

# Klinisk utprøving av WISC-V

## TEKST

Jens Egeland

Olaf Lund

Tor Herman Andreassen

PUBLISERT 1. mars 2022

DOI <https://doi.org/10.52734/e4t74RZ6>

## ABSTRACT:

Two-hundred and thirty-six WISC-V protocols from children assessed in the Child Psychiatric Clinic of Vestfold Hospital Trust took part in the study. We analyzed four issues known either to be specifically important or to be problematic in previous Wechsler adaptations. The first issue relates to whether the items on the translated verbal tests show the expected decreasing frequency of correct answers. Second, we check for invariance of the five-factor structure across age and gender groups. Third, we assess whether the new visual span test varies sufficiently from digit span to defend having two span tests. Fourth, we report the Cronbach's alpha levels in the clinical sample and check for possible verbal items decreasing reliability. The analyses showed small insufficiencies, and thus we conclude that the test also maintains the necessary psychometric capacity in a clinical sample.

*Keywords:* WISC-V, invariance of factor structure, fluid reasoning, working memory

## EMNER

WISC-V

faktorstruktur

Flytende resonnering

Arbeidshukommelse

Wechsler Intelligence Scale for Children (WISC) har i mange år vært den mest brukte kognitive testen for barn i Norge (Ryder, 2021; Vaskinn & Egeland, 2012). Testen har utviklet seg betydelig i de 72 årene som har gått fra den første utgaven i USA. Opprinnelig målte den et relativt homogent konstrukt kalt intelligens, riktignok oppdelt i verbale evner og utføringsevner. Fra 70tallet har testen vært gjenstand for mange faktoranalyser. I den reviderte versjonen (WISCR) avdekket Kaufman (1979) en tredje faktor kalt Freedom from distractibility. Han lanserte slagordet *intelligent testing* heller enn intelligenstesting ved å hevde at det er relative styrker og svakheter i evneprofilen som er mest nyttig for å forstå pasienten og kunne treffe målrettede tiltak. Målsettingen var ikke primært å klassifisere, men å forstå læreforutsetningene til pasienten. Denne tenkningen er brakt videre i ytterligere revisjoner og har nådd sitt foreløpige høydepunkt ved lanseringen av den femte versjonen i USA i 2014 og i Norge fra 2017. WISCV er bygd opp for å måle fem aspekter ved generell evne. I tillegg til de fire faktorene som allerede var innbakt i WISCIV, pretenderer testen nå å måle Flytende resonnering. Dette begrepet er tatt fra Cattell Horns dikotomi mellom flytende og

krystalliserte ferdigheter (Schneider & McGrew, 2012). Mens krystalliserte ferdigheter er faktisk kunnskap tilegnet gjennom skole og akkultureringsprosesser, er potensialet for utvikling av flytende resonnering medfødt og biologisk forankret (Kent, 2017). Femfaktorløsningen er funnet i faktoranalyser foretatt av testutgiverne av mer enn ti ulike nasjonale standardiseringer verden over (Van de Vijver et al., 2019), mens en rekke uavhengige forskere har rapportert at de finner at den gamle firefaktorløsningen fra WISCIV fortsatt passer best (Canivez et al., 2017).

Mens alle Wechsler-testene frem til nå har undersøkt arbeidshukommelse kun med språklige tester, har WISCV nå inkorporert en visuell arbeidshukommelsestest kalt Bildehukommelse. Begrepet arbeidshukommelse har over tid blitt mer viktig for å forstå læreforutsetninger. I en rekke høyfrekvente tilstander hos barn er vansker med arbeidshukommelse vanlig og har stor betydning for å forstå funksjonsvanskene. Mest fokus har det vært på ADHD (Fosco et al., 2020). Imidlertid er svikt i verbal arbeidshukommelse typisk for lese og skrivevansker, som jo er en mer høyfrekvent tilstand enn ADHD (Ne'eman & Shaul, 2021). Måling av arbeidshukommelse kun i den auditive modaliteten kan lede til falskt positive diagnostiseringer av oppmerksomhetsvansker. Tidligere versjoner av WISC har da også vist at barn med ADHD ikke skårer svakere på Arbeidshukommelsesindeksen enn på andre indekser (Egeland et al., 2006). I analyse av de auditive minnespennprøvene fra WAISIV sammen med visuelt minnespenn fra Wechslers Memory Test fremsto modalitet (auditivt vs. visuelt) som den viktigste distinksjonen, og det ble rådet til å ikke si noe generelt om arbeidshukommelse uten å ha testet i begge modaliteter (Egeland, 2015). Det gir WISCV for første gang mulighet til. Vi skal i denne kliniske utprøvingen av WISCV se hva introduksjonen av en visuell test betyr i praksis for vurdering av arbeidshukommelsesferdigheter.

Det er meget komplisert å standardisere en norsk, eller i dette tilfellet: en skandinavisk, versjon av WISC. Psykologer som bruker testen, vil lese seg opp på internasjonale forskningsartikler for å lære å fortolke testen, og det er da viktig at den lokale versjonen måler det samme overalt i verden. Den må være godt normert, den må ha samme faktorstruktur som den opprinnelige versjonen, og den må kunne måle riktig på ulike kognitive ferdighetsnivåer. Den må være valid og reliabel, det vil si at den må måle det den pretenderer å måle, og den må gjøre det slik at ulike testere vil komme til samme resultat. Dette er særlig viktig for evnetester som gir grunnlag for fordeling av milliarder av kroner i spesialundervisning, og det er en forutsetning for diagnostisering av psykisk utviklingshemming, som kan oppleves som stigmatiserende. Det er mange eksempler på hvordan utilstrekkelig standardisering kan gi målefeil med store konsekvenser. Det mest dramatiske eksempelet er at WISCIII overdiagnostiserte psykisk utviklingshemming (Sundberg et al., 2006), noe som fikk konsekvenser for mange voksne med egne barn, da barnevernet overtok omsorgen med henvisning til denne diagnosen. Graves et al. (2020) understreker at forskere har en etisk forpliktelse til å sjekke kvaliteten av mye brukte tester uavhengig av de kommersielle eierne.

Med dette som bakteppe vender vi tilbake til WISC V og til de forhold vi undersøker i denne kliniske valideringen av testen.

## **1. Er det jevnt over økende vanskegrad av de norske verballedene?**

Ved oversettelse av en evnetest må de språklige leddene tilpasses den kulturelle konteksten. Leddene må oversettes, og noen må tas ut fordi de ikke passer i en skandinavisk sammenheng. Det må lages nye ledd, og leddene må rangeres ut fra vanskegrad. Både i WISCV og i tidligere Wechsler-testnormeringer har rangeringen av de språklige leddene vært gjort på normeringsutvalget. Det er problematisk fordi nær sagt hele normeringsutvalget skal greie de innledende leddene. Å bruke friske gjennomsnittspersoner til å rangere vanskegraden av ledd som skal besvares riktig av 95–98 %, er problematisk fordi det krever et svært stort normeringsutvalg for å få til en reliabel rangering. Feil i rangeringen kan føre til at personer kan nå kuttepunktet for avslutning av testingen prematurt, altså at de hadde greid flere ledd som de ikke blir spurt om, fordi de faller ut på vanskelige inngangsledd. Dette er et problem med WAIS-IV, der det i starten blir stilt en rekke geografisørsmål som viser seg å være særlig vanskelige for personer med lærevansker (Egeland, 2014). Siden det i flere Wechsler-tester er problematisk å rangere vanskegrad, har vi tidligere foreslått at rangeringen må gjøres i et klinisk utvalg, slik at den blir mer reliabel for personer med lærevansker (Egeland et al., 2009). Dette er ikke hensyntatt i WISCV, men i denne kliniske utprøvingen gjør vi det.

## **2. Er det invarians i faktorstruktur?**

WISCV pretenderer å måle fem aspekter ved intellektuell funksjon. Det har vært omstridt om den faktisk gjør det. En rekke studier av Canivez og Watkins gruppe (Canivez et al., 2017; Canivez et al., 2020) har ikke gitt støtte for femfaktormodellen, mens studiene gjort i forbindelse med de ni ulike nasjonale normeringene som er gjort så langt, har gitt støtte for faktorstrukturen (Weiss, Holdnack et al., 2019). Det kan virke overraskende og lite tillitvekkende at uavhengige faktoranalyser ikke gir støtte, mens de som er publisert av eierne, gjør det, men det handler i vesentlig grad om at to skoleretninger bruker ulik type statistikk. De nåværende utgiverne av Wechsler-testene er svært påvirket av et nevropsykologisk paradigme, der det er viktig i klinikken å se på styrker og svakheter og skreddersy tiltak for spesifikke lærevansker, arbeidshukommelsesvikt eller nedsatt tempo. De tester ut hierarkiske eller korrelerte faktormodeller som har til hensikt å avdekke gruppefaktorer under en samlet G-faktor. Alternativt kan man anvende bifaktormodeller der deltestene kan lade både på eventuelle gruppefaktorer eller direkte på G. Metoden «stjeler» varians fra gruppefaktorene og reduserer dermed muligheten for å identifisere dem. I faktoranalysen i den norske versjonen ble bare hierarkiske modeller testet, og da med alle 15 delprøvene. Firefaktorløsningen som tillot delprøven Regning å lade på flere indekser, ga da best samsvar mellom data og modell. Det var likevel såpass godt samsvar mellom femfaktormodellen og de observerte skårene at man la den til grunn for den norske versjonen (Wechsler, 2017). Vi har imidlertid prøvd ut faktorstrukturen i

en serie konfirmatoriske faktoranalyser, der vi sammenlignet både hierarkiskemodeller, bifaktormodeller og korrelerte tre, fire og femfaktormodeller i det samme kliniske utvalget denne artikkelen er basert på. Vi fant at en korrelert femfaktormodell uten G var den som passet data best (Egeland et al., 2021). At løsningen uten G var best, er i tråd med synspunktene til forskergruppen bak WISCV, som anbefaler at indeksene skal være det primære fortolkningsnivået i WISCV (Weiss, Saklofske et al., 2019). Faktisk indikerer faktoranalysen at bruk av fullskalaIQ i en barnepsykiatrisk sammenheng ofte kan gi et kunstig gjennomsnitt som er lite til nytte. I denne artikkelen går vi litt videre og belyser spørsmålet om invarians i faktorløsning. Det er viktig for bruk av testen at faktorløsningen er gyldig for ulike grupper som skal eksponeres for testen. Scheiber (2016) fant samme faktorstruktur på tvers av etnisitet og kjønn i den amerikanske originalversjonen, og Van de Vijver et al. (2019) fant invarians mellom ni ulike nasjonale versjoner. Graves et al. (2020) fant likevel litt ulik testladning på faktorene mellom svarte og hvite barn. Bildehukommelse og Figurveker ladet ikke fullt så høyt på sine respektive faktorer hos svarte sammenlignet med hvite barn. I denne undersøkelsen splitter vi utvalget i to ved 12 års alder og ser om faktorstrukturen er lik for de yngste og de eldste. Tilsvarende ser vi om det er forskjell på faktorstrukturen for gutter og jenter.

### **3. Styrkes Arbeidshukommelsefaktoren med innføring av bildehukommelse?**

Vi vet fra faktoranalysene av alle de ti kjernedelprøvene i WISCV at Bildehukommelse og Tallhukommelse lader på samme faktor, presumptivt arbeidshukommelse. Her ser vi på fordelingen av visuell arbeidshukommelse hos personer som skårer lavt på auditiv arbeidshukommelse, for å få et mål på hvor mange som kunne risikert feilklassifisering hvis testen fortsatt bare hadde målt arbeidshukommelse i den auditive modaliteten.

### **4. Er det tilstrekkelig intern konsistens i delprøvene?**

Delprøvene skal måle enhetlige egenskaper, og da er det viktig at det er høy korrelasjon mellom leddene. Vi ser på samsvaret leddene imellom målt med Cronbachs alfa. For de språklige delprøvene som jo er mest tilpasset, ser vi om det er enkeltledd som svekker konsistensen, det vil si om samsvaret bedres hvis enkeltledd hadde vært tatt ut.

## **Metode**

### **Deltagere**

To hundre og trettisju barn deltok i studien og ble fortløpende undersøkt med WISCV i forbindelse med henvisning og oppstart av behandling ved Barne og ungdomspsykiatrisk avdeling (BUPA) i Sykehuset i Vestfold. Prosjektet ble godkjent av Norsk senter for forskningsdata i 2018 (nummer 57463). Alder og IQmål på deltagerne fremkommer av tabell 1. Trettisju prosent av utvalget var jenter.

## **Tabell 1**

*Alder og IQ-fordeling av 237 deltagere*

	Spenn	Gjennomsnitt (SA)
Alder	6-16	11.6 (3.2)
Fullskala-IQ	46-129	90.3 (13.6)
Verbal forståelse	45-136	94.6 (13.9)
Visuo-spatial	57-138	94,9 (15.4)
Flytende resonnering	51-128	93.3 (15.1)
Prosesseringshastighet	49-135	91.6 (15.2)
Arbeidshukommelse	51-127	91.0 (14.5)

*Merknad.* SA: standardaavik.

Utover WISCV protokollen har vi ikke ytterligere data på deltagerne. De antas å gjenspeile pasientfor delingen i klinikken. De ICDdiagnoser som ble satt i 2019 og første halvdel av 2020, var følgende: F90– F98 Atferdsforstyrrelser (inklusive ADHD) 52 %; F40– F48 Nevrotiske, belastningsrelaterte og somatoforme lidelser (angst, fobier og OCD) 27 %; F50–F59 Atferdsforstyrrelser forbundet med fysiologiske forstyrrelser og fysiske faktorer (i hovedsak spiseforstyrrelser) 9.2 % og F30–F39 Affektive lidelser 6.6 %.

## Mål

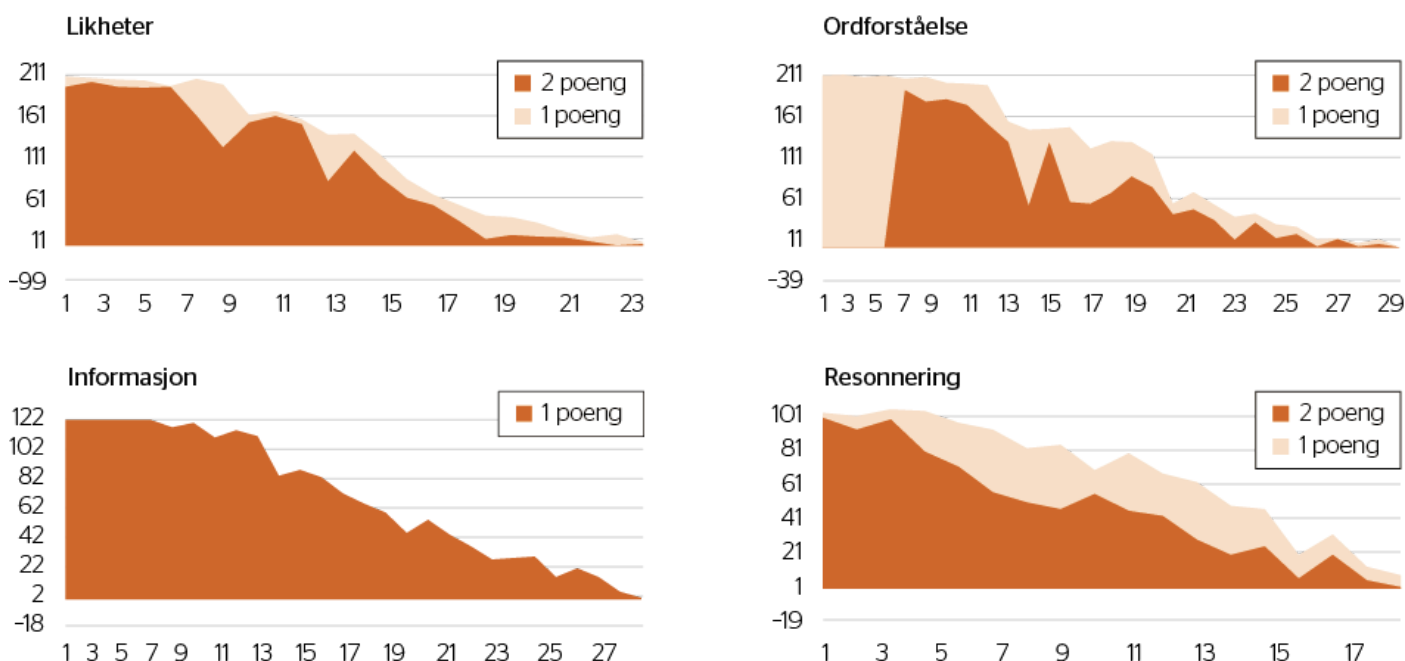
WISCV (Wechsler, 2014, 2017) består av 15 delprøver. De følgende sju er nødvendig for å beregne fullskala IQ: Likheter, Ordforståelse, Terningmønster, Matriser, Figurvekter, Tallhukommelse og Koding. For å beregne de fem indeksskårene trengs i tillegg følgende tre: Visu elle puslespill, Bildehukommelse og Symbolleting.

## Statistiske analyser

Løsningsfrekvens av leddene i språkopp gavene er enkle forekomster av null, en og to-poengsvar. Siden deltestene avsluttes etter tre nullpoengssvar, får vi ikke vite om personen hadde kunnet svare på enkeltledd etter dette kuttepunktet. Testing av invarians i faktorstruktur er gjort med eksploratorisk faktoranalyse med varimaxrotasjon der fem faktorer er spesifisert. Utvalget er splittet på kjønn og delt etter aldersmedianen i utvalget i yngre barn (6–12 år) og eldre barn (13–16 år).

Sammenhengen mellom Tallhukommelse og Bildehukommelse illustreres med et scatterplot som viser de enkelte del tagernes skalerte skårer på begge tester, og det gjøres en frekvensopptelling av variasjon i Bildehukommelse ved skår sju eller lavere på Tallhukommelse. Cronbachs alfa beregnes for tester som ikke er tidsavhengige. Cronbachs alfa øker ved antallet ledd. For de oversatte språklige prøvene ser vi om Cronbachs alfa kan bedres hvis man tar ut enkeltledd som er mindre konsistente.

## Resultater



**FIGUR 1** Løsningsfrekvens av de fire verbalprøvene i WISC-V

### 1. Økende vanskegrad av verballeddene

Figur 1 viser løsningsfrekvensene av de enkelte ledd i hver av de fire verbalprøvene. Ideelt sett skulle figuren vise en perfekt trekant. På Likheter er det særlig ledd 7 og 11 som avviker fra en rett hypotenus i trekanten. Lav løsningsfrekvens av disse leddene og høyere løsningsfrekvens av de påfølgende sees i både 6–7årsgruppen og gruppene 8–11 og 12–16 år.

I Ordforståelse er det særlig ledd 11 og leddene 13 til 15 som avviker negativt fra ideell kurve, mens ledd 12 avviker positivt. Dette gjelder alle tre aldersgrupper. Andelen to-poengssvar hos de yngste på leddene 10–16 fordeler seg slik: 14, 3, 36, 3, 6, 6 og 8 %. For de eldste er tilsvarende løsningsfrekvens 80, 41, 76, 41, 44, 45, 60 og 60 %.

For Informasjon er det kun mindre avvik fra en perfekt fallende hypotenus. I Resonnering er det særlig ledd 9 som avviker med noe høyere frekvens av topoengssvar enn de to foregående ledd.

**Tabell 2**

Invarians i faktorstruktur på tvers av kjønn. Faktorladninger gutter (n = 149) og jenter (88)

	1.faktor		2.faktor		3.faktor		4.faktor		5.faktor	
	Gutter	Jenter	Gutter	Jenter	Gutter	Jenter	Gutter	Jenter	Gutter	Jenter
Likheter	.334	.221	.043	<b>.807</b>	<b>.818</b>	.147	-.029	.156	.271	.174
Ordforståelse	.099	.017	.174	<b>.882</b>	<b>.766</b>	.134	.441	.040	.043	-.069
Terningmønster	<b>.829</b>	<b>.726</b>	.177	.183	.174	.354	.167	.054	.059	.191
Visuelle puslespill	<b>.889</b>	<b>.850</b>	.034	.074	.090	.292	.106	.156	.155	.013
Matriser	.666	.288	.221	.114	.256	<b>.814</b>	.275	.169	.135	.091
Figurvekter	.442	.294	.204	.241	.296	<b>.724</b>	<b>.635</b>	.192	-.022	.099
Tallhukommelse	.154	.119	.122	.268	.198	.132	.135	<b>.742</b>	<b>.892</b>	.246
Bildehukommelse	.236	.099	.028	-.019	.081	.186	<b>.707</b>	<b>.882</b>	<b>.500</b>	.092
Symbolleting	.049	-.032	<b>.859</b>	.025	.012	.272	.218	.127	.220	<b>.896</b>
Koding	.250	.429	<b>.851</b>	.075	.169	-.158	-.041	.297	-.058	<b>.697</b>
% forklart varians	23.8	16.8	16.3	16.2	15.2	15.9	12.8	15.5	12.2	14.5
Eigenverdi	3.92	3.91	1.42	1.36	1.20	1.12	0.90	0.85	0.61	0.65

**TABELL 2** Invarians i faktorstruktur på tvers av kjønn. Faktorladninger gutter (n = 149) og jenter (88)**Tabell 3**

Invarians i faktorstruktur på tvers av kjønn. Faktorladninger barn 6–12 (n = 97) og 13–16 år (140)

	1.faktor		2.faktor		3.faktor		4.faktor		5.faktor	
	6–12	13–16	6–12	13–16	6–12	13–16	6–12	13–16	6–12	13–16
Likheter	.337	.283	.191	<b>.803</b>	<b>.738</b>	.047	.069	-.022	.167	.280
Ordforståelse	.013	.125	.103	<b>.899</b>	<b>.879</b>	.036	.198	.150	-.018	-.001
Terningmønster	<b>.750</b>	<b>.843</b>	.114	.134	.195	.148	.333	.069	.146	.050
Visuelle puslespill	<b>.824</b>	<b>.849</b>	.203	.060	.106	.037	.289	.191	.067	-.034
Matriser	.358	.725	.352	.258	.073	.116	<b>.708</b>	.057	.102	.149
Figurvekter	.259	.540	.088	.385	.256	.216	<b>.777</b>	<b>.534</b>	.121	-.048
Tallhukommelse	.329	.043	<b>.720</b>	.185	.231	.133	.157	.260	.110	<b>.873</b>
Bildehukommelse	.050	.174	<b>.892</b>	.051	.092	.095	.151	<b>.880</b>	.162	.298
Symbolleting	-.023	.138	.139	.016	.012	<b>.753</b>	.310	.020	<b>.883</b>	<b>.462</b>
Koding	.476	.141	.188	.069	.161	<b>.910</b>	-.164	.137	<b>.717</b>	-.052
% forklart varians	18.9	24.2	16.0	17.3	15.3	15.1	15.1	12.1	14.1	11.8
Eigenverdi	4.26	3.92	1.14	1.42	0.94	1.20	0.88	0.90	0.71	0.61

**TABELL 3** Invarians i faktorstruktur på tvers av kjønn. Faktorladninger barn 6–12 (n = 97) og 13–16 år (140)

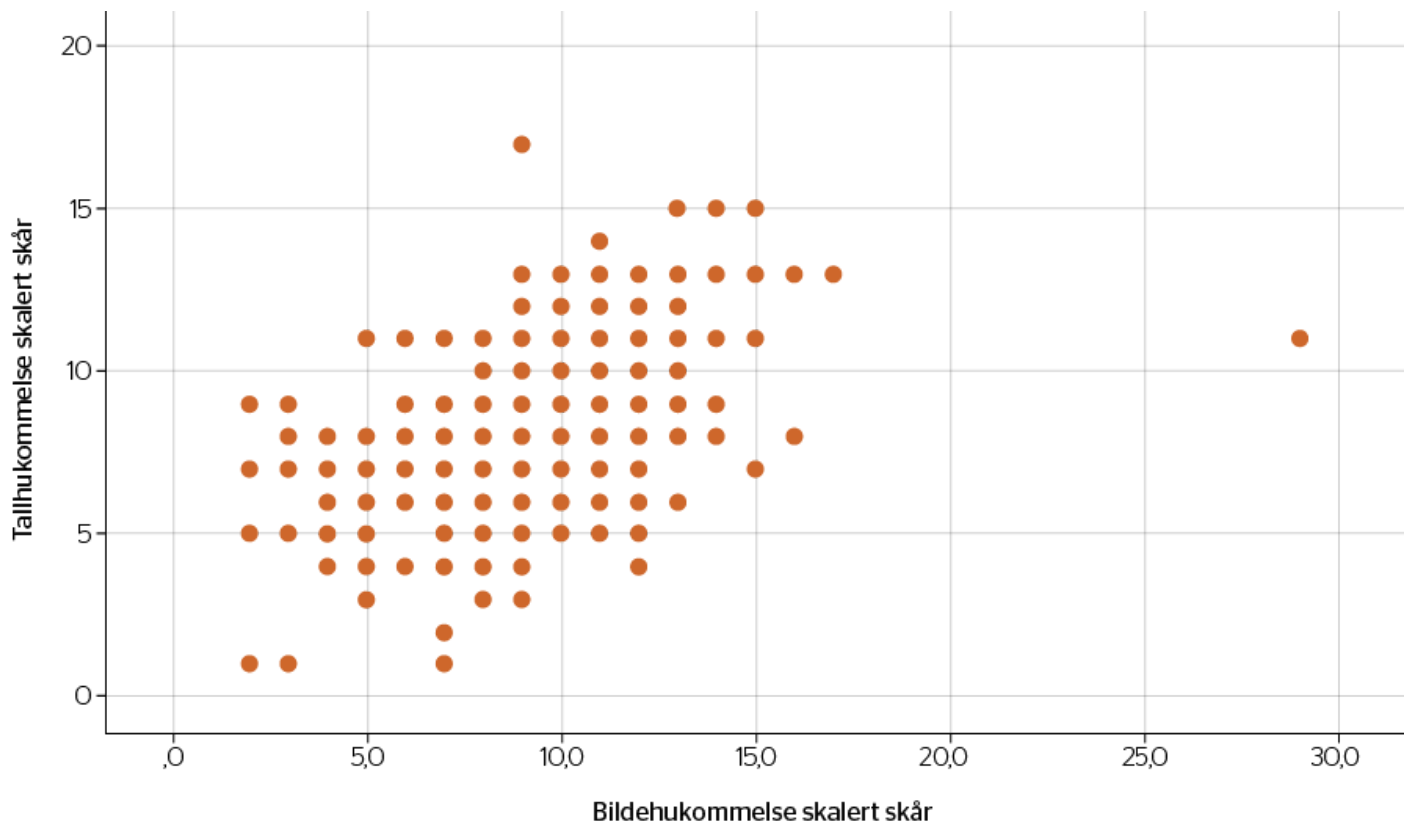
## 2. Invariasjon i faktorstruktur

Tabellene 2 og 3 viser faktorstrukturen i forhold til kjønn og alder. Håndbokens femfaktormodell fremkommer for guttene og for de yngste barna, mens for jentene og de eldste barna lader Figurvekter sammen med Bildehukommelse i stedet for å lade sammen med Matriser slik man ville forvente ut fra håndboken.

## 3. Styrkes Arbeidshukommelsesfaktoren med innføring av bildehukommelse?

Figur 2 viser fordelingen av Tallhukommelse og Bildehukommelse. Ett hundre og elleve barn hadde skår sju eller lavere på Tallhukommelse. Førtitre prosent av disse hadde

også (minst) lett redusert Bildehukommelse. Tyve prosent hadde gjennomsnittlig eller bedre Bildehukommelse.



**FIGUR 2** Scatterplot av skalerte skårer for Tallhukommelse og Bildehukommelse

#### 4. Indre konsistens

Cronbachs alfa varierte fra .864 for Likheter til .915 for Resonnering og Figurvekter. For Likheter var det samlede alfanivået like høyt eller høyere enn det ville være hvis noen av leddene ble tatt ut. For Ordforståelse kunne alfanivået øke fra .891 til .897 hvis de fire første leddene og det siste leddet ble tatt ut. For Informasjon kunne alfanivået øke fra .894 til .899 hvis de første fem leddene ble tatt ut. For Resonnering kunne alfanivået øke fra .915 for hele testen til .938 hvis ledd 10 ble tatt ut.

#### Diskusjon

Hensikten med denne artikkelen er å validere WISCV i et klinisk utvalg ved å se på fire ulike psykometriske forhold som enten vurderes som særlig viktig eller erfaringsmessig har vært problematisk med Wechslertestene.

Når det gjelder hvorvidt leddene viser økende vanskegrad utover i de språklige delprøvene, har vi identifisert noen få enkeltledd som avviker fra den forventede trekantformen på figuren. Faren ved avvikende vanskelige ledd tidlig i en test er at det øker risikoen for at personer når kuttepunktet for avslutning av testen etter tre nullpoengssvar, før potensialet deres er uttømt. Mens kuttepunktet i den amerikanske originalversjonen er bestemt empirisk, er det ikke gjort tilsvarende analyse av den norske versjonen. Det er likevel liten grunn til å tro at de små avvikene som er



identifisert her, har noen betydning så lenge nedgangen i topoengssvar kompenseres ved økt forekomst av ettpoengssvar.

Siden indeksskårene er det primære fortolkningsnivået i testen, er det viktig at faktorløsningen er gyldig i ulike aldersgrupper og for begge kjønn. Vi har testet faktorstrukturen i en konfirmatorisk faktoranalyse for hele utvalget tidligere (Egeland et al., 2021), så spørsmålet nå er ikke om femfaktormodellen er best, men om den er lik på tvers av kjønn og alder. For fire av faktorene finner vi invarians på tvers av disse inndelingene, men den flytende resonneringsfaktoren gjenfinnes ikke for de eldste og for jentene. Der lader derimot Figurveker sammen med Bildehukommelse. I sin studie av invarians på tvers av etniske grupper fant Graves et al. (2019) at Figurveker har en verbal komponent, mens vi altså finner at i den grad den avviker fra standardmodellen, er det fordi den lader sammen med testen av visuell arbeidshukommelse. For de eldste barna kan faktoranalysen få som konsekvens at det da ikke fremkommer en fortolkningsbar Arbeidshukommelsesindeks, siden Tallhukommelse lader sammen med Symbolleting og ikke Bildehukommelse. Det handler i bunn og grunn om at Matriser ikke helt oppfører seg som intendert: For jentene og de eldste lader den på en visuell persepsjonsfaktor sammen med Terningsmønster og Visuelle puslespill. Det har vært reist kritikk om at Matriser, som historisk sett har vært sentral i differensieringen mellom flytende og krystallisert intelligens, likevel viser seg å være mindre flytende og mer av en krystallisert visuoperseptuell test enn man tenkte (Tranel et al., 2008).

Hadde Bildehukommelse og Tallhukommelse korrelert perfekt, ville førstnevnte ikke bidratt med ny informasjon om arbeidshukommelse: Alle som gjorde det godt på den ene testen, ville da gjøre det godt på den andre. Den empiriske fordelingen i materialet her viser at under halvparten av de med lett redusert Tallhukommelse også har tilsvarende lett redusert Bildehukommelse. Tyve prosent skårer over middels på Bildehukommelse til tross for en lett redusert Tallhukommelse. Studier av spesifikke lærevansker viser nettopp tilsvarende variasjon i minnespenntester ut fra modalitet (Giofrè et al., 2018). I dette materialet vet vi ikke om variasjonen skyldes at barna har ulike lærevansker, men det er noe som burde undersøkes nærmere. Så langt kan vi konkludere med at Bildehukommelse i hele utvalget lader på en arbeidshukommelsesfaktor. Samvariasjonen med Figurveker for de eldste og guttene er en indikasjon på at den måler noe annet enn Tallhukommelse. Testen bidrar til å nyansere arbeidshukommelsesfaktoren sammenlignet med hva som hadde vært tilfellet hvis man fortsatt bare undersøkte i den auditive modaliteten.

Cronbachs alfaverdier er sjelden noe som vekker entusiasme hos klinikerne, men de er viktige mål som viser deltestenes indre konsistens som må dokumenteres. Her er verdiene høye og viser gjennomgående god reliabilitet. Selv om inngangsleddene i et par av oppgavene svekker reliabiliteten antydningssvis, er det ikke tilstrekkelig grunn til å fjerne dem for de få som ikke greier dem. I analysene her er det kun identifisert et enkeltledd i Resonnering (kalenderledd 10) som ser ut til å svekke delprøven, men ikke mer enn at alfanivået er komfortabelt høyt også med det leddet. I denne kliniske valideringen av WISCV har vi gått gjennom fire forhold som enten er svært viktig for

bruk av testen, eller som historisk har vist seg å være svake punkt i Wechsler-testene. Hvordan vil vi nå samlet vurdere testen? Deler av forfattergruppen har gjennom de siste 20 årene foretatt kliniske uavhengige valideringer av Wechsler-testene og påvist alt fra store mangler i WISCIII (Sundberg et al., 2006) til mindre problemer som klinikerne allikevel må være klar over i WAISIV (Egeland, 2014). Vår vurdering er at WISC-V er den mest vellykkede Wechsler-versjonen etter at det offentlige sluttet å finansiere normeringer (dvs. etter WISC-R). Brøndbo og Egeland (2019) gjennomgang av WISC-V i PsykTest Barn diskuterte den manglende dokumentasjonen av gyldigheten av femfaktorløsningen og anbefalte at klinisk profiltolkning bør gjøres med forsiktighet. Utsagnet kunne så tvil om den måten mange klinikere bruker evnetesten på, nemlig å analysere sterke og svake sider. Utenfor habiliteringsfeltet vil det å kun fortolke fullskala-IQ gjøre testen lite verdifull. Publikasjonen av en konfirmatorisk faktoranalyse (Egeland et al., 2021) og ganske god invarians på tvers av alder og kjønn i denne artikkelen tilsier at femfaktormodellen har noe for seg, og at det spesielle forbeholdet til Brøndbo og Egeland (2019) kan trekkes tilbake. Det gjenstår mye klinisk nyttig utprøving, så som å se hva den nye indeksen Flytende resonnering faktisk predikerer, og hva diskrepans mellom Tallhukommelse og Bildehukommelse kan fortelle om eventuelle lese og skrivevansker eller non-verbal problematikk.

Begrensningene i denne studien er hovedsakelig knyttet til at dette er en arkivstudie der vi mangler klinisk informasjon om deltagerne. Hadde vi kjent diagnosene, kunne vi eksempelvis regnet på kjennetegn ut fra modalitetsforskjeller i arbeidshukommelse. Vi skulle helst ha fortsatt utprøvingen utover håndbokens kuttepunkt på tre nullpoengssvar, slik at vi mer overbevisende kunne dokumentere at rangeringen av leddene i språkprøvene ikke påvirker totalskåren. At det er få jenter i datamaterialet, legger en begrensning på funnet om at femfaktorløsningen er problematisk for dem. Noen få uvanlige skårer kan ha bidratt til det avvikende funnet.

*Teksten sto på trykk første gang i Tidsskrift for Norsk psykologforening, Vol 59, nummer 3, 2022, side 198-205*

#### TEKST

**Jens Egeland**

KONTAKT: jens.egeland@siv.no

**Olaf Lund**

**Tor Herman Andreassen**, Klinikk psykisk helse og avhengighet, Sykehuset i Vestfold HF

+ **Vis referanser**

Brøndbo, H. & Egeland, J. (2019) Måleegenskaper ved den norske versjonen av Wechsler Intelligence Scale for Children - Fifth Edition (WISC-V) [Measurement properties of the Norwegian version of Wechsler Intelligence Scale for Children-Fifth Version (WISC-V)]. *PsykTestBarn*, 9(5). <https://doi.org/10.21337/0064>

Canivez, G. L., McGill, R. J., Dombrowski, S. C., Watkins, M. W., Pritchard, A. E. & Jacobson, L. A. (2020). Construct Validity of the WISC-V in Clinical Cases: Exploratory and Confirmatory Factor

Analyses of the 10 Primary Subtests. *Assessment*, 27(2), 274-296.

Canivez, G. L., Watkins, M. W. & Dombrowski, S. C. (2017). Structural validity of the Wechsler Intelligence Scale for Children – Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*, 29(4), 458-472.  
<https://doi.org/10.1037/pas0000358>

Egeland, J. (2014). Klinisk utprøving av WAIS-IV: Leddanalyser, stopp-kriterier og arbeidsminnemåling. *Tidsskrift for Norsk psykologforening*, 51(11), 911-919.

Egeland, J. (2015). Measuring Working Memory with Digit Span and the Letter-Number Sequencing subtests from the WAIS-IV: Too Low Manipulation Load and Risk of Underestimating Modality Effects. *Applied Neuropsychology: Adult*, 22(6). <https://doi.org/10.1080/23279095.2014.992069>

Egeland, J., Andreassen, T. H. & Lund, O. (2021). Factor structure of the new Norwegian WISC-V version: support for a five-factor model. *Scandinavian Journal of Psychology*.  
<https://doi.org/10.1111/sjop.12780>

Egeland, J., Bosnes, O. & Johansen, H. (2009). Utprøving av WAIS-III i et klinisk utvalg. *Tidsskrift for Norsk psykologforening*, 46, 658-666.

Egeland, J., Sundberg, H., Andreassen, T-H. & Stensli, O. (2006). Reliability and validity of Freedom from Distractibility and Processing Speed Factors in the Norwegian WISC-III-version. *Nordic Psychology*, 58, 136-149.

Fosco, W. D., Kofler, M. J., Groves, N. B., Chan, E. S. M. & Raiker, J. S. (2020). Which 'Working' Components of Working Memory aren't Working in Youth with ADHD?. *Journal of Abnormal Child Psychology*, 48, 647-660. <https://doi.org/10.1007/s10802-020-00621-y>

Giofrè, D., Donolato, E. & Mammarella, I. C. (2018). The differential role of verbal and visuospatial working memory in mathematics and reading. *Trends in Neuroscience and Education*, 12, 1-6.  
<https://doi.org/10.1016/j.tine.2018.07.001>

Graves, S. L Jr., Smith, L. V. & Nichols, K. D. (2020). Is the WISC-V a Fair Test for Black Children: Factor Structure in an Urban Public School Sample. *Contemporary School Psychology*.  
<https://doi.org/10.1007/s40688-020-00306-9>

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. John Wiley & Sons.

Kent, P. (2017). Fluid intelligence: A brief history. *Applied Neuropsychology: Child*, 6(3), 193-203.  
<https://doi.org/10.1080/21622965.2017.1317480>

Ne'eman, A. & Shaul, S. (2021). Readiness or Impairment: Cognitive and Linguistic Differences Between Children Who Learn to Read and Those Who Exhibit Difficulties With Reading in Kindergarten Compared to Their Achievements at the End of First Grade. *Frontiers in Psychology*, 12, 456. <https://doi.org/10.3389/fpsyg.2021.614996>

Ryder, T. (2021). Testkvalitetsprosjektet - del 1: Norske psykologers testholdninger og testbruk. *Tidsskrift for Norsk psykologforening*, 58(1), 28-37.

Scheiber, C. (2016) Is the Cattell-Horn-Carroll-Based Factor Structure of the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) Construct Invariant for a Representative Sample of African-American, Hispanic, and Caucasian Male and Female Students Ages 6 to 16 Years? *Journal of Pediatric Neuropsychology*, 2, 79-88, DOI 10.1007/s40817-016-0019-7

Schneider, W. J. & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. I D. P. Flanagan & P. L. Harrison (red.), *Contemporary intellectual assessment: Theories, tests, and issues* (3. utg., s. 99-144). New York: Guilford.

Schreiber, J. B., Nora, A., Stage, F. K, Barlow, E. A. & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research*, 99(6), 323-338. <https://doi.org/10.3200/JOER.99.6.323-338>

Sundberg, H., Egeland, J., Andreassen, T. H. & Stensli, O. (2006). Sammenligning av WISC-R- og WISC-III-skårer i et klinisk utvalg: Svakheter i WISC-III-normer eller foreldede WISC-R-normer? *Tidsskrift for Norsk psykologforening*, 43, 476-481.

Tranel, D., Manzel, K. & Anderson, S. W. (2008). Is the prefrontal cortex important for fluid intelligence? A neuropsychological study using Matrix Reasoning. *The Clinical neuropsychologist*, 22(2), 242-261. <https://doi.org/10.1080/13854040701218410>

Van de Vijver, F. J. R., Weiss, L. G., Saklofske, D. H., Batty, A. & Prifitera, A. (2019). A cross-cultural analysis of the WISC-V. I L. G. Weiss, D. H. Saklofske, J. A. Holdnack & A. Prifitera (red.), *WISC-V. Clinical use and interpretation* (s. 223-224), Academic Press.

Vaskinn, A. & Egeland, J. (2012). Testbruksundersøkelsen: En oversikt over tester brukt av norske psykologer. *Tidsskrift for Norsk psykologforening*, 49, 658-665.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children - Fifth Edition. Technical and interpretive manual*. San Antonio, TX: NCS Pearson Assessment.

Wechsler, D. (2017). *Wechsler intelligence scale for children - Fifth Edition (WISC-V). Manual, Del 1, norsk versjon*. Stockholm: Pearson Assessment.

Weiss, L. G., Holdnack, J. A, Saklofske, D. H. & Prifitera, A. (2019). *Theoretical and Clinical Foundations of the Wechsler Intelligence Scale for Children - Fifth Edition*. I L. G. Weiss, D. H. Saklofske, J. A. Holdnack & A. Prifitera (red.), *WISC-V. Clinical use and interpretation* (s. 97-128). Academic Press.

Weiss, L. G., Saklofske, D. H., Holdnack, J. A & Prifitera, A. (2019). *Advances in the Assessment of Intelligence*. I L. G. Weiss, D. H. Saklofske, J. A. Holdnack & A. Prifitera (red.), *WISC-V. Clinical use and interpretation* (s. 3-23). Academic Press.