



**Regionsenter for barn og
unges psykiske helse**

Helseregion Øst og Sør

What is a normal SDQ?

SDQ Workshop

Aarhus March 17th 2016

Sonja Heyerdahl, Senior researcher emerita
MD, PhD, Child and adolescent psychiatrist
Center for Child and Adolescent Mental Health
Eastern and Southern Norway, Oslo

RBUP

Why do we need to know «What is a normal SDQ?»»

- We need to find/identify children with mental health problems
- We need good methods for the assessment
- What is «good»?
 - Fits the purpose
 - Good psychometric properties
 - Relevant norms and cut off values
- The web site <http://www.psyktestbarn.no/> evaluates assessment methods in child mental health

PSYKTEST

BARN Måleegenskaper ved tester og kartleggingsverktøy

- **Aim:**
- To provide systematic reviews of Norwegian versions of tests and assessment methods in child and adolescent mental health
- Review psychometric psychometric properties
- Provide knowledge which can help clinicians to choose good quality assessment methods in their clinical work
- Publish the reviews as online papers



Velkommen til PsykTestBarn

PsykTestBarn er et elektronisk tidsskrift som utgir artikler om måleegenskaper ved norske versjoner av tester og kartleggingsverktøy for å undersøke psykisk helse, psykososiale problemer, evner og ferdigheter hos barn og ungdom. Ansvarlig for tidsskriftet og nettstedet er RBUP i samarbeid med de regionale kunnskapssentrene for barn og unge (RKBU Nord, RKBU Midt og RKBU Vest). Vi bruker systematiske metoder for å finne forskningsbasert dokumentasjon og for å evaluere måleegenskapene. Målet med PsykTestBarn er å publisere artikler for tester og kartleggingsverktøy som brukes i norsk praksis og forskning. Tips oss gjerne om relevante tester som mangler!



MELD DEG PÅ
VÅRT NYHETSBREV

Siste utgave av PsykTestBarn

Vineland Adaptive Behavior Scales
Sonja Heyerdahl, Svein Eikeseth
2014, 1:2 (publisert 18.mars 2014)

CORS - Child Outcome Rating Scale
Hege Kornør, Ketil Hanssen-Bauer
2014, 1:1 (publisert 28.januar 2014)

[Flere utgaver av PsykTestBarn](#)

Tester under evaluering

ADBB - Alarm Distress Baby Scale

BRIEF - Behavior Rating Inventory of Executive Function

CGAS - Children's Global Assessment Scale

CSRS - Child Session Rating Scale

K-SADS-PL - Schedule for Affective Disorders and Schizophrenia for School Aged Children - Present and Lifetime version

KIDSCREEN

LOI-CV - Leyton obsessive Inventory Child Version

YSR - Youth Self Report

[Flere tester under evaluering](#)

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

Hege Kornør¹ (hege.kornor@r-bup.no)

Sonja Heyerdahl¹ (sonja.heyerdahl@r-bup.no)

PsykTestBarn 2014, 2:5

ISSN 1893-9910

Mottatt dato 27.06.2014

Publisert dato 15.10.2014

Artikkelens URL <http://www.r-bup.no/CMS/ptb.nsf/pages/sdq>

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.

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.

¹ RBUP Øst og Sør

Paper on SDQ-T

22 publications included here

For the three versions P – T – S

70 Norwegian papers with relevant data
(psychometric, norms etc) included

Most of the papers are general
population studies

What about Norwegian norms?

«What is a normal SDQ?»

Unfortunately - We do not have national Norwegian norms

We have good regional datasets from the general population

- Good procedures and response rates
- But – most often limited age groups
- Most often selected informants

Can we use British norms?

Parent reported problems – SDQ

Consistency between studies?

Total problems Mean (SD)	Bergen Child Study Age 7-9 years N=5 997	Akershus (county - around Oslo) N=5 410
Age 7-9 years	5.5 (4.5)	6.1 (4.8)
Age 10-13	5.3	Boys 6.6 (5.2) Girls 5.7 (4.8)

Parent reported problems – Consistency between countries?

Total problems Mean (SD)	Bergen Child Study N=5 997/4 344	Akershus (county around Oslo) N=5 410/7 553	Denmark N=56 531/6 559	Britain N=2 954/2 191
Age 7-9 years	Age 7-9 5.5 (4.5)	Age 7-9 6.1 (SD 4.8)	Age 5-7 Boys 6.4 (4.8) Girls 5.5 (4.2)	Age 5-10 8.6 (5.7)
Age 10-13	5.3	Boys 6.6 (5.2) Girls 5.7 (4.8)	Age 10-12 Boys 6.0 (5.2) Girls 5.3 (4.7)	Age 11-15 Boys 8.8 (5.9) Girls 7.6 (5.6)

Differences in data reported, by gender etc

Cut off values Total difficulties – SDQ-P

	Akershus (county around Oslo) Age 10-13	Denmark Age 10-12	Britain Age 4-17
Normal	0-10	0-10 boys 0-8 girls	0-13
Borderline	11-13	11-13 boys 9-11 girls	14-16
Abnormal	14-40	14-40 boys 12-40 girls	17-40

The Nordic advantage in child mental health: separating health differences from reporting style in a cross-cultural comparison of psychopathology

Einar Heiervang,¹ Anna Goodman,² and Robert Goodman³

Can we use British norms and cut off values in Norway?

Parents and teachers report less problems than in Great Britain

Parent reported problems

	Britain	Norway
Total problemer	8.4 (SD 6.4)	5.9 (SD 4.9)

What does this mean?

Less problems or differences in rating thresholds in the parental reports?

Results from the Norwegian/British comparison

- Information from SDQ and the DAWBA interview was compared across countries.
- Norwegian children had lower prevalence of externalising disorders (in accordance with SDQ results)
- Norwegian children had similar rates of emotional disorders, (although SDQ results were lower) indicating that Norwegian parents underreported these problems

Complicated!

- Appropriate Norwegian norms are lacking
- British results differ from Norwegian results, and British cut off values for parental reports are higher
- Still, in clinical work, we use British cut off values
- Cut off values should be used cautiously, consider how high the scores are

Table 1 Mean SDQ scores and standard deviations in Norway and Britain, according to parents and teachers

SDQ score	Parent report			Teacher report		
	Britain Mean (SD) <i>N</i> = 4898	Norway Mean (SD) <i>N</i> = 6658	Difference in mean (Britain – Norway) (95%CI)	Britain Mean (SD) <i>N</i> = 3836	Norway Mean (SD) <i>N</i> = 6184	Difference in mean (Britain – Norway) (95%CI)
Total symptoms	8.37 (6.39)	5.87 (4.87)	2.50 (2.29,2.72)	6.57 (6.07)	4.12 (4.79)	2.46 (2.23,2.68)
Emotional	1.92 (2.09)	1.30 (1.71)	.62 (.55,.69)	1.52 (1.97)	.64 (1.32)	.88 (.81,.96)
Conduct	1.53 (1.84)	.96 (1.28)	.58 (.52,.64)	.90 (1.66)	.61 (1.26)	.29 (.23,.35)
Hyperactivity	3.47 (2.82)	2.65 (2.15)	.82 (.72,.91)	2.85 (2.79)	2.12 (2.36)	.73 (.62,.84)
Peer	1.45 (1.81)	.96 (1.54)	.49 (.42,.55)	1.30 (1.75)	.74 (1.46)	.55 (.49,.62)
Prosocial (higher more desirable)	8.81 (1.62)	8.52 (1.53)	.30 (.24,.36)	7.59 (2.48)	8.44 (2.01)	-.85 (-.94,-.76)
Impact	.46 (1.33)	.16 (.94)	.30 (.25,.34)	.43 (.94)	.20 (.83)	.22 (.18,.26)
Burden	.37 (.78)	.19 (.55)	.17 (.15,.20)	.37 (.071)	.23 (.60)	.15 (.12,.18)

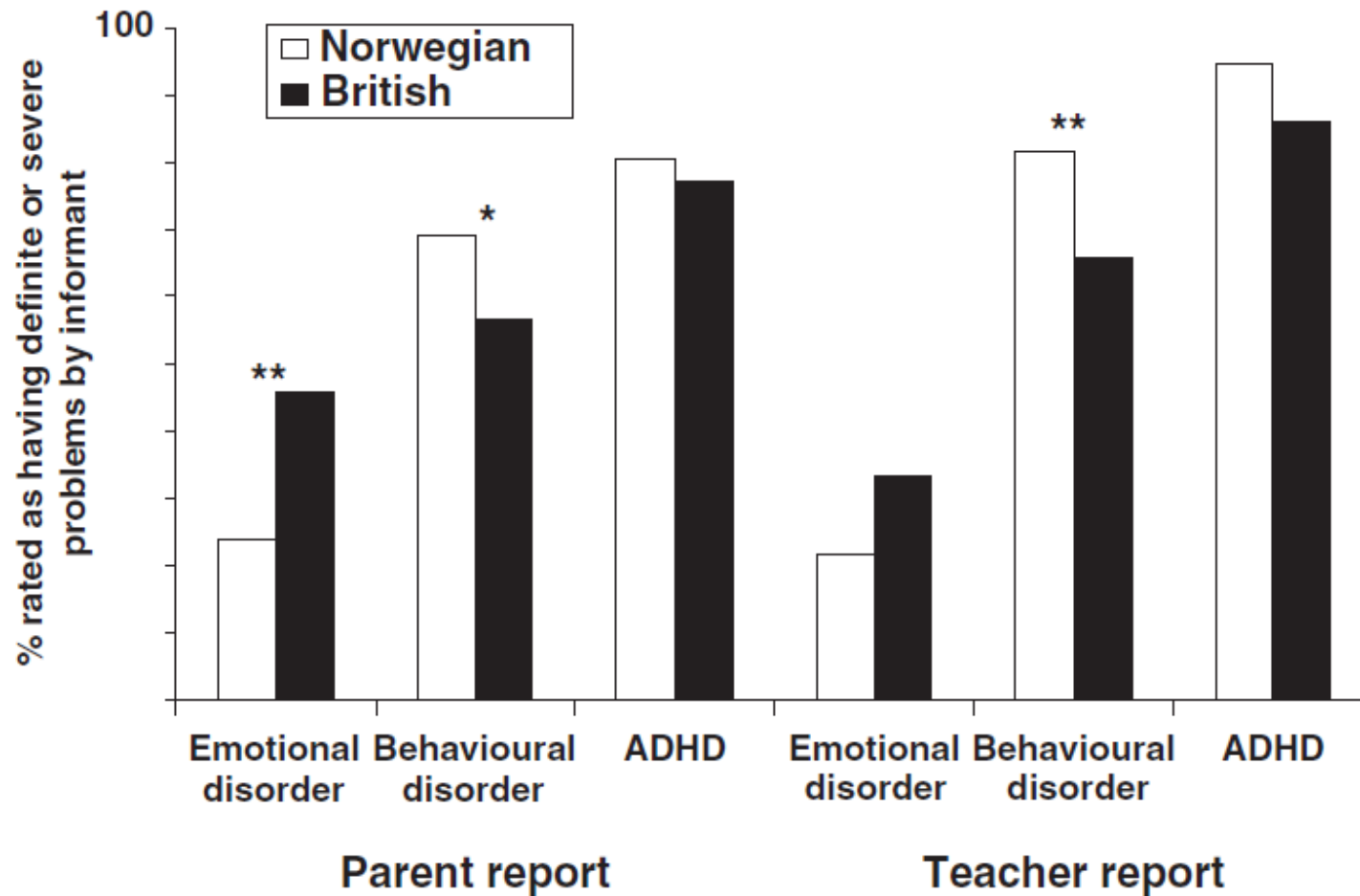
All Norway–Britain differences in means significant at $p < .001$, adjusted for clustered sample design.

Table 2 Weighted prevalence of DSM-IV disorders in 8–10-year-olds in Norway and Britain

	Britain (percent and 95% CI)	Norway (percent and 95% CI)	Difference (Britain – Norway), and 95%CI	<i>p</i> -value of difference
Emotional disorder(anxiety/depression)	3.0 (2.5–3.5)	3.2 (2.0–4.4)	.3 (–1.5–1.0)	.699
Behavioural disorder (oppositional/conduct)	4.8 (4.2–5.4)	2.5 (1.9–3.0)	2.3 (1.5–3.2)	< .001
ADHD ¹	2.5 (2.1–2.9)	1.3 (.8–1.7)	1.2 (.6–1.8)	< .001
ASD ²	.7 (.5–1.0)	.4 (.2–.6)	.3 (.00–.07)	.048
Any disorder	8.7 (7.9–9.5)	6.1 (4.7–7.4)	2.6 (1.0–4.2)	.001

¹ADHD = attention-deficit/hyperactivity disorder; ²ASD = autistic spectrum disorder.

All 4,898 British children had detailed psychiatric diagnostic measures. Only 1,024 of the 6,658 Norwegian children had detailed psychiatric diagnostic measures – prevalences weighted back to the 6,658 adjusting for selective sampling and participation. Confidence intervals and *p*-values adjust for clustered sample design in Britain and weight for selective sampling in Norway.



Difference between Norway and Britain: * $p < .05$ ** $p < .01$

Figure 1. Proportion of Norwegian and British children with different groups of disorders who are reported to have definite or severe problems by parents and teachers



**Regionsenter for barn og
unges psykiske helse**

Helseregion Øst og Sør

How do we get the best out of more informants?

SDQ Workshop

Aarhus March 17th 2016

Sonja Heyerdahl, Senior researcher emerita
MD, PhD, Child and adolescent psychiatrist
Center for Child and Adolescent Mental Health
Eastern and Southern Norway, Oslo

RBUP

SDQ Informants – parents, teachers, selfreport

- To get information from many informants is of great value
 - Different perspectives
 - Different contexts (home, school)
- Do they tell the same story?
 - Agreement between parents, teachers and adolescents is usually not high
 - Both agreement and disagreement between informants about problems and competencies is important information in clinical work

SDQ results from 2 Norwegian studies with more informants

	Bergen Child Study Age 7-9 years		Akershus study Age 10-13 years	
Total problems Mean (SD)	Parent SDQ N= 6 658	Teacher SDQ N= 6 184	Parent SDQ N= 4 279/4 238	SDQ Self-report N= 4 101/4 096
	5.9 (4.9)	4.1 (4.8)	Boys 6.6 (5.2) Girls 5.7 (4.8)	Boys 10.1 (5.2) Girls 10.0 (5.1)
	<p>For both informants lower SDQ scores seemed to reflect real differences in prevalence of behavioral disorders between Britain and Norway, while for emotional disorders lower SDQ scores seemed to reflect underreporting.</p> <p><i>(Heiervang et al 2008)</i></p>		<p>Self-reported results were higher than parent reported results, and in accordance with British results.</p> <p><i>(Van Roy et al 2008)</i></p>	

Norske normalutvalg

SDQ Selvrappport

Tabell 3. Grenseverdier for Totale vansker

	Normal	Grenseområde	Antatt klinisk område
Britiske normer	0-15	16-19	20-40
Akershus	0-15	16-18	19-40
Nord-Norge	0-14	15-17	18-40
Nye britiske normer <i>2014</i>	0-14	15-17	18-40

For foreldre-rapporterte og lærer-rapporterte problemer er ikke grenseverdier angitt for de norske studiene

How is agreement between SDQ results for different informants and diagnosis?

- 5 Norwegian papers have studied this
- Different study populations
 - Age groups
 - Prevalence of disorders
- Different diagnostic methods

Agreement between a test and diagnosis Screening

	Diagnose +	Diagnose -	
Test +	a TP	b FP	T=True P=Positive N=Negative
Test -	c FN	d TN	
Sensitivity	$a/a+c$	Positive predictive ability $a/a+b$	
Specificity	$d/b+d$	Negative predictive ability $d/c+d$	

- **Sensitivitet** er en tests evne til å identifisere de som har en diagnose eller definerte problemer
- **Spesifisitet** er en tests evne til å identifisere de som ikke har en diagnose eller problemer
- For en perfekt test vil sensitivitet og spesifisitet være 1 (eller 100%). Alle med problemer identifiseres, og alle uten problemer identifiseres til ikke å ha problemer

Fem norske studier har sett på samsvar mellom SDQ og diagnose

1. Indredavik et al. 2005. Barn med lav fødselsvekt SDQ-P, S, T
2. Ullebø et al 2011. Normalutvalg, SDQ-P, T. Evne til å identifisere barn med ADHD
3. Brøndbo et al. 2011. Barn i BUP. SDQ skårings-algoritme – kombinerer P-S-T.
4. Sveen et al. 2013. Førskolebarn Normalutvalg SDQ-P, T
5. Lehmann et al. 2014. Barn i fosterhjem. SDQ-P, T og SDQ skåringsalgoritme

Diagnostisk metode som SDQ er sammenlignet med

1. Indredavik et al. 2005. Barn med lav fødselsvekt SDQ-P, S, T
Kiddie SADS- Intervju
2. Ullebø et al 2011. Normalutvalg, Evne til å identifisere barn med ADHD
SNAP-IV spørreskjema
3. Brøndbo et al. 2011. Barn i BUP. SDQ skårings-algoritme –
kombinerer P-S-T. **DAWBA- Intervju**
4. Sveen et al. 2013. Førskolebarn Normalutvalg SDQ-P, T. **PAPA-
Intervju**
5. Lehmann et al. 2014. Barn i fosterhjem. SDQ-P, T og SDQ
skåringsalgoritme. **DAWBA- Intervju**



RESEARCH

Open Access

The Strengths and Difficulties Questionnaire as a Screening Instrument for Norwegian Child and Adolescent Mental Health Services, Application of UK Scoring Algorithms

Per Håkan Brøndbo^{1,2*}, Børge Mathiassen^{1,2}, Monica Martinussen², Einar Heiervang³, Mads Eriksen⁴,
Therese Fjeldmo Moe⁵, Guri Sæther⁶ and Siv Kvernmo^{1,2}

Detecting Psychiatric Disorders in Preschoolers: Screening With the Strengths and Difficulties Questionnaire

Trude Hamre Sveen, Psy.D., Turid Suzanne Berg-Nielsen, Ph.D., Stian Lydersen, Ph.D.,
Lars Wichstrøm, Ph.D.

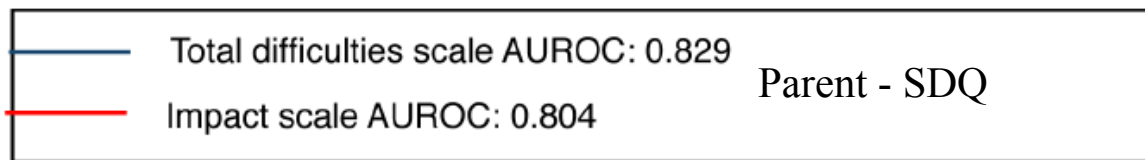
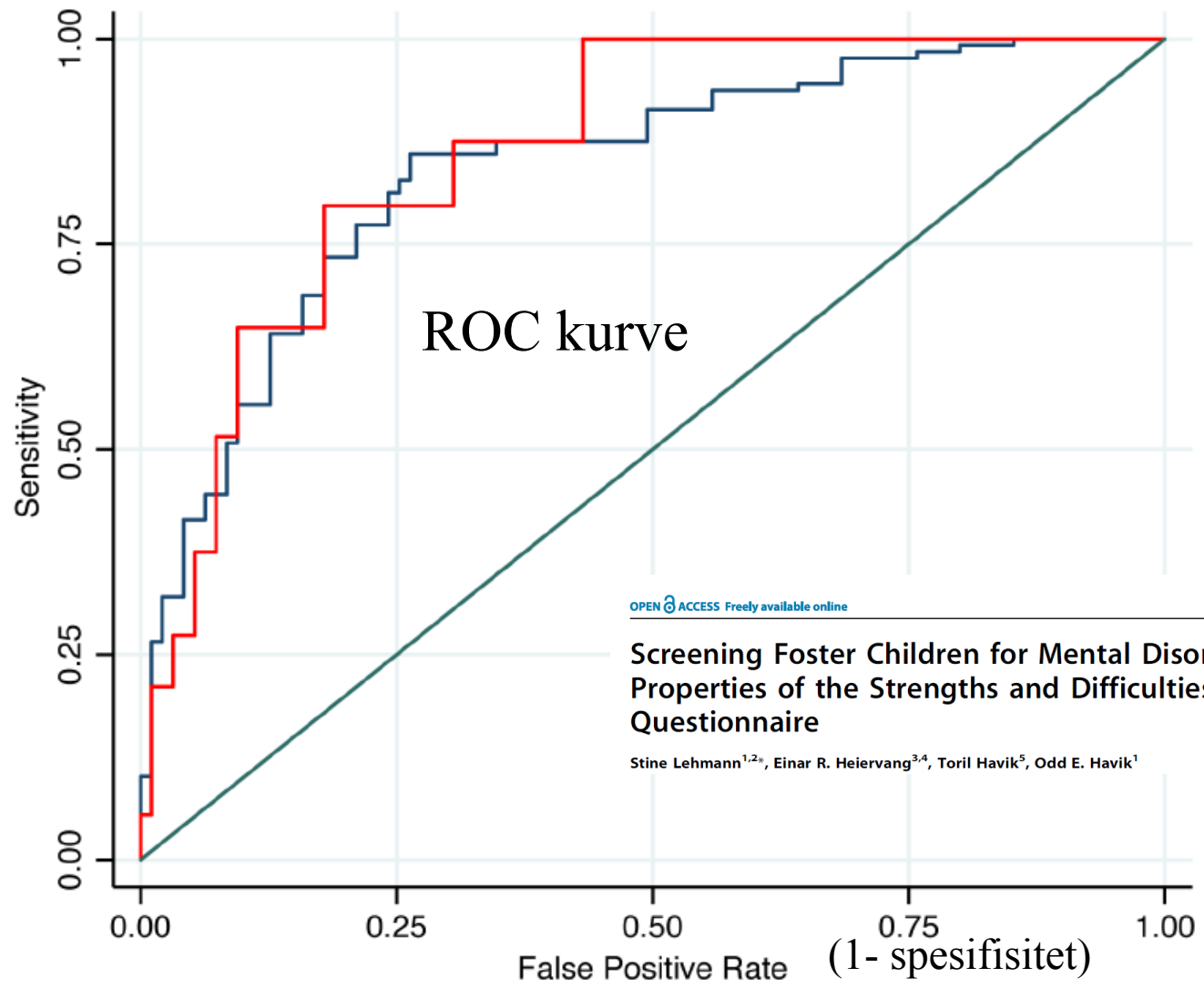
JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY
VOLUME 52 NUMBER 7 JULY 2013

Screening Foster Children for Mental Disorders: Properties of the Strengths and Difficulties Questionnaire

Stine Lehmann^{1,2*}, Einar R. Heiervang^{3,4}, Toril Havik⁵, Odd E. Havik¹

Studiene – ikke sammenlignbare

- Store forskjeller
- Forskjellige utvalg
- Stor forskjell i prevalens av problemer og type problemer
- Forskjell i alder
- Forskjell i diagnostisk metode
 - Og om diagnostikk gjøres av en annen enn den som fyller ut spørreskjema



«Areal under kurven»

AUC

- Viser hvor god prediksjon SDQ har i forhold til diagnose (kriteriet). Ved optimal prediksjon er AUC 1.
- Et overordnet mål, som sammenfatter sensitivitet og spesifisitet

AUC for Total problemer i forhold til 'Any Disorder'

	AUC	Kommentar
Barn med lav fødselsvekt <i>Indredavik et al</i>	0.67-0.85	Liten studie, undergrupper Vurdert for ulike informanter
Barn i BUP <i>Brøndbo et al</i>	-	Vurdert algoritme
Førskolebarn <i>Sveen et al</i>	SDQ-P 0.76 SDQ-T 0.65	Høyere for spesifikke lidelser (emosj. og atferd)
Barn i fosterhjem <i>Lehmann et al</i>	SDQ-P 0.83 SDQ-T 0.77	Vurdert som god 'akseptabel'

Screening properties

Discriminative ability with different cut off values

SDQ Total Difficulties

In relation to at least one mental health diagnosis

	SDQ-P Score (cut off)		SDQ-T Score (cut-off)	
	≥13	≥14	≥12	≥13
Sensitivity	0.83	0.81	0.86	0.86
Specificity	0.74	0.75	0.71	0.77
PPV	0.81	0.81	0.74	0.78
NPV	0.76	0.74	0.84	0.86

Lehmann et al 2014, Foster children

Screening-properties

Discriminative ability with different cut off values

	SDQ-P Score (cut off)		SDQ-T Score (cut-off)	
	≥ 13	≥ 14	≥ 12	≥ 13
Sensitivity	0.83	0.81	0.86	0.86
Speccificity	0.74	0.75	0.71	0.77
British cut off values	Abnormal 17 Borderline 14		Abnormal 16 Borderline 12	

Norwegian cut off values that provide good sensitivity and specificity are considerably lower than the British

Lehmann et al 2014, Fosterbarn

SDQ- Ability to identify children with a mental health diagnosis

- «Screening»-properties for **total difficulties** are reasonably good for Parent and Teacher SDQ. Still there will be quite a few false positives and false negatives. We lack information for SDQ Selfreport.
- How many false positives and false negatives we can tolerate depends on the purpose for the assessment
- The cutoff values with best prediction are lower than the British cutoff values for SDQ-P and SDQ-T, but probably close to the British for SDQ-S

Hva med subskalaene og Impact-skåren? Og funn for ulike informanter?

•Table 1. Area Under the Receiver Operating Curve for SDQ Scales.

SDQ scales on DAWBA diagnostic groups	Caregiver SDQ (n = 223)	Teacher SDQ (n = 195)
	AUROC 95% CI	AUROC 95% CI
Total difficulties on Any disorder	.83 [.78, .88]	.77 [.71, .86]
Impact on Any disorder	.80 [.75, .86]	.75 [.68, .82]
Emotional subscale on Emotional disorder	.82 [.76, .88]	.74 [.66, .82]
Conduct subscale on Behavioral disorder	.89 [.84, .93]	.86 [.80, .93]
Hyperactive subscale on ADHD	.81 [.74, .87]	.80 [.72, .87]

doi:10.1371/journal.pone.0102134.t001

Lehmann S, Heiervang ER, Havik T, Havik OE (2014) Screening Foster Children for Mental Disorders: Properties of the Strengths and Difficulties Questionnaire. PLoS ONE 9(7): e102134. doi:10.1371/journal.pone.0102134
<http://www.plosone.org/article/info:doi/10.1371/journal.pone.0102134>

Sveen vurderte om Impact-skåren og lærerskåren økte den prediktive verdi i forhold til foreldrevurdering av totale vansker for førskolebarn.

De fant ingen vesentlig økning ved å inkludere andre skårer

The SDQ multi-informant-algorithm

- A nice initiative to put together scores from several informants
- Predicts whether disorder is «unlikely», «possible» or «probable»
 - Emotional disorder
 - Behavioral disorder
 - Hyperactivity or concentration disorder
 - Any diagnosis
- An algorithm summarizes results from several informants for both Problemscores and Impact scores.

'Diagnostisk prediksjon'

- Algoritmen gir prediksjon for
 - Emotional disorder
 - Behavioral disorder
 - Hyperactivity or concentration disorder
 - Any diagnosis
- Konklusjonen angir
 - Lav risiko for diagnose - 'unlikely' ca 80%
 - mulig risiko - 'possible' ca 10%
 - høy risiko - 'probable' ca 10%

SDQ Screeningegenskaper i norsk BUP.

SDQ Any disorder

- DAWBA diagnose for 66%, SDQ Any disorder for 72%

- 1. **Screen +** 'Probable'. **Screen -** 'Possible' og 'Unlikely'
- 2. **Screen +** 'Probable' og 'Possible'. **Screen -** 'Unlikely'

	'Probable'	'Probable' + 'possible'
Sensitivitet	0.85	0.96
Spesifisitet	0.52	0.24

- Lav spesifisitet betyr mange falske positive

• Brøndbo et al. *SDQ as a screening instrument for Norwegian CAMHS*,

• *Child and Adolescent Psychiatry and Mental Health*, 2011

SDQ Screening abilities

SDQ Any disorder

- Results from Brøndbo et al. SDQ as a screening instrument for Norwegian CAMHS,
•Child and Adolescent Psychiatry and Mental Health, 2011
- Resultats from Lehmann et al. Screening foster children. PLOS 2014
- Resultats from Goodman et al. Using the SDQ multi-informant algorithm to screen looked-after children. Eur Child Adolesc Psychiatry 2004

	'Probable'	'Probable' + 'possible'
Sensitivitet	0.85 0.68 0.85	0.96 0.84
Spesifisitet	0.52 0.82 0.80	0.24 0.58

•

Results for the algorithm

- Both studies conclude that the algorithm can not be recommended to identify children with a mental health disorder.
- This does not mean that it is not useful to combine information from several informants
- To the contrary – Information from more informants and information on the impact of difficulties, is very useful
- Results must be interpreted by considering the profile and abnormality in the results from different informants and agreement/disagreement between informants