

Sammenligning av WISC-R- og WISC-III-skårer i et klinisk utvalg: Svakheter i WISC-III-normer eller foreldete WISC-R-normer?

Det er viktig at svakheter ved den tredje versjonen av Weschlers evnestest for barn kartlegges systematisk, slik at norske fagfolk tolker testresultatene på en forsvarlig måte. Denne studien stiller spørsmålet om normene er for strenge for ungdommer med lavt evnenivå.

TEKST

Hilde Sundberg

Jens Egeland

Tor Herman Andreassen

Odd Stensli

PUBLISERT 1. mai 2006

ABSTRACT:

Comparison of WISC-R- and WISC-III-scores in a clinical sample: Are WISC-III norms flawed, or are WISC-R norms outdated?

Whether the Norwegian WISC-III translation has the same psychometric properties as the Swedish version, or whether it measures intellectual function differently from the WISC-R, seems unclear. This study compares the performance of 50 subjects on both tests. Significantly lower IQs were calculated using the WISC-III. Older children obtained a particularly poor WISC-III PIQ. Diminished VIQs were calculated for children with a score above 85. WISC-R to WISC-III conversion has serious implications for the classification of mild mental retardation. According to the new test, over 40 % of the child psychiatric population would have a comorbid MMR. This expansion of the diagnosis may lead to a greater focus on cognitive function. It might also augment false positive classifications of mental retardation.

EMNER

WISC-III

Normer

svakheter

Artikkelen er skrevet som en del av WISC-R/WISC-III-prosjektet i Vestfold, og støttet med bevilgning til forfattergruppen fra R-BUP.

Innledning

Den norske versjonen av Wechsler Intelligence Scale for Children – Third Edition (Ellertsen & Johnsen, 2003) ble utgitt i 2003, og erstattet forgjengeren Wechsler Intelligence Scale for Children-Revised, norsk utgave (Undheim, 1978). WISC-III er en

renormering, oppdatering og utvidelse av WISC-R. Mens WISC-R er normert i Norge, baserer «den norske» WISC-III seg på svenske normer og tilhørende reliabilitets- og validitetsstudier fra 1998 (Sonnander, Ramund & Smedler, 1999). Den svenske versjonen tok utgangspunkt i britisk testmaterieell da verbalprøvene skulle tilpasses svenske forhold, mens den norske er en oversettelse fra svensk og britisk utgave. Kun nye verballedd er testet ut blant norske barn, for å finne rett vanskegrad innen delprøvene. Vi vet altså ikke hvordan den norske WISC-III samlet faller ut i et representativt utvalg norske barn, noe som er en svakhet ved testen, og en kilde til uro i fagfeltet (Strand, 2005). Bjørn Ellertsen, som er medansvarlig for den norske utgaven, stiller i intervju med Tidsskriftet (Strand, 2005) spørsmålet om WISC-III-normene er for strenge for de eldste barna og de med lavt evnenivå. Usikkerheten rundt kvaliteten på normene gjør at mange foretrekker WISC-R, som anses som en mer pålitelig evnetest. Nilsen (2005), fra distributøren Assessio Norge AS, bekrefter at WISC-III har svakheter ved at den ikke differensierer testtakere som skårer to standardavvik eller mer fra gjennomsnittet. Dessuten gir den ofte for lave skårer i de eldste aldersgruppene. Han begrunner dette med at for det første gir et større antall ledd for de eldste barna flere muligheter for tilfeldige feil. For det andre gir økt vektlegging av tidsbonus også ofte for lave skårer hos de eldste.

I denne studien undersøker vi blant annet om det er empirisk grunnlag for skepsisen fra fagfeltet når det gjelder om WISC-III-normene er for strenge for barn med lavt IQ-nivå. En alternativ forklaring er at erfarne klinikere er blitt vant til «snille» WISC-R-normer, der egentlige gjennomsnittsprestasjoner fremstår som særskilt sterke, og barn med lærevansker sjelden skårer under testens normalområde.

At befolkningen gjennomsnittlig stadig presterer bedre på intelligenstagter, kalles «Flynn-effekten». Flynn (1984) beregnet økningen i IQ-poeng til omtrent tre for hvert tiende år. Et barn som retestes med en revidert evnetest, vil derfor typisk prestere svakere enn ved tidligere testing. Vi vet lite om hvordan «Flynn-effekten» slår ut for ulike kliniske grupper, men når WISC-III nå har erstattet WISC-R, forventer vi at barna vi møter, vil skåre lavere på WISC-III enn på WISC-R. De 20 årene som er gått siden normeringen av WISC-R i Norge, tilsier en Flynn-effekt på omtrent seks poeng, og gir altså en forventning om rundt seks poeng gjennomsnittlig reduksjon fra WISC-R til WISC-III. Vi kjenner ikke til norske eller svenske studier som har sammenlignet WISC-R med WISC-III i samme utvalg, men i USA finnes en rekke slike studier. I en metaanalyse av 26 studier fant Zimmerman og Woo-Sam (1997) en gjennomsnittlig nedgang på omtrent seks IQ-poeng mellom WISC-R og WISC-III. Variasjonen studiene imellom var stor – helt fra 0.4 til 13.6 poeng. Endringen i verbal- og utførings-skårene mellom de to prøvene var på samme nivå (henholdsvis 5.3 og 5.9 poeng). Rangkorrelasjonen mellom totalskårene, verbalskårene og utføringskårene var som oftest høy. For total IQ-skår var den gjennomsnittlig på .79, men også denne varierte en del. Høy korrelasjon tyder på høy begrepsmessig overlapping mellom de to amerikanske WISC-versjonene.

Ulike utvalg viser noe ulike reduksjoner ved retesting med WISC-R/WISC-III. I Wechslers standardiseringsutvalg (1991) var forskjellen i total-, verbal- og utføringskala

henholdsvis 5.3, 2.4 og 7.4 poeng. Eventuelle øvelseeffekter var kontrollert for i rotert eller balansert design, der det ble vekslet på hvilken test som ble gitt først. Resultatet ga en forventning om en reduksjon på omtrent fem poeng mellom de to testene, med størst reduksjon i utføringsskalaen. Sevier, Bain og Hildman (1994) fant metastudiens største reduksjoner fra WISC-R til WISC-III i et utvalg av særlig evnerike barn. Totalreduksjonen var mer enn dobbelt så stor som i normeringsutvalget, og profilen var motsatt, nemlig størst reduksjon i verbalskalaen. I en longitudinell undersøkelse av særlig evnerike barn fant Wechsler (1991) samme type profil, men en liten total reduksjon. I likhet med dette finner Sabatino og Spangler (1995) i en studie med balansert design små forskjeller mellom testversjonene på alle IQ-mål i et utvalg av særskilt evnerike barn. En forklaring på at store reduksjoner kan finnes i enkelte studier av særskilt evnerike barn, kan være at det er statistiske artefakter i form av regresjon mot gjennomsnittet.

Når det gjelder studier av barn med lærevansker, finner noen to til tre poengs større reduksjoner enn hos normalbegavede barn, og større reduksjoner i utføringsskalaen enn i verbalskalaen (Bolen, Aichinger, Hall & Webster, 1995; Lyon, 1995; Slate, 1995; Wechsler, 1991). Andre studier, derimot, finner ingen slik differensiell reduksjon hos læresvake barn (Gunter, Sapp & Green, 1995; Sapp, Abbot & Hinkley, 1997). I disse studiene er utvalgene ofte lite beskrevet, og omfatter gjerne hele den varierte gruppa elever som får spesialtilpasset opplæringstilbud i skolen. I en studie av barn med utviklingshemning var reduksjonene også noe større enn forventningene ut fra normgruppen (Wechsler, 1991). Sapp et al. (1997) derimot fant ingen økt reduksjon i skårer mellom de to WISC-versjonene i et liknende utvalg. De økte reduksjoner hos barn med lærevansker og utviklingshemning som noen studier finner, er det motsatte av hva vi ville forvente ved regresjon mot gjennomsnittet. De fleste av disse studiene er basert på rutinemessige retestinger som gjøres hvert tredje år, slik at tidsintervallet dermed er forholdsvis stort. Lavere skårer på WISC-III kan dermed gjenspeile at disse sårbare barna vokser seg inn i ytterligere funksjonshemning utover i ungdomsalderen, heller enn at WISC-III har spesifikt strengere normer enn WISC-R for svaktfungerende barn. En studie av Stavrou (1990) støtter denne hypotesen. Han fant en kontinuerlig nedgang i IQ hos læresvake barn ved gjentatt testing med WISC-R over en seksårsperiode.

Andre typer utvalg som er undersøkt, er barn med språkvansker eller dysleksi. Også her spriker resultatene mellom ulike studier. Ackerman, Weir, Holloway og Dykman (1995) fant en reduksjon på hele 12 poeng, mens andre har funnet reduksjoner på et gjennomsnittlig nivå (Doll & Boren, 1993; Newby, Recht, Caldwell & Schaeffer, 1993). Når det gjelder studier med blandede kliniske utvalg, finner to store undersøkelser med henholdsvis 104 og 96 doble protokoller ingen avvik i IQ-reduksjonene, som ligger rundt fem til seks poeng (Sapp et al., 1997; Wechsler, 1991). Reduksjonene i verbal- og utføringsskalaene er jevne. Den ene studien benyttet rotert design, og hadde et utvalg av ulike lære- eller lesevansker, ADHD og emosjonelle forstyrrelser. Det andre utvalget bestod av barn med lærevansker, utviklingshemning og atferdsvansker, og studien hadde en longitudinell design, med et intervall på tre år. Det var altså en svært stor

overensstemmelse i resultatene til tross for ulik design. Dette kan bety at verken øvingseffekter eller en øket funksjonshemming over tid i forhold til normgruppen gjør seg særlig gjeldende for denne type utvalg, selv når intervallet er tre år.

Problemstillinger

Vi ønsket å undersøke hvordan den norske oversettelsen av den svenske versjonen av WISC-III faller ut i et blandet klinisk utvalg der sannsynligheten for å bli testet med WISC-III er størst. Vi undersøkte resultatene i et blandet klinisk utvalg av barn fra PP-tjeneste, habiliteringstjeneste og psykisk helsevern, i forhold til følgende spørsmål: 1) Hvor stor er endringen i IQ-skårer fra WISC-R til WISC-III? 2) Er det noen differensiell effekt av alder eller initialt IQ-nivå på denne endringen? Nærmere spesifisert, vil endringen i IQ-skårer være større for de eldste barna, og større for barn med lærevansker, enn for barn som intellektuelt presterte nær middelverdien?

Metode

Datamaterialet er hentet fra egne arkiver og anonymisert. Til sammen 50 doble protokoller ble funnet. I alle tilfeller er WISC-R administrert før WISC-III. 33 var gutter og 17 jenter. Gjennomsnittsalderen ved WISC-III var 11,8 år, fra 8 år 6 måneder til 16 år 1 måned (S.A. = 2,1 år). Den svenske utgaven av WISC-III er kun normert opp til 16 år, men en protokoll med 16,1 års alder ble likevel inkludert, da forskjellen på 30 dager ble vurdert som liten og uten betydning for resultatene. Det er gjennomsnittlig 1 år og 9 måneder mellom administreringen av testene. Testintervallet varierer fra 3 til 43 måneder, og 47 saker faller i intervallet 6 til 36 måneder.

Barna var alle henvist til Habiliteringstjenesten i Vestfold, Barne- og Ungdomspsykiatrisk Poliklinikk, PP-tjenesten eller nevropsykologisk privatpraksis i Vestfold. WISC-III og WISC-R var administrert av over 35 ulike psykologer og pedagoger i PP-tjenester, Barne- og Ungdomspsykiatrisk poliklinikk, Habiliteringstjenesten, nevropsykologisk privatpraksis, og i ett tilfelle ved et statlig kompetansesenter. WISC-R var i 48 saker tatt i PPT, i ett tilfelle i BUPA, og ved en anledning i et statlig kompetansesenter. WISC-III var i 26 tilfeller tatt i BUPA, 17 i Habiliteringstjenesten, fem i PPT og to i privat nevropsykologisk praksis.

Av de 50 barna hadde 44 barn diagnoser. Ni barn hadde to diagnoser. Seksten barn hadde diagnoser innen hyperkinesi-spekteret. Ti barn hadde fått diagnosen psykisk utviklingshemning. Åtte barn hadde diagnoser innen autismespekteret, fem hadde spesifikke språkvansker, fire generelle lærevansker, tre belastnings- og tilpasningsrelaterte forstyrrelser, to lese- og skrivevansker, to ikke-verbale lærevansker, en alvorlig somatisk sykdom, en uspesifisert forstyrrelse av psykologisk utvikling og en tilknytningsforstyrrelse.

Dataanalyser

Endring i total IQ (TIQ), verbal IQ (VIQ), og utførings-IQ (UIQ) ble analysert med t-test for parede utvalg. For å analysere alderseffekter ble utvalget delt i tre omtrent like store

grupper etter alder ved WISC-III; fra 8 til 10.4 år (n = 16), 10.5–12.9 år (n = 17) og fra 13 til 16 år (n = 17). Begrunnelsen for aldersinndelingen var å lage store nok grupper til å fange opp signifikante forskjeller i datamaterialet, og samtidig skille ut en gruppe for de eldste barna for å undersøke problemstillingen vedrørende dem. Gruppeforskjellene i TIQ, VIQ og UIQ ble analysert med variansanalyse (ANOVA), med posthoc-analyse (LSD). Mulig effekt av forskjeller i tidsintervall mellom retestingene innen de tre aldersgruppene ble kontrollert for i en kovariansanalyse (ANCOVA). For å analysere eventuelle forskjeller ut fra initialt IQ-nivå ble utvalget delt inn to, TIQ < 85 poeng (n = 30), og TIQ = 85 poeng (n = 20). Disse gruppene ble også sammenlignet ved hjelp av variansanalyse med $p < .05$.

Resultater

Gjennomsnittlig Total IQ (TIQ), Verbal IQ (VIQ), Utførings-IQ (UIQ) og gjennomsnittlig skalerte skårer på alle delprøver på WISC-R og WISC-III fremkommer av Tabell 1. TIQ for WISC-R ligger omtrent midt i «nedre normalområde» (80–89 poeng) med 84,3 poeng, mens den er redusert med 9.9 poeng på WISC-III og ligger da midt i «lavt område» (70–79 poeng) med 74.3 poeng. Alle tre IQ-verdiene er statistisk signifikant redusert fra WISC-R til WISC-III ($p < .001$). Reduksjonen i VIQ på 10.8 poeng, er statistisk signifikant større enn reduksjonen i UIQ på 6.7 poeng ($t = 2.02, p = .049$). Tolv prosent (seks barn) skåret under IQ 70 på WISC-R, mens 40 % (20 barn) skåret under samme verdi på WISC-III. Ti prosent av barna (5 barn) skåret mer enn 20 IQ-poeng svakere på WISC-III, mens 12 % (seks barn) forbedret sine skårer.

Tabell 1. WISC-R og WISC-III-resultater for et blandet klinisk utvalg (N = 50).

MÅL	WISC-R		WISC-III		DIFFERANSER WISC-R - WISC-III	
Total IQ (TIQ)	84	18	74	19	.86**	- 9.9**
Verbal IQ (VIQ)	87	20	76	18	.81**	- 10.8**
Utførings-IQ (UIQ)	85	15	78	18	.81**	- 6.7**
Bildeutfylling	9	3	8	3	.58**	-1.1*
Informasjon	7	3	6	3	.65**	-.7
Koding	7	3	6	3	.70**	-.9**
Likheter	9	4	7	3	.53**	-2.7**
Tegneserier	8	4	8	4	.56**	.1
Regning	7	4	6	3	.71**	-.7
Terningmønstre	8	4	7	3	.73**	-1.4**

MÅL	WISC-R		WISC-III		DIFFERANSER WISC-R - WISC-III	
Ordforståelse	8	4	7	3	.71**	-1.9**
Puslespill	8	3	7	3	.57**	-.5
Resonnering	9	4	6	3	.53**	-2.5**
Tallhukommelse	7	3	7	2	.71**	.1
Labyrinter (N = 40)	10	3	8	4	.50**	-1.4*
Symbolleting (N = 48)	-	-	6	4	-	-

Aldersgrupper ved WISC-III

Utvalget ble inndelt i tre omtrent like store alderskategorier ved WISC-III, 8–10.4 år (n = 16), 10.5–12.9 år (n = 17) og 13–16 år (n = 17). Gjennomsnitt, gjennomsnittlige differanser og korrelasjoner i hver av de tre aldersgruppene framkommer i Tabell 2. Den eldste gruppa skåret statistisk signifikant svakere enn den yngste med hensyn til WISC-R verbal IQ ($p = .022$). Ingen andre WISC-R-resultater var statistisk signifikant forskjellige mellom de tre aldersgruppene. Ved WISC-III var det signifikante forskjeller mellom aldersgruppene når det gjaldt alle tre IQ-verdier ($p < .012$). IQ-verdiene i de to yngste gruppene er gjennomsnittlig over 70 poeng, mens i den eldste gruppa faller alle WISC-III IQ-verdier under 70 poeng.

Hos den eldste gruppa sees statistisk signifikant større reduksjoner fra WISC-R til WISC-III, når det gjelder TIQ sammenlignet med yngste gruppe, og UIQ sammenlignet med de to yngste gruppene. Når en kontrollerer for forskjellen i retestings-tidsintervall mellom gruppene, viser ANCOVA at UIQ-reduksjonen fortsatt er statistisk signifikant forskjellig ($p = .001$).

Tabell 2. WISC-R- og WISC-III-resultater differensiert på 3 aldersgrupper.

	8-10,4 ÅR (N = 16)			10,5-13,9 ÅR (N = 17)			13-16 ÅR (N = 17)		
Total IQ (TIQ)	89 (18)	83 (17)	- 6.3*	86 (20)	77 (20)	- 9.1**	78 (14)	64 (14)	- 14.3**
Verbal IQ (VIQ)	94 (17)	84 (16)	- 9.4**	90 (22)	78 (18)	-11.5**	78 (16)	67 (15)	- 11.5**
Utførings-IQ (UIQ)	87 (19)	85 (16)	-1.5	86 (13)	82 (19)	-4.2	82 (15)	68 (15)	-14.2**

Kategorier etter IQ-nivå ved WISC-R

For å analysere eventuelle forskjeller i reduksjoner i IQ-verdier fra WISC-R til WISC-III avhengig av initialt IQ-nivå ble utvalget splittet i to ut fra WISC-R IQ over/under 85 poeng. Tabell 3 viser resultatene. Det ble funnet en statistisk signifikant større nedgang i Verbal IQ for de med WISC-R skår > 85 poeng.

Tabell 3. WISC-R- og WISC-III-differanser ut fra initial IQ over/under 85.

	WISC-R TIQ < 85 (N = 30) M (S.D.)	WISC-R TIQ > 85 (N = 20) M (S.D.)
Differanse TIQ	8.6 (10.6)	12.0 (7.6)
Differanse VIQ **	7.7 (11.8)	15.5 (9.9)
Differanse UIQ	7.7 (11.3)	5.3 (9.0)

Korrelasjoner

Korrelasjonsanalyser ble utført for å finne ut i hvilken grad den innbyrdes rangeringen innen gruppene var den samme ved WISC-III som ved WISC-R, og dermed få mål på begrepsvaliditeten til WISC-III. Korrelasjonene mellom de ulike IQ-målene framkommer av Tabell 4. Alle korrelasjonene er statistisk signifikante.

Tabell 4. Korrelasjoner mellom WISC-R og WISC-III.

	WISC-R TIQ	WISC-R VIQ	WISC-R UIQ	WISC-III TIQ	WISC-III VIQ
WISC-R TIQ					
WISC-R VIQ	,928(**)				
WISC-R UIQ	,852(**)	,598(**)			
WISC-III TIQ	,861(**)	,838(**)	,684(**)		
WISC-III VIQ	,752(**)	,810(**)	,489(**)	,944(**)	
WISC-III UIQ	,866(**)	,754(**)	,813(**)	,922(**)	,745(**)

Diskusjon

Undersøkelsen viser som forventet reduserte IQ-verdier fra WISC-R til WISC-III. Nedgangen i total IQ og særlig verbal IQ var imidlertid større enn den beregnede Flynn-effekten skulle tilsi. Som det framkommer i Tabell 2, viser undersøkelsen at de eldste barna har større nedgang mellom testene enn yngre barn. Når det gjelder alder framkommer imidlertid ingen spesifikk effekt av språkprøvene, snarere er det utføringsprøvene som synes særlig vanskelig for de eldste. Skiller man på de best

fungerende barna og barna med lette til moderate lærevansker, får de best fungerende barna større nedgang i verbalskalaen enn de andre. At WISC-III har betydelig strengere normer enn WISC-R, får som konsekvens at andelen barn som presterer under IQ 70, øker fra 12 til 40 %. Samtidig synes den innbyrdes rangeringen av barna i utvalget å være ganske stabil mellom de to testene. Dette framkommer av Tabell 4, hvor rangkorrelasjonene mellom tilsvarende IQ-mål på WISC-R og WISC-III er høye, og på nivå med tilsvarende rangkorrelasjoner mellom henholdsvis verbal-, utførings- og total IQ innenfor WISC-R og WISC-III hver for seg. Begrepsvaliditeten til WISC-III synes altså å være god på IQ-nivå målt ved sammenligning med WISC-R.

Flere forhold kan ha bidratt til at undersøkelsen viste større reduksjoner mellom WISC-R og WISC-III, særlig for total og verbal IQ, enn forventet ut fra amerikanske studier. En mulighet er at barna har blitt relativt sett mer avvikende med alderen. Som nevnt innledningsvis fant Stavrou (1990) reduserte skårer med tiden hos læresvake barn. Når vi har en seriell design i undersøkelsen, er det mulig at slike endringer kan ha skjedd i vårt utvalg. Imidlertid fant de to amerikanske undersøkelsene med blandede kliniske utvalg som er omtalt innledningsvis (Sapp et al., 1997; Wechsler, 1991), reduksjoner med lignende størrelse, selv om den ene studien hadde et intervall som var større enn i denne undersøkelsen, og den andre hadde en balansert design. Dette, i tillegg til at tidsintervallet i den foreliggende undersøkelsen er forholdsvis lite, gjør at sannsynligheten for reelle forandringer i utvalget vurderes som liten.

En annen forklaring på vårt funn med større reduksjoner enn forventet særlig i verbal-skalaen kan være at WISC-III-normene i verbalprøver er for strenge. Tabell 1 viser at de tre mest reduserte delprøvene (likheter, resonnering og ordforståelse) alle er verbale, og dette gir en stor samlet verbal reduksjon. Disse tre delprøvene ligger høyt i forhold til snittet av delprøver på WISC-R og noe under snittet ved WISC-III. Dette kan tyde på at den store verbale reduksjonen mest skyldes foreldete WISC-R-normer, men også i noen grad for strenge normer på WISC-III. Når vi samtidig finner at resonnering og likheter har lavest rangkorrelasjon av alle delprøver, kun etterfulgt av labyrinter, tyder dette på at disse verbalprøvene nå måler noe litt annet i WISC-III enn de gjorde i WISC-R.

Mens verbalreduksjonen er jevnt stor i de tre ulike aldersgruppene, er reduksjonen i utføringsskår størst hos de eldste. Dette stemmer med fagfolks erfaringer og Nilsens påpekning (1995) om at WISC-III-normene ofte gir feilaktig lave skårer for de eldste barna, på grunn av økt antall ledd i delprøvene og høyere tempokrav.

Undersøkelsen tyder på at det ikke er for de laveste evnenivåene at IQ-reduksjonene er størst, men i gruppa over 85 poeng. Forskjellen ligger her i verbalskalaen, ikke i utføringsskalaen. Påstanden om at WISC-III oftere gir feilaktig lave skårer for barn med lave IQ-nivåer enn WISC-R, blir altså ikke støttet av denne studien.

Med innføring av WISC-III, som har strengere normer enn WISC-R, vil et større antall barn enn tidligere skåre under 70 poeng, slik også vår studie viser. Dette betyr at mens vi tidligere nærmest betraktet lett psykisk utviklingshemning som en differensialdiagnose til psykisk lidelse, vil vi nå måtte tenke at en stor gruppe har

komorbide lærevansker. Overgangen til WISC-III kan bidra til en berettiget øket vektlegging av kognitiv fungering i psykologisk og psykiatrisk behandling. Siden WISC-R de siste årene har hatt foreldete normer har vi nå et «diagnostisk etterslep» der barn og unge med lærevansker og lett utviklingshemning ikke i tilstrekkelig grad har blitt fanget opp og forstått, med mangelfull tilrettelegging i skole og hjem som sannsynlig konsekvens. Personer med utviklingshemning som ikke har blitt oppdaget, har heller ikke kunnet nyte godt av rettigheter i trykkesystemet og kommunale tjenester. Mangelfull tilpasning øker risikoen for sekundære problemer som uro, angst og depresjon, og kan føre til henvisning til hjelpeapparatet på grunn av psykiske og emosjonelle vansker. Ved innføring av WISC-III er imidlertid faren for overdiagnostisering også aktuell, særlig ved at WISC-III-normene kan være for strenge for de eldste barna.

Denne studiens styrke ligger i typen utvalg, i den forstand at det er representativt for de barna som undersøkes med WISC i Norge. Klinikeren som har vært vant til å bruke WISC-R, kan ut fra denne studien bedre fortolke hva som representerer stabilitet i intellektuell funksjon, og hvilke testmessige endringer på individnivå som må forstås som reell forandring. Studien har et stort mangfold av administratorer og behandlingstyper, noe som er en begrensning for en eksperimentell design, men som reflekterer den praksis som finnes og den virkeligheten praktikere må forholde seg til i sin tolkning av testskårene. Den longitudinelle designen åpner for mulige øvingeffekter som vi ikke har kontrollert for. Øvingeffektene er likevel trolig minimale gitt tidsintervallet mellom WISC-R og WISC-III. Hvis øvingeffekter likevel har preget resultatene, har dette eventuelt gitt mindre reduksjoner enn ellers, og da antagelig mest på utføringskalaen, som antas mest påvirkelig for øving (Kaufman, 1994). Funnet med store reduksjoner i verbalskåre står altså enda sterkere hvis studien er påvirket av øvingeffekter.

Ved overgangen WISC-R/WISC-III vil klinikere kunne møte en større andel barn og unge enn ellers som er henvist på grunn av psykiske og emosjonelle vansker, hvor en må vurdere om vanskene er utviklet gjennom kombinasjonen lave kognitive forutsetninger og for høye forventninger fra omgivelsene. Som kliniker bør en være oppmerksom på at de store reduksjonene i IQ-verdier mellom WISC-R og WISC-III kan skyldes både svakheter i WISC-III-normer – særlig utføringsprøvene for de eldste barna, og foreldete WISC-R-normer – spesielt innen verbalskalaen. Reduksjonen kan videre skyldes forandringer i testenens innhold eller oversettelse og administreringsfeil, og altså ikke nødvendigvis at barnet har svekket kognitiv fungering i forhold til barn på samme alder. WISC-resultater må som før vurderes i lys av annen diagnostisk informasjon og ut fra en helhetsvurdering, og reduserte total IQ-skårer på 10–20 IQ-poeng fra WISC-R til WISC-III vil ikke være uvanlig. Manglende norske WISC-III-normer kompliserer tolkningen av testresultatene og skaper flere feilkilder enn nødvendig. Det mest brukte utredningsverktøyet for barn og unge – som skal hjelpe oss til å finne gode tiltak for deres videre utvikling – bør kvalitetssikres så godt som mulig. Det er viktig at svakheter ved WISC-III kartlegges systematisk, slik at norske fagfolk kan få mer

utfyllende informasjon om hvordan WISC-III kan tolkes på en forsvarlig måte. At WISC-III normeres på norsk, er også en viktig forutsetning for kvalitativ god forskning videre på feltet i Norge.

Hilde Sundberg

Nordgata 29

3295 Helgeroa

Tlf 33 48 88 00

E-post hilde.sundberg@piv.no

Teksten sto på trykk første gang i Tidsskrift for Norsk psykologforening, Vol 43, nummer 5, 2006, side 476-481

TEKST

Hilde Sundberg

Jens Egeland, Klinikk for Psykisk Helse og Rusbehandling, Sykehuset i Vestfold, Psykologisk institutt, Universitetet i Oslo

KONTAKT: jens.egeland@siv.no

Tor Herman Andreassen

Odd Stensli

+ [Vis referanser](#)

Referanser

Ackerman, P. T., Weir, N. L., Holloway, C. A., & Dykman, R. A. (1995). Adolescents earlier diagnosed as dyslexic show major IQ declines on the WISC-III. *Reading and Writing*, 7, 163-170.

Bolen, L. M., Aichinger, K. S., Hall, C. W., & Webster, R. E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. *Journal of Clinical Psychology*, 51, 89-95.

Doll, B., & Doren, R. (1993). Performance of severely language-impaired students on the WISC-III, language scales, and academic achievement measures. *Journal of Psycho-educational Assessment*, 81, 77-86.

Ellertsen, B., & Johnsen, I. M. B. (2003). WISC-III. Wechsler Intelligence Scale for Children-Third Edition. Norsk versjon. Administrasjonsveiledning. Oslo: Assesio Norge AS.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51.

Gunter, C. M., Sapp, S. G., & Green, A. (1995). Comparison of scores on WISC-III and WISC-R of urban disabled students. *Psychological Reports*, 77, 473-474.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: John Wiley & Sons.

Lyon, M. A. (1995). A comparison between WISC-III and WISC-R scores for learning disabilities reevaluations. *Journal of Learning Disabilities*, 28, 4, 253-255.

Newby, R. F., Recht, D. R., Caldwell, J., & Schaeffer, J. (1993). Comparison of WISC-III and WISC-R IQ changes over a 2-year time span in a sample of children with dyslexia. *Journal of Psychoeducational Assessment*, 81, 87-93.

Nilsen, D. Ø. E. (1995). WISC-III og normer. *Tidsskrift for Norsk Psykologforening*, 42, 532-533.

- Sabatino, D. A., & Spangler, R. S. (1995). The relationship between the Wechsler Intelligence Scale for Children-Revised and the Wechsler Intelligence Scale for Children-III scales and subtests with gifted children. *Psychology in the Schools*, 32, 18-23.
- Sapp, G. L., Abbot, G., Hinkley, R., & Rowell, A. (1997). Examination of the validity of the WISC-III with urban exceptional students. *Psychological Reports*, 81, 1163-1168.
- Sevier, R. C, Bain, S. K., & Hildman, L. K. (1994). Comparison of WISC-R and WISC-III for gifted children. *Roeper Review*, 17, 39-42.
- Slate, J. R. (1995). Two investigations of the validity of the WISC-III. *Psychological Reports*, 76, 299-306.
- Sonnander, K., Ramund, B., & Smedler, A. (1999). Wechsler Intelligence Scale for Children - Third Edition. Svensk versjon. Manual. Stockholm: Psykologiforlaget AB.
- Stavrou, E. (1990). The long-term stability of WISC-R scores in mildly retarded and learning-disabled children. *Psychology in the Schools*, 27, 101-111.
- Strand, N. (2005). WISC-III skaper uro. *Tidsskrift for Norsk Psykologforening*, 42, 519-521.
- Zimmerman, I. L., & Woo-Sam, J. M. (1997). Review of the criterion-related validity of the WISC-III: the first five years. *Perceptual and Motor Skills*, 85, 531-546.
- Undheim, J. O. (1978). Håndbok. WISC-R. Wechsler Intelligence Scale for Children - Revised. Norsk utgave. Trondheim: Psykologisk Institutt.
- Wechsler, D. (1991). Wechsler Intelligence Scale for Children - III. San Antonio: Psychological Corp.