

Ergoterapi analyse og Vurdering av Aktivitet-EVA konstruktvaliditet og interrater reliabilitet

EN METODESTUDIE

Av Lene Bobakk, Kristine Risum, Kjersti Aabel Tvedte & Helen Bull

Abstract

Aim: The aim of this study was to investigate the psychometric properties of the Norwegian Ergoterapi analyse og Vurdering av Aktivitet (EVA) using data from the Children's Cooking Task (CCT) to assess the construct validity of EVA process skills and the interrater reliability of EVA motor- and EVA process skills.

Method: Data were video recordings of 30 children aged 10 to 17 with acquired brain injury who had performed the CCT in a hospital setting. Construct validity was assessed through hypothesis testing, comparing EVA process skills ratings to CCT executive cognitive function ratings. Interrater reliability was investigated by two Occupational therapists' ratings of the 30 CCT performances.

Results: The hypothesis testing indicated good construct validity of EVA process skills compared to CCT. EVA process skills had a weighted kappa value of 0.78 (95 percent CI 0.63-0.92), indicating moderate-to-excellent interrater reliability. EVA motor skills had a weighted kappa value of 0.59 (CI 0.39-0.80), indicating moderate to good interrater reliability.

Conclusion: The study demonstrated good construct validity of EVA process skills compared to CCT executive cognitive function, excellent interrater reliability on EVA process skills and good to moderate interrater reliability on EVA motor skills for children aged 10-17 with acquired brain injury.

Keywords: Ergoterapi analyse og vurdering av aktivitet-EVA, psychometric properties, Construct validity, interrater reliabilitet, activity



Lene Bobakk er ergoterapispesialist i barns helse, Msc, og arbeider ved Seksjon for ortopedisk rehabilitering, Ortopedisk klinikk, Oslo universitetssykehus HF, Rikshospitalet. E-post: lenbob@ous-hf.no.



Kristine Risum er spesialist i barnefysioterapi, PhD, ved Ortopedisk klinikk, Oslo universitetssykehus HF, Rikshospitalet og førsteamanuensis ved Institutt for rehabiliteringsvitenskap og helsevitenskap, Fakultet for helsevitenskap, OsloMet-Storbyuniversitetet.



Kjersti Aabel Tvedte er spesialergoterapeut ved Seksjon for hodeskade, Avdeling for traumatiske hodeskader, Sunnaas sykehus HF.



Helen Bull er førsteamanuensis ved Avdeling for ergoterapi, Institutt for rehabiliteringsvitenskap og helsevitenskap, Fakultet for helsevitenskap, OsloMet-Storbyuniversitetet.

Introduksjon

I denne studien vurderer vi noen av de psykometriske egenskapene til Ergoterapi analyse og Vurdering av Aktivitet (EVA). I dagens helsevesen stilles det strenge krav til effektivitet, dokumentasjon og kunnskapsbasert praksis (Drageset & Ellingsen, 2010; Helsebiblioteket, 2016). Bruk av validerte kartleggingsredskap bidrar til kvalitet i pasientbehandlingen, da kartleggingsredskap med gode psykometriske egenskaper gir pålitelig grunnlag for å planlegge, gjennomføre og evaluere den behandlingen som gis (Mokkink et al., 2016).

Ergoterapeuter ser på samspillet mellom menneske, aktivitet og omgivelser (MAO) og bruker redskap som kartlegger mennesker i aktivitet, uavhengig av diagnose, alder og type aktivitet (Ergoterapeutene, 2017). Utførelsen av aktiviteter vurderes med tanke på hva som kreves av motoriske-, prosess- og kommunikasjonsferdigheter, og hva som kreves av omgivelsene (Ellingham et al., 2025). Begrepet *Menneske* inkluderer personlige faktorer som motivasjon, kroppsfunksjoner og ferdigheter som er relatert til den aktuelle aktiviteten. *Aktivitet* er de observerbare handlingene i aktiviteten, og hva aktiviteten krever av personen. *Omgivelse* er det fysiske rommet der aktiviteten finner sted (Bass et al., 2024). Metodisk bruk av aktivitet i kartlegging, implementering og evaluering forutsetter en systematisk analyse av de aktuelle aktivitetene (Ellingham et al., 2025). Ergoterapi analyse og Vurdering av Aktivitet (EVA) er et ikke-validert kartleggingsredskap som ivaretar ergoterapeuters perspektiv på aktivitet (Ellingham et al., 2025). EVA er tilgjengelig på EVA- blogg med instruksjonsvideoer, manualer og skjemaer, og krever ikke sertifisering (Ellingham et al., 2025; Eva-blogg, 2025).

EVA ble utviklet til bruk i undervisningen av aktivitetsanalyse for ergoterapeutstudenter ved Høgskolen i Oslo (nå OsloMet) (Ellingham et al., 2025). I 2011–2012 ble det utført en klinisk anvendbarhetsstudie innenfor somatikk, psykiatri og primærhelsetjenesten, hvor EVA ble vurdert som et nyttig redskap for å observere og analysere aktiviteter i de ulike kliniske fagområdene (Clark et al., 2014). EVA bidro til å tydeliggjøre ergoterapeuters aktivitetsperspektiv, styrke felles begrepsbruk og dokumentasjon, fremme rapportering og kommunikasjon i tverrfaglig team samt å kunne bidra til å øke etterspørsel etter ergoterapi (Clark et al., 2014). Vurdering av validitet og reliabilitet av EVA er et nødvendig skritt på veien til å kvalitetssikre våre er-

goterapeutiske redskap. Validitet kan defineres som at et instrument virkelig måler det som det er ment å måle, mens reliabilitet handler om hvor stabile målingene er (De Vet et al., 2011).

Vi har validert EVA mot en kjøkkenaktivitet, Children's Cooking Task (CCT). CCT er en standardisert, aktivitetsbasert test som vurderer eksekutive kognitive funksjoner hos barn (Chevignard et al., 2009). Eksekutive funksjoner blir beskrevet som en samling funksjoner som gjør det mulig for den enkelte å delta i aktiviteter på en effektiv og hensiktsmessig måte, med kompleks, meningsfull og målrettet problemløsning (Chevignard et al., 2010). Å undersøke eksekutive funksjoner hos barn har vist seg å være utfordrende. Nevropsykologiske skrivebords-tester har blitt adaptert fra studier med voksne pasienter, og det stilles spørsmål ved testenenes økologiske validitet. Økologisk validitet referer til i hvilken grad aktivitetsutførelsen på standardiserte tester samsvarer med den faktiske utførelse i naturlige omgivelser. Testsituasjoner bør ligne på naturlig oppførsel for å kunne forutsi fungering i dagligdagse aktiviteter (Barkley, 1991; Chaytor et al., 2006; Franzen & Wilhelm, 1996). Både EVA og CCT er aktivitetsbaserte kartleggingsredskaper, hvor man vurderer barnet mens de utfører en aktivitet. Barn med ervervet hjerne-skade opplever ofte utfordringer med eksekutive kognitive funksjoner som fører til aktivitetsutfordringer i hverdagen. Barna kan ha utfordringer med å planlegge og gjennomføre aktiviteter, noe som gir dem utfordringer både med tanke på ADL-aktiviteter og aktiviteter i skolen (Chevignard et al., 2009; Finnanger et al., 2022).

De eksekutive kognitive funksjonene i CCT og prosessferdighetene i EVA baserer seg på vurdering av mange av de samme ferdighetene:

1. om barna utfører aktiviteten på en effektiv, trygg og selvstendig måte uten behov for intervensjon fra en voksen
2. om barnet velger det utstyret/redskapene som er satt fram, om redskapene brukes slik som de er tiltenkt, og om de håndteres på en trygg og sikker måte
3. hvordan barnet søker informasjon underveis, om barnet leser oppskriften
4. om barnet følger riktig rekkefølge, initierer alle stegene underveis, og om barnet har vansker med å avslutte noen av stegene i oppgaven
5. om barnet holder seg til oppgaven og fullfører den slik som den er tiltenkt (Chevignard et al., 2009; Ellingham et al., 2025)

Disse likhetstrekkene gir et godt grunnlag for sammenligning av EVA og CCT for å vurdere konstruktvaliditet for prosessferdighetene i EVA. Da CCT kun vurderer eksekutiv kognitiv funksjon hos barn og ikke motoriske eller kommunikasjonsferdigheter er testen kun egnet for å validere prosessferdigheter i EVA.

Hovedmålet med studien var å undersøke konstruktvaliditet på EVA prosessferdigheter opp mot de aktivitetsbaserte målene på eksekutive kognitive funksjoner i CCT. Konstruktvaliditet viser i hvilken grad skårene fra måleinstrument vi undersøker samsvarer med skår på andre instrumenter som måler det samme konstruktet (De Vet et al., 2011). Videre vurderte vi interrater reliabilitet på EVA motoriske ferdigheter og EVA prosessferdigheter ved å skåre EVA motorikk- og prosessferdigheter på gjennomføringen av CCT. Interrater reliabilitet vurderes ved at to eller flere personer utfører målinger ved bruk av samme måleinstrument på en pasient (De Vet et al., 2011).

Metode

STUDIEDESIGN OG DELTAGERE

Studien var en del av en randomisert kontrollert studie (CORE pABI) som blant annet brukte CCT for å måle eksekutiv kognitiv fungering hos barn med ervervet hjerneskade (Brandt et al., 2021; Hypher et al., 2019). Deltagerne i vår studie var 30 barn med ervervet hjerneskade som deltok i CORE pABI-studien. Barna ble inkludert på grunnlag av foreldrerapporterte utfordringer med eksekutive kognitive funksjoner, minst 12 måneder siden skade, sykdom og siste kreftbehandling (Brandt et al., 2021; Hypher et al., 2019). Skriftlig informert samtykke ble signert av deltagere over 16 år og av foresatte til deltakere under 16 år. CORE pABI-studien er godkjent av regional komité for medisinsk og helsefaglig forskningsetikk (REK) i Norge (2017/772/REK). Endringsmelding til REK godkjente bruk av filmene fra CORE pABI-studien også i vår studie (Finnanger et al., 2022).

INSTRUMENTENE

Ergoterapi analyse og vurdering av aktivitet - EVA

EVA systemet består av flere skjemaer og brukes til analyse av ulike deler av samspillet mellom menneske, aktivitet og omgivelser (Ellingham et al., 2025). Ved bruk av EVA vurderes motoriske-, prosess- eller kommunikasjonsferdigheter i en aktivitet. Kommuni-

kasjonsferdigheter var ikke aktuelle i denne studien og beskrives ikke i artikkelen. Motoriske ferdigheter klassifiseres som en observerbar handling når en person beveger kroppen og/eller flytter og bruker gjenstander i en aktivitet. Prosessferdigheter klassifiseres som utførelsesferdigheter som en person bruker for å strukturere og styre aktivitetsutførelsen (Ellingham et al., 2025). Motoriske ferdigheter vurderes gjennom 16 ferdigheter, gruppert under fire temaer: «kroppens stilling», «få tak i og holde gjenstander», «forflytte seg selv og gjenstander» og «opprettholde utførelse».

Prosessferdigheter vurderes gjennom 21 ferdigheter, gruppert under fem temaer: «opprettholde utførelse», «anvende kunnskap», «organisering av tid», «organisering av rom og gjenstander» og «tilpasse utførelsen». Under hver ferdighet graderes kvaliteten av aktivitetsutførelsen som ergoterapeuten har observert. Det settes et kryss på en linje som en indikasjon på vurderingen av utførelsesferdighet. Ferdighetene blir vurdert ut fra grad av selvstendighet, trygghet og/eller effektivitet som personen viser i aktivitetsutførelsen (Ellingham et al., 2025).

Etter å ha vurdert de enkelte ferdighetene vurderes aktivitetsutførelsen samlet. Også her brukes graderingsskalaen, i tillegg til en skriftlig oppsummering, hvor både ressurser og utfordringer registreres. I oppsummeringen inkluderes også hypoteser eller kunnskap om bakenforliggende årsaker til eventuelle utfordringer som personen opplever i aktivitetsutførelsen (Ellingham et al., 2025).

Children's Cooking Task - CCT

CCT er en standardisert, aktivitetsbasert test som vurderer eksekutiv kognitiv funksjon hos barn. CCT ble utviklet som en økologisk valid standardisert test som undersøkte eksekutive kognitive funksjoner hos barn gjennom aktivitet fremfor gjennom tradisjonelle skrivebordstester (Chevignard et al., 2010; Chevignard et al., 2009). Tidligere studier, inkludert en norsk studie har vist at CCT har god konstruktvaliditet vurdert opp mot valide og reliable instrumenter på eksekutiv funksjon hos barn med ervervet hjerneskade, og god interrater reliabilitet (Chevignard et al., 2010; Chevignard et al., 2009; Finnanger et al., 2022). I gjennomføringen av CCT får barnet i oppgave å *bake en sjokoladecake* og å *lage en fruktdrikk*. Barnet får muntlig instruksjon i tillegg til skriftlig veiledning og oppskrifter som er tilgjengelige gjennom hele testen. Barnet blir instruert i å oppføre seg som om det er alene og kan

1. Det er negativ korrelasjon mellom totalskår på EVA prosessferdigheter og totalt antall feil på CCT	Vi antar at barn med høy totalskår feil på CCT vil oppnå en lavere total skår på EVA prosessferdigheter
2. Det er negativ korrelasjon mellom totalskår på EVA prosessferdigheter og totalt antall feil på nevropsykologisk klassifikasjon på CCT	Vi antar at barn med høy totalskår på antall feil på nevropsykologisk klassifikasjon på CCT, vil oppnå en lavere total skår på EVA prosessferdigheter
3. Barna som bruker lengre tid på CCT oppnår en lavere totalskår på EVA prosessferdigheter	Vi antar at desto lengre tid man bruker på oppgaven desto større rom blir det for at man gjør feil, og derfor oppnår en lavere totalskår på EVA prosessferdigheter
4. Det er høyere på totalskår på EVA prosessferdigheter om barnet har fått JA på måloppnåelse på CCT	Vi antar at barn som har oppnådd målet med oppgaven vil oppnå høyere totalskår på EVA prosessferdigheter enn de som ikke oppnådde målet
75 prosent av hypotesene måtte bekreftes for å kunne si at EVA prosessferdigheter har god konstruktvaliditet. Korrelasjonskoeffisient > 0.3 måtte oppnås for at hypotesene kunne verifiseres, p-verdi < 0.05 for hypoteser analysert med Mann-Whitney U test EVA = Ergoterapianalyse og Vurdering av Aktivitet CCT = Children's Cooking Task	

Tabell 1: Predefinerte hypoteser for konstruktiv validitet av EVA prosessferdigheter.

bare be om hjelp hvis det står i oppskriften. Skåringen av CCT består av en kvantitativ og en kvalitativ del. I den kvantitative delen telles det hvor mange feil barnet gjør underveis. En feil kan for eksempel være at barnet bytter om på stegene i oppskriften. Feilene barnet gjør klassifiseres både som deskriptive og som nevropsykologiske feil. Deskriptive vil si at man bare beskriver feilene barnet gjør uten å klassifisere feilene. Nevropsykologiske feil registrerer mulig underliggende årsak til at feilen oppstår. Hvis barnet eksempelvis bytter om på stegene i oppskriften (har sukker i bollen før egg) vil den deskriptive beskrivelsen av feilen være feil i rekkefølge mens den nevropsykologiske klassifiseringen vil være en kontrollfeil. Feilene summeres i en totalskår. Høyere totalskår indikerer at barnet har større utfordringer med eksekutiv kognitiv funksjon.

I den kvalitative delen av skåringen vurderes det om de hadde erfaring med å lage mat (ja eller nei), om barnet har oppnådd målet med oppgaven; om de har bakt en sjokoladecake og laget en fruktdrikk (ja eller nei), om det oppsto noen farlig oppførsel underveis, som at barnet prøvde å ta kaken ut av ovnene selv (ja eller nei), samt hvor lang tid barnet brukte på oppgaven (Chevignard et al., 2010).

DATAINNSAMLING

De 30 inkluderte barna med ervervet hjerneskade ble filmet mens de gjennomførte CCT på ergoterapiavdelingens kjøkken på Oslo universitetssykehus. Alle videoene var skåret på antall feil på CCT, bruk av tid og måloppnåelse. Alle videoene av CCT ble deretter skåret på EVA motoriske- og prosessferdigheter av to uavhengige skårere. Begge skårerne var

ergoterapeuter som hadde kjennskap til EVA fra tidligere, men noe ulik erfaring med bruk i klinisk praksis. Den ene (skårer A) var prosjektmedarbeider i CORE pABI-studien og hadde møtt barna som deltok i studien, vært til stede under CCT-testen, i tillegg til å ha sett videoene i etterkant for skåring av CCT. Den andre (skårer B) hadde fått undervisning om EVA som en del av ergoterapistudiet på OsloMet, og hadde klinisk erfaring med bruk av EVA på voksne pasienter med hjerneblødning. Begge ergoterapeutene hadde erfaring med bruk av ulike standardiserte kartleggingsredskaper i klinisk praksis.

Prosedyre/gjennomføring

For vurdering av konstruktvaliditet og interrater reliabilitet har vi fulgt The original Consensus-based Standard for the selection of health Measurement Instruments (COSMIN), et ledende miljø innen helsefaglige metodestudier (De Vet et al., 2011).

KONSTRUKTVALIDITET

EVA prosessferdigheter ble vurdert mot CCT. I forkant av analysene utarbeidet vi fire hypoteser basert på hele skåringskjemaet til CCT, med skåring på deskriptive og nevropsykologiske feil, måloppnåelse og total tid (Tabell 1).

I tråd med COSMINs anbefalinger måtte 75 prosent av hypotesene bekreftes for å oppnå god konstruktvaliditet, med en korrelasjonskoeffisient på > 0.3 , samt p-verdi < 0.05 for hypoteser analysert med Mann-Whitney U test (De Vet et al., 2011).

INTERRATER RELIABILITET

Interrater reliabilitet sier noe om hvorvidt variasjon i

en testskår skyldes forandringer hos pasienten eller at det er en annen ergoterapeut som har utført skåringen (Laver Fawcett, 2014). Jo likere man skårer jo høyere interrater reliabilitet har testen (De Vet et al., 2011). Dette er relevant i klinisk praksis, når det ikke bestandig er den samme ergoterapeuten som vurderer pasienten.

For å vurdere interrater reliabiliteten av EVA motoriske- og prosessferdigheter brukte vi filmene av barna hvor de gjennomførte CCT. Barnas aktivitet utførelse i aktivitetene *å bake en sjokoladecake* og *å lage en fruktdrikk* ble vurdert ved bruk av EVA. Standardiseringen i CCT påvirket aktivitetsformen. Barna bakte ikke bare en kake og lagde en fruktdrikk slik som de vanligvis ville ha gjort det, men de hadde i forkant fått instruksjoner i henhold til gjennomføringen av CCT. Som første steg arbeidet skårerne i fellesskap fram en trinnanalyse av aktivitetsformen *å bake en sjokoladecake* og *å lage en fruktdrikk*. Dette ble gjort for å ha en felles forståelse av hvilke trinn aktivitetene inneholder, og hvilke ferdigheter den krever.

Skårerne gjorde individuelle skåringer av den filmede aktivitet utførelsen hos 30 barn. Skåringene ble gjennomført på forskjellige tidspunkt, og for å sikre uavhengige skåringer ble resultatene ikke diskutert under skåringen. Ved bruk av videomateriale vet vi at skårerne ser samme gjennomføring av aktiviteten, variasjonene i vurderingene skyldes da skårerne (De Vet et al., 2011). Skårerne vurderte EVA prosessferdigheter og EVA motoriske ferdigheter hos barna. Ved gradering av ferdighetene kan krysset ifølge testmanualen til EVA plasseres fritt på linjen mellom 0 og 1, men i vår studie valgte vi å

sette kryssene som vist i Tabell 2. Dette ble gjort for å få et tallmateriale som kunne benyttes i statistiske analyser.

Vi valgte å omgjøre det slik:

0 = Helt uselvstendig/utrygg/ineffektiv, eventuelt behov for å gripe inn

1 = Uselvstendig/utrygg/ineffektiv, eventuelt behov for hjelp

2 = Delvis ineffektiv/uselvstendig/utrygg

3 = Gjennomfører, men så vidt selvstendig/trygt/effektivt

4 = Mindre selvstendig, trygg og effektiv, noen svakheter

5 = Helt selvstendig, trygg og effektiv

STATISTISKE ANALYSER

De statistiske analysene ble gjennomført i IBM SPSS, versjon 27.0 (Pallant, 2010). Da variablene er kategoriale og tallmaterialet ikke er normalfordelt ble ikke-parametriske tester benyttet.

For å vurdere konstruktvaliditeten ble median av totalskår fra skårer A og B på EVA prosessferdigheter brukt i analysene. Spearman's korrelasjonskoeffisient ble brukt for å vurdere korrelasjon mellom total skår på EVA prosessferdigheter og hypotesene: 1) totalt antall feil på CCT, 2) totalt antall feil klassifisert nevropsykologisk på CCT og 3) hvor lang tid barna brukte på CCT. Begrunnelsen for valg av hypotesene er beskrevet i Tabell 1.

For å vurdere grad av korrelasjon tok vi utgangspunkt i Cohens vurderingsskala: rho = .10-.29 lav korrelasjon, rho = .30-.49 moderat korrelasjon, rho = .50-1.0 høy korrelasjon (Pallant, 2010). Mann-Whitney U test ble brukt for hypotese 4 for å undersøke forskjeller i totalskår på EVA prosessferdigheter mellom barna som hadde og ikke hadde måloppnåelse på CCT. Interrater reliabiliteten av EVA ble vurdert med Weighted kappa med 95 prosent konfidens

	Tallverdi	Forklarende tekst fra EVA
0 ——— ——— x 1	5	Helt selvstendig, trygg og effektiv
0 ——— x ——— 1	4	Mindre selvstendig, trygg og effektiv, noen svakheter
0 ——— x ——— 1	3	Gjennomfører, men så vidt selvstendig/trygt/effektivt
0 ——— x ——— 1	2	Delvis ineffektiv/uselvstendig/utrygg
0 — x ——— ——— 1	1	Uselvstendig/utrygg/ineffektiv, eventuelt behov for hjelp
0 x ——— ——— 1	0	Helt uselvstendig/utrygg/ineffektiv, evt. behov for å gripe inn
Ifølge testmanualen til EVA plasseres x fritt på linjen mellom 0 og 1. For å få et tallmateriale som kunne benyttes i statistiske analyser, satte skårene kryssene som vist i tabellen. EVA = Ergoterapianalyse og Vurdering av Aktivitet		

Tabell 2: Gradering av aktivitet utførelse i EVA.

	N=30
Gutt/jente n (%)	14/16 (46.7/53.3)
Alder ved testtidspunkt Gjennomsnitt (SD)	12.7 (2.2)
Ervervet hjerneskade n (%)	
- Hjerneturmor	19 (63.3)
- Traumatisk hodeskade	2 (6.7)
- Anoksi	2 (6.7)
- Blødning/infarkt	5 (16.7)
- Hjernebetennelse	2 (6.7)
SD standardavvik	

Tabell 3: Demografiske karakteristikk av deltagerne.

	Median (25th-75th percentil)	Minimum-maksimum
<i>Cooking habits</i>		
Has already cooked yes/no, n (%)	30 (100)	
Total time required to complete the task, minutes	34 (25-38)	22-66
<i>Errors</i>		
Total errors	26 (16-38)	8-101
Total errors without comments and questions	18 (14-24)	5-59
<i>Descriptive level</i>		
Additions	6 (4-10)	0-34
Omissions	5 (2-8)	0-18
Commentaries and questions	7 (3-10)	0-70
Substitutions-Sequence errors	2 (1-7)	0-16
Estimation errors	3 (1-4)	0-6
<i>Neuropsychological classification</i>		
Control errors	9 (6-12)	1-30
Context neglect	9 (5-12)	2-72
Environmental adherence	2 (1-4)	0-26
Purposeless actions and displacements	3 (2-6)	0-14
Dependency	3 (1-5)	0-23
Inappropriate behavior	0 (0-1)	0-3
Prospective memory score (1-15)	14 (12-15)	8-15
<i>Qualitative rating (yes/no)</i>		
Goal achieved, n (%)	23 (77)	
Occurrence of dangerous behavior, n (%)	5 (17)	
Need for an adult intervention, in order to avoid a complete failure, or following a sixth request for help, n (%)	10 (33)	
Number of forgotten ingredients	0 (0-0)	0-2
Number of supplementary ingredients used	0 (0-0)	0-0
Number of omitted or partially omitted steps	2 (1-3)	0-8
Number of inversions in recipe steps	1 (0-3)	0-8
Number of forgotten dirty utensils (not put in the sink) before cueing	0 (0-2)	0-5
Number of forgotten waste items (not put in the rubbish bin) before cueing	0 (0-1)	0-3
Da skåringskjema på CCT ikke var oversatt til norsk i hovedstudien, har vi valgt å bruke de engelske klassifiseringene i denne tabellen. EVA = Ergoterapianalyse og Vurdering av Aktivitet CCT = Children's Cooking Task		

Tabell 4: Skåringene på Children's Cooking task (CCT), inkludert totalskår på kvalitative, deskriptive og nevropsykologiske vurderinger.

intervall (KI). Weighted kappa ble kalkulert for total skår på EVA prosessferdigheter og EVA motoriske ferdigheter. Weighted kappa vurderer ikke bare om skårene er uenige, men også graden av uenighet. Uenighet mellom kategorier som er nær hverandre, vektet som mindre alvorlig enn ved kategorier som er lengre fra hverandre (De Vet et al., 2011). Totalskår ble angitt etter en gradering på skala fra 0-5. Det finnes ingen standarder på hvilke verdier som angir akseptabel reliabilitet, bare guidelines. De verdiene som ligger til grunn for tolkningen av Weighted kappa i denne studien følger COSMINs anbefalinger: < 0.4 indikerer lav reliabilitet, 0.4 til 0.75 indikerer moderat til god reliabilitet, > 0.75 utmerket reliabilitet (De Vet et al., 2011).

Resultat

De demografiske karakteristikkene av deltagerne er presentert i tabell 3. Det var en relativt jevn fordeling mellom kjønn, selv om det var noen flere jenter enn gutter. Gjennomsnittsalderen på barna ved testtidspunktet var 12.7 år. Det var flest barn med hjernetumor. Resultater fra CCT er oppsummert i tabell 4. Alle barna hadde erfaring med å lage mat fra før. De brukte gjennomsnittlig 34 minutter på å gjennomføre aktiviteten og gjennomsnitt på totalt antall feil var 24.

KONSTRUKTVALIDITET

Alle de predefinerte hypotesene ble bekreftet. Det var negativ korrelasjon mellom totalskår på EVA prosessferdigheter og totalt antall feil på CCT (Spearman's rho -0.74, $p < 0.001$). Videre var det negativ korrelasjon mellom totalskår på EVA prosessferdigheter og totalt antall feil på nevropsykologisk klassifisering på CCT (Spearman's rho -0.69, $p < 0.001$). Vi fant også at barna som brukte lengre tid på CCT oppnådde en lavere totalskår på EVA prosessferdigheter (Spearman's rho = -0.59, $p < 0.001$). Det var statistisk signifikant forskjell på totalskår på EVA prosessferdigheter mellom barn som lyktes med oppgaven på CCT ($n=23$) og de som ikke lyktes ($n=7$), median (25-75 percentilen) 4.0 (3.5-5.0) versus 2.0 (0.5-3.0), $p < 0.001$. Dette indikerer god konstruktvaliditet på EVA prosessferdigheter.

INTERRATER RELIABILITET

Interrater reliabilitet på totalskår på EVA prosessferdigheter ble vurdert til moderat til utmerket. Weighted kappa var 0.78 (95 prosent KI 0.63-0.92).

Interrater reliabilitet på totalskår på EVA motoriske ferdigheter ble vurdert til moderat-god med Weighted Kappa på 0.59 (95% KI.39-0.80), men det brede konfidensintervallet introduserer en grad av usikkerhet. Resultatene på interrater reliabilitet på EVA prosessferdigheter og EVA motoriske ferdigheter er presentert i Tabell 5.

Totalskårene på EVA prosessferdigheter og EVA motoriske ferdigheter for de to skårene er presentert i henholdsvis Tabell 6 og 7. Fet skrift viser antall tilfeller skårene har vært helt enige. På EVA prosessferdigheter var skårer A og B enig 14 ganger og uenig 16. Det skilte 1 kategori mellom skårene i 13 av tilfellene og 2 kategorier i 3 av tilfellene. Skårer A har jevnt over vært strengere i sine vurderinger enn skårer B. Hele skalaen er benyttet av skårer A, mens skårer B ikke benyttet skåren 0. Det er en overvekt av skåringer i kategoriene 3-5. Det var absolutt enighet i 47 prosent av tilfellene mellom skårene.

På EVA motoriske ferdigheter var skårene enige 25 ganger og uenig 5 ganger. Det skilte 1 kategori i 4 av tilfellene og 2 kategorier i 1 av tilfellene. I 22 tilfeller har begge skårene benyttet skår 5. Skår 4 og 5 på skalaen ble benyttet flest ganger, men også skår 2 og 3 ble benyttet, henholdsvis 1 og 2 ganger. Det var absolutt enighet i 83 prosent av tilfellene mellom skårene.

Diskusjon

Oss bekjent er dette den første studien som vurderer konstruktvaliditet på EVA prosessferdigheter og interrater reliabilitet av EVA prosessferdigheter og EVA motoriske ferdigheter. Resultatene fra studien viste god konstruktvaliditet for EVA prosessferdigheter vurdert mot CCT, moderat til utmerket interrater reliabilitet for EVA prosessferdigheter og moderat til god interrater reliabilitet for EVA motoriske ferdigheter. En tidligere anvendbarhetsstudie har vist at EVA kan være et nyttig redskap i klinisk

	Weighted Kappa	95 % konfidensintervall
Total skår EVA prosessferdigheter	0.78	0.63-0.92.
Total skår EVA motoriske ferdigheter	0.59	0.39-0.80
EVA = Ergoterapianalyse og Vurdering av Aktivitet		

Tabell 5: Interrater reliabilitet.

Skårer A Totalskår EVA prosessferdigheter							
		0 Helt uselvstendig /utrygg/ineffektiv, evt. behov for å gripe	1 Uselvstendig/ utrygg/ ineffektiv, eventuelt behov for hjelp	2 Delvis ineffektiv/ uselvstendig/ utrygg	3 Gjennomfører, men så vidt selvstendig /trygt/ effektivt	4 Mindre selvstendig, trygg og effektiv, noen svakheter	5 selvstendig trygg og effektiv
Skårer B Totalskår EVA prosessferdigheter	0 Helt uselvstendig /utrygg/ineffektiv, evt. behov for å gripe inn	0	0	0	0	0	0
	1 Uselvstendig/utrygg/ ineffektiv, eventuelt behov for hjelp	2	1	0	0	0	0
	2 Delvis ineffektiv/ uselvstendig/ utrygg	0	0	0	0	0	0
	3 Gjennomfører, men så vidt selvstendig /trygt/ effektivt	0	1	2	1	0	0
	4 Mindre selvstendig, trygg og effektiv, noen svakheter	0	0	1	5	6	2
	5 Selvstendig trygg og effektiv	0	0	0	1	2	6
Det var 46.6% absolutt enighet mellom skårer A og skårer B							
EVA = Ergoterapianalyse og Vurdering av Aktivitet							

Tabell 6: Totalskår EVA prosessferdigheter (skår 0-5). Cellene viser antallet av de ulike kombinasjoner av skår hos de to skårere. Fet skrift på grå bakgrunn viser antall tilfeller med absolutt enighet mellom skårerne. Det var totalt 30 barn som ble vurdert av begge skårere.

Skårer A Totalskår EVA motoriske ferdigheter							
		0 Helt uselvstendig /utrygg/ineffektiv, evt. behov for å gripe	1 Uselvstendig/ utrygg/ ineffektiv, eventuelt behov for hjelp	2 Delvis ineffektiv/ uselvstendig/ utrygg	3 Gjennomfører, men så vidt selvstendig /trygt/ effektivt	4 Mindre selvstendig, trygg og effektiv, noen svakheter	5 selvstendig trygg og effektiv
Skårer B Totalskår EVA motoriske ferdigheter	0 Helt uselvstendig /utrygg/ineffektiv, evt. behov for å gripe inn	0	0	0	0	0	0
	1 Uselvstendig/utrygg/ ineffektiv, eventuelt behov for hjelp	0	0	0	0	0	0
	2 Delvis ineffektiv/ uselvstendig/ utrygg	0	0	0	0	1	0
	3 Gjennomfører, men så vidt selvstendig /trygt/ effektivt	0	0	0	0	0	0
	4 Mindre selvstendig, trygg og effektiv, noen svakheter	0	0	0	2	3	2
	5 Selvstendig trygg og effektiv	0	0	0	0	0	22
Det var 83.3% absolutt enighet mellom skårer A og skårer B							
EVA = Ergoterapianalyse og Vurdering av Aktivitet							

Tabell 7: Totalskår EVA motoriske ferdigheter (skår 0-5). Cellene viser antallet av de ulike kombinasjoner av skår hos de to skårere. Fet skrift på grå bakgrunn viser antall tilfeller med absolutt enighet mellom skårerne. Det var totalt 30 barn som ble vurdert av begge skårere.

praksis i tillegg til i undervisning av studenter (Clark et al., 2014). Vår studie understøtter at EVA også er et valid og reliabelt redskap til bruk i klinisk praksis, selv om videre forskning er nødvendig.

KONSTRUKTVALIDITET

I denne studien har vi vurdert konstruktvaliditet for EVA prosessferdigheter. Vi framsatte fire hypoteser der vi antok at det var sammenheng mellom totalskår på EVA prosessferdigheter og totalskår på CCT deskriptive feil, totalskår CCT nevropsykologiske feil samt tidsbruk på CCT, i tillegg til forskjell mellom total skår på EVA på de med og de uten måloppnåelse på CCT. Ved at vi benyttet flere deler av skåringskjema av CCT i utviklingene av hypotesene, fikk vi god bredde i sammenligning mellom EVA prosessferdigheter og CCT. Dette er en styrke i vurderingen av konstruktvaliditet av EVA prosessferdigheter da alle hypotesene ble verifisert med tilfredsstillende korrelasjon og signifikante resultater.

INTERRATER RELIABILITET

Weighted kappa-analysene viste utmerket interrater reliabilitet på EVA prosessferdigheter og moderat-god på EVA motoriske ferdigheter. Ved bruk av ikke-parametriske tester har man ikke muligheten for å vurdere målefeil. Imidlertid oppgis absolutt enighet mellom skårene (Mokkink et al., 2016), og i vår studie var enighet høyere på EVA motoriske ferdigheter enn EVA prosessferdigheter. Når vi ser nærmere på tallene fra interrater reliabilitet ser vi at på EVA prosessferdigheter var skårene hovedsakelig enige eller at det skilte en kategori, med 3 tilfeller som viste 2 kategorier uenighet. På EVA motoriske ferdigheter var skårene også hovedsakelig enige, med kun et tilfelle der det skilte 2 kategorier. I en klinisk hverdag vil uenighet på 1 kategori få mindre konsekvenser for videre oppfølging av barnet. En uenighet på 2 kategorier viser dårlig stabilitet i vurderingene, og kan få konsekvenser for om adekvate ergoterapeutiske tiltak igangsettes. Weighted kappa vurderer antall ulike svar og avstanden mellom svarene, og forutsetter at det er noe variasjoner i det som måles, men at det er en likhet mellom de som måler. Ved et heterogent utvalg vil skårene som regel fordele seg jevnt utover skåringsskalaen. Ved et homogent utvalg tenderer skåringene å samle seg, og bare deler av skåringsskalaen blir tatt i bruk. Når hele skåringsskalaen tas i bruk blir det enklere å skille både skårene og skårene fra hverandre. Når det er

vanskelig å skille skårene og skåringene fra hverandre kan dette føre til en lavere skår på Weighted Kappa (De Vet et al., 2011).

Inklusjonskriteriene i hovedstudien medførte at barna som deltok i studien hadde ervervet hjerneskade med utfordringer med eksekutiv kognitiv fungering, men stort sett uten motoriske utfall. Dette gjenspeiles i skåringene av EVA motoriske ferdigheter, hvor den store majoriteten skåret beste kategori, mens det var mer variasjon i skårene på EVA prosessferdigheter. Gruppen vi har undersøkt vurderes som en homogen gruppe når det gjelder motoriske ferdigheter. Vurderingen av interrater reliabilitet av EVA motoriske ferdigheter fikk derfor lavere Weighted kappa verdi enn EVA prosessferdigheter, selv om det var større enighet mellom skårene på EVA motoriske ferdigheter enn på EVA prosessferdigheter. Konfidensintervallet på EVA motoriske ferdigheter var også bredt, med en nedre grense som tilsvarer lav reliabilitet til en øvre grense som tilsvarer utmerket reliabilitet. Imidlertid styrker den høye graden av enigheten mellom skårene at EVA motoriske ferdigheter var akseptabel. EVA prosessferdigheter får en høyere kappa verdi fordi man hadde en større spredning i skåringene og hele måleskalaen ble tatt i bruk. Ved vurdering av EVA motoriske ferdigheter ser man at skårene ikke har brukt hele skalaen, og de fleste har fått en totalskår på 4 og 5, noe som gir lite variasjon og gjør det mer utfordrende å skille skårene fra hverandre. Muligens kunne vi oppnådd høyere kappaverdi hvis vi hadde hatt en mer heterogen gruppe og man hadde brukt hele skalaen slik som på EVA prosessferdigheter.

Weighted kappa er kalkulert ut fra totalskår på EVA prosessferdigheter og EVA motoriske ferdigheter. Vi vet derfor ikke hvordan enigheten mellom skårer A og B fordeler seg på enkeltferdighetene i EVA. Undersøkelse av skåring på enkeltferdigheter kan være aktuelt i videre studier.

ØKOLOGISK RELEVANS

EVA tar utgangspunkt i at en aktivitetsform vil variere ut fra konteksten aktiviteten utføres i (Ellingham et al., 2025), noe som tilsier at CCT aktivitet utførelsen vil være påvirket av at barna utførte en standardisert aktivitet på et ukjent kjøkken uten å ha valgt aktiviteten selv. Den standardiserte situasjonen kan redusere den økologiske relevansen, ettersom konteksten avviker fra barnas vante miljø og deres egne aktivitetsvalg. Dette kan potensielt påvirke både motivasjon, mestring og strategibruk, og dermed

gi et ufullstendig bilde av barnas faktiske funksjon i hverdagen.

Å vurdere pasienters funksjon gjennom mer eller mindre standardiserte aktiviteter i en sykehussetting er imidlertid en del av den kliniske hverdagen. Med bakgrunn i at alle barna hadde bakt før og hadde kjennskap til aktiviteten, vurderer vi utfordringene de viste i aktivitetsutførelsen som reelle, og at resultatene fra denne standardiserte aktiviteten er egnet til å gi nyttig informasjon om barnets funksjon.

STYRKER OG BEGRENSNINGER MED STUDIEN

Styrker ved studien er at vi fulgte anbefalingene fra COSMIN for evaluering av konstruktvaliditet og interrater reliabilitet, skårerne var erfarne terapeuter med kjennskap til standardiserte tester og EVA spesifikt, og skårerne hadde også relevant klinisk erfaring innen barneergoterapi. Selv om vi hadde færre barn/deltakere (n=30) enn det som anbefales i slike metodestudier (n>50) har vi imidlertid et relativt høyt antall barn inkludert med en sjelden diagnose. I tillegg utfører barna samme aktivitet som gir en unik mulighet til å vurdere konstruktvaliditet og interrater reliabilitet av EVA.

En annen styrke er at vi for vurderingen av interrater reliabilitet av EVA prosessferdigheter og EVA motoriske ferdigheter brukte videomaterialet fra CCT i CORE pABI studien. Ved bruk av videomaterialet vet vi at skårerne ser samme utførelse av oppgavene, variasjonene i vurderingene skyldes da skårerne (De Vet et al., 2011). Skårer A var også involvert i CORE pABI studien og var fysisk til stede da barna ble filmet og gjennomførte CCT testen, i tillegg til å skåre testen i etterkant. Skårer A sitter derfor inne med en annen type informasjon enn skårer B, som kan være med på å påvirke skåringene i EVA. Når vi har sett nærmere på skåringene fra skårer A og skårer B ser vi at skårer A jevnt over har vært strengere i sine vurderinger. Variasjonene kan skyldes informasjon som skårer A har fått med seg ved å fysisk være til stede som ikke er tilgjengelig for skårer B som bare har sett videomaterialet av aktivitetsutførelsene. Begge skårerne hadde god kjennskap til EVA, mens skårer B hadde mer erfaringer med bruk av EVA i klinisk praksis. Ideelt sett skal man ha så lik erfaring som mulig, men i klinisk praktisk vil det ofte være forskjeller. Enkelte mener at antall skårere er like viktig for resultatet som antall deltagere man inkluderer i en interrater reliabilitetsstudie (Koo & Li, 2016). Det kunne styrket

studien å ha flere skårere, men det var ikke praktisk gjennomførbart.

Generaliserbarheten til studien begrenses av at EVA ble validert mot en gruppe barn med ervervet hjerneskade, og ikke mot en bredere populasjon, samt at barna gjennomførte en standardisert oppgave og ikke en selvvalgt en. Samtidig er det en styrke at barna gjennomførte en kjøkkenaktivitet så selvstendig som mulig etter oppskrift. Dette er en vanlig fremgangsmåte i mange kartleggingssituasjoner, og styrker studiens generaliserbarhet til andre strukturerte aktiviteter.

Konklusjon

Studien viste god konstruktvaliditet for EVA prosessferdigheter målt mot CCT, moderat til utmerket interrater reliabilitet på EVA prosessferdigheter og moderat til god interrater reliabilitet på EVA motoriske ferdigheter for barn med ervervet hjerneskade. Disse resultatene indikerer at EVA er et valid og reliabelt kartleggingsredskap. Det er en styrke at valideringen er gjort mot en spesifikk gruppe. Det vil være behov for å undersøke egenskaper hos EVA med andre brukergrupper og aktiviteter. Det er også behov for validering av EVA motoriske ferdigheter og EVA kommunikasjonsferdigheter.

Finansiering

Prosjektet er støttet av Ergoterapeutenes FoU-midler.

INTERESSEKONFLIKT

Forfatterne erklærer at forskningen ble gjennomført uten noen kommersielle eller økonomiske forhold som kan tolkes som en potensiell interessekonflikt. Ingen av forfatterne har heller vært med i utviklingen av EVA.

Takk til

Vi ønsker å takke barna og foreldrene for deres deltagelse i prosjektet, som gjorde det mulig for oss å gjennomføre denne studien. Takk også til CORE pABI prosjektet for tillatelse til å bruke data.

Referanser

- Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology*, 19(2), 149–178. <https://doi.org/10.1007/BF00909976>

- Bass, J. D., de Sam Lazaro, S. L., Baum, C. M., & Marchant, J. K. (2024). Editorial: The Person-Environment-Occupation-Performance (PEOP) Model—An OTJR Focused Issue. *OTJR: Occupational Therapy Journal of Research*, 44(3), 449–454. <https://doi.org/10.1177/15394492241252578>
- Brandt, A. E., Finnanger, T. G., Hypher, R. E., Rø, T. B., Skovlund, E., Andersson, S., Risnes, K., & Stubberud, J. (2021). Rehabilitation of executive function in chronic paediatric brain injury: a randomized controlled trial. *BMC Medicine*, 19(1), 253. <https://doi.org/10.1186/s12916-021-02129-8>
- Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Arch Clin Neuropsychol*, 21(3), 217–227. <https://doi.org/10.1016/j.acn.2005.12.002>
- Chevignard, M., Catroppa, C., Galvin, J., & Anderson, V. (2010). Development and Evaluation of an Ecological Task to Assess Executive Functioning Post Childhood TBI: The Children's Cooking Task. *Brain Impairment*, 11(2), 125–143. <https://doi.org/10.1375/brim.11.2.125>
- Chevignard, M. P., Servant, V., Mariller, A., Abada, G., Pradat-Diehl, P., & Laurent-Vannier, A. (2009). Assessment of executive functioning in children after TBI with a naturalistic open-ended task: A pilot study. *Developmental Neurorehabilitation*, 12(2), 76–91. <https://doi.org/10.1080/17518420902777019>
- Clark, E. G., Jespersen, L. F., Ellingham, B., & Brovold, T. (2014). Fra skolepulten til kjøkkenbenken , Klinisk anvendbarhet av et pedagogisk utviklet aktivitetsanalyse-system. *Ergoterapeuten*, 1, 26–34.
- De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine, A practical guide*, Cambridge University Press.
- Drageset, S., & Ellingsen, S. (2010). Forståelse av kvantitativ helseforskning - en introduksjon og oversikt. *Nordisk Tidsskrift for Helseforskning*, 5. <https://doi.org/10.7557/14.244>
- Ellingham, B., Jespersen, L. F., & Clark, E. G. (2025). *EVA 2025 Ergoterapi analyse, vurdering og dokumentasjon av aktivitetsutførelse Manual for analyse, vurdering og dokumentasjon av aktivitetsutførelse*. EVA-blogg. Retrieved 25.09 from <https://eva933.wordpress.com/wp-content/uploads/2025/08/eva-2025-manualen-aktivitetsutforelse.pdf>
- Ergoterapeutene (2017). *Alle skal kunne delta - Ergoterapeuters kjernekompetanse - Yrkesetiske retningslinjer* https://ergoterapeutene.sharepoint.com/Arkiv/7-Kommunikasjon%20og%20markedsf%C3%B8ring/75-Trykksaker/Kjernekompetanse%202017/Ergoterapeuters_kjernekompetanse_Web_enkeltsider.pdf?ga=1
- Eva-blogg (2025). *Eva-blogg* Retrieved 25.09. from <https://eva933.wordpress.com/nedlasting-av-eva-dokumenter-og-skjemaer/>
- Finnanger, T. G., Andersson, S., Chevignard, M., Johansen, G. O., Brandt, A. E., Hypher, R. E., Risnes, K., Rø, T. B., & Stubberud, J. (2022). Assessment of Executive Function in Everyday Life—Psychometric Properties of the Norwegian Adaptation of the Children's Cooking Task [Original Research]. *Frontiers in Human Neuroscience, Volume 15 - 2021*. <https://doi.org/10.3389/fnhum.2021.761755>
- Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In *Ecological validity of neuropsychological testing*. (s 91–112). Gr Press/St Lucie Press, Inc.
- Helsebiblioteket (2016). *Kunnskapsbasert praksis*. Retrieved 25.09 from <https://www.helsebiblioteket.no/kunnskapsbasert-praksis>
- Hypher, R. E., Brandt, A. E., Risnes, K., Rø, T. B., Skovlund, E., Andersson, S., Finnanger, T. G., & Stubberud, J. (2019). Paediatric goal management training in patients with acquired brain injury: study protocol for a randomised controlled trial. *BMJ Open*, 9(8), e029273. <https://doi.org/10.1136/bmjopen-2019-029273>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Laver Fawcett, A. (2014). A Routine standardised outcome measurement to evaluate the effectiveness of occupational therapy interventions: essential or optional? *Ergoterapeuten*, 4, 28–37.
- Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., Vet, H. C. W. d., & Terwee, C. B. (2016). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy*, 20.
- Pallant, J. (2010). *SPSS survival manual: a step by step guide to data analysis using SPSS* (4 ed.).