

EVA – Analyse av fysiske omgivers påvirkning på aktivitetsutførelse: Inter-rater reliabilitet ved et generisk vurderingsverktøy

Av Mikkel M. Thørrisen, Lene Fogtmann Jespersen, Brian Ellingham & Ellen Gjertsen Clark

Abstract

EVA – Analysis of the impact of the physical environment on occupational performance: Inter-rater reliability for a generic assessment instrument

Background: Physical aspects of the environment influence occupational performance, and occupational therapists are key in identifying environmental factors that affect performance skills. This study aimed to explore the inter-rater reliability of EVA physical environment, a generic assessment instrument for analyzing the impact of the physical environment on occupational performance.

Methods: In this cross-sectional study, 23 occupational therapists analyzed two occupational performances using EVA physical environment. Inter-rater reliability was estimated using intraclass correlation coefficients from two-way mixed effects consistency models. Analyses were conducted at the instrument level, and separately for different sections and items, with subgroup analyses based on the therapists' clinical field, amount of preparation, and analytical competence.

Results: EVA physical environment demonstrated good to excellent inter-rater reliability across all levels (instrument, section, item). Reliability was highest among well-prepared and competent therapists, and therapists in specialized and somatic health care. Acceptable reliability was also found among less prepared and competent therapists and those in municipal and mental health care.

Conclusions: EVA physical environment demonstrated satisfactory inter-rater reliability for assessing the impact of the physical environment on occupational performance.

Keywords: Activity analysis; Physical environment; Psychometrics; Inter-rater reliability; Occupational performance



Mikkel M. Thørrisen er førsteamanuensis i helsevitenskap ved Fakultet for helsevitenskap, OsloMet – storbyuniversitetet. E-post: mithor@oslo-met.no.



Lene Fogtmann Jespersen er ergoterapistspesialist i somatisk helse ved Oslo Universitetssykehus HF.



Brian Ellingham er universitetslektor ved Fagsenter for seniorer, OsloMet – storbyuniversitetet.



Ellen Gjertsen Clark er spesialergoterapeut ved Diakonhjemmet sykehus.

Forfatterne bekrefter at det ikke er uenigheter med hensyn til rettigheter til artikkelen, og at ingen av forfatterne har interessekonflikter.

Introduksjon

Deltakelse og utførelse av aktiviteter formes gjennom et dynamisk samspill mellom individets forutsetninger, aktivitetens krav og påvirkningen fra omgivelsene (Law et al., 1996). Dette samspillet påvirkes av individets motivasjon, kapasitet, sosiale deltakelse og aktivitetsmønstre, i tillegg til de mulighetene og begrensningene omgivelsene gir (Kielhofner, 2010). Omgivelsene omfatter en rekke aspekter som fysiske, sosiale, psykologiske, temporale, politiske, økonomiske og kulturelle faktorer (Fisher & Marterella, 2019). Særlig viktig er hvordan de fysiske omgivelsene, som romforhold, objekter og klimatiske forhold, påvirker personens ferdigheter som benyttes i aktivitetsutførelsen.

Ergoterapeuter har en nøkkelrolle i å identifisere de enkelte elementene eller faktorene i de umiddelbare omgivelsene hvor en aktivitet utføres som kan påvirke en eller flere utførelsesferdigheter. Ergoterapeuter kartlegger ressurser og utfordringer i samspillet mellom menneske, aktivitet og omgivelser (MAO) (Law et al., 1996) og bistår med tilpassning eller tilrettelegging av elementer og relasjoner innenfor MAO-systemet for å fremme aktivitetsutførelse og deltakelse (Bonsaksen & Ellingham, 2017). Grunnleggende kompetanse i aktivitetsanalyse er sentral for ergoterapeuter i dette arbeidet (Feldborg & Røhl, 2019; Hersch et al., 2005; Thomas, 2023).

Flere konseptuelle modeller fremhever omgivelsenes betydning for aktivitetsutførelse, herunder Model of Human Occupation (MOHO) (Kielhofner, 2010; Taylor, 2017), Canadian Model of Occupational Performance and Engagement (CMOP-E) (Polatajko et al., 2007) og den transaksjonelle modellen for aktivitet (Fisher & Martella, 2019). Disse modellene omtaler gjerne omgivelsene i form av generelle kategorier uten spesifikk terminologi for konkrete aktivitetssituasjoner. Det er også utviklet kartleggingsverktøy for analyse og vurdering av omgivelsene, eksempelvis Housing Enabler (Iwarsson & Slaug, 2001, 2010), Residential Environment Impact Scale (REIS) (Fisher et al., 2014) og Work Environment Impact Scale (WEIS) (Corner et al., 1997; Ellingham & Jakobsen, 2003). Disse verktøyene er imidlertid domenespesifikke, i den forstand at de er knyttet til spesifikke aktivitetskontekster (for eksempel hjem, skole eller arbeid). Det har manglet et generisk kartleggingsverktøy for de fysiske omgivelsene, altså et verktøy som gjør det mulig å analysere og vurdere påvirkning av aktivitetsomgivelsene i aktivitetsutførelse ved hjelp av konkret terminologi, uavhengig av aktivitetskontekst.

EVA (Ergoterapianalyse og Vurdering av Aktivitet) benyttes av ergoterapeuter og ergoterapistudenter ved flere utdanninger i Norge. EVA er et system for systematisk vurdering og dokumentasjon av ulike faktorer i samspillet mellom menneske, aktivitet og omgivelser med hensyn til individets utførelse av aktiviteter (Clark et al., 2014; Jespersen et al., 2014). EVA-systemet består av verktøy (analyseskjemaer) for analyse av aktivitetsform og aktivitetsutførelse i vurdering av trinn og handlingssekvenser, utførelsesferdigheter og kroppsfunksjoner (Ellingham et al., 2023). EVA – Analyse av fysiske omgivelers påvirkning på aktivitetsutførelse (EVA fysiske omgivelser) er et nylig utviklet verktøy som kartlegger de fysiske omgivelsenes påvirkning på aktivitetsutførelse, og som kan benyttes i ulike aktivitetskontekster (Jespersen et al., 2025). I EVA fysiske omgivelser kartlegges til sammen 26 omgivelsesfaktorer innenfor tre omgivelseskategorier, hver faktor graderes på en linje fra hemmende til fremmende, det gjøres en samlet helhetsvurdering på tvers av faktorene og det er rom for kvalitative beskrivelser, som underbygger og eksemplifiserer vurderingene (Jespersen et al., 2025). En studie fra 2024 viste at 16 av 22 ergoterapeuter (73 prosent) opplevde EVA fysiske omgivelser som nyttig eller svært nyttig, og 21 av 22 (95 prosent) mente analyseskjemaet dekker aktuelle og relevante fysiske faktorer i omgivelsene i stor eller svært stor grad (Jespersen et al., 2025).

Selv om det er nødvendig at et kartleggingsverktøy oppleves som nyttig, er det alene ikke tilstrekkelig. Verktøyet må også kunne demonstrere tilfredsstillende måleegenskaper (Gagnier et al., 2021), og inter-rater reliabilitet er særlig viktig for kartleggingsverktøy som benyttes i forbindelse med observasjoner i naturlige kontekster (Mokkink et al., 2023). Reliabilitet handler om i hvilken grad et verktøy er fri for målefeil, mens inter-rater reliabilitet mer spesifikt er relatert til i hvilken grad det er enighet mellom observatører (ratere) som har gjort en vurdering av samme person eller situasjon ved hjelp av samme måleverktøy (Asunta et al., 2019; Gagnier et al., 2021). Et verktøy har høy inter-rater reliabilitet når to eller flere observatører vurderer samme person eller situasjon likt når de bruker verktøyet i observasjon. Motsatt vil verktøyet ha lav inter-rater reliabilitet dersom observatørene vurderer personen eller situasjonen ulikt. Tidligere forskning har undersøkt og funnet moderat til høy inter-rater reliabilitet ved verktøyene Housing Enabler (Helle et al., 2014; Iwarsson et al., 2005), REIS (Harrison et al., 2023)

og WEIS (Kielhofner et al., 1999). Måleegenskaper ved EVA fysiske omgivelser har ikke tidligere blitt undersøkt. Formålet med denne studien var derfor å undersøke inter-rater reliabilitet ved EVA fysiske omgivelser.

Metode

DESIGN

Studien ble gjennomført som en tverrsnittstudie der 23 ergoterapeuter vurderte påvirkning av fysiske omgivelser på to aktivitetsutførelser ved hjelp av verktøyet EVA – analyse av fysiske omgivelser påvirkning på aktivitetsutførelse (EVA fysiske omgivelser).

DELTAKERE

Deltakerne ble rekruttert blant ergoterapeuter som har gjennomført EVA-kurs, blant deltakere på en EVA-workshop på fagkongress i ergoterapi i Stavanger 2022 samt via åpen påmelding distribuert gjennom sosiale medier og EVA-bloggen (Ellingham et al., u.å.). Til sammen 23 ergoterapeuter deltok i studien. Kjennetegn ved deltakerne er presentert i Tabell 1.

Ergoterapeutene i utvalget hadde betydelig klinisk erfaring (M = 14 år), og ergoterapeuter fra både spesialisthelsetjeneste (n = 14; 61 prosent) og kommunehelsetjeneste (n = 8; 35 prosent) var representert. De fleste ergoterapeutene jobbet innenfor somatisk helse (n = 18; 78 prosent).

INSTRUMENT: EVA FYSISKE OMGIVELSER ANALYSESKJEMA

Kjernen i studien var at deltakerne, uavhengig av hverandre, skulle vurdere de fysiske omgivelsernes påvirkning på aktivitetsutførelse ved hjelp av analyseskjemaet EVA fysiske omgivelser. I analyseskjemaet skal 26 omgivelsesfaktorer graderes på en skala fra hemmende til fremmende, med kommentarfelt der eksempler og vurderinger som underbygger graderingen kan beskrives. Omgivelsesfaktorene er organisert i tre kategorier:

- 1 rom/spatiale forhold (bygde/naturlige forhold der aktiviteten utføres)
- 2 objekter, utstyr, materiale og møbler (gjenstander som anvendes i aktiviteten)
- 3 sensoriske, klimatiske og andre forhold

En oversikt over omgivelsesfaktorer og kategorier er presentert i Tabell 2.

Kjennetegn ved deltakerne (N = 23)	
Erfaring som ergoterapeut (år), M (SD)	14 (10)
Arbeidssted	
Spesialisthelsetjeneste, n (%)	14 (61)
Kommunehelsetjeneste, n (%)	8 (35)
Fagfelt¹	
Somatisk helse, n (%)	18 (78)
Psykisk helse, n (%)	4 (17)
Barns helse, n (%)	3 (13)
Annet, n (%)	2 (9)
Analysekompetanse²	
Ingen, n (%)	0 (0)
Begrenset, n (%)	0 (0)
Moderat, n (%)	11 (48)
God, n (%)	11 (48)
Svært god, n (%)	1 (4)
Grad av forberedelser³	
Ikke lest, n (%)	1 (4)
Bladd igjennom, n (%)	2 (9)
Lest noe av materialet, n (%)	3 (13)
Lest grundig, n (%)	14 (61)
Lest svært grundig, n (%)	3 (13)

Tabell 1. Kjennetegn ved deltakerne.

M = gjennomsnitt; SD = standardavvik; ¹Samlet % overstiger 100 som følge av at deltakerne kunne krysses av for flere alternativer; ²Selvurdert kompetanse om analyse med bruk av EVA-systemet; ³Hvor grundig deltakerne hadde forberedt seg i forkant av studien (i hvilken grad de hadde lest brukerveiledning og øvrig tilsendt informasjon om EVA fysiske omgivelser)

Hver omgivelsesfaktor graderes på en visuell linje fra -1 til +1, delt med en midtstreke. Desto lenger ut fra midtstreken faktoren graderes (i begge retninger), desto sterkere påvirkning vurderes faktoren å ha. Eksempelvis vil en gradering mot ytterpunkt -1 indikere at faktoren i betydelig grad hemmer aktivitetsutførelsen, mens en gradering rett til venstre for midtstreken indikerer at faktoren hemmer aktivitetsutførelsen i noen, men liten grad.

I forbindelse med denne studien ble det benyttet en digital versjon av analyseskjemaet der graderingene ble gjort på en skala fra -3 til +3, med et nøytralt nullpunkt. Dette muliggjorde kvantifisering av graderinger foretatt innenfor henholdsvis det hemmende (-3 til -1) og det fremmende (+3 til +1) domenet. Figur 1 viser graderingssystemet, både slik

Oversikt over kategorier og faktorer i EVA fysiske omgivelser	
Kategorier	Faktorer
1. Rom/spatiale forhold (bygde/naturlige) og arkitektoniske elementer	1.1 Adkomst 1.2 Dimensjoner – spatiale 1.3 Underlag 1.4 Overflater 1.5 Utforming/planløsning 1.6 Faste bevegelige elementer 1.7 Faste flyttbare elementer 1.8 Tekniske innredninger 1.9 Annet*
2. Objekter, utstyr, materiale og møbler	2.1 Tilgjengelighet 2.2 Dimensjoner – objekter 2.3 Tyngde 2.4 Overflater og teksturer 2.5 Form, fasong og design 2.6 Plassering 2.7 Kompleksitet 2.8 Automatisering 2.9 Annet*
3. Sensoriske, klimatiske og andre forhold	3.1 Lys 3.2 Farger/kontraster 3.3 Lyd 3.4 Luft 3.5 Temperatur 3.6 Lukt 3.7 Temporale faktorer 3.8 Annet*

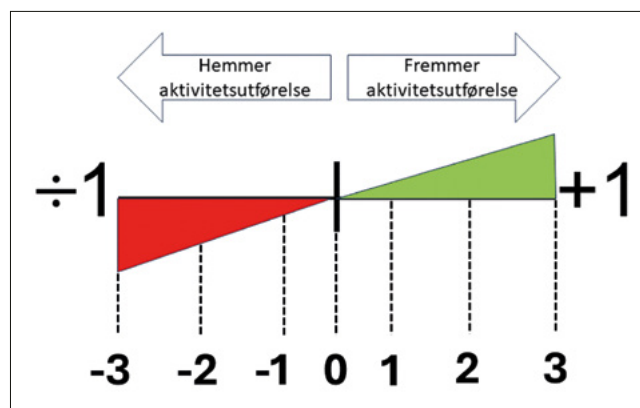
Tabell 2. Oversikt over kategorier og faktorer i EVA fysiske omgivelser.

*Analyseskjemaet åpner for at observatøren kan vurdere relevante forhold utover de faktorene som er angitt. Slike forhold er ikke inkludert i denne studien ettersom graderinger av «andre forhold» ikke vil være sammenliknbare på tvers av observatørene; EVA = Ergoterapianalyse og Vurdering av Aktivitet.

det fremstår i papirskjemaet, og slik det ble benyttet i den digitale versjonen.

DATAINNSAMLING OG PROSEDYRE

Datainnsamlingen ble gjennomført i form av et webinar med varighet på tre timer. Webinaret ble gjennomført på Microsoft Teams i juni 2024, og i forkant fikk deltakerne tilsendt en brukerveiledning for EVA fysiske omgivelser, et ikke-utfyllt analyseskjema og et eksempel på et utfyllt analyseskjema. Hver deltaker deltok individuelt i webinaret, slik at det ikke var mulighet for å samarbeide. I webinaret ble det vist to forhåndsinnspilte videoer. Videoene viste to ulike aktiviteter utført av forskjellige aktører i ulike omgivelser. Hver video hadde en varighet på



Figur 1. Graderingsskala for omgivelsesfaktorer i analyseskjemaet EVA fysiske omgivelser. Øverst i figuren illustreres den visuelle linjen slik den forekommer i papirversjonen. Nederste talllinje viser graderingsskalaen benyttet i denne studien (den digitale versjonen).

cirka ti minutter. Aktivitet 1 viste en kvinne som lager et glass sjokolademelk på kjøkkenet sitt. Aktivitet 2 viste en mann som tar på sengetøy og rer opp en seng i huset sitt.

Hver video ble vist to ganger. Ved første visning ble deltakerne bedt om å observere aktørens aktivitetsutførelse uten å fylle ut analyseskjemaet EVA fysiske omgivelser. Ved andre visning ble deltakerne bedt om å fylle ut analyseskjemaet på papir. I etterkant av andre visning ble deltakerne gitt tid til å fullføre analyseskjemaet. Det ble fylt ut ett analyseskjema for hver aktivitetsutførelse. Etter å ha fylt ut papirskjemaene for begge observasjonene ble deltakerne bedt om å overføre sine graderinger til et digitalt spørreskjema som var konstruert på samme måte som papirskjemaet. Graderingene i det digitale spørreskjemaet utgjorde studiens datamateriale.

Det digitale spørreskjemaet inneholdt også noen bakgrunnsspørsmål (se tabell 1) om erfaring som ergoterapeut (antall år), arbeidssted (spesialisthelsetjeneste; kommunehelsetjeneste; annet) og fagfelt (somatisk helse; psykisk helse; barns helse; annet). Deltakerne ble i spørreskjemaet også bedt om å indikere sin kompetanse om analyse av aktivitetsutførelse med bruk av EVA-systemet (ingen kompetanse; begrenset kompetanse; moderat kompetanse; god kompetanse; svært god kompetanse), samt hvor grundig de hadde forberedt seg ved å lese brukerveiledningen for EVA fysiske omgivelser og analyseskjema i forkant av studien (ikke lest; bladd igjennom; lest noe av materialet; lest grundig; lest svært grundig).

ANALYSE

Inter-rater reliabilitet ble undersøkt ved hjelp av

Inter-rater reliabilitet på verktøynivå				
		95 % KI for ICC		
	ICC ¹	Nedre	Øvre	p
Hovedanalyser:				
Aktivitet 1 og 2 samlet (n = 46/23)	0,90	0,85	0,94	<0,001
Aktivitet 1 (n = 23/23)	0,88	0,79	0,94	<0,001
Aktivitet 2 (n = 23/23)	0,92	0,86	0,96	<0,001
Subgruppeanalyser²:				
Spesialisthelsetjeneste (n = 46/14)	0,83	0,75	0,89	<0,001
Kommunehelsetjeneste (n = 46/8)	0,75	0,63	0,85	<0,001
Somatisk helse (n = 46/18)	0,87	0,81	0,92	<0,001
Psykisk helse (n = 46/4)	0,63	0,42	0,77	<0,001
Barns helse (n = 46/3)	0,47	0,14	0,69	0,005
Moderat analysekompetanse (n = 46/11)	0,74	0,62	0,84	<0,001
God/svært god analysekompetanse (n = 46/12)	0,87	0,80	0,92	<0,001
Lite forberedt (n = 46/6)	0,69	0,53	0,81	<0,001
Godt forberedt (n = 46/17)	0,88	0,82	0,92	<0,001

Tabell 3. Inter-rater reliabilitet på verktøynivå.

¹Intraklassekorrelasjonskoeffisient (ICC), estimert gjennom two-way mixed effects consistency modell; ²Subgruppeanalyser utført for aktivitet 1 og 2 samlet; KI = konfidensintervall; n = antall observasjoner (faktorgraderinger)/antall observatører (ratere)

intraklassekorrelasjonsanalyser med estimering av intraklassekorrelasjonskoeffisienter (ICC). ICC er et standardisert mål på enighet mellom to eller flere observatører og varierer på en skala fra 0 til 1, der høyere verdi indikerer høyere grad av enighet (sterkere inter-rater reliabilitet). Følgende terskelverdier ble lagt til grunn i denne studien: <0,40 = svak (poor); 0,40 – 0,59 = akseptabel (fair); 0,60 – 0,74 = god (good); og 0,70 – 1,00 = utmerket (excellent) (Cicchetti, 1994). Det finnes flere varianter av ICC. I denne studien benyttet vi en two-way mixed effects consistency modell for flere observatører/målinger (multiple raters/measurements) (McGraw & Wong, 1996), også kjent som ICC (3,k) (Shrout & Fleiss, 1979). ICC estimeres da ved hjelp av formelen $MS_R - MS_E$: MS_R (MSR = mean square for rows, MSE = mean square for error). I en slik modell behandles observasjonene (faktorgraderingene) som konstante (fixed factors) mens observatørene behandles som tilfeldige (random factors). Videre fokuserte vi på grad av konsistens mellom observatørenes graderinger (consistency model) snarere enn absolutt graderingssamsvar (absolute agreement). Det ble beregnet 95 prosent konfidensintervaller for ICC, og signifikansnivå ble a priori fastsatt til $p < .05$. Alle analyser ble gjort i IBM SPSS versjon 28.

Verktøynivå

Hovedanalysene ble gjort på verktøynivå, hvilket innebar at det ble estimert ICC for hele verktøyet, først for begge aktiviteter samlet og deretter separat for hver aktivitet. Videre ble det gjort subgruppeanalyser (for begge aktiviteter samlet) basert på kjennetegn ved utvalget (observatørene). Slike separate analyser ble gjort i henhold til arbeidssted (spesialisthelsetjeneste; kommunehelsetjeneste), fagfelt (somatisk helse; psykisk helse; barns helse), analysekompetanse (moderat; god/svært god), og grad av forberedelser (ikke lest/bladd igjennom/lest noe; lest grundig/svært grundig). Disse analysene ble gjort for å undersøke om det eksempelvis var ulik inter-rater reliabilitet blant ergoterapeuter som jobbet i spesialisthelsetjeneste sammenliknet med ergoterapeuter som jobbet i kommunehelsetjeneste.

Kategorinivå

Omgivelsesfaktorene i EVA fysiske omgivelser er inndelt i tre kategorier (rom/spatiale forhold; objekter, utstyr, materiale og møbler; sensoriske, klimatiske og andre forhold). Det ble utført separate analyser for hver av kategoriene, basert på graderinger for begge aktiviteter samlet. Disse analysene ble

Inter-rater reliabilitet på kategorinivå				
		95 % KI for ICC		
Kategori ¹	ICC ²	Nedre	Øvre	p
1. Rom/spatiale forhold (n = 23)	0,92	0,84	0,97	<0,001
2. Objekter, utstyr, materiale og møbler (n = 23)	0,84	0,69	0,93	<0,001
3. Sensoriske, klimatiske og andre forhold (n = 23)	0,95	0,90	0,98	<0,001

Tabell 4. Inter-rater reliabilitet på kategorinivå.

¹Analysert utført for aktivitet 1 og 2 samlet;

²Intraklassekorrelasjonskoeffisient (ICC), estimert gjennom two-way mixed effects consistency modell; KI = konfidensintervall; n = antall observatører (ratere)

gjort for å undersøke om inter-rater reliabilitet var forskjellig for de ulike kategoriene i verktøyet.

Faktornivå

Vi beregnet standardavvik (SD) for hver omgivelsesfaktor basert på graderinger for begge aktiviteter samlet. SD er et spredningsmål og størrelsen på SD indikerer hvor stor spredning det var i observatørenes graderinger av den aktuelle faktoren (høyere SD = større spredning/større uenighet). Disse analysene ble gjort for å identifisere hvilke fysiske omgivelsesfaktorer det var minst/mest enighet om.

FORSKNINGSETIKK

Alle deltakerne ga skriftlig informert samtykke til deltakelse. Studien ble gjennomført anonymt. Det ble altså ikke samlet inn opplysninger som gjør det mulig å identifisere enkeltpersoner. Deltakerne fikk tilgang til det digitale spørreskjemaet gjennom en lenke distribuert av forskerne. Spørreskjemaet var utformet som en anonym løsning i Nettskjema, hvilket innebar at det ikke ble registrert informasjon om e-post eller IP-adresse. Studien var derfor ikke meldepliktig til Sikt. Informert samtykke ble innhentet ved at deltakerne måtte svare bekreftende på følgende innledende spørsmål i det digitale spørreskjemaet for å kunne komme videre: «Jeg bekrefter at jeg har fått informasjon om studien og at jeg samtykker til å delta». Samtykke ble således innhentet uten registrering av navn eller andre personidentifiserende opplysninger.

Resultater

INTER-RATER RELIABILITET PÅ VERKTØYNIVÅ

Tabell 3 viser inter-rater reliabilitet for EVA – Analyse av fysiske omgivelers påvirkning på aktivitetsutførelse (EVA fysiske omgivelser) på verktøynivå.

Hovedanalysene viste at EVA fysiske omgivelser

hadde utmerket inter-rater reliabilitet på verktøynivå, både når begge aktivitetene ble analysert samlet (ICC = 0,90, $p < 0,001$), og når aktivitetene ble vurdert hver for seg (ICC_{Aktivitet1} = 0,88, $p < 0,001$; ICC_{Aktivitet2} = 0,92, $p < 0,001$). Subgruppeanalysene viste utmerket reliabilitet både blant ergoterapeuter som jobbet i spesialisthelsetjeneste (ICC = 0,83, $p < 0,001$) og blant de som jobbet i kommunehelsetjeneste (ICC = 0,75, $p < 0,001$). Når det gjaldt arbeidssted viste verktøyet utmerket reliabilitet blant de som jobbet innen somatisk helse (ICC = 0,87, $p < 0,001$), god reliabilitet blant de som jobbet innen psykisk helse (ICC = 0,63, $p < 0,001$) og akseptabel reliabilitet blant de som jobbet innen barns helse (ICC = 0,47, $p = 0,005$). Reliabiliteten var noe sterkere blant de som oppga å ha god eller svært god analysekompetanse (ICC = 0,87, $p < 0,001$), sammenliknet med de som mente å ha moderat analysekompetanse (ICC = 0,74, $p < 0,001$). Reliabiliteten var også sterkere i gruppen som hadde forberedt seg grundig eller svært grundig i forkant (ICC = 0,88, $p < 0,001$), sammenliknet med de som var mindre forberedt (ICC = 0,69, $p < 0,001$).

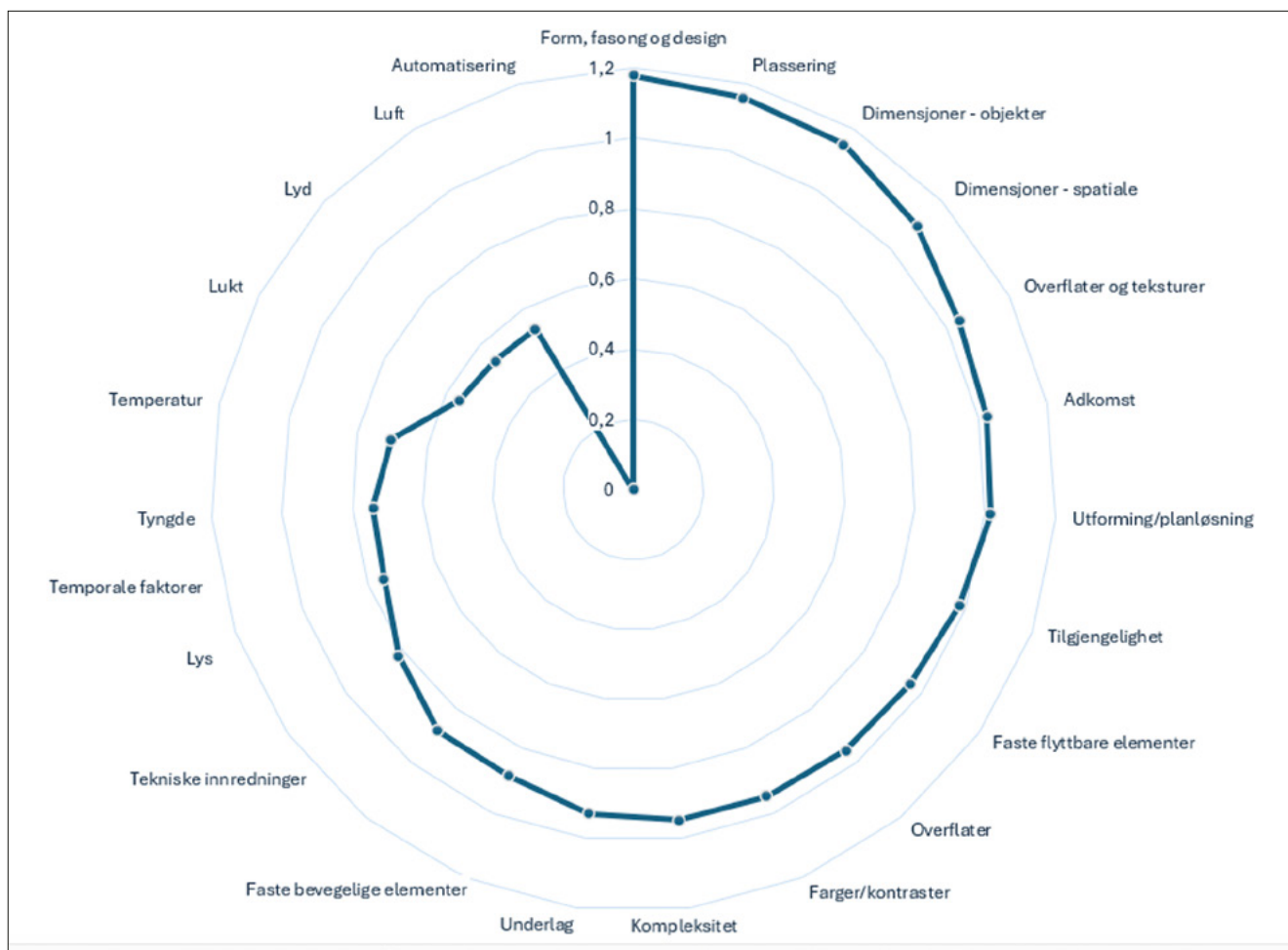
INTER-RATER RELIABILITET PÅ KATEGORINIVÅ

Tabell 4 viser inter-rater reliabilitet for EVA fysiske omgivelser på kategorinivå.

EVA fysiske omgivelser demonstrerte utmerket inter-rater reliabilitet innenfor alle de tre omgivelsekategoriene. Reliabiliteten var sterkest for kategorien sensoriske, klimatiske og andre forhold (ICC = 0,95, $p < 0,001$), etterfulgt av rom/spatiale forhold (ICC = 0,92, $p < 0,001$) og objekter, utstyr, materiale og møbler (ICC = 0,84, $p < 0,001$).

INTER-RATER RELIABILITET PÅ FAKTORNIVÅ

Analysen på faktornivå viste at det gjennomgående var betydelig enighet mellom observatørene om i hvilken grad omgivelsesfaktorene påvirket



Figur 2. Spredning i vurdering av omgivelsesfaktorer for aktivitet 1 og 2 samlet. Verdiene representerer standardavvik (SD) på en syvpunktsskala (-3 til +3, inkludert 0).

aktivitetsutførelse på tvers av de to aktivitetene. Hver faktor ble skåret på en syvpunktsskala (-3 til +3, inkludert 0), og 16 av de 23 faktorene hadde et SD på mindre enn 1,00. Det var størst enighet om faktoren automatisering (SD = 0,00), etterfulgt av sensoriske/klimatiske forhold som luft (SD = 0,54), lyd (SD = 0,54), lukt (SD = 0,56) og temperatur (SD = 0,70). Det var minst enighet om faktorene form, fasong og design (SD = 1,18), plassering (SD = 1,16) og dimensjoner - objekter (SD = 1,15), som alle sorterer i kategorien objekter, utstyr, materiale og møbler. Spredningen på samtlige omgivelsesfaktorer er illustrert i Figur 2.

Diskusjon

Formålet med denne studien var å undersøke inter-rater reliabilitet ved verktøyet EVA – Analyse av fysiske omgivelers påvirkning på aktivitetsutførelse (EVA fysiske omgivelser), et nylig utviklet generisk verktøy for analyse av de fysiske omgivel-

senes påvirkning på aktivitetsutførelse. En studie gjennomført i 2024 (Jespersen et al., 2025) antyder at ergoterapeuter opplever EVA fysiske omgivelser som nyttig og brukervennlig, men det har ikke tidligere blitt undersøkt hvorvidt verktøyet er i stand til å generere reliabel vurdering og dokumentasjon av fysiske aktivitetsomgivelser på tvers av ulike ergoterapeuter. I foreliggende studie ble det undersøkt om EVA fysiske omgivelser viste tilfredsstillende inter-rater reliabilitet når verktøyet ble benyttet av 23 ergoterapeuter til å vurdere to ulike aktivitetsutførelser. Det ble også undersøkt om ulike deler av verktøyet fungerte forskjellig, og derfor ble det utført separate analyser for hver av de tre delene (omgivelseskategoriene) i EVA fysiske omgivelser, så vel som for hver av omgivelsesfaktorene knyttet til de tre kategoriene. Det var også et poeng å undersøke om verktøyet fungerte ulikt avhengig av kjennetegn ved ergoterapeutene som gjennomførte kartleggingen. Derfor ble det utført subgrup-

peanalyser basert på ergoterapeutenes arbeidssted, fagfelt, analysekompetanse og grad av forberedelser i forkant av studien.

Resultatene viste at EVA fysiske omgivelser gjennomgående hadde god til utmerket inter-rater reliabilitet. Dette innebar at ulike observatører, uavhengig av hverandre, vurderte omgivelsene likt da de observerte de samme aktivitetene. Sagt på en annen måte: Verktøyet syntes å være godt i stand til å kartlegge de fysiske omgivelsenes påvirkning på aktivitetsutførelse. Det sterke samsvaret mellom graderinger foretatt av ulike observatører tyder på at kartleggingsresultatene reflekterer de faktiske faktorene i omgivelsene snarere enn individuelle variasjoner i hvordan ergoterapeuter vurderer omgivelser som påvirker aktivitetsutførelse. Tidligere forskning som har undersøkt inter-rater reliabilitet ved andre liknende kartleggingsverktøy har funnet moderat til høy reliabilitet, eksempelvis knyttet til verktøyene Housing Enabler (Helle et al., 2014; Iwarsson et al., 2005), REIS (Harrison et al., 2023) og WEIS (Kielhofner et al., 1999). Resultatene for EVA fysiske omgivelser er sammenliknbare og antyder gjennomgående noe sterkere inter-rater reliabilitet for dette verktøyet.

I motsetning til andre instrumenter benyttet for å kartlegge omgivelser er EVA fysiske omgivelser ment å være et generisk verktøy (Jespersen et al., 2025). Verktøy som Housing Enabler, REIS og WEIS er knyttet til spesifikke aktivitetskontekster (f.eks. hjem, skole eller arbeid) (Corner et al., 1997; Ellingham & Jakobsen, 2003; Fisher et al., 2014; Iwarsson & Slaug, 2001, 2010), mens EVA fysiske omgivelser tar sikte på å gjøre det mulig å analysere og vurdere omgivelsenes påvirkning på tvers av aktivitetskontekster. Verktøyet må derfor være reliabelt på tvers av aktivitetstyper, -kontekster, bruker-/pasientgrupper og ergoterapeutgrupper. Studiens resultater antyder at EVA fysiske omgivelser kan fungere tilfredsstillende som et generisk kartleggingsverktøy. Inter-rater reliabiliteten var utmerket for de to ulike aktivitetsutførelsene og de tre delene (kategoriene) i verktøyet fungerte tilnærmet like godt. Reliabiliteten var utmerket både blant ergoterapeuter som jobbet i spesialisthelsetjeneste og blant de som jobbet i kommunehelsetjeneste. EVA fysiske omgivelser fungerte også tilfredsstillende på tvers av ergoterapeutenes fagfelt, analysekompetanse og grad av forberedelse i forkant av studien.

EVA fysiske omgivelser var mest reliabelt blant ergoterapeuter som i utgangspunktet var godt

forberedt, som hadde god analysekompetanse og som jobbet i spesialisthelsetjenesten innenfor somatisk helse. Verktøyet viste likevel god reliabilitet blant ergoterapeuter i kommunehelsetjeneste og innen psykisk helse, og også blant terapeuter som hadde moderat analysekompetanse og som var lite forberedt i forkant av studien. Resultatene viste noe svakere reliabilitet blant ergoterapeuter som jobbet med barns helse. Dette kan skyldes at aktivitetsutførelsene som ble vist i liten grad samsvarer med aktiviteter og -kontekster som terapeuter i dette fagfeltet jobber med til daglig. Alternativt kan det skyldes metodiske utfordringer knyttet til at det i denne studien kun var tre ergoterapeuter i utvalget som oppga å jobbe med barns helse. Når inter-rater reliabilitet estimeres på grunnlag av få observasjoner er det risiko for at resultatet blir mindre nøyaktig og pålitelig som følge av større sannsynlighet for at tilfeldige variasjoner påvirker resultatene, hvilket kan føre til lavere reliabilitetsestimater (Mehta et al., 2018).

I separate analyser av de enkelte fysiske omgivelsesfaktorene var det gjennomgående stor enighet mellom ergoterapeutene med hensyn til hvordan faktorene skulle graderes. Det var imidlertid en svak tendens til at det var mest enighet om gradering av sensoriske og klimatiske faktorer, og noe mindre enighet om faktorer relatert til objekter og gjenstander som ble benyttet i aktivitetene. En mulig forklaring på denne tendensen kan være at forhold som objektform/-fasong, plassering og objektdimensjonering er mer komplekse enn faktorer som luft, lyd og lukt, hvilket kan ha ført til noe mer variasjon i vurdering av objektrelaterte faktorer. Tendensen kan også skyldes at den faktiske kompleksiteten ved sensoriske/klimatiske faktorer er vanskelig å få frem i videoformat, hvilket kan ha resultert i at disse faktorene ble vurdert som noe mer konstante enn dersom aktivitetsutførelsene hadde blitt observert fysisk.

METODISKE BETRAKTNINGER

Dette er den første studien som undersøker måleegenskaper ved EVA fysiske omgivelser. Det er en styrke ved denne studien at den bygger på data fra mange observatører (N = 23) som har vurdert omgivelsenes påvirkning i to ulike aktivitetsutførelser i form av et betydelig antall uavhengige faktorgraderinger (N = 23-46). Den statistiske styrken i en studie som undersøker inter-rater reliabilitet bestemmes av forholdet mellom antall observatører og antall

observasjoner. Studien hadde tilfredsstillende styrke, gitt en styrkeberegning basert på $\alpha = 0,05$ og $\beta = 0,20$ (Walter et al., 1998). Det relativt store utvalget gjorde det mulig å gjennomføre subgruppeanalyser i tillegg til hovedanalysene. Subgruppeanalyser var helt nødvendig i og med at EVA fysiske omgivelser er utviklet som et generisk verktøy som skal kunne brukes på tvers av aktiviteter og kontekster av ergoterapeuter innen ulike fagfelt og med ulik erfaring og kompetanse.

Studien har imidlertid noen viktige begrensninger. Ergoterapeutene som deltok i studien var ikke tilfeldig utvalgt. Deltakerne ble rekruttert blant ergoterapeuter som hadde gjennomført EVA-kurs, blant deltakere på en EVA-workshop samt via åpen påmelding distribuert gjennom sosiale medier og EVA-bloggen (Ellingham et al., u.å.). Denne rekrutteringsstrategien kan ha medført at deltakerne var noe mer interesserte og kompetente i aktivitetsanalyse og EVA-systemet enn ergoterapeuter flest, hvilket kan ha gitt en seleksjonsskjevhet. Inter-rater reliabiliteten identifisert i denne studien kan dermed fremstå som noe sterkere enn dersom deltakerne hadde blitt tilfeldig rekruttert fra ergoterapeutpopulasjonen. På den annen side viste resultatene at EVA fysiske omgivelser var reliabel også blant ergoterapeuter som kun hadde moderat analysekompetanse og som var lite forberedt før studien.

Deltakerne observerte aktivitetsutførelsene i form av video og var ikke til stede i de faktiske fysiske omgivelsene. Deltakerne ble vist to ulike aktivitetsutførelser, men disse er ikke nødvendigvis dekkende for alle aktiviteter ergoterapeuter møter i sin kliniske praksis. Videre forskning kan med fordel fokusere på ekstern validitet ved å teste verktøyets reliabilitet i reelle omgivelser som er aktuelle i klinisk praksis.

Ved beregning av inter-rater reliabilitet benyttet vi en estimeringsmodell som fokuserte på konsistens mellom observatører snarere enn absolutt samsvar mellom observatører. Erfaringsmessig kan konsistensmodeller gi noe høyere reliabilitetsestimater enn modeller basert på absolutt samsvar (Koo & Li, 2016). Det ble vurdert som viktigere å få et bilde av konsistensen (graderingstendensen) mellom observatørene heller enn i hvilken utstrekning de landet på nøyaktig samme gradering. Dette dels fordi graderingene ble gjort på en syvpunktsskala, og dels fordi det i komplekse vurderinger av omgivelsenes påvirkning på aktivitetsutførelse nødvendigvis vil være noe innslag av subjektivitet. Å kreve at observatørene skulle lande på identiske graderinger ville vært urimelig strengt.

Det er viktig å merke seg at inter-rater reliabilitet er én av flere viktige måleegenskaper ved et kartleggingsverktøy (Gagnier et al., 2021). Videre forskning kan med fordel utforske flere egenskaper ved EVA fysiske omgivelser, eksempelvis intra-rater reliabilitet, altså hvor like graderingene blir når samme observatør vurderer samme aktør og aktivitetsutførelse flere ganger med relativt korte mellomrom. Det kan også være interessant å undersøke verktøyets konvergente validitet, altså i hvilken grad vurdering av aktivitetsutførelse gjort ved hjelp av EVA fysiske omgivelser samsvarer med vurdering av samme aktivitetsutførelse med andre relevante kartleggingsverktøy. Inter-rater reliabilitet er imidlertid en svært viktig måleegenskap for kartleggingsverktøy som benyttes ved observasjon i naturlige kontekster. Denne studien utgjør derfor et viktig bidrag.

EVA fysiske omgivelser er begrenset til de fysiske omgivelsene. Videre fagutviklingsarbeid kan med fordel fokusere på utvikling av et generisk verktøy for analyse av de psykososiale omgivelsenes påvirkning på aktivitetsutførelse.

KONKLUSJON

Formålet med denne studien var å undersøke inter-rater reliabilitet ved verktøyet EVA – Analyse av fysiske omgivelers påvirkning på aktivitetsutførelse (EVA fysiske omgivelser). Resultatene viste gjennomgående at verktøyet hadde god til utmerket reliabilitet samt at verktøyet fungerte godt med hensyn til ulike aktiviteter og kontekster, både blant ergoterapeuter på ulike arbeidsteder, innen ulike fagfelt og med ulik analysekompetanse. Resultatene indikerer dermed at EVA fysiske omgivelser er et reliabelt generisk verktøy for analyse av fysiske omgivelers påvirkning på aktivitetsutførelse. Videre forskning bør teste verktøyets reliabilitet i flere situasjoner innen ulike praksisområder og for flere aktiviteter.

Referanser

- Asunta, P., Viholainen, H., Ahonen, T., & Rintala, P. (2019). Psychometric properties of observational tools for identifying motor difficulties - a systematic review. *BMC Pediatrics*, 19(1). <https://doi.org/10.1186/s12887-019-1657-6>
- Bonsaksen, T., & Ellingham, B. (2017). *Klinisk resonering i ergoterapi*. Mankoni media.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. <https://doi.org/10.1037/1040-3590.6.4.284>

- Clark, E. G., Jespersen, L. F., Ellingham, B., & Brovold, T. (2014). Fra skolepulten til kjøkkenbenken. *Ergoterapeuten*, 57(1), 26-34.
- Corner, R. A., Kielhofner, G., & Lin, F.-L. (1997). Construct validity of a work environment impact scale. *Work*, 9(1), 21-34. <https://doi.org/10.3233/WOR-1997-9104>
- Ellingham, B., Clark, E. G., & Jespersen, L. F. (u.å.). *EVA-blogg. Ergoterapi analyse og vurdering av aktivitet (EVA)*. <https://eva933.wordpress.com/>
- Ellingham, B., & Jakobsen, K. (2003). *Work Environment Impact Scale. Brukermanual, norsk versjon* (HiO-rapport 2003 nr. 38). Høgskolen i Oslo.
- Ellingham, B., Jespersen, L. F., & Clark, E. G. (2023). *EVA 2023. Ergoterapi analyse og vurdering av aktivitet. Manual for analyse, vurdering og dokumentasjon av aktivitetsutførelse*. <https://eva933.wordpress.com/nedlasting-av-eva-dokumenter-og-skjemaer/>
- Feldborg, M. B., & Røhl, M. L. R. (2019). Aktivitetsanalyse i ergoterapi. I Å. Brandt, H. Peoples, & U. Pedersen (red.), *Basisbog i ergoterapi - aktivitet og deltakelse i hverdagslivet* (4. utg., s. 181-196). Munksgaard.
- Fisher, A. G., & Marterella, A. (2019). *Powerful practice. A model for authentic occupational therapy*. Center for innovative OT solutions.
- Fisher, G., Forsyth, K., Harrison, M. B., Angarola, R., Kayhan, E., Noga, P. L., Johnson, C., & Irvine, L. (2014). *Residential Environment Impact Scale (REIS) version 4.0*. Model of Human Occupation Clearinghouse: University of Illinois at Chicago.
- Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guidelines for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research*, 30, 2197-2218. <https://doi.org/10.1007/s11136-021-02822-4>
- Harrison, M., Forsyth, K., Murray, A. L., Angarola, R., HENDERSON, S., Irvine Fitzpatrick, L., & Fisher, G. (2023). Establishing the measurement properties of the Residential Environment Impact Scale (Version 4.0). *Scandinavian Journal of Occupational Therapy*, 30(6), 898-907. <https://doi.org/10.1080/11038128.2022.2143891>
- Helle, T., Nygren, C., Slaug, B., Brandt, A., Pikkarainen, A., Hansen, A.-G., Petursdottir, E., & Iwarsson, S. (2014). The Nordic Housing enabler: inter-rater reliability in cross-Nordic occupational therapy practice. *Scandinavian Journal of Occupational Therapy*, 21, 71-79. <https://doi.org/10.3109/11038128.2014.952907>
- Hersch, G. I., Lamport, N. K., & Coffey, M. S. (2005). *Activity analysis: application to occupation*. Slack.
- Iwarsson, S., Nygren, C., & Slaug, B. (2005). Cross-national and multi-professional inter-rater reliability of the Housing Enabler. *Scandinavian Journal of Occupational Therapy*, 12(1), 29-39. <https://doi.org/10.1080/11038120510027144>
- Iwarsson, S., & Slaug, B. (2001). *The housing enabler. An instrument for assessing and analysing accessibility problems in housing*. Vetén & Skapen HB & Slaug Data Management.
- Iwarsson, S., & Slaug, B. (2010). *Screeningverktøyet housing enabler - kortfattet manual*. Vetén & Skapen HB & Slaug Data Management.
- Jespersen, L. F., Clark, E. G., & Ellingham, B. (2014). Klinisk bruk av ferdighetsanalyse fra EVA-systemet i observasjon og dokumentasjon av aktivitetsutførelse. *Ergoterapeuten*, 57(4), 60-69.
- Jespersen, L. F., Clark, E. G., Thørrisen, M. M., & Ellingham, B. EVA - Analyse av fysiske omgivers påvirkning på aktivitetsutførelse. Utvikling og nytteverdi av et kartleggingsverktøy. (2025) *Ergoterapeuten*, 68(2), 20-30.
- Kielhofner, G. (2010). *MOHO – Modellen for menneskelig aktivitet: ergoterapi til utdanning og praksis* (2. utg.). Munksgaard.
- Kielhofner, G., Lai, J. S., Olson, L., Haglund, L., Ekbadh, E., & Hedlund, M. (1999). Psychometric properties of the work environment impact scale: a cross-cultural study. *Work*, 12(1), 71-77.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Law, M., Cooper, B., Strong, S., Stewart, D., Rigby, P., & Letts, L. (1996). The person-environment-occupational model: a transactive approach to occupational performance. *Canadian Journal of Occupational Therapy*, 63(1), 9-23. <https://doi.org/10.1177/000841749606300103>
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30. <https://doi.org/10.1037/1082-989X.1.1.30>
- Mehta, S., Bastero-Caballero, R. F., Sun, Y., Zhu, R., Murphy, D. K., Hardas, B., & Koch, G. (2018). Performance of intraclass correlation coefficient (ICC) as a reliability index under various distributions in scale reliability studies. *Statistics in Medicine*, 37(18), 2734-2752. <https://doi.org/10.1002/sim.7679>
- Mokkink, L. B., Eekhout, I., Boers, M., van der Vleuten, C. P., & de Vet, H. C. (2023). Studies on reliability and measurement error of measurements in medicine – from design to statistics explained for medical researchers. *Patient Related Outcome Measures*, 14, 193-212. <https://doi.org/10.2147/PROM.S398886>
- Polatajko, H. J., Townsend, E. A., & Craik, J. (2007). Canadian model of occupational performance and engagement (CMOP-E). I E. A. Townsend & H. J. Polatajko (red.), *Enabling occupation II: Advancing an occupational therapy vision of health, well-being, & justice through occupation* (s. 22-36). CAOT publications.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420. <https://doi.org/10.1037/0033-2909.86.2.420>
- Taylor, R. (2017). Introduction to the model of human occupation. I R. R. Taylor (red.) *Kielhofner's model of human occupation: Theory and application* (s. 57-73). Lippincott Williams and Wilkins.
- Thomas, H. (2023). *Occupational and activity analysis* (3. utg.). Slack.
- Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine*, 17(1), 101-110. [https://doi.org/10.1002/\(SICI\)1097-0258\(19980115\)17:1%3C101::AID-SIM727%3E3.0.CO;2-E](https://doi.org/10.1002/(SICI)1097-0258(19980115)17:1%3C101::AID-SIM727%3E3.0.CO;2-E)