

Måleegenskaper ved den norske versjonen av Ages and Stages Questionnaires (ASQ)

Monica Martinussen¹ (monica.martinussen@uit.no)

Lisbeth Valla² (lisbeth.valla@r-bup.no)

PsykTestBarn 2013, 1:6

ISSN 1893-9910

Mottatt dato 23.06.2013

Publisert dato 01.07.2013

Artikkelens URL <https://psyktestbarn.r-bup.no/no/artikler/asq-ages-and-stages-questionnaire>

DOI 10.21337/0023

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.

¹ RKB Nord, Universitetet i Tromsø

² RBUP Øst og Sør

Denne artikkelen skal siteres på følgende måte:

Martinussen, M. & Valla, L. (2013). Måleegenskaper ved den norske versjonen av Ages and Stages Questionnaires (ASQ). *PsykTestBarn*. 1:6. doi: 10.21337/0023

Sammendrag

Beskrivelse. Ages & Stages Questionnaires ble utviklet på 80-tallet i USA som et screeningsverktøy for sped- og småbarn for å kartlegge utviklingsnivå (Squires, Bricker & Potter, 1997). Instrumentet ble oversatt til norsk i 2001-2002 av Harald Janson, Beathe Enoksen og Kristin Furuholmen ved Psykologisk institutt, UiO. Ulike versjoner av spørreskjemaet (til sammen 19) finnes for aldersgruppene 4 mnd til 60 mnd. Den norske testen og manualen forhandles av RBUP Øst og Sør. Testen består av til sammen 30 spørsmål som summeres opp i seks delskalaer og eventuelt en totalskåre. Det tar 10-20 minutter å besvare spørsmålene for foreldrene. Det stilles krav om relevant utdanning og opplæring for å kunne anvende testen. Testen er tenkt brukt både klinisk og til forskning.

Litteratursøk. Det systematiske litteratursøket resulterte i 24 treff hvorav seks oppfylte inklusjonskriteriene. I tillegg ble den norske manualen lagt til grunn. De inkluderte studiene rapporterte resultater fra fire ulike studier. Søket etter dokumentasjon for de svenske og danske versjonene resulterte i én artikkel som belyste testens validitet.

Psykometri. De inkluderte publikasjonene rapporterte gjennomsnittsskårer for testen, alfaverdier og test-retest reliabilitet. Median alfaverdi for totalskåren > 0,80 og test-retest reliabilitet > 0,70. Det foreligger norske normer, men utvalgene som disse er basert på er for enkelte aldersgrupper noe små.

Konklusjon. Testens totalskåre har god reliabilitet, mens for enkeltskalaer er denne utilfredsstillende. Det finnes norske normer basert på et populasjonsutvalg med foreslåtte grenseverdier. Testens validitet er kun i liten grad undersøkt på norske eller skandinaviske utvalg.

Abstract

Description. The Ages & Stages Questionnaires (ASQ) were developed in the 80's in the USA as a developmental screening tool for babies and young children (Squires, Bricker & Potter, 1997). The set of questionnaires was translated to Norwegian in 2001-2002 by Harald Janson, Beathe Enoksen, and Kristin Furuholmen at the Department of Psychology, University of Oslo. There are different versions of the questionnaire (a total of 19) for different age groups from 4 to 60 months. The Norwegian version, including the manual, may be purchased at www.r-bup.no. The test includes a total of 30 questions grouped into six subscales and a total score. It takes 10 – 20 minutes for parents to complete the questionnaire. Relevant education and training is required for professionals in order to use the test. The ASQ may be used for clinical and research purposes.

Literature search. The systematic literature search resulted in 24 articles of which six met the inclusion criteria and included in addition to the test manual. The included studies reported the results from four independent studies. The search for Swedish and Danish studies resulted in one article examining test validity.

Psychometrics. The included studies reported mean scores, alpha values, and test-retest reliability. Median alpha value for the total score was > .80 and test-retest reliability was >.70. Norwegian norms are available, but for some age groups based on relatively small samples.

Conclusion. The reliability of the total score was good, but unsatisfactory for some of the subscales. There are Norwegian norms based on a sample from the general population with suggested cut-off values. The test validity has only to some extent been examined based on Norwegian or Scandinavian samples.

Ages & Stages Questionnaires (ASQ, versjon 2, Bricker & Squire, 1999) er et foreldrebasert screeninginstrument som kartlegger barns utvikling innen fem områder: kommunikasjon, grovmotorikk, finmotorikk, problemløsningsevner, og personlig/sosial utvikling. ASQ består av 19 ulike spørreskjemaer for alderen 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 27, 30, 33, 36, 42, 48, 54 og 60 måneder. Hvert skjema inneholder til sammen 30 spørsmål, med 6 spørsmål innen hvert av de fem utviklingsområdene. I tillegg til de 30 spørsmålene, er det på slutten av hvert skjema generelle spørsmål som tar opp bekymringer som foreldre kan ha om barns utvikling og generelle helse. Barnets primære omsorgsperson er den som fyller ut skjemaet. Hvert spørsmål tar for seg ferdigheter som forventes å være utviklet i den angitte perioden, og disse besvares med enten, «Ja», «av og til» eller «ikke ennå». Spørsmålene innenfor hvert utviklingsområde skåres ved å sette 10 poeng for hvert «ja», «5» poeng for hvert av og til og «0» poeng for ikke ennå. Skåring gjøres på separate skåringsark. Hvis skåren for ett eller flere av utviklingsområdene er lik eller lavere enn grenseverdien som vises i skåringsarket, antyder dette at en nærmere undersøkelse er nødvendig. Hvis skåren er i nærheten av grenseverdien, bør resultatet følges opp. Hvis skåren er høyere enn grenseverdien, ser det ut til at barnet på det nåværende tidspunkt trolig er i rute på dette området. Det vil si, barnet ligger sannsynligvis innenfor den vanlige spredningen av mestring av ferdigheter innenfor det aktuelle område på det nåværende tidspunkt. Spørsmålene er skrevet i et enkelt og direkte språk og skal kunne forstås av personer med leseferdigheter tilsvarende 8. klassenivå. ASQ kan fylles ut på 10-20 minutter. Mindre lesevante foreldre vil kunne bruke mer tid, og skjemaene for eldre barn (36 måneder og mer) inneholder generelt noen flere aktiviteter som det tar tid å prøve med barnet. ASQ har norske normer

Innledning

med aldersspesifikke grenseverdier (Janson & Smith, 2003).

ASQ er ifølge manualen (Bricker & Squire, 1999) normert på 2008 amerikanske barn (1068 gutter og 940 jenter) i alderen fra 4 til 36 måneder. Test-retest reliabilitet er estimert med et intervall på 2 uker mellom måletidspunktene og viser en korrelasjon på 0,94 for totalskåren (N = 175). Inter-observer reliabilitet viser en korrelasjon på 0,94 (N = 112). Indre konsistens er estimert med Cronbachs alfa. Gjennomsnittlig indre konsistens varierte mellom aldersgruppene fra 0,63 til 0,75 for kommunikasjon, fra 0,53 til 0,87 for grovmotorikk, fra 0,79 til 0,49 for finmotorikk, fra 0,52 til 0,75 for problemløsning og fra 0,52 til 0,68 for personlig/sosial utvikling. ASQ har en sensitivitet på 0,75, og en spesifisitet på 0,86 ved 2 standardavvik under gjennomsnittet, og en sensitivitet på 0,82 og spesifisitet på 0,74 ved 1,5 standardavvik under gjennomsnitt (Bricker & Squire, 1999).

Metode

Det ble søkt etter dokumentasjon på testens psykometriske egenskaper i databasene Norart, BIBSYS, SveMed+, CRISTin.no, Cochrane Library, PsycINFO, Medline, Embase og PubMed (søkedato: 26.03.2012). Søkestrategien er tilgjengelig på <http://www.r-bup.no/CMS/ptb.nsf/pages/asq>. Vi kontaktet også personer i Norge som har brukt testen 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 skandinaviske utvalg:

- normdata for testen

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

I tillegg, og kun for norske versjoner av ASQ, inkluderte vi publikasjoner som rapporterte gjennomsnittsskårer og/eller forekomster for henholdsvis generelle populasjoner og kliniske 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.

Resultater

Litteratursøk

Norske studier: Det systematiske søket etter dokumentasjon på psykometriske egenskaper for den norske versjonen av ASQ resulterte i 24 treff, hvorav 11 ble vurdert i fulltekst for inklusjon. Av disse var det tre artikler som ble utelukket da de var basert på kortversjoner av utvalgte ASQ-skalaer (Brandlistuen, Stene-Larsen, Holmstrøm, Landolt, Eskedal & Vollrath, 2010; Brandlistuen, Stene-Larsen, Holmstrøm, Landolt, Eskedal & Vollrath, 2011b; Schølberg, Eadie, Zachrisson, Øyen & Prior, 2011), én artikkel ble utelukket fordi den ikke

inneholdt psykometriske analyser eller deskriptiv statistikk for ASQ (Janson, 2003) og én artikkel fordi den ikke omhandlet ASQ (Rostad, Nyberg & Sivberg, 2008). Dette resulterte i seks artikler basert på til sammen fire ulike studier som ble inkludert i oppsummeringen (se tabell 1) (Alvik & Grøholt, 2011; Brandlistuen, Stene-Larsen, Holmstrøm, Landolt, Eskedal & Vollrath, 2011a; Henriksen et al., 2008; Janson & Squires, 2004; Richter & Janson, 2007; Westerberg, Schei, Henriksen, Smith, Drevon & Iversen, 2011). I tillegg ble den norske manualen for ASQ lagt til grunn (Janson & Smith, 2003) der ytterligere detaljer om den norske populasjonsstudien presenteres (se også Janson & Squires, 2004; Richter & Janson, 2007).

Studier fra øvrige skandinaviske land: Litteratursøket resulterte i 28 skandinaviske studier hvorav 11 ble bestilt i fulltekst. Alle unntatt tre artikler omhandlet ASQ, men dette var i stor grad studier der ASQ var brukt til å studere forskjeller mellom grupper av premature barn for enten å avdekke utviklingsforstyrrelser eller studere effekten av en intervensjon. To artikler omhandlet ASQ-SE som er å betrakte som en egen test (Salomonsen & Sandell, 2011; Salomonsen & Slead, 2010). Én artikkel rapporterte resultater som direkte belyser testens validitet (Klamer, Lando, Pinborg & Greisen, 2005). Denne studien sammenlignet ASQ med andre måleinstrumenter noe som belyser begrepsvaliditeten til instrumentet.

Tabell 1. Inkluderte studier fra Norge

Referanse	Design	Populasjon	N	Mål	Rapporterte egenskaper, verdier
Alvik & Grøholt, 2011	Survey	Populasjon fra Oslo- 6 mnd	1053	Alle subskalaer + Total	M, SD + Cronbachs alfa for alle skalaer
Brandlistuen, et al., 2011a	Kohort-studie	Barn m/hjertefeil + kontrollgruppe fra MoBA-studien (3 år)	175 + 43929	Kun skala for kommunikasjon er brukt i sin helhet. Øvrige skalaer er kortutgaver brukt.	Theta (tilsvarer Cronbachs alfa) for hele gruppen.
Janson & Squires, 2004; Richter & Janson, 2007	Survey	Populasjonsutvalg- mange aldergrupper fra 4 mnd-5 år	1172/1461	Alle skalaer	M, SD for ulike aldergrupper. Alfa og test-retest reliabilitet
Henriksen et al., 2008; Westerberg, et al., 2011	Eksperiment (emærning)	Premature < 1500g (målt ved 6 mnd og 20 mnd)	92-105	Alle skalaler	M, SD. Alfa er ikke rapportert

Note: M = aritmetisk gjennomsnitt, SD = standardavvik, N = antall

Tabell 2. Middelerverdier, kliniske og andre undergrupper

Referanse	Utvalg/ Gruppe	N	Kommuni- kasjon	Grovmotor.	Finmotor.	Problem- løsning	Personlig- sosial	Total skåre
			M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Alvik & Grøholt, 2011	6 mnd gamle barn fra Oslo	1035-1054	44,1 (9,5)	36,9 (11,7)	40,1 (13,2)	50,8 (9,6)	47,2 (11,0)	219,1 (36,4)
Henriksen et al., 2008	Premature (6 mnd) tilskudd ernæring	50 interv.	45,4 (7,9)	33,3 (11,5)	45,2 (10,7)	53,4 (7,0)	43,2 (7,0)	221,0 (32,0)
		55 kontroll	46,6 (9,1)	30,9 (11,1)	45,8 (14,3)	49,5 (9,5)	49,5 (9,5)	215,0 (39,0)

Note: M = aritmetisk gjennomsnitt, SD = standardavvik, N = antall

Middelerverdier i kliniske undergrupper

I tabell 2 presenteres det gjennomsnittstall for to grupper; den ene er 6 mnd gamle barn fra Oslo og den andre er to grupper av premature barn som inngår i en evalueringsstudie om ernæring.

Normer

Utvalgte gjennomsnittsverdier og standardavvik for en større norsk studie (Janson & Smith, 2003; Janson & Squires, 2004; Richter & Janson, 2007) er presentert i tabell 3.

Validitet

Den norske populasjonsbaserte studien undersøkte en del sammenhenger mellom ASQ og bla sosiodemografiske forhold hos mor, barnets kjønn og om barnet var prematurt eller ikke (Richter & Janson, 2007). Resultatene indikerte lavere skårer hos premature barn noe som må sies å støtte begrepsvaliditeten til instrumentet. Det ble også avdekket forskjeller mellom jenter og gutter der jentene skåret signifikant høyere enn guttene på alle områder unntatt grovmotorikk. Forskjellene var størst for skalaene

Tabell 3. Gjennomsnittsverdier (M) og standardavvik (SD) fra representative utvalg av barn i Norge

Referanse	Utvalg/ Gruppe	N	Kommuni- kasjon	Grovmoto- r.	Finmotor.	Problem- løsning	Personlig/ -sosialt	Total
			M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Janson & Squires, 2004	4 mnd	176	50 (8)	55 (7)	50 (11)	55 (6)	50 (10)	261 (29)
	8 mnd	165	53 (7)	47 (13)	56 (7)	52 (8)	51 (8)	259 (27)
	12 mnd	145	42 (14)	46 (16)	52 (9)	51 (10)	44 (12)	236 (46)
	16 mnd	146	42 (13)	57 (9)	54 (8)	54 (9)	48 (10)	255 (34)
	20 mnd	138	47 (15)	57 (6)	52 (9)	50 (9)	51 (9)	258 (34)
	24 mnd	128	53 (10)	56 (6)	53 (8)	50 (9)	51 (8)	264 (30)
	30 mnd	134	57 (7)	56 (6)	50 (13)	52 (9)	53 (7)	267 (31)
	36 mnd	126	54 (7)	56 (7)	52 (10)	54 (9)	53 (8)	269 (31)
	48 mnd	100	56 (6)	54 (9)	50 (13)	54 (9)	56 (7)	271 (28)
60 mnd	82	55 (5)	55 (6)	51 (10)	52 (9)	56 (7)	269 (24)	

Note. Den norske manualen inneholder deskriptiv statistikk for flere aldersgrupper (Janson & Smith, 2003).

Kommunikasjon og Personlig/Sosial og utgjorde omtrent 1/3 standardavvik i størrelse. Forfatterne foreslår på bakgrunn av dette at det utvikles kjønnsespesifikke normer.

En av de skandinaviske studiene sammenlignet ASQ med WIPPSI-R som er en test på generell intelligens (Klamer, et al., 2005). Denne studien var basert på 30 barn i alderen 36-42 mnd. Korrelasjonen mellom totalskårene for begge tester resulterte i en korrelasjon på 0,48 noe som støtter at ASQ også måler kognitive evner (Klamer et al., 2005). En annen mulig fortolkning av denne korrelasjonen er at kognitiv utvikling sammenfaller i tid med de aspekter ved utvikling som ASQ måler. Dette kunne vært bedre belyst dersom korrelasjoner mellom delskalaer fra ASQ og WIPPSI-R hadde vært rapportert i den aktuelle artikkelen.

En større studie av barn i Oslo (6 mnd) (Alvik & Grøholt, 2011) der mødrene ble rekruttert under svangerskapet har indikerte visse problemer med grenseverdiene (cut-off) for klinisk område sammenlignet med tilsvarende grenser i den norske og den amerikanske manualen. Studien inkluderte nesten alle mødre (92 %) som møtte til ultralydundersøkelse ved Ullevål universitetssykehus i en periode. Studien kan dermed sies å være basert på et svært representativt utvalg av mødre i Oslo. Dersom man hadde anvendt grenseverdiene i manualene (norsk eller amerikansk) så ville omtrent en tredjedel av barna ha skårer under grenseverdien i ett eller flere områder, noe som indikerer behov for ytterligere kartlegging. Forfatterne problematiser dette og mener dette indikerer dårlig spesifisitet for denne aldersgruppen (mange falske positive) (Alvik & Grøholt, 2011). Normeringsstudien (Janson & Smith, 2003) og studien til Alvik og Grøholt (2011) er ulike mht utvalgene både hva størrelse og rekruttering angår som kanskje kan forklare disse forskjellene. Den norske populasjonsstudien er basert på et relativt lite utvalg for hver aldersgruppe (f.eks. $N = 169$ for aldersgruppen 6 mnd) men utvalget er til gjengjeld trukket fra hele Norge. Imidlertid er svarprosenten langt dårligere i populasjonsstudien (61 %) sammenlignet med studien fra Oslo (over 90 %) noe som kan indikere større problemer med seleksjonsbias i

populasjonsstudien. En annen forskjell var at i populasjonsstudien ble ASQ gjennomført hjemme hos familien, mens i studien basert på mødre fra Oslo, ble den gjennomført i en klinisk setting. Den amerikanske studien er basert på foreldre som selv velger å logge seg på en nettside og gjennomføre testen via denne, og det er dermed ikke mulig å anslå noen svarprosent.

Reliabilitet

Reliabilitet i form av Cronbachs alfa er rapportert i enkelte studier (se tabell 4).

Diskusjon og konklusjon

Det foreligger noe dokumentasjon på testens psykometriske egenskaper basert på norske og skandinaviske utvalg inklusive en større norsk studie basert på mødre fra hele Norge (Janson & Smith, 2003). Det foreligger en norsk manual som beskriver administrering, skåring og fortolkning av testresultatene (Janson & Smith, 2003).

Når det gjelder testens reliabilitet (Cronbachs alfa) så indikerer studiene noe varierende alfa både mellom delskalaer og aldersgrupper. For totalskåren er denne god og basert på et stort utvalg (over 0,80). Tilsvarende viste også test-retest reliabiliteten for totalskåren seg å være tilstrekkelig til god for totalskåren basert på alle aldersgrupper unntatt for de yngste (4-6 mnd)

(Janson & Smith, 2003). For delskalene var alfa generelt noe lav (medianverdier < 0,70). Dette kan ifølge forfatterne muligens skyldes lite variasjon i skårene i utvalget der man for enkelte aldersgrupper har svært mange barn som har full eller neste full skåre på delskalaer. En annen mulig forklaring på de noe lave alfaverdiene er at de ulike indikatorene (hvilke aktiviteter barnet kan på et gitt alderstrinn) slår inn noe ulikt i tid for den aktuelle aldersgruppen. I tillegg er antallet testledd innenfor hvert område relativt lite noe som også påvirker alfaverdiene.

Når det gjelder testens validitet er dette undersøkt ved enkelte gruppesammenligninger (f.eks. premature vs. fulltermin barn) og ved en korrelasjon med WIPPSI-R. Disse resultatene underbygger at testen måler utviklingsnivå. Normeringsutvalget er stort totalt sett, men for enkelte aldersgrupper i det minste laget (dvs < 150). Manualen inneholder ikke gjennomsnittsverdier for gutter og jenter separat selv om forfatterne i en publikasjon indikerer at dette er ønskelig (Richter & Janson, 2007). Imidlertid må de observerte forskjellene mellom jenter og gutter beskrives som små (standardiserte gjennomsnittsforskjeller i

Tabell 4. Indre konsistens (alfaverdier) og test-retest reliabilitet fra de tre norske studiene

Referanse	Utvalg gruppe	N	Skalaer	Alfa-verdier + test-retest
Alvik & Grøholt, 2011	Populasjon fra Oslo-6 mnd	1053	Alle subskalaer + Total	Kommunikasjon = 0,36, Grovmotorikk = 0,53; Finmotorikk = 0,65; Problemløsning = 0,57; Personlig-sosial = 0,78
Brandistuen, et al., 2011a	Barn m/hjertefeil + kontrollgruppe fra MoBA-studien (3 år)	175 + 439/29	Kun skala for kommunikasjon er brukt i sin helhet.	Theta (tilsvarer Cronbachs alfa) for Totalskåre = 0,93
Janson & Smith, 2003	Populasjonsutvalg-mange aldersgrupper fra 4 mnd-5 år	2633	Alle skalaer + totalt	alfa varierer mellom skalaer og aldersgrupper. Medianverdi over aldersgruppene (fra 4 til 60 mnd) var: Kommunikasjon = 0,66; Grovmotorikk = 0,63; Finmotorikk = 0,59; Problemløsning = 0,52; Personlig sosialt = 0,51; Total skåre = 0,82
Janson & Smith, 2003	Populasjonsutvalg	1079	Totalskåre	Stabilitet over 2-6 mnd basert på alle aldersgrupper. Median $r = 0,74$ (varierte mellom 0,48 og 0,89)

området 0,20-0,30). Svarprosenten i normeringsstudien var noe lav (60 %), og den var lavere for de eldste barna og for mødre med lavt utdanningsnivå (Janson, 2003). En studie basert på mødre fra Oslo med svært høy svarprosent resulterte i generelt noe lavere gjennomsnittsverdier på ASQ enn studien basert på mødre fra hele Norge. Dette kan skyldes flere forhold bla som større problemer med seleksjonsbias i normstudien eller kanskje ulikheter i måten testen ble administrert på (hjem vs klinikk). En tredje mulighet er jo at det er reelle forskjeller mellom barn i Oslo og fra landet generelt, men dette er antakelig mindre sannsynlig. Betydningen dette har i forhold til å anvende grenseverdiene i manualen kan være at man må utvise noe forsiktighet, og at testen dermed kan få en noe lav spesifisitet spesielt for enkeltskalaer – dvs at man får for mange falske positive.

Oppsummert kan man si at for testen totalskåre foreligger det god støtte for testens reliabilitet (både alfa og test-retest) og det foreligger norske normer. Testens validitet er i mindre grad undersøkt basert på norske eller nordiske studier.

Referanser

Alvik, A. & Grøholt, B. (2011). Examination of the cut-off scores determined by the Ages and Stages Questionnaire in a population-based sample of 6 month-old Norwegian infants. *BMC Pediatrics*, 11: 117. doi: 10.1186/1471-2431-11-117

Brandlistuen, R. E., Stene-Larsen, K., Holmstrøm, H., Landolt, M., Eskedal, L. T. & Vollrath, M. E. (2010). Motor and social development in 6-month-old children with congenital heart defects. *The Journal of Pediatrics*, 156, 265-269. doi: 10.1016/j.jpeds.2009.08.035

Brandlistuen, R. E., Stene-Larsen, K., Holmstrøm, H., Landolt, M. A., Eskedal, L. T. & Vollrath, M. E. (2011a). Occurrence and predictors of developmental impairments in 3-year-old children with congenital heart defects. *Journal of Developmental and Behavioral Pediatrics*, 32, 526-532.

Brandlistuen, R. E., Stene-Larsen, K., Holmstrøm, H., Landolt, M. A., Eskedal, L. T. & Vollrath, M. E. I. (2011b). Symptoms of communication and social impairment in toddlers with congenital heart defects. *Child: Care, Health and Development*, 37, 37-43. doi: 10.1111/j.1365-2214.2010.01148.x

Bricker, D. & Squires, J. (1999). *Ages and Stages Questionnaires: A parent-completed, child monitoring system*. Baltimore: Brooks Publishing Co.

Henriksen, C. K., Haugholt, K., Lindgren, M., Aurvåg, A. K., Rønnestad, A., Grønn, M.,...Drevon, C. A. (2008). Improved cognitive development among preterm infants attributable to early supplementation of human milk with docosahexaenoic acid and arachidonic acid. *Pediatrics*, 121, 1137-1145. doi: 10.1542/peds.2007-1511

Janson, H. (2003). Influences on participation rate in a national Norwegian child development screening questionnaire study. *Acta Paediatrica*, 92, 91-96.

Janson, H. & Smith, L. (2003). Norsk manualsupplement til Ages and Stages Questionnaires. Oslo: RBUP Øst og Sør.

Janson, H. & Squires, J. (2004). Parent-completed developmental screening in a Norwegian population sample: A comparison with US normative data. *Acta Paediatrica*, 93, 1525-1529.

Klamer, A., Lando, A., Pinborg, A. & Greisen, G. (2005). Ages and Stages Questionnaire used to measure cognitive deficit in children born extremely preterm. *Acta Paediatrica*, 94, 1327-1329.

Richter, J. & Janson, H. (2007). A validation study of the Norwegian version of the Ages and Stages Questionnaires. *Acta Paediatrica*, 96, 748-752.

Rostad, A. M., Nyberg, P. & Sivberg, B. (2008). Predicting developmental deficiencies at the age of four based on data from the first seven months of life. *Infant Mental Health Journal*, 29, 588-608. doi: 10.1002/imhj.20194

Salomonsson, B. & Sandell, R. (2011). A randomized controlled trial of mother-infant psychoanalytic treatment: I. Outcomes on self-report questionnaires and external ratings. *Infant Mental Health Journal, 32*, 207-231. doi: 10.1002/imhj.20291

Salomonsson, B. & Sled, M. (2010). The ages & stages questionnaire: Social-emotional: A validation study of a mother-report questionnaire on a clinical mother-infant sample. *Infant Mental Health Journal, 31*, 412-431.

Schjølberg, S., P., Eadie, P., Zachrisson, H. D., Øyen, A.-S. & Prior, M. (2011). Predicting language development at age 18 months: Data from the

Norwegian mother and child cohort study. *Journal of Developmental and Behavioral Pediatrics, 32*, 375-383. doi: 10.1097/DBP.0b013e31821bd1dd

Squires, J., Bricker, D. & Potter, L. (1997). Revision of a parent-completed developmental screening tool: Ages and Stages Questionnaires. *Journal of Pediatric Psychology, 22*, 313-328. doi: 10.1093/jpepsy/22.3.313

Westerberg, A. C., R., Schei, Henriksen, C., Smith, L., Drevon, C. A. & Iversen, P. O. (2011). Attention among very low birth weight infants following early supplementation with docosahexaenoic and arachidonic acid. *Acta Pædiatrica, 100*, 47-52.