

WISC-III og normer

TEKST

Bjørn Rishovd Rund

Dag Øyvind Engen Nilsen

PUBLISERT 1. juni 2005

Testutviklingen i Norge har i de seneste tiår vært på etterskudd. Det har vært publisert få nye instrumenter, og mange av testene som har vært benyttet har vært utdatert i forhold til både normering og relevans. Assessio Norge AS har i løpet av de siste tre årene hatt som oppgave å gjøre noe med dette, og vi kan ikke annet enn å ønske debatten om standardiseringen av WISC-III velkommen.

Flere har stilt spørsmål ved normene som anvendes på WISC-III; altså normene som ble utarbeidet i den svenske standardiseringen i 1997. Vi har derfor funnet det nødvendig å presentere noen av premisene for evnetesting med Wechslerskalaene (WISC, WAIS og WPPSI).

Mange har observert at WISC-III skiller seg markant fra de britiske normene når det gjelder lave skårer, og man tar for gitt at de britiske normene er «bedre».

Svensk normering

Den svenske WISC-III er normert på byene Stockholm og Uppsala. Hver av de 20 alderskohortene inneholdt fra 50–63 barn. Utvalgets representativitet ble vurdert i forhold til foreldrenes utdanningsnivå. Dette var noe høyere enn for den svenske befolkning for øvrig, og skårene er derfor nedjustert i forhold til dette. Den svenske normeringen er gjort etter alle kunstens regler, og de feilkilder som finnes er belyst og kontrollert for i den grad det lar seg gjøre. Både de svenske og britiske normene er gode. Vi har valgt å bruke de svenske fordi disse er nyere, og vi antar at det er større kulturell nærhet mellom Norge og Sverige, enn mellom Sverige og Storbritannia. Vi vil aldri få noe endelig bevis for hvilke normer som er best. Det viktigste når man tolker psykologiske tester, er at man setter seg inn i normgrunnlaget og forholder seg til dette. Når enkelte opplever at testresultater «stemmer dårlig med virkeligheten», har dette like mye å gjøre med Wechslerskalaenes konstruksjon som med normene.

Svakheter i testkonstruksjonen

Grunnen til at Wechslerskalaene i noen tilfeller gir store avvik mellom testtakers oppnådde skåre og «faktiske evner», har å gjøre med skalaenes konstruksjon. For det første er normgrunnlaget for lave skårer så lite at testene ikke er egnet til å differensiere testtakere som skårer utenfor to standardavvik over eller under gjennomsnittet (over 130 eller under 70). I WISC-III består normgrunnlaget av 1036 barn. I et tilfeldig utvalg

vil vi forvente at ca. 21 (2 %) av disse barna har en IQ som er lavere enn 70. Disse 21 barnas resultater skal så danne normgrunnlaget for skårer som ligger mellom 55 og 70, i 20 ulike alderskohorter. Da sier det seg selv at normene for lave (og høye) skårer blir lite reliable. I tillegg vet vi at reliabiliteten i en test er avhengig av antall ledd som administreres. WISC-III er konstruert slik at de som skårer lavt får færre oppgaver enn de som skårer høyt. Testen er altså mest reliabel for skårer som ligger rundt gjennomsnittet, og resultater som ligger langt unna gjennomsnittet er mindre reliable. Konsekvensen av dette er at WISC-III ikke er et godt egnet instrument for å avgjøre grad av psykisk utviklingshemning.

Vi minner også om at WISC-III har en generell svakhet i forhold til de eldste aldersgruppene. Problemet her er at det kreves svært mange korrekte svar for å få normal skåre, og et par tilfeldige feil kan gjøre at barnet får «for lav» skåre. På utføringsoppgavene er det også slik at de eldste barna må ha tidsbonus for å skåre normalt, noe som gjør at barn som arbeider langsomt, skårer uforholdsmessig lavt i forhold til jevnaldrende som arbeider i normalt tempo. Dette handler for øvrig ikke om den norske og svenske versjonen, men generelt i alle land hvor WISC-III finnes. Den britiske versjonen er heller ikke egnet til å differensiere skårer under 70.

Konfidensintervall

Et viktig poeng i evnetesting er man skal bruke resultatet med varsomhet, bl.a. fordi testene inneholder målefeil. Derfor står alle skårer oppgitt med et konfidensintervall. En IQ-skåre omgis med et konfidensintervall på 12 IQ-poeng. En IQ-skåre på 68 kan med 90 % konfidensintervall like gjerne være 64 som 76 poeng. I tillegg er reliabiliteten lavere og konfidensintervallet større i det nedre ytterpunktet av skalaen. Dette skyldes som nevnt at svaktpresterende barn testes med færre ledd, og feilmarginen blir dermed større. Bommer man på et ledd, og befinner seg i de nedre persentiler, får dette større konsekvenser enn om man bommer på et ledd og ligger rundt gjennomsnittet.

Hva skal man forvente seg?

WISC-III er en test som gir forholdsvis valide mål på et barns evnenivå innenfor normalområdet. Så lenge skårene varierer innenfor normalområdet (80–120), kan man danne hypoteser i forhold til spesifikke problemer, i tillegg til å anslå nivå på de latente variablene Total IQ, Verbal IQ, Utførings-IQ, Verbal forståelse, Perseptuell organisering, Hurtighet og Oppmerksomhet. Når IQ-skåren nærmer seg 70, vil resultatet være mindre reliabelt, og når IQ-skårene faller under 70, bør man forholde seg til resultatet som en skåre under 70, og ikke som en skåre på 57 eller 63.

Diskreditering av testresultater

Mens jeg studerte, administrerte jeg WAIS på en ung kvinne som oppnådde en IQ-skåre på 42. Dette skulle ikke være så urimelig, all den tid vi vet at noe under en promille av befolkningen har evner som ligger i dette sjiktet, og at mange av disse før eller siden kommer til en psykolog for utredning. Problemet var her at testtaker hadde sertifikat,

og grensen for strafferettslig tilregnelighet går ved en IQ på 55. Burde denne kvinnen fått sitt førerkort inndratt som følge av WAIS-resultatet? Svaret er nei. Dersom en person har bestått både den praktiske og teoretiske delen av førerprøven, er det liten sannsynlighet for at personen har en IQ på 42.

Supplerende informasjon

Wechsler-testene er, som alle andre psykologiske tester, hypotese-genererende instrumenter. For å gjøre en vurdering av intellektuelle evner, skal en samle informasjon fra ulike kilder. Disse kan være skoleresultater, sosial fungering og ADL-ferdigheter. Generelt er testresultater den mest valide informasjonen, fordi man sammenligner testtaker med en norm. Men man må alltid ta høyde for at ekstern informasjon kan diskreditere et testresultat. Hvorfor skal en vite om barnet har 70 eller 50 i IQ, annet enn for å sette en diagnose? Dette synes å være vel så mye et byråkratisk problem som et klinisk. Man burde i stedet samle inn informasjon fra eksterne kilder som på dette nivået ofte er like valide som et WISC-III resultat. Dette kan foregå ved å undersøke for eksempel ADL-ferdigheter og sosial fungering, som er mer relevant i forhold til hvilke tiltak en skal settes inn for barnet. En alternativ tilnærming til testing av mennesker med psykisk utviklingshemninger, er å bytte WISC-III med WPPSI-r. Da kan en sammenligne barnet med yngre barn, og gjøre et anslag av barnets mentale alder.

Teksten sto på trykk første gang i Tidsskrift for Norsk psykologforening, Vol 42, nummer 6, 2005, side

TEKST

Bjørn Rishovd Rund, Psykologisk institutt, Universitetet i Oslo

Dag Øyvind Engen Nilsen