

## Ytringer

# Tester for barn og unge i Norge bør kvalitetssikres bedre

En stor del av de psykologiske testene som finnes tilgjengelig for bruk i Norge, er ikke godt nok kvalitetssikret, skriver Muirne Paap og kolleger.



**TESTKVALITET:** Testing av barn og unge kan ha mange fordeler, men det forutsetter at testen er godt nok kvalitetssikret. Det bør vi oftere sjekke om er tilfellet, f.eks. ved å gå til tidsskriftet PsykTestBarn, skriver Muirne Paap og kolleger. Illustrasjon: Gunnvor Bakke / Flickr.

**Muirne C. S. Paap, Monica Martinussen, Reidar Jakobsen, Thomas Jozefiak, Brynhildur Axelsdottir & Simon-Peter Neumer**

Sist oppdatert: 27.08.18 Publisert: 24.06.18

## + Forfatterinfo

Det er problematisk at så mange psykologiske tester og kartleggingsverktøy som blir brukt i Norge, ikke er tilfredsstillende kvalitetssikret (Kornør, Jozefiak, Hanssen-Bauer & Hysing, 2015; Lindgren, 2018; Strand, 2018).

Testresultater danner grunnlaget for vurderinger som kan ha store konsekvenser for personen som blir testet. De kan ha betydning for hvorvidt man får en diagnose og behandling, og for vurderinger om tiltak har den tiltenkte effekten.

## **LEDIG STILLING: Folkehelseinstituttet søker forsker – psykisk helse hos barn og unge**

Testing kan ha mange fordeler, som en mer sikker måling av egenskaper, evner eller ferdigheter, men økt testing i bredere forstand i skolen har også blitt kritisert i mediene de siste årene. Blant annet har fagfolk uttrykt bekymring for at utstrakt bruk av testing vil medføre en mer stresset hverdag for elevene (Øverli, 2013).

### **Tester trenger mer dokumentasjon**

Når en test blir brukt for å si noe om et enkeltindivid, er det vesentlig at det har blitt dokumentert at testen har høy nok kvalitet, og at den er egnet til å anvendes i målgruppen man er interessert i.

Med kvalitet menes det dokumentasjon på testens *reliabilitet*, *validitet* og *normer* basert på studier av populasjonen testen er tenkt brukt på (European Federation of Psychologists' Associations, 2013):

- Hvis en test har høy **reliabilitet** i en viss populasjon, som f.eks. barn i alderen 3 til 6 år, betyr det at skårene er målt med høy nøyaktighet i denne populasjonen.
- **Validitet** har å gjøre med hvorvidt testskårene reflekterer egenskapen man ønsker å måle, og om testen egner seg til det aktuelle formålet. Høy reliabilitet og god validitet for en populasjon, betyr ikke at testen nødvendigvis er god for en annen populasjon, f.eks. barn i alderen 9 til 12 år.
- **Normer** trengs for å kunne tolke om testskåren er høy eller lav, og disse er som regel basert på testskårer samlet i et stort og representativt utvalg som personens skåre kan sammenlignes med. Tilsvarende er det viktig at personen som skal anvende testen, har kunnskap om og ferdigheter i å bruke den og kjenner til eventuelle begrensninger med testen.

### **Tidsskriftet PsykTestBarn gir tilgang**

**PsykTestBarn**, som drives av de fire regionsentrene for barn og unges psykiske helse (RBUP/RKBU) på oppdrag fra Helsedirektoratet, er et nettbasert tidsskrift som utgir fritt tilgjengelige artikler om måleegenskaper ved norske versjoner av tester og kartleggingsverktøy for å undersøke psykisk helse, psykososiale problemer, funksjonsnivå, livskvalitet, og evner og ferdigheter hos barn og ungdom. Hovedmålet

med våre artikler er å gi leserne rask tilgang til objektiv informasjon av høy kvalitet som de kan bruke når de skal avgjøre hvilken test de skal anvende for å måle en viss egenskap.

For hver test gjennomføres det en systematisk litteraturgjennomgang. Dette innebærer at fagbibliotekarer gjennomfører systematiske litteratursøk etter artikler der den aktuelle testen er undersøkt, basert på skandinaviske utvalg. Deretter gjennomgår to forskere resultatet fra litteratursøket og utarbeider en artikkel. Artikkelen vurderes så av to fagfeller og redaktøren før den revideres og endelig versjon publiseres. Artiklene er bygget opp etter et standardisert format og inneholder bl.a. informasjon om testen, litteratursøket og hva som finnes av dokumentasjon om den norske utgaven av testen vedrørende reliabilitet, validitet og normer. Testens psykometriske egenskaper vurderes i tråd med retningslinjene fra den europeiske foreningen for psykologforeninger ([EFPA, 2013](#)).

Så langt er det publisert 56 artikler i [PsyKTestBarn](#), og over 100 tester venter på vurdering. En følge av dette arbeidet er at artiklene samlet sett avdekker hvor lite informasjon det i mange tilfeller finnes om viktige kvalitetsaspekter ved den norske versjonen av testen: 69 % av artiklene viste at det ikke forelå norske normer for den aktuelle testen, 42 % at det ikke fantes norske studier vedrørende testens reliabilitet og 44 % at det ikke fantes norske studier vedrørende validitet.

### **Vi kan lære av Nederland**

Våren 2018 reiste redaksjonen i PsyKTestBarn til Nederland for å besøke en nederlandsk komite med et liknende formål som vårt: «COTAN (Commissie Testaangelegenheiten Nederland) – komite for testsaker i Nederland). Hensikten med møtet var å sammenlikne våre arbeidsmåter, lære av hverandre og diskutere mulig framtidig samarbeid. Det ble et lærerikt og inspirerende møte.

COTAN ble dannet i 1959 og publiserte sine første vurderinger av tester i 1969 ([Evers, Sijtsma, Lucassen & Meijer, 2010](#)). Opp gjennom årene har de sett en betydelig økning i testkvaliteten på testene de vurderer. Testutviklere og utgivere har også i utviklingsfasen i økende grad tatt hensyn til COTANs kriterier for å vurdere tester. Man kan konkludere med at testvurderingene som COTAN publiserer, har hatt en positiv effekt på feltet. De påpekte imidlertid at den økte kvalitetshevingen ikke nødvendigvis gjaldt testbrukere. Det vil si at testbrukere ikke alltid tenkte på testkvalitet når de skulle velge en test. De var kanskje mer opptatt av praktiske overveielser, som f.eks. hvorvidt en test ble hyppig brukt av kollegaer eller var lett tilgjengelig og gratis.

Dette er også noe som har blitt påpekt i en nylig publisert systematisk oversikt over tester brukt i norske skoler ([Arnesen, Braeken, Ogdén & Melby-Lervåg, 2018](#)), der testbrukere ikke alltid er klar over hva slags kvalitetsaspekter som er relevante for en test.

Hva trenges for å øke kvaliteten på tester brukt i Norge og for å gjøre test brukerne mer bevisst på hva som er egnet bruk av tester? Våre kollegaer i COTAN har opplevd at retningslinjer fra myndighetene kan ha en stor innflytelse i så henseende. Da

Kunnskapsdepartementet i Nederland innførte krav om at tester som brukes for å kartlegge barn med spesielle behov, må være kvalitetssikret av COTAN ([The Dutch Ministry of Education Culture and Science, 2012](#)), fikk dette viktige følger: Økt bruk av de testene som var av høy nok kvalitet, større bevissthet hos brukere angående testkvalitet når man skulle velge en test, og på sikt også forbedring av tester som ikke holdt høy nok kvalitet.

## LES OGSÅ: – Norge er et «verstingland» på psykologiske tester

### Still strengere krav til tester

En viktig forskjell mellom COTAN og PsykTestBarn er at COTAN vurderer tester som blir innmeldt av testutgiveren. Det er testutgiveren (eller forfatteren) som ber COTAN gjennomføre en vurdering basert på innlevert dokumentasjon. Dette betyr at testen får karakter basert på det COTAN anser som de viktigste psykometriske egenskapene (kvalitetsaspekter) og dokumentasjonen testutgiver har lagt ved.

PsykTestBarn, derimot, publiserer artikler som er basert på systematiske litteratursøk, der dokumentasjonen som belyser den psykometriske kvaliteten, blir presentert og diskutert. PsykTestBarn oppsummerer ikke vurderingen i form av en karakter, men formulerer en skriftlig konklusjon der status for den norske versjonens psykometriske egenskaper er beskrevet og vurdert. Det er mulig for brukerne å foreslå tester til vurdering, men det er PsykTestBarn som selv avgjør hvilke tester som skal undersøkes. Som initiativtaker er PsykTestBarn mer uavhengig i sin prosedyre.

Ved å publisere artikler som oppsummerer måleegenskaper ved tester som blir brukt for barn og ungdom i Norge, blir det synlig hva slags empiriske studier vi trenger. Det kreves antakelig mer enn informasjon for å motivere testforlag og testbrukere til å verdsette testkvalitet i høyere grad og få til en endring av praksis. Her spiller de ulike tjenestene, beslutningstakerne og myndighetene en viktig rolle.

---

Det bør stilles strengere kvalitetskrav til tester som brukes på barn og unge i Norge.

---

Vi mener det ikke er akseptabelt å basere viktige avgjørelser på testskårer som ikke er kvalitetssikret. Vi oppfordrer testbrukere og myndighetene til å stille kvalitetskrav når det gjelder tester som skal anvendes på barn og unge.

### Kilder

Arnesen, A., Braeken, J., Ogden, T. & Melby-Lervåg, M. (2018). Assessing children's social functioning and reading proficiency: A systematic review of the quality of

educational assessment instruments used in Norwegian elementary schools. *Scandinavian Journal of Educational Research*, 1–26.  
doi:[10.1080/00313831.2017.1420685](https://doi.org/10.1080/00313831.2017.1420685)

European Federation of Psychologists' Associations, EFPA (2013). EFPA Review model for the description and evaluation of psychological and educational tests: Test review form and notes for reviewers. Version 4.2.6. Hentet fra <http://www.efpa.eu/download/650d0d4ecd407a51139ca44ee704fda4>

Evers, A., Sijtsma, K., Lucassen, W. & Meijer, R. R. (2010). The Dutch review process for evaluating the quality of psychological tests: History, procedure, and results. *International Journal of Testing*, 10(4), 295–317.  
doi:[10.1080/15305058.2010.518325](https://doi.org/10.1080/15305058.2010.518325)

Kornør, H., Jozefiak, T., Hanssen-Bauer, K. & Hysing, M. (2015). Forsvarlig testbruk. *Tidsskrift for Norsk psykologforening*, 52(7), 600–601. Hentet fra <http://www.psykologtidsskriftet.no/pdf/2015/600-601.pdf>

Lindgren, R. M. B. (2018, 24. februar). Testing i skolen: ditt barn blir feilvurdert i skole og barnehage. *Dagbladet Meninger*. Hentet fra <https://www.dagbladet.no/kultur/ditt-barn-blir-feilvurdert-i-skole-og-barnehage/69549367>

Strand, N. (2018). Sjekk om testen er god. *Tidsskrift for Norsk psykologforening*, 55(4), 254–257. Hentet fra [http://www.psykologtidsskriftet.no/index.php?seks\\_id=508514&a=3](http://www.psykologtidsskriftet.no/index.php?seks_id=508514&a=3)

The Dutch Ministry of Education Culture and Science. (2012). Regeling screenings-en testinstrumenten LWOO en PrO schooljaar 2012–2013. Hentet fra <http://wetten.overheid.nl/BWBR0030361/2011-08-20>

Øverli, L. (2013, 27. juni). Utdanningsforbundet: Altfor mye testing i skolen. *Aftenposten*. Hentet fra <https://www.aftenposten.no/norge/i/BRR47/Utdanningsforbundet-Altfor-mye-testing-i-skolen>