a total difficulties score ranging 0–40. On the prosocial scale, higher scores stand for desirable, positive behavioural traits, whereas on the other four scales higher scores reflect increasing risks of mental disorders (22).

Goodman (6) has presented the 80th and 90th percentiles as provisional cut-offs for "borderline" and "abnormal". Because the number and quality of psychiatric symptoms in this age group in Finland is insufficiently known, the lower cut-off was used.

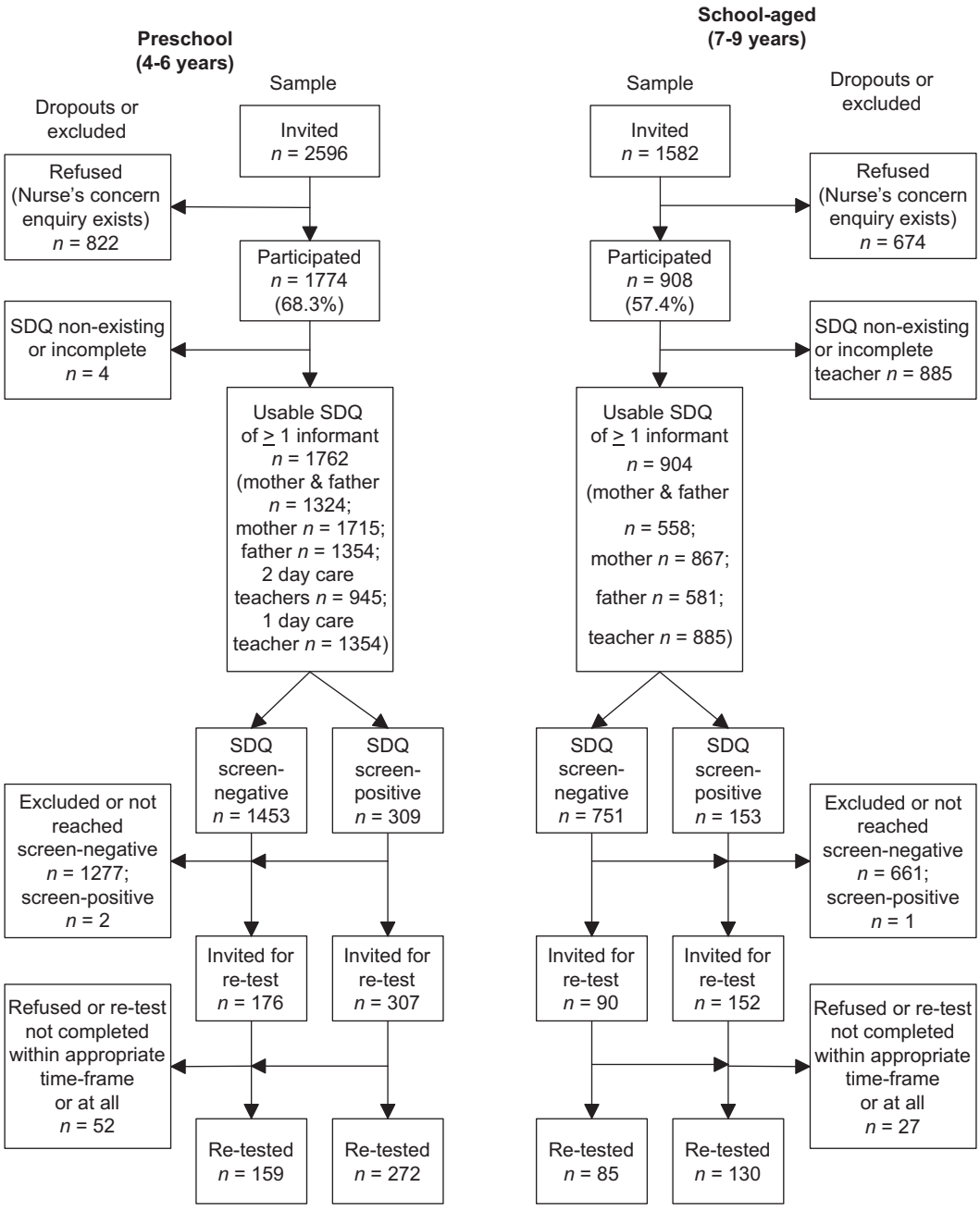


Fig. 1. Subject flow chart of the Strengths and Difficulties Questionnaire (SDQ) reliability study.

In the impact supplement, the respondent is asked whether the child has any emotional or behavioural difficulties and if so, about the duration or chronicity of the difficulties, overall distress, social impairment and burden to others. The items on overall distress and social impairment are scored from 0 to 2. The total impact score, ranging from 0 to 10 for parent-rated and from 0 to 6 for teacher-rated SDQ, can be classified as normal for a score of 0, as borderline for a score of 1 and as abnormal for a score of 2 or more (22).

In the *Nurse's concern enquiry*, public health nurses assessed, based on clinical evaluation, whether she thinks the child has any emotional or behavioural difficulties, irrespective of the first question of the parent and teacher SDQ impact supplement. Of the non-participating children, the nurse also recorded age and gender.

The *socio-demographic form* was developed for this study and completed by parents (Table 1).

Statistical methods

The internal consistency of the SDQ total and subscale scores was examined by Cronbach's alpha. The inter-rater (between teachers in preschool-aged children), cross-informant (between parents, between parents and teachers)

Table 1. Socio-demographic data of the participating children and families in preschool-aged (4–6-year-old) and school-aged (7–9-year-old) children.

	4–6-year-olds ($n = 1774$)		7–9-year-olds ($n = 908$)			All ($n = 2682$)	
	Mean	Range	Mean	Range	p	Mean	Range
Age (years)	5.1	4–7	8.0	6–10	< 0.001	6.1	4–10
	%		%			%	
Hospital district					< 0.001		
Pirkanmaa (PSHP)	71.9		57.5			67.0	
South Karelia (EKSHP)	28.1		42.5			33.0	
Gender					ns		
Boys	49.7		47.9			49.1	
Girls	50.3		52.1			50.9	
Family constitution					< 0.001		
Nuclear family	85.9		75.0			82.2	
Living in reconstituted family	4.3		8.6			6.0	
Living with one parent	7.4		13.8			9.6	
Adoption and foster parents	1.4		1.2			1.4	
Other	0.7		1.3			0.9	
Number of siblings					ns		
1	12.6		13.6			12.9	
2	52.5		44.6			48.8	
3	25.2		29.0			26.5	
4 or more	9.7		12.8			10.6	
Day-care arrangements					–		
Care in home	20.2						
Care in a small group of children	11.1						
Day-care centre	61.4						
Something else	7.8						
Mother's professional education					0.001		
No formal education/vocational courses	6.2		8.7			7.1	
Vocational/secondary school	26.6		31.9			28.4	
College/polytechnic school	39.9		37.4			39.0	
University	26.2		20.8			24.4	
Other	1.1		1.2			1.1	
Father's professional education					0.001		
No formal education/vocational courses	8.9		13.3			10.3	
Vocational/secondary school	37.3		41.8			38.7	
College/polytechnic school	28.3		25.3			27.3	
University	24.2		18.9			22.4	
Other	1.3		0.9			1.1	
Family income					ns		
Fully satisfactory	50.8		51.9			51.1	
Somewhat/fairly satisfactory	40.1		36.7			38.9	
Somewhat inadequate/inadequate	7.9		10.1			8.6	
Decline to answer	1.3		1.4			1.3	

and test–retest reliability, based on continuous scores were examined with Spearman's correlation coefficient.

The SDQ cross-informant reliability between parents and inter-rater reliability was good (see results). Therefore the scores of the mother and the father were combined into a single parent score, and those of the two day-care teachers to a single teacher score: if values of both informants were available, the combined score was the mean of the scores of the informants, otherwise the score of the only informant.

The scores of the SDQ subscales and the total difficulty score were non-normally distributed. Consequently, medians and the 80th and 90th percentiles were used. Because of non-normality, the significance of the SDQ score differences between groups defined by age and gender were examined by the Mann–Whitney U test; p-values < 0.05 were considered statistically significant.

The possible interaction of gender and age group was examined, for SDQ total score only, by two methods to see the effects of SDQ non-normality, yet enabling comparability with earlier studies. In Cox regression, all cases were considered having had the "event" and the inverse of the SDQ score represented the "survival time". The other method was univariate analysis of variance (ANOVA).

The distributions of categorized variables are described as percentages and those of continuous variables (except SDQ scores explained above) as medians and quartiles.

The statistical analyses were accomplished using SPSS v. 15.

Results

Comparison of the participating and non-participating children

Of the non-participants (n = 1,496), 50.9% were boys (vs. 49.1% of participants, n.s.). The proportion of preschoolers was 54.9% (vs. 68.3%, $p < 0.001$). The public health nurses considered 6.6% of participants and 12.0% of non-participants to have definite or severe difficulties. When the age groups were compared, the concern was definite or severe in 6.6% of the participants in both groups and in 12.7% of the younger and 11.1% of the older non-participants. The public health nurses rated their concern about the child to be minor in 17.2% of participants and 20.4% of non-participants in the total sample, and chose the option "cannot say" or "do not know" in 3.0% of participants and 1.1% of non-participants. The differences between the public health nurse's concerns about the participants and the non-participants were all statistically significant ($p < 0.001$).

Internal consistency of the SDQ-Fin

The internal consistency of the SDQ total difficulty score in the entire sample in all informants' reports was satisfactory to good (Cronbach's alpha = 0.77–0.86; Table 2). Highest internal consistencies ($\alpha = 0.68$–0.87) were found in teacher reports except in the impact score ($\alpha = 0.56$; parent reports $\alpha = 0.67$–0.72). Regarding the subscales in parent reports, the lowest internal consistencies were in peer problems ($\alpha = 0.46$–0.52) and in emotional symptoms ($\alpha = 0.53$–0.56) scales. Internal consistencies of the subscales except peer problems and prosocial skills were generally lower in the younger than in the older group as rated by parents. In all informants' reports, the internal consistencies for boys were higher than for girls except in emotional symptoms. This trend of higher internal consistencies for boys than for girls was also confirmed in parents' reports in both age groups, with certain exceptions.

Inter-rater reliability

The inter-rater reliability (Spearman's correlation) of the total difficulty scores was 0.81 between the day-care teachers ($n = 945$) (Table 3). The lowest correlation ($r = 0.59$) was found in the emotional symptoms score, whereas the correlations related to the other subscores were 0.67–0.81 and that of the impact score was 0.77.

Cross-informant agreement

The cross-informant agreements between mothers and fathers are shown in Table 3. In the total sample, the correlation for the total score was 0.65, between 0.50 and 0.65 for the subscales and 0.54 for the impact score. When comparing the average of parents' ratings (or the rating of a single parent) with the teacher rating, cross-informant correlations between parents were seen to be higher than parent and teacher correlations on all scales. Parents' evaluations were fairly similar in both age groups, except in the conduct and in the impact score in which their agreements were lower in younger than in older children. The agreement in preschool children between parent and teacher reports in the total score was higher ($r = 0.44$) than in school-aged children (0.39). The parents, and parents and teachers found the highest agreement in evaluating child's hyperactivity, and reached higher levels of agreement in evaluating boys than girls in all scores. The agreement in evaluating girls was lower between parents and teachers ($r = 0.21$–0.38) than between parents' reports (0.41–0.61).

Test–retest reliability

The stability in the parents' reports over 12 weeks according to Spearman correlation was 0.76 for total difficulties, 0.60–0.68 for emotional, conduct and peer problems and the prosocial subscore, 0.79 for hyperactivity, and 0.45 for impact scores. When the stability of ratings over the time interval of 2 weeks was compared with that of over 17 weeks, the impact scores correlation fell from 0.75 to 0.57, but unexpectedly no remarkable differences were found on other scales.

The descriptive properties of the distributions of the Finnish SDQ parent and teacher reports

Figure 2 shows the distributions of parent and teacher-reported total scores and subscale scores. Parents and teachers reported higher total scores for boys than girls. Parents rated higher total scores for boys and girls than did teachers. Concerning the subscales, both informants rated higher for boys than for girls in both age groups on the conduct and hyperactivity subscales and in the preschool children on the peer problems subscale. The results for the prosocial behaviour were reversed. Compared with the international borderline and clinical cut-off points, the 80th and 90th percentiles in the Finnish study were generally lower on all the SDQ scales except the parent-reported conduct and peer problems subscales in younger boys, parent-reported hyperactivity subscale in older boys and teacher-reported conduct subscale in younger boys. On the prosocial subscale, parents reported lower scores for school-aged boys and teachers reported higher scores for girls than the international cut-off points suggest.

Table 2. Internal consistency (Cronbach alpha, α) of the Finnish version of the Strengths and Difficulties Questionnaire (SDQ-Fin) total score, subscores and impact value according to parents' and teachers' ratings in preschool-aged (4–6-year-old) and school-aged (7–9-year-old) boys and girls.

	4–6-year-olds			7–9-year-olds			Total sample		
	Boys α	Girls α	All α	Boys α	Girls α	All α	Boys α	Girls α	All α
Mother's evaluation									
Scale scores	(*n* = 822–848)	(*n* = 838–860)	(*n* = 1660–1708)	(*n* = 402–414)	(*n* = 437–450)	(*n* = 839–864)	(*n* = 1224–1261)	(*n* = 1275–1308)	(*n* = 2499–2567)
Total score	0.79	0.71	0.76	0.81	0.79	0.81	0.80	0.74	0.78
Emotional symptoms	0.55	0.45	0.51	0.65	0.60	0.62	0.59	0.52	0.56
Conduct problems	0.58	0.50	0.55	0.69	0.62	0.66	0.62	0.55	0.59
Hyperactivity	0.79	0.73	0.77	0.80	0.77	0.80	0.79	0.74	0.78
Peer problems	0.57	0.42	0.51	0.59	0.49	0.55	0.57	0.44	0.52
Prosocial skills	0.69	0.66	0.69	0.66	0.69	0.69	0.68	0.67	0.69
Impact	(*n* = 236)	(*n* = 160)	(*n* = 396)	(*n* = 181)	(*n* = 119)	(*n* = 300)	(*n* = 417)	(*n* = 279)	(*n* = 696)
	0.68	0.52	0.64	0.66	0.75	0.70	0.68	0.66	0.67
Father's evaluation									
Scale scores	(*n* = 661–675)	(*n* = 662–676)	(*n* = 1323–1350)	(*n* = 269–275)	(*n* = 300–306)	(*n* = 569–580)	(*n* = 930–949)	(*n* = 962–982)	(*n* = 1892–1929)
Total score	0.78	0.72	0.76	0.79	0.79	0.79	0.78	0.74	0.77
Emotional symptoms	0.49	0.49	0.49	0.60	0.58	0.59	0.53	0.53	0.53
Conduct problems	0.57	0.52	0.55	0.63	0.58	0.61	0.59	0.54	0.57
Hyperactivity	0.77	0.72	0.75	0.79	0.75	0.78	0.77	0.73	0.76
Peer problems	0.51	0.39	0.46	0.47	0.39	0.43	0.51	0.39	0.46
Prosocial skills	0.72	0.69	0.71	0.68	0.63	0.66	0.71	0.67	0.70
Impact	(*n* = 174)	(*n* = 135)	(*n* = 309)	(*n* = 103)	(*n* = 66)	(*n* = 169)	(*n* = 277)	(*n* = 201)	(*n* = 478)
	0.75	0.69	0.73	0.73	0.62	0.69	0.74	0.68	0.72
Teacher's evaluation*									
Scale scores	(*n* = 659–668)	(*n* = 673–683)	(*n* = 1332–1350)	(*n* = 392–428)	(*n* = 409–455)	(*n* = 801–880)	(*n* = 1052–1096)	(*n* = 1083–1131)	(*n* = 2135–2230)
Total score	0.87	0.83	0.86	0.85	0.85	0.86	0.86	0.84	0.86
Emotional symptoms	0.55	0.71	0.65	0.74	0.70	0.72	0.65	0.70	0.68
Conduct problems	0.75	0.72	0.74	0.73	0.66	0.72	0.75	0.71	0.74
Hyperactivity	0.89	0.85	0.88	0.87	0.82	0.87	0.88	0.84	0.87
Peer problems	0.76	0.65	0.72	0.68	0.60	0.65	0.73	0.63	0.69
Prosocial skills	0.81	0.79	0.81	0.82	0.81	0.83	0.81	0.80	0.81
Impact	(*n* = 197)	(*n* = 115)	(*n* = 312)	(*n* = 156)	(*n* = 74)	(*n* = 230)	(*n* = 353)	(*n* = 189)	(*n* = 542)
	0.64	0.49	0.61	0.50	0.49	0.49	0.57	0.53	0.56

*For 4–6-year-olds the mean of the evaluations of two teachers, if available, otherwise and for 7–9-year-olds the only teacher's evaluation.

Table 3. Agreement (Spearman's rho, ρ) between both parents, between parents and teachers, and between two teachers in day-care in preschool-aged (4–6-year-old) and school-aged (7–9-year-old) boys and girls.

	4–6-year-olds								7–9-year-olds								Total sample							
	Mother–father (n = 1325–1322)			Parent–teacher* (n = 1346–1339)			Teacher 1–2 (n = 945–922)			Mother–father (n = 558–555)			Parent–teacher* (n = 871–865)			Mother–father (n = 1883–1874)			Parent–teacher* (n = 2217–2204)					
	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ	Boys ρ	Girls ρ	All ρ			
Total score	0.68	0.59	0.64	0.50	0.36	0.44	0.83	0.78	0.81	0.68	0.65	0.67	0.46	0.28	0.39	0.68	0.61	0.65	0.48	0.34	0.43			
Emotional symptoms	0.57	0.51	0.54	0.32	0.34	0.33	0.60	0.58	0.59	0.51	0.58	0.54	0.31	0.20	0.26	0.55	0.53	0.54	0.31	0.27	0.29			
Conduct problems	0.56	0.57	0.57	0.34	0.34	0.33	0.79	0.71	0.76	0.68	0.65	0.66	0.36	0.25	0.31	0.61	0.60	0.61	0.36	0.30	0.34			
Hyperactivity	0.66	0.60	0.64	0.52	0.38	0.47	0.83	0.77	0.81	0.69	0.62	0.67	0.49	0.35	0.46	0.67	0.61	0.65	0.51	0.38	0.47			
Peer problems	0.57	0.40	0.49	0.45	0.31	0.38	0.69	0.65	0.67	0.52	0.52	0.52	0.35	0.21	0.28	0.56	0.44	0.50	0.41	0.27	0.34			
Prosocial skills	0.58	0.50	0.55	0.30	0.26	0.31	0.73	0.64	0.71	0.60	0.49	0.56	0.18	0.15	0.21	0.59	0.50	0.56	0.25	0.21	0.26			
Impact score	0.52	0.21	0.42	0.31	0.15	0.27	0.79	0.72	0.77	0.74	0.66	0.71	0.34	0.28	0.33	0.60	0.41	0.54	0.32	0.21	0.30			

*For 4–6-year-olds the mean of the evaluations of two teachers, if available, otherwise and for 7–9-year-olds the only teacher's evaluation.

To enable comparison with earlier studies also, the means and standard deviations of the SDQ-Fin scores were calculated and are obtainable from the corresponding author.

The comparisons between the age groups revealed significant differences in all parent-rated scales and in all teacher-rated scales except in peer relationships and prosocial skills. The same was true of parents' and teachers' evaluations of girls in these age groups. Among boys, the differences between age groups were non-significant on the parent-rated hyperactivity and prosocial skills subscale, and in teacher-rated total score, hyperactivity and peer relationship subscale but significant on all other scales. Comparison of boys and girls within age groups and regardless of age group or informant revealed significant differences on all scales except the emotional symptoms subscale (Fig. 2).

According to both Cox regression and ANOVA, the interaction of age group and gender on parent-rated SDQ total score was non-significant (Table 4). Both models showed an independent effect of gender, boys having significantly higher scores. Age group as a whole had an independent effect, too, although the size of it differed in the models. In the case of teacher-rated total score, age group, gender and their interaction were significant in both models. School-aged girls had the lowest scores, while boys had the highest scores regardless of age group (Fig. 2).

Conclusion

The main results of the present study confirmed good reliability properties of the SDQ-Fin in the community sample of 4–9-year old children tested by internal consistencies, cross-informant agreements and test–retest reliability. The distributions of the scores of the SDQ-Fin indicated significant and clinically important gender differences.

In the present study, the internal consistencies of the total score in all informants' reports was good to excellent ($\alpha = 0.77$–0.86) falling between the earlier Finnish results ($\alpha = 0.71$) in 7–15-year-old children (21) and previous British, Dutch, German and Swedish results ($\alpha = 0.80$–0.87; 9, 23–25). In concordance with earlier studies, the internal consistencies in teacher reports were higher than in parent reports (9, 26, 27) and also reports on boys had higher values than those on girls (25). Internal consistencies of the emotional subscale rated by parents were lower than in any of the 26 studies reviewed in 2010 (28). The items of the emotional symptoms may thus reflect a somewhat different construct among Finnish parents. Also peer relationships subscale had low alphas and one explanation for this may be that adults find it difficult to observe child's behaviour in these areas. From this perspective, it is more reliable to use

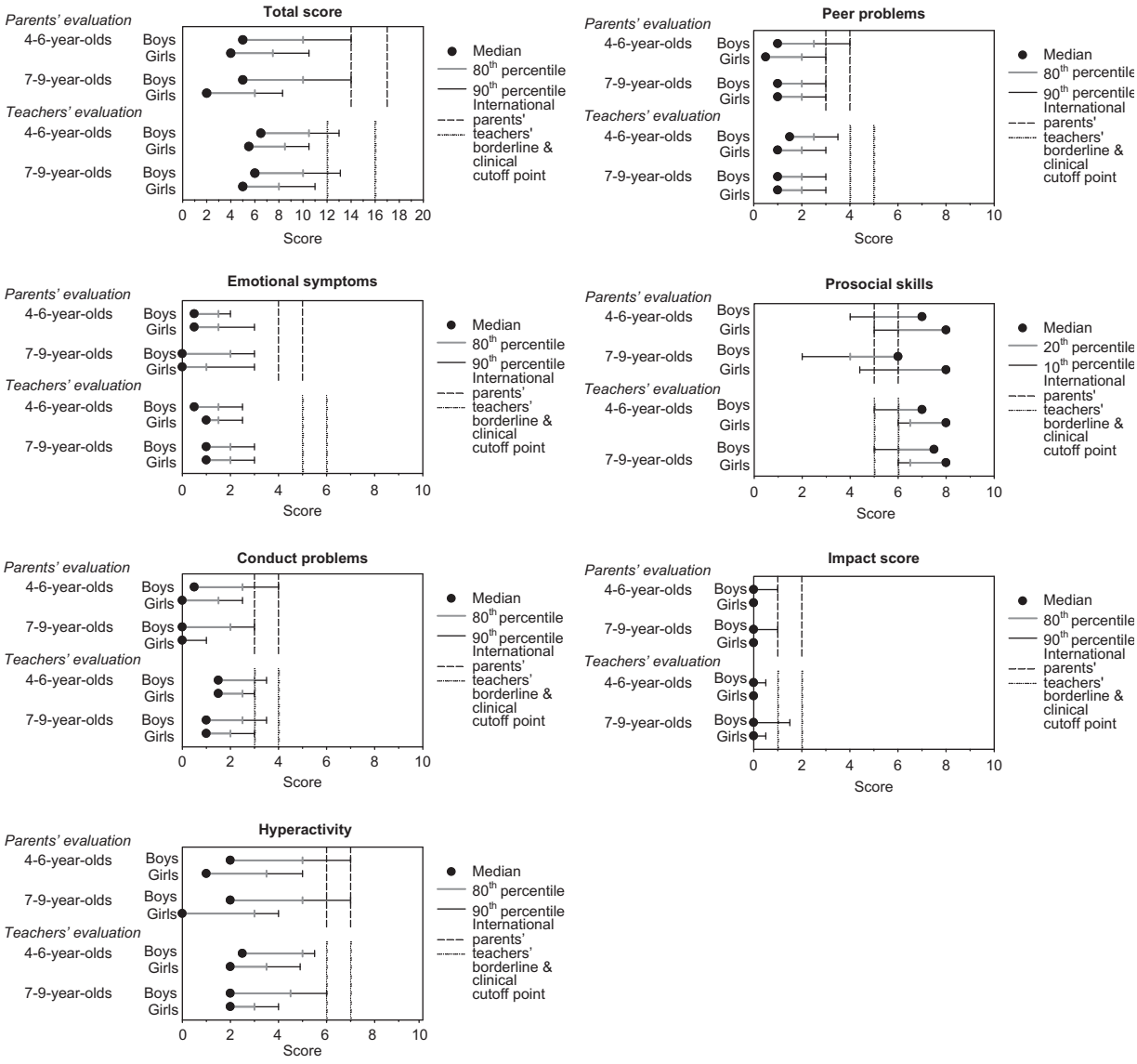


Fig. 2. The Strengths and Difficulties Questionnaire (SDQ) total and subscale as well as impact score medians and quartiles used in defining cut-off points, for preschool-aged (4–6-year-old) and school-aged (7–9-year-old) Finnish boys and girls according to parents' and teachers' reports. Parent scores are averages of scores reported by the mother and the father, if both were available, otherwise scores of the only parent. For 4–6-year-olds, teachers' scores are averages of scores reported by two day-care teachers, if both were available, otherwise scores of the only teacher. Number of cases: in parents' reports 4–6-year-old boys $n = 864$, girls $n = 881$, 7–9-year-old boys $n = 427$, girls $n = 462$–463; in teachers' reports 4–6-year-old boys $n = 670$, girls $n = 685$–687, 7–9-year-old boys $n = 427$–429, girls $n = 453$–456.

and interpret the total score than the subscales separately in screening.

All informants reached higher levels of agreement in evaluating boys than girls in all scores but previous studies with similar study design were hard to find. All informants also found the best agreement in evaluating child's hyperactivity, which is in accordance with earlier studies (9, 21, 26, 27). Parents and teachers found higher levels of agreement in evaluating preschool than school-aged children, possibly because of children's less context-dependent behaviour, and parents' and day-care teachers' more close collaboration in preschool stage. The correlations between parents' and teachers' evaluations in total scores were somewhat lower in the present study than in earlier studies (9, 21, 27) but not as low as in a Danish study with 5–7-year-olds (11). Inter-rater agreement between two teachers was excellent ($r = 0.81$) and in this respect the SDQ is a reliable method in the day-care.

The activity of both parents to participate was unexpectedly high (73%) as, according to clinical impressions,

Table 4. Results of the Cox regression and univariate analysis of variance (ANOVA) used to examine the interaction of age group and gender on parent and teacher-rated Strengths and Difficulties Questionnaire (SDQ) total difficulty scores.

	Cox regression			Univariate ANOVA		
	p	OR	95% CI	p	B	95% CI
Parent-rated SDQ						
Age group	0.002			0.039		
4–6-year-olds (vs. 7–9-year-olds)		1.20	1.07–1.34		0.29	−0.16 to 0.73
Gender	0.001			<0.001		
Boys (vs. girls)		1.25	1.10–1.43		1.17	0.65–1.69
Age group and gender interaction	0.930			0.764		
Teacher-rated SDQ						
Age group	<0.001			0.001		
4–6-year-olds (vs. 7–9-year-olds)		1.44	1.28–1.62		1.27	0.69–1.84
Gender	<0.001			<0.001		
Boys (vs. girls)		1.67	1.46–1.90		2.72	2.08–3.35
Age group and gender interaction	0.005			0.010		

OR, odds ratio; CI confidence interval.

fathers often have a peripheral role when evaluating the well-being of the child with professionals. The agreement in total scores between parents was good in both age groups and genders. These results support and encourage professionals to invite both parents to co-operate. This is warranted especially when evaluating girls, where agreement between parents and teachers was harder to find than in the case of boys. Yet, it should be noted that the informants' differing views of the child in everyday situations can both enrich the information and challenge the attempts to find a common understanding of the mental health of the child.

The stability of the SDQ was good in parents' reports in test–retest analyses. In previous studies where the time interval was shorter than in the present study the parents' test–retest correlations were higher (7, 24) and vice versa (9). However, this study also suggests high stability of the method in the long term in line with the findings of an Australian community study (29).

Clinically important and statistically significant differences were found between genders in the distributions of the scores of the SDQ-Fin in accordance with earlier studies (11, 14, 21, 22, 24, 25, 29). On average, parents gave higher scores than teachers and both sets of informants reported higher scores for boys than for girls. In addition, girls were reported to have fewer problems than boys in their prosocial behaviour. Only on the emotional symptoms subscale did parents and teachers not report significant gender differences.

The results of the SDQ showed that the gender of the child has more clinical importance than the age group, but the meaning of age must also be noted. Fewest difficulties of all were reported for school-aged girls, except in emotional symptoms.

The 80th and 90th percentiles of the SDQ scores were generally lower (with certain exceptions) than the respective international ones (22). In addition, the mean scores of the SDQ scales were found to be 1–2 points lower than in British, American, Australian and German community samples (5, 22, 29), but on the other hand the parent-rated mean scores were found to be slightly higher than in Nordic and Dutch studies (11, 14, 21, 24). The present study replicates the findings of parents reporting lower scores for hyperactivity problems in the Nordic countries than in other cultures (11, 14). Both Finnish parents and teachers gave lower scores for emotional symptoms and prosocial behaviour than is previously reported in all the above-mentioned studies. When generalizing the results of prevalences according to the SDQ, the possible differences in recognizing and reporting children's psychosocial difficulties between Finnish and other cultures have also to be taken into account.

To control and reduce the possible effects of the moderate participating rates, information on the mental health of the both participants and non-participants was collected. The public health nurses reported that non-participants had more commonly definite or severe difficulties than the participants (12% vs. 7%). Thus the descriptive results of the study may be underestimates, especially in the school-aged sample. Potential methodological limitations in the study design should also be noticed. First, teachers and day-care personnel may have answered both more thoroughly and more cautiously knowing that the parents could see their answers. Secondly, potential differences in parents' answers between the first and second completion could be affected by the rather wide variation in the time interval between the completions and by the different completion methods. The strengths of this study are in the multifaceted design, the large sample size and fairly homogenous age distribution of children for the study aims. The high

participation rates of the fathers gave further information on the agreement of parents in evaluating their child's mental health.

Further research is needed to assess prevalence of psychiatric disorders in this age group of Finnish children: is it true that children do not suffer from emotional symptoms or is it more a matter of how parents and teachers recognize and report their evaluations (30). A validation study of the SDQ-Fin against a diagnostic assessment method is needed to adjust the optimal cut-off points for this age group. It is ethical and efficient to recognize the children in need of psychosocial support as early as possible. The results of this study support the position of the SDQ for a screening method in primary healthcare.

Although this study supported earlier findings of good internal consistency, inter-rater agreement and test–retest of the SDQ, noteworthy differences from earlier studies were also found both in the reliability and descriptive properties of the method in this young age group of Finnish children. At least the gender and the age of the child, the number of informants and cultural differences in evaluating and reporting styles affect the results of the reliability of the method and observed prevalence of the mental health problems in children. These findings confirm the need to re-evaluate the questionnaire in the culture and population in question.

Acknowledgements—We thank the participating families, the public health nurses and the teachers in day-care and in schools participating in the study; the project co-ordinator Sari Miettinen at the Child Psychiatric Clinic, Tampere University Hospital, for collaboration in data collection. The data were collected on the project "Developing children's mental health work, 2007–2009", funded by the hospital districts of Pirkanmaa and South-Karelia and the Ministry of Social Affairs and Health, Finland. This study was supported by grants from the Finnish Child Psychiatric Research Foundation, the Medical Research Foundation of Tampere University Hospital and the Emil Aaltonen Foundation.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Costello EJ, Egger H, Angold A. 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. J Am Acad Child Adolesc Psychiatry 2005;44:972–986.
2. Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. Pediatrics 2010;125:75–81.
3. Sanne B, Torsheim T, Heiervang E, Stormark KM. The Strengths and Difficulties Questionnaire in the Bergen Child Study: A conceptually and methodically motivated structural analysis. Psychol Assess 2009;21:352–364.
4. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003;60:837–844.
5. Ravens-Sieberer U, Wille N, Erhart M, Bettge S, Wittchen HU, Rothenberger A, et al. Prevalence of mental health problems among children and adolescents in Germany: Results of the BELLA study within the National Health Interview and Examination Survey. Eur Child Adolesc Psychiatry 2008;17 Suppl 1:22–33.
6. Goodman R. The Strengths and Difficulties Questionnaire: A research note. J Child Psychol Psychiatry 1997;38:581–586.
7. Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. J Child Psychol Psychiatry 1999;40: 791–799.
8. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. Eur Child Adolesc Psychiatry 1998;7:125–130.
9. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry 2001;40: 1337–1345.
10. Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? J Abnorm Child Psychol 1999;27:17–24.
11. Elberling H, Linneberg A, Olsen EM, Goodman R, Skovgaard AM. The prevalence of SDQ-measured mental health problems at age 5–7 years and identification of predictors from birth to preschool age in a Danish birth cohort: The Copenhagen Child Cohort 2000. Eur Child Adolesc Psychiatry 2010;19:725–735.
12. Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. J Am Acad Child Adolesc Psychiatry 2003;42:1203–1211.
13. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. Br J Psychiatry 2000;177:534–539.
14. Heiervang E, Goodman A, Goodman R. The Nordic advantage in child mental health: Separating health differences from reporting style in a cross-cultural comparison of psychopathology. J Child Psychol Psychiatry 2008;49:678–685.
15. Klasen H, Woerner W, Wolke D, Meyer R, Overmeyer S, Kaschnitz W, et al. Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. Eur Child Adolesc Psychiatry 2000;9:271–276.
16. Koskelainen M, Sourander A, Vauras M. Self-reported strengths and difficulties in a community sample of Finnish adolescents. Eur Child Adolesc Psychiatry 2001;10:180–185.
17. Marzocchi GM, Capron C, Di Pietro M, Duran Tauleria E, Duyme M, Frigerio A, et al. The use of the Strengths and Difficulties Questionnaire (SDQ) in Southern European countries. Eur Child Adolesc Psychiatry 2004;13 Suppl 2:II40–6.
18. Obel C, Heiervang E, Rodriguez A, Heyerdahl S, Smedje H, Sourander A, et al. The Strengths and Difficulties Questionnaire in the Nordic countries. Eur Child Adolesc Psychiatry 2004;13 Suppl 2:II32–9.
19. Woerner W, Fleitlich-Bilyk B, Martinussen R, Fletcher J, Cucchiaro G, Dalgalarrondo P, et al. The Strengths and Difficulties Questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. Eur Child Adolesc Psychiatry 2004;13 Suppl 2: II47–54.
20. Skovgaard AM, Houmann T, Christiansen E, Landorph S, Jorgensen T, CCC 2000 Study Team, et al. The prevalence of mental health problems in children 1(1/2) years of age—The Copenhagen Child Cohort 2000. J Child Psychol Psychiatry 2007;48:62–70.
21. Koskelainen M, Sourander A, Kaljonen A. The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. Eur Child Adolesc Psychiatry 2000;9:277–284.
22. Youthinmind. 2010; Accessed April/10, 2010.
23. Malmberg M, Rydell AM, Smedje H. Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-Swe). Nord J Psychiatry 2003;57:357–363.
24. Muris P, Meesters C, van den Berg F. The Strengths and Difficulties Questionnaire (SDQ)—Further evidence for its reliability and

validity in a community sample of Dutch children and adolescents. Eur Child Adolesc Psychiatry 2003;12:1–8.
25. Rothenberger A, Becker A, Erhart M, Wille N, Ravens-Sieberer U, BELLA study group. Psychometric properties of the parent strengths and difficulties questionnaire in the general population of German children and adolescents: Results of the BELLA study. Eur Child Adolesc Psychiatry 2008;17 Suppl 1:99–105.
26. Van Leeuwen K, Bosmans G, De Medts L, Braet C. The Strengths and Difficulties Questionnaire in a community sample of young children in Flanders. Eur J Psychol Assess 2006; 22:189–197.
27. van Widenfelt BM, Goedhart AW, Treffers PD, Goodman R. Dutch version of the Strengths and Difficulties Questionnaire (SDQ). Eur Child Adolesc Psychiatry 2003;12:281–289.
28. Stone LL, Otten R, Engels RC, Vermulst AA, Janssens JM. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: A review. Clin Child Fam Psychol Rev 2010;13:254–274.
29. Hawes DJ, Dadds MR. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. Aust N Z J Psychiatry 2004;38:644–651.
30. Myers K, Winters NC. Ten-year review of rating scales. II: Scales for internalizing disorders. J Am Acad Child Adolesc Psychiatry 2002;41:634–659.

Anne-Mari Borg, M.D., University of Tampere, School of Medicine, Department of Child Psychiatry, Lääketieteen yksikkö, 33014 Tampereen yliopisto, Finland; Tampere University Hospital, Department of Child Psychiatry, Tampereen yliopistollinen sairaala, Lastenpsykiatrian vastuualue, PL 2000, 33521 Tampere, Finland.
Pälvi Kaukonen, M.D., Tampere University Hospital, Department of Child Psychiatry, Tampereen yliopistollinen sairaala, Lastenpsykiatrian vastuualue, PL 2000, 33521 Tampere, Finland.
Raili Salmelin, M.Sc., Ph.D., University of Tampere, School of Health Sciences, Terveystieteiden yksikkö, 33014 Tampereen yliopisto, Finland; Tampere University Hospital, Department of Child Psychiatry, Tampereen yliopistollinen sairaala, Lastenpsykiatrian vastuualue, PL 2000, 33521 Tampere, Finland.
Matti Joukamaa, M.D., Ph.D., University of Tampere, School of Health Sciences, Terveystieteiden yksikkö, 33014 Tampereen yliopisto, Finland; Tampere University Hospital, Department of Psychiatry, Tampereen yliopistollinen sairaala, Psykiatrian vastuualue, PL 2000, 33521 Tampere, Finland.
Tuula Tamminen, M.D., Ph.D., University of Tampere, School of Medicine, Department of Child Psychiatry, Lääketieteen yksikkö, 33014 Tampereen yliopisto, Finland; Tampere University Hospital, Department of Child Psychiatry, Tampereen yliopistollinen sairaala, Lastenpsykiatrian vastuualue, PL 2000, 33521 Tampere, Finland.

informa
healthcare

Finnish norms for young children on the Strengths and Difficulties Questionnaire

ANNE-MARI BORG, PÄLVI KAUKONEN, MATTI JOUKAMAA, TUULA TAMMINEN

Borg A-M, Kaukonen P, Joukamaa M, Tamminen T. Finnish norms for young children on the Strengths and Difficulties Questionnaire. Nord J Psychiatry 2014;68:433–442.

Background: Early recognition of children's mental health problems is crucial. Although the Strengths and Difficulties Questionnaire (SDQ) is a commonly used screening method, further research is needed on its validity and norms for young children. *Aims:* The aims of the study were to confirm the adjusted lower (normal/borderline) and upper (borderline/abnormal) cut-offs for the SDQ in a Finnish community sample of 4–9-year-old children, and to explore the SDQ's ability to identify the children with mental health problems. *Methods:* Parents and teachers completed the SDQs ($n = 2666$). The Development and Well-Being Assessment (DAWBA) was administered to parents and teachers of 646 children. *Results:* The overall participation rate was 57%. The suggested cut-offs for the SDQ total difficulties scale rated by parents and teachers were 2–5 points lower than the corresponding published British norms. The sensitivity for the total score normal/borderline cut-off (9/10) was 76% in the parent and 66% in the teacher reports and for the borderline/abnormal cut-off (11/12) 90% and 70% respectively. The respective specificity values were 69%, 63%, 74% and 66%. The area under curve (AUC) values of the higher cut-offs were good for parent (0.87) and satisfactory for teacher rated (0.76) total scores. The presence of a DAWBA-rater assigned diagnosis in the abnormal group compared with the normal group was sixfold in the parent and threefold in the teacher reported SDQs. *Conclusions:* The suggested cut-offs were clearly lower than the British norms. Yet the properties of the method's discriminative validity were acceptable. Population specific norms, taking into account both the culture and children's age, seem necessary for screening and for international comparisons of the method's validity properties.

• *Child psychiatry, Cut-off points, Finnish, Screening questionnaire, Strengths and Difficulties Questionnaire, Validity.*

Anne-Mari Borg, M.D., Department of Child Psychiatry, School of Medicine, University of Tampere; Lääketieteen yksikkö, 33014 Tampereen yliopisto, Finland, E-mail: anne-mari.borg@uta.fi; Accepted 4 October 2013.

Assessing children's mental health is central to health monitoring in primary healthcare. Mental health problems are common among children of all ages (1–7). Primary care workers face substantial challenges and responsibility to identify children in need of support. Consistent, structured questionnaires may aid in recognizing the children at risk for psychopathology.

The Strengths and Difficulties Questionnaire (SDQ) is a well-documented short enquiry having adequate psychometric qualities (8). The SDQ is commonly used in both research settings and in screening 4–16-year-old children's mental health in community and clinical settings in different countries and cultures (2, 9–15).

Even though the SDQ is extensively used in research and in numerous language versions, the norms of the method are only available for six countries (16). However, it is not possible to estimate the prevalence of child mental disorders cross-nationally without population specific norms (17). To adjust the norms for a method often requires diagnostic assessments.

Earlier studies on the SDQ have mostly been carried out in samples of upper primary school-aged children and adolescents, and therefore further research focusing on younger age groups is needed. It is also noteworthy, that in an overview of the psychometric properties of the SDQ for 4–12-year-olds (18), only few study designs involved a diagnostic assessment as a gold standard.

The earlier Finnish results on the good reliability properties of the SDQ support its position as a screening

DOI: 10.3109/08039488.2013.853833

method in community samples of 4–15-year-old children (19, 20). The concurrent validity of the SDQ against the Achenbach System of Empirically Based Assessment (ASEBA) questionnaires (21) in 15-year-olds has been found to be satisfactory (19). However, there is a lack of a Finnish validity study on younger children and against a diagnostic assessment method. The aims of the present study were: 1) to confirm the adjusted lower and upper cut-offs of the SDQ in Finnish 4–9-year-old children, and 2) to explore the capacity of the SDQ to identify the children suffering from mental health problems or disorders among this age group by comparing the parent- and teacher-rated SDQs with a diagnostic assessment method.

Materials and Methods

Study design

The study was a part of a project "Developing children's mental health work, 2007–2009", conducted in two hospital districts in Finland. Altogether 154 child health clinics and school healthcare clinics in 25 municipalities participated. The local ethics committees approved the study. Informed consent was obtained from all participating parents.

In the two-phase study design, parents and teachers in the day-care or at school first completed the SDQ questionnaires in the context of the 4–9-old children's regular medical check-ups. Families not speaking Finnish were excluded from the study. The design and the sample of the first phase of the study have been described in more detail elsewhere (20).

For the second phase, the SDQs were used to divide the children into screen-positive (scoring at or above the British 80th percentile cut-off, according to any informant) and screen-negative (scoring below the British 80th percentile cut-off, according to every informant) subgroups. Every parent of a screen-positive child was invited to the Development and Well-Being Assessment (DAWBA) interview. For every two screen-positive cases (at the beginning of the study for every such case), a parent of a screen-negative child, matched for age group and gender, was invited to the DAWBA. One parent per child was interviewed by phone 2–17 weeks after the child's check-up visit. In 92% of the cases, the same parent (mother or father) as the one who completed the initial SDQ in the first phase was interviewed in the second phase. With the parent's permission the child's teacher in day-care or at school was also asked to complete DAWBA as a questionnaire.

Sample

Altogether 4178 eligible children (49.5% girls) and their parents were invited to participate in the study, and the participation rate in the first phase was 63.8% ($n = 2666$) (Fig. 1). Of these, 17.3% had scored over the 80th percentile British cut-off in the parent or teacher report or both, the respective individual proportions being 9.4% and 13.0%. In the subgroup of the SDQ screen-positives, DAWBA information was obtained from 93% (431/462) of the children. In the subgroup of those screen-negatives who were invited to the interview, DAWBA information was obtained from 81% (215/266) of the children. Altogether 647 parents were interviewed, of whom 67% had preschool-aged children. After the interview, one family decided to withdraw from the second stage of the study. Consequently, the participation rate in the second phase was 89% (646/728). The overall participation rate was 57%. Ninety-one per cent of the DAWBA interviewed respondents were mothers, 7% were fathers and 2% some other person. Of the participants, 75% ($n = 486$) had a teacher's report available. The proportion of boys among DAWBA participants was 66%. The study sample is not considered an epidemiological one.

Measures

The SDQ is a screening questionnaire for 3–16-year-olds to be completed by parents, teachers and 11–16-year-old children themselves (8, 22, 23). In this study, the Finnish version of the method, including the impact supplement, was used (16).

The symptom part of the SDQ consists of 25 items forming five subscales: emotional symptoms, conduct problems, hyperactivity/inattention difficulties, peer relationship problems and prosocial behaviour. The items were originally scored as 1 for "somewhat true" and, depending on the item's phrasing, as 0 or 2 for "not true" or for "certainly true" Before scoring all items were reversed to be positively phrased. The scores from all the scales except the prosocial scale are summed to a total difficulties score ranging 0–40.

Goodman (22) has proposed the 80th and 90th percentile as provisional cut-offs for "borderline" and "abnormal". To indicate a screen-positive case, the 80th British percentile was used in this study because the particular aim was to cover possible cases extensively, as the number and quality of psychiatric symptoms in this age group in Finland is insufficiently known.

In the impact supplement, the respondent is asked whether the child has any emotional or behavioural difficulties and if so, about the duration or chronicity of the difficulties, overall distress, social impairment, and burden to others. The items on overall distress and social impairment are scored from 0 to 2. The total impact score, ranging from 0 to 10 for parent-rated and from 0 to 6 for teacher-rated SDQ, can be classified as normal for a score of 0, as borderline for a score of 1 and as abnormal for a score of 2 or more (16).

The diagnostic assessment was conducted using the Finnish version of the DAWBA method (9). The semi-structured interview can be administered to parents of

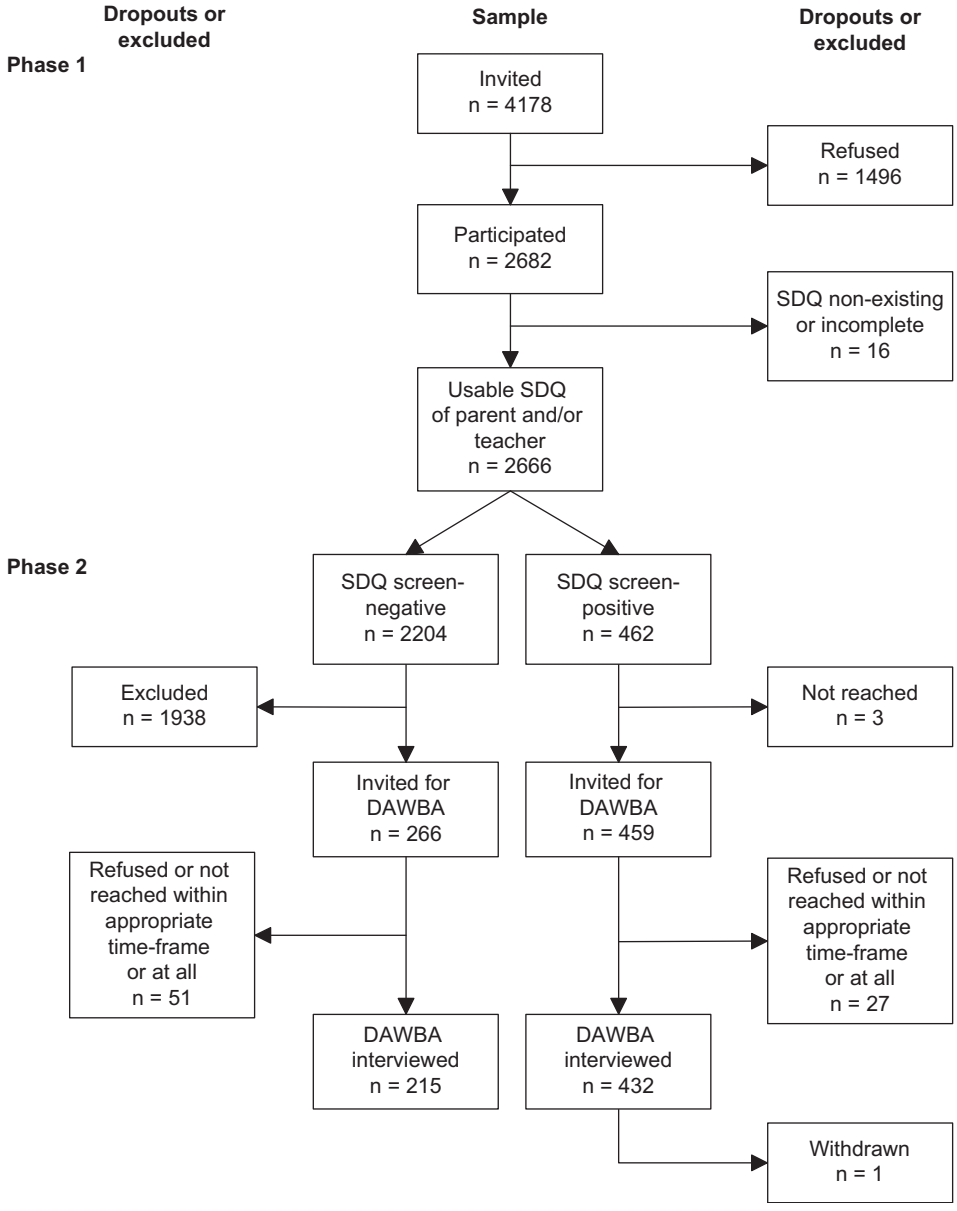


Fig. 1. Subject flow chart of the validity study of the Strengths and Difficulties Questionnaire (SDQ) against the Development and Well-Being Assessment (DAWBA) interview in 4–9-year old Finnish children.

children aged 5–17-years and to children over 11-years, and a briefer questionnaire version can be administered to teachers.

The interview consists of structured questions about psychiatric symptoms and their impact, and, if definite symptoms are identified, parents are asked to describe the problems in more detail by open-ended questions. The structured questions in separate sections cover most child psychiatric disorders and follow closely the diagnostic criteria according to the ICD-10 (24) and DSM-IV (25).

Based on the structured questions, the DAWBA program assigns each child to a level of ordered-categorical measures, called bands, which represent the prevalence of any diagnosis and several specific diagnosis categories in epidemiological samples. Each band has up to six levels ($<0.1\%$, $\approx 0.5\%$, $\approx 3\%$, $\approx 15\%$, $\approx 50\%$, $> 70\%$), representing the proportion of children in the respective band level having the particular disorder. Bands are defined separately for parent, teacher, child and multi-informant assessments (26). To decide on definitive diagnoses a clinical rater then reviews all relevant information: the structured, closed and open accounts of all available informants and the prevalence level produced by computer algorithms. The categorization of the

predictive measure (prevalence of any disorder) offered to the rater slightly differs from the band categorization, being < 1% (very low), < 5% (low), 20% (moderate), 75% (high) (27).

The DAWBA has shown satisfactory inter-rater reliability, and discrimination and predictive validity (9, 14, 28–30).

One of the authors (A-M B), blind to the SDQ screening status of the children, reviewed all the interviews and assigned the diagnoses according to ICD-10. The rater was trained by practising the cases in the training manual (27) and participating in an intensive 2-day training course arranged by an adolescent psychiatric team (Turku University Hospital), experienced with the DAWBA. During the rating process, the rater had regular supervision meetings with the professor of child psychiatry (TT) and also consulted the contributor of the method. When the diagnoses were uncertain, a consensus diagnosis was obtained by a consultation group of four experienced child psychiatrists. The frequency of diagnoses set by the rater was compared with the computer-predicted level of prevalence of any disorder. The associations were statistically significant ($P < 0.001$) between all pairs of the following groups: in the group of very low (< 5%) prevalence 11% of children had been assigned diagnoses, in the group of moderate (≥ 20%) prevalence 45% and in the group of high (≥ 75%) prevalence 93%.

Statistical analyses

According to an earlier study of the same study sample, the SDQ inter-rater reliability on the one hand between parents and on the other hand between the two day-care teachers was good (20). Therefore the scores of the mother and the father were combined into a single parent score, and those of the two day-care teachers to a single teacher score: if values of both informants were available, the combined score was the mean of the scores of the informants, otherwise the score of the only informant.

Two measures based on the DAWBA were used as the gold standard when examining the SDQ. The first measure, the computer-generated predictive measure, i.e. the prevalence of any disorder was utilized: 1) in assessing the discriminating capacity of the SDQ by receiver operating characteristics (ROC) analysis, 2) in exploring the SDQ's lower and upper cut-offs by ROC analysis and 3) in evaluating the sensitivity and specificity values resulting from the various adjusted cut-off candidates. The four-level measure offered to the clinical rater was used instead of the respective six-level band because, according to Goodman (personal communication 5 April 2007), the latter is correct for individual disorders but too low for any disorder at all. The number of categories was reduced to three (< 5% prevalence = low probability of disorder, 20% = moderate, 75% = high) in order to get large enough case numbers per category for statistical analyses. The distribution of cases into these categories was almost identical to the respective, similarly recategorized band variable.

The second gold standard DAWBA measure, the existence of any DAWBA-rater assigned diagnoses, a dichotomous variable, was exploited 1) as an extra control test for adjusting the Finnish upper cut-offs of the SDQ and 2) in estimating the diagnostic discriminative capacity of the method in the study sample. In addition, the DAWBA-rater assigned diagnoses were utilized in assessing the frequencies of psychiatric diagnoses in the study sample.

The discriminating capacity of SDQ total score and subscale scores was examined by area under curve (AUC) in ROC analysis. The analysis were performed separately for gold standard cut-offs discriminating cases with low disorder probability from those of moderate or high probability and cases with high disorder probability from those with moderate or low probability. AUC values of 0.90–1 were considered excellent, values of 0.80–0.90 good, 0.70–0.80 fair, 0.60–0.70 poor and values of 0.60 downwards were considered fail (31). The possibility of dose–response association between the three computer-generated disorder probability categories and SDQ total score was examined by curve estimation.

The parent and teacher SDQ total score's lower and upper cut-offs for Finnish population were sought by two methods: defining the 80th and 90th percentiles of the SDQ scores and by ROC analysis as described above. The sensitivity and specificity resulting from the various cut-off candidates were evaluated against the respective dichotomized computer-generated disorder probabilities (low vs. moderate or high for the lower SDQ cut-off and low or moderate vs. high for the higher cut-off). Regarding the upper cut-off, the evaluation was made even against the existence of any DAWBA-rater assigned diagnosis. The proportions of children with any DAWBA-rater assigned diagnosis or subgroup of diagnoses in the categories defined by the suggested Finnish SDQ cut-offs were examined to estimate diagnostic discriminative capacity.

The results are presented without stratifying by gender or age group as the stratified analyses did not alter the findings. The statistical analyses were accomplished with SPSS v. 19.

Results

The distributions of parent and teacher reported Finnish SDQ scores (of the first phase sample of the study) in comparison with the British cut-offs have been presented earlier (20).

In the present study, among the children assessed by the DAWBA ($n = 646$) the frequencies of the computer-generated probabilities of any disorder were: 71% ($n = 456$) had a low probability (prevalence in the category < 5%) for any diagnosis, 18% ($n = 116$) had a moderate probability (prevalence ≥ 20%) and 12% ($n = 74$)

had a high probability (prevalence ≥ 75%). At least one psychiatric diagnosis was assigned by the DAWBA-rater to 20% ($n = 128$) of children. The diagnoses assigned (among all children assessed by the DAWBA) were divided into four groups: the frequency of emotional disorders was 9%, that of conduct disorders 7%, of hyperactivity 5% and that of other diagnoses 4% (tic/Tourette, pervasive developmental disorders, not otherwise specified mental disorders).

An estimate of the overall frequency of any DAWBA-rater assigned diagnosis in the entire first phase sample was calculated by applying to that sample the proportions observed in the DAWBA sample. The 29.4% proportion of children with any diagnosis in the SDQ positive DAWBA group and their 4.1% proportion in the SDQ negative group resulted in 90 and 136 children with a diagnosis in the entire first phase sample. Consequently, the frequency of any diagnosis in that sample was 8.5% (226/2666).

Capacity of the SDQ to discriminate the severity of symptoms

The AUC values of the SDQ were generally higher in the parent reports than in the teacher reports (Table 1). The AUC values of the higher cut-offs were from satisfactory to good for total score, conduct problems, hyperactivity, prosocial behaviour and for impact score. However, for teacher-rated emotional symptoms and for parent-rated peer problems the AUCs were low.

Table 1. Areas under the curve (AUC), generated by ROC analyses, for the cut-off points of the total and subscores of the Strengths and Difficulties Questionnaire (SDQ) by parents and teachers in Finnish 4–9-year-old children, calculated against the Development and Well-Being Assessment (DAWBA) disorder probabilities.

	Lower cut-off point* (AUC)	Higher cut-off point† (AUC)
Parent reported SDQ		
Total score	0.79	0.87
Emotional symptoms	0.61	0.70
Conduct problems	0.76	0.83
Hyperactivity	0.74	0.81
Peer problems	0.66	0.68
Prosocial behaviour	0.65	0.73
Impact score	0.71	0.82
Teacher reported SDQ		
Total score	0.71	0.76
Emotional symptoms	0.51	0.54
Conduct problems	0.73	0.78
Hyperactivity	0.68	0.73
Peer problems	0.65	0.70
Prosocial behaviour	0.68	0.71
Impact score	0.68	0.71

*Between SDQ normal and borderline group, DAWBA low and moderate risk group.

†Between SDQ borderline and abnormal group, DAWBA moderate and high risk group.

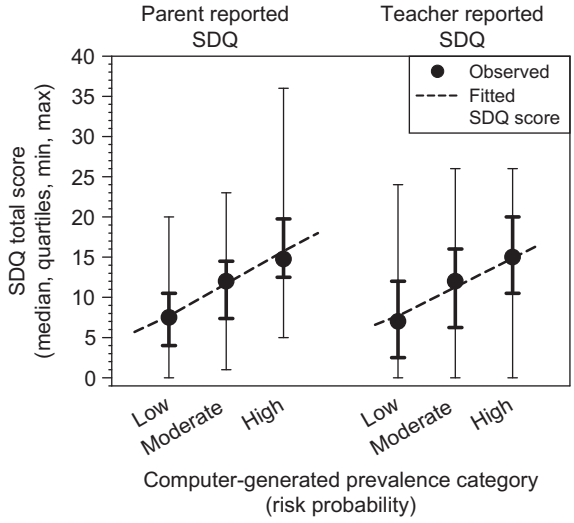


Fig. 2. The relationships between the parent and teacher reported Strengths and Difficulties Questionnaire (SDQ) total scores and the DAWBA computer predictions in Finnish 4–9-year-old children ($n = 646$).

There was also a significant, positive linear dose–response relationship between the SDQ scores and the DAWBA computer predictions ($R^2 = 0.28$, $P < 0.001$) (Fig. 2).

Selecting the Finnish cut-offs of the SDQ

The suggested Finnish parent and teacher reported cut-offs, sought by the 80th and 90th percentiles and by ROC analysis, together with the British norms for the SDQ, are presented in Table 2. Both methods suggested the same lower (9/10) and higher (11/12) cut-offs for total scores in the parent reports and for teacher reported higher cut-off (11/12). In the teacher report the 80th percentile suggested 8/9 and the ROC method 9/10 for the lower cut-off. Stratification by gender and age of the child (not shown) suggested an even wider set of cut-offs, which, applied to the entire sample, are also included in Table 3.

The sensitivity and specificity against the DAWBA computer-generated disorder probabilities were examined using the suggested British cut-offs (www.sdqinfo.com), which are 4–5 points higher in the parent-rated and 2–4 points higher in the teacher-rated total scores than in the present study (Table 2). The sensitivity in the total scores for the lower cut-off was 46% in the parent and 59% in the teacher reports and for the upper cut-off 36% and 49% respectively. The specificity for the lower cut-off was 89% in the parent reports and 71% in the teacher reports, while the corresponding values for the upper British cut-off were 97% and 86%.

The sensitivity and specificity for the nationally adjusted lower and upper cut-off options are presented in

Table 2. Parent and teacher reported cut-off points, suggested by the 80th and 90th percentiles and by ROC analysis, compared with the international norms for the SDQ, in the sample of Finnish 4–9-year-old children.

	Lower cut-off point			Higher cut-off point		
	Finnish		British‡	Finnish		British‡
	80th percentile	ROC*	80th percentile	90th percentile	ROC†	90th percentile
Parent reported SDQ	(*n* = 2635)	(*n* = 641)		(*n* = 2635)	(*n* = 641)	
Total score	9.5	10.25	14	12.0	12.25	17
Emotional symptoms	2.0	1.25	4	2.5	1.25	5
Conduct problems	2.5	2.75	3	3.5	2.75	4
Hyperactivity	4.0	3.75	6	5.0	4.75	7
Peer problems	2.0	1.75	3	3.0	2.25	4
Teacher reported SDQ	(*n* = 2242)	(*n* = 573)		(*n* = 2242)	(*n* = 573)	
Total score	8.5	10.25	12	12.0	12.25	16
Emotional symptoms	1.5	0.75	5	2.5	0.75	6
Conduct problems	2.0	1.75	3	3.0	2.75	4
Hyperactivity	4.0	4.25	6	6.0	4.75	7
Peer problems	2.0	1.75	4	3.0	2.75	5
Procosial skills	20th percentile	ROC*	20th percentile	10th percentile	ROC†	10th percentile
Parent	6.0	6.75	5	5.5	6.75	4
Teacher	5.0	5.75	5	4.0	5.25	4

*Analysed against the computer-predicted probability of any disorder, generated by the DAWBA, low vs. moderate or high probability.
†Analysed against the computer-predicted probability of any disorder, generated by the DAWBA, low or moderate vs. high probability.
‡www.sdqinfo.com.

Table 3. Sensitivity and specificity for the considered nationally adjusted lower and upper total score cut-off points of the parent and teacher reported Strengths and Difficulties Questionnaire (SDQ) in Finnish 4–9-year-old children (*n* = 646).

	Sensitivity		Specificity	
	Reference method 1* (%)	Reference method 2† (%)	Reference method 1* (%)	Reference method 2† (%)
Lower SDQ total score cut-off				
Parent reported SDQ				
9/10	76	–	69	–
Teacher reported SDQ				
6/7	80	–	48	–
8/9	70	–	57	–
9/10	66	–	63	–
Upper SDQ total score cut-off				
Parent reported SDQ				
10/11	90	76	68	70
11/12	90	72	74	76
12/13	79	61	79	80
Teacher reported SDQ				
9/10	76	72	58	61
11/12	70	63	66	69
13/14	57	50	79	81

*Calculated against the Development and Well-Being Assessment (DAWBA) computer-generated disorder prevalence category (probability), low vs. moderate or high probability for the lower SDQ cut-off and low or moderate vs. high probability for the higher SDQ cut-off.
†Calculated against the existence of any DAWBA-rater assigned ICD-10 diagnosis.

Table 3. Examining against the existence of clinically rated diagnoses produced lower values for sensitivity and higher values for specificity than comparisons against the DAWBA computer-generated probabilities. In the present community sample, the balance between recognizing psychiatric symptoms and avoiding false positives was targeted by choosing scores with high sensitivity for the lower cut-off and high specificity for the higher cut-off, while requiring both of them to be at least 60% when calculated against both the DAWBA computer-generated probabilities and the clinically rated diagnoses, when possible. According to these principles, the optimal cut-offs for total scores in both informants' reports were 9/10 for the lower and 11/12 for the upper cut-off. With these cut-offs, the sensitivity for the lower cut-off was 76% in the parent report and 66% in the teacher report and the specificity 69% and 63% respectively, when measured against computer-generated probabilities. For the upper cut-off, the sensitivity against the same reference was 90% in the parent report and 70% in the teacher report, while the specificity was 74% and 66%. When calculated against the clinical disorder rating, the sensitivity for the upper cut-off was 72% in the parent and 63% in the teacher report while the specificity was 76% in the parent and 69% in the teacher report.

With the selected Finnish cut-offs, parents scored 18% and teachers 15% of the children at or over the lower cut-off, and both informants scored 11% of the children at or over the upper cut-off. Either in the parent or in the teacher reported SDQ or both, 24% of the children were scored at or over the lower and 17% at or over the upper cut-off. Altogether 6% of the children were scored at or above the lower cut-off according to both informants' reports, and the respective proportion regarding the upper cut-off was 3%.

Capacity of the SDQ to distinguish disorders

Frequencies of DAWBA-rater assigned diagnoses in groups defined by the Finnish cut-offs for informant-rated total scores of the SDQ are presented in Table 4. In the normal group, defined by the combined parent report 93% of the children had no diagnosis while the respective proportion defined by the combined teacher report was 90%. In the parent reported abnormal group, on the other hand, the proportion of children having at least one diagnosis was 41% and 34% in the corresponding teacher reported group. In addition, the frequencies of diagnoses increased when the parent reported borderline and abnormal group were compared. Both the capacity to distinguish diagnoses between the borderline and abnormal group and the prevalence rates in the abnormal group were lower in the teacher reports than in the parent reports.

Conclusion

The cut-offs of the SDQ in the Finnish community sample of 4–9-year-old children were sought against the diagnostic assessment method of the DAWBA. The adjusted cut-offs indicated significant differences compared with the corresponding British norms. Yet the properties of the method's discriminative validity were acceptable.

Very low sensitivity values were found with the suggested British norms for the SDQ (16) in this sample. In order to identify psychiatric symptoms in a Finnish community sample of young children, adjusted cut-offs must be sought. The Finnish cut-offs were clearly lower than

Table 4. Frequencies of DAWBA-rater assigned diagnoses, based on the Development and Well-Being Assessment (DAWBA), in the normal, borderline and abnormal groups, defined by the Finnish, adjusted cut-off points for parent and teacher rated total scores of the Strength and Difficulties Questionnaire (SDQ).

	SDQ total score			
	< 10 Normal (%)	10–11 Borderline (%)	≥ 12 Abnormal (%)	*P*
Parent reported SDQ	(*n* = 349)	(*n* = 58)	(*n* = 209)	
At least one diagnosis	7	12	41	< 0.001
Emotional disorder	3	2	20	< 0.001
Conduct disorder	2	5	16	< 0.001
Hyperactivity	1	5	11	< 0.001
Other diagnosis	1	0	9	< 0.001
Teacher reported SDQ	(*n* = 307)	(*n* = 35)	(*n* = 212)	
At least one diagnosis	10	28	34	< 0.001
Emotional disorder	6	17	12	0.024
Conduct disorder	3	11	13	< 0.001
Hyperactivity	1	0	14	< 0.001
Other diagnosis	3	7	6	ns

the British norms. Yet, with the nationally adjusted cut-offs for total scores, the sensitivity and specificity values were acceptable. A large proportion of the sample was under school age (4–6-year-olds) and thus confirmed the previous findings of younger children having lower score distributions in normative SDQ data from Britain, America and Spain (16). Other explanations for the low cut-off scores in the present study can be sought in differences in recognizing and reporting styles among Finnish parents and teachers compared with other cultures and from the high level of social welfare and well-being of children in Finland. To confirm the results, further studies on the validity of the SDQ in young children are needed in different cultures.

Surprisingly, no significant differences were found in the SDQ norms between genders in the present study even though the Finnish parents and teachers had reported higher scores for boys than for girls in an earlier study of the same sample (20). Also, in the British (5–15-year-old children), American (4–17-year-old children) and Danish (5–12-year-old children) normative data, significantly higher percentages of boys than girls scored above the original percentile-based cut-offs of the SDQ (16, 32, 33). Common cut-offs are recommended for British and US boys and girls, whereas the Dane recommend separate bandings for SDQ cut-off scores for the genders both in the age groups of 5–7 and 10–12-years (16). The present study, searching the cut-offs both against the percentiles of the SDQ scores and the diagnostic assessment tool offered a wide set of cut-off candidates. The analyses stratified by gender did not alter the findings of the best options for the lower and upper Finnish cut-offs, regarding sensitivity and specificity.

With the selected Finnish cut-offs, 24% of the sample evaluated by parents or teachers was scored as borderline and 17% as abnormal on the total scores of the SDQ. Previously in Finland, in the epidemiological prevalence study in 1989, 24% of 8–9-year-old children were evaluated by parents, teachers or themselves to have psychiatric symptoms (34). According to Rutter's parent or teacher questionnaire, in a sample of 8-year-olds in 2005, 20% of the boys and 10% of the girls scored screen-positive (35). The sample of the present study does not meet the criteria of an epidemiological study, though results assessed in such a large sample can yield representative information also from the screening perspective. The present results suggest that in young children, too, a significant proportion is in need of mental health support.

The results of the AUC values indicated that the SDQ was adequate for differentiating between low-risk and high-risk children in the sample. The parent reports differentiated these groups better than the teacher reports did. The capacity to distinguish between children with and without diagnoses was approximately on the level of chance on the emotional symptoms subscale reported by teachers and low on the peer problems subscale reported by parents. In these areas, adults may have an outsider role in the child's life and have difficulties in evaluating them. The findings reassert the perception that it is more reliable to use and interpret the SDQ total score in screening than to interpret the subscales separately (18). On the total scores of the SDQ, the values of AUC were the same in the Finnish parent reports (0.87) and lower in the Finnish teacher reports (0.76 vs. 0.83) than the values of AUC weighted averages in earlier studies presented by Stone's review (18). Different comparison methods in defining the AUCs for the SDQ make it difficult to compare the present results to earlier findings.

The diagnostic discrimination capacity of the SDQ functioned well in the sample. The presence of a DAWBA-rater assigned diagnosis in the high-risk group compared with the low-risk group was sixfold in the parent and threefold in the teacher reported SDQs. In addition, the SDQ's capacity to distinguish the diagnostic spectrum seemed credible when compared with earlier findings.

The present sample suggests an overall frequency of psychiatric disorders of about 8.5% in the sample. Earlier studies considering mental disorders in children have presented prevalence rates ranging from 7% to 17% (1, 7, 14, 17, 28, 29, 36–38). The present results suggest that young children suffer from psychiatric disorders as much as older children do (3, 5, 7).

It would be most sensible to validate an instrument in representative samples of national cohorts. Unfortunately, in Finland as in many other countries, this has not been possible. This is considered a definite limitation of the study. Doing several studies in selected groups of children and adolescents will introduce differences in methods, biases and confounding through differences in sampling strategies, attrition and assessments.

The overall partition rate of the study was only 57%, which is an important limitation and which weakens seriously the representativeness of the study sample. Another problem of the representativeness was that the British cut-offs were used to divide the sample into screen-positive and screen-negative subgroups, although, according to the subsequent analysis, the cut-offs showed very low sensitivity in the sample. The representativeness of the sample however was improved by the fact that not only the 90th but also the 80th percentile were used in defining caseness. Preschool-aged (4–6-year-old) children and boys were overrepresented in the sample but in the validity results the differences between age-groups and genders did not play an important role.

Using single combined SDQ mean scores for parents and for teachers in the analyses instead of maximum score values in each informant class might have introduced a limitation in the study, as it may have affected the adjusted cut-offs or sensitivities. On the other hand, the mean scores of the mother and the father SDQ

reports and of the two day-care teachers' reports may better depict the different views of the child's well-being than the maximum score alternative. In order to adhere to a multi-informant approach in evaluating the child's mental health, the combined SDQ mean scores were tested against the multi-informant diagnostic evaluation with the DAWBA.

Several limitations should be noted in the interview phase of the study. Firstly, it is a definite limitation not to be able to arrange an inter-rater reliability testing of the DAWBA. Having a single DAWBA-rater removed inter-rater bias but may have caused a systematic distortion in the diagnoses assigned. Several arrangements were made, though, to ensure the competence of the DAWBA-rater: adequate training of the method, regular supervision meetings and possibilities to consult a group of experienced colleagues in diagnostic problems.

Secondly, it should be noted that the DAWBA-rater was not blinded to the SDQ scores collected as part of the DAWBA assessment in the second phase. This may have introduced bias and a tendency to circularity in calculating the cut-off scores for the SDQ, even though the cut-offs were based on the first phase SDQ, for which the rater was blind. Furthermore, in the DAWBA assessment, we used the skip-rules, which are directly linked to the SDQ score and this may also have increased the calculated sensitivity of the screening instrument.

In addition, the DAWBA-rater was aware of the children's DAWBA computer-predicted disorder probabilities when assigning diagnoses and may therefore have been influenced by these. Nevertheless, in earlier studies the results of sensitivity analyses have been similar whether raters have been blind to the DAWBA probabilities or not (26).

The strengths of the study were the high participation rates in the diagnostic assessment and multi-informant reports in most cases. Effects of different ways of administration of diagnostic interviews have mostly been studied among adults. Preliminary results of parents' face-to-face and telephone administration of a child psychiatric interview has been comparable (39). However, most previous surveys have either used face-to-face or web-completion of the DAWBA. Compared with the approach with personal interviews, the web-based surveys with DAWBA may be faster to perform at lower costs, but they may also introduce bias related to selective participation and partial or low response rates (40, 41). In the present study, decent participation rates were achieved by the telephone administration and parents also gave informally positive feedback on their experiences of being interviewed by phone.

In conclusion, the study confirmed the necessity of seeking adjusted cut-offs of the SDQ for the population in question. The norms are needed for the method to achieve appropriate capacity to distinguish the high-risk children with sufficient sensitivity but simultaneously to avoid raising unwarranted concerns about low-risk children. Because young children suffer from psychiatric disorders as much as older children do, the necessity of early recognition becomes emphasized. For this purpose, the SDQ stands for a standardized and validated screening method in Finland. International comparisons of the method's validity properties and the frequencies of children's mental health problems call also for population specific norms, taking into account not only the culture but also the child's age.

Acknowledgements—We thank the participating families, the public health nurses and the teachers in day-care and in the schools participating in the study; the DAWBA interviewers; the project co-ordinator Sari Miettinen at the Child Psychiatric Clinic, Tampere University Hospital, for collaboration in data collection. We are grateful to Raili Salmelin, PhD, at the University of Tampere, School of Health Sciences for her guidance and help in producing tables and figures. We thank the DAWBA cases' consulting group of child psychiatrists; Ilona Luoma, Mirjami Mäntymaa, Reija Latva and Merja Mäki. The data were collected on the project "Developing children's mental health work, 2007–2009", funded by the hospital districts of Pirkanmaa and South Karelia and the Ministry of Social Affairs and Health, Finland. This study was supported by grants from the Finnish Child Psychiatric Research Foundation, the Emil Aaltonen Foundation, the Medical Research Foundation of Tampere University Hospital and the Finnish Foundation for Pediatric Research.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003;60:837–44. doi: 10.1001/archpsyc.60.8.837.
2. Du Y, Kou J, Coghill D. The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China. Child Adolesc Psychiatry Ment Health 2008;2:8. doi: 10.1186/1753-2000-2-8.
3. Skovgaard AM, Houmann T, Christiansen E, Landorph S, Jorgensen T, CCC 2000 Study Team, et al. The prevalence of mental health problems in children 1(1/2) years of age—The Copenhagen Child Cohort 2000. J Child Psychol Psychiatry 2007;48:62–70. doi: 10.1111/j.1469-7610.2006.01659.x.
4. Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. Pediatrics 2010;125:75–81. doi: 10.1542/peds.2008-2598.
5. Gleason MM, Zamfirescu A, Egger HL, Nelson CA,3rd, Fox NA, Zeanah CH. Epidemiology of psychiatric disorders in very young children in a Romanian pediatric setting. Eur Child Adolesc Psychiatry 2011;20:527–35. doi: 10.1007/s00787-011-0214-0; 10.1007/s00787-011-0214-0.
6. Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, et al. Child and adolescent mental health worldwide: Evidence for action. Lancet 2011;378:1515–25. doi: 10.1016/S0140-6736(11)60827-1.
7. Wichstrom L, Berg-Nielsen TS, Angold A, Egger HL, Solheim E, Sveen TH. Prevalence of psychiatric disorders in preschoolers. J Child Psychol Psychiatry 2011 Dec 29. doi: 10.1111/j.1469-7610.2011.02514.x.
8. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry 2001;40:1337–45. doi: 10.1097/00004583-200111000-00015.
9. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: Description and initial

validation of an integrated assessment of child and adolescent psychopathology. J Child Psychol Psychiatry 2000;41:645–55.
10. Obel C, Heiervang E, Rodriguez A, Heyerdahl S, Smedje H, Sourander A, et al. The Strengths and Difficulties Questionnaire in the Nordic countries. Eur Child Adolesc Psychiatry 2004;13 Suppl 2:II32–9. doi: 10.1007/s00787-004-2006-2.
11. Marzocchi GM, Capron C, Di Pietro M, Duran Tauleria E, Duyme M, Frigerio A, et al. The use of the Strengths and Difficulties Questionnaire (SDQ) in Southern European countries. Eur Child Adolesc Psychiatry 2004;13 Suppl 2:II40–6. doi: 10.1007/s00787-004-2007-1.
12. Woerner W, Fleitlich-Bilyk B, Martinussen R, Fletcher J, Cucchiaro G, Dalgalarrondo P, et al. The Strengths and Difficulties Questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. Eur Child Adolesc Psychiatry 2004;13 Suppl 2:II47–54. doi: 10.1007/s00787-004-2008-0.
13. Hawes DJ, Dadds MR. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. Aust N Z J Psychiatry 2004;38:644–51.
14. Goodman R, Slobodskaya H, Knyazev G. Russian child mental health—A cross-sectional study of prevalence and risk factors. Eur Child Adolesc Psychiatry 2005;14:28–33.doi: 10.1007/s00787-005-0420-8.
15. Rothenberger A, Becker A, Erhart M, Wille N, Ravens-Sieberer U, BELLA study group. Psychometric properties of the parent strengths and difficulties questionnaire in the general population of German children and adolescents: Results of the BELLA study. Eur Child Adolesc Psychiatry 2008;17 Suppl 1:99–105. doi: 10.1007/s00787-008-1011-2.
16. Youthinmind. SDQ, Information for researchers and professionals about the Strengths and Difficulties Questionnaires [updated 1 January 2012; cited 13 June 2012]. Available from: http://www.sdqinfo.com/.
17. Goodman A, Heiervang E, Fleitlich-Bilyk B, Alyahri A, Patel V, Mullick MS, et al. Cross-national differences in questionnaires do not necessarily reflect comparable differences in disorder prevalence. Soc Psychiatry Psychiatr Epidemiol 2011 Oct 28. doi: 10.1007/s00127-011-0440-2.
18. Stone LL, Otten R, Engels RC, Vermulst AA, Janssens JM. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: A review. Clin Child Fam Psychol Rev 2010;13:254–74. doi: 10.1007/s10567-010-0071-2.
19. Koskelainen M, Sourander A, Kaljonen A. The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. Eur Child Adolesc Psychiatry 2000;9:277–84.
20. Borg AM, Kaukonen P, Salmelin R, Joukamaa M, Tamminen T. Reliability of the Strengths and Difficulties Questionnaire among Finnish 4–9-year-old children. Nord J Psychiatry 2012 Mar 8. doi: 10.3109/08039488.2012.660706.
21. Achenbach TM RL. Manual for the ASEBA school-age forms & profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families; 2001.
22. Goodman R. The Strengths and Difficulties Questionnaire: A research note. J Child Psychol Psychiatry 1997;38:581–6.
23. Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. J Child Psychol Psychiatry 1999;40:791–9.
24. World Health Organisation editor. The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research. Geneva: WHO; 1994.
25. American Psychiatric Association editor. Diagnostic and statistical manual of mental disorders, 4th edition (DSM-IV). Washington, DC: APA; 1994.
26. Goodman A, Heiervang E, Collishaw S, Goodman R. The 'DAWBA bands' as an ordered-categorical measure of child mental health: Description and validation in British and Norwegian samples. Soc Psychiatry Psychiatr Epidemiol 2011;46:521–32. doi: 10.1007/s00127-010-0219-x.
27. Youthinmind. DAWBA, Information for researchers and clinicians about the Development and Well-Being Assessment. [updated 2011 Februari 11; cited 2012 June 13]. Available from: http://www.dawba.com/.
28. Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. J Am Acad Child Adolesc Psychiatry 2003;42:1203–11. doi: 10.1097/01.chi.0000081820.25107.ae.
29. Heiervang E, Stormark KM, Lundervold AJ, Heimann M, Goodman R, Posserud MB, et al. Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. J Am Acad Child Adolesc Psychiatry 2007;46:438–47. doi: 10.1097/chi.0b013e31803062bf.
30. Frigerio A, Rucci P, Goodman R, Ammaniti M, Carlet O, Cavolina P, et al. Prevalence and correlates of mental disorders among adolescents in Italy: The PrISMA study. Eur Child Adolesc Psychiatry 2009;18:217–226. doi: 10.1097/chi.0b013e31803062bf.
31. Tape T. University of Nebraska Medical Center. Interpreting Diagnostic Tests. [cited 2012 Jan 26]. Available from: http://gim.unmc.edu/dxtests/roc3.htm
32. Bourdon KH, Goodman R, Rae DS, Simpson G, Koretz DS. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. J Am Acad Child Adolesc Psychiatry 2005;44:557–64. doi: 10.1097/01.chi.0000159157.57075.c8.
33. Niclasen J, Teasdale TW, Andersen AM, Skovgaard AM, Elberling H, Obel C. Psychometric properties of the Danish Strength and Difficulties Questionnaire: The SDQ assessed for more than 70,000 raters in four different cohorts. PLoS One 2012 Feb 27. doi: 10.1371/journal.pone.0032025.
34. Almqvist F, Kumpulainen K, Ikaheimo K, Linna SL, Henttonen I, Huikko E, et al. Behavioural and emotional symptoms in 8–9-year-old children. Eur Child Adolesc Psychiatry 1999;8 Suppl 4:7–16.
35. Sourander A, Niemela S, Santalahti P, Helenius H, Piha J. Changes in psychiatric problems and service use among 8-year-old children: A 16-year population-based time-trend study. J Am Acad Child Adolesc Psychiatry 2008;47:317–27. doi: 10.1097/CHI.0b013e318160b98f.
36. Mullick MS, Goodman R. The prevalence of psychiatric disorders among 5–10 year olds in rural, urban and slum areas in Bangladesh: An exploratory study. Soc Psychiatry Psychiatr Epidemiol 2005;40:663–71. doi: 10.1007/s00127-005-0939-5.
37. Pillai A, Patel V, Cardozo P, Goodman R, Weiss HA, Andrew G. Non-traditional lifestyles and prevalence of mental disorders in adolescents in Goa, India. Br J Psychiatry 2008;192:45–51. doi: 10.1192/bjp.bp.106.034223.
38. Almqvist F, Puura K, Kumpulainen K, Tuompo-Johansson E, Henttonen I, Huikko E, et al. Psychiatric disorders in 8–9-year-old children based on a diagnostic interview with the parents. Eur Child Adolesc Psychiatry 1999;8 Suppl 4:17–28.
39. Paing WW, Weller RA, Dixon TA, Weller EB. Face-to-face versus telephone administration of the parent's version of the children's interview for psychiatric syndromes (P-ChIPS). Curr Psychiatry Rep 2010; 12:122–6. doi: 10.1007/ s11920-010-0101-2.
40. Heiervang E, Goodman R. Advantages and limitations of web-based surveys: Evidence from a child mental health survey. Soc Psychiatry Psychiatr Epidemiol 2011;46:69–76. doi: 10.1007/s00127-009-0171-9.
41. Goodman R. Online child and adolescent mental health surveys can be good enough. Soc Psychiatry Psychiatr Epidemiol 2013;48:1317–25. doi: 10.1007/s00127-013-0658-2.

Anne-Mari Borg, M.D., Department of Child Psychiatry, School of Medicine, University of Tampere, 33014 Tampere, and Department of Child Psychiatry, Tampere University Hospital, 33521 Tampere, Finland.
Pälvi Kaukonen, M.D., Department of Child Psychiatry, Tampere University Hospital, 33521 Tampere, Finland.
Matti Joukamaa, M.D., Ph.D., School of Health Sciences, University of Tampere, 33014 Tampere, and Department of Psychiatry, Tampere University Hospital, 33521 Tampere, Finland.
Tuula Tamminen, M.D., Ph.D., Department of Child Psychiatry, School of Medicine, University of Tampere, 33014 Tampere, and Department of Child Psychiatry, Tampere University Hospital, 33521 Tampere, Finland.

Journal of Child and Adolescent Mental Health 2014, 26(3): 229–238

JOURNAL OF
CHILD & ADOLESCENT MENTAL HEALTH
ISSN 1728-0583 EISSN 1728-0591
http://dx.doi.org/10.2989/17280583.2014.923432

Research Article

Feasibility of the Strengths and Difficulties Questionnaire in assessing children's mental health in primary care: Finnish parents', teachers' and public health nurses' experiences with the SDQ

Anne-Mari Borg[1,3]*, Raili Salmelin[2,3], Pälvi Kaukonen[3], Matti Joukamaa [2,4], Tuula Tamminen[1,3]

[1] University of Tampere, School of Medicine, Department of Child Psychiatry, Lääketieteen yksikkö, 33014 Tampereen yliopisto, Finland
[2] University of Tampere, School of Health Sciences, Terveystieteiden yksikkö, 33014 Tampereen yliopisto, Finland
[3] Tampere University Hospital, Department of Child Psychiatry, Tampereen yliopistollinen sairaala, Lastenpsykiatrian vastuualue, PL 2000, 33521 Tampere, Finland
[4] Tampere University Hospital, Department of Psychiatry, Tampereen yliopistollinen sairaala, Psykiatrian vastuualue, PL 2000, 33521 Tampere, Finland
**Corresponding author, email: anne-mari.borg@uta.fi*

Background: **The aim of the study was to evaluate the feasibility of the Strengths and Difficulties Questionnaire (SDQ).**
Method: Following the administration of the SDQ in medical check-ups of 4–9 year-old children (*n* = 2 682) the involved parents, teachers and public health nurses were asked to complete a feedback questionnaire of the SDQ.
Results: **Parents took a maximum of 10–15 minutes to complete the SDQ, and only the public health nurses reported that its use was rather burdensome. The SDQ was an age-appropriate method and it was helpful in increasing information and agreement about the child's mental health and need for support. Using the SDQ was a positive experience for parents, but they expected more dialogue with the professionals about the child's situation. The respondents criticised the questionnaire somewhat for being difficult to interpret and complete.**
Conclusions: **The SDQ was found to be a feasible method for screening children's mental health in primary health care together with parents, teachers and public health nurses. Using the SDQ was a positive experience for parents. However, they reminded the professionals of the importance of sensitive dialogue when assessing the mental health of the child.**

Introduction

Front-line professionals play an important role in identifying children suffering from mental health problems, whose frequency is known to be high (Costello, Egger and Angold 2005, Skovgaard *et al*. 2007, Merikangas *et al*. 2010, Kieling *et al*. 2011, Wichstrom *et al*. 2012). Standardised assessment methods ensure reliability and validity in recognising children in need of psychosocial support. However, such an assessment method also has to meet many practical criteria to be accepted and used in everyday clinical practice.

The concept of feasibility has been defined as a psychometric property of an assessment schedule being suitable for routine clinical practice (Slade, Thornicroft and Glover 1999, Slade *et al*., 2001). Slade *et al*. (1999) also suggest that a feasible method is one that incorporates six desirable elements: brevity, simplicity, relevance, acceptability, availability and value. Very few studies on the feasibility properties of the methods used in assessing children's mental health can be found. These feasibility studies have mostly focused on assessing the compliance (response rates) of the measurement (Gold *et al*. 2009, Allen *et al*. 2010) and professionals' evaluations and the acceptability of the method using focus group methodology (Oesterheld and Haber 1997, Slade *et al*. 2001, Sorensen, Thomsen and Bilenberg 2007, Martin *et al*. 2011).

Feasibility criteria for selecting a questionnaire from professionals' point of view have been proposed to be ease of use, helpfulness of the information and appropriateness of the method to the informant's abilities and to the desired evaluation (Myers and Winters 2002). Earlier studies have shown little interest in parents' and other informants' experiences of assessing children's mental health using screening methods.

The Strengths and Difficulties Questionnaire (SDQ) is a well-documented short enquiry for assessing 4−16-year-old children's mental health (Goodman 2001). The SDQ is commonly used both in research settings and for screening in community and clinical settings in different countries and cultures (Hawes and Dadds 2004, Marzocchi *et al*. 2004, Obel *et al*. 2004, Woerner *et al*. 2004, Goodman, Slobodskaya and Knyazev 2005, Du, Kou and Coghill 2008, Rothenberger *et al*. 2008). The method was designed mindful of practical aspects to simplify its use: a short one to two-page format, a single form suitable for both parents and teachers, free of charge and easy availability (Goodman 1997, Youthinmind 2013). A few qualitative studies on the SDQ have been carried out assessing the method's acceptability, face validity and cultural appropriateness (Williamson *et al*. 2010, Moran *et al*. 2012, Stasiak *et al*. 2013). However, to the best of the authors' knowledge no studies in naturalistic settings are available on how the SDQ is experienced by informants and accepted by professionals as a screening method when assessing children's psychosocial health in routine clinical practice.

The earlier Finnish results on the reliability and validity of the SDQ support its position as a screening method in community samples of children from preschool age (under 7-year-olds) to adolescence (4−15 year olds) (Koskelainen, Sourander and Kaljonen 2000, Borg *et al*. 2012). The aims of this study were to evaluate the feasibility of the method in Finnish primary health care by collecting feedback on the SDQ from parents, teachers in pre-school education and at schools and from public health nurses.

Method

Study design

The study was a part of a project entitled *Developing children's mental health work*, 2007–2009, conducted in two hospital districts in Finland from March 2008 to March 2009. Altogether, 154 child health clinics and school healthcare clinics in 25 municipalities participated (according to their willingness and resources). The local ethics committees approved the study and therefore the study was conducted in accordance with the ethical standards of the Declaration of Helsinki. Informed consent was obtained from all participating parents.

Parents were asked to complete the SDQ before attending their child's regular medical check-up at their child health clinic, child's age 4−6 years or at the school healthcare clinic, child's age 7−9 years. Participating parents also asked the child's teacher in pre-school education or at school to complete the SDQ. The parents returned all the questionnaires to the public health nurse when attending the medical check-up. The public health nurses and the pre-school teachers could use the SDQ in discussion with the parents but the sum scores or reports produced by the method were not at their disposal. After the check-up visit a feedback questionnaire on the feasibility of the SDQ was completed anonymously by the participating parents and only once by each teacher and public health nurse involved in the process. Otherwise the design and the sample of the study have been described in more detail elsewhere (Borg *et al*. 2012).

Before the study started, brief training sessions on the SDQ were offered to medical professionals in primary health care together with pre-school education personnel and school teachers in the participating municipalities. The training session comprised an overview of the method and its use, and of instructions on implementing the study design in the context of routine medical check-ups. Participants were encouraged to contact the study coordinator or the researcher (the first author) in case of any further questions. The response rate of the feedback questionnaires was increased by reminders to public health nurses and by supporting their administrative tasks during the study process.

Sample

Altogether 4 178 eligible children (50% girls) and their parents were invited to participate in the study; the participation rate was 64% (*n* = 2 682). The response rate for the Feasibility Questionnaire among parents was 58% (1 546/2 682). In the participating 154 child health clinics and school healthcare clinics, 225 public health nurses were involved in the study process. Of the nurses 70% (*n* = 156) completed and returned the Feasibility Questionnaire of the SDQ. The nurses involved in the study process completed the Feasibility Questionnaire only once and not per every child they met in regular health checks. Of the parental respondents 1 009 had pre-school-aged and 537 had school-aged children. A total of 107 of the nurses worked at child health clinics and 49 at school healthcare clinics. The exact number of participating pre-school education personnel could not be determined, but 1 286 Feasibility Questionnaires were returned by pre-school teachers. At schools, the number of teachers (of 7–9 year-old children) involved in the study process was 123. Teachers also completed the Feasibility Questionnaire only once during the study procedure, even when one teacher usually had many pupils who participated in the study. Four teachers did not return the Feasibility Questionnaire and thus the response rate was 97%.

Measures

The SDQ is a screening questionnaire for 3−16 year-olds to be completed by parents, teachers and 11−16 year-old children themselves (Goodman 1997, 2001). The first page of the questionnaire consists of 25 items on the child's symptoms and social skills (emotional symptoms, conduct problems, hyperactivity/inattention difficulties, peer relationship problems, prosocial behaviour). The second page of the SDQ consists of global assessments of the impact of the child's mental difficulties (Goodman 1999). The Finnish version of the two-page, extended version of the SDQ was used (Youthinmind 2013).

The SDQ items are scored 1 for 'somewhat true' and, depending on the item, 0 or 2 for 'not true' or for 'certainly true'. In the impact supplement the respondent is asked whether the child has any emotional or behavioural difficulties and if so, about the duration or chronicity of the difficulties, overall distress, social impairment, and burden to others. The items on overall distress and social impairment are scored from 0 to 2.

The Feasibility Questionnaire of the SDQ was developed for this study to collate experiences and opinions of parents, public health nurses and teachers in pre-school education and at school in the context of medical check-ups in primary health care (see Table 1). The questionnaire was developed based on both earlier research (Slade *et al*. 1999, 2001, Myers and Winters 2002) and clinical considerations.

Using close-ended questions, all respondents were asked to report how long it took to complete the SDQ and how burdensome they found it. They were also asked to estimate how age-appropriate (to the developmental stage) this method was in assessing the psychosocial well-being of the child. Teachers at schools answered only these first three questions. The next five items, answered on a five-step scale (totally agree, somewhat agree, somewhat disagree, totally disagree, can't say), on the parent's, public health nurse's and on the pre-school teacher's form assessed the impact of using the SDQ in cooperation between parents and public health nurses or pre-school teachers. These items collated how useful the SDQ was in gathering information and in reaching a common understanding about the child's mental health and possible need for support. All these respondents were asked how parents felt about the method and how burdensome a tool

Table 1: The contents of the Feasibility Questionnaire of the SDQ for Finnish parents, public health nurses and teachers in preschool education and at school

Question/statement	Type	Options/unit	Target respondents
How long a time did it take to complete the SDQ?	Open-ended	min	Parents, Public health nurses, Teachers in preschool education, Teachers at school
How age-appropriate was the SDQ in assessing the psychosocial well-being of the child?	Close-ended	Very good / Fairly good / Not good, not poor / Rather poor / Very poor	Parents, Public health nurses, Teachers in preschool education Teachers at school
How burdensome did you find the SDQ?	Close-ended	Not at all / Slightly / Rather / Very	Parents, Public health nurses, Teachers in preschool education, Teachers at school
The SDQ increases co-operation	Close-ended	Totally agree / Somewhat agree / Somewhat disagree / Totally disagree / Can't say	Parents, Public health nurses, Teachers in preschool education
The SDQ increases knowledge			
The SDQ helps to reach agreement			
The SDQ is too burdensome a tool			
The using of the SDQ is a positive experience for parents			
Please, compare your observations of the SDQ and the earlier methods you have used when evaluating children's psychosocial well-being.	Open-ended	-	Public health nurses, Teachers in preschool education
Please, give informal feedback on your experiences of the SDQ.	Open-ended	-	Parents, Public health nurses, Teachers in preschool education, Teachers at school

they found it as part of the medical check-up of the child and in the conversations between parents and pre-school teachers.

The open-ended questions asked public health nurses and pre-school teachers to list what methods they had previously used to evaluate children's psychosocial well-being, social skills and possible need for support and to compare their observations of the SDQ with these earlier methods. In addition, all respondents were offered a chance to write feedback in their own words on their experiences of the method.

Data analyses

The results of the Feasibility Questionnaire of the SDQ are mostly reported for all age groups together because the number of participating public health nurses at school health clinics was small (n = 49) and thus their answers were distributed into small categories. The Feasibility Questionnaires of the SDQ were completed anonymously and only the respondents' status and the age group of the child concerned were identified. Therefore it was not possible to combine data from the Feasibility Questionnaires of the SDQ with other data on the children in the study.

The open-ended questions included in the Feasibility Questionnaire of the SDQ were analysed by inductive and qualitative data dissection of the written responses. The first author analysed the data manually and had no preconceived hypotheses. The written responses of all informants were first explored thoroughly. The written answers were next coded according to the main themes emerging from different respondents. To systematise the analysis, the incidents of these codes were then tabulated in quantitative form as per cent proportions of both respondents and all responses. This categorisation was made more compact by combining parallel themes. The steering group of the

study assessed and discussed the choices made in the analyses. The most frequently emerged themes are represented in this study.

Results

Most of the parents spent a maximum of 10 minutes on completing the SDQ (Table 2). Teachers, on average, took for 33% a maximum of 10 minutes and for 76% up to 15 minutes per child to complete the SDQ. Of the public health nurses 56% went through the questionnaire with parents in 10 minutes or less and 89% in 15 minutes or less.

The SDQ was considered very or fairly age-appropriate for assessing the child's psychosocial well-being by 86% of the parents, 83% of the teachers and 80% of the public health nurses (Table 2).

Almost all the parents (97%) and 93% of the teachers found the method not at all or not much burdensome. A total of 37% of the public health nurses felt that the use of the SDQ in medical check-ups was rather burdensome (Table 2).

From half to over two-thirds of the parents, public health nurses and teachers reported that the SDQ increased cooperation between parents and professionals (Figure 1). The method substantially increased what was known of the mental health and social skills of the child according to three-quarters or more of all respondent groups. According to most of all respondents, the method also helped reach agreement about the child's mental health and possible need for support (Figure 1). Using the SDQ was a positive experience for 88% of the parents; this was also reported by half the pre-school teachers and by 81% of the public health nurses. Parents and teachers did not consider the SDQ too burdensome a tool for medical check-ups or for use in the conversations in

Table 2: Distribution of the feasibility aspects of the Strengths and Difficulties Questionnaire among Finnish parents, teachers and public health nurses

	Parents[a]		Public health nurses[c]	Teachers
	Mother (*n* = 2 595) (%)	Father (*n* = 1 911) (%)	(*n* = 93) (%)	(*n* = 1 406) (%)
Time (min)				
<5	5.5	6.4	7.5	2.4
5–10	82.3	83.4	48.4	30.8
10–15	9.6	8.7	33.3	43.1
15–20	1.8	0.8	9.7	21.2
>20	0.9	0.7	1.1	2.5
	Parents[b] (*n* = 1546) (%)		Public health nurses[c] (*n* = 156) (%)	Teachers (*n* = 1406) (%)
Appropriateness				
Very good	11.9		8.2	15.1
Fairly good	74.5		71.4	67.8
Not good, not poor	9.6		16.3	10.5
Rather poor	3.8		3.1	5.8
Very poor	0.1		1.0	0.8
Burden				
Not at all	56.2		11.1	44.8
Not much	40.8		50.5	48.4
Rather a lot	2.9		37.4	6.1
Very much	0.1		1.0	0.8

[a] Parents answered the question when completing the SDQ (separately from The Feasibility Questionnaire).

[b] Respondents answered the Feasibility Questionnaire anonymously.

[c] Exact information on the proportions of public health nurses who used the SDQ in conversations with parents during the study is lacking. However, 93 out of the 156 public health nurses who completed the Feasibility Questionnaire answered the question on the time needed to go through the SDQ.

pre-school education, whereas 70% of the public health nurses agreed somewhat (43%) or totally (27%) with the statement that the method was too burdensome.

In the open-ended feedback on the SDQ, the public health nurses (n = 88) and the pre-school teachers (n = 1 078) listed the most important methods they had previously used when evaluating children's psychosocial health. The most common methods for the public health nurses were locally developed questionnaires for the children's medical check-ups (63%), interviews with the parents (54%) and feedback from the child's teacher (41%). The most common methods for the pre-school

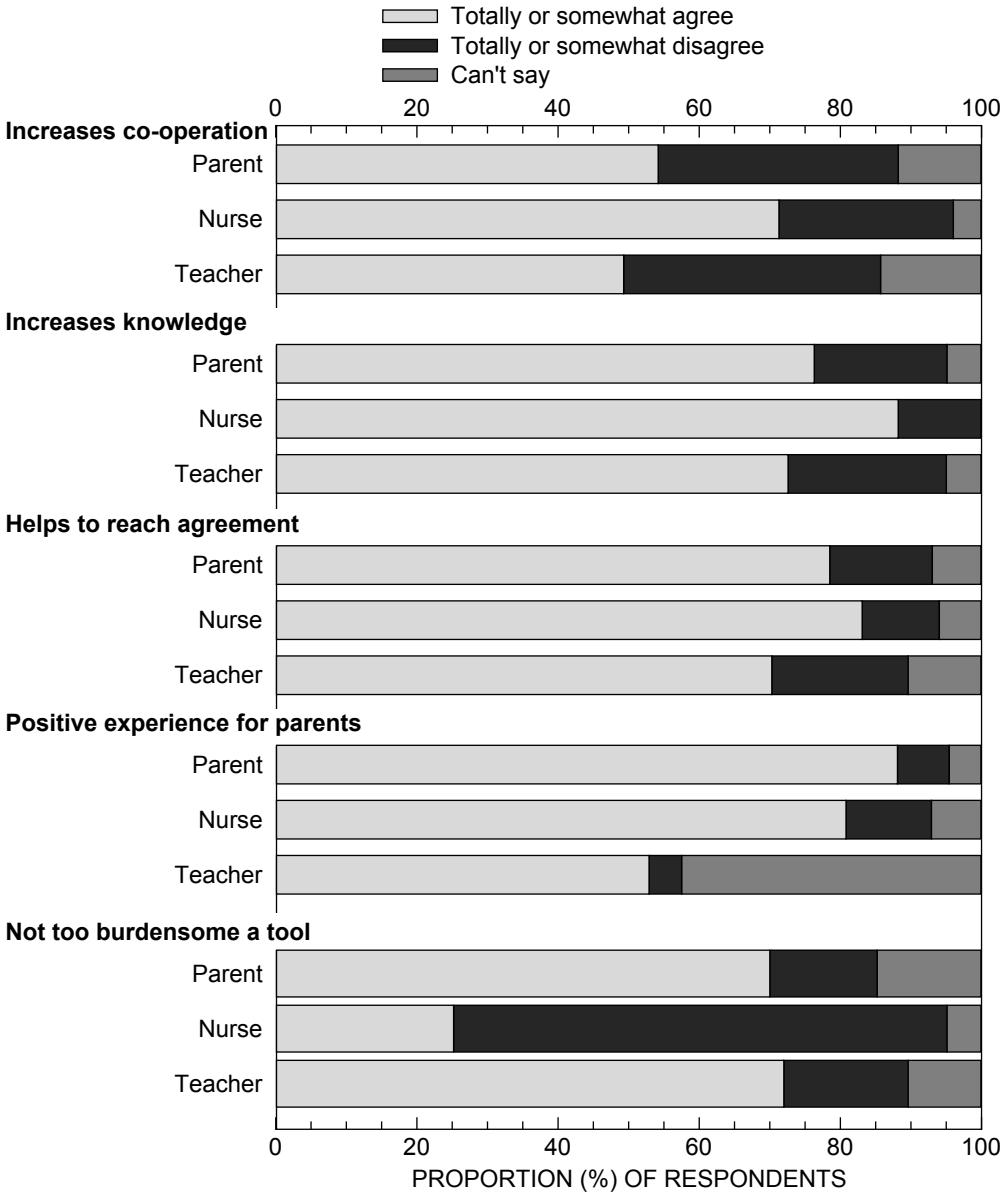


Figure 1. Opinions of the parents (n = 1546) of 4–9-year-old children, public health nurses (n = 156) and pre-school teachers (n = 1 286) about the feasibility of the Strengths and Difficulties Questionnaire

teachers were observation of the child (45%), discussions with parents (45%), making an annual early education plan for the child with parents (40%) and written feedback on the development of the child to the public health nurse (33%). Consultation with early education specialists, teamwork with other professionals and planning special guidance were mentioned as methods used in preschool education if concern about the child had emerged.

When the experiences of the SDQ and previous methods were compared, 43% of the public health nurses (n = 82) and 33% of the pre-school teachers (n = 863) reported that using the SDQ increased focusing on the mental health of the child. The public health nurses also praised the SDQ for yielding more information on the child's health (17%), and the pre-school teachers reported that the method enhanced structure and clarity in assessing the child (16%). Both respondent groups criticised the questionnaire for being difficult to interpret; open text fields would have facilitated answering according to 20% of the teachers and 8% of the public health nurses.

In addition, a proportion of parents (n = 449), public health nurses (n = 77), teachers in pre-school education (n = 653) and at school (n = 67) wrote direct and informal feedback on their experiences of the SDQ. In these open text fields some of the comments recapitulated the features elicited in the close-ended questions and items about contributing to collaboration between informants, the suitability and burden of the method. Teachers in pre-school education (24%) and at school (34%) also reported difficulties in filling in the questionnaire: they would have needed more open-ended options, the three-step scale of the item scoring was found too narrow and they wanted the strengths of the child to be pointed out more clearly. In addition, 9% found the formulation of the questionnaire better suited to school-aged children than to children in pre-school education. The theme most commonly (36%) mentioned by the parents was the contradictory and confusing experiences of using the SDQ in the medical check-ups: lack of dialogue with and interest on the part of professionals, and lack of feedback on the thoughts of the professionals about the child and the family. Secondly (31%), the parents gave positive and supportive feedback on paying attention to the mental health of the child and they pointed out the importance of including this in regularly administered health check-ups. They also felt it was good to focus on the well-being of their child. Parents reported similar difficulties in filling in the questionnaire (21%) as did the teachers.

Discussion

To the best of our knowledge this is the first study to carefully examine the feasibility of the SDQ in naturalistic settings in primary health care, in pre-school education and at school. The impact of using the screening questionnaire in assessing children's mental health on cooperation between parents and professionals was also evaluated.

The desirable aspects of the method, brevity, simplicity and availability, were mostly subscribed to in the present findings. It took a maximum of 10–15 minutes to complete the SDQ, and only the public health nurses reported the use of the SDQ to be rather burdensome. As respondents become more familiar with the method, there may be a practice effect on reducing administration time and on experiences of simplicity. However, all respondent groups criticised the questionnaire to some extent for being difficult to interpret and complete. Comparing the simplicity of the SDQ with other similar methods is difficult because so few standardised questionnaires are available and used in assessing children's mental health in primary health care in the first place. The SDQ has been made available free of charge on the Internet and it has also been translated into more than 70 languages (Youthinmind 2013). One carefully validated translation (including, e.g., back-translation process) per language will secure the reliability properties of the method and make it possible to compare the psychometric and feasibility properties of the method in different cultures. It also helps informants interpret the questions if the translation is fluent and easy to understand.

The public health nurses play an integral role in monitoring children's mental health and therefore their critical feedback considering the SDQ to be somewhat burdensome and difficult to interpret must be addressed. These difficulties may suggest that more training is needed before using the SDQ. In addition, given the attempt to introduce the SDQ into routine practice as a screening method, it is essential to ascertain how to support and treat the children identified to

have mental health problems. Early support and treatment call for developing networks of child mental health work and treatment plans for children's mental health problems. Multiple aspects, such as symptoms, functioning in various environments and family aspects, need to be considered when evaluating a child's need for treatment. Therefore, the nurses need opportunities to consult specialised mental health professionals. It takes efforts to comprehensively monitor and support the development and health of children in regular health checks and it is, presumably, challenging to allocate time and resources for screening mental health problems. Thus, many aspects need to be considered to enhance the use of the SDQ in busy primary care clinics.

The relevance, acceptability and value of the SDQ were evaluated as high in Finnish routine medical check-ups for children. The method had also earlier attracted wide acceptance in the studies aimed to ascertain what service users think about routine outcome measurement in child and adolescent mental health services (Moran *et al*. 2012, Stasiak *et al*. 2013). In our study, the SDQ was found relevant to the parents and front-line professionals in evaluating children's mental well-being. The SDQ was also accepted as an age-appropriate method for assessing the psychosocial well-being of the child by parents, public health nurses and teachers. However, the response rates of the SDQ were moderate, if acceptability is evaluated by response activity (Fitzpatrick *et al*. 1998). A measure will be seen as valuable if it measures what it is intended to measure and results in a more comprehensive assessment than would otherwise have been the case (Slade *et al*. 1999). The value of the SDQ among public health nurses and pre-school education personnel was demonstrated in the higher level of confidence in the SDQ than in previously used methods. The professionals reported that the SDQ helped them focus on the mental health of the child in a more structured way than before as there has been a lack of national standards and uniform methods in primary health care to evaluate child's psychosocial health. In addition, the finding was that the SDQ helped to learn more and to reach agreement between parents and professionals about the child's mental health and possible need for support. Parents also appreciated and supported the inclusion of mental health in the children's regularly administered health check-ups.

The findings of this study also gave further information about the impact of using the SDQ in cooperation between parents and front-line professionals. Assessing the mental health of the child is a sensitive matter for parents, thus their experiences of the method and its use are extremely important. Using the SDQ was a positive experience for parents and this was also noticed by the professionals. Parents seemed to have approved of the use of a questionnaire, a structured measure, when assessing the mental health of the child. They also knew that the SDQ was equally administrated to all the other parents visiting the health checks. These aspects might have introduced a neutral atmosphere to evaluate both the child's strengths and difficulties. However, the parents also expected a more interested, conversational and forthcoming approach on the part of the professionals than they experienced when the SDQ was used to assess the mental health of the child. Education and training are needed for public health nurses to further improve their collaboration with families. Also in earlier studies, the consumers in child and adolescent mental health services have considered the collaboration with clinicians more important than the choice of individual measures (Stasiak *et al*. 2013).

Some design limitations must be considered when interpreting the results of the study. The possible effects of the moderate participation rates in the first phase of the study, when collecting the SDQ questionnaires, have been discussed elsewhere (Borg *et al*. 2012). No data on reasons for non-completion of the Feasibility Questionnaire of the SDQ were collected, nor were non-respondents characterised. However, the evaluation of the method by non-respondents may have differed from that of respondents. If, for example, participants in the study and respondents to the Feasibility Questionnaire were more favourably disposed in the first place, this might have biased the results towards positive evaluations. The SDQ was a new approach for the participants in assessing children's mental health, which may have burdened the professionals more than their normal routines in clinical work. For these reasons the results may be biased towards negative evaluations of the feasibility of the SDQ. The study was intended to be carried out in as natural a clinical setting as possible, but it has been discussed before that a clinical setting which becomes the subject of research immediately becomes atypical (Slade *et al*. 1999). The participating clinics

in this study may have been atypical to begin with, and awareness of being monitored may also have changed the participants' behaviour.

The strengths of this study are the large sample size and the multi-informant approach. The feasibility of the screening method was evaluated by those front-line services informants in key roles in assessing children's well-being. To outweigh conceptual difficulties of assessing feasibility we tried to carefully evaluate numerous aspects of the feasibility of the method under consideration in routine clinical practice. The findings of this study are relevant in implementing screening methods in child mental health services because little published information is available on the feasibility aspects of any methods in this field.

According to our results, the SDQ is a feasible method for monitoring 4–9-year-old Finnish children's mental health in collaboration with public health nurses, parents and teachers in pre-school education and at school. However, the feasibility of a measurement method is context dependent (Slade *et al*. 1999) and these findings of the feasibility of the SDQ can thus be generalised only partly across cultures. Further research is also needed on the feasibility of the SDQ among older children and adolescents and in other operating environments. Presumably, the practical properties of the brevity, simplicity and availability of the SDQ are mostly applicable across cultures and have significance specifically for low income countries. However, comparative results are needed considering parents' and professionals' experiences about the relevance, acceptability and value of the SDQ and its impact on cooperation in other cultures.

Conclusions

The SDQ was found to be a feasible method for assessing children's mental health and met the practical criteria for a good screening questionnaire in primary health care: ease of use, helpfulness in increasing knowledge and common understanding about the child's mental health and possible need for support. The public health nurses reported, however, the SDQ to be somewhat burdensome. Using the SDQ was a positive experience for parents, but they also reminded the professionals of the importance of sensitive dialogue when assessing the mental health of the child.

References

Allen JL, Blatter-Meunier J, Ursprung A, Schneider S. 2010. The Separation Anxiety Daily Diary: Child version: Feasibility and psychometric properties. *Child Psychiatry and Human Development* 41: 649–662.

Borg AM, Kaukonen P, Salmelin R, Joukamaa M, Tamminen T. 2012. Reliability of the Strengths and Difficulties Questionnaire among Finnish 4–9-year-old children. *Nordic Journal of Psychiatry* 66: 403–413.

Costello EJ, Egger H, Angold A. 2005. 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry* 44: 972–986.

Du Y, Kou J, Coghill D. 2008. The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China. *Child and Adolescent Psychiatry and Mental Health* 2: 8, doi: 10.1186/1753-2000-2-8.

Fitzpatrick R, Davey C, Buxton MJ, Jones DR. 1998. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 2: 1–74, doi: 10.3310/hta2140.

Gold J, Buonopane RJ, Caggiano RA, Picciotto M, Vogeli C, Kanner NT, Babcock RL, Li Z, Jellinek M, Drubner SJ, Sklar KJ, Murphy JM. 2009. Assessing outcomes in child psychiatry. *The American Journal of Managed Care* 15: 210–216.

Goodman R. 1997. The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 38: 581–586.

Goodman R. 1999. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 40: 791–799.

Goodman R. 2001. Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry* 40: 1337–1345.

Goodman R, Slobodskaya H, Knyazev G. 2005. Russian child mental health — a cross-sectional study of prevalence and risk factors. *European Child & Adolescent Psychiatry* 14: 28–33.

Hawes DJ, Dadds MR. 2004. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *The Australian and New Zealand Journal of Psychiatry* 38: 644–651.

Kieling C, Baker-Henningham H, Belfer M, Conti G, Ertem I, Omigbodun O, Rohde LA, Srinath S, Ulkuer N, Rahman A. 2011. Child and adolescent mental health worldwide: Evidence for action. *Lancet* 378: 1515-1525.

Koskelainen M, Sourander A, Kaljonen A. 2000. The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *European Child & Adolescent Psychiatry* 9: 277–284.

Martin AM, Fishman R, Baxter L, Ford T. 2011. Practitioners' attitudes towards the use of standardized diagnostic assessment in routine practice: A qualitative study in two child and adolescent mental health services. *Clinical Child Psychology and Psychiatry* 16: 407–420.

Marzocchi GM, Capron C, Di Pietro M, Duran Tauleria E, Duyme M, Frigerio A, Gaspar MF, Hamilton H, Pithon G, Simoes A, Therond C. 2004. The use of the strengths and difficulties questionnaire (SDQ) in southern European countries. *European Child & Adolescent Psychiatry* 13 2: II40–46.

Merikangas KR, He JP, Brody D, Fisher PW, Bourdon K, Koretz DS. 2010. Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics* 125: 75–81.

Moran P, Kelesidi K, Guglani S, Davidson S, Ford T. 2012. What do parents and carers think about routine outcome measures and their use? A focus group study of CAMHS attenders. *Clinical Child Psychology and Psychiatry* 17: 65–79.

Myers K, Winters NC. 2002. Ten-year review of rating scales. I: Overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child and Adolescent Psychiatry* 41: 114–122.

Obel C, Heiervang E, Rodriguez A, Heyerdahl S, Smedje H, Sourander A, Guethmundsson OO, Clench-Aas J, Christensen E, Heian F, Mathiesen KS, Magnusson P, Njarethvik U, Koskelainen M, Ronning JA, Stormark KM, Olsen J. 2004. The Strengths and Difficulties Questionnaire in the Nordic countries. *European Child & Adolescent Psychiatry* 132: II32–39.

Oesterheld JR, Haber J. 1997. Acceptability of the Conners parent rating scale and Child Behavior Checklist to Dakotan/Lakotan parents. *Journal of the American Academy of Child and Adolescent Psychiatry* 36: 55–63; discussion 63-64 WHAT IS THIS?.

Rothenberger A, Becker A, Erhart M, Wille N, Ravens-Sieberer U, BELLA study group. 2008. Psychometric properties of the parent Strengths and Difficulties Questionnaire in the general population of German children and adolescents: Results of the BELLA study. *European Child & Adolescent Psychiatry* 17: 99–105.

Skovgaard AM, Houmann T, Christiansen E, Landorph S, Jorgensen T, CCC 2000 Study Team, Olsen EM, Heering K, Kaas-Nielsen S, Samberg V, Lichtenberg A. 2007. The prevalence of mental health problems in children 1(1/2) years of age - the Copenhagen Child Cohort 2000. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 48: 62–70.

Slade M, Cahill S, Kelsey W, Powell R, Strathdee G, Valiakalayil A. 2001. Threshold 3: The feasibility of the Threshold Assessment Grid (TAG) for routine assessment of the severity of mental health problems. *Social Psychiatry and Psychiatric Epidemiology* 36: 516–521.

Slade M, Thornicroft G, Glover G. 1999. The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology* 34: 243–249.

Sorensen MJ, Thomsen PH, Bilenberg N. 2007. Parent and child acceptability and staff evaluation of K-SADS-PL: A pilot study. European Child & Adolescent Psychiatry 16: 293–297.

Stasiak K, Parkin A, Seymour F, Lambie I, Crengle S, Pasene-Mizziebo E, Merry S. 2013. Measuring outcome in child and adolescent mental health services: Consumers' views of measures. *Clinical Child Psychology and Psychiatry* 18: 519–535.

Wichstrom L, Berg-Nielsen TS, Angold A, Egger HL, Solheim E, Sveen TH. 2012. Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 53: 695–705.

Williamson A, Redman S, Dadds M, Daniels J, D'Este C, Raphael B, Eades S, Skinner T. 2010. Acceptability of an emotional and behavioural screening tool for children in Aboriginal community controlled health services in urban NSW. *The Australian and New Zealand Journal of Psychiatry* 44: 894–900.

Woerner W, Fleitlich-Bilyk B, Martinussen R, Fletcher J, Cucchiaro G, Dalgalarrondo P, Lui M, Tannock R. 2004. The Strengths and Difficulties Questionnaire overseas: Evaluations and applications of the SDQ beyond Europe. *European Child & Adolescent Psychiatry* 13: II47–54.

Youthinmind. 2013. SDQ, information for researchers and professionals about the Strengths & Difficulties Questionnaires. Available at http://www.sdqinfo.org/ [accessed at May 2014].

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 286939, 11 pages
http://dx.doi.org/10.1155/2014/286939

Research Article

Cutting a Long Story Short? The Clinical Relevance of Asking Parents, Nurses, and Young Children Themselves to Identify Children's Mental Health Problems by One or Two Questions

Anne-Mari Borg,[1,2] Raili Salmelin,[3] Matti Joukamaa,[3,4] and Tuula Tamminen[1,2]

[1] *Department of Child Psychiatry, School of Medicine, University of Tampere, 33014 Tampere, Finland*
[2] *Department of Child Psychiatry, Tampere University Hospital, PL 2000, 33521 Tampere, Finland*
[3] *School of Health Sciences, University of Tampere, 33014 Tampere, Finland*
[4] *Department of Psychiatry, Tampere University Hospital, PL 2000, 33521 Tampere, Finland*

Correspondence should be addressed to Anne-Mari Borg; anne-mari.borg@uta.fi

Received 4 July 2014; Revised 28 November 2014; Accepted 7 December 2014; Published 31 December 2014

Academic Editor: Veit Roessner

Background and Aims. Assessing young children's mental health is a crucial and challenging task. The aim of the study was to evaluate the clinical relevance of asking parents, nurses, and young children themselves to identify children's mental health problems by only one or two questions. *Methods*. In regular health check-ups of 4- to 9-year-old children ($n = 2682$), parents and public health nurses assessed by one question whether the child had any emotional or behavioral difficulties. The child completed a self-evaluation enquiry on his/her emotional well-being. A stratified proportion of the participating parents were invited to a diagnostic interview. *Results*. Sensitivities were fairly good for the parents' (68%), nurses' (65%), and their combined (79%) one-question screens. Difficulties identified by parents and nurses were major risks (OR 10–14) for any child psychiatric disorders ($P < 0.001$). The child's self-evaluation was related to 2-fold to 3-fold risks ($P < 0.05$) for any psychiatric diagnosis, for any emotional diagnosis, and for negative situational factors. *Conclusion*. The one-question screen for parents and public health nurses together quite adequately identified the young children with mental health problems. The child's self-evaluation provided relevant and complementary information on his/her mental health and especially emotional problems.

1. Introduction

Assessing young children's mental health is a challenging task in primary services. Children's mental health problems are a global burden [1, 2] but, in general, their comprehensive screening is still in its infancy. It is necessary to develop and document validated and appropriate methods of screening for children's early mental health problems.

There are many special challenges in evaluating young children's mental health. Firstly, it is important to anchor the child's socioemotional and behavioural problems within the context of the child's developmental level [3]. Front line workers and parents may find it difficult to identify the child's psychopathology from the typical course of psychosocial development. Secondly, the well-being of the child is dependent on her/his family support and it cannot be evaluated in isolation from the well-being of the family. In addition to parents' reports, information on the child's symptoms and impairment in other significant social environments is needed [4]. Thirdly, interpretation and integrating all of the multi-informant and multimethod data is difficult and time consuming.

Every informant's evaluations count because no single informant's ratings can be used as "a gold standard" by which to measure psychopathology in children [5]. Discrepancies are common in different informants' ratings of child psychopathology [5]. Children's behaviour is known to be context dependent and parents' and teachers' reports are usually assessed. However, young children are rarely asked to self-evaluate their well-being.

Standardized self-report questionnaires are usually validated for school-aged children over 11 years old [6–8]. Young

children's ability to provide reliable and useful information on their moods and feelings has been questioned. Evidence suggests that any agreement between the child's and adults' reports is stronger with respect to externalizing than to internalizing problems [4]. In addition, children tend to report more emotional symptoms than do their parents or teachers [4, 9–11]. It has recently been concluded that using only parental reports for assessing children's emotions results in an underestimation of emotional problems [11]. Thus it seems necessary to further develop and assess self-report methods also for young children.

Standardized screening questionnaires for children's mental health problems have been developed and some of them are well documented [6, 12, 13]. The number of screening tools available for assessing social-emotional functioning in the infant-toddler period and in preschool-aged children has also grown [3]. However, standardized questionnaires are seldom used regularly and comprehensively in monitoring children's mental health [14, 15]. Instead, asking ordinary questions of "How are you, how do you feel?" or "Do you have some difficulties or concerns?" seems to be the prevalent practices among health care professionals. Yet there is little evidence on how reliable and valid such ordinary concern questions are in identifying the children at risk for mental health problems.

Asking parents and teachers very shortly, by only one or a few questions, about their perceptions of the child's behavioural and emotional difficulties has been proven useful in recognising the children with mental health problems [13, 16, 17]. Ford et al. (2005) have found high values of specificity and negative predictive power for parental concerns evaluated by four questions, in screening child psychiatric disorders [16]. In that study, about half of the children of whom the parents reported at least one problem had a psychiatric disorder. The Strengths and Difficulties Questionnaire (SDQ) is a widely used short questionnaire in assessing children's mental health [12, 13]. In the first question on the SDQ impact supplement, the respondents are asked to evaluate whether the child has difficulties in one or more of the following areas: emotions, concentration, behaviour, or being able to get on with other people [18]. Notably, this one question has discriminated between community and clinical samples almost as well as the whole SDQ measure, and it has also predicted child psychiatric diagnosis quite accurately [13, 17].

Screening cost effectively for early problems in large groups of children necessitates multistage screening procedures [3]. In the present study, the focus of interest was on developing and testing as brief, simple, and easy-to-use a first-stage screening assessment tool as possible to identify children at elevated risk for mental health problems. The specific aims of the present study were

(1) to assess the reliability and validity of a one-question screen presented to parents and public health nurses in everyday clinical practice in identifying children suffering from mental health problems,

(2) to assess the clinical relevance of directly asking a young child to evaluate his/her emotional well-being.

2. Methods

2.1. Study Design. The study was a part of a project called "*Developing Children's Mental Health Work*, 2007–2009," conducted in two hospital districts in Finland from March 2008 to March 2009. Altogether 154 child health clinics and school health care clinics participated in 25 municipalities. The respective local ethics committees approved the study. Informed consent was obtained from all participating parents.

Public health nurses introduced the study to parents making appointments for their 4–9-year-old children's regular health check-up. Prior to the visit to the clinic the study information and questionnaires were sent to interested parents at home: an informed consent form and a sociodemographic questionnaire including a parent's one-question screen and Strengths and Difficulties Questionnaires (SDQ) for both parents. The participating parents also asked the child's teacher in preschool education or at school to complete the SDQ. The parents returned all these forms to the public health nurse when attending the check-up. The design and sample of this phase of the study have been described in more detail elsewhere [19].

During the health check-up the child completed a self-evaluation enquiry about his/her well-being with the help of the public health nurse. In addition, the public health nurses completed a nurse's one-question screen for every child having a health check-up. After the check-up visit a feedback questionnaire on the feasibility of the child's self-evaluation enquiry was completed anonymously by the participating parents and once by each public health nurse involved in the process.

A subgroup of the participating parents was invited to a Development and Well-Being Assessment (DAWBA) interview. The SDQs were used to divide the children into screen-positive (scoring at or above the British 80th percentile cutoff, according to any informant) and screen-negative (scoring below the British 80th percentile cutoff, according to every informant) subgroups after the check-up visit. Every parent of a screen-positive child was invited to the DAWBA interview. For every two screen-positive cases (at the beginning of the study for every such case) a parent of a screen-negative child, matched for child's age group and gender, was invited to the DAWBA. With the parent's permission the child's teacher was also asked to complete DAWBA as a questionnaire. The interview phase of the study has been described in more detail elsewhere [20].

2.2. Sample. The sample consisted of 4- to 6-year-old preschoolers in child health clinics and 7- to 9-year-old children in school health care. Families not speaking Finnish were excluded from the study. Altogether 4,178 eligible children (49.5% girls) and their parents were invited to participate in the study, 3/5 of them being preschoolers (n = 2,596), the rest being school-aged (n = 1,582). The participation rate in the total sample was 64.2% (n = 2,682).

The participating parents filled in the parent's one-question screen in 98.9% of the cases. Of the 2,682 participating children 97.8% completed the self-evaluation enquiry.

Both of these enquiries were available for 96.8% (n = 2,595) of the participating children. The public health nurses returned the nurse's one-question screen for 99.3% of participants.

Altogether 646 parental DAWBA interviews were available. Of these participants 67% were preschool-aged and 66% were boys. A teacher's report was available for 75% (n = 486) of the DAWBA participants.

Fifty-five percent of the participating parents and 68% of the public health nurses involved in the process (154/225) completed the feedback questionnaire on the feasibility of the child's self-evaluation enquiry.

2.3. Measures. The SDQ is a screening questionnaire for 4- to 16-year-olds to be completed by parents, teachers, and by 11- to 16-year-old children themselves [7, 12, 13]. In this study the Finnish version of the method, including both the symptom questionnaire and the impact supplement, was collected [18].

The SDQ symptom questionnaire consists of 25 items forming five subscales: emotional symptoms, conduct problems, hyperactivity/inattention difficulties, peer relationship problems, and prosocial behaviour. The items are scored as 1 for "somewhat true" and, depending on the item, as 0 or 2 for "not true" or "certainly true" and for analysis they were recoded as 0 to 2 for increasing severity. The scores from all the subscales except for the prosocial scale are summed to a total difficulties score in the range 0–40. Goodman [7] has proposed the 80th and 90th percentiles as provisional cutoffs for "borderline" and "abnormal."

The first question on the parent and teacher version of the SDQ impact supplement asks, "Overall, do you think that your child/this child has difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people?" The answering alternatives are No, Yes—minor difficulties, Yes—definite difficulties, and Yes—severe difficulties. The rest of the impact supplement questions enquire, if difficulties are reported, about the duration or chronicity of the difficulties, overall distress, social impairment, and burden to others. The first question was used in the analyses of the present study. Otherwise, the reliability and validity properties of the extended version of the SDQ in the sample of Finnish 4–9-year-old children have been represented elsewhere [19, 20].

In the parent's one-question screen, parents were asked to assess whether their child had any emotional problems or any difficulties in behaviour, concentration, or social skills. The enquiry was answered on a four-step scale (no difficulties, not many difficulties, quite many difficulties, and very many difficulties). The enquiry was slightly modified from the first question on the parent's SDQ impact supplement.

In the nurse's one-question screen, public health nurses assessed, based on clinical evaluation, whether the child had, overall, difficulties in one or more of the following areas: emotions, behaviour, concentration, or being able to get on with other people. This was consistent with the first question in the parent's and teacher's SDQ impact supplement. The enquiry was answered on a five-step scale (no, yes/minor difficulties, yes/definite difficulties, yes/severe difficulties, and cannot say). The last answering option was added to the original alternatives of the abovementioned question on the SDQ.

The child's self-evaluation enquiry on emotional well-being was developed for this study and consisted of two questions; see Figure 1. The written response alternatives had visual analogues in the form of facial expressions. The public health nurse read the questions and response alternatives to the child even if he/she could read. The child chose and marked with a cross the answer best describing his/her feelings.

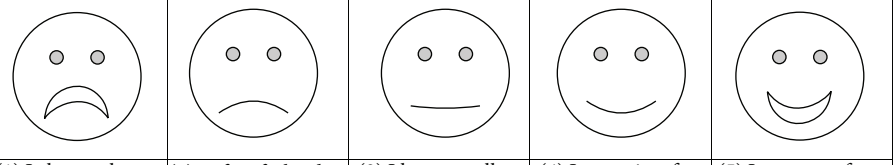
The *DAWBA* method [21] consists of a semistructured interview, which can be administered to the parents of children aged 5 to 17 and to children over 11 years themselves; there is also a briefer questionnaire version for teachers. The structured questions cover most child psychiatric disorders and closely follow the diagnostic criteria according to the ICD-10 and DSM-IV. If definite symptoms are identified, parents are asked to describe the problems in more detail.

According to the responses of all available informants on the structured questions, the DAWBA program assigns each child to a level of an ordinal-scale measure which represents the prevalence of any diagnosis in epidemiological samples [22]. The categorization of this predictive measure offered to the clinical rater is <1% (very low), <5% (low), ≥20% (moderate), and ≥75% (high) [23]. To decide on definitive diagnoses a clinical rater then reviews all relevant information: the structured, closed, and open accounts of all available informants and the computer-predicted level of prevalence of any diagnosis.

The first author reviewed all the interviews and assigned the diagnoses according to ICD-10. The diagnoses were placed in five categories: emotional, conduct, hyperactivity, and other diagnoses (Tic/Tourette, pervasive developmental disorders, and not otherwise specified mental disorders) and situational factors (Z61 problems related to negative life events in childhood, Z62 other problems related to upbringing, and Z63 other problems related to primary support group, including family circumstances). The rater was trained by practising with the cases in the training manual [23] and participating in a two-day training course. When the diagnoses were uncertain, a consensus diagnosis was obtained by a consultation group of four experienced child psychiatrists. The frequency of diagnoses set by the rater was compared with the computer-predicted level of prevalence of any diagnosis. The associations were statistically highly significant ($P < 0.001$) between all pairs of the following groups: in the low prevalence group (<5%) 3% of the children were assigned to diagnoses, in the moderate prevalence group (≥20%) 38%, and in the high prevalence group (≥75%) 93% of the children.

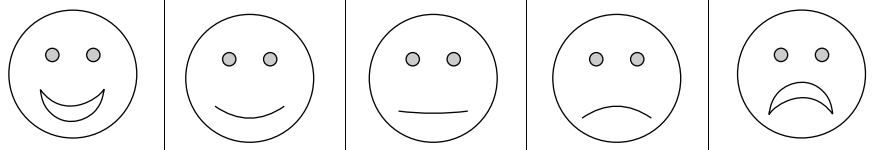
In the *feedback questionnaire* on the feasibility of the child's self-evaluation enquiry, parents and public health nurses were asked how appropriate this method was in assessing the psychosocial well-being of the child (very good/fairly good/not good, not poor/rather poor/very poor). In addition, the public health nurses were asked to report how long, on average, to the nearest five minutes, it took to complete the child's self-evaluation enquiry and how burdensome they found it (not at all/not very/rather/very burdensome).

Child's self-evaluation enquiry on emotional well-being
How are you?
Please choose the picture that best describes your life and feelings. Then mark the answer with a cross (X) above the picture you have chosen.



(1) I almost always feel sad or miserable.	(2) I often feel sad or miserable.	(3) I have equally many happy and miserable moments.	(4) I am quite often happy and in a good mood.	(5) I am very often happy and in a good mood.

What do you expect for your near future? What will your life be like?
Please choose the picture that best describes how you feel. Then mark the answer with a cross (X) above the picture you have chosen.



(1) I think my future will be very nice and happy.	(2) I think my future will be fairly nice and happy.	(3) I don't really worry about future.	(4) I suspect that some bad things are going to happen to me.	(5) I suspect that many bad things are going to happen to me.

Thank you for your answers!

Figure 1: Child's self-evaluation enquiry on emotional well-being.

2.4. Statistical Analyses. The distributions of the multicategory parent's and nurse's one-question screens and the child's self-evaluation questions are expressed as percentages, and cross-informant agreements between them were examined with the γ coefficient. Because of the requirements of further analysis (examination of validity properties and logistic regressions) these questions were dichotomised in such a way that the upper category would include children with the strongest concerns and still be large enough for the analysis (see Table 1). Consequently, the categories were as follows: the parent's one-question screen, no/not many difficulties versus quite many/very many difficulties; the nurse's one-question screen, no/minor difficulties versus definite/severe difficulties; the first question of the child's self-evaluation enquiry (How are you?), very/quite often happy/as many happy as miserable moments versus often/almost always sad; and the second question (What do you expect?), very/fairly nice and happy future, do not really worry versus some bad/many bad things are going to happen to me. Furthermore, the answers of the two child's self-evaluation questions were combined as positive in both questions versus other combinations.

The DAWBA computer-predicted level of any child psychiatric diagnosis, dichotomised, according to the same principles as above, as <75% versus ≥75% (high prevalence level), was used as the gold standard in assessing the validity properties (sensitivity, specificity, and positive and negative predictive values [PPV, NPV]) of the abovementioned one-question screen for the parent, nurse, and child. All relevant two-variable analyses were conducted by comparing both age and gender groups or stratifying by them.

The DAWBA variable and the existence of any or selected specific diagnoses assigned by the rater were, one at a time, used as outcome variables in a set of logistic regression analyses. In the first set the only explanatory variable was one of the one-question screen variables (parent, nurse, and child) at a time, and the enter-method used thus produced unadjusted odds ratios (OR) for them. In the second set of logistic regression analyses, the explanatory variables comprised all one-question screens, gender, age group, and their interaction. To determine the strongest factors affecting the respective outcome variable, backwards stepwise method was used.

P values < 0.05 are considered to show statistical significance. The statistical analyses were accomplished with SPSS v. 19.

Table 1: Distributions of answers in the parents' and nurses' one-question screen and child's self-evaluation in a community sample of Finnish 4–9-year-old children.

	Total sample			χ^2	Preschoolers			School-aged children			χ^2
	All	Boys	Girls	df[1]	All	Boys	Girls	All	Boys	Girls	df[1]
	%	%	%	P[2]	%	%	%	%	%	%	P[3]
Parent's one-question screen	(n = 2652)	(n = 1295)	(n = 1357)	80.79	(n = 1757)	(n = 867)	(n = 890)	(n = 895)	(n = 428)	(n = 467)	13.88
No difficulties	47.0	38.7	55.0	2	46.3	38.2	51.2	48.5	39.7	56.5	2
Not many difficulties	46.6	52.4	41.0	<0.001	48.4	54.4	42.6	43.0	48.4	30.1	0.001
Quite many/very many difficulties	6.4	8.9	4.0		5.3	7.4	3.3	8.5	11.9	5.4	
Nurse's one-question screen[4]	(n = 2602)	(n = 1269)	(n = 1333)	81.73	(n = 1722)	(n = 847)	(n = 875)	(n = 880)	(n = 422)	(n = 458)	0.37
No difficulties	75.5	67.8	82.8	2	75.2	68.5	81.7	76.1	66.6	84.9	2
Yes—minor difficulties	17.7	22.5	13.1	<0.001	18.0	22.1	14.1	17.0	23.2	11.4	ns.
Yes—definite/severe difficulties	6.8	9.7	4.1		6.8	9.4	4.2	6.8	10.2	3.7	
Child's self-evaluation											
How are you?	(n = 2623)	(n = 1287)	(n = 1336)	22.86	(n = 1739)	(n = 860)	(n = 879)	(n = 884)	(n = 427)	(n = 457)	446.33
Very often happy	54.0	51.4	56.4	3	68.0	66.3	69.6	26.5	21.5	31.1	3
Quite often happy	28.9	28.0	29.7	<0.001	18.5	17.1	19.8	49.3	49.9	48.8	<0.001
Equally many happy and lousy moments	15.1	17.7	12.5		10.9	13.0	8.9	23.2	27.2	19.5	
Often/almost always sad	2.1	2.9	1.3		2.6	3.6	1.7	1.0	1.4	0.7	
What do you expect for your near future?	(n = 2620)	(n = 1284)	(n = 1336)	10.16	(n = 1737)	(n = 858)	(n = 879)	(n = 883)	(n = 426)	(n = 457)	115.01
Very nice and happy days	50.1	48.5	51.6	3	55.8	55.4	56.3	38.8	34.7	42.7	3
Quite nice and happy days	29.4	28.8	30.0	<0.017	23.0	22.3	23.8	42.0	42.0	42.0	<0.001
Not bothering	15.7	18.0	13.5		15.4	17.0	13.9	16.2	20.0	12.7	
Some/many bad things are going to happen	4.8	4.7	4.9		5.7	5.4	6.0	2.9	3.3	2.6	

[1]df: degrees of freedom.
[2]Tested between genders.
[3]Tested between age groups.
[4]Answering alternatives "do not know" and "cannot say" extracted from analyses (n = 80).

3. Results

3.1. Distributions of Parents' and Nurses' Perceptions and Children's Self-Evaluations. Table 1 shows the distributions of parents' and nurses' perceptions of the child's difficulties and children's self-evaluations in the total sample, stratified by gender and age groups. Six to seven percent of the children were evaluated by both the parents and the public health nurses to have definite or severe difficulties. According to both sets of informants' reports the proportion of boys having such difficulties was at least twice that of girls ($P < 0.001$) in the total sample and in both age groups. In addition, in parent's evaluations, the school-aged children were evaluated to have more commonly difficulties (8.5%) than the preschool-aged children (5.3%). The public health nurses could not say or did not know about the child's situation in 3.0% ($n = 80$) of the cases.

Of the children 2.1% evaluated themselves as feeling often or almost always sad or miserable (Table 1). Boys reported such negative feelings twice as commonly as girls ($P < 0.001$). Boys also reported more commonly than girls having "as many happy and miserable moments." In the second question, 4.8% of children expected some or many bad things to happen. Younger children reported more commonly than older children negative feelings and future expectations.

Cross-informant agreement between the parents' and nurses' perceptions was fairly good ($\gamma = 0.73$) in the total sample. The agreements between child's self-evaluation and adults' evaluations were very low (child-parent $\gamma = 0.10$ and child-nurse $\gamma = 0.15$).

The agreement between the parent-rated one-question screen and the first question on the SDQ impact supplement was $\gamma = 0.92$.

3.2. Validity of the One-Question Screen against the Diagnostic Assessment. The sensitivity of the dichotomized parent's and nurse's one-question screen against the DAWBA computer-predicted high prevalence level of any diagnosis was fairly good (68% and 65%, resp., Table 2). The respective specificities were high (87-88%). PPVs were low and NPVs high. The sensitivity and PPV of the child's self-evaluation enquiries were very low (7–26%) and the specificity and NPV high (89–98%).

The sensitivities of the adult informants' perceptions were considerably higher and the specificities somewhat lower for boys than for girls (Table 2). In addition, the sensitivities were higher for older than for younger children. There were no differences in the values between the genders regarding the child's self-evaluation questions except in the second question, where the sensitivity for girls was higher than for boys (18% versus 6%). The PPV and NPV of the nurse's one-question screen and the child's self-evaluation questions were lower for boys than for girls, contrary to the results of the parent's responses.

Combining two or three of the informants' reports produced higher sensitivity than any of the respective single informants' reports. In the total sample, the sensitivity of the combined child's self-evaluation was 14% and the specificity was 93%. The sensitivity of both the combination of the parent's and nurse's perceptions and that of combining all three informants' reports was 79%, the respective specificities being 80% and 75%.

3.3. Risks for Child Psychiatric Disorders Related to the One-Question Screens. If parents or nurses identified difficulties the odds ratios for any and selected specific child psychiatric disorders were all statistically highly significant ($P < 0.001$, Table 3). The highest odds ratio (OR) related to difficulties identified by parents was that for a DAWBA computer-predicted high prevalence level (≥75%) of a child psychiatric diagnosis (OR 14.4), and the lowest (OR 4.5) was that for an emotional diagnosis. Nurse's assessment of definite or severe difficulties was most strongly associated with a hyperactivity diagnosis (OR 34.4) and least strongly with an emotional diagnosis (OR 4.0).

The negative rating in the combined child's self-evaluation was statistically significantly ($P < 0.05$) associated with a DAWBA computer-predicted high prevalence level of any diagnosis (OR 2.2), with any DAWBA-rater assigned diagnosis (OR 2.4), with any emotional diagnosis (OR 3.0), and with negative situational factors (OR 3.2).

Examining the effects of the evaluations of all three informants simultaneously by backwards stepwise logistic regression revealed that the difficulties identified by parents and nurses remained the strongest and significant risk factors for all child outcomes (OR 2.7–7.1), except in predicting a hyperactivity diagnosis, where only difficulties identified by the nurses remained statistically significant (OR 20.9); see Table 4. The child's self-evaluation remained a statistically significant risk factor for any emotional diagnosis (OR 2.7) and for negative situational factors (OR 2.9). Girls had higher risk than boys for any emotional diagnosis (OR 2.3) and school-aged children had higher risk than preschoolers for any assigned diagnoses (OR 1.8).

3.4. Feedback on the Child's Self-Evaluation Enquiry. The child's self-evaluation enquiry was considered to be very or fairly age-appropriate for assessing the child's psychosocial well-being by 63% of the parents and by 71% of the public health nurses in the total sample. Eight percent of the parents and 9% of the public health nurses evaluated the appropriateness of the enquiry to be rather or very poor. Fourteen percent of the parents had no opinion on the subject.

Most (96%) public health nurses found the two questions not very or not at all burdensome. Almost all (99%) the public health nurses completed the enquiry with the children in 10 minutes or less.

4. Discussion

The main results of the study suggested that the one-question screen presented to parents and public health nurses offers a valid and clinically relevant guide in identifying children suffering from mental health problems. It is also useful to hear the young child's own perspective when trying to identify children at high risk, especially for emotional problems.

Table 2: Sensitivity, specificity, and positive (PPV) and negative (NPV) predictive values of the parent's and nurse's one-question screen and child's self-evaluation questions (no or mild concerns/problems versus more severe options) calculated against the DAWBA computer-predicted prevalence level of any diagnosis (<75% versus ≥75%) in a sample of Finnish 4–9-year-old children.

	High prevalence level of diagnosis																			
	Total sample				Boys				Girls				Preschoolers				School-aged children			
	Sens[1] %	Spec[2] %	PPV[3] %	NPV[4] %	Sens[1] %	Spec[2] %	PPV[3] %	NPV[4] %	Sens[1] %	Spec[2] %	PPV[3] %	NPV[4] %	Sens[1] %	Spec[2] %	PPV[3] %	NPV[4] %	Sens[1] %	Spec[2] %	PPV[3] %	NPV[4] %
Parent's one-question screen		($n = 637$)				($n = 423$)				($n = 214$)				($n = 424$)				($n = 213$)		
	68	87	41	95	73	86	45	96	50	89	30	95	70	90	44	96	65	82	39	93
Nurse's one-question screen		($n = 622$)				($n = 411$)				($n = 211$)				($n = 414$)				($n = 208$)		
	65	88	41	95	69	85	42	95	56	92	40	96	72	88	42	97	55	87	40	92
Child's self-evaluation																				
How are you?		($n = 629$)				($n = 418$)				($n = 211$)				($n = 418$)				($n = 211$)		
	7	98	26	89	8	97	25	88	6	99	33	92	10	97	29	91	3	98	20	86
What do you expect for your near future?		($n = 628$)				($n = 417$)				($n = 211$)				($n = 418$)				($n = 210$)		
	9	94	16	89	6	95	14	87	18	93	19	93	10	94	15	91	7	96	20	86
Combined Child's self-evaluation		($n = 629$)				($n = 418$)				($n = 211$)				($n = 418$)				($n = 211$)		
	14	93	20	90	13	93	21	88	18	93	19	93	18	92	19	91	10	95	23	86
Combined parent's and nurse's one-question screen		($n = 613$)				($n = 405$)				($n = 208$)				($n = 407$)				($n = 206$)		
	79	80	34	97	85	77	36	97	61	84	27	96	84	82	35	98	72	75	32	94

[1]Sensitivity; [2]specificity; [3]positive predictive value; [4]negative predictive value.

Table 3: The odds ratios (OR) for child outcomes related to parent's and nurse's evaluation of the child's difficulties and child's self-evaluation of emotional well-being according to DAWBA assessment in a sample of Finnish 4–9-year-old children ($n = 646$). The OR of each separate evaluation (no or mild difficulties/concerns versus more severe options) for each outcome measure is shown.

	Computer-predicted prevalence[1]	Rater-assigned child psychiatric ICD-10 diagnosis					
		Any	Emotional	Conduct	Hyperactivity	Other[2]	Situational factors[3]
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Parent's concern enquiry	14.4** (8.4–24.9)	9.9** (6.3–15.6)	4.5** (2.6–19.2)	9.9** (5.1–18.9)	8.1** (3.9–16.8)	9.7** (4.0–23.2)	7.3** (3.8–14.0)
Nurse's concern enquiry	13.6** (7.8–23.5)	12.4** (7.8–19.7)	4.0** (2.2–7.1)	10.0** (5.2–19.3)	34.3** (12.9–91.1)	5.8** (2.5–13.3)	8.2** (4.2–16.0)
Child's self-evaluation (two questions combined)	2.2* (1.1–4.7)	2.4* (1.3–4.5)	3.0* (1.4–6.5)	2.0 (0.85–5.0)	1.2 (0.4–4.2)	1.1 (0.3–5.0)	3.2* (1.4–7.5)

[1] Prevalence level <75%/≥75%.
[2] Tic/Tourette, pervasive developmental disorders, and not otherwise specified mental disorders.
[3] Factors influencing health status and contact with health services (ICD-10): Z61 problems related to negative life events in childhood, Z62 other problems related to upbringing, and Z63 other problems related to primary support group, including family circumstances.
* $P < 0.05$.
** $P < 0.001$.

Table 4: The odds ratios (OR) for child outcomes related to the combined effects of the parent's and nurse's one-question screen and the child's self-evaluation of emotional well-being as well as child's gender and age group. The OR for the variables remaining in the model at the last step of each backwards stepwise logistic regression are shown.

Variables entered into each model	Computer-predicted prevalence[1] ($n = 68$)	Rater-assigned child psychiatric ICD-10 diagnosis					
		Any ($n = 117$)	Emotional ($n = 53$)	Conduct ($n = 41$)	Hyperactivity ($n = 32$)	Other[2] ($n = 23$)	Situational factors[3] ($n = 38$)
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Parent's one-question screen[4]	6.7** (3.6–12.7)	4.3** (2.5–7.4)	2.7* (1.3–5.5)	4.4* (2.0–9.6)	2.1 (0.9–5.1)	4.7* (1.7–13.0)	4.0* (1.8–8.9)
Nurse's one-question screen[4]	6.6** (3.5–12.5)	7.1** (4.1–12.1)	2.9* (1.4–5.9)	4.9** (2.3–10.7)	20.9** (7.2–60.4)	3.1* (1.1–8.4)	3.5* (1.6–7.7)
Child's self-evaluation[4] (combined)		2.1 (1.0–4.6)	2.7* (1.2–6.2)				2.9* (1.1–7.4)
Child's gender[5]			2.3* (1.3–4.3)		0.4 (0.1–1.1)		
Child's age[6]		1.8* (1.1–2.9)				2.5 (1.0–6.2)	
Gender ∗ age	—[7]					—[7]	

[1] Prevalence level <75%/≥75%.
[2] Tic/Tourette, pervasive developmental disorders, and not otherwise specified mental disorders.
[3] Factors influencing health status and contact with health services (ICD-10): Z61 problems related to negative life events in childhood, Z62 other problems related to upbringing, and Z63 other problems related to primary support group, including family circumstances.
[4] No or mild difficulties/concerns versus more severe options.
[5] Girls versus boys.
[6] School-aged versus preschool children.
[7] The variable remained in the model but OR could not be computed because there were too few cases in some of the subgroups.
* $P < 0.05$.
** $P < 0.001$.

Of the children having regular health check-ups 6-7% were evaluated by the parents or the public health nurses in this study to have definite or severe difficulties. In a British epidemiological sample 9.5% of parents reported concerns about their child's emotions, behaviour, or activity level [24]. The present finding compares closely with earlier 5–24% frequencies of psychiatric symptoms or disorders in population samples of young children [20, 25–28]. As in some earlier studies boys were more commonly than girls reported to have difficulties [25, 28, 29] or any disorder [28].

The present finding that older children had more parent-rated difficulties than younger children should be considered preliminary. The earlier findings on differences between the score distributions split by comparable age groups have been inconsistent according to the computerised multicultural norms of the SDQ [18].

In the present study only 2% of children evaluated themselves as feeling often or almost always sad or miserable. The child's self-reported frequency of emotional problems in this study was lower than in earlier studies with validated assessment methods [30, 31]. Recently, 12–16% of Belgian 5- to 10-year-old children reported emotional problems, such as anger, anxiety, and sadness, in a short self-report questionnaire [11]. Further studies are needed on young children's self-reported frequencies of emotional problems in community and clinic samples.

The cross-informant agreement between the parent's and public health nurse's perceptions was fairly good but the agreements between child's self-evaluation and adults' evaluations were very low. In the present study, it was not our purpose to compare the child's and the adults' reports because adult informants answered a similar question on overall difficulties whereas the child's questions focused on his/her emotional well-being and expectations. However, the weak correlation between the parent's and the child's evaluations was in concordance with earlier studies [4, 10, 11]. As expected, the agreement between the parent's one-question screen and the first question on the SDQ impact supplement was high. The present one-question screen was only slightly modified from the original abovementioned question on the SDQ.

The single question for the parent and public health nurse had an adequate capacity to discriminate between the low-risk and high-risk children in the sample. The child's self-evaluation, however, was not sensitive for identifying high-risk children. The one-question screen for the parent and public health nurse detected two-thirds of the children with a psychiatric disorder and the specificity, the proportion of true negatives, of the adults' evaluations was high. Of the children identified as having difficulties, 41% had a computer-predicted DAWBA diagnosis and of those children identified as having no difficulties only 5% had a respective diagnosis. Thus the parent's and public health nurse's perceptions of difficulties were found to be fairly good and evaluations of no concern about the child's situation were quite accurate.

The present values of sensitivity and specificity for the one-question screen for parents concur closely with earlier results on the validity values of screening questionnaires [6, 12, 32]. We replicated the earlier finding that a single question on whether the child has emotional or behavioural difficulties discriminates almost as well as a whole questionnaire comprising many items between low-risk and high-risk children [13]. Further information was also gained on the effect of combining two or three different informants' answers, which was found to produce higher sensitivity values compared to a single informant's report. The parents and nurses together identified four-fifths of the children with a psychiatric diagnosis. Obviously, the combined parent's and public health nurse's one-question screen seemed to be a good indication for a more comprehensive evaluation of the child's mental health.

Difficulties identified by parents and nurses were found to be strong and statistically significant risk factors for any child psychiatric disorders. The highest odds ratio for parental perception of difficulties was found for any child psychiatric diagnosis (OR 14.4) and for public health nurse's respective perception for a hyperactivity diagnosis (OR 34.4). Both informants' concerns had the lowest OR for an emotional diagnosis, being still a fourfold to fivefold risk. Difficulties identified by parents and nurses remained the strongest risk factors for most of the child's outcomes when all the predictors, including child's age and gender, were taken into account. The present strong association between difficulties identified by parents and a child's psychiatric diagnosis is comparable to the earlier finding of a strong association (OR 16) between high scores on the SDQ parent report and a child psychiatric disorder [12]. The present results suggest that the one-question screen for parents and public health nurses yields validated and supplementary information on the child's risks for mental disorders.

When the young children reported low mood or negative expectations this was related ($P < 0.05$) to elevated risks for a psychiatric disorder, emotional disorders, and negative situational factors in the family. When taking all the risk factors into account, the child's self-evaluation remained as a statistically significant threefold risk for any emotional diagnosis and for negative situational factors. For an emotional diagnosis girls were found to have a twofold risk compared to boys. Older children had a twofold risk for any assigned diagnoses compared to younger children. The present findings confirmed that the young child's self-evaluation yields relevant and complementary information on the child's emotional well-being from the child's inner perspective.

Although the screening properties of the one-question screen were quite adequate, one-fifth of the children with a child psychiatric disorder were not identified by their parents or public health nurses. The clinicians should remember that even if they use a standardized screening method there will remain a proportion of these "false negative" children. It is a special challenge to try to identify these children in need of psychosocial support. In addition, whatever screening method was used it needs to be administered systematically in order to produce reliable results.

Several limitations should be noted in the study. The child's self-evaluation enquiry had not been tested before, and therefore further studies are needed on the psychometric screening properties of similar brief screening assessments for children. Possible effects of the moderate participation rates in the first phase and in the interview phase of the study as well as limitations related to diagnostic procedures have been discussed elsewhere [19, 20].

The strengths of the study were the multi-informant approach and the large sample of young children. The study was conducted in an everyday clinical setting of children's regular health check-ups, thus improving the usability of the results. The discriminative validity properties of the single questions were assessed against a diagnostic assessment as

a gold standard. This made it possible to explore the child's symptoms and level of impairment in a multi-informant approach and in different contexts, also including the child's teacher's report. The study presents further information about the very brief screening assessments for parents and public health nurses.

The present study generated new information about directly asking a young child to evaluate his/her emotional well-being by two pictorial questions in a clinical setting. We found no earlier corresponding studies. Some pictorial self-report questionnaires for preschool and for young schoolchildren are available [4, 33–35]. The correspondence between self-reports of children and the reports of parents and teachers was not altered when pictorial self-report questionnaires were used instead of traditional verbal self-report instruments [4]. The use of pictures combined with verbal questions, however, was assumed to help children in communicating their opinions in the present study. The parents and public health nurses gave positive feedback on the feasibility of the child's self-evaluation enquiry in the context of regular health check-ups.

5. Conclusions

The results suggest that the one-question screen for parents and public health nurses together adequately identifies those young children with mental health problems and can thus be considered as a first step screening assessment in everyday clinical front-line practice. In addition, the young child's self-evaluation questions yielded complementary and relevant information on their mental health and especially emotional problems, speaking for the importance of directly asking the child's own perspective.

Conflict of Interests

The authors declare no conflict of interests in preparing this paper.

Acknowledgments

The authors thank the participating families and the public health nurses for participating in the study and the project coordinator, Nurse Sari Miettinen at the Child Psychiatric Clinic, Tampere University Hospital, for collaboration in data collection. They are also grateful to Pälvi Kaukonen, M.D., the Chief Medical Officer at the Department of Child Psychiatry in the Tampere University Hospital, for her general guidance, support, and feedback during various stages of the study. The data were collected on the project "Developing Children's Mental Health Work, 2007–2009," funded by the hospital districts of Pirkanmaa and South Karelia and the Ministry of Social Affairs and Health, Finland. This study was supported by the Finnish Brain Foundation and the Research Foundation of Tampere.

References

[1] E. J. Costello, H. Egger, and A. Angold, "10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 44, no. 10, pp. 972–986, 2005.

[2] C. Kieling, H. Baker-Henningham, M. Belfer et al., "Child and adolescent mental health worldwide: evidence for action," *The Lancet*, vol. 378, no. 9801, pp. 1515–1525, 2011.

[3] A. S. Carter, M. J. Briggs-Gowan, and N. O. Davis, "Assessment of young children's social-emotional development and psychopathology: recent advances and recommendations for practice," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 45, no. 1, pp. 109–134, 2004.

[4] E. M. Ederer, "Mental health problems in young children: self-reports and significant others as informants," *Psychology Science*, vol. 46, supplement 1, pp. 123–140, 2004.

[5] A. de los Reyes and A. E. Kazdin, "Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study," *Psychological Bulletin*, vol. 131, no. 4, pp. 483–509, 2005.

[6] T. M. Achenbach and L. A. Rescorla, *Manual for the ASEBA School-Age Forms & Profiles*, Research Center for Children, Youth & Families, University of Vermont, Burlington, Vt, USA, 2001.

[7] R. Goodman, "The Strengths and Difficulties Questionnaire: a research note," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 38, no. 5, pp. 581–586, 1997.

[8] M. Kovacs, "Rating scales to assess depression in school-aged children," *Acta Paedopsychiatrica*, vol. 46, no. 5-6, pp. 305–315, 1981.

[9] I. Seifige-Krenke and F. Kollmar, "Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behaviour: a longitudinal analysis of parent-adolescent agreement on internalising and externalising problem behaviour," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 39, no. 5, pp. 687–697, 1998.

[10] B. van Roy, B. Groholt, S. Heyerdahl, and J. Clench-Aas, "Understanding discrepancies in parent-child reporting of emotional and behavioural problems: effects of relational and socio-demographic factors," *BMC Psychiatry*, vol. 10, article 56, 2010.

[11] N. Michels, B. Vanaelst, E. Stoppie et al., "Parental and children's report of emotional problems: agreement, explanatory factors and event-emotion correlation," *Child and Adolescent Mental Health*, vol. 18, no. 3, pp. 180–186, 2013.

[12] R. Goodman, "Psychometric properties of the Strengths and Difficulties Questionnaire," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 40, no. 11, pp. 1337–1345, 2001.

[13] R. Goodman, "The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 40, no. 5, pp. 791–799, 1999.

[14] J. Gold, R. J. Buonopane, R. A. Caggiano et al., "Assessing outcomes in child psychiatry," *The American Journal of Managed Care*, vol. 15, no. 4, pp. 210–216, 2009.

[15] M. J. Batty, M. Moldavsky, P. S. Foroushani et al., "Implementing routine outcome measures in child and adolescent mental health services: from present to future practice," *Child and Adolescent Mental Health*, vol. 18, no. 2, pp. 82–87, 2013.

[16] T. Ford, K. Sayal, H. Meltzer, and R. Goodman, "Parental concerns about their child's emotions and behaviour and referral to specialist services: general population survey," *The British Medical Journal*, vol. 331, no. 7530, pp. 1435–1436, 2005.

[17] A. Goodman and R. Goodman, "Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 52, no. 1, pp. 100–108, 2011.

[18] Youthinmind, "SDQ, Information for researchers and professionals about the Strengths and Difficulties Questionnaires," 2013, http://www.sdqinfo.org/.

[19] A.-M. Borg, P. Kaukonen, R. Salmelin, M. Joukamaa, and T. Tamminen, "Reliability of the Strengths and Difficulties Questionnaire among Finnish 49-year-old children," *Nordic Journal of Psychiatry*, vol. 66, no. 6, pp. 403–413, 2012.

[20] A. M. Borg, P. Kaukonen, M. Joukamaa, and T. Tamminen, "Finnish norms for young children on the Strengths and Difficulties Questionnaire," *Nordic Journal of Psychiatry*, vol. 68, no. 7, pp. 433–442, 2014.

[21] R. Goodman, T. Ford, H. Richards, R. Gatward, and H. Meltzer, "The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 41, no. 5, pp. 645–655, 2000.

[22] A. Goodman, E. Heiervang, S. Collishaw, and R. Goodman, "The "DAWBA bands" as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples," *Social Psychiatry and Psychiatric Epidemiology*, vol. 46, no. 6, pp. 521–532, 2011.

[23] Youthinmind, "DAWBA, Information for researchers and clinicians about the Development and Well-Being Assessment," 2012, http://www.dawba.com/.

[24] T. Ford, R. Goodman, and H. Meltzer, "The British child and adolescent mental health survey 1999: the prevalence of DSM-IV disorders," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 42, no. 10, pp. 1203–1211, 2003.

[25] C. Glazebrook, C. Hollis, H. Heussler, R. Goodman, and L. Coates, "Detecting emotional and behavioural problems in paediatric clinics," *Child: Care, Health and Development*, vol. 29, no. 2, pp. 141–149, 2003.

[26] H. L. Egger and A. Angold, "Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 47, no. 3-4, pp. 313–337, 2006.

[27] H. Elberling, A. Linneberg, E. M. Olsen, R. Goodman, and A. M. Skovgaard, "The prevalence of SDQ-measured mental health problems at age 5–7 years and identification of predictors from birth to preschool age in a Danish birth cohort: the Copenhagen Child Cohort 2000," *European Child and Adolescent Psychiatry*, vol. 19, no. 9, pp. 725–735, 2010.

[28] L. Wichstrøm, T. S. Berg-Nielsen, A. Angold, H. L. Egger, E. Solheim, and T. H. Sveen, "Prevalence of psychiatric disorders in preschoolers," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 53, no. 6, pp. 695–705, 2012.

[29] M. J. Zimmer-Gembeck, T. A. Hunter, A. M. Waters, and R. Pronk, "Depression as a longitudinal outcome and antecedent of preadolescents' peer relationships and peer-relevant cognition," *Development and Psychopathology*, vol. 21, no. 2, pp. 555–577, 2009.

[30] A. P. Ringoot, P. W. Jansen, J. Steenweg-de Graaff et al., "Young children's self-reported emotional, behavioral, and peer problems: the Berkeley Puppet Interview," *Psychological Assessment*, vol. 25, no. 4, pp. 1273–1285, 2013.

[31] A. Sourander, A. Haavisto, J. A. Ronning et al., "Recognition of psychiatric disorders, and self-perceived problems. A follow-up study from age 8 to age 18," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 46, no. 10, pp. 1124–1134, 2005.

[32] R. Goodman, T. Ford, H. Simmons, R. Gatward, and H. Meltzer, "Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample," *The British Journal of Psychiatry*, vol. 177, pp. 534–539, 2000.

[33] J.-P. Valla, L. Bergeron, H. Berube, N. Gaudet, and M. St-Georges, "A structured pictorial questionnaire to assess DSM-III-R-based diagnoses in children (6–11 years): development, validity, and reliability," *Journal of Abnormal Child Psychology*, vol. 22, no. 4, pp. 403–423, 1994.

[34] E. Klimkeit, C. Graham, P. Lee, M. Morling, D. Russo, and B. Tonge, "Children should be seen and heard: self-report of feelings and behaviors in primary-school-age children with ADHD," *Journal of Attention Disorders*, vol. 10, no. 2, pp. 181–191, 2006.

[35] D. R. Martini, J. M. Strayhorn, and J. Puig-Antich, "A symptom self-report measure for preschool children," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 29, no. 4, pp. 594–600, 1990.