

## Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, lærerrapport (SDQ-T)

Hege Kornør<sup>1</sup> (hege.kornor@r-bup.no)

Sonja Heyerdahl<sup>1</sup> (sonja.heyerdahl@r-bup.no)

PsykTestBarn 2014, 2:5

**ISSN** 1893-9910

**Mottatt dato** 27.06.2014

**Publisert dato** 15.10.2014

**Artikkelens URL** <https://psyktestbarn.r-bup.no/no/artikler/sdq-t-strengths-and-difficulties-questionnaire-laererrapport>

**DOI** 10.21337/0036

Som alle artikler i PsykTestBarn, kan denne fagfellevurderte artikkelen lastes ned, skrives ut og distribueres fritt for alle slags formål på følgende vilkår: korrekt referanse skal oppgis (se under), ingen kommersiell bruk, og ingen bearbeidelse av tekst eller innhold.

---

<sup>1</sup> RBUP Øst og Sør

**Denne artikkelen skal siteres på følgende måte:**

Kornør, H. & Heyerdahl, S. (2014). Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, lærerrapport (SDQ-T). *PsykTestBarn*, 2:5. doi: 10.21337/0036

**Sammendrag**

**Beskrivelse.** Strengths and Difficulties Questionnaire (SDQ) er et sett spørreskjema for utfylling av foreldre, lærere og barn. SDQ ble opprinnelig utgitt på engelsk av Robert Goodman i 1997. I 1999 kom de norske versjonene, basert på en oversettelse og tilbakeoversettelse ved Einar Heiervang og kolleger. Robert Goodman eier rettighetene til testen. Lærerversjonen (SDQ-T) har 25 ledd fordelt på fem delskalaer. Det er et skjema som skal fylles ut av lærere for å kartlegge psykiske helse, vennerelasjoner og prososial atferd hos barn i alderen 4-17 år. Det tar ~~under 30~~**noen få** minutter å fylle ut. Forskere, klinikere og pedagoger kan administrere og tolke SDQ-T.

**Litteratursøk.** Våre systematiske søk etter dokumentasjon for måleegenskapene ved den norske versjonen av SDQ-T førte til 252 treff, hvorav 22 publikasjoner fra ni studier ble inkludert. I alt 18 av publikasjonene rapporterte fra store befolkningsstudier i Bergen, Trondheim, Bodø, Østfold og Oslo-området.

**Psykometri.** De store befolkningsstudiene bidro med regionale normdata for aldersgruppene 4 år og 7-9 år. I to av studiene undersøkte forskerne faktorstrukturen til SDQ-T, og tre studier rapporterte verdier for diagnostisk nøyaktighet. Gruppesammenlikninger mellom subgrupper og kontrollgrupper belyste også validiteten ved instrumentet, spesielt resultater for ulike diagnostiske grupper. Vi fant mål på indre konsistens (Cronbachs alfa) fra sju studier.

**Konklusjon.** SDQ-T har regionale normdata, men bare for barn under 9 år, og nasjonale data mangler. Forventede gruppeforskjeller og konfirmatoriske faktoranalyser støtter opp under instrumentets begrepsvaliditet. Skalaene har gjennomgående tilfredsstillende til god indre konsistens. Evnen til å fange opp barn med en psykiatrisk diagnose synes å være ganske god. Likevel er ikke den diagnostiske nøyaktighet så god at SDQ-T kan anbefales som universelt screening-instrument. Den norske versjonen av SDQ-T kan være egnet til å screene for psykiatriske diagnoser i selekterte utvalg med risiko for psykiske helseproblemer.

**Abstract**

**Description.** The Strengths and Difficulties Questionnaire (SDQ) is a set of parent-, teacher- and youth-reported questionnaires originally published in English by Robert Goodman in 1997. The Norwegian versions were published in 1999, based on a translation and back-translation by Einar Heiervang and colleagues. Robert Goodman holds the copyright. The teacher version (SDQ-T) contains 25 items organized in five subscales. It is a teacher report form to assess mental health, peer relations and prosocial behavior in children aged 4-17 years. Completion takes a few minutes. Clinicians and teachers can administrate and interpret the SDQ-T.

**Literature search.** Our systematic searches for psychometric evidence for the Norwegian SDQ-T version identified 252 references, of which 22 publications from nine studies were included. Eighteen publications were reports from large population-based studies in Bergen, Trondheim, Bodø, Østfold and the Oslo area.

**Psychometrics.** The large population-based studies contributed with regional norm data for 4 year-olds and 7-9 year-olds. Researchers assessed the SDQ-T factor structure in two studies, and

three studies reported diagnostic accuracy data. Group comparisons between subgroups and controls also addressed the instrument's validity, especially data for various diagnostic groups. We identified measures of internal consistency (Cronbach's alpha) from seven studies.

**Conclusion.** The SDQ-T has regional norms, but only for children below the age of nine, and there is a lack of national data. Expected group differences and confirmatory factor analyses support the construct validity of the instrument. The scales have adequate to good internal consistencies. The ability to detect children with a psychiatric diagnosis seems to be quite good. Still, the diagnostic accuracy is not sufficiently good to recommend the SDQ-T as a universal screening instrument. The Norwegian version of the SDQ-T may be appropriate for screening for psychiatric diagnoses in selected populations with high risk of mental health problems.

## Innledning

Strengths and Difficulties Questionnaire (SDQ) (<http://www.sdqinfo.org/>) er utviklet av Robert Goodman (Goodman, 1997). SDQ ble oversatt til norsk av Einar Heiervang, Irene Elgen og Sonja Heyerdahl, og tilbakeoversatt av Melanie Young i 1999. På norsk heter instrumentet Sterke og svake sider, men vi vil i denne sammenhengen referere til det som SDQ.

SDQ er et kortfattet spørreskjema, laget for å kartlegge psykisk helse, vennerelasjoner og prososial atferd hos barn og unge. Det tar ~~under~~ ~~30~~**noen få** minutter å fylle ut. Hovedskjemaet har 25 ledd (5 delskalaer med 5 ledd i hver). Det er fire problemskalaer: Atferdsproblemer, Hyperaktivitet – oppmerksomhetsproblemer, Emosjonelle symptomer og Venneproblemer. I tillegg er det en skala som kartlegger prososial atferd. Man angir om utsagnene i skjema stemmer for barnet (0-stemmer ikke, 1-stemmer delvis eller 2-stemmer helt). Skårene summeres for hver delskala, og de fire problemskalaene summeres til en skåre for totale vansker. Man kan også summere skala for emosjonelle symptomer og venneproblemer til en internaliserende skåre og skala for atferdsproblemer og hyperaktivitet-oppmerksomhetsproblemer til en eksternaliserende skåre (Goodman, Lamping & Ploubidis, 2010). Dette er særlig relevant i utvalg med relativt lite problemer. I et tilleggsskjema («impact supplement») spør man om barnet har vansker med følelser, konsentrasjon, oppførsel eller med å komme overens med andre mennesker. I så fall spørres om varighet av problemene og

hvordan vanskene virker inn på barnets liv (fem spørsmål). Tre av spørsmålene som angir vanskenes innvirkning på barnets/ungdommens liv (0-ikke i det hele tatt, 0-bare litt, 1-en god del, 2-mye), kan summeres til en skåre for vanskenes innvirkning («impact»).

Lærerversjonen (SDQ-T) var beregnet for barn/ungdom i alderen 4-16 år (Goodman, 1997). Aldersspennet er senere endret til 4-17 år (<http://www.sdqinfo.org/>). SDQ-T inngår i kartleggingspakken Development and Well-Being Assessment (DAWBA), som finnes både i papir- og elektronisk versjon (Goodman et al., 2000). Spørreskjema foreligger også i foreldreversjon for den samme aldersgruppen og for barn 3-4 år (SDQ-P), samt i selvrappportversjon (SDQ-S) for barn/ungdom 11-16 år (Goodman, 1999). Lærerversjon for barn 3-4 år foreligger foreløpig ikke på norsk.

Det foreligger også en egen versjon for oppfølging etter behandling. Denne har med to spørsmål om behandlingen (Ford, Hutchings, Bywater, Goodman & Goodman, 2009).

## Oversettelser

SDQ er brukt i svært mange studier, både internasjonalt og i Norge. Pr 2014 er SDQ oversatt til mer enn 72 språkvarianter. Spørreskjema på norsk kan lastes ned fra <http://www.sdqinfo.org/>. SDQ foreligger også som nettversjon [www.youthinmind.info](http://www.youthinmind.info), men foreløpig ikke på

norsk. Papirversjonen kan brukes fritt for ikke-kommersielle formål, men man må ha tillatelse fra Robert Goodman for å benytte skjemaet elektronisk og betale en mindre avgift for dette.

### Skåring

På <http://www.sdqinfo.org/> er det informasjon om både håndskåring og elektronisk skåring (skåringsyntax og internettbasert skåring). Det er angitt grenseverdier for antatt avvikende skårer, basert på britisk normalmateriale. Skårer over ca 90-percentilen angitt som «antatt klinisk område» (på engelsk: «abnormal»), og skårer mellom ca 80- og 90-percentilen er kategorisert som «grenseområde» («borderline range»). For Totale vansker på SDQ-T er de britiske grenseverdiene 16 og 12 for henholdsvis antatt klinisk område og grenseområdet. Robert Goodman har utviklet en PC-basert algoritme for å sammenfatte resultater fra ulike informanter, med prediksjon av om barnepsykiatrisk diagnose er sannsynlig på grunnlag av resultatene (Goodman, Ford, Simmons, Gatward & Meltzer, 2000).

### Måleegenskaper internasjonalt

Mange internasjonale studier har studert reliabilitet og validitet ved SDQ-T. Goodman (2001) fant at femfaktorstrukturen ble bekreftet i data fra et britisk epidemiologisk utvalg. Han vurderte den interne konsistens som tilfredsstillende; Cronbachs alfa var 0,87 for Totale vansker og  $\geq 70$  for subskalaer og Impact-skalaen. SDQ-T-skårer over 90-percentilen predikerte klart økt sannsynlighet for uavhengig diagnostisert psykiatrisk lidelse (gjennomsnittlig odds ratio 15,2). Spesifisitet og negativ prediktiv verdi for prediksjon av psykiatrisk lidelse var høye, mens sensitivitet og positiv prediktiv verdi var lavere. Generelt hadde SDQ-T psykometriske egenskaper på nivå med eller bedre enn SDQ-P (foreldrerapport) og bedre enn SDQ-S (selvrapport). I en valideringsstudie hvor også skåren for vanskenes betydning («impact»-skåren) var inkludert, fant Goodman (1999) at skåren for vanskenes betydning diskriminerte bedre enn

symptom-skårer mellom kliniske utvalg og befolkningsutvalg.

Achenbach og kolleger (2008) har oppsummert funn på psykometriske egenskaper for både SDQ og tilsvarende skjemaer i Achenbach System of Empirically Based Assessment (ASEBA) i en oversiktsartikkel.

### Metode

Vi søkte etter dokumentasjon på testens psykometriske egenskaper i databasene Norart, BIBSYS, SveMed+, CRISStin.no, Cochrane Library, PsycINFO, Medline, Embase og PubMed. Det første søket ble gjort 25.11.2010, og for å få et så oppdatert bilde som mulig av litteraturen søkte vi igjen 23.04.2013. Søkestrategiene er tilgjengelig på <http://www.r-bup.no/CMS/ptb.nsf/pages/sdq>. Vi kontaktet også førsteforfattere av inkluderte artikler, samt oversetterne av SDQ, for å identifisere dokumentasjon som eventuelt ikke ble fanget opp av det systematiske søket.

Vi inkluderte alle publikasjoner av studier som har undersøkt og rapportert minst ett av følgende i norske utvalg:

- normdata for testen
- reliabilitet: indre konsistens, test-retest eller endringssensitivitet
- validitet: samsvar med liknende testskårer, samsvar med referansestandard eller annet kriterium, og/eller faktorstruktur

I tillegg inkluderte vi publikasjoner som rapporterte gjennomsnittsskårer for henholdsvis generelle populasjoner og for kliniske og andre undergrupper.

To forskere gikk gjennom sammendragene til alle identifiserte publikasjoner, uavhengig av hverandre, etter at dubletter var fjernet. Alle publikasjoner som kunne virke relevante ble bestilt inn i fulltekst, og prosessen over ble gjentatt for fulltekstrapportene.

To forskere vurderte normering, validitet og reliabilitet, uavhengig av hverandre, ved hjelp av

en tilpasset versjon av Test review form and notes for reviewers (European Federation of Psychologist Association (EFPA), 2008).

## Resultater

### Litteratursøk

De to systematiske litteratursøkene etter dokumentasjon for den norske versjonen av SDQ-T identifiserte til sammen 252 unike publikasjoner. Av disse virket 90 relevante og ble gjennomgått i fulltekst. Totalt 22 publikasjoner oppfylte inklusjonskriteriene våre, og ble inkludert. Vi ekskluderte de resterende 68 publikasjonene fordi de ikke rapporterte relevante data, rapporterte data fra ikke-norske utvalg, kun rapporterte data fra andre SDQ-versjoner enn lærerversjonen, eller ikke hadde brukt SDQ i det hele tatt (Vedlegg 1).

Vi kontaktet også forskere knyttet til de største inkluderte studiene for å identifisere ytterligere relevant dokumentasjon. På den måten kom vi over én relevant publikasjon, som vi inkluderte (Sveen, Berg-Nielsen, Lydersen & Wichstrom, 2013). Vi søkte ikke etter dokumentasjon for

svenske og danske versjoner av SDQ.

### Inkluderte studier

De totalt 22 inkluderte publikasjonene rapporterte data fra til sammen ni studier (Tabell 1). Der hvor det fantes flere publikasjoner med data fra én og samme studie, har vi valgt å gjengi data fra de publikasjonene som hadde størst utvalg, og flest skalaer og psykometriske verdier. Fra kohortstudier med to eller flere datainnsamlinger angir vi resultater fra de ulike bølgeene, med anmerkning om at dette er en longitudinell studie med samme populasjon.

Halvparten av de inkluderte artiklene rapporterte SDQ-T-data fra den longitudinelle befolkningsbaserte studien Barn i Bergen (Ekornås, Heimann, Tjus, Heyerdahl & Lundervold, 2011; Heiervang, Goodman & Goodman, 2008; Heiervang et al., 2007; Lundervold, Heimann & Manger, 2008; Munkvold, Lundervold, Lie & Manger, 2009; Munkvold, Lundervold & Manger, 2011; Posserud et al., 2008; Sanne, Torsheim, Heiervang & Stormark, 2009; Stormark, Heiervang, Heimann, Lundervold & Gillberg, 2008; Sørensen, Hugdahl &

**Tabell 1.** Inkluderte studier

Førsteforfattere år	Studiedesign	Populasjon	N <sup>1</sup>	Svarprosent
Ekornås 2011, Heiervang 2007, Heiervang 2008, Lundervold 2008, Munkvold 2009, Munkvold 2011, Posserud 2008, Sanne 2009, Stormark 2008, Sørensen 2008, Sørensen 2012, Ullebø 2011	Kohort (Barn i Bergen)	7-9-åringer 11-13 <sup>2</sup> -åringer	9155 4607	97 % 66 %
Holen 2012a, Holen 2012b	RCT (Zippys venner)	7-åringer	1308	88 %
Sveen 2013	Tverrsnitt	4-åringer	845	68 %
Eivers 2010, Eivers 2012	Kohort	3-7-åringer	203 <sup>3</sup>	43 %
Indredavik 2005	Kasus-kontroll	14-åringer med svært lav fødselsvekt, liten for gestsjonalder, kontroller	44+52+72	52 %
Javo 2009	Kohort	11-12-åringer	173	58 %
Nordhov 2012	RCT	5-åringer med lav fødselsvekt, kontroller	176	72 %
Eilertsen 2011	Kasus-kontroll	6-20-åringer, kreftbehandlede, kontroller	36 + 19	35 %
Hanssen-Bauer 2007	Tverrsnitt	9-15-åringer med epilepsi	38	35 %

Note: <sup>1</sup>Antall med SDQ-T-data; <sup>2</sup>Den samme kohorten ved oppfølging; <sup>3</sup>Data fra Eivers 2012.

Lundervold, 2008; Sørensen, Plessen & Lundervold, 2012; Ullebø, Posserud, Heiervang, Gillberg & Obel, 2011). Barn i Bergen har fire runder. Datainnsamlingen til den første runden hadde samtlige 9430 2.-4. klassinger (7-9-åringer) i Bergen som målgruppe, og fant sted skoleåret 2002/2003. Denne runden fikk forskerne inn SDQ-T-data for 9155 av 9430 (97 %) barn i målgruppen (Heiervang et al., 2007). Runde 2 fant sted fire år senere, da barna var 11-13 år, og omfattet 7007 barn med foreldresamtykke for nye undersøkelser i runde 1 (Ekornås et al., 2011). I runde 3 og 4 var barna henholdsvis 14-16 og 16-18 år.

De inkluderte artiklene rapporterte gjennomsnittsskårer for utvalgsstørrelser som varierte fra 6150 (svarprosent 65 %), (Munkvold et al., 2009) til 9072 (svarprosent 96 %), (Lundervold et al., 2008) (Heiervang et al., 2008; Lundervold et al., 2008; Munkvold et al., 2009; Munkvold et al., 2011). Kjønnsforskjeller og forskjeller mellom ulike subgrupper og kontroller var også rapportert, samt grenseverdier, faktorstruktur, diagnostisk nøyaktighet, forhold mellom SDQ-T-skårer og andre faktorer, indre konsistens og stabilitet. Dataene omfattet alle SDQ-T-skalaene, men noen av artiklene fokuserte på spesifikke skalaer.

En annen stor studie var et klyngerandomisert kontrollert forsøk som inkluderte andreklassinger (gjennomsnittsalder 7,3 år) ved 35 skoler i Trondheim, Bodø og Østfold (Holen, Lervåg, Waaktaar & Ystgaard, 2012; Holen, Waaktaar, Lervåg & Ystgaard, 2012). I alt 18 skoler (745 andreklassinger) ble trukket ut til å få det universal-forebyggende programmet Zippys venner, som skal forebygge psykiske problemer hos barn ved å gi dem økte mestringsferdigheter. De øvrige 17 skolene (738 andreklassinger) ble trukket ut til å utgjøre en kontrollgruppe som ikke fikk noe tiltak. Forskerne fikk inn SDQ-T-data for 1308 barn (88 %) i denne studien, og rapporterte data omfattet z-skårer, faktorstruktur, korrelasjoner og indre konsistens.

Sveen og kolleger rapporterte data fra den longitudinelle befolkningsbaserte studien Tidlig trygg i Trondheim, med Trondheims rundt 3000 4-åringer som målgruppe (Sveen et al., 2013).

Foreldre og barn ble rekruttert gjennom helsestasjonen, og 2475 familier samtykket til å delta. Til sammen 1250 av disse ble tilfeldig valgt ut fra fire strata av SDQ-skårer for å delta i det diagnostiske intervjuet Preschool Age Psychiatric Assessment (PAPA) (Egger et al., 2006) med en foresatt. Den inkluderte artikkelen rapporterte tverrsnittsdata for 845 (svarprosent 68 %) barn med komplette datasett (foreldreintervju, SDQ-P og SDQ-T). Artikkelen rapporterte gjennomsnittsskåre og *Area Under the Curve* (AUC) for samsvar mellom SDQ-T Totale vansker og PAPA.

Eivers og kolleger (Eivers, Brendgen & Borge, 2010; Eivers, Brendgen, Vitaro & Borge, 2012) brukte SDQ-subskalaene Prososial atferd og Atferdsproblemer i tre undergrupper fra en kohortstudie. Hele utvalget besto av 619 barnehagebarn utenfor Oslo (62 % deltakelse). De tre undergruppene med SDQ-T-data var 203 3-5-åringer, 130 4-6-åringer og 118 5-6-åringer. Den siste gruppen var barn som begynte på skolen i året etter første datainnsamling, og ble sammenliknet med den nest siste gruppen, som fortsatt gikk i barnehagen. Artiklene rapporterte gjennomsnittsskårer, aldersforskjeller, endring, korrelasjoner, regresjonsanalyser og indre konsistens. De har brukt en skåringsprosedyre som er forskjellig fra den som angis på <http://www.sdqinfo.org/>; resultatene kan derfor ikke sammenlignes med andre studier og presenteres ikke som normdata.

De øvrige fem inkluderte studiene hadde mindre utvalg fra kliniske eller etniske populasjoner (Eilertsen, Rannestad, Indredavik & Vik, 2011; Hanssen-Bauer, Heyerdahl & Eriksson, 2007; Indredavik, Vik, Heyerdahl, Kulseng & Brubakk, 2005; Javo, Rønning, Handegard & Rudmin, 2009; Nordhov, Rønning, Ulvund, Dahl & Kaaresen, 2012). Fire av disse studiene sammenliknet utvalgene med kontrollgrupper. Den femte og minste studien (Hanssen-Bauer et al., 2007) manglet kontrollgruppe, og data fra denne studien blir derfor ikke rapportert her.

## Normer

Ingen av de inkluderte studiene rapporterte nasjonale normer, men vi fant gjennomsnittsskårer for et stort utvalg fra Bergen, fra et utvalg fra Bodø, Trondheim og Østfold, og fra et utvalg 4-åringer fra Trondheim (Tabell 2).

Blant 7-9-åringer i Bergen var det en statistisk signifikant kjønnsforskjell for samtlige skalaer, bortsett fra Emosjonelle symptomer (Lundervold et al., 2008). Jentene fikk gjennomgående lavere skårer enn guttene på problemskalaene, og de hadde høyere skårer enn guttene på Prosocial atferd. Dataene er hentet fra en artikkel som undersøker forskjeller mellom barn med språkvansker og den generelle populasjonen. Den generelle populasjonen er derfor «supernormal». Heiervang og kolleger (Heiervang et al., 2007) rapporterer gjennomsnitt for SDQ-T Totale vansker for 6297 barn hvor både lærer og foreldreskjema var brukt: 4,1 (SD 4,8) og for 2858 barn hvor bare lærerskjema var fylt ut: 5,4 (SD 5,4). På grunnlag av dette har vi beregnet at gjennomsnitt for Totale vansker for 9155 barn fra Barn i Bergen-studien var 4,5.

Oppfølgingsdata som rapporteres i Tabell 2 fra Barn i Bergen, da deltakerne var blitt 11-13 år, kom

også fra et «supernormalt» utvalg: kontrollgruppen med 'ikke sannsynlig psykiatrisk diagnose' (Ekornås et al., 2011). Underskalaen Problemer med venner var den eneste skalaen som ble brukt i denne aldersgruppen.

Heiervang og kolleger (2008) sammenliknet SDQ-T-skårer hos britiske barn ( $N = 3836$ ; 8-10 år) og Barn i Bergen-deltakere ( $N = 6184$ ; 7-9 år). De britiske barna skåret høyere enn de norske på samtlige skalaer, bortsett fra Prosocial atferd, hvor norske barn skåret høyere. Den gjennomsnittlige forskjellen var for Totale vansker 2,5 (95 % konfidensintervall 2,2 – 2,7); Gjennomsnittsskåre for de norske barna var 4,1 ( $SD = 4,8$ ); for de britiske barna 6,6 ( $SD = 6,1$ ). Alle forskjellene var statistisk signifikante ( $p < 0,001$  – justert for «clustered sample design»). Studien brukte både spørreskjema (SDQ-P og SDQ-T) og intervju (DAWBA) og undersøkte om forskjeller i SDQ-skårer mellom landene reflekterte reelle forskjeller i mental helse eller forskjeller i rapporteringsstil. For barn med diagnostiserte atferdsproblemer, ADHD og autisme-spekter forstyrrelser, var det ikke forskjeller mellom landene i SDQ-T Totale vansker og heller ikke SDQ-P Totale vansker. Dette indikerer at forskjellene i SDQ-skårene for disse

**Tabell 2.** SDQ-T-skårer i generelle populasjoner

Referanse	Lundervold 2008			Ekornås 2011	Holen 2013	Sveen 2013
Populasjon	Barn i Bergen, 7-9 år <sup>1</sup>			Barn i Bergen, 11-13 <sup>2</sup>	Barn i Bodø, Trondheim og Østfold, 7 år <sup>3</sup>	4-åringer i Trondheim
	Total	Jenter	Gutter			
<b>N</b>	7469	3906	3563	4344	1310	845
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
<b>Totale vansker</b>	4,0 (4,4)	4,0 (3,6)	5,0 (4,9)*			5,1 (4,8-5,4) <sup>4</sup>
<b>Emosjonelle symptomer</b>	0,6 (1,3)	0,6 (1,3)	0,6 (1,2)		0,9 (1,6)	
<b>Atferdsproblemer</b>	0,6 (1,3)	0,4 (0,9)	0,8 (1,4)*		0,7 (1,3)	
<b>Hyperaktivitet</b>	2,1 (2,3)	1,5 (1,8)	2,8 (2,6)*		2,5 (2,7)	
<b>Problemer med venner</b>	0,7 (1,3)	0,6 (1,2)	0,8 (1,4)*	0,6 (1,2)	0,9 (1,5)	
<b>Prosocial atferd</b>	8,4 (2,0)	8,9 (1,6)	7,9 (2,2)*		7,9 (2,2)	
<b>Impact</b>	0,3 (1,2)	0,2 (0,8)	0,5 (1,5)*		0,4 (1,1)	

Note: \* $p < 0,001$ ; <sup>1</sup> Gjelder barn uten språkvansker; <sup>2</sup>Oppfølgingsdata, barn uten psykiatrisk diagnose; <sup>3</sup>Pretestskårer, intervensjonsgruppe, <sup>4</sup>95% konfidensintervall.

**Tabell 3.** Korrelasjoner mellom SDQ-T subskalaer og tre skalaer i Autism Screening Questionnaire (Posserud et al., 2008;  $N = 6229$ )

	Emosjonelle symptomer	Atferdsproblemer	Hyperaktivitet	Problemer med venner	Prososial atferd
<i>Social difficulties</i>	0,46	0,63	0,54	0,71	-0,51
<i>Motor/tics/OCD</i>	0,33	0,29	0,33	0,39	-0,24
<i>Autistic style</i>	0,21	0,16	0,11	0,30	-0,13

Note: Samtlige korrelasjoner var statistisk signifikante ( $p < 0,001$ )

SDQ-T Impact-skalaen korrelerte med Preschool Age Psychiatric Assessment (PAPA) Total Incapacity-skalaen ( $r = 0,26$ ;  $p < 0,001$ ) blant de 845 4-åringene i Tidlig trygg i Trondheim (Sveen et al., 2013).

gruppene reflekterer reelle forskjeller i problemer. For emosjonelle lidelser, derimot, var SDQ-T Totale vansker, Impact og belastning signifikant lavere for norske barn enn for britiske barn, hvilket kan indikere at det er forskjeller i rapporteringsstil eller terskel for rapportering av emosjonelle problemer mellom lærere i de to landene. Det var tilsvarende funn for emosjonelle lidelser for foreldrenes SDQ skårer. Det var ikke forskjell i prevalens av emosjonelle forstyrrelser mellom de to landene.

Skårene til 7-åringer i Bodø, Trondheim og Østfold (Holen, Waaktaar, et al., 2012) lå jevnt over høyere enn for 7-9-åringene i Bergen, bortsett fra Sosial atferd, som lå høyere i Bergen. Forskjellene var riktignok små, og er ikke statistisk testet.

## Validitet

### Begrepsvaliditet

Ved hjelp av faktoranalyser, korrelasjoner og gruppesammenlikninger fant vi data som kunne brukes til vurdering av begrepsvaliditeten ved SDQ-T-skalaene.

### Faktorstruktur

I studiene Barn i Bergen ( $N = 4483$ ) og Zippys venner ( $N = 1308$ ) undersøkte forskerne SDQ-Ts faktorstruktur ved hjelp av konfirmatoriske

faktoranalyser (Holen, Lervåg, et al., 2012; Holen, Waaktaar, et al., 2012; Sanne et al., 2009). RMSEA (root mean square error of approximation)-indeksen varierte fra 0,05 til 0,09 (median = 0,06). Tilsvarende CFI (comparative fit index) varierte fra 0,86 til 0,97 (median = 0,93). I tillegg brukte Sanne og kolleger (2009) eksplorerende faktoranalyse, med RMSEA = 0,036 ( $N = 4516$ ) for femfaktormodellen (Sanne et al., 2009).

### Korrelasjoner mellom SDQ-T- og tilsvarende skalaer fra andre instrumenter

Posserud og kolleger (2008) undersøkte forholdet mellom SDQ-T-subskalaene og de tre skalaene i Autism Spectrum Screening Questionnaire (Ehlers, Gillberg & Wing, 1999). Utvalget besto av 6229 deltakere i Barn i Bergen med komplett utfylte ASSQ, både lærer- og foreldreversjon. Samtlige korrelasjoner var statistisk signifikante ( $p < 0,001$ ), og korrelasjonene var høyest mellom SDQ-T-subskalaene, spesielt Problemer med venner, og ASSQ-skalaen Sosiale vansker ( $r: -0,51 - 0,71$ ) (Tabell 3).

Eivers og kolleger (2010) undersøkte hvordan SDQ-T-skårer på henholdsvis Atferdsproblemer og Sosial atferd, forholdt seg til vurderinger som jevnaldrende førskolebarn hadde gjort av hverandre med hensyn til de samme atferdene (vennevurdert atferd). De 248 4-6-åringene ble



bedt om å si navnene til inntil seks barn som ofte viste en eller flere av tre forskjellige prososiale og tre forskjellige antisosiale atferder. De barna som ble nevnt, og som deltok i studien fikk en skåre for hver av atferdene. På grunnlag av dette beregnet forskerne en sammenlagt skåre for henholdsvis antisosial og prososial atferd. SDQ-T-skalaen Atferdsproblemer korrelerte positivt med de vennevurderte antisosiale atferdene ( $r = 0,29 - 0,43$ ;  $p < 0,01$ ), og negativt med vennevurdert hjelpeatferd ( $r = -0,17$ ;  $p < 0,05$ ), samt sammenlagt vennevurdert prososial atferdsskåre ( $r = -0,14$ ;  $p < 0,05$ ). SDQ-T-skalaen Prososial atferd korrelerte ikke med de vennevurderte prososiale atferdene, men korrelerte negativt med de antisosiale atferdene ( $r = -0,28 - -0,20$ ;  $p < 0,01$ ).

generelle populasjoner og undergrupper (Tabell 4). Se Vedlegg 2 for gruppeforskjeller i subskalaskårer. De fleste subgruppene med somatiske eller psykiatriske sykdommer og barna med språkvansker hadde mer totale vansker enn kontrollgruppene. Lundervold og kolleger (2008), og Indredavik og kolleger (2012) hadde testet forskjellene statistisk.

I tillegg sammenliknet Stormark og kolleger (2008) deltakere ( $N = 6611$ ) og ikke-deltakere ( $N = 2544$ ) i Barn i Bergen med hensyn til Totale vansker, Emosjonelle symptomer, Atferdsproblemer, Hyperaktivitet og Problemer med venner. Ikke-deltakerne hadde gjennomgående høyere skårer enn deltakerne ( $p < 0,001$ ).

### Gruppeforskjeller

Mange av de inkluderte publikasjonene rapporterte om forskjeller i Totale vansker mellom

### Kriterievaliditet

Tre studier undersøkte SDQ-T-skårenes diagnostiske nøyaktighet ved å sammenholde dem med ulike referansestandarder (Indredavik et al.,

**Tabell 4.** Totale vansker hos subgrupper sammenliknet med kontrollgrupper

Referanse	Alder (år)	Subgruppe			Kontrollgruppe	
		Kjennetegn	N	M (SD)	N	M (SD)
Lundervold 2008 (Barn i Bergen)	7-9	Språkvansker	540	12,0 (7,3)	8532	4,0 (4,4)*
		Jenter	174	9,0 (6,2)		3,1 (3,6)
		Gutter	366	13,5 (7,4)		5,0 (4,9)
Heiervang 2008 <sup>1,2</sup>	8-10	Emosjonell forstyrrelse		6,9 (4,8-9,0)		3,5 (3,2-3,7) <sup>3</sup>
		Atferdsforstyrrelse		16,8 (15,0-18,6)		
		ADHD		19,1 (17,3-20,9)		
		Autismespekterforstyrrelse		16,4 (13,2-19,7)		
Eilertsen 2011	6-20	Kreftbehandlet	36	7,1 (7,0)	19	5,4 (6,8) <sup>3</sup>
Indredavik 2005	14	Svært lav fødselsvekt	44	7,1 (4,9)	72	4,1 (4,2)*
		Liten for svangerskapsalder	52	6,0 (6,9)		
Nordhov 2012	5	For tidlig fødte med tiltak	60	7,0 (5,7)	60	4,4 (3,6) <sup>3</sup>
		For tidlig fødte uten tiltak	56	7,1 (6,5)		

Note: \* $p < 0,001$  for forskjell subgruppe/kontrollgruppe; <sup>1</sup>N ikke angitt for subgruppene, Total N i intervjuundersøkelsen som er grunnlag for diagnosene er 1024; <sup>2</sup>95 % konfidensintervall i parentesene. <sup>3</sup>p-verdi ikke angitt

**Tabell 5.** SDQ-T/Totale vansker: evne til å skille mellom barn med og uten psykiatrisk diagnose (Area Under the Curve (95 % konfidensintervall))

	Sveen 2013		Indredavik 2005		Ullebø 2011
Referansestandard	PAPA		Kiddie-SADS-PL		SNAP IV
	N = 845	Svært lav fødselsvekt (N = 44)	Liten for svangerskapsalder (N = 51)	Kontroll (N = 72)	N = 6233
Minst en psykiatrisk diagnose	0,65 (0,58-0,71)	0,80 (0,65-0,96)	0,82 (0,62-1,00)	0,67 (0,48-0,86)	
Emosjonell forstyrrelse	0,70 (0,61-0,79)				
Atferdsforstyrrelse	0,76 (0,68-0,84)				
Emosjonell og atferdsforstyrrelse	0,70 (0,62-0,77)				
ADHD					0,95 (0,94-0,95)

2005; Sveen et al., 2013; Ullebø et al., 2011). Data var rapportert som sensitivitet, spesifisitet og *Area Under the Curve* (AUC) i Receiver Operator Curve- (ROC-) analyser (Tabell 5).

Fra studien av 4-åringer, Tidlig trygg i Trondheim, rapporterte Sveen og kolleger (2013) AUC for henholdsvis Totale vansker med terskelverdi 8-14 og Impact, og fire forskjellige diagnostiske kategorier: minst én psykiatrisk diagnose (N = 75), emosjonell forstyrrelse (N = 39), atferdsforstyrrelse (N = 40), emosjonell og atferdsforstyrrelse (N = 63). Referansestandard var det diagnostiske intervjuet Preschool Age Psychiatric Assessment (PAPA) (Egger et al., 2006), med intervjuere blindet for SDQ-skårene. AUC for terskelverdi 8-14 varierte fra 0,65 til 0,76 (Tabell 5). For Impact varierte AUC fra 0,60 til 0,67.

Barn i Bergen (N = 6233) brukte Swanson, Nolan og Pelhams spørreskjema for lærere og foreldre - versjon IV (SNAP-IV) (Swanson et al., 2001) som referansestandard for å undersøke den diagnostiske nøyaktigheten til SDQ-T Hyperaktivitet-skårer med hensyn til ADHD (Ullebø et al., 2011). 90-percentilen for SDQ-T

Hyperaktivitet var 5. Ved denne grenseverdien var sensitiviteten 84 %, spesifisiteten 88 % og den positive prediktive verdien 0,28 (prevalens: 5 %). AUC var 0,95 (95 % konfidensintervall: 0,94 – 0,95).

Med Schedule for Affective Disorders and Schizophrenia for School Aged Children (K-SADS) (Ambrosini, 2000) som referansestandard og 90-percentilen som terskelverdi for minst én psykiatrisk diagnose, hadde SDQ-T en sensitivitet på 17 % - 60 %, og en spesifisitet på 83 % - 89 % i en relativt liten studie av barn med lav fødselsvekt, barn som var små for gestasjonsalder og kontrollgruppe (Indredavik et al., 2005). AUC var 0,67-0,82.

## Reliabilitet

Sju av artiklene rapporterte indre konsistens (Cronbachs  $\alpha$ ) for SDQ-T-skalaene (Tabell 6) (Eivers et al., 2010; Eivers et al., 2012; Holen, Lervåg, et al., 2012; Posserud et al., 2008; Sanne et al., 2009; Sveen et al., 2013; Sørensen et al., 2012).

## Diskusjon og konklusjon

Litteratursøket identifiserte 22 publikasjoner fordelt på ni studier som hadde brukt SDQ-T. Det har vært omfattende norsk forskning på SDQ-T, med fokus på psykometriske egenskaper og data fra befolkningsutvalg.

Det foreligger resultater fra tre relativt store befolkningsbaserte undersøkelser med høye svarprosent, men de har alle bare med barn under 9 år og lokalt avgrensede populasjoner. Dataene kan derfor ikke brukes som norske nasjonale normer. Kjønnsforskjellene i skårer støtter annen empiri om mer problematferd hos gutter enn hos jenter, og mer prososial atferd hos jenter enn hos gutter. En studie sammenlignet resultater fra Bergen og Storbritannia for 7-9 år gamle barn (Heiervang et al., 2008). Man fant lavere SDQ-T og SDQ-P-skårer for barna i Bergen, som samsvarte med lavere prevalenstall for de norske barna for ikke-emosjonelle forstyrrelser. For emosjonelle forstyrrelser derimot, hvor også de norske foreldrene og lærerne skåret barna lavere på SDQ enn de britiske, var det ikke forskjell i prevalens mellom landene. Dette kan indikere at

norske lærere og foreldre underrapporterer emosjonelle symptomer. Det viser også at krysskulturelle forskjeller i spørreskjema-resultater ikke nødvendigvis kan tolkes som reelle forskjeller i mental helse.

Når SDQ skåres ved håndskåring eller på nettet, brukes britiske normer, og britisk 80 og 90 percentil benyttes som grenseverdier. Forskjellene i funn mellom Norge og Storbritannia innebærer at de britiske normene sannsynligvis ikke er korrekte for norske forhold. Resultater som skåres med dette normgrunnlaget, må tolkes med forsiktighet.

Konfirmatoriske faktoranalyser og forventete gruppeforskjeller støttet opp under instrumentets begrepsvaliditet. Barn i Bergen-forskerne fant at SDQ-T Problemer med venner korrelerte som forventet med ASSQ Social Difficulties ( $r = 0,71$ ), mens SDQ-T Prososial atferd var svakere korrelert ( $r = -0,51$ ).

Når det gjelder kriterievaliditet varierte AUC-verdiene noe med type utvalg og type diagnostikk. Blant 4-åringene kunne Totale vansker med akseptabel nøyaktighet skille mellom barn i og utenfor diagnostiske grupper, men ikke hvem som hadde minst én psykiatrisk diagnose. Dette viser at jo mer avgrenset populasjonen og/eller diagnosene er, desto større diagnostisk treffsikkerhet. Dette er i tråd med funn fra Barn i Bergen, hvor SDQ-T Hyperaktivitet-skalaen samsvarte godt med ADHD-diagnose ved SNAP-IV

**Tabell 6.** Indre konsistens (Cronbachs  $\alpha$ )

Skala	Antall studier	Median $\alpha$	Spredning
Totale vansker	2	0,87	0,86 – 0,87
Emosjonelle symptomer	3	0,73	0,73 - 0,76
Atferdsproblemer	4	0,72	0,64 – 0,75
Hyperaktivitet	2	0,84	0,80 – 0,88
Problemer med venner	2	0,72	0,71 – 0,72
Prososial atferd	4	0,84	0,82 – 0,85

(AUC 0,95).

For undergrupper av 14-åringer med lav fødselsvekt eller som var små for svangerskapsalder hadde Totale vansker utmerket evne til å skille mellom individer med og uten minst én psykiatrisk diagnose, mens denne evnen var mindre pålitelig for en jevn gammel kontrollgruppe. Dette var en liten studie, og forekomsten av problemer var lav, ikke minst i kontrollgruppen.

Målene på indre konsistens var basert på få studier, men store utvalg. Skalaene hadde gjennomgående tilfredsstillende til god indre konsistens, til forskjell fra den norske selvrappoteringsversjonen av SDQ (SDQ-S), hvor kun én av skalaene hadde akseptabel indre konsistens (Kornør & Heyerdahl, 2013). Dette er i tråd med funn fra en nasjonal britisk studie, hvor også SDQ-T-skalaene hadde god indre konsistens, mens SDQ-S skalaene hadde noe dårligere reliabilitet (Goodman, 2001).

Impact-skåren kan være et nyttig supplement til problemskåren, men med det foreliggende norske datagrunnlaget må Impact-skåren tolkes med forsiktighet.

Den norske versjonen av SDQ-T kan være egnet til å screene for psykiatriske diagnoser i selekterte utvalg med risiko for psykiske helseproblemer. Det er ikke grunnlag for å anbefale SDQ-T for universell screening. Man må være oppmerksom på at den prediktive verdi kan være lav slik at det kan bli relativt mange falske positive og/ eller negative hvis instrumentet brukes til screening i normal-befolkning.

Instrumentet har dokumentert støtte for god diskriminant begrepsvaliditet. Mer forskning er imidlertid nødvendig for å få dokumentert den sammenfallende begrepsvaliditeten bedre ved å undersøke hvordan SDQ-T-skårer forholder seg til skårer fra tilsvarende, validerte instrumenter. Det er også behov for bedre normdata, ikke minst for eldre barn og ungdommer, og også studier som bruker SDQ for både foreldre, lærere og ungdom.

## Referanser

- Achenbach, T. M., Becker, A., Dopfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C. et al. (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.
- Ambrosini, P. J. (2000). Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (K-SADS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 49-58.
- Egger, H. L., Erkanli, A., Keeler, G., Potts, E., Walter, B. K. & Angold, A. (2006). Test-Retest Reliability of the Preschool Age Psychiatric Assessment (PAPA). *J Am Acad Child Adolesc Psychiatry*, 45(5), 538-549. doi: 10.1097/01.chi.0000205705.71194.b8
- Ehlers, S., Gillberg, C. & Wing, L. (1999). A Screening Questionnaire for Asperger Syndrome and Other High-Functioning Autism Spectrum Disorders in School Age Children. *Journal of Autism and Developmental Disorders*, 29(2), 129-141. doi: 10.1023/A:1023040610384
- Eilertsen, M.-E. B., Rannestad, T., Indredavik, M. S. & Vik, T. (2011). Psychosocial health in children and adolescents surviving cancer. *Scandinavian Journal of Caring Sciences*, 25(4), 725-734. doi: 10.1111/j.1471-6712.2011.00883.x
- Eivers, A. R., Brendgen, M. & Borge, A. I. (2010). Stability and change in prosocial and antisocial behavior across the transition to school: Teacher and peer perspectives. *Early Education and Development*, 21(6), 843-864. doi: <http://dx.doi.org/10.1080/10409280903390684>
- Eivers, A. R., Brendgen, M., Vitaro, F. & Borge, A. I. (2012). Concurrent and longitudinal links between children's and their friends'

- antisocial and prosocial behavior in preschool. *Early Childhood Research Quarterly*, 27(1), 137-146. doi: <http://dx.doi.org/10.1016/j.ecresq.2011.05.001>
- Ekornås, B., Heimann, M., Tjus, T., Heyerdahl, S. & Lundervold, A. J. (2011). Primary school children's peer relationships: Discrepancies in self-perceived social acceptance in children with emotional or behavioral disorders. *Journal of Social and Clinical Psychology*, 30(6), 570-582. doi: <http://dx.doi.org/10.1521/jscp.2011.30.6.570>
- European Federation of Psychologist Association (EFPA). (2008). EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers (3.42 utg.): EFPA.
- Ford, T., Hutchings, J., Bywater, T., Goodman, A. & Goodman, R. (2009). Strengths and Difficulties Questionnaire Added Value Scores: evaluating effectiveness in child mental health interventions. *The British Journal of Psychiatry*, 194(6), 552-558. doi: 10.1192/bjp.bp.108.052373
- Goodman, A., Lamping, D. & Ploubidis, G. (2010). When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology*, 38(8), 1179-1191. doi: 10.1007/s10802-010-9434-x
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38(5), 581-586. doi: 10.1111/j.1469-7610.1997.tb01545.x
- 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*, 40(5), 791-799. doi: 10.1111/1469-7610.00494
- Goodman, R., Ford, T., Richards, H., Gatward, R. & Meltzer, H. (2000). The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41, 645-55.
- Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337-1345. doi: 10.1097/00004583-200111000-00015
- Goodman, R., Ford, T., Simmons, H., Gatward, R. & 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*, 177(6), 534-539. doi: 10.1192/bjp.177.6.534
- Hanssen-Bauer, K., Heyerdahl, S. & Eriksson, A.-S. (2007). Mental health problems in children and adolescents referred to a national epilepsy center. *Epilepsy & Behavior*, 10(2), 255-262. doi: 10.1016/j.yebeh.2006.11.011
- Heiervang, E., Goodman, A. & 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*, 49(6), 678-685. doi: <http://dx.doi.org/10.1111/j.1469-7610.2008.01882.x>
- Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M.-B. et al. (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 & Adolescent Psychiatry*, 46(4), 438-447. doi: <http://dx.doi.org/10.1097/chi.0b013e31803062bf>
- Holen, S., Lervåg, A., Waaktaar, T. & Ystgaard, M. (2012). Exploring the associations between coping patterns for everyday stressors and mental health in young schoolchildren. *Journal of School*

- Psychology*, 50(2), 167-193. doi: <http://dx.doi.org/10.1016/j.jsp.2011.10.006>
- Holen, S., Waaktaar, T., Lervag, A. & Ystgaard, M. (2012). The effectiveness of a universal school-based programme on coping and mental health: A randomised, controlled study of Zippy's Friends. *Educational Psychology*, 32(5), 657-677. doi: <http://dx.doi.org/10.1080/01443410.2012.686152>
- Indredavik, M. S., Vik, T., Heyerdahl, S., Kulseng, S. & Brubakk, A.-M. (2005). Psychiatric symptoms in low birth weight adolescents, assessed by screening questionnaires. *European Child & Adolescent Psychiatry*, 14(4), 226-236. doi: 10.1007/s00787-005-0459-6
- Javo, C., Rønning, J. A., Handegard, B. H. & Rudmin, F. W. (2009). Cross-informant correlations on social competence and behavioral problems in Sami and Norwegian preadolescents. *European Child & Adolescent Psychiatry*, 18(3), 154-163. doi: <http://dx.doi.org/10.1007/s00787-008-0714-8>
- Kornør, H. & Heyerdahl, S. (2013). Måleegenskaper ved den norske versjonen av Strengths and Difficulties Questionnaire, selvrappport (SDQ-S). *PsykTestBarn*(2), 13.
- Lundervold, A., Heimann, M. & Manger, T. (2008). Behaviour-emotional characteristics of primary-school children rated as having language problems. *British Journal of Educational Psychology*, 78(4), 567-580. doi: <http://dx.doi.org/10.1348/000709908X320480>
- Munkvold, L., Lundervold, A., Lie, S. A. & Manger, T. (2009). Should there be separate parent and teacher-based categories of ODD? Evidence from a general population. *Journal of Child Psychology and Psychiatry*, 50(10), 1264-1272. doi: <http://dx.doi.org/10.1111/j.1469-7610.2009.02091.x>
- Munkvold, L. H., Lundervold, A. J. & Manger, T. (2011). Oppositional defiant disorder-Gender differences in co-occurring symptoms of mental health problems in a general population of children. *Journal of Abnormal Child Psychology*, 39(4), 577-587. doi: <http://dx.doi.org/10.1007/s10802-011-9486-6>
- Nordhov, S., Rønning, J. A., Ulvund, S. E., Dahl, L. B. & Kaaresen, P. I. (2012). Early intervention improves behavioral outcomes for preterm infants: Randomized controlled trial. *Pediatrics*, 129(1), e9-e16. doi: <http://dx.doi.org/10.1542/peds.2011-0248>
- Posserud, B., Lundervold, A. J., Steijnen, M. C., Verhoeven, S., Stormark, K. M. & Gillberg, C. (2008). Factor analysis of the Autism Spectrum Screening Questionnaire. *Autism*, 12(1), 99-112. doi: <http://dx.doi.org/10.1177/1362361307085268>
- Sanne, B., Torsheim, T., Heiervang, E. & 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: <http://dx.doi.org/10.1037/a0016317>
- Stormark, K. M., Heiervang, E., Heimann, M., Lundervold, A. & Gillberg, C. (2008). Predicting Nonresponse bias from teacher ratings of mental health problems in primary school children. *Journal of Abnormal Child Psychology*, 36(3), 411-419. doi: <http://dx.doi.org/10.1007/s10802-007-9187-3>
- Sveen, T. H., Berg-Nielsen, T. S., Lydersen, S. & Wichstrom, L. (2013). Detecting psychiatric disorders in preschoolers: Screening with the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(7), pp. doi: 10.1016/j.jaac.2013.04.010 23800486
- Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B. et al. (2001). Clinical relevance of the

primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment.

*Journal of the American Academy of Child & Adolescent Psychiatry*, 40(2), 168-179.

Sørensen, L., Hugdahl, K. & Lundervold, A. J. (2008).

Emotional symptoms in inattentive primary school children: A population-based study. *Journal of Attention Disorders*, 11(5), 580-587. doi:

<http://dx.doi.org/10.1177/1087054707308491>

Sørensen, L., Plessen, K. J. & Lundervold, A. J.

(2012). The impact of inattention and emotional problems on cognitive control in primary school children. *Journal of Attention Disorders*, 16(7), 589-599. doi:

<http://dx.doi.org/10.1177/1087054711417394>

Ullebø, A. K., Posserud, M.-B., Heiervang, E.,

Gillberg, C. & Obel, C. (2011). Screening for the attention deficit hyperactivity disorder phenotype using the strength and difficulties questionnaire. *European Child & Adolescent Psychiatry*, 20(9), 451-458. doi:

<http://dx.doi.org/10.1007/s00787-011-0198-9>