



Mia C. Heller
Stipendiat ved Institutt for pedagogikk, Universitetet i Oslo. Heller disputerer den 02. oktober, og forsvarer da avhandlingen: *Second-language learning in the early elementary years: A researcher-practitioner approach*. Hun jobber nå i Statped sørøst.

Mia.cecilie.heller@statped.no



Arne Lervåg
er professor ved Institutt for pedagogikk, Universitetet i Oslo. Lervåg forsker på språk, læring og leseutvikling. Han har arbeidet spesielt mye med store longitudinelle studier som har fulgt barn over lang tid.
a.o.lervag@iped.uio.no

KARTLEGGING AV NORSKSPRÅKLIGE FERDIGHETER FOR ELEVER PÅ 1.-4.TRINN: EN VURDERING AV KARTLEGGINGSVERKTØYET NORSK SOM LÆRINGSSPRÅK (NSL)

Sammendrag

Kartlegging av norskspråklige ferdigheter kan bidra til tidlig identifisering av minoritetsspråklige barn som trenger språkstøtte i skolen, og dermed være veiledende for innholdet i undervisningen. Kvaliteten ved kartleggingsverktøy er viktig ved kartlegging. Denne studien gjennomførte analyser av reliabilitet, begrepsvaliditet og kriterievaliditet i kartleggingsverktøyet «Norsk Som Læringspråk» (NSL), som er en læreplanorientert kartlegging av elevers norskspråklige ferdigheter for 1.-4.trinn. NSL vurderer ferdigheter innenfor kategorisering, vokabular (bredde og dybde) og grammatikk. Konfirmerende faktoranalyser identifiserte NSL som en fire-faktormodell. Det ble også funnet korrelasjoner mellom NSL og andre standardiserte språkkartlegginger som underbygger validiteten i kartleggingsverktøyet. Implikasjoner og potensielle utfordringer ved NSL diskuteres.

Innledning

Gode språkferdigheter er et viktig grunnlag for læring og deltakelse i et sosialt fellesskap. I skolen er det en økende andel elever som har et annet morsmål enn norsk. Begrepet minoritetsspråklig defineres ofte i skolesammenheng som barn med et annet førstespråk enn norsk eller samisk (Opplæringslovens § 2-8). Tall for 2018 (Statistisk sentralbyrå, 2019) viser at andelen barn med vedtak om særskilt norskopplæring utgjør 6,7 % av elevtallet i grunnskolen i Norge. I Oslo utgjør denne gruppen elever 19,5 %. Minoritetsspråklige elever som gruppe har svakere ferdigheter på skolens språk enn sine enspråklige medelever (Melby-Lervåg & Lervåg, 2014). Forskjellene i språkferdigheter tydeliggjør utfordringene skolen står overfor når det gjelder tilrettelegging av undervisningen for elevene. Kunnskap om hva språk er og hvordan språkferdigheter utvikles og måles er avgjørende for å identifisere minoritetsspråklige elever som trenger språkstøtte. Betydningen av barns norskspråklige ferdigheter og hvordan barnehagen og skolen kan bidra til å støtte barnas språkutvikling løftes frem i stortingsmeldingen om tidlig innsats og inkluderende fellesskap (Kunnskaps-

departementet, 2019). Innenfor en norsk kontekst finnes det foreløpig få kartleggingsverktøy som skolen kan benytte for vurdering av språkferdigheter. Det er ingen føringer for hvilke kartleggingsverktøy som skal ligge til grunn utover de nasjonale og kommunale prøvene som er obligatoriske. Disse kartleggingene måler barnas leseferdigheter. Det kan derfor være en utfordring for skolen å sette inn relevante tiltak for å styrke språkutviklingen hos minoritetsspråklige barn som viser mangelfulle norskspråklige ferdigheter.

En av grunnene til at *Norsk som læringspråk* ble utviklet var mangelen på verktøy som måler elevenes norskspråklige ferdigheter (NSL; Frøyen, Ahmadinia, Heller & Skjåk, 2011; Frøyen, Ahmadinia, Heller, Skjåk & Namvar, 2015). Man så behov for et læreplanorientert instrument som kartla minoritetsspråklige barns norskspråklige ferdigheter uten å stille krav til lese- og skriveferdigheter. Innenfor en norsk kontekst finnes det få studier av kvaliteten til kartleggingsverktøy som benyttes for å vurdere barns skolefaglige ferdigheter. Arnesen, Braeken, Ogden og Melby-Lervåg (2019) gjennomfør-

te imidlertid nylig en undersøkelse som belyste kvaliteten på kartleggingsverktøy brukt til vurdering av sosial fungering og leseferdigheter i barneskolen, og hvorvidt resultatene ga føringer for innholdet i undervisningen. Resultatene viste at mange av kartleggingsverktøyene hadde svak validitet og reliabilitet eller at det ikke var foretatt noen kvalitetsvurdering av dem i det hele tatt. Dette er spesielt bekymringsfullt sett i lys av at flere av disse kartleggingsverktøyene er hyppig brukt i skolen. Når det gjaldt nytteverdien indikerte funn at føringer for praksis i stor grad var basert på lærernes uformelle vurderinger fremfor informasjonen som fremkom gjennom resultatene fra kartleggingene. Denne artikkelen vurderer komponentene NSL består av gjennom analyser av reliabilitet, begrepsvaliditet og kriterievaliditet, med påfølgende implikasjoner for praksis.

Språkets betydning for skolefaglig utvikling

Språk er et komplekst fenomen. Det deles ofte i muntlige ferdigheter (ekspressive) og lytteferdigheter (impressive). Språkets struktur beskrives ut ifra komponentene semantikk (kunnskap om ord og setningers meningsinnhold og betydning), grammatikk (syntaks og morfologi), fonologi (språkllyder) og pragmatikk (bruk av språk) (Berko Gleason, 2009). Disse komponentene påvirker hverandre gjensidig og er samtidig utsatt for innflytelse fra omgivelsene. Barn utvikler språket gjennom samspill, både som aktive deltakere og lyttere (Hoff, 2006). Vokabular danner grunnlaget for språkutviklingen gjennom muligheter for å kommunisere tanker, meninger, følelser og behov, men også ved evnen til å systematisere og kategorisere ord i systemer. Det skilles ofte mellom vokabularbredde og vokabulardybde (Schmitt, 2014). Vokabularbredde handler om antall ord barn har kjennskap til, mens vokabulardybden viser til hvor dyp kunnskap barnet har om ord. Ordinnlæringen henger nært sammen med barns grammatiske utvikling ved at evnen til å sette sammen ord til enkle ytringer forutsetter at en viss mengde ord er lært. Antall ord et barn forstår og kan bruke er til hjelp for å hente ut mening av en ytring. Samtidig kan en ytring bidra med kontekstuell støtte for å forstå nye ord eller bidra til dypere kunnskap om et ord. Denne parallelle utviklingen og relasjonen mellom vokabular og grammatikk bekreftes gjennom signifikant høye korrelasjoner i studier av språkutvikling (Klem et al., 2015; Lervåg, Hulme & Melby-Lervåg, 2018).

Tvillingstudier viser at gener påvirker språktilegnelsen (Hayiou-Thomas, Dale & Plomin, 2012), samtidig som

miljøet barnet er en del av spiller en sentral rolle (Hart & Risley, 1995). Mengden av og variasjonen i språkstimuleringen barn blir eksponert for i tidlig alder har sammenheng med senere språkferdigheter (Hoff, 2006; Marulis & Neuman, 2010; Rowe, 2012). Samtidig har studier vist at det eksisterer en sammenheng mellom graden av språkstimulering og de sosioøkonomiske forholdene barn vokser opp under, noe som blant annet medfører at det tidlig etableres forskjeller i størrelsen på barns ordforråd (Hart & Risley, 1995). Disse forskjellene har vist seg å vedvare over tid (Melby-Lervåg et al., 2012; Storch & Whitehurst, 2002). Variasjon i størrelsen på ordforrådet er bekymringsfullt sett i lys av at skolerelaterte ferdigheter i stor grad predikeres av barnets ordforråd i førskolealder (Dickinson & Tabors, 2001; Dickinson & Porche, 2011; Hjetland, Brinchmann, Scherer & Melby-Lervåg, 2017).

Barn møter allerede fra skolestart et mer akademisk språk preget av abstraksjon, og som dessuten kjennetegnes av et avansert vokabular, kompleks syntaks og krav til varierte pragmatiske ferdigheter (Snow & Uccelli, 2009). Det krever ferdigheter som ikke er temaspesifikke, men som gir muligheter for å forstå faginnhold på tvers av fagområder. Ferdighetene er betydningsfulle for tilegnelse av kunnskap, spesielt med tanke på at det akademiske språket også er undervisningsspråket i skolen (Snow, 2010). Det stilles ofte uforholdsmessige store krav til hva barn som er på vei til å tilegne seg et andrespråk skal få med seg i løpet av skoledagen: elevene er både i en prosess hvor et nytt språk skal læres og mestres (tospråklig utvikling), samtidig som den skolefaglige kunnskapen blir formidlet og skal tilegnes gjennom et språk de foreløpig ikke behersker godt nok. Teoretisk sett skilles det mellom en simultan- eller en suksessiv tospråklig utvikling (Paradis, Genesee & Crago, 2011). Simultan tospråklig utvikling innebærer at barnet eksponeres for og utvikler første- og andrespråket parallelt (for eksempel at foresatte snakker hvert sitt språk med barnet). Ved en suksessiv tospråklig utvikling blir barnet først eksponert for morsmålet, for så å tilegne seg andrespråket på et senere tidspunkt (for eksempel ved oppstart i barnehagen eller skolestart). I Norge har mange minoritetsspråklige elever en suksessiv tospråklig utvikling (Bjerkan, Monsrud & Thurmann-Moe, 2013).

Hvordan kan språk måles?

Måling av språkferdigheter stiller krav til en operasjonalisering av begrepet språk. Operasjonaliseringen innebærer å gjøre begrepet målbart ved bruk av ulike kartleggingsverktøy som representerer de språklige komponentene.

Kartlegging av barn som trenger språkstøtte handler ofte om å identifisere hvilke ord (vokabular) og hvilken type setninger (grammatikk) et barn forstår og kan bruke, samt fortellerferdigheter (pragmatikk). Det er vanlig å bruke standardiserte kartleggingsverktøy for å sammenlikne barns språkferdigheter med hva som er forventet for alderen. Eksempler på slike kartleggingsverktøy er British Picture Vocabulary Scale (BPVS-II; Dunn, Dunn, Whetton & Burley, 1997; Lyster, Horn & Rygvold, 2010) som vurderer impressivt vokabular og Test for Reception of Grammar (TROG-2; Bishop, 2003; Lyster & Horn, 2009) som måler grammatisk forståelse. Selv om hensikten med kartleggingsverktøy er å gi presise målinger av begrepet språk, vil det forekomme målefeil. Målefeil kan være systematiske (skjev måling av begrepet på grunn av for snever eller for bred operasjonalisering) eller tilfeldige (tilfeldigheter man ikke kan kontrollere slik som elevens dagsform under kartleggingen).

Et springende punkt ved utvikling av kartleggingsverktøy handler om instrumentets kvalitet. Kvalitet vurderes ut ifra spesifikke kriterier for validitet og reliabilitet, som baseres på analyser av komponentene instrumentet består av. Reliabilitet gir informasjon om hvor stabilt og reproduserbart måleinstrumentet er (målesikkerhet) (Thorndike & Thorndike-Christ, 2014). Validitet uttrykker i hvilken grad teorien og empirien som ligger til grunn støtter tolkningen av testresultatene innenfor en spesifikk kontekst (Messick, 1995). For at et kartleggingsverktøy skal ha nytteverdi, både for forskning og praksisfeltet, må det være reliabelt og valid. Dersom en test viser seg å ha lav reliabilitet, vil den heller ikke egne seg til bruk da den ikke er valid (gyldig). Samtidig er det viktig å påpeke at god reliabilitet ikke er en garanti for validitet, da en test kan vise seg å være pålitelig uten at den faktisk måler det den er utviklet for å måle (Thorndike & Thorndike-Christ, 2014). I denne artikkelen fokuseres det på begrepsvaliditet og kriterievaliditet i evalueringen av validitet i kartleggingsverktøyet NSL.

Begrepsvaliditet dreier seg om i hvilken grad komponentene måleinstrumentet består av er representative for begrepet de er ment å måle. Ved utvikling av et kartleggingsverktøy vurderes begrepsvaliditet gjennom estimering av faktormodeller (Brown, 2015). For å vurdere strukturen i kartleggingsverktøyet brukes faktoranalyser som undersøker den felles variasjonen mellom variablene. Analysene gir informasjon om hvorvidt begrepet, i dette tilfellet språk, er best forklart som én felles dimensjon (språk) eller om det representerer separate dimensjoner (for eksempel gramma-

tikk og vokabular). Flere studier finner støtte for at variasjon i yngre barns språkferdigheter, da spesielt vokabular, grammatikk og språkbruk, kan forklares gjennom en endimensjonal språkfaktor (Klem et al., 2015; Tomblin & Zhang, 2006). Hos eldre barn er det funnet støtte for en flerdimensjonal språkfaktor (Tomblin & Zhang, 2006). Siden normer brukes som retningslinjer for tolkning av elevens prestasjoner, er det viktig å vite hvorvidt kartleggingsverktøyet er representativt for barngruppen det brukes på. Invarians handler om i hvilken grad det operasjonaliserte begrepet kan sies å måle det samme på tvers av grupper eller over tid (Kline, 2016). I forskningslitteraturen vises det til grad av invarians, og det måles basert på faktorstrukturen (nivå 1; konfigural), faktorladningene (nivå 2; metrisk) og intersept gjennomsnittet (nivå 3; skalar). Det laveste nivået, faktorstrukturen, må være oppfylt for å kunne vurdere invarians over faktorladningene som igjen må være oppfylt for kunne vurdere invarians over interseptene (Putnik & Bornstein, 2016). Skalar invarians er en forutsetning for å kunne sammenlikne gjennomsnitt mellom grupper (Kline, 2016). Det brukes ofte for å sammenlikne om det operasjonaliserte begrepet (språk) måles på samme måte i forhold til faktorer som alder, kjønn, sosioøkonomisk bakgrunn og språkstatus. Dersom det ikke er skalar invarians må det brukes forskjellige normer, da begrepet ikke måles likt mellom gruppene.

Kriterievaliditet vurderer hvorvidt kartleggingsverktøyet samsvarer med andre instrumenter som måler det samme begrepet, og om instrumentet tilfører ny informasjon som ikke fanges opp av andre verktøy som allerede benyttes for samme formål. I tillegg belyses det hvor godt måleinstrumentet klarer å identifisere elevgruppen det ble designet for (Thorndike & Thorndike-Christ, 2014). Kriterievaliditet kan enten være samtidig eller prediktiv, noe som avhenger av hvorvidt testene er foretatt på samme tidspunkt eller benyttes for å forutsi fremtidige prestasjoner.

Norsk som læringsspråk (NSL)

NSL er et læreplanorientert screeningsverktøy og ikke et diagnostisk kartleggingsverktøy (slik som for eksempel BPVS-II og Trog-2). Hensikten med NSL er å identifisere minoritetsspråklige elever som har behov for å styrke norskspråklige ferdigheter i barneskolen (1.-4.trinn), for deretter å tilrettelegge opplæringen basert på resultatene som foreligger. Komponentene i NSL baseres på teoretiske og praktiske tilnærminger til språk for tilegnelse av kunnskap. Den teoretiske tilnærmingen baseres på teorier og studier om språkutvikling, både generell og tospråklig

språkutvikling, og ligger til grunn for utvelgelse av komponentene i NSL (kategorisering, vokabular og grammatikk). Den praktiske tilnærmingen tok utgangspunkt i at kartleggingsverktøyet skulle være læreplanorientert. Prosjektmedarbeiderne foretok analyser av hvilke krav som stilles til elevenes språkferdigheter i Kunnskapsløftet (LK06; Utdanningsdirektoratet, 2006). Resultatene ble kombinert med analyser av ord fra lærebøker i fagene norsk, matematikk, samfunnsfag, naturfag og KRLE. Lærere fra 1.-4. trinn vurderte deretter ordlistene og rangerte hvert ord basert på grad av relevans og viktighet i undervisningen. Disse analysene lå til grunn for utvelgelsen av ordene i NSL. De utvalgte ordene var ord barn ikke nødvendigvis eksponeres for gjennom dagligdagse samtaler (for eksempel tolke, moden, kolonne, likhet). Den første versjonen av NSL ble ferdigstilt i 2011. Kartleggingsverktøyet hadde da vært igjennom en pilotering (94 barn ble kartlagt), hvor sammenhenger mellom resultater på NSL opp mot resultater på de standardiserte språktestene TROG-2 (Bishop, 2003; Lyster & Horn, 2009) og delprøven Ordforståelse fra Wechsler Intelligence Scale for Children, 4th edition (WISC-IV; Wechsler, 2003) ble undersøkt for å måle kriterievaliditet. Det ble funnet sterke korrelasjoner, både mellom NSL og TROG-2 ($r = .70$) og mellom NSL og Ordforståelse ($r = .79$). NSL ble deretter normert på 245 barn (Frøyen et al., 2011). Etter at NSL hadde vært i bruk i over tre år, ble det foretatt en ny normering ($N = 373$). Denne renormeringen var hovedsakelig basert på tilbakemeldinger fra praksisfeltet og med et formål om å supplere antall elever i normeringsgrunnlaget. Det ble i tillegg foretatt mindre endringer av et utvalg bilder og i protokollen, samt enkelte presiseringer av instruksjonene i manualen. Den nye versjonen av NSL ble ferdigstilt i 2015 (Frøyen et al., 2015).

Formålet med studien

Intensjonen med denne studien var å gjennomføre en vurdering av komponentene i NSL (Frøyen et al., 2011; 2015) og grad av samsvar mellom NSL og andre språkkartleggingsinstrumenter. I den reviderte utgaven av NSL (Frøyen et al., 2015) ble det ikke foretatt nye sammenlikninger med andre standardiserte språktester. I forbindelse med en intervensjonsstudie ble NSL imidlertid brukt for å identifisere minoritetsspråklige barn med behov for språkstøtte. Her ble det også benyttet andre språkkartleggingsverktøy for å vurdere barnas ferdigheter (Heller, Lervåg, & Grøver, 2019). Ved å inkludere utvalget fra intervensjonsstudien sammen med normeringsutvalget til NSL fra 2014, forsøker denne studien spesielt på å besvare følgende forsknings spørsmål:

1. I hvilken grad er den indre konsistensen i NSL tilfredsstillende (reliabilitet)?
2. Hvordan er strukturen i NSL best forklart – gjennom en én-faktormodell eller en flerfaktormodell (begrepsvaliditet)?
3. I hvilken grad viser NSL invarians over språkstatus, kjønn og klassetrinn (begrepsvaliditet)?
4. I hvilken grad er det samsvar mellom NSL og andre standardiserte språkkartlegginger (kriterievaliditet)?

METODE

Deltakere

I denne studien ble det benyttet to utvalg. Det første utvalget var normeringsgrunnlaget til NSL fra 2014 som bestod av 373 elever (1.-4.trinn) fordelt på 41 skoler. 60.1 % av barnegruppen var enspråklige (47,5 % jenter, gjennomsnittsalder = 7 år og 7 måneder, *standardavvik* = 14,00). Denne selekteringen var basert på at utvalget skulle gjenspeile den flerkulturelle populasjonen i skolen. Det ble gjennomført et tilfeldig utvalg av åtte skoler fra hver skolegruppe i Oslo (19 bydeler er fordelt på seks skolegrupper med totalt 48 skoler), med krav om at andelen minoritetsspråklige elever skulle utgjøre 20-40 % av utvalget. Åtte elever fra hver skole fordelt på to trinn (1. og 3.trinn eller 2. og 4.trinn) ble tilfeldig valgt ut etter innhenting av foresattes samtykke. Totalt 40 minoritetsspråk var representert, hvor urdu, engelsk, arabisk, somali, vietnamesisk og spansk utgjorde 44,1 %. Åtte barn ble kartlagt på hver av skolene, enten en kombinasjon av 1. og 3.trinnselever eller av 2. og 4.trinnselever. Ingen av barna hadde vedtak om spesialundervisning i norsk (Opplæringslovens § 5-1), og ingen av de minoritetsspråklige barna hadde vedtak om særskilt norskopplæring (Opplæringslovens § 2-8). Barn med kjente språkvansker eller mangelfulle erfaringer med norsk språk ble ekskludert siden formålet med NSL er å vurdere elevenes norskspråklige ferdigheter opp mot hva som er forventet for alderen.

Utvalg II bestod av 137 barn (1. og 2.trinn) fordelt på 16 skoler, alle minoritetsspråklige elever (54 % jenter, gjennomsnittsalder = 6 år og 3 måneder, *standardavvik* = 6,20). NSL ble benyttet som screener for å identifisere elever med mangelfulle norskspråklige ferdigheter for deltakelse i en randomisert kontrollert studie som vurderte effekten av en språkindervensjon (Heller et al., 2019). Minoritetsspråklige barn som skåret 1,5 standardavvik eller mer under gjennomsnittet på NSL ble invitert til å delta i studien. Åtte til tolv barn fra hver av skolene ble tilfeldig valgt ut etter innhenting av foresattes samtykke. Totalt 31 språk var repre-

sentert, hvor urdu, arabisk, somali, kurdisk, tyrkisk og tamilsk utgjorde 67,9 %. Alle elevene kvalifiserte for særskilt norskopplæring etter Opplæringslovens § 2-8.

Prosedyre

Datagrunnlaget for denne studien ble innhentet over to perioder, høsten 2014 og høsten 2016. Elevene i utvalg I ble kartlagt med NSL høsten 2014. Kartleggingen ble foretatt på skolen og gjennomført av pedagogisk-psykologiske rådgivere med testkompetanse. Samtlige testledere hadde gjennomgått opplæringskurset i NSL og benyttet kartleggingsverktøyet i praksis i forkant av datainnsamlingen. Utvalg II ble kartlagt med NSL og de fem standardiserte språktestene BPVS-II (Dunn et al., 1997; Lyster et al., 2010), Ekspressivt Ordforråd fra Clinical Evaluation of Language Fundamentals, 4th edition (CELF-4; Semel, Wiig & Secord, 2003), Ordforståelse fra WPPSI-IV/WISC-IV (Wechsler, 2003, 2012), TROG-2 (Bishop, 2003; Lyster & Horn, 2009) og Bus Story (Renfrew, 1991) høsten 2016. Kartleggingen ble gjennomført på skolen av førsteforfatter, som tidligere har jobbet som pedagogisk-psykologisk rådgiver.

Måleinstrumenter

Norsk som læringspråk (NSL; Frøyen et al., 2011, 2015) består av 195 testledd fordelt på 10 delprøver og måler elevens ferdigheter innenfor kategorisering, vokabular (bredde og dybde) og grammatikk. Oppgavene stiller krav til språkforståelse (impressive ferdigheter) og muntlige språkferdigheter (ekspressive ferdigheter). Kategorisering vurderer evnen til å se relasjoner og systematisere ord ved at barnet peker på hvilke bilder som hører sammen i semantiske kategorier. Ferdighetene måles gjennom delprøvene matching av kategorier (MK, 45 testledd) og generalisering av kategorier (GK, 20 testledd). Vokabularbredden måles gjennom delprøvene *ordforståelse impressiv* (OFI, 29 testledd) og *ordforståelse ekspressiv* (OFE, 39 testledd), hvor elevene peker på eller benevner ett og ett bilde. Vokabulardybden vurderes ved hjelp av delprøvene *antonymer* (ANT, 14 testledd) og *synonymer* (SYN, 16 testledd) som stiller krav til å sammenlikne likheter og forskjeller mellom ord på et mer abstrakt nivå. Grammatiske ferdigheter vurderes gjennom impressiv setningsforståelse, hvor eleven får opplest én og én setning for så å peke ut riktig bilde til setningen. Det er delprøver for *adjektiv* (ADJ, 13 testledd), *preposisjoner* (PRE, 7 testledd), *pronomen* (PRO, 7 testledd) og *adverb* (ADV, 5 testledd). Alle delprøvene i NSL er bildebaserte med unntak av synonymer. Ett poeng gis for hvert riktig svar og null poeng for feil svar. Kartleggingen tar i gjennomsnitt 30 minutter og administreres individuelt. NSL er utviklet til

bruk av lærere som et verktøy i tilretteleggingen av opplæringen, mens de fem språktestene som beskrives nedenfor hovedsakelig er utviklet for diagnostisk bruk og har andre krav til testkompetanse.

British Picture Vocabulary Scale, 2nd edition (BPVS-II; Dunn et al., 1997; Lyster et al., 2010) er en impressiv språktest hvor barnet får i oppgave å identifisere et ord som blir lest opp av testleder ved å peke ut riktig bilde basert på fire alternativer. Det gis ett poeng for hvert riktig svar og null poeng for feil svar. Den norske utgaven av BPVS-II består av 144 testledd. Delprøven *impressiv ordforståelse* i NSL stiller i likhet med BPVS-II krav til at eleven peker ut riktig bilde etter at testleder har sagt ordet.

Clinical Evaluation of Language Fundamentals, 4th edition (CELF-4; Semel et al., 2003). Delprøven *Ekspressivt Ordforråd* kartlegger elevens muntlige benevningsferdigheter. Testleder viser ett og ett bilde som barnet deretter skal benevne. Ett poeng gis for hvert riktig svar og null poeng for feil svar. Delprøven har 20 testledd. I NSL er delprøven *ordforståelse ekspressiv* konstruert på tilsvarende måte ved at eleven sier ordet til bildet som presenteres.

Wechsler Preschool and Primary Scale of Intelligence, 4th edition (WPPSI-IV; Wechsler, 2012) / *Wechsler Intelligence Scale for Children, 4th edition* (WISC-IV; Wechsler, 2003). Delprøven *Ordforståelse* måler elevens evne til å gi en definisjon eller en forklaring av et ord (ekspressiv). Testleder leser opp ett og ett ord for barnet. Det gis inntil to poeng for hvert riktige svar og null poeng for feil svar. Delprøven inneholder 30 testledd. Med bakgrunn i barnas språkferdigheter, ble en kombinasjon bestående av ord fra *Ordforståelse* i WPPSI-IV og WISC-IV benyttet. Denne versjonen har tidligere blitt brukt i norske studier som vurderte effekten av språkintervensjoner (Hagen, Melby-Lervåg, & Lervåg, 2017; Rogde, Melby-Lervåg, & Lervåg, 2016).

Test for Reception of Grammar, 2nd edition (TROG-2; Bishop, 2003; Lyster & Horn, 2009) er en impressiv språktest som kartlegger elevens forståelse av grammatikk. Setninger blir lest opp én etter én, og barnets oppgave er å peke ut riktig bilde basert på fire alternativer. Ett poeng gis for hvert riktig svar og null poeng for feil svar. Testen består av 80 testledd. Grammatikkoppgavene i NSL har samme format som TROG-2.

Bus Story (Renfrew, 1991) er en narrativ test som måler elevens evne til å gjenfortelle en historie etter å ha fått den

opplest (ekspressiv). Testleder leser historien for eleven ved hjelp av bildestøtte (12 bilder). Det gis inntil to poeng for hver ytring (basert på presisjonsnivået i barnets gjentelling) og null poeng for feil svar. Testen innehar 32 testledd.

Analyser

Råskårer ble benyttet i alle analysene. For å besvare det første forskningsspørsmålet, ble det først foretatt analyser av den indre konsistensen (samvariasjon mellom testleddene) i NSL ved bruk av Cronbachs alpha (verdier over .70 vurderes som tilfredsstillende; Kline, 2016). Det andre forskningsspørsmålet ble adressert gjennom konfirmerende faktoranalyser for å undersøke strukturen i NSL. Analysene ble utført i *Mplus* versjon 8 (Muthén & Muthén, 2017). Resultatene fra de konfirmerende faktoranalysene ble deretter vurdert basert på kriterier for faktorladninger (faktorladninger $\geq .30$ eller $.40$ blir ofte regnet som betydningsfulle) og tilpassningsindekser som gir informasjon om i hvilken grad den estimerte faktormodellen passer til de observerte dataene (Brown, 2015). Modelltilpasning ble vurdert på bakgrunn av *kji*-kvadrat verdier, samt gjennomsnittsverdier for the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), the Root Mean Square of Error of Approximation (RMSEA) og the Standardized Root Mean Square Residual (SRMR). Indikasjoner på god modelltilpasning er CFI $\geq .95$, TLI $\geq .95$, RMSEA $\leq .06$ og SRMR $\leq .08$ (Hu og Bentler, 1999). *Kji*-kvadrat differansetest ble brukt for å teste forskjellen mellom faktormodellene, det vil si å teste hvorvidt korrelasjonene mellom faktorene er signifikant forskjellig fra 1. Siden det er diskutabelt hvor god *kji*-kvadrat differansetesten er til å teste mellom modeller med ulikt antall faktorer (det vil si testing på grensen av det mulige området – hvorvidt en korrelasjon er forskjellig fra 1.00) ble det supplert med Bayesian informasjonskriterium (BIC). BIC brukes for å sammenlikne modeller, hvor

modellen med lavest BIC verdi har best tilpasning (Kline, 2016). Det tredje forskningsspørsmålet ble besvart gjennom analyser av invarians ved bruk av flergroupe konfirmerende faktoranalyser. For 2014-utvalget ble det gjennomført analyser for språkstatus, kjønn og alderstrinn, mens det for 2016-utvalget kun var mulig å analysere invarians for kjønn på grunn av utvalgsstørrelsen (7,8 % var 1.trinnselever). Invarians er funnet når *kji*-kvadrat differansetesten ikke viser signifikante forskjeller mellom tilsvarende modeller hvor faktorladninger (metrisk invarians) og/eller intersept (skala invarians) er enten tvunget til å være 1) lik eller 2) fritt estimert på tvers av gruppene. Ved invarians konkluderer en med at faktorene/begrepene defineres eller forstås på samme måte på tvers av gruppene. Det fjerde forskningsspørsmålet om kriterievaliditet ble besvart ved å gjennomføre korrelasjonsanalyser mellom NSL og de fem standardiserte språktestene (BPVS-II, Ekspressivt Ordforråd fra CELF-4, Ordforståelse fra WPPSI-IV/WISC-IV, TROG-2 og Bus Story).

Resultater

Deskriptiv informasjon om gjennomsnitt, standardavvik, skjevhet, kurtose og reliabilitet er fremstilt i Tabell 1, mens Tabell 2 viser korrelasjonene mellom delprøvene i NSL. Som det fremkommer av Tabell 2 er det forskjeller mellom de to utvalgene når det gjelder korrelasjoner mellom delprøvene i NSL. Delprøvene *matching av kategorier* og *generalisering av kategorier* korrelerer i liten grad med de øvrige delprøvene i 2016-utvalget (minoritetsspråklige elever med mangelfulle norskspråklige ferdigheter) sammenliknet med 2014-utvalget (enspråklige og minoritetsspråklige elever uten vedtak om § 5-1 i norsk eller § 2-8). Når det gjelder det første forskningsspørsmålet, i hvilken grad den indre konsistensen i NSL er tilfredsstillende, viser resultatene tilfredsstillende indre konsistens for begge utvalgene med Cronbachs alfa over .70 (se Tabell 1).

Tabell 1: Gjennomsnitt, standardavvik og reliabilitet for 2014-utvalget og 2016-utvalget (én-faktormodell og fire-faktormodell)

| | 2014-utvalget | | | | | 2016-utvalget | | | | |
|---------------------|---------------|-------|----------|---------|--------------------|---------------|-------|----------|---------|--------------------|
| | Gj.snitt | SA | Skjevhet | Kurtose | Cronbachs α | Gj.snitt | SA | Skjevhet | Kurtose | Cronbachs α |
| Én-faktor modell | 129.99 | 23.55 | -.378 | -.298 | .959 | 64.01 | 19.80 | -.512 | .584 | .939 |
| Firefaktor modell | | | | | | | | | | |
| Kategorisering | 42.57 | 10.04 | -.790 | .381 | .922 | 24.87 | 9.47 | .040 | .064 | .911 |
| Vokabular | 47.72 | 8.50 | -.496 | .203 | .899 | 19.92 | 10.50 | -.139 | -1.015 | .934 |
| Antonymer/synonymer | 11.09 | 6.23 | .423 | -.519 | .898 | 1.68 | 2.15 | 1.840 | 4.435 | .730 |
| Grammatikk | 28.62 | 3.12 | -.827 | 1.026 | .751 | 17.45 | 6.45 | -1.261 | 1.293 | .877 |

Tabell 2: Korrelasjoner mellom delprøver for 2014-utvalget (under diagonalen) og 2016-utvalget (over diagonalen)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1. Matching av kategorier (MK) | 1 | .195* | .003 | .008 | .138 | .175* | .067 | .054 | .076 | .140 |
| 2. Generalisering av kategorier (GK) | .391** | 1 | .080 | .044 | .085 | .169* | .157 | .266** | .148 | .214* |
| 3. Ordforståelse impressiv (OFI) | .448** | .356** | 1 | .489** | .191* | .121 | .452** | .264** | .508** | .394** |
| 4. Ordforståelse ekspressiv (OFE) | .446** | .355** | .517** | 1 | .263** | .420** | .514** | .395** | .534** | .473** |
| 5. Antonymer (ANT) | .538** | .395** | .585** | .624** | 1 | .385** | .237** | .154 | .244** | .172* |
| 6. Synonymer (SYN) | .420** | .237** | .369** | .418** | .553** | 1 | .275** | .247** | .329** | .293** |
| 7. Adjektiv (ADJ) | .484** | .401** | .545** | .530** | .574** | .381** | 1 | .624** | .610** | .562** |
| 8. Preposisjoner (PRE) | .481** | .361** | .563** | .503** | .537** | .333** | .575** | 1 | .604** | .605** |
| 9. Pronomen (PRO) | .422** | .369** | .402** | .325** | .417** | .282** | .426** | .412** | 1 | .477** |
| 10. Adverb (ADV) | .352** | .300** | .419** | .436** | .464** | .340** | .447** | .433** | .282** | 1 |

Merknad: * $p < .05$, ** $p < .01$.

Hvordan er strukturen i NSL best forklart – gjennom en én-faktormodell eller en fler-faktormodell?

For å kunne vurdere strukturen i NSL, ble det først estimert faktormodeller for 2014-utvalget, hvor en én-faktormodell for språk bestående av alle de 10 delprøvene i NSL ble estimert. Modelltilpasningen for én-faktormodellen var god ($\chi^2(29, N = 373) = 65.868$, $p < .0012$, CFI = .979, TLI = .972, SRMS = .030, og RMSEA = .049 (90% CI = .030-.067)).

Resultatet betyr at indikatorene, de ti delprøvene i NSL, er meningsfullt relatert til faktoren. Denne én-faktormodellen ble så sammenliknet med en to-faktormodell (kategorisering og språk), deretter en tre-faktormodell (kategorisering, vokabular og grammatikk) og til slutt en fire-faktormodell (kategorisering, vokabular, antonymer/synonymer og grammatikk) (se Tabell 3). Modellsammenlikningen

basert på tilpasningsindeksene og Bayesian informasjonskriterium indikerte at fire-faktormodellen bestående av faktorene kategorisering, vokabular, antonymer/synonymer og grammatikk hadde signifikant bedre modelltilpasning enn modellene med færre faktorer ($p = < .05$) støttet opp av et mindre BIC. Tilpasning til data for denne fire-faktormodellen var svært god ($\chi^2(29, N = 373) = 31.520$, $p < .3413$, CFI = .998, TLI = .997, SRMS = .023, og RMSEA = .015 (90% CI = .000-.043)). Figur 1 viser at faktorladningene ligger godt over anbefalte kriterier ($\geq .30$ eller $.40$) med verdier fra .55 til .91. Korrelasjonene mellom faktorene i fire-faktormodellen varierte fra .82 til .98.

De samme faktormodellene ble deretter estimert for 2016-utvalget (se Tabell 4). I motsetning til 2014-utvalget,

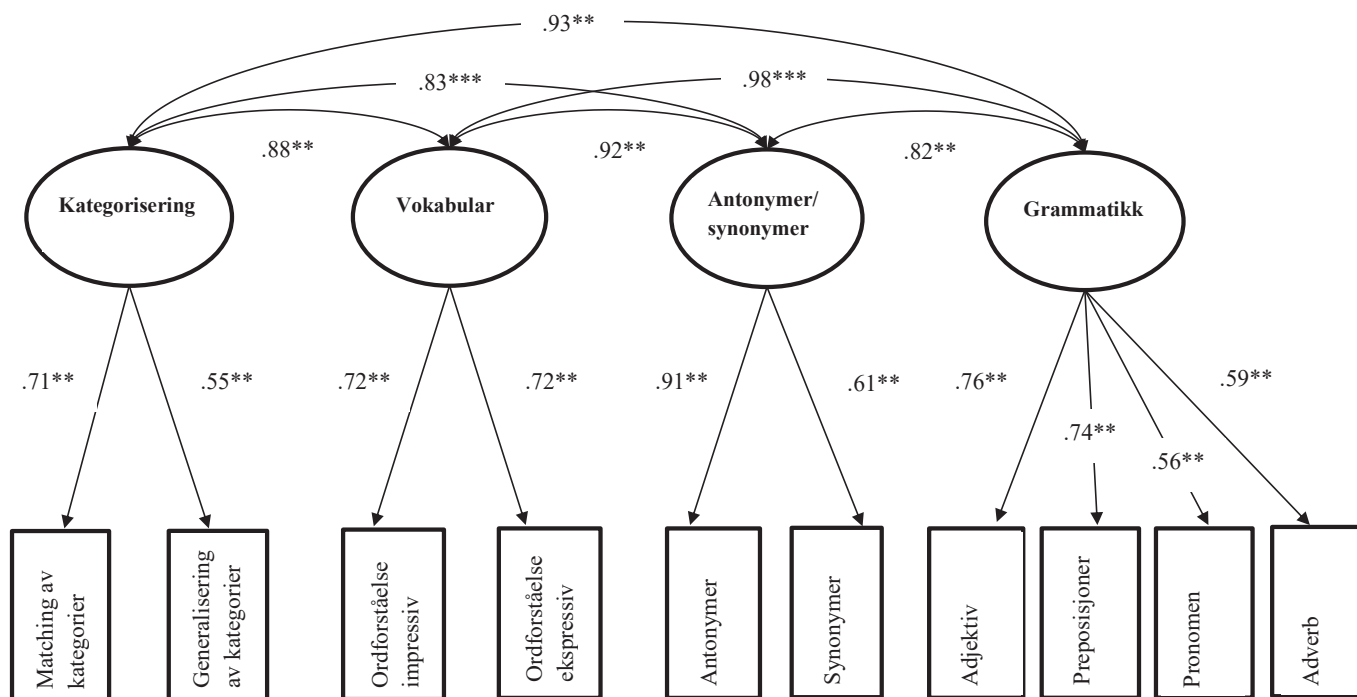
Tabell 3: Faktormodeller for 2014-utvalget

| | χ^2 (df), p | CFI | TLI | RMSEA | SRMR | Adj. BIC |
|----------------|-------------------|------|------|-------|------|----------|
| Faktormodell 1 | 65.868 (35), .001 | .979 | .972 | .049 | .030 | 15785 |
| Faktormodell 2 | 63.283 (34), .002 | .980 | .973 | .048 | .029 | 15786 |
| Faktormodell 3 | 50.048 (32), .022 | .987 | .982 | .039 | .027 | 15778 |
| Faktormodell 4 | 31.520 (29), .341 | .998 | .997 | .015 | .023 | 15768 |

Tabell 4: Faktormodeller for 2016-utvalget

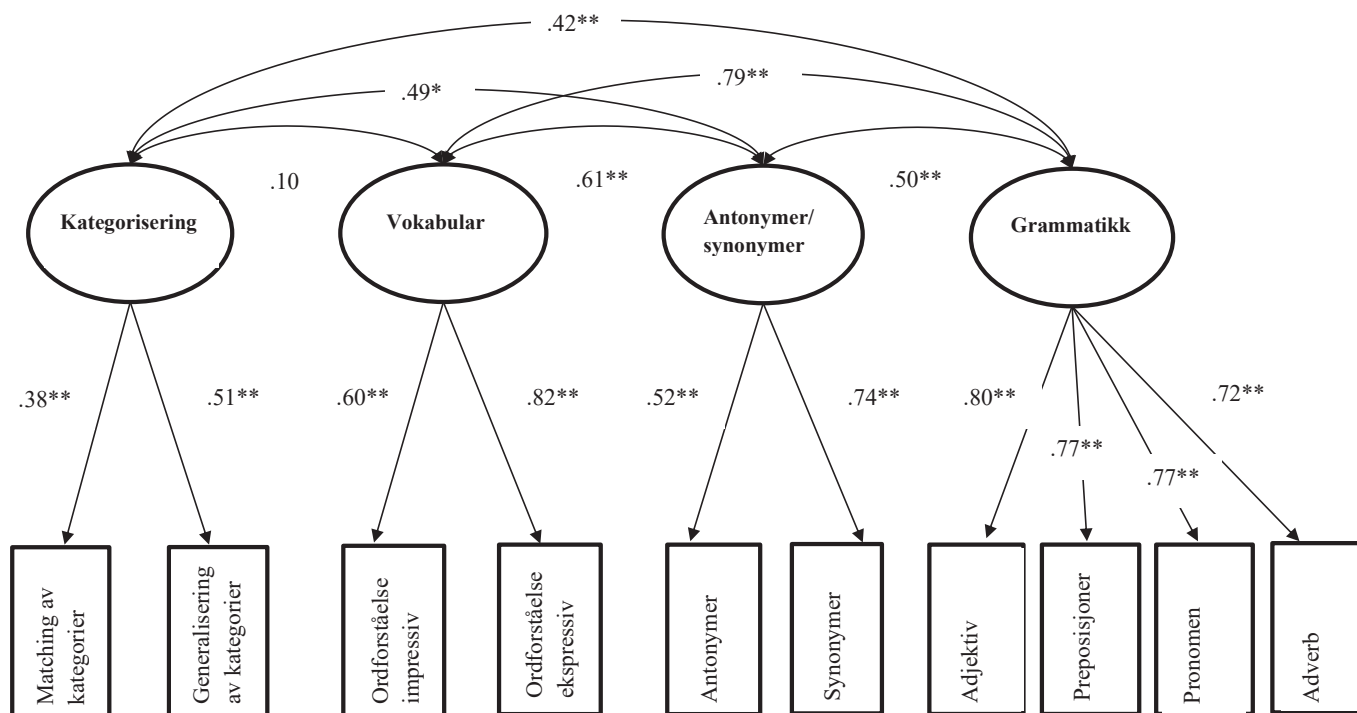
| | χ^2 (df), p | CFI | TLI | RMSEA | SRMR | Adj. BIC |
|----------------|-------------------|------|------|-------|------|----------|
| Faktormodell 1 | 82.871 (35), .000 | .880 | .845 | .100 | .066 | 6264 |
| Faktormodell 2 | 78.556 (34), .000 | .888 | .852 | .098 | .063 | 6261 |
| Faktormodell 3 | 64.347 (32), .001 | .919 | .886 | .086 | .058 | 6251 |
| Faktormodell 4 | 48.213 (29), .014 | .952 | .925 | .070 | .044 | 6240 |

Figur 1: Fire-faktormodell for 2014-utvalget



Merknad: * $p < .05$, ** $p < .01$.

Figur 2: Fire-faktormodell for 2016-utvalget



Merknad: * $p < .05$, ** $p < .01$.

viste én-faktormodellen svak modelltilpasning ($\chi^2(29, N = 137) = 82.871, p < .000, CFI = .880, TLI = .845, SRMS = .066, \text{ og } RMSEA = .100$ (90 % CI = .072-.128)). Også her hadde fire-faktormodellen en signifikant bedre modelltilpasning enn modellene med færre faktorer ($p < .05$) støttet opp av et mindre BIC. Tilpasning til data for denne modellen var god ($\chi^2(29, N = 137) = 48.123, p < .0140, CFI = .952, TLI = .925, SRMS = .044, \text{ og } RMSEA = .070$ (90 % CI = .031-.103)). Som det fremkommer av Figur 2, ligger faktorladningene over de anbefalte kriteriene ($\geq .30$ eller $.40$) med verdier fra $.38$ til $.82$. Korrelasjonene mellom faktorene varierte fra $.10$ til $.79$, hvor faktoren kategorisering ikke korrelerte med vokabular.

I hvilken grad viser NSL invarians over språkstatus, kjønn og klassetrinn?

Invarians, altså i hvilken grad begrepet måles på samme måte over tid eller grupper, ble estimert for hvert av utvalgene. For 2014-utvalget ble det funnet skalar invarians for både kjønn ($\Delta\chi^2(6) = 8.98, p = .174$) og språkstatus ($\Delta\chi^2(6) = 11.369, p = .078$), men ikke klassetrinn ($\Delta\chi^2(12) = 55.571, p = .000$). For 2016-utvalget ble det funnet metrisk invarians for kjønn ($\Delta\chi^2(6) = 7.202, p = .303$), men ikke skalar invarians ($\Delta\chi^2(6) = 17.789, p = .007$).

Funn av skalar invarians i 2014-utvalget gjorde det derfor mulig å sammenlikne gjennomsnittene for kjønn og språkstatus. Barnas språkstatus (enspråklig eller minoritetspråklig) viste statistisk signifikante forskjeller i favør av de enspråklige elevene på vokabular ($d = -.895, p = .000$), antonymer/synonymer ($d = -.446, p = .000$) og grammatikk ($d = -.448, p = .000$), mens det ikke forekom signifikante

forskjeller knyttet til kategorisering ($d = -.203, p = .092$). Enspråklige barn gjorde det bedre på ordforråd og setningsforståelse, mens det ikke var signifikante forskjeller mellom enspråklige og minoritetspråklige elever når det gjaldt kategorisering, en ferdighet som primært stiller krav til logisk resonnering (evnen til å se relasjoner og systematisere ord i semantiske kategorier). Det var ingen signifikante forskjeller mellom jenter og gutter verken på kategorisering ($d = -.009, p = .949$), vokabular ($d = -.045, p = .722$), antonymer/synonymer ($d = -.038, p = .740$) eller grammatikk ($d = -.151, p = .196$).

I hvilken grad er det samsvar mellom NSL og standardiserte språkkartlegginger?

Tabell 5 viser korrelasjonene mellom fire-faktormodellen til 2016-utvalget og de fem standardiserte språkstestene (BPVS-II, Ekspresivt Ordforråd fra CELF-4, Ordforståelse fra WPPSI-IV/WISC-IV, TROG-2 og Bus Story). Som det fremkommer av tabellen, er det moderate til sterke korrelasjoner mellom faktoren vokabular og de standardiserte språkstestene, samt mellom faktoren grammatikk og de standardiserte språkstestene. For faktoren antonymer/synonymer og de standardiserte språkstestene er korrelasjonene svakere, men tilfredsstillende. Det er derimot ikke korrelasjon mellom faktoren kategorisering og språkstestene Ekspresivt Ordforråd fra CELF-4, Ordforståelse fra WPPSI-IV/WISC-IV og Bus Story. Kategorisering viser svake korrelasjoner med BPVS-II og TROG-2.

Diskusjon

Hensikten med denne studien var å undersøke komponentene til kartleggingsverktøyet NSL gjennom analyser av

Tabell 5: Korrelasjoner mellom komponentene i fire-faktormodellen i Norsk som læringsspråk (2016-utvalget) og de standardiserte språkstestene

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--|--------|--------|--------|--------|--------|-------|--------|--------|---|
| 1. British Picture Vocabulary Scale II (BPVS-II) | 1 | | | | | | | | |
| 2. Ekspresivt Ordforråd (Clinical Evaluation of Language Fundamentals, CELF-4) | .557** | 1 | | | | | | | |
| 3. Ordforståelse (WPPSI-IV/WISC-IV) | .541** | .487** | 1 | | | | | | |
| 4. Test for Reception of Grammar (TROG-2) | .643** | .372** | .535** | 1 | | | | | |
| 5. Bus Story | .494** | .522** | .563** | .457** | 1 | | | | |
| 6. Kategorisering (Norsk som læringsspråk) | .228** | .004 | .007 | .303** | .065 | 1 | | | |
| 7. Vokabular (Norsk som læringsspråk) | .531** | .395** | .632** | .537** | .437** | .035 | 1 | | |
| 8. Antonymer/synonymer (Norsk som læringsspråk) | .260** | .211* | .407** | .413** | .302** | .206* | .322** | 1 | |
| 9. Grammatikk (Norsk som læringsspråk) | .545** | .359** | .589** | .621** | .473** | .168* | .617** | .336** | 1 |

Merknad: * $p < .05$, ** $p < .01$.

reliabilitet, begrepsvaliditet og kriterivaliditet. Det ble funnet tilfredsstillende reliabilitet, og resultatene fra de konfirmerende faktoranalysene indikerte at NSL skiller mellom fire komponenter; kategorisering, vokabular, antonymer/synonymer og grammatikk, siden fire-faktormodellen viste den beste modelltilpasningen til dataene i begge utvalgene som inngikk i denne studien. I denne fire-faktormodellen ble det funnet invarians. Det operasjonaliserte begrepet språk ble målt på samme måte for kjønn og språkstatus, men ikke klassetrinn i 2014-utvalget (enspråklige og minoritetsspråklige elever uten vedtak om § 5-1 i norsk eller § 2-8). Det ble i tillegg påvist signifikante positive korrelasjoner mellom NSL og standardiserte språktester (BPVS-II, Ekspresstivt Ordforråd fra CELF-4, Ordforståelse fra WPPSI-IV/WISC-IV, TROG-2 og Bus Story) i 2016-utvalget (minoritetsspråklige elever med mangelfulle norskspråklige ferdigheter).

Konfirmerende faktoranalyser sammenliknet én-faktormodellen (de ti delprøvene i NSL) med tre alternative flerfaktormodeller (to-faktormodell, tre-faktormodell eller fire-faktormodell), hvor fire-faktormodellen viste den beste modelltilpasningen. Det var signifikante og høye korrelasjoner mellom komponentene kategorisering, vokabular, antonymer/synonymer og grammatikk i 2014-utvalget, noe som viser at komponentene er nært relaterte ferdigheter. Ut ifra disse korrelasjonene, kan det diskuteres hvor forskjellige komponentene faktisk er. Det var imidlertid lavere korrelasjoner mellom disse komponentene i 2016-utvalget. Særlig kategorisering viste svakere korrelasjoner med de tre øvrige komponentene, da spesielt med vokabular. En årsak til forskjellen i korrelasjonene mellom utvalgene kan være kriteriene for seleksjon. Mens 2014-utvalget består av hele elevvariasjonen på klassetrinnene som inngår i analysene, har 2016-utvalget kun med språksvake elever. Variasjonen mellom barna i 2016-utvalget er dermed mindre enn i 2014-utvalget, noe som påvirker korrelasjonene mellom faktorene siden disse baseres på variasjonen som finnes i utvalget.

Ved å benytte et utvalg bestående av både enspråklige og minoritetsspråklige elever (2014-utvalget), var det mulig å gjennomføre analyser av invarians i NSL for å se om språk måles på samme måte i begge gruppene. NSL differensierte ikke mellom språkstatus. Tilsvarende manglende forskjeller ble funnet for kjønn (språk ble målt på samme måte hos jenter og gutter), mens NSL ikke målte det samme begrepet over klassetrinn (alder). Sistnevnte er naturlig siden det stilles ulike krav til elevenes ferdigheter, kunnskap og

læringsmål på hvert enkelt klassetrinn. Begrepene som er en del av NSL bygger på det som læres på tvers av fagene norsk, matematikk, samfunnsfag, naturfag og KRLE på ulike trinn. Samtidig er det påvist betydelige forskjeller i språkferdigheter mellom enspråklige og minoritetsspråklige elever (Melby-Lervåg & Lervåg, 2014). I denne studien ble det funnet signifikante forskjeller mellom enspråklige og minoritetsspråklige elever innenfor vokabular, antonymer/synonymer og grammatikk, mens det ikke ble påvist tilsvarende forskjeller for kategorisering. Kategorisering skiller seg fra de tre andre komponentene ved at det primært stilles krav til logisk resonneringsevne i oppgavene. Det innebærer å forstå relasjonen mellom forskjellige gjenstander/objekter og plassere dem inn i abstrakte kategorier (for eksempel fire bilder hvorav tre hører til kategorien kjøretøy). Siden kategorisering ikke differensierer mellom enspråklige og minoritetsspråklige elever, kan det diskuteres hvorvidt det er hensiktsmessig å inkludere denne typen oppgaver i et kartleggingsverktøy hvor formålet er å vurdere språkforskjeller mellom elevene. Samtidig har inkluderingen av kategorisering i denne studien vist at forskjellene mellom enspråklige og minoritetsspråklige elever først og fremst er knyttet til vokabular (bredde og dybde) og grammatikk. En viktig funksjon ved bruk av kartlegging i skolen handler om å få informasjon om elevenes ferdigheter for å tilrettelegge opplæringen på best mulig måte. Siden det per dags dato ikke finnes liknende kartleggingsverktøy til bruk i skolen, er det nyttig å ha med kategorisering i kartleggingen av elevene da den gir informasjon om elevenes evne til å se relasjoner og systematisere ord i semantiske kategorier. Forskjellene i elevenes norskspråklige ferdigheter i NSL indikerer et behov for direkte arbeid med vokabular og grammatiske ferdigheter i opplæringen.

Et annet viktig aspekt ved kvalitetsvurderinger er hvorvidt måleinstrumentet samsvarer med andre tester som er ment å måle det samme teoretiske begrepet (kriterievaliditet) – i dette tilfellet språk. NSL viste tilfredsstillende korrelasjoner med øvrige standardiserte språktester (BPVS-II, Ekspresstivt Ordforråd fra CELF-4, Ordforståelse fra WPPSI-IV/WISC-IV, TROG-2 og Bus Story) som ble foretatt på tre av fire faktorer i 2016-utvalget (vokabular, grammatikk og antonymer/synonymer). Den fjerde faktoren, kategorisering, korrelerte med BPVS-II og Trog-2.

Begrensninger

Det er begrensninger ved studien som bør belyses. 2014-utvalget ble ikke kartlagt med supplerende språktester, og analyser av kriterievaliditet ble derfor kun gjennomført på

2016-utvalget. Dette er et utvalg elever som befant seg i den ene ytterkanten av normalfordelingskurven (1.5 SD eller mer under gjennomsnittet). Samtidig er det viktig å påpeke at det ble funnet kriterievaliditet under piloteringen av NSL gjennom signifikant høye korrelasjoner mellom NSL og TROG-2 samt mellom NSL og delprøven Ordforståelse fra WISC-IV. Disse resultatene var basert på et normalfordelt utvalg (Frøyen et al., 2011).

Oppsummerende refleksjoner og pedagogiske implikasjoner

Språkferdigheter er viktige byggesteiner for barns tilegnelse av kunnskap ved at det gjør det mulig å uttrykke seg og forstå hva et enkeltord betyr, informasjonen som formidles i en setning, innholdet i en tekst eller i samtaler. Et springende punkt innenfor dagens utdanningssystem handler om utvikling av kartleggingsverktøy som lærerne kan bruke for identifisering av barn som trenger språkstøtte på andrespråket (Kunnskapsdepartementet, 2019). Dersom elever med mangelfulle norskspråklige ferdigheter ikke fanges opp tidlig, vil det kunne ha konsekvenser for læringsutbyttet, da videreutvikling av skolefaglige ferdigheter bygger på det språklige grunnlaget et barn har. For å gi et bedre opplæringstilbud som på sikt kan bidra til å redusere det akademiske gapet mellom elevene, er det viktig å kartlegge hvilke språkferdigheter det bør gis direkte opplæring i.

En forutsetning for at kartleggingsverktøy skal fungere som et redskap for å identifisere barn som har behov for språkstøtte i skolen, er at slutningene som trekkes fra resultatene er valide. Denne studien viste at NSL har tilfredsstillende reliabilitet, begrepsvaliditet og kriterievaliditet og dermed så langt kan sees som egnet til kartlegging av norskspråklige ferdigheter på 1.-4.trinn. NSL viser at det er spesielt faktoren vokabular som skiller mellom enspråklige og minoritetsspråklige elever, tett etterfulgt av faktorene grammatikk og antonymer/synonymer. Det er imidlertid ingen forskjeller mellom elevgruppene når det gjelder faktoren kategorisering. KLASSETRINNBASERTE normer er nødvendige ved tolkning av resultatene. Læreplanorienterte kartleggingsverktøy som NSL kan derfor være et viktig supplement i opplæringen når det gjelder å finne ut hvilke språkferdigheter en elev behersker, hvilke ferdigheter eleven er på vei til å mestre og hvilke tiltak som bør iverksettes for å redusere det språklige gapet. Når en elev får lave resultater på NSL, bør skolen sette inn tiltak for å støtte utviklingen av norskspråklige ferdigheter. Elevens utbytte av tiltakene vil avgjøre behovet for ytterligere tilpasninger av opplæringstilbudet og om språkferdighetene bør utredes nærmere.

Referanser

- Amesen, A., Braeken, J., Ogden, T. & Melby-Lervåg, M. (2019). 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*, 63(3), 465-490. doi:10.1080/00313831.2017.1420685
- Berko Gleason, J. (2009). The development of language: an overview and a preview. I J. Berko Gleason & N.B. Ratner (Red): *The Development of Language* (7. utgave) (s. 1-34). Boston, MA: Pearson Education.
- Bishop, D. M. (2003). *Test for Reception of Grammar version 2 (Trog-2 manual)* [Norwegian translating and norms: Lyster, S.-A.H. & Horn, E., 2009]. Stockholm, Sverige: Pearson
- Bjerkas, K. M., Monsrud, M. B. & Thurmann-Moe, A. C. (2013). *Ordforråd hos flerspråklige barn: Pedagogiske og spesialpedagogiske utfordringer*. Oslo: Gyldendal Akademisk.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press.
- Dickinson, D. K. & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, MD: Brookes.
- Dickinson, D. K. & Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development*, 82(3), 870-886. doi:10.1111/j.1467-8624.2011.01576.x
- Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997). *British Picture Vocabulary Scale II*. Windsor, UK: National Foundation for Educational Research-Nelson.
- Frøyen, W., Ahmadinia, S., Heller, M. C. & Skjåk, O. (2011). *Norsk som lærings-språk*. NSL, Oslo, Norge: PPT Oslo, Utdanningsetaten.
- Frøyen, W., Ahmadinia, S., Heller, M. C., & Skjåk, O. & Namvar, N. (2015). *Norsk som læringsspråk redigert, NSL-R*. Oslo, Norge: PPT Oslo, Utdanningsetaten.
- Hagen, Å. M., Melby-Lervåg, M., & Lervåg, A. (2017). Improving language comprehension in preschool children with language difficulties: A cluster randomized trial. *Journal of Child Psychology and Psychiatry*, 58(10), 1132-1140. doi:10.1111/jcpp.12762
- Hart, B. & Risley, T. (1995). *Meaningful differences in the everyday life of young American children*. Baltimore: Paul Brookes.
- Hayiou-Thomas, M. E., Dale, P. S. & Plomin, R. (2012). The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Developmental Science*, 15(2), 233-249. doi:10.1111/j.1467-7687.2011.01119.x
- Heller, M. C., Lervåg, A., & Grøver, V. (2019). Oral language intervention in Norwegian schools serving young language-minority learners: a randomized trial. *Reading Research Quarterly*, 54(4), 531-552. doi:10.1002/rrq.248
- Hjetland, H., Brinchmann, E. I., Scherer, R. & Melby-Lervåg, M. (2017). Preschool predictors of later reading comprehension ability: a systematic review. *Campbell Systematic Reviews*, 13(1). doi:10.4073/csr.2017.14
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55-88. doi:10.1016/j.dr.2005.11.002
- Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. doi:10.1080/10705519909540118
- Klem, M., Melby-Lervåg, M., Hagtvet, B. E., Lyster, S. A. H., Gustafsson, J. E. & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, 18, 146-154.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th edition). New York: The Guilford Press.
- Kunnskapsdepartementet (2019). Tett på – tidlig innsats og inkluderende fellesskap i barnehage, skole og SFO. Meld. St. 6 (2019-2020). Oslo: Departementets servicesenter.

- Lervåg, A., Hulme, C. & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: it's simple but complex. *Child Development*, 89(5), 1821-1838. doi:10.1111/cdev.
- Lyster, S. A. H. & Horn, E. (2009). Trog-2: Norsk versjon [Test for Reception of Grammar version 2 (Trog-2) manual]. Windsor, UK: GL Assessment.
- Lyster, S. A. H., Horn, E. & Rygvold, A. L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge: Resultater fra en utprøving av British Picture Vocabulary Scale, Second Edition (BPVS-II). *Spesialpedagogikk*, 9, 35-43.
- Marulis, L. M. & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300-355. doi:10.3102/0034654310377087
- Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: a meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, 140(2), 409-433. doi:10.1037/a0033890
- Melby-Lervåg, M., Lervåg, A., Lyster, S. A. H., Klem, M., Hagtvet, B. E. & Hulme, C. (2012). Nonword repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, 23, 1092-1098. doi:10.1177/0956797612443833
- Messick, S. (1995). Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. doi:10.1037/0003-066x.50.9.741
- Muthén, L. K. & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Paradis, V., Genesee, F. & Crago, M. B. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2. utgave). Baltimore, Maryland: Paul Brookes Publishing
- Putnik, D. L. & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review*, 41, 71-90. doi:10.1016/j.dr.2016.06.004
- Renfrew, C. E. (1997). *Bus story test - A test of narrative speech* (4th ed.). Bicester, England: Winslow Press.
- Rogde, K., Melby-Lervåg, M., & Lervåg, A. (2016). Improving the general language skills of second-language learners in kindergarten: A randomized controlled trial. *Journal of Research on Educational Effectiveness*, 9(1), 150-170. doi:10.1080/19345747.2016.1171935
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83(5), 1762-1774. doi:10.1111/j.1467-8624.2012.01805.x
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: what the research shows. *Language Learning*, 64(4), 913-951. doi:10.1111/lang.12077
- Semel, E., Wiig, E. H. & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.) [Translated and adapted to Norwegian]. Toronto, Canada and Assens, Norway: The Psychological Corporation/Pearson Assessment Company.
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328(5977), 450-452. doi:10.1126/science.182597
- Snow, C. E. & Uccelli, P. (2009). The challenge of academic language. In D.R. Olson and N. Torrance (Eds), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press.
- Statistisk sentralbyrå (2019). Elever med morsmålsopplæring, tospråklig fagopplæring, tilrettelagt opplæring og særskild norskopplæring. <https://www.ssb.no/utdanning/statistikk/utgrs>
- Storch, S. A. & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934-947. doi:10.1037/0012-1649.38.6.934
- Thorndike, R. M. & Thorndike-Christ, T. (2014). *Measurement and evaluation in psychology and education* (8th edition), Harlow: Pearson.
- Tomblin, J. B. & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research*, 49(6), 1193-1208. doi:10.1044/1092-4388(2006)086
- Utdanningsdirektoratet (2006). Læreplanverket Kunnskapsløftet (LK06). Oslo: Utdanningsdirektoratet.
- Wechsler, D. (2003). *Wechsler intelligence scale for children administration manual* (4th ed.) [Norwegian administration manual]. San Antonio, TX and Stockholm, Sweden: The Psychological Corporation/Pearson Assessment Company.
- Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence administration manual* (4th ed.) [Norwegian administration manual]. San Antonio, TX and Stockholm, Sweden: The Psychological Corporation/Pearson Assessment Company.

UTLYSNING AV STIPEND FRA NLL

NORSK
LOGOPEDLAG



Visste du at Norsk logopedlag har en stipendordning for støtte til materiellutvikling eller oversetterarbeid?

Norsk logopedlag kan dele ut ett eller flere stipend per landsmøteperiode. Stipendordningen skal være en økonomisk støtte til materiellutvikling, oversetterarbeid eller annet utviklingsarbeid innenfor fagfeltet logopedi. Søkerne må være medlem av Norsk logopedlag. Innkomne søknader behandles og avgjøres av styret i Norsk logopedlag i samarbeid med aktuelle utvalg, i henhold til søknadens faglige innhold. Hver søker kan motta inntil kr. 10.000, og det enkelte medlem kan motta stipendet bare én gang per prosjekt. Tildeling av stipend må til enhver tid vurderes i henhold til lagets økonomi. På NLLs nettside og i tidsskriftet opplyses det om tildelte stipend.

Standard søknadsskjema fylles ut, skannes og sendes til leder i NLL.

Mer informasjon om ordningen og søknadsskjema finner du på NLLs nettsider:

<https://norsklogopedlag.no/stipendordning/>