

# California Verbal Learning Test-II (CVLT-II) utprøvd i et klinisk utvalg i Norge: Psykometriske egenskaper og krysskulturell validitet

Denne studien tar for seg en ny norsk utgave av en av de sentrale metodene for å undersøke hukommelse hos voksne. Har CVLT-II tilsvarende psykometriske egenskaper som originalen?

TEKST

Ole Bosnes

PUBLISERT 1. juli 2007

ABSTRACT:

## The Norwegian version of the California Verbal Learning Test - II (CVLT-II): Psychometric characteristics and crosscultural validity

The California Verbal Learning Test (CVLT, 1987) has been widely used both in research and clinical settings. Due to doubts about how representative the norms are, a second version of the California Verbal Learning Test (CVLT-II, 2000) was published. Studies comparing the psychometric properties of a Norwegian translation (2004) to the original version are lacking. This study reports results of 128 patients tested with the Norwegian version of CVLT-II, comparing the results with comparative data reported in the original manual. Factor analyses of the Norwegian sample resulted in six factors, compared to five factors in the American sample, but apart from this difference, factor structure and loadings corresponded. The study indicates that the Norwegian version of CVLT-II has retained the most important assets of the original version. Test results from the translated CVLT-II were also compared to results from the translated Wechsler Memory Scale-Revised (1987), showing nearly identical results when data were corrected for the «Flynn effect».

Keywords: memory, assessment, The California Verbal Learning Test, crosscultural validity

EMNER

Hukommelse

vurdering

The California Verbal Learning Test

tverrkulturell validitet

## Innledning

Kognitiv svikt, særlig sviktende hukommelse, har i de siste år i økende grad kommet i fokus både hos fagfolk og politikere, både internasjonalt og i Norge. I tråd med dette har man ved ulike sykehus opprettet såkalte «hukommelsesklinikker», hvor den kanskje mest kjente her i landet ble bygd opp ved Ullevål sykehus på 1980-tallet. Parallelt har det blitt utviklet en rekke metoder for å avdekke eventuell hukommelsessvikt. Dette har sammenheng med at man har fått medikamenter som i noen grad hevdes å kunne bremse alvorlig og progressiv hukommelsessvikt – demensutvikling. Det antas at jo tidligere man starter behandling, desto

bedre er behandlingseffekten. Tidlig diagnostikk og behandling har følgelig blitt sentralt, og anstrengelsene for å utvikle reliable og valide diagnostiske metoder har økt tilsvarende.

Innenfor klinisk nevropsykologi har man sett en tilsvarende økende interesse for metoder for å undersøke hukommelse hos voksne/eldre. Tester som hyppig benyttes i klinisk virksomhet i dag, har sin forhistorie tidlig i forrige århundre. Dette gjelder blant annet Wechsler Memory Scale fra 1945 og California Verbal Learning Test fra 1987, som har sine forløpere i tester utviklet av henholdsvis Binet og Simon (1916) og Claparede (1919). De siste utgavene av disse to testene er California Verbal Learning Test-II (CVLT-II) fra 2000 og Wechsler Memory Scale-III (WMS-III) fra 1997. Kun CVLT-II foreligger i norsk oversettelse ved Lundervold og Sundet (2004), slik at man i klinisk bruk av Wechslers test foreløpig er henvist til å bruke WMS-R, som foreligger i norsk oversettelse ved Gimse (1992).

I omtalen av den første utgaven av California Verbal Learning Test (CVLT-1) refererer Spreen og Strauss (1998) forskning som har reist tvil om testens normer er representative for den amerikanske befolkning, og som har hevdet at normene er for strenge (bl.a. Elwood, 1995; Randolph, Gold, Kozora, Hermann & Wyler, 1994; Stallings, Boake & Sherer, 1995). Dette var noe av bakgrunnen for at The Psychological Corporation besluttet å foreta en ny normering med et langt større og mer representativt utvalg. Samtidig ble det også foretatt endringer i instruksjoner og innført enkelte nye måter å analysere data fra testen på, i tråd med erfaringer med den første utgaven.

Bosnes og Ellertsen (2000) sammenlignet prestasjoner på CVLT-1 og WMS-R i et norsk klinisk utvalg og fant bedre prestasjoner på WMS-R enn på CVLT-1, tilsvarende det blant annet Randolph et al. (1994) rapporterte fra USA. De psykometriske egenskaper ved CVLT-1 ble også undersøkt, og for en gruppe pasienter med multippel sklerose fant man en faktorstruktur og faktorladninger nesten identiske med det som var oppgitt for et tilsvarende amerikansk pasientutvalg.

Egeland et al. (2005) sammenlignet skårer på CVLT-1 i et norsk normalutvalg med de opprinnelige normene og med normene i CVLT-II. De fant at normene i CVLT-II trolig var for «snille» i Norge, og at de opprinnelige normene i CVLT-1 passet bedre. De fant samtidig at normene for kvinner og menn slo noe forskjellig ut, ved at normene for kvinner kunne være for strenge i Norge.

## **Problemstilling**

CVLT-II tilbys nå norske psykologer av Harcourt Assessment Sweden AB, og ut fra antall internasjonale publikasjoner som har benyttet CVLT-1 (se oversikt hos Lezak, 2004; Mitrushina, Boone, Razani & D'Elia, 2005), er det grunn til å tro at også CVLT-II vil bli benyttet i Norge, både i forskning og klinisk praksis. Ved gjennomgang av den norske manualen og ved litteratursøk ble det ikke funnet publikasjoner som har undersøkt psykometriske egenskaper, blant annet faktorstruktur og faktorladninger, i den norske versjonen av CVLT-II. Man vet følgelig ikke om oversettelsen har medført endringer i testens psykometriske egenskaper. Vi fant det derfor av interesse å undersøke prestasjoner på CVLT-II hos en norsk blandet pasientgruppe, for å undersøke om de psykometriske egenskaper som er rapportert fra et blandet klinisk utvalg i USA, også gjelder i Norge.

Ut fra tidligere erfaring med og forskning på CVLT-1 og WMS-R, som stadig benyttes mye, fant vi det også av interesse å undersøke om man finner den samme normative prestasjonsforskjell

mellom CVLT-II og WMS-R som vi fant mellom CVLT-1 og WMS-R.

Problemstillingene for studien var:

1. Viser den norske oversettelsen av CVL-II tilsvarende psykometriske egenskaper som originalen?
2. Hvordan er det normmessige samsvaret mellom CVLT-II og WMS-R?

## Metode

For å besvare spørsmål 1 benyttet man testdata fra 128 pasienter som i forbindelse med omfattende nevropsykologisk undersøkelse ved Sykehuset Namsos hadde gjennomgått CVLT-II.

Tabell 1 viser gjennomsnittlig alder, utdanningsnivå og WAIS Total IQ, samt kjønnsfordeling for pasientgruppen. Tabellen viser at de undersøkte mennene er noe yngre og har lavere utdanningsnivå og Total IQ enn kvinnene, men forskjellene mellom kvinners og menns alder, utdanning og IQ er ikke signifikante (t-test for uavhengige utvalg)

**TABELL 1. DELTAKERNES ALDER, UTDANNING OG WAIS TOTAL IQ**

	<b>Antall</b>	<b>Alder</b>	<b>Utdanning</b>	<b>Total IQ/SD</b>
Kvinner	62	41.8	11.74	89/14.3
Menn	66	39	10.87	87/15.4

n = 128

Tabell 2 viser pasientgruppens diagnostiske fordeling. Gruppen består både av pasienter med nevrologisk sykdom, psykisk lidelse og utviklingsforstyrrelse, og kan følgelig med rette betegnes som en blandet klinisk gruppe.

**TABELL 2. DIAGNOSEFORDELING FOR DELTAKERE**

<b>Demens</b>	<b>MCI</b>	<b>HI</b>	<b>Slag</b>	<b>ADHD</b>	<b>PU</b>	<b>LD</b>	<b>Dep.</b>	<b>Schiz.</b>	<b>Rus</b>	<b>PD</b>	<b>Angst</b>	<b>Spisef.</b>	<b>Somatof.</b>
8	24	5	3	3	8	4	11	2	7	19	21	1	12

n = 128  
(Demens = demens, MCI = mild kognitiv svikt, HI = hodeskade, Slag = hjerneslag, ADHD = AD/HD, PU = psykisk utviklingshemning, LD = lærevansker, PD = personlighetsforstyrrelse, Spisef. = spiseforstyrrelse, Somatof. = somatoform lidelse)

For å besvare spørsmål 2 benyttet man testdata fra 100 pasienter som hadde gjennomgått både CVLT-II og WMS-R i forbindelse med omfattende nevropsykologisk undersøkelse ved Sykehuset Namsos. (Disse pasientene inngår også i datagrunnlaget for hypotese 1.) For disse pasientene valgte man på et bestemt tidspunkt å endre rekkefølgen testene ble gitt i, slik at man fikk et tilnærmet «counter-balanced design» og dermed mindre sannsynlighet for en

eventuell skjevhet på grunn av rekkefølgen testene var gitt i. I alt 49 pasienter gjennomgikk WMS-R først og 51 CVLT-II først.

Tabell 3 viser gjennomsnittlig alder, utdanning og WAIS Total IQ for gruppen. Det fremgår av tabellen at kvinnene var noe eldre og hadde bedre utdanning og høyere IQ enn mennene, men t-testing viste ingen signifikante forskjeller. Alle pasienter ble testet av erfaren nevropsykolog, eller av erfaren testtekniker under veiledning av nevropsykolog. Prosjektet var godkjent av Regional Etisk Komite (REK).

**TABELL 3. DELTAKERNES ALDER, UTDANNING OG WAIS TOTAL IQ**

	<b>Antall</b>	<b>Alder</b>	<b>Utdanning</b>	<b>Total IQ/SD</b>
Kvinner	37	45.54 / 16.8	12.01 / 3.4	92.54 / 15.77
Menn	63	38.13 / 15.4	10.42 / 2.3	85.94 / 15.22

n = 100

### **Beskrivelse av CVLT-II og WMS-R**

Både CVLT-1 og CVLT-II bygger på Rey Auditory Verbal Learning Test fra 1956 og består av en liste ord (Liste A) som skal gjentas av testpersonen, etter at den har blitt lest høyt av testlederen. Samme prosedyre gjentas fem ganger og etterfølges så av en ny liste med ord (Liste B), hvor testpersonen skal gjenta så mange ord som mulig fra denne nye listen etter at den er lest høyt. Man ber så personen gjengi så mange ord hun/han kan huske fra den første listen, først uten og deretter med stikkord («cuing»). Dette gjentas så etter 20 minutter og etterfølges av et forsøk med gjenkjenning og etter ytterligere ti minutter av et forsøk hvor personen skal angi hvilket av to ord som var med i den første listen («forced choice»). For mer detaljert beskrivelse av testen og prosedyrer vises til manualen (2000).

WMS-R består av 12 deltester, hvor man har som hensikt å få mål på umiddelbar og utsatt hukommelse med verbalt og visuelt materiale. Testen har også deltester som er ment å gi mål på oppmerksomhet og konsentrasjon. De verbale deltestene består av innlæring av ordpar og historier, med gjengivelse umiddelbart og etter 30 minutter, mens de visuelle deltestene består av umiddelbar gjenkjenning av figurer, samt gjengivelse (tegning) av figurer umiddelbart og etter 30 minutter. For mer detaljert beskrivelse vises til manualen (1987).

### **Resultater**

#### **Faktoranalyse av CVLT-II**

Tabell 4 viser psykometriske egenskaper basert på samme faktoranalyse (prinsippal komponent analyse, varimax-rotering) av resultater på CVLT-II for den norske pasientgruppen (n = 128) på de samme 16 variablene som er oppgitt i manualen for en amerikansk blandet pasientgruppe (n = 128). De amerikanske tallverdiene står i parentes. For oversiktens skyld har vi bare tatt med faktorladninger > .50.

**TABELL 4. FAKTORMATRISSE BASERT PÅ 16 CVLT-II VARIABLER**

TABELL 4. FAKTORMATRISER BASERT PÅ 16 CVLT-II VARIABLER

Variabel	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6
Trial 5 korrekte	.91(.88)					
Fri gjenkalling- Kort intervall	.81(.83)					
Fri gjenkalling- Langt intervall	.81(.81)					
Total læringskurve - Trial 2-5	.75(.79)					
Konsistens gjenkalling - Trial 1-5	.68(.73)					
Total gjenkjenning-diskriminabilitet	.73(.71)					
Fri gjenkalling - total intrusjoner		(-.75)	.88			
Total respons-bias		(.69)				.96
Intrusjoner - gjenkalling med stikkord		(-.68)	.68			
Gjentakelser - Total (alle trials)			.57(.77)			
Trial 1 korrekte		.75	(.71)			
Trial B korrekte		.71	(.53)			
Seriell gruppering forover (chance adjusted) Trial 1-5				(-.91)	.76	
Semantisk gruppering (chance adjusted) Trial 1-5				(.83)	-.76	
% gjenkalling - Primacy region				.91	(.87)	
% gjenkalling - Recency region				-.73	(-.79)	
Forklart varians	27,95%	13,75%	10,25%	9,53%	8,65%	6,63%
n = 128						
Analyse gjort med SSPS versjon 13, Prinsipal komponent analyse, Varimax- rotering med Kaiser Normalisering, Eigenvalues > 1 extracted.						

Tabell 4 viser at mens man i den amerikanske gruppen fant fem faktorer, finner man i det norske utvalget seks faktorer med eigenverdi >1. Inspeksjon av datamatriksen viser at dette skyldes at i det norske utvalget utgjør variabelen kalt «Total respons-bias» en sterk egen faktor (faktor 6), mens variabelen i det amerikanske utvalget lader moderat høyt på faktor 2, sammen med «Fri gjenkalling – total intrusjoner» og «Intrusjoner – gjenkalling med stikkord». Tabell 4 viser videre at faktor 1, som må antas å være en generell innlærings- og hukommelsesfaktor, er

den dominerende faktoren, idet den alene forklarer 28% av variansen. Variablene som inngår, lader også svært likt i begge pasientutvalg.

Rekkefølgen av faktorene er noe forskjellig, men ettersom egen-verdiene for alle faktorer, bortsett fra faktor 1 og faktor 2, er nærmest identiske (mellom 1.0 og 1,6), har dette liten betydning. Det bemerkes ellers at de variablene som inngår i faktor 4 og faktor 5, lader høyt i begge pasientgrupper.

### CVLT-II versus WMS-R

Tabell 5 viser prestasjoner på CVLT-II og WMS-R for de 100 pasientene som gjennomførte begge tester. Man velger her å presentere de indeksene som det er mest vanlig å benytte ved vurdering av hukommelse i klinisk praksis. CVLT-II inneholder kun verbale ledd, mens WMS-R inneholder ledd med både verbalt og visuelt innhold. Som følge av dette, og fordi indeksene i WMS-R generelt er av mer «global» karakter og er mindre opptatt av hvilke læringsstrategier som benyttes og hvilke typer feil forsøkspersonen gjør, enn det man ser i CVLT-II, er ikke indekser fra de to testene direkte sammenlignbare. Vi vil imidlertid hevde at følgende indekser er av de vanligste ved vurdering av innlæring og hukommelse.

**TABELL 5. RESULTATER WMS-R OG CVLT-II**

	<b>Kvotient / z- og T-skårer</b>	<b>T-skårer</b>
Generell hukommelse	93.86 / 20.86	45.99
Utsatt hukommelse	92.86 / 20.36	45.24
Totalt korrekte 1-5	39.53 / 10.87	39.53
Utsatt fri gjenkalling	- , 92 / 1.37	40.80

n = 100

For CVLT-II antar man at dette er:

- Totalt antall korrekte for forsøkene 1–5 («Totals 1–5»).
- Antall korrekte etter langt intervall uten stikkord («long delay free recall»).

For WMS-R antar man at følgende indekser er viktigst:

- Generell (umiddelbar) hukommelse («general memory»).
- Utsatt hukommelse («delayed memory»).

Totalt antall korrekte i de fem første forsøkene («Totals 1–5») i CVLT-II tilsvarer i prinsippet indeksen for Generell hukommelse i WMS-R, mens indeksen for totalt antall korrekte etter langt intervall uten stikkord tilsvarer indeksen Utsatt hukommelse i WMS-R.

Ettersom CVLT-II opererer med både z-skårer og T-skårer (gjennomsnitt henholdsvis 0 og 50 og SD henholdsvis 1 og 10), og WMS-R bruker kvotientskårer (gjennomsnitt =100 og SD =15), har man for å bedre oversikten og lette sammenligningen omgjort alle skårer til T-skårer.

Tabell 5 viser at prestasjoner på CVLT-II ligger omtrent fem T-skårer, eller et halvt standardavvik, lavere enn prestasjonene på WMS-R. Testing av Pearsons rangkorrelasjon viste høyt statistisk signifikant samsvar mellom totalt antall korrekte i de fem første forsøkene på CVLT-II og generell hukommelse i WMS-R ( $r = .64, p < .0001$ ). Det var også høyt statistisk signifikant samsvar mellom fri gjengivelse etter langt intervall i CVLT-II og utsatt hukommelse i WMS-R ( $r = .58, p < .0001$ ).

## Diskusjon

Denne studien ga to hovedfunn: For det første er faktorstrukturen i CVLT-II i et norsk blandet klinisk utvalg noe forskjellig fra det som er oppgitt for de samme variablene i CVLT-II i et blandet amerikansk utvalg, men faktorladningene er svært like. For det andre er samsvaret mellom prestasjoner på CVLT-II og WMS-R bedre enn samsvaret mellom CVLT-1 og WMS-R.

Forskjellen i faktorstruktur skyldes at variabelen kalt «Total response bias», som dreier seg om hvorvidt man har tendens til å svare bekreftende eller benektende når man er usikker, i det amerikanske utvalget lader på samme faktor som variablene kalt «Fri gjenkalling – totalt intrusjoner» og «Intrusjoner – gjenkalling med stikkord», mens variabelen «Total response bias» står fram som en egen faktor i det norske utvalget (faktor 6).

Dessuten består faktor 2 i det norske utvalget av to variabler som vanligvis oppfattes som et mål på konsentrasjon/minnespenn. I det amerikanske utvalget lader disse variablene på faktor 3, sammen med totalt antall gjentakelser.

Hvorvidt dette avspeiler at den norske oversettelsen fungerer forskjellig fra den amerikanske, er det ikke mulig å besvare sikkert, blant annet fordi diagnostiske pasientdata for den amerikanske pasientgruppen ikke fremgår av manualen. Man vet følgelig ikke om det dreier seg om to direkte sammenlignbare pasientgrupper.

I begge utvalg finner vi likevel at de samme seks av i alt 16 utvalgte variabler lader høyt på en og samme hovedfaktor (faktor 1), og denne faktoren forklarer alene 28% av variansen i det norske utvalget. Dette, sammen med at disse variablene alle har med ulike sider av læring og hukommelse å gjøre, indikerer at faktoren må antas å avspeile en underliggende generell innlærings- og hukommelsesfaktor («g-faktor» for hukommelse).

En finner også tre andre faktorer i begge pasientutvalg som kan antas å være uttrykk for ulike strategier for innlæring og hukommelse (faktor 3, faktor 4 og faktor 5). Faktorenes innhold har delvis å gjøre med feiltyper og læringsfeil, som tendensen til å si andre ord enn dem man var bedt om å huske, og delvis å gjøre med ulike strategier for å huske ordlisten, som å forsøke å gjengi ordlisten slik den ble lest, og at man grupperer ord etter innhold eller etter plassering på listen. Teoretisk har disse faktorene stor interesse, fordi de har sin parallell i velkjente begreper fra kognitiv psykologi, nemlig «serieposisjonseffekt» og «primacy-recency»-effekt.

Delis, Freeland, Kramer og Kaplan (1988) og Klatzky (1980) har vist at tendens til å benytte semantisk gruppering korrelerer høyt med god langtidshukommelse. Baldo, Delis, Kramer og Shimamura (2002) har funnet at redusert nøyaktighet i hukommelsen, i form av økt tendens til å rapportere ord som ikke står på listen, redusert bruk av semantisk gruppering og redusert evne til å gjenkjenne ord er vanlig hos pasienter med frontale skader.

Delis, Kramer, Kaplan og Ober (2000) refererer studier som har vist at uttalt tendens til å si ja og svare bekreftende ved gjenkjenningforsøk forekommer ofte hos pasienter med alvorlige innkodningsvansker, slik man ser ved Alzheimers sykdom. Til sammenligning har depressive

pasienter en uttalt tendens til å si nei og svare benektende i gjenkjenningsforsøk. Delis et al. (2005) fant også at en ny indeks i CVLT-II kalt «Gjenkjenning – diskriminabilitet», som dreier seg om evne til å skille mellom ord som forekom/ikke forekom i ordlisten ved gjenkjenningsforsøk, skilte mellom personer med Alzheimers sykdom og personer med Huntingtons sykdom.

De tre nevnte faktorene kan således antas å være vesentlige for å skille mellom ulike typer hukommelsessvikt. Når de gjenfinnes som sterke faktorer både i et amerikansk og et norsk klinisk utvalg, er dette en indikasjon på at den oversatte versjonen ivaretar vesentlig klinisk informasjon.

For klinisk undersøkelse av hukommelse er det ikke nødvendigvis vesentlig at man finner en viss forskjell i antall faktorer og faktorladninger mellom den amerikanske og den norske versjonen, ettersom man for begge versjoner finner en svært sterk og lik førstefaktor, som ut fra de indeksene som inngår, kan antas å tilsvare en sentral generell faktor for innlæring og hukommelse. At det i begge grupper også finnes tre faktorer som kan antas å avspeile ulike underliggende strategier for innlæring og hukommelse, styrker antagelsen om at den norske oversatte versjonen av CVLT-II har beholdt sine viktigste egenskaper som diagnostisk instrument.

Wiegner og Donders (1999) fant fire faktorer hos en gruppe pasienter med traumatisk hodeskade og ikke fem eller seks, slik manualen til CVLT-1 oppga. Det kan bety at antall variabler avhenger av hvilken klinisk gruppe som undersøkes.

Det ville styrke validiteten av den norske versjonen om man fant en tilsvarende sammenheng mellom prestasjoner på testen og demografiske variabler som Delis et al. (2000) har påvist for den amerikanske versjonen. Det er tidligere godt dokumentert av blant annet Kramer, Delis og Daniel (1988) og Luszcz og Bryan (1999) at kjønn og alder, samt i noe mindre grad utdanning, er positivt korrelert med prestasjoner på den første versjonen av CVLT. Det ville derfor vært av interesse å undersøke om yngre skårer bedre enn eldre, om kvinner skårer bedre enn menn, og om høyt utdannende skårer noe bedre enn lavere utdannede også i Norge. Ettersom vår gruppe består av pasienter og antallet er relativt lite, fant man det imidlertid ikke hensiktsmessig å gjennomføre en slik analyse. Når det er samlet et større datamateriale, vil en slik analyse likevel være verdifull, og dette er noe vi planlegger å gjennomføre.

Resultatene fra faktoranalysen kan videre tyde på at man også ved bruk av CVLT-II bør basere vurderinger på flere variabler når man undersøker hukommelse hos kliniske grupper. Dette er i tråd med Reitan og Wolfsons (1993) tilråding om at det ved kliniske vurderinger som en hovedregel bør tas utgangspunkt i brede samleskårer («composite scores»), ettersom konsistente mønstre av skårer vanligvis er mer reliable og valide enn enkeltskårer. Samtidig vil det å studere spredning i underliggende enkeltskårer kunne gi verdifull klinisk informasjon om testpersonens særtrekk og måte å tenke på. Det samme resonnementet gjør også Kaufman og Lichtenberger (1999) seg til talsmenn for ved tolkning av WAIS-III. I praksis vil også de fleste klinikere vanligvis basere seg på flere mål på hukommelse enn det som fremkommer på en ordlisteprøve som CVLT-II.

Det andre hovedfunnet i denne studien er at det normmessige samsvaret i prestasjonsnivå mellom CVLT-II og WMS-R er vesentlig bedre enn samsvaret mellom CVLT-1 og WMS-R. Dette kan tyde på at den kritikken som ble reist mot normene i CVLT-1 av blant annet Randolph et al. (1994) og av Bosnes og Ellertsen (2000), var berettiget, selv om Egeland et al. (2005) fant at i



et norsk normalutvalg kunne normene i CVLT-II være for «snille». Egeland et al. brukte en gruppe som bestod av 50 personer i alderen 20–49 år (gjennomsnittsalder 33), samtidig som estimert gjennomsnittlig IQ var 105, slik at deres gruppe både er vesentlig mindre og yngre og har atskillig bedre IQ enn vår kliniske gruppe. Lundervold, Reinvang og Lundervold (1994) fant at 20 friske kontroller med gjennomsnittsalder på 39 år presterte i overkant av gjennomsnittet i normgruppen for CVLT-1 på indeksen Totalt antall korrekte under innlæring («Totals 1–5»), og Øie, Sundet og Rund (1999) fant at 30 normale ungdommer med gjennomsnittsalder på knapt 16 år presterte tilsvarende gjennomsnittet for personer i gruppen 17–34 år på CVLT-1.

En styrke ved de refererte norske studiene er bruken av normale kontrollgrupper, men en svakhet er at utvalgene består av mest yngre personer og normalgruppene er relativt små. Egeland et al. fant heller ikke at prestasjonsmålene i CVLT korrelerte statistisk signifikant med alder, noe som ellers er et konsistent funn for hukommelsestester. Etter vår vurdering medfører dette at deres konklusjoner må betraktes som noe usikre, i alle fall for høyere aldersgrupper, noe også Egeland et al. selv anfører. Hvorvidt de publiserte normene for CVLT-II er for «snille» eller ikke, kan likevel ikke besvares ut fra vår studie. Dette kan bare en studie av et større og representativt normalmateriale avgjøre.

Ut fra studier av Flynn (1987) kan man vente en årlig prestasjonsbedring på Wechsler-skalaene på ca. 0,3 poeng. Hvorvidt det samme gjelder for andre tester av kognitiv funksjon, er ikke klart. Normene i WMS-R kom i 1987, mens normene i CVLT-II kom i 2000. Hvis «Flynn-effekten» gjelder, ville man altså vente ca. fire poeng høyere prestasjon på WMS-R enn på CVLT-II. Den observerte forskjellen er ca. fem poeng og passer altså meget godt med den forventede prestasjonsbedringen.

Dette betyr at de mest sammenlignbare skårer på de to testene *på gruppenivå* kan forventes å være like, forutsatt at man tar hensyn til at normene i CVLT-II er av 13 år nyere dato enn normene i WMS-R. Ved svake prestasjoner på en hukommelsestest bør man i klinisk praksis med enkeltpasienter som prinsipp også gjøre andre hukommelsestester, for å se om det svake resultatet kan bekreftes med en annen metode. Det ansees å være god praksis å stole mer på mønster i prestasjoner enn på enkeltskårer og enkelttester. Slik sett er våre funn av interesse, ettersom de bidrar til økt sikkerhet om hvor stort samsvar som kan ventes mellom ulike metoder, som hver for seg er tenkt å avspeile noenlunde de samme underliggende funksjonene, i dette tilfellet innlæring og hukommelse.

Ole Bosnes

Psykiatrisk avdeling, Sykehuset Namsos

7800 Namsos

Tlf 74 21 54 00

E-post [ole.bosnes@hnt.no](mailto:ole.bosnes@hnt.no)

*Teksten sto på trykk første gang i Tidsskrift for Norsk psykologforening, Vol 44, nummer 7, 2007, side 887-892*

**TEKST**

**Ole Bosnes**

## Referanser

- Baldo, J. V., Delis, D., Kramer, J. & Shimamura, A. P. (2002). Memory performance on the California Verbal Learning Test-II: Findings from patients with focal frontal lesions. *Journal of the International Neuropsychological Society*, 8, 539-546.
- Binet, A. & Simon, T. (1916). *New methods for the diagnosis of the intellectual level of subnormals. I The development of intelligence in children*. Baltimore: Williams & Wilkin Company.
- Bosnes, O. & Ellertsen, B. (2003). California Verbal Learning Test og Wechsler Memory Scale-Revised: Prestasjoner i et klinisk utvalg i Norge. *Tidsskrift for Norsk Psykologforening*, 40, 590-597.
- Claparede, E. (1919). Percentiles of some aptitude tests. *Archives de Psychologie*, 17, 313-324.
- Delis, D. C., Freeland, J., Kramer, J. H. & Kaplan, E. (1988). Integrating clinical assessment with cognitive neuroscience: Construct validity of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology*, 56, 123-130.
- Delis, D. C., Kramer, J. H., Kaplan, E. & Ober, B. A. (1987). *California Verbal Learning Test: Adult Version Manual*. San Antonio, TX: The Psychological Corporation.
- Delis, D. C., Kramer, J. H., Kaplan, E. & Ober, B. A. (2000). *California Verbal Learning Test Second Edition. Adult Version. Manual*. The Psychological Corporation.
- Delis, D. C., Wetter, S. R., Jacobson, M. W., Peavy, G., Halmilton, J., Gongvatana, A., Kramer, J. H., Bondi, M. W., Corey-Bloom, J. & Salmon, D. P. (2005). Recall discriminability: Utility of a new CVLT-II measure in the differential diagnosis of dementia. *Journal of the International Neuropsychological Society*, 11, 708-715.
- Egeland, J., Sundet, K., Landrø, N. I., Rund, B. R., Asbjørnsen, A., Hugdahl, K., Lund, A., Roness, A. & Stordal, K. (2005). Validering av normer for oversatte tester av oppmerksomhet og hukommelse i et norsk normalutvalg. *Tidsskrift for Norsk Psykologforening*, 42, 99-105.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Gimse, R. (1992). *Wechsler Memory Scale-Revised. Håndbok*. Oslo: Norsk Psykologforening.
- Kaufman, A. S. & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. New York: John Wiley & Sons.
- Klatzky, R. L. (1980). *Human memory: Structure and processes*. San Francisco: Freeman.
- Kramer, J. H., Delis, D. C. & Daniel, M. (1988). Sex differences in verbal learning. *Journal of Clinical Psychology*, 44, 907-915.
- Lezak, M. D., Howieson, D. E. & Loring, D. W. (2004). *Neuropsychological assessment (4th ed.)*. New York: Oxford University Press.
- Lundervold, A. L., Reinvang, I. & Lundervold, A. (1994). Characteristic patterns of verbal memory functions in patients with Huntington's disease. *Scandinavian Journal of Psychology*, 35, 38-47.
- Lundervold, A. J. & Sundet, K. S. (2004). *Norsk Versjon. Manualsopplement. CVLT-II*. Psykologiforlaget AB.
- Luszcz, M. A. & Bryan, J. (1999). Toward understanding age-related memory loss in late adulthood. *Gerontology*, 45, 2-9.
- Mitrushina, M., Boone, K. B., Razani, J. & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment (2nd ed.)*. New York. Oxford University Press.
- Randolph, C., Gold, J. M., Kozora, E., Cullum, C. M., Hermann, B. P. & Wyler, A. R. (1994). Estimating memory function: Disparity of Wechsler Memory Scale-Revised and California Verbal Learning Test indices in clinical and normal samples. *The Clinical Neuropsychologist*, 8, 99-108.
- Reitan, R. M. & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation (2nd ed.)*. Tucson, AZ: Neuropsychology Press.
- Spreen, A. & Strauss, B. (1998). *A compendium of neuropsychological tests (2nd ed.)*. New York: Oxford University Press.
- Wechsler, D. (1987). *Wechsler Memory Scale-Revised. Manual*. Psychological Corporation.

Wiegner, S. & Donders, J. (1999). Performance on the California Verbal Learning Test after traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 23, 159-170.

Øie, M., Sundet, K. & Rund, B. R. (1999). Contrasts in memory functions between adolescents with schizophrenia or ADHD. *Neuropsychologia*, 37, 1352-1358.