

Validering av C-BiLLT

*En valideringsstudie av en ny språkforståelsestest
tilpasset barn med store tale- og bevegelsesvansker*

Sara Fiske



Masteroppgave i spesialpedagogikk
Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

Vår 2018

Validering av C-BiLLT

*En valideringsstudie av en ny språkforståelsestest
tilpasset barn med store tale- og bevegelsesvansker*

Av Sara Fiske

Masteroppgave ved Institutt for spesialpedagogikk
UNIVERSITETET I OSLO

Vår 2018

© Sara Fiske

2018

Validering av C-BiLLT; En valideringsstudie av en ny språkforståelsestest tilpasset barn med store tale- og bevegelsesvansker

Sara Fiske

<http://www.duo.uio.no>

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Bakgrunn, formål og problemstilling: Flere studier viser at mange barn med store tale- og bevegelsesvansker ikke har mottatt ordinær kartlegging av evner som språkforståelse, til tross for at disse barna har behov for tiltak og alternativ og supplerende kommunikasjon (ASK) for å få tilgang til læring og mulighet for å delta i sitt miljø. Da det ikke er et en-til-en-forhold mellom hva barn med store tale- og bevegelsesvansker kan uttrykke av språk og hva de forstår av språk, kan antagelser om disse barnas språkforståelsesnivå både undervurdere og overvurdere barnets faktiske evner og tiltak som bygger på slike antagelser kan bli lite tilpasset barnets behov. *Computer-Based Instrument for Low motor Language Testing* (C-BiLLT) er en nederlandsk språkforståelsestest utviklet ut fra behovet for en test egnet i kartleggingen av barn hvis store tale- og bevegelsesvansker vanskeliggjør testbesvarelse på måten tester er standardisert. C-BiLLT er digital og kan besvares med ulike responsmodaliteter, som fingerpeking, blikkpeking eller skanning ved hjelp av brytere, og den krever dermed ikke verbal respons og minimal motorisk innsats for å besvares. Selv om testen er tilpasset barn med store tale- og bevegelsesvansker kan den også benyttes i kartleggingen av barn uten slike vansker, noe som tilsier at C-BiLLT også kan anvendes for å realisere tidlige innsats. Det er i dag mangel på standardiserte og normerte norske tester som kartlegger både forståelse av vokabular og grammatikk, som kan dekke aldersspennet 1;6 til 7;6 år og som er egnet for kartlegging av barn med store tale- og bevegelsesvansker. Ullevål Universitetssykehus, seksjon for nevrohabilitering – barn, har derfor oversatt C-BiLLT til norsk. Denne studien har som hovedformål å undersøke problemstillingen: *Hvor egnet er den norske versjonen av C-BiLLT til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 til 7;6 år?* Studien er en del av prosjektet ”Bedre liv med Bliss” og prosjektleder er Kristine Stadskleiv.

Metode og materiale: For å undersøke problemstillingen var det ønskelig å statistisk analysere 1;6 til 7;6 år gamle barns skårer på C-BiLLT samt å sammenligne disse skårene med skårer på andre tester av språkforståelse og ikke-verbal resonneringsevne. En kvantitativ tilnærming med et kryss-seksjonelt design er derfor benyttet, da dette gjør det mulig å samle inn testskårer fra barn i ulike aldre. Deltakerne i denne studien er 187 typisk utviklede barn i alderen 20 til 95 måneder. Alle disse barna ble testet med C-BiLLT. Avhengig av barnets alder ble barna også testet med språkforståelsesdelen av Reynell (Hagtvet & Lillestølen, 1989), Ordgjenkjenning fra Wechsler Preschool and Primary Scale of Intelligence – fjerde

versjon (WPPSI-IV; Wechsler, 2015), og Test of Reception of Grammar-2 (TROG-2; Bishop 2009) som alle er tester av språkforståelse, samt Ravens Coloured Progressive Matrices (CPM; Raven, 2008) og Terningmønster fra WPPSI-IV som begge er tester av ikke-verbal resonneringsevne. All data er analysert med den statistiske programvarepakken IBM SPSS. Deskriptiv statistikk, korrelasjonsanalyser, en enveis variansanalyse (ANOVA) og en eksplorerende faktoranalyse ble gjennomført for å besvare denne oppgavens problemstilling.

Resultater og konklusjoner: Korrelasjonsanalysen viste at korrelasjonen mellom sumskårer på C-BiLLT og de andre testene av språkforståelse var høyere enn korrelasjonen mellom C-BiLLT og CPM, og C-BiLLT og Terningmønster for barna over fire år. Dette ble tolket som at det som måles med C-BiLLT også måles med de andre testene av språkforståelse, men ikke av testene av ikke-verbal resonnering, og at C-BiLLT dermed kan beskrives som et rent mål på barns språkforståelse. Den eksplorerende faktoranalysen resulterte i to ekstraherte faktorer. Disse faktorene korrelerte høyt med hverandre og forklarte begge variasjoner i alle seksjoner. De to faktorene ble tolket som å representere henholdsvis vokabularforståelse og grammatisk forståelse, mens korrelasjonen mellom dem ble tolket som en indikasjon på at begge representerer aspekter ved et bredere språkforståelseskonstrukt og at en enhetlig språklig evne bidrar til variasjoner i alle C-BiLLT-seksjoner. Dette støtter at skårer på seksjonene i C-BiLLT kombineres til én sumskåre tenkt å representere variasjoner i én dimensjon; en generell språkforståelsesevne. Resultatene fra enveis ANOVA viste at det var signifikante forskjeller på C-BiLLT-sumskårer mellom alle helårsgrupper. Deskriptiv statistikk viste videre at ingen av deltakerne oppnådde høyest eller lavest skåre, men at spredningen var noe lav innad i aldersgruppene for barna over fem år. Dette ble tolket som at C-BiLLT er egnet til å differensiere mellom språkforståelsesnivået til barn mellom ulike aldersgrupper og innad i de yngste aldersgruppene samt språklig lavtfungerende barn. Den lave spredningen i aldersgruppene over fem år ble tolket som å enten representere at C-BiLLT er mindre egnet til å differensiere mellom språkforståelsesnivået til de eldste og språklig høytfungerende barna, eller at deltakerne i denne studien var noe homogene og at standardavvikene i disse gruppene dermed ble unaturlig lave.

Til sammen støtter dette at C-BiLLT er godt egnet til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år. Samtidig må det heftes en noe større usikkerhet til hvor egnet C-BiLLT er til å differensiere mellom språkforståelsesnivået til barn over fem år, før C-BiLLT er prøvd ut på et større og mer heterogent utvalg.

Forord

Tusen takk til min hovedveileder og prosjektleder psykologspesialist og ph.d. Kristine Stadskleiv. Jeg er veldig glad for at jeg har fått muligheten til å bidra til dette prosjektet. Jeg er takknemlig både for all praktisk og teoretisk erfaring om kartlegging og testutvikling jeg har fått gjennom dette prosjektet, og motivasjonen i det å vite at alt arbeidet ville munne ut i noe det er et stort behov for. Takk for din veiledning og alt du har delt av din erfaring og kunnskap.

Jeg vil også rette en stor takk til alle barna som deltok i denne studien. Det har vært en glede å møte og bli kjent med dere alle sammen under kartleggingen. Tusen takk til foreldre, barnehagepersonale og skolepersonale for deres positive innstilling og imøtekommenhet.

Tusen takk til min biveileder logoped og ph.d. Melanie Kirmess for konstruktiv kritikk og gode innspill.

Takk til logoped og ph.d. Joke Geytenbeek og logoped Jael Bootsma for hyggelig og informativ kursing i administrering av C-BiLLT og kartlegging av barn med store tale- og bevegelsesvansker. Og tusen takk for tilgang til nederlandsk data.

Tusen takk til min medstudent og venn Anne Lise Haddeland for hyggelig selskap i datainnsamlingsperioden samt for gode faglige diskusjoner underveis i skriveprosessen.

Takk til Mamma og Pappa for støtte og bilturer til avsidesliggende skoler og barnehager. Og takk til Erik for all støtte og oppmuntring i skriveperioden.

Oslo, mai 2018

Sara Fiske

Innholdsfortegnelse

1	Innledning.....	1
1.1	Bakgrunn og formål.....	1
1.1.1	Problemstilling.....	2
1.2	Oppgavens disposisjon og avgrensning.....	3
2	Språkforståelse.....	6
2.1	Komponenter i språkforståelse.....	6
2.2	Distinkte eller integrerte språkkomponenter.....	9
2.2.1	Vokabular og grammatikk; styrt av distinkte mekanismer.....	9
2.2.2	Språklig spesialisering; fremkommende snarere enn medfødt.....	10
2.3	Barn med komplekse kommunikasjonsbehov.....	14
2.3.1	Barn med CP.....	15
2.3.2	Barn med CP og kommunikasjon.....	15
3	Kartlegging av språkforståelse.....	19
3.1	Testteoretiske utfordringer.....	19
3.1.1	Validitet og validering.....	20
3.2	Konkrete kartleggingsutfordringer.....	25
3.3	Språkforståelsestesten C-BiLLT.....	27
4	Metode.....	28
4.1	Design.....	28
4.2	Deltakere.....	29
4.3	Kartleggingsinstrumenter.....	31
4.3.1	C-BiLLT.....	32
4.3.2	Språkforståelsesdelen av Reynell.....	39
4.3.3	Ordgjenkjenning fra WPPSI-IV.....	40
4.3.4	TROG-2.....	40
4.3.5	CPM.....	41
4.3.6	Terningmønster fra WPPSI-IV.....	41
4.4	Digitale verktøy.....	42
4.5	Prosedyre.....	43
4.5.1	Testoversettelsesprosessen.....	43
4.5.2	Testing.....	45
4.5.3	Retesting.....	47
4.6	Statistikk.....	48
4.6.1	Inspisering av data og normalitet.....	48
4.6.2	Valg av hovedanalyser.....	50
4.7	Etikk.....	52
5	Resultater.....	55
5.1	Undersøkelse av korrelasjon mellom tester.....	55
5.2	Utforskning av faktorstrukturen til C-BiLLT.....	56
5.3	Testskårer ved ulike aldersgrupper.....	59
6	Diskusjon.....	62
6.1	Hovedfunnene fra reliabilitetsundersøkelsen.....	62
6.2	Evidens basert på relasjoner til andre variabler.....	63
6.2.1	Konvergerende evidens.....	63

6.2.2	Diskriminerende evidens	66
6.3	Evidens basert på indre struktur.....	68
6.3.1	Endimensjonalitet eller flerdimensjonalitet?	68
6.3.2	Er kryssladninger et problem for testtolkning?.....	72
6.3.3	Hva reflekterer lav kommunalitet?	75
6.4	Evidens basert på forskjeller mellom og innad i aldersgrupper	76
6.5	Sammenstilling av evidens.....	81
6.6	Styrker og begrensninger	83
6.7	Kliniske implikasjoner	85
7	Avslutning	87
7.1	Veien videre.....	89
	Litteraturliste	90
	Vedlegg 1: Samtykkeskjema	102
	Vedlegg 2: Informasjonsskriv.....	105
	Vedlegg 3: Godkjennelse fra NSD.....	106
	Vedlegg 4: Nederlandsk data.....	108
	Vedlegg 5: Korrelasjoner mellom seksjoner	109
	Vedlegg 6: Gjennomsnittlig riktig svarprosent for alle testledd i C-BiLLT	110

Figurliste

Figur 1.	Modell over nivåer og relaterte prosesser i språkforståelse.....	6
Figur 2.	En forenklet modell av de tidlige nivåene i språkforståelse som viser interaksjoner mellom alle nivåer.....	12
Figur 3.	Testledd 44 fra C-BiLLT	32
Figur 4.	Læringsmodulen fra C-BiLLT	33
Figur 5.	Testledd 1 fra Del 1 av C-BiLLT	34
Figur 6.	Testledd P9 fra parallellseksjonen av C-BiLLT.....	34
Figur 7.	Testledd 38 fra Del 2 av C-BiLLT.....	35
Figur 8.	Øvelsesoppgave 2 fra Del 2 av C-BiLLT	35
Figur 9.	Testoppsett 1.....	42
Figur 10.	Testoppsett 2 (t.v.) og Testoppsett 3 (t.h.).....	43
Figur 11.	Testledd 10 fra C-BiLLT	44
Figur 12.	Nye bilder av en postkasse (øverst) og brunost (nederst) i den norske versjonen av C-BiLLT	45
Figur 13.	Boxplot over C-BiLLT sum av råskårer i ulike helårsgrupper	50
Figur 14.	Scree plot fra den eksplorerende faktoranalysen	56
Figur 15.	Linjediagram over fordeling av gjennomsnittlig sumskårer i helårsgrupper	61
Figur 16.	Testledd 42 fra Seksjon 6 av C-BiLLT.....	74
Figur 17.	Testledd 39 fra Seksjon 5 av C-BiLLT.....	74
Figur 18.	Testledd 47 fra C-BiLLT	79

Tabelliste

Tabell 1. Fordeling av barn i ulike aldersgrupper	30
Tabell 2. Oversikt over foreldre rapporterte tilstander	30
Tabell 3. Oversikt over andre morsmål hos studiens deltakere	31
Tabell 4. Seksjoner fra C-BiLLTs datatest med eksempler på testledd.	36
Tabell 5. Reliabilitet og validitet tilknyttet den nederlandske versjonen av C-BiLLT	38
Tabell 6. Testrekkefølge A og B	46
Tabell 7. Fordeling av testadministratorer og responsmodaliteter	47
Tabell 8. Deskriptiv statistikk og skjevhets- og kurtosisverdier for C-BiLLT sumskårer fordelt på helårsgrupper	48
Tabell 9. Deskriptiv statistikk og skjevhets- og kurtosisverdier for z-skårer	49
Tabell 10. Oversikt over korrelasjoner mellom råskårer på ulike tester	55
Tabell 11. Mønstermatrise fra den eksplorerende faktoranalysen	57
Tabell 12. Strukturmatrisen fra den eksplorerende faktoranalysen	58
Tabell 13. Oversikt over forskjeller i gjennomsnittlig C-BiLLT sumskåre mellom halvårsgrupper	59
Tabell 14. Oversikt over forskjeller i gjennomsnittlig C-BiLLT sumskåre mellom helårsgrupper	59
Tabell 15. Økning i gjennomsnittlig sumskåre på C-BiLLT i halvårsgrupper	60
Tabell 16. Økning i gjennomsnittlig sumskåre på C-BiLLT fra den nederlandske utprøvingen av C-BiLLT	60
Tabell 17. Utklipp fra Vedlegg 6 – Testledd 35 til 44 fra Seksjon 5 og 6	73
Tabell 18. Resultater fra den nederlandske faktoranalysen	108
Tabell 19. Resultater fra den nederlandske variansanalysen	108
Tabell 20. Korrelasjoner mellom de 12 seksjonene i C-BiLLT (N = 187)	109

Liste over forkortelser

AERA	American Educational Research Association
APA	American Psychological Association
ASK	Alternativ og Supplerende Kommunikasjon
BPVS	British Picture Vocabulary Scale
C-BiLLT	The Computer-Based Instrument for Low motor Language Testing
CELF	Clinical Evaluation of Language Fundamentals
CFCS	Communication Function Classification System
CP	Cerebral parese
CPM	Coloured Progressive Matrices
GMFCS	Gross Motor Function Classification System
LARRC	Language and Reading Consortium
MACS	Manual Ability Classification System
NCME	National Council on Measurement in Education
NESH	The National Committee for Research Ethics in the Social Sciences and the Humanities
PPVT	Peabody Picture Vocabulary Scale
SCPE	Surveillance of Cerebral Palsy in Europe
TROG-2	Test of Reception of Grammar-2

1 Innledning

1.1 Bakgrunn og formål

Evnen til å forstå og bruke språk i kommunikasjon er en svært sentral del av barns utvikling. Slike språkferdigheter danner et grunnlag for akademisk læring og for sosiale interaksjoner og deltakelse i samfunnet og arbeidslivet (Tomblin, 2008; Murray & Goldbart, 2009).

Vansker med språkforståelse kan således få store konsekvenser for generell læring, utvikling, sosiale relasjoner og arbeidsmuligheter. Stabiliteten i barns rangordnede språkevne har videre en tendens til å øke over tid, noe som fordrer at tiltak og intervensjoner settes inn tidlig (Tomblin, Nippold, Fey, & Zhang, 2014; Bornstein, Hahn, Putnick, & Suwalsky, 2013; Bornstein, Hahn, & Putnick, 2016). For at tiltak skal tilpasses barnets behov bør de bygge på en detaljert kartlegging (Sattler, 2008). En hovedpilar i en slik kartlegging er datainnsamling gjennom administrering og skåring av standardiserte og normerte tester.

En gruppe barn som ofte har behov for kartlegging og tiltak tidlig i sin utvikling er barn med store tale- og bevegelsesvansker. Barn med store tale- og bevegelsesvansker utgjør en heterogen gruppe med variert kommunikativt og kognitivt funksjonsnivå. Forekomsten av motoriske talevansker er hyppig i denne gruppen, og noen barn har så store talevansker at deres talespråk er fraværende (Odding, Roebroek, & Stam, 2006). Flere studier viser imidlertid at barnets språkforståelse utvikles selv når barnet ikke snakker, og at det dermed kan eksistere en diskrepans mellom hva barnet med store tale- og bevegelsesvansker kan uttrykke av språk og hva de forstår av språk (Bishop, Brown, & Robson, 1990; Frampton, Yude, & Goodman, 1998; Straub & Obrzut, 2009; Geytenbeek, Heim, Knol, Vermeulen, & Oostrom, 2015). En kartlegging av barnets språkforståelse er dermed et viktig grunnlag både for å planlegge tilpassede intervensjoner og for å tilrettelegge og implementere Alternativ og Supplerende Kommunikasjon (ASK). Slike tiltak er fundamentalt for at disse barna skal få tilgang til læring og erfaring ved å gjøre det mulig for dem å kommunisere og være aktive deltakere i sitt miljø (Light, Parsons, & Drager, 2002; Murray & Goldbart, 2009; Batorowicz, Statdskleiv, von Tetzchner, & Missiuna, 2016).

Kartlegging av språkforståelse hos barn med store tale- og bevegelsesvansker kan imidlertid være utfordrende, da utredning ofte krever tilrettelegging av standardiserte tester. Som et resultat er det mange barn med store tale- og bevegelsesvansker som ikke kartlegges

(Andersen et al., 2008; Stadskleiv, Vik, Andersen, & Lien, 2015). Den nederlandske språkforståelsestesten C-BiLLT er utviklet av logoped og ph.d. Geytenbeek, logoped Bootsma, barnenevrolog Vermeulen og barnenevropsykolog Oostrom utfra behovet om en språkforståelsestest som valid og reliabelt kan kartlegge språkforståelse hos barn med store tale- og bevegelsesvansker. Testen kartlegger både forståelsen av ord og setninger og har normer for nederlandske barn i alderen 1;6 til 7;0 år. C-BiLLT er en digital test og kan besvares gjennom flere responsmodaliteter som krever minimal motorisk innsats, som fingerpeking, blikkpeking eller skanning ved hjelp av brytere. Selv om testen er særlig utviklet for barn med store tale- og bevegelsesvansker kan den også brukes med barn uten talevansker eller motoriske vansker.

Per i dag finnes det ikke en normert og standardisert norsk test av språkforståelse som kan dekke aldersspennet 1;6 til 7;6 år, som kartlegger både forståelse av vokabular og grammatikk, og som egner seg for barn med alvorlige tale- og bevegelsesvansker. Ullevål Universitetssykehus, seksjon for nevrohabilitering - barn, har derfor oversatt den nederlandske versjonen av C-BiLLT til norsk. Formålet med denne studien er å undersøke validiteten til de slutningene som kan trekkes om språkforståelse hos barn i alderen 1;6 til 7;6 år på bakgrunn av den norske versjonen av C-BiLLT. Validiteten tilknyttet en test må ses i sammenheng med testens reliabilitet. For den norske versjonen av C-BiLLT undersøkes reliabiliteten i masteroppgaven *Språkforståelsestesten C-BiLLT: Vurdering av reliabilitet og sammenligning av ulike responsmodaliteter. En test tilpasset barn med store tale- og bevegelsesvansker* av Anne Lise Haddeland, som leveres ved institutt for spesialpedagogikk vår 2018. Validerings- og reliabilitetsstudien er begge en del av prosjektet "Bedre liv med Bliss" og prosjektleder er psykologspesialist og ph.d. Kristine Stadskleiv.

1.1.1 Problemstilling

Med utgangspunkt i studiens formål er problemstillingen:

Hvor egnet er den norske versjonen av C-BiLLT til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år?

Validitet er et mangefasettert begrep. For å kunne besvare denne problemstillingen har jeg skissert ulike forventninger som i denne oppgaven vil utdypes og diskuteres:

- 1) Det er forventet en høyere korrelasjon mellom skårene på C-BiLLT og skårer på andre tester av språkforståelse enn mellom skårene på C-BiLLT og skårer på tester av ikke-verbal resonnering.
- 2) Det er forventet at en eksplorerende faktoranalyse på C-BiLLT resulterer i samme faktorløsning som for den nederlandske versjonen av C-BiLLT; en endimensjonal faktorløsning som representerer en overordnet språkforståelsesevne.
- 3) Det er forventet at C-BiLLT er sensitiv nok til å fange opp forskjeller i språkforståelsesnivå mellom barn i ulike aldersgrupper og barn i samme aldersgruppe. Dette bør reflekteres i signifikante forskjeller i gjennomsnittlig sumskåre mellom ulike aldersgrupper og fravær av gulveffekter og takeffekter innad i ulike aldersgrupper.

1.2 Oppgavens disposisjon og avgrensning

Denne oppgaven består av syv kapitler. Kapittel 2 til 3 vil ta for seg oppgavens teoretiske bakgrunn. En diskusjon av validitet er avhengig av en forståelse for hva testen er utformet for å måle, som for C-BiLLT er forståelse av tale. I kapittel 2 vil jeg derfor begynne med å redegjøre for fenomenet språkforståelse. Jeg har her valgt å ta utgangspunkt i Bishops (2014) modell for språkforståelse ettersom modellen er egnet både som et rammeverk for redegjørelsen av ulike nivåer i språkforståelse, og for en diskusjon rundt hvor overlappende og interagerende de ulike nivåene er. C-BiLLT er utformet for å kartlegge språkforståelse gjennom både forståelse av ord, eller vokabularforståelse, og setninger, eller forståelse av grammatikk. Ettersom det er forventet at faktoranalysen av C-BiLLT vil resultere i en endimensjonal modell, vil jeg i kapittel 2 også redegjøre for ulike perspektiver og studier på språkevnsens dimensjonalitet. Hovedfokuset vil her ligge på ulike perspektiver på hvorvidt forståelse av ords semantiske innhold, eller vokabularforståelse, og forståelse av grammatikk anses som separate ferdigheter eller sammenfallende komponenter av en overordnet språkevne. C-BiLLT er utformet spesifikt for gruppen barn med store tale- og bevegelsesvansker, og kapittel 2 vil derfor avsluttes med en presentasjon av gruppen barn med store tale- og bevegelsesvansker. Hovedfokuset vil ligge på vansker med funksjonell

kommunikasjon og hvorfor disse barna ofte har behov for en individuell kartlegging av blant annet språkforståelse.

I kapittel 3 vil fokus flyttes fra hva og hvem som kartlegges til testteoretiske og konkrete problemstillinger tilknyttet hvordan å kartlegge språkforståelse. Kapittelet vil begynne med en redegjørelse for testteoretiske problemstillinger tilknyttet begrepet validitet og hvordan en valideringsprosess kan styrke tillitten vi kan ha til slutninger som kan trekkes på bakgrunn av testskårer. Jeg har her valgt å ta utgangspunkt i validitetsstandardene utformet av American Educational Research Association (AERA), American Psychological Association (APA) og National Council on Measurement in Education (NCME) (2014) der validitet anses som et resultat av en sammenstilling av ulike teoretiske og empiriske evidensformer. Forventningene beskrevet over vil her utdypes gjennom tre former for validitetsevidens; evidens basert på relasjon til eksterne variabler, evidens basert på indre struktur, og evidens basert på konsekvenser og bruk av testskårer. Evidens basert på testens innhold er en del av operasjonaliseringen og testutviklingen og vil ikke diskuteres direkte da den norske versjonen av C-BiLLT kun er oversatt og ikke tillagt noen nye testledd. Denne evidensformen vil likevel presenteres da begreper og antagelser tilknyttet testens innhold er relevant å trekke inn i diskusjonen av andre former for evidens. I kapittelets siste del vil konkrete utfordringer ved kartlegging av språkforståelse hos barn med store tale- og bevegelsesvansker belyses. Kapittelet vil avsluttes med en kort redegjørelse for språkforståelsestesten C-BiLLT.

Kapittel 4 vil ta for seg denne studiens metode ved å beskrive dens design, deltakere, kartleggingsinstrumenter, digitale verktøy, prosedyre og valg av statistiske analyser. Resultatene og diskusjonen fra reliabilitetsundersøkelsen vil behandles i Haddelands (2018) masteroppgave. En kort redegjørelse for reliabilitetsstudien og dens hovedfunn vil imidlertid presenteres også i denne oppgaven, da en vurdering av reliabilitet er en forutsetning for en diskusjon av validitet.

Resultatene fra denne studiens hovedanalyser vil presenteres i kapittel 5. Kapittel 6 utgjør oppgavens diskusjonsdel. Her vil jeg svare på problemstillingen ved å først diskutere de ovennevnte forventningene hver for seg opp mot studiens empiriske resultater og teori gjennomgått i Kapittel 2 og 3, før jeg sammenstiller de ulike evidensformene. Til slutt i dette kapittelet vil styrker og svakheter ved studien samt kliniske implikasjoner av funnene beskrives.

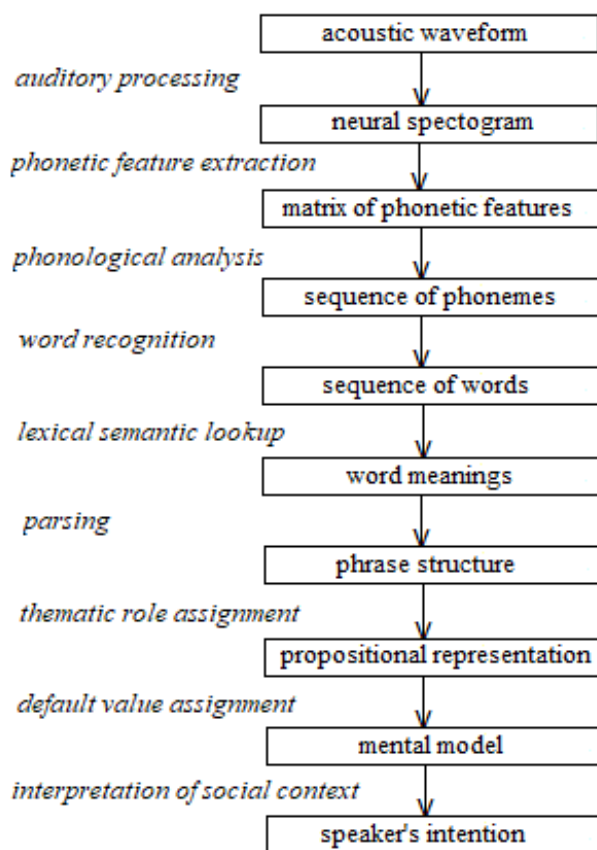
Oppgaven avsluttes med Kapittel 7. Her vil jeg gi en oppsummering med fokus på hovedfunnene i denne studien. Jeg vil avslutte med noen tanker rundt veien videre.

I tråd med APA 6th vil alle tall med desimaler i denne oppgaven rapporteres med punktum istedenfor komma, og verdier som per definisjon ikke kan overskride 1.0, som korrelasjoner, p -verdier og faktorladninger, vil rapporteres uten null foran punktum. Begrepet ”totalt antall forespurte” vil i denne oppgaven betegne alle barna som ble sendt forespørsel om å delta i studien, mens begrepene ”utvalg” og ”deltakere” vil brukes synonymt for å betegne barna som faktisk deltok i denne studien.

2 Språkforståelse

2.1 Komponenter i språkforståelse

Forståelse av språk er noe de fleste barn utvikler med letthet og blir etter hvert en så naturlig og integrert del av deres hverdag at det er lett å glemme at det egentlig er snakk om et svært komplekst fenomen. Å få innsikt i språkforståelse som prosess er sentralt for å forstå hva som testes med en språkforståelsestest som C-BiLLT som er utformet for å teste forståelse av ord og setninger. Bishop (2014) beskriver språkforståelse som en sammensatt prosess der informasjon forandres fra en type representasjon til en annen. Hun har videre samlet de ulike representasjonsnivåene og involverte prosesser i en modell for språkforståelse (Figur 1) som gir et nyttig rammeverk for en redegjørelse av språkforståelse som prosess.



Figur 1. Modell over nivåer og relaterte prosesser i språkforståelse. Etter Bishop, 2014, s. 4
Note. Nivåene er avbildet til høyre og prosessene er avbildet til venstre i modellen

Når noen sier et ord eller en setning oppstår lydbølger. Dette akustiske signalet inneholder mye fonetisk informasjon, men ikke alle fonetiske variasjoner er relevante for å forstå det som sies fordi de ikke alle er betydningsskillende i språket. De fire første nivåene i Bishops

(2014) modell beskriver overgangen fra en slik artikulatorisk fonetisk representasjon av talen i form av lydbølger, til ørets auditive analyse og omgjørelse av denne informasjonen til en indre nevrologisk representasjon, og videre til hjernens tolkning og omgjørelse av denne nevrologiske representasjonen til en abstrakt representasjon av talen som sekvenser av fonemer. Fonemer er betegnelsen for en abstrakt lingvistisk forenkling av fonetiske variasjoner som kan være artikulatorisk ulike, men funksjonelt like i hvordan de signaliserer mening i språket. Fonemer utgjør således språkets minste betydningsskillende enhet (Bjerkan, 2011). Lust (2006) fremhever denne omgjøringsprosessen som grunnleggende for språkforståelse, da den endelige fonologiske representasjonen kun inneholder de variasjoner i det akustiske signalet som er relevant for å forstå det lingvistiske budskapet, mens lingvistisk irrelevant informasjon, som for eksempel skyldes forskjell mellom talere eller bakgrunnsstøy i rommet, siles bort.

De to neste nivåene i Bishops (2014) modell beskriver en overgang fra en fonologisk til en semantisk representasjon av informasjon som gjør at ords betydningsinnhold kan forstås. Mentalt leksikon eller vokabular betegner barns mentale langtidslager både av ords fonologiske uttrykk, som ordlyden /hund/, og ordets semantiske betydningsinnhold, som at en hund er et dyr med fire bein som bjeffer. Prosessen å forstå talte ord beskrives her som avhengig av at barn kan gjenkjenne og segmentere ut ord fra talestrømmen som så gjør kontakt med og aktiverer tilsvarende fonologiske representasjoner og deres semantiske referent i mentalt leksikon (Jusczyk & Luce, 2002; Bishop, 2014). Hvorvidt ord fra talestrømmen gjenkjennes og forstås er dermed avhengig både av bredden på barns vokabular, antallet lagrede ordformer, og dybden i barns vokabular, i form av en dypere forståelse av ord og uttrykks semantiske betydningsinnhold (Lahey, 1988; Lust; 2006).

For å forstå en talt setning er det ikke nok å basere seg på forståelsen av ordene i setningen hver for seg, da en setnings betydning også avhenger av hvordan ord er kombinert og bygget opp. Det syvende og åttende nivået i Bishops (2014) modell beskriver hvordan en mental representasjon av setningers betydningsinnhold også er avhengig av en grammatisk analyse av setningen. Lust (2006) fremhever i denne sammenhengen at forståelse av setninger forutsetter at barnet har implisitt kunnskap om det språklige skjelettet som ligger under hver setning, også kalt syntaks. Dette innebærer en forståelse av at setninger er oppbygd av fraser eller setningsledd, en samling ord som har ulik funksjon i setningen som subjekt, verbal og objekt, og at ulike rekkefølger av setningsledd gir ulik betydning. Å vite at setningsledd får

ulike tematiske roller eller betydninger i setningen avhengig av deres plassering rundt verbalet er for eksempel avgjørende for å forstå hvem som gjør hva mot hvem i setninger av typen ”Katten jager hunden”. Innenfor leddsetninger vil også ords oppbygning brukes for å signalisere mening, betegnet i lingvistikken som morfologi (Simonsen & Theil, 2011). Kunnskap om språkets morfologi er dermed viktig for å forstå hvordan endringer i ord, som bøyning, avledning og sammensetning av ord, påvirker betydningen ordet har i setningen, for eksempel ved at ordet ”hund” har en annen betydning enn ”hunder” og ”gitar” har en annen betydning enn ”gitarist”.

Miller og Paul (1995) beskriver hvordan formelle tester av barns språkforståelse ofte er utformet for å måle de fonologiske, semantiske og grammatiske komponentene beskrevet hittil. Språkforståelse blir da ensbetydende med lingvistisk forståelse eller lingvistisk avkodning. Selv om fonologiske, semantiske og grammatiske prosesser er avgjørende for barns språkforståelse, er det viktig å presisere hva som ikke testes av slike tester. De siste nivåene av Bishops (2014) modell beskriver hvordan språkforståelse i hverdagen også avhenger av å integrere den betydningen som ligger i språket med miljømessig kontekst og generell kunnskap for å trekke slutninger om hva som ikke direkte blir sagt og danne en mental modell eller forestilling av denne meningen. I hverdagen er det ofte ikke et en-til-en forhold mellom den lingvistiske formen på og meningen i det som sies fordi språk i hverdagen brukes som kommunikasjon (Kristofferesen, 2011). Forståelse av sosial kontekst og andre menneskers intensjoner er dermed sentralt for å forstå den betydningen som ligger i et språklig uttrykk (Tomasello, 2003). I denne sammenhengen er det vist at selv om to år gamle barn er i stand til å forstå forskjellen i betydningsinnhold som skyldes ordrekkefølge i setninger av typen ”kakemonsteret kiler den store fuglen” og ”den store fuglen kiler kakemonsteret” (Golinkoff, Cauley, Hirsh-Pasek, & Gordon, 1987) vil barn under tre år i stor grad basere seg på enkeltord og hva som er den mest sannsynlige tolkningen av betydningsinnholdet i en setning basert på kontekst og erfaring. Strohner og Nelson (1974) demonstrerte for eksempel at i oppgaver der to til fem år gamle barn skulle bruke objekter til å spille ut både aktive og passive setninger med et usannsynlig meningsinnhold, som ”katten blir jaget av musen”, tolket barn under tre år slike setninger utfra det mest sannsynlige scenarioet, at musen jages av katten, mens fem år gamle barn i større grad kunne basere seg på setningsstrukturen og det språklige uttrykket i seg selv, noe som førte til riktig tolkning av setningen. Ettersom mange tester av språkforståelse forsøker å minimere rollen sosial kontekst og kulturell erfaring har for språkforståelse, fremhever Bishop (2014) at det ofte kan

eksistere et gap mellom barnets skårer på språkforståelsestester og hva foreldre og voksne rundt barnet opplever at barnet forstår i hverdagen.

2.2 Distinkte eller integrerte språkkomponenter

Bishop (2014) beskriver språkforståelse som målt med tester som en sammensatt prosess bestående av integrering av prosesser og informasjon fra fonologiske, semantiske og grammatiske nivåer. Når språktester konstrueres foreligger det ofte ifølge Tomblin og Zhang (2006) implisitte antagelser om at slike språklige komponenter representerer distinkte språklige evner da de fleste testbatterier består av deltester tenkt å reflektere ulike underdomener av språkevnen. Language and Reading Research Consortium (LARRC, 2015) fremhever imidlertid at det er uenighet rundt om barns språkevne best bør kartlegges som et sett av distinkte dimensjoner, og testtolkning og diagnostisering av språkvansker hos barn dermed bør gjøres på bakgrunn av deltester, eller om språkevnen kan kartlegges som en enhetlig og endimensjonal evne der sumskåren danner grunnlag for testtolkning. Det er dermed ulike syn på språkevnen dimensjonalitet eller struktur; hvorvidt den er sammensatt av separate underdomener og dermed flerdimensjonal, eller om den best kan beskrives som en helhetlig dimensjon eller evne (Slocum-Gori & Zumbo, 2011). Slike syn kan videre relateres til ulike teorier på hvordan barn utvikler språk. En redegjørelse for slike teoretiske standpunkt danner dermed et grunnlag for en diskusjon rundt hvordan en språktest bør konstrueres og tolkes.

2.2.1 Vokabular og grammatikk; styrt av distinkte mekanismer

Barn viser tidlig en imponerende språklige kompetanse og en bratt språklige utviklingskurve (Rowe, Raudenbush, & Goldin-Meadow, 2012; Simonsen, Kristoffersen, Bleses, Wehberg, & Jørgensen, 2014). Dette har av visse teoretikere blitt tolket som bevis for at barn har en medfødt og domenespesifikk språklig predisposisjon. Nativister er i denne sammenhengen betegnelsen på teoretikere som antar at hjernen gjennom evolusjonen har utviklet spesialiserte moduler tilpasset ulike funksjoner og styrt av ulike systemer og læringsmekanismer (Lust, 2006). Barns forståelse av ords semantiske betydningsinnhold, deres vokabularutvikling, antas for eksempel å styres av et medfødt assosiativt minnesystem som fremmer koblinger mellom ordlyder og deres semantiske referent (Pinker, 1998). Ifølge Markman (1990) vil et slikt system utstyre barnet med universelle regler og begrensninger for hvilke semantiske referenter en ordlyd sannsynligvis refererer til, noe som reduserer

utfordringen med å finne ut hva en ordlyd referer til i virkeligheten. For eksempel vil regelen kalt *whole object assumption* bidra til at barn antar at en ordlyd sannsynligvis referer til et helt objekt snarere enn deler av det. Barns grammatiske utvikling og forståelse av setninger anses i nativismen videre å styres av en annen medfødt og regelbasert mekanisme. Denne mekanismen inneholder abstrakte lingvistiske kategorier som letter oppdelingen av talestrømmen i grammatiske kategorier som setningsledd, og dermed ekstraksjonen av grammatiske regler fra talestrømmen (Chomsky, 1959; Pinker, 1998).

Fra et nativistisk perspektiv vil dermed grammatikk og vokabular representere to distinkte og spesialiserte språklige komponenter hvis utvikling er genetisk determinert av ulike medfødte og spesialiserte læringsmekanismer. Barn kan da forventes å ha selektive svekkelser i ulike språklige komponenter som en konsekvens av en skade eller anomali i ett av disse modulære systemene. Flere teoretikere argumenterer for eksempel for eksistensen av en undergruppe barn med spesifikke språkvansker som har store selektive svekkelser i grammatikk, mens deres forståelse av enkeltord er mer preservert eller tilnærmet normal (van der Lely, Rosen, & McClelland, 1998; van der Lely, 2005). Utfra et slikt syn vil det være hensiktsmessig å skille vokabular og grammatikk i kartleggingen, ved å kartlegge forståelse av ord og forståelse av setninger hver for seg, snarere enn å kombinere mål på grammatikk og vokabular til en overordnet sumskåre ettersom en slik sumskåre blir vanskelig å tolke da den påvirkes på uvis måte av variasjoner i to fenomener og ikke ett (Slocum-Gori & Zumbo, 2011). At analysen av faktorstrukturen til språktester resulterer i flerdimensjonale modeller der flere overordnede faktorer må til for å forklare variasjoner i testledd (Petscher, Connor, & Otaiba, 2012) er videre å teoretisk forvente utfra synet på språkevnen som sammensatt av separate komponenter styrt av separate mekanismer.

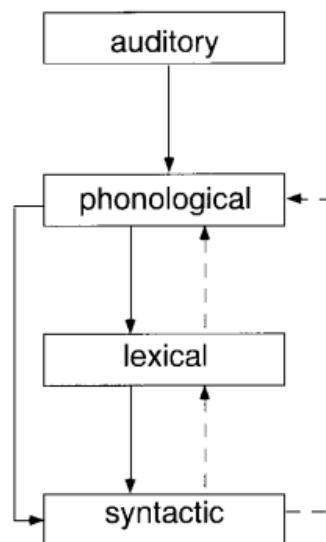
2.2.2 Språklig spesialisering; fremkommende snarere enn medfødt

Nativismen har blitt kritisert for å ha et genetisk deterministisk syn på språkutvikling, da språkets utviklingsbane i virkeligheten er påvirket både av genetiske og miljømessige faktorer (Hayiou-Thomas, Dale, & Plomin, 2012). Karmiloff-Smith (2009) legger i denne sammenhengen et nevrokonstruktivistisk perspektiv til grunn som beskriver barns utvikling som en kompleks prosess der gener, hjerne, kognisjon og miljø gjensidig interagerer. Hjernen beskrives ikke her som å inneha ulike domenespesifikke moduler fra fødselen av, men domenerrelevante forskjeller bestående av små strukturelle ulikheter mellom ulike

hjerneregioner som passer visse typer prosessering bedre enn andre. En språklig spesialisering er fra et slikt perspektiv ikke medfødt, men en fremkommende egenskap der mer generelle prosesseringssystemer blir mer finjusterte i møte med språklig input. Flere studier har i denne sammenhengen trukket frem evnen til generell mønsterpersepsjon og assosiativ læring som viktig både for utvikling av vokabular og grammatikk. Ved å simulere menneskelig læring gjennom å bruke konneksjonistiske nettverk har det for eksempel blitt vist at det er mulig å ekstrahere regler fra språklig input, selv uten medfødte abstrakte lingvistiske kategorier (Plunkett & Marchman, 1993). Det er videre demonstrert at åtte måneder gamle spedbarn er oppmerksomme på statiske regulariteter i språklig input, for eksempel hvilke stavelser som sannsynligvis opptrer sammen og danner ord, noe som letter segmenteringen av ord fra talestrømmen (Saffran, Aslin, & Newport, 1996), samt at mål på fire til seks år gamle barns evne til implisitt statistisk læring er positivt korrelert med læring av nye syntaktiske strukturer (Kidd, 2012). I denne sammenhengen har det blitt foreslått at barns generelle evne til statistisk læring, evnen til å ekstrahere statistiske regler fra verden rundt dem, spiller en rolle i utvikling både av vokabular og grammatikk (Romberg & Saffran, 2010, Arciuli & Torkildsen, 2012), noe som står i motsetning til antagelsen om at utviklingen av vokabular og grammatikk er avhengig av separate prosesser og læringsmekanismer.

Dersom en språklig spesialisering er fremkommende snarere enn medfødt, vil også spedbarns hjerne og dens funksjoner kunne beskrives som nært forbundet heller enn bestående av separat fungerende moduler fra fødselen av (Karmiloff-Smith, 2009). Flere studier viser i tråd med dette hvordan mål på yngre barns kognitive evner, som språklige og ikke-verbale resonneringsevner, er høyere positivt korrelert og dermed mindre differensiert enn hos eldre barn over fire år (Colledge et al., 2002; Tideman & Gustafsson, 2011). Bishop (2014) kritiserer i denne sammenheng nativismen for å utelate andre kognitive evner i beskrivelsen av språkforståelse og at den som teori derfor blir for snever til å forklare hvordan grammatisk kunnskap brukes i forståelsen av innkommende tale. Bishop (2014) fremhever at avkoding av grammatiske strukturer ikke skjer når setningen er ferdigsagt, men underveis mens setningen høres. Forståelse av setninger er ikke da kun avhengig av abstrakt grammatisk kunnskap, men også av evnen til å holde innkommende tale i minnet mens meningen i det som er hørt avkodes. Arbeidsminne, delen av minnet som lagrer informasjonen i en kort tidsperiode slik at den kan bearbeides, er i denne sammenhengen fremhevet som betydningsfull både for forståelse av syntaktiske strukturer (Norbury, Bishop, & Briscoe, 2002) og tidlig vokabularlæring (Gathercole, Willis, Emslie & Baddeley, 1992).

Bishop (1997; 2014) fremhever videre nativismen som lite egnet i forklaringen av den dynamiske utviklingen av språkforståelse hos barn. For selv om pilene i Bishops (2014) språkforståelsesmodell (Figur 1) går én vei, fremhever hun at det i virkeligheten vil være interaksjoner mellom alle nivåene og prosessene i modellen (som illustrert i Figur 2). Begrepet bootstrapping brukes i denne sammenhengen for å beskrive hvordan læring i en språklig dimensjon muliggjør læring i andre dimensjoner og hvordan barns vokabularforståelse og grammatiske forståelse gjensidig interagerer i utviklingen av språkforståelse. Teorien om leksikalsk bootstrapping beskriver for eksempel hvordan barns vokabular danner grunnlaget for grammatisk utvikling, blant annet av bøyningsmorfologi, ved at grammatiske mønstre og regulariteter vil ekstraheres når vokabularet når en kritisk masse (Marchman & Bates, 1994; Moyle, Weismer, Evans, & Lindstrom, 2006). Gleitman (1994) kritiserer videre antagelsen om at barn lærer ord kun ved å gjenkjenne kjeder av fonemer i talen og dedusere mening fra miljømessig kontekst, ettersom dette ikke egner seg som forklaring på hvordan barn lærer betydningen av verb eller abstrakte substantiver. Teorien om syntaktisk bootstrapping beskriver i denne sammenhengen hvordan barn kan utnytte grammatisk kunnskap om ords oppbygning og plassering i setningen for å utlede betydningen til ord de ikke før har hørt (Gleitman, 1994; Abend, Kwiatowski, Smith, Goldwater, & Steedman, 2017).



Figur 2. En forenklet modell av de tidlige nivåene i språkforståelse som viser interaksjoner mellom alle nivåer. Hentet fra Bishop, 1997, s. 906

Dersom språklige komponenter i utviklingen av språkforståelse ikke er distinkte og modulære, men nært sammenhengende og interagerende, vil vansker på ett nivå i

språkforståelsesmodellen sannsynligvis føre til vansker i de andre nivåene. Bishop (2014; 1997) fremhever i denne sammenheng at det ved utviklingsmessige vansker oftere er komplekse mønstre av assosierte vansker som må forklares snarere enn selektive svekkelser i en språklig komponent som grammatikk. Dette kan videre belyses med LARRC, Jiang, Logan og Jias (2018) longitudinelle studie. De undersøkte utviklingsbanene (Eng. developmental trajectories) til grammatikk og vokabular ved å kartlegge 420 barn mellom 44 og 73 måneder med standardiserte tester av både reseptiv og ekspressiv grammatikk og vokabular opptil fem ganger over en femårsperiode. Ved sammenligning av de to utviklingsbanene fremstod de som nært sammenhengende; begge viste en initial akselererende økning etterfulgt av en viss redusert vekstøkning fra barnehagealder til syvårsalder. De lå dermed på nær linje med hverandre og var høyst korrelerte ferdigheter hos barna i dette aldersspennet. Dette ble tolket som å representere endimensjonaliteten i språkferdigheter hos yngre barn, og brukt som et argument for at barn som har en svekket ordforståelse sannsynligvis også vil ha svekket grammatisk forståelse, samt at barn som har et avansert reseptivt vokabular sannsynligvis også vil vise avansert reseptiv grammatikk, noe som sammenfaller med hva Bishop (1997; 2014) fremhever.

Dersom en språklig spesialisering er fremkommende snarere enn medfødt, og forståelse av enkeltord og setninger kan anses som avhengig av nært sammenhengende ferdigheter, vil videre oppfordringen til å ikke blande mål på vokabular og grammatikk til en sumskåre i kartleggingen av barn være mindre begrunnet. I tråd med dette har flere kryss-seksjonelle studier av faktorstrukturen til språktester vist hvordan variasjoner i barns språkevne, målt med standardiserte tester av både reseptivt og ekspressivt vokabular og grammatikk, kan forklares av én overordnet faktor eller én generell språkevne (Colledge et al., 2002; Hayiou-Thomas et al., 2006; Anthony, Davis, Williams, & Anthony, 2014; Klem, Gustafsson, & Hagtvat, 2014). Ekstraksjonen av en endimensjonal modell som representerer en enhetlig språkevne kan videre tolkes som en indikasjon på at individuelle forskjeller i komponenter som grammatikk og vokabular avhenger av lignende og ikke separate mekanismer, og at lignende prosesser, både miljømessige og genetiske, ligger under individuelle forskjeller i flere språkdomener hos yngre barn (Hayiou-Thomas et al., 2006). LARRC og kollegaer (2018) beskriver videre hvordan utviklingsbanene for grammatikk og vokabular i deres longitudinelle studie i økende grad divergerte etter syvårsalder ved at barnas vokabularferdigheter fortsatte å vise en stabil økning, mens økningen i barns grammatiske evne viste en nedflatning. Denne divergeringen ble fortsatt tolket som at vokabular- og

grammatikkferdigheter er nært sammenhengende, men at mens ordlæringen fortsetter hele livet etter hvert som nye ord møtes vil grammatisk forståelse, som i større grad er avhengig av tilegnelsen av et finitt sett av grammatiske regler, nå ett platå. Dette sammenfaller med andre studier som viser at korrelasjonen mellom barns vokabular- og grammatikkferdigheter gradvis minker med økende alder og dermed blir gradvis mer differensierte ferdigheter (Tomblin & Zhang, 2006; Lyster, Horn, & Rygvold, 2010; LARRC, 2015) og hvordan en flerdimensjonal faktorløsning der vokabular og grammatikk anses som to distinkte men korrelerte ferdigheter gir en bedre forklaring på variasjoner i språkferdigheter hos de eldre skolebarna (Tomblin & Zhang, 2006; LARRC, 2015). En slik utvikling vil videre være i tråd med antagelsen om at en språklig spesialisering og differensiering er fremkommende snarere enn medfødt.

2.3 Barn med komplekse kommunikasjonsbehov

Språkforståelse kan anses som en kompleks prosess avhengig av en interaksjon mellom flere språklige komponenter som fonologi, semantikk og grammatikk. Disse komponentene kan videre anses som nært sammenhengende i utviklingen av språkforståelse både fordi utviklingen i en språklig komponent kan anses å avhenge av utviklingen i andre komponenter (Bishop, 2014; LARRC et al., 2018), og fordi variasjoner i både grammatikk og vokabular hos yngre barn kan forklares av en endimensjonal modell som representerer en overordnet språkevne (Colledge et al., 2002; Hayiou-Thomas et al., 2006; Tomblin & Zhang, 2006; Anthony et al., 2014; Klem et al., 2014; LARRC, 2015). At språklige komponenter og språklige og kognitive evner er mindre differensiert hos yngre barn kan videre tolkes som at en spesialisering av evner og ferdigheter er et fremkommende resultat fra en dynamisk prosess der gener, hjerne, kognisjon og miljø samspiller (Karmiloff-Smith, 2009; Hayiou-Thomas et al., 2006). Etersom utviklingsmessige vansker oppstår mens barnet er i utvikling, snarere enn i et ferdig utviklet og spesialisert system, er videre assosierte og komplekse mønstre av vansker å forvente (Karmiloff-Smith, 2009; Bishop, 2014). Barn med store tale- og bevegelsesvansker er en gruppe barn som ofte viser et komplekst mønster av assosierte vansker, og C-BiLLT er utformet særlig for å kartlegge språkforståelse hos denne gruppen barn. Cerebral Parese (CP) er den vanligste årsaken til tale- og bevegelsesvansker i barndom (Cans, De-la-Cruz, & Memet, 2008) og derfor vil hovedfokus i redegjørelsen under rettes mot denne gruppen og hvorfor disse barna kan ha et behov for kartlegging av språkforståelse.

2.3.1 Barn med CP

CP kan defineres som en samlebetegnelse for en gruppe permanente vansker med kroppsholdning og bevegelse som skyldes en ikke-progressiv skade eller anomali i foster- eller spedbarnshjernen (Rosenbaum et al., 2007). Ifølge populasjonsbaserte studier er prevalensen av barn med CP omtrent 2 per 1000 levende fødte barn (Surveillance of Cerebral Palsy in Europe (SCPE), 2002; Andersen et al., 2008; Himmelmann, Hagberg, & Uvebrant, 2010). Det kan ofte være vanskelig å fastslå den nøyaktige årsaken til CP, og noen ganger ser det videre ikke ut til å være en enkeltårsak, men flere risikofaktorer som kumulativt fører til CP. Dette har ført til diskusjoner rundt kausale veier og sammensatte årsaksmekanismer, fremfor en enkeltårsak til CP (Himmelmann et al., 2010). Vanskens heterogenitet bør også ses i sammenheng med at hjerneskaden eller –anomalien forekommer i en hjerne i utvikling. Mønsteret av symptomer kan dermed endres med alder og få nye konsekvenser på grunn av plastisitet og modning i sentralnervesystemet. Det er dermed sjeldent et en-til-en-forhold mellom hjerneskade og kognitiv fungering (Sankar & Mundkur, 2005; Karmiloff-Smith, 2009).

Ettersom CP er en såpass heterogen vanske er det utviklet flere klassifiseringssystemer for å beskrive type og alvorlighetsgrad av CP. Det er vanlig å skille mellom fire hovedtyper CP utfra motoriske symptomer (spastisk, dyskinetisk og ataktisk) og etter vanskens omfang; om én side av kroppen er rammet (unilateralt) eller om begge sider av kroppen er rammet (bilateralt) (SCPE, 2000). Barnets grov- og finmotoriske funksjonsnivå, og hvordan dette funksjonsnivået påvirker aktivitet og deltakelse, klassifiseres utfra henholdsvis Gross Motor Function Classification System (GMFCS) (Palisano et al., 1997) og Manual Ability Classification System (MACS) (Eliasson et al., 2006). I begge systemene gjøres klassifisering på bakgrunn av fem nivåer som gir beskrivelser av barnets evne til å sitte og gå (grovmotorikk), eller hvordan barnet bruker hendene til å håndtere gjenstander i dagliglivet (finmotorikk), samt behov for hjelpemidler. Rundt 70% av barna klassifiseres som å ha mild til moderat grov- og finmotorisk fungering (nivå I-III), mens rundt 30% klassifiseres i de mer alvorlige nivåene (nivå IV-V) (Himmelmann, Beckung, Hagberg, & Uvebrant, 2006).

2.3.2 Barn med CP og kommunikasjon

Mens tidlige definisjoner utelukkende beskrev CP som en motorisk vanske, fremhever nyere definisjoner at flertallet av personene med CP i tillegg til motoriske vansker også har

tilleggsvansker som forstyrrelser i sansing, persepsjon, kognisjon, kommunikasjon og atferd (Rosenbaum et al., 2007). Det fremheves videre at slike vansker for noen kan være mer begrensende for deltakelse og aktivitet enn den motoriske vansken (Himmelman et al., 2010). Funksjonell kommunikasjon klassifiseres utfra Communication Function Classification System (CFCS) som er tilsvarende og kompletterende til GMFCS og MACS (Hidecker et al., 2011). Funksjonell kommunikasjon defineres her som å oppstå når samtalepartnere etablerer delt forståelse, og denne delte forståelsen beskrives som avhengig av en interaksjon mellom aktivitet, deltakelse, og kroppslige strukturer og funksjoner (som kognisjon, språk, tale og hørsel) og kontekstuelle faktorer (som samtalepartners samtaleferdigheter, type ASK, setting og kulturell bakgrunn). Flere av disse faktorene kan vanskeliggjøre kommunikasjon.

For det første er vansker med funksjonell kommunikasjon mer frekvent hos barna klassifisert på de høyere GMFCS-nivåene IV og V (85-100%) (Voorman, Dallmeijer, Eck, Schuengel, & Becher, 2010; Himmelman, Lindh, & Hidecker, 2012). Dette kan sees i sammenheng med at de motoriske vanskene hos barn med CP også påvirker taleapparatet og at barna i de høye GMFCS-nivåene også har størst frekvens av talevansker (80%) (Andersen, Mjøen, & Vik, 2010). Den vanligste talevansken hos barn med CP er dermed motoriske talevansker (dysartri) som hemmer normal uttale av språklyder slik at barnets tale blir vanskelig å forstå for samtalepartner. Hos noen barn er den motoriske talevansken så alvorlig at barnet ikke har tale (anartri) (Odding et al., 2006).

Hva gjelder reseptive språkvansker hos barn med CP er det vanskelig å si noe spesifikt om i hvor stor grad dette påvirker funksjonell kommunikasjon, da rapportert prevalens av språkvansker hos barn med CP varierer fra 28-74% (Voorman et al., 2010; Parkes, Hill, Platt & Donnelly, 2010; Pirila et al., 2007). Mei og kollegaer (2016) knytter slike varierende tall både til mangel på studier av forekomsten av språkvansker hos barn med CP, i tillegg til at eksisterende studier er vanskelig å sammenligne grunnet metodologi, fordi det anvendes ulike og ikke-standardiserte kartleggingsinstrumenter, og heterogenitet i utvalget barn med CP hva gjelder motorisk type CP, alvorlighetsgrad av motoriske vansker og alder. Noen studier beskriver imidlertid språkvanskene hos barn med CP som å gå på tvers av språkdomener og innebære vansker med å integrere informasjon fra ulike språklige domener som semantikk, syntaks og morfologi for å forstå mer komplekse setninger (Mei et al., 2016; Geytenbeek et al., 2015). Slike vansker vil i så fall være i tråd med Bishops (2014) definisjon av

språkforståelse som avhengig av informasjonsintegrering fra flere språknivåer, samt ikke-modulære utviklingsteorier om at vansker på ett språklig nivå vil vise seg i vansker på andre nivåer.

Barn med CP som har anartri blir ofte omtalt som ikke-verbale, noe som kan føre til antagelsen om at disse barna også har kognitive og reseptive språkvansker som påvirker deres funksjonelle kommunikasjon (Nordberg, Miniscalco, Lohmander, & Himmelmann, 2013). Selv om utvikling av språkforståelse og ekspressivt språk vanligvis forløper samtidig, er det viktig å presisere at det kan eksistere en dissosiasjon i utvikling av kommunikative evner hos barn med CP fordi reseptiv og ekspressiv språkutvikling kan følge to ulike utviklingsbaner. For mens den ekspressive språkutviklingen hos barn med CP er relatert til motoriske vansker kan barns reseptive språkutvikling i større grad relateres til barnets kognitive fungering, taleprosessering og verbal intelligens (Vos et al., 2014; Mei et al., 2016, Choi, Choi, & Park, 2017). I tråd med dette har flere studier vist at reseptivt vokabular og grammatikk utvikles selv når barnet ikke snakker og at flere ikke-talende barn med CP har aldersadekvat reseptivt språk (Bishop et al., 1990; Frampton, Yude, & Goodman, 1998; Straub & Obrzut, 2009; Geytenbeek et al., 2015). Det kan dermed eksistere en diskrepans mellom hva barnet med CP kan uttrykke av språk og hva de forstår av språk.

Dersom barnet har vansker med å uttrykke språk kan dette få store konsekvenser for barnets språkutvikling, da denne er avhengig av interaksjonen mellom biologiske og miljømessige faktorer. Tomasello (2003) fremhever for eksempel at bruk av språk i sosiale interaksjoner er avgjørende for språkutviklingen. Språk og kommunikasjonsferdigheter utvikles dermed ikke i et vakuum, men gjennom deltakelse i miljøet. Den språklige interaksjonen mellom foreldre og barn er for eksempel i flere longitudinelle studier blitt markert som en viktig faktor i språkutviklingen, ved at foreldreforskjeller i antall ord og ordtyper som brukes er relatert til forskjeller i barns vokabularstørrelse (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Hoff, 2003) og frekvensen av grammatiske strukturer i foreldres språklige interaksjoner med sine barn er positivt relatert til utvikling av slike strukturer i barns språk (Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010). Dersom barnet med store tale- og bevegelsesvansker utvikler en språkvanske eller en vanske med funksjonell kommunikasjon må dette også sees i sammenheng med at barnet kan ha hatt manglende språklige og kommunikative erfaringer fra deltakelse i ulike samhandlinger med voksne og jevnaldrende grunnet manglende miljømessig tilrettelegging for slike erfaringer (Pennington &

McConachie, 2001; Rosenbaum et al., 2007). Språkforståelsesvansker eller vansker med funksjonell kommunikasjon hos barn med CP bør dermed beskrives både som en følge av etiologien bak CP, og som en sekundær konsekvens av aktivitets- og deltakelsesbegrensninger som begrenser læring og erfaring. Dette synliggjør behovet barn med store tale- og bevegelsesvansker har for ASK, andre språklige uttrykksformer enn tale, og hvordan slike tiltak er fundamentalt for etableringen av sosiale relasjoner, funksjonell kommunikasjon og deltakelse (Light et al., 2002; Batorowicz et al., 2016) samt språklig og kognitiv læring og utvikling (Tetzchner & Martinsen, 2002; Murray & Goldbart, 2009).

Ettersom barn med CP er en heterogen gruppe med variert språklig og kommunikativt funksjonsnivå, vil de også vil ha ulike behov og forutsetninger for ASK. Barnets motoriske funksjonsnivå vil for eksempel påvirke om barnet bør benytte hjulpet kommunikasjon, der barnet selv produserer språklige tegn for eksempel med hendene, eller ikke-hjulpet kommunikasjon, der språklige uttrykk foreligger i fysisk form utenfor brukeren og velges av brukeren. Videre vil barnets språkforståelsesnivå påvirke valg av eventuelle tegnsystemer da visse systemer, som Bliss-systemet, stiller høyere krav til språkforståelse enn andre, som Piktogrammer (Tetzchner & Martinsen, 2002). Det vil også være stor variasjon i hvilket kompleksitetsnivå kommunikasjonshjelpemiddelet bør ha for å være best tilpasse barnets behov. Et slikt behov kan variere fra at barnet skal velge fra noen få objekter på en gang til at barnet bør kunne ha muligheten til å uttrykke komplekse setninger tilsvarende de produsert av jevnaldrende (Stadskleiv, 2015).

For at ASK-løsningen skal bidra positivt til funksjonell kommunikasjon og barnets språklige og kommunikative utvikling er det dermed sentralt at valg bygger på en detaljert kartlegging som omfatter barnets interesser og motivasjon, miljøforutsetninger samt barnets styrker og svakheter innenfor områder som somatikk, motorikk, språk og kognisjon. En slik kartlegging bør videre igangsettes tidlig, slik at ASK-tiltak kan iverksettes så tidlig som mulig, og kontinuerlig, slik at tiltak kan justeres og videreutvikles i tråd med barnets språklige og kommunikative utvikling (Siegel & Cress, 2002; Tetzchner & Martinsen, 2002; Stadskleiv, Schjørbeck, Seeland, & Lillehaug, 2014). Kunnskap om barnets språkforståelsesnivå gir videre føringer for hvilke andre tiltak som vil være nyttige for barnet samt hvordan voksne og jevnaldrende bør interagere med barnet for å tilpasse seg barnets språklige nivå (Pennington & McConachie, 2001; Geytenbeek, Heim, Vermeulen, & Oostrom, 2010a). En viktig del av en slik detaljert utredning er dermed kartlegging gjennom normerte tester av språkforståelse.

3 Kartlegging av språkforståelse

3.1 Testteoretiske utfordringer

Kartlegging kan defineres som en sammensatt prosess som innebærer å integrere data fra ulike kartleggingsverktøy og kartleggingsmetoder for å danne seg et helhetlig bilde av barnet og dets kontekst, og dette gir et grunnlag for å utforme tilpassede tiltak og intervensjoner (Sattler, 2008). En grunnstein i en slik kartlegging er datainnsamling gjennom administrering og skåring av normerte tester. Slike tester har en standardisert administrering og skåring som sørger for at testen presenteres og skåres likt for alle barn. Det oppgis videre normer for typiske testskårer utfra hva som er representativt for barn i samme alder. Slike tester gir således et sammenligningsgrunnlag slik at barnets skåre kan tolkes utfra hvilken rangering den har i normgruppen. Barnets skåre på en slik test kan dermed brukes til å si noe om barnets funksjonsnivå på den testede evnen relativt til andre barn (Price, 2017).

Sattler (2008) fremhever imidlertid at man ikke uten videre kan akseptere barnets testskåre som et sant bilde på barnets evner, men at resultatet av slike tester bør tolkes varsomt for å vurdere hva det kan antyde om barnets faktiske evner. Dette er relatert til en sentral testteoretisk problemstilling; hvordan kan vi være sikre på at testen måler det den er uformet for å måle? Grunnen til denne problemstillingen er at mange av de fenomener som studeres i spesialpedagogikken ikke lar seg direkte observere (Nunnally & Bernstein, 1994). Bishops (2014) beskrivelse av språkforståelse som en sammensatt prosess avhengig av interaksjoner mellom fonologi, semantikk og grammatikk er en beskrivelse som befinner seg på det teoretiske plan. I testteori kalles slike beskrivelser for konstrukter, og konstrukter vil per definisjon ikke være mulig å måle direkte med en test. Måling defineres i tråd med dette som regler for å tildele symboler til konstrukter for å representere deres attributter numerisk (Nunnally & Bernstein, 1994). Konstruktet språkforståelse måles dermed ikke direkte med en språkforståelsestest. Det som representeres numerisk med en testsumskåre er konstruktets antatte attributter, eller egenskaper og uttrykk, som at barnet etter å ha hørt en talt setning kan peke på det riktige bildet som tilsvarer setningens betydningsinnhold.

Flere faktorer påvirker gyldigheten til slutningene som tas om et konstrukt på bakgrunn av målingen av dets antatte attributter. En grunnsetning i klassisk testteori er at en observert skåre = sann skåre + målefeil (Friborg, 2010). En sumskåre oppnådd på en

språkforståelsestest (observert skåre) er ifølge en slik grunnsetning satt sammen av personens faktiske språkforståelsesevne (sanne skåre) og ulike feil ved målingen (målefeil). Målefeil kan være systematiske eller usystematiske. Systematiske målefeil er noe som konstant legges til eller trekkes fra den sanne skåren. Slike målefeil gir samme avvik hver gang noe måles og trekker dermed den observerte skåren i samme retning unna den sanne skåren. Dette svekker gyldigheten i de slutningene som kan trekkes på bakgrunn av den observerte skåren.

Tilfeldige målefeil varierer på en usystematisk måte og trekker dermed den observerte skåren i ulike retninger fra den sanne skåren, noe som svekker påliteligheten til en måling. På samme måte som et konstrukt ikke lar seg direkte observere er den sanne skåre aldri kjent. I vektleggingen av at den observerte skåre inneholder målefeil er det dermed heftet en usikkerhet ved kunnskapen vi kan ha om fenomener i virkeligheten (Kleven, 2008). Sattlers (2008) presisering om at barnets testskåre ikke uten videre kan aksepteres som et sant bilde på barnets evner blir i lys av dette en begrunnet reservasjon i tolkning av testresultater.

Tester er på tross av dette et mye brukt verktøy i kartlegging da de likevel kan gi et riktigere bilde av barnets funksjonsnivå enn ved gjetting og antagelser. Ettersom resultatene på tester kan ha direkte innvirkning på og få konsekvenser for menneskers liv da tiltak uformes på bakgrunn av slike resultater, er det viktig å forsøke å minimere usikkerheten knyttet til resultatene av en måling slik at det kan heftes større sikkerhet til slutningene som trekkes på bakgrunn av testen (Messick, 1995). En måte å øke tilliten vi kan ha til slutninger som trekkes på bakgrunn av en test, er at testen går gjennom en valideringsprosess.

3.1.1 Validitet og validering

Validitet er et mangefasettert begrep tradisjonelt beskrevet som å omhandle i hvilken grad en test måler det den er utformet for å måle (Nunnally & Bernstein, 1994). I beskrivelsen av dette mangefasetterte begrepet har det historisk vært vanlig å skille mellom innholdsvaliditet, kriterievaliditet og begrepsvaliditet. Flere teoretikere fremhever imidlertid at en slik oppdelingen av validitet kan føre til en fragmentert tilnærming til det som egentlig burde være en koherent og overordnet validitetsvurdering (Nunnally & Bernstein, 1994; Benson, 1998; Price, 2017). I denne sammenhengen har AERA, APA og NCME (2014) sammenfattet en validitetsstandard for å fremme en felles tolkning og anvendelse av det sammensatte validitetsbegrepet. Validitet defineres i denne standarden som "the degree to which evidence and theory supports the interpretations of test scores for proposed uses of tests" (AERA et al.,

2014, s. 11). Validitet beskrives her som en prosess i testutvikling bestående av å samle teoretisk og empirisk evidens som til sammen støtter den intenderte tolkningen og bruken av testskårer. Validitet blir da ikke en egenskap ved selve testinstrumentet, men en kvalitet ved de slutningene som kan trekkes på bakgrunn av skåren på testinstrumentet. Det fremheves videre at de ulike evidenskildene bidrar til å belyse ulike sider ved validitet, men at de ikke representerer distinkte typer validitet. Validitet blir da ikke en enten-eller-egenskap uttrykt med ett statistisk mål eller én teoretisk ekspertbedømmelse, men et gradsspørsmål basert på integrering av flere evidenskilder.

En kilde til validitetsevidens kommer fra vurdering av relasjonen mellom innholdet i testen og konstruktet testen er utformet for å måle (AERA et al., 2014). En slik vurdering sammenfaller med det første steget i testutvikling som omhandler å spesifisere det teoretiske innholdsdomenet til konstruktet det skal trekkes slutninger om, og deretter definere hvilke attributter som er mest relevante å inkludere i testen, og hvordan disse attributtene skal måles, for å sikre at testen tapper innholdsdomenet på en presis og dekkende måte. Benson (1998) beskriver dette som å trekke forbindelseslinjer mellom det teoretiske og empiriske planet slik at det teoretiske fenomenet operasjonaliseres. Å trekke slike forbindelseslinjer, ved å velge hvilke attributter som skal inngå i en test og hvordan de skal måles, er utfordrende da det empiriske domenet består av alle mulige potensielle attributter samt alle mulige måter disse attributtene kan måles på. Messick (1995) fremhever i denne sammenheng at to kilder til systematisk målefeil truer validiteten i operasjonaliseringsprosessen. Konstrukt-
underrepresentasjon beskriver i hvilken grad en test systematisk mislykkes i å fange opp viktige aspekter ved konstruktet. Et eksempel kan være om testleddene som velges for å representere konstruktet språkforståelse er for få til å tappe alle domenets fasetter. Videre beskriver konstruktirrelevant varians i hvilken grad testskårer er påvirket av prosesser som er urelatert til testens intenderte bruk. Dersom for eksempel flere av testleddene som inkluderes i en språkforståelsestest egentlig tapper aspekter ved ikke-verbal resonnering vil disse testleddene være irrelevante for det egentlige innholdsdomenet og dermed føre til systematisk irrelevant variasjon i testskårer. Innholdsbasert evidens som støtter tolkningen og bruken av testskårer kommer fra eksperters bedømmelser og teoretiske overveielser om hvorvidt attributtene som velges for å utgjøre testen er representative for det teoretiske domenet slik det er definert (Nunnally & Bernstein, 1994; AERA et al., 2014). Målet er både å sikre et representativt utvalg testledd og at metoden for testing er fornuftig hva gjelder ordlegging, testleddformat og svarmåte (Nunnally & Bernstein, 1994).

Konstruktunderrepresentasjon og konstruktirrelevant varians bør ikke kun belyses gjennom teoretiske overveielser, men også empiriske undersøkelser. Prosessen å spesifisere konstruktets teoretiske innholdsdomene og tilhørende attributter medfører også å plassere konstruktet i det som kalles et nomologisk nettverk der forventede og lovmessige relasjoner med andre konstrukter og ytre variabler defineres (Cronbach & Meehl, 1955). En annen kilde til validitetsevidens er i denne sammenhengen evidens basert på testskårerers relasjoner med variabler som er eksterne til testen (AERA et al., 2014). Valg av slike eksterne variabler avhenger av den intenderte bruken av testskårer, men et vanlig valg er skårer på andre valide og reliable tester antatt å måle det samme, relaterte eller urelaterte konstrukt. Validitetsstøtte avhenger av hvorvidt relasjonene mellom testskårer på testen og tester på andre relaterte og urelaterte konstrukter er konsise med teoretiske forventninger om konstruktet som ligger under antatt testskåretolkning. I denne studien er for eksempel den første forventningen at barnets skåre på språkforståelsestesten C-BiLLT samvarierer med barnets skåre på andre validerte språkforståelsestester ettersom disse er utformet for å trekke slutninger om samme konstrukt, og at samvariasjonen mellom skåren på C-BiLLT og skårer på tester ment å måle ikke-verbale resonneringsevner er mindre sterk da testene er utformet for å trekke slutninger om ulike konstrukter. Dersom slike forventninger ikke innfris kan dette tyde på at konstruktunderrepresentasjon eller konstruktirrelevant varians påvirker testskåren til testen som undersøkes. Disse evidensformene omtales ofte som henholdsvis konvergerende og diskriminerende evidens og er fremhevet som særlig nyttige for å nyansere tolkningen og bruken av testskårer (AERA et al., 2014).

I tillegg til teoretiske forventninger om hvordan testen er relatert til ytre variabler, er det også teoretiske forventninger til hvordan testleddene i testen er relatert til hverandre, med andre ord testens indre struktur. AERA, APA og NCME (2014) fremhever i denne sammenhengen at en viktig kilde til validitetsevidens kommer fra en strukturanalyse, da denne kan indikere hvorvidt relasjonen mellom testledd er i tråd med konstruktet testskåretolkningen er basert på. En mye brukt metode i denne sammenhengen er faktoranalyse (Benson, 1998).

Faktoranalyse bygger på forventningen om at dersom testledd, eller attributter, antas å tappe samme teoretiske konstrukt vil de også påvirkes på samme forutsigbare måte av dette konstruktet og dermed korrelere med hverandre. Samvariasjonen mellom alle testledd beskriver i hvilken grad testen har indre konsistens; hvorvidt alle testledd kan relateres til det samme og dermed måler det samme (Nunnally & Bernstein, 1994). Tall på en tests indre

konsistens forklarer imidlertid ikke de komplekse mønstre av korrelasjoner som kan eksistere mellom ulike testledd eller hvorvidt ulike grupperinger av testledd kan relateres til mer generelle konstrukter. En faktoranalyse er et sett statistiske prosedyrer utformet for å kunne si noe mer spesifikt om en tests indre struktur og hvor mange overordnede konstrukter, kalt faktorer, som er nødvendig for å forklare og representere mønsteret av korrelasjoner mellom testledd (Fabrigar & Wegener, 2012). En faktoranalyse gir også tall på styrken og retningen av påvirkningen hver faktor har på hvert testledd, kalt testleddets faktorladning, samt et tall på hvor stor del av variansen i et testledd som forklares av de ekstraherte faktorene, kalt ekstrahert kommunalitet. Er antall ekstraherte faktorer og faktorladninger i tråd med teoretiske forventninger om konstruktet vil dette støtte den intenderte tolkningen av testskårer og testskårebruk (AERA et al., 2014). I denne studien er den andre forventningen at variasjonen i skårer kan forklares av én overordnet faktor, eller en endimensjonal modell, beskrevet som språkforståelse. En faktoranalyse kan videre også gi en indikasjon på hvorvidt noen testledd fører til at konstruktet underrepresenteres eller feilrepresenteres. Dersom et testledd har lav ekstrahert kommunalitet og faktorladning indikerer dette at variansen i skårene på testleddet ikke kan forklares av ekstraherte faktorer, men antas å skyldes påvirkning av unike faktorer (Fabrigar & Wegener, 2012). Lave ekstraherte kommunaliteter og faktorladninger kan dermed være et tegn på at testledd påvirkes på en systematisk måte av andre konstrukter enn det testen er utformet for å måle.

Hovedfokuset i valideringsprosessen beskrevet hittil har vært på systematiske målefeil i form av konstruktunderrepresentasjon og konstruktirrelevant varians. Det er imidlertid viktig å presisere at en forutsetning for en vurdering av validitet er at innvirkningen usystematiske målefeil har på barnets skåre er vurdert. Usystematiske målefeil knyttes til testens reliabilitet; hvorvidt testens måling er konsistent på tvers av ulike målinger og dermed kan vurderes som pålitelig (AERA et al., 2014). En standardisert administrering og skåring er utformet for å sikre at variasjoner i testadministrering og skåring ikke fører til usystematiske målefeil i barnets skårer. Visse aspekter ved testprosedyren vil imidlertid alltid variere, som tidspunktet testingen gjøres på og hvem som tester barnet. Ettersom en antagelse bak måling av evner som språk er at variasjonen i testskårer utformet for å måle språk er relativt stabil, vil store forskjeller i barnets skårer for eksempel mellom ulike måletidspunkt og ulike testere, ofte beskrevet som henholdsvis lav test-retest- og lav inter-rater-reliabilitet, indikere at usystematiske målefeil påvirker barnets testskåre (Nunnally & Bernstein, 1994). Testleddene som utgjør en test kan også betraktes som ulike målinger av det samme underliggende

konstruert. Dersom variasjonen i skårer på disse testleddene inneholder mye tilfeldig målefeil, for eksempel fordi spørsmålene er tvetydig formulert, vil dette reflekteres i lave korrelasjoner mellom testleddene. Indre konsistens er dermed også et reliabilitetsmål (Nunnally & Bernstein, 1994). Dersom en test har lav reliabilitet kan ikke skårene brukes til å si noe nøyaktig om det som måles, og slutningene som kan trekkes på bakgrunn av barnets testskårer blir dermed mindre sikre. På så måte er en vurdering av reliabilitet en forutsetning for validitet.

Tester er vanligvis utformet med den forventning at tolkningen og bruken av testskårer vil føre til positive konsekvenser for barnet som testes, som riktig diagnose som utløser rettigheter og valg av effektive intervensjoner og tiltak. Messick (1995) fremhever i denne sammenheng at validitet blir en sosial verdi som har en betydning også utenfor kartleggingssituasjonen. Spredning i typisk utviklede barns språkevne er naturlig å forvente (Simonsen et al., 2014) og i denne studien er den tredje forventningen at C-BiLLT er sensitiv nok til å fange opp slike forskjeller i språkforståelsesnivået både til barn i ulike aldersgrupper og til barn innad i samme aldersgruppe. Dette er viktig nettopp fordi en presis forståelse av barns språkforståelsesnivå er grunnleggende for å utforme tilpassede tiltak. Ifølge AERA, APA og NCME (2014) er en vurdering av hvorvidt slike positive konsekvenser faktisk vil realiseres ved bruken av testskårer også en kilde til validitetsevidens. Særlig viktig er å vurdere hvilke uintenderte konsekvenser konstruktunderrepresentasjon og konstruktirrelevant variasjon kan ha for intendert bruk av testskårer, og om dette kan føre til ulike konsekvenser for ulike undergrupper. Dersom for eksempel en test er for vanskelig kan dette føre til at testen ikke differensierer mellom de yngste eller de lavtfungerende barna ettersom alle vil få minimumskåre eller skåre på samme lave nivå. På samme måte vil en test som er for lett ikke differensiere mellom de eldste eller de høytfungernde barna ettersom alle vil få maksimumskåre eller skårer på samme høye nivå. Dette kalles henholdsvis gulveffekter og takeffekter (Price, 2017). Slike effekter påvirker bruken av testskårer til å presist beskrive individuell variasjon i domenet testen er utformet for å måle, noe som vil vanskeliggjøre eller føre til feilaktig bruk av testskårer for noen undergrupper.

I tråd med fokuset på konsekvensene av testtolkning fremhever Nunnally og Bernstein (1994) at valideringsprosessen teknisk sett omhandler bruken av måleinstrumentet, ikke måleinstrumentet i seg selv. Slutningene som kan trekkes på bakgrunn av en test er ikke valide i seg selv, men valide for et spesielt formål og en spesifikk gruppe. Dette forsterker

behovet for at testen valideres på et utvalg som er representativt for den populasjonen testen er intendert for. Price (2017) fremhever også valideringsprosessen som en prosess uten ende som innebærer en pågående akkumulering av evidens for testens bruksområde i nye tider og sammenhenger. De fleste tester må med tiden reevalueres og ny evidens fra utprøving av tester på nye grupper kan gi forslag til modifisering av eksisterende tester eller utvikling av alternative testtilnærminger. I kartlegging av barn med store tale- og bevegelsesvansker er det for eksempel ofte behov for tilrettelegging av standardiserte tester.

3.2 Konkrete kartleggingsutfordringer

Hva fremkommer av redegjørelsen over er at kartleggingsinstrumenter utformet for å kartlegge språkforståelse hos barn bør være grundig undersøkt hva gjelder validitet og reliabilitet slik at det kan heftes sikkerhet til tolkningen og bruken av testskårer. Barn med store tale- og bevegelsesvansker har et behov for ASK og tilpassede intervensjoner som bygger på en kartlegging av blant annet språkforståelse. I sin litteraturgjennomgang av standardiserte kartleggingsinstrumenter benyttet for å kartlegge språkforståelse hos barn med CP, konkluderte imidlertid Geytenbeek og kollegaer (2010b) med at få standardiserte og normerte kartleggingsinstrumenter var egnet til å trekke valide slutninger om språkforståelse hos barn med store tale- og bevegelsesvansker. Dette fordi alle språkforståelsestestene var standardisert på typisk utviklede barn og stilte krav til barnets verbale og motoriske evne da testene skulle besvares enten gjennom verbal respons, peking på bilder eller objekter som tilsvarer det testleder sier eller ved å manipulere objekter i samsvar med et spørsmål eller handlingsoppfordring. For barn med manglende eller uforståelig talespråk og med store motoriske vansker er slike tester uegnet fordi det er vanskelig å avgjøre om resultatene på testen skyldes nedsatt motorisk funksjonsevne eller nedsatt språkforståelse (Geytenbeek et al., 2010b; Sabbadini, Bonani, Carlesimo, & Caltagirone, 2001; Pirila et al., 2007). På grunn besvaringsmåten testene har oppstår det dermed konstruktirrelevant variasjon i testskårer for gruppen barn med store tale- og bevegelsesvansker. Slutningene som trekkes på bakgrunn av slike tester blir dermed usikre og kan føre til en underestimering av barnets faktiske språkforståelse, noe som igjen kan ha negative miljøeffekter i form av deprivasjon samt en negativ innvirkning på tilpasning av tiltak og ASK-hjelpemidler. Å bedømme barnets språkforståelse uten kartlegging kan videre føre både til en undervurdering og en overvurdering av barnets evner (Geytenbeek et al., 2010b).

Mangel på tester egnet for å kartlegge barn med store tale- og bevegelsesvansker kan få store konsekvenser. Registerstudier angir for eksempel at kun 29-32% av CP-populasjonen i Norge hadde mottatt ordinær kartlegging av kognitive evner, mens hos de resterende barna var vurdering av kognitiv fungering gjort på bakgrunn av klinisk bedømming eller et begrenset antall tester (34-54%) eller så var kognitivt funksjonsnivå ukjent (17-35%) (Andersen, et al., 2008; Stadskleiv et al., 2015). For svært mange av disse barna utformes dermed tiltak basert på antagelser om barnets funksjonsnivå, snarere enn en grundig kartlegging. Det er videre vist at det særlig er de barna med store tale- og bevegelsesvansker (GMFCS- og MACS-nivå IV-V) som ikke kartlegges til tross for at denne gruppen har et stort behov for tilpassede tiltak og ASK (Sigurdadottir et al., 2008; Sherwell et al., 2014). I denne sammenheng rapporterte Andersen og kollegaer (2008) at kun 34% av de barna som hadde alvorlige talevansker eller manglet tale helt brukte ASK-hjelpemidler. I en annen studie brukte 54% av barna med talevansker ASK-hjelpemidler, men et stort antall av disse brukte ASK-systemer som var lite tilpasset og lite funksjonelle i kommunikasjon (Andersen et al., 2010).

Vektleggingen av at eventuelle tiltak bør settes inn tidlig i barns utvikling (Tomblin et al., 2014; Bornstein et al., 2013; 2016) fordrer videre en tidlig kartlegging. En annen konkret kartleggingsutfordring er i denne sammenhengen mangelen på tilgjengelige normerte og standardiserte tester for barn i barnehagealder i Norge. WPPSI-IV har normer for barn fra 2;6 til 7;7 år og kartlegger verbal forståelse som en av de kognitive dimensjonene som utgjør generell intelligens. Dimensjonen kalles Verbal Forståelsesindeks og utgjøres av ulike deltester som til sammen måler barnets evne til begrepsdannelse, verbal resonnering og ervervet kunnskap. Den norske versjonen av British Picture Vocabulary Scale (BPVS; Lyster et al., 2010) måler barnets reseptive vokabular, og har normer fra 3;0 til 16;1 år. Det er imidlertid kun den andre versjonen av BPVS som har norske normer, ikke den tredje og nyeste, og det er dermed ikke mulig å kjøpe den nyeste versjonen av testen. TROG-2 kartlegger barns forståelse av setninger og grammatikk og har norske normer fra 4;0 til 16;5 år. Clinical Evaluation of Language Fundamentals (CELF; Semel, Wigg, & Secord, 2013) kartlegger også språkforståelse, men egner seg ikke for barn under fem år. Til slutt eksisterer Reynells Språktest som kartlegger språkforståelse for barn i alderen 1;6 til 6;11 år. Reynells tilhørende normer er imidlertid gamle. Alle disse testene besvares videre med verbal respons eller motorisk respons i form av peking eller manipulering av objekter og er dermed uegnet for barn med store tale- og bevegelsesvansker.

Til sammen fremmer dette et behov for en språkforståelsestest som dekker et større og yngre aldersspenn og som kan brukes til å trekke valide og reliable slutninger om språklig funksjonsnivå hos barn med store tale- og bevegelsesvansker.

3.3 Språkforståelsestesten C-BiLLT

C-BiLLT er en nederlandsk språkforståelsestest utviklet av Geytenbeek og kollegaer utfra behovet om en språkforståelsestest egnet i kartleggingen av barn med store tale- og bevegelsesvansker på GMFCS nivå IV og V. C-BiLLT er en databasert test og kan besvares med ulike responsmodaliteter som fingerpeking, blikkpeking og skanning med brytere. Den krever dermed ingen verbal respons og minimal motorisk respons for å besvares. Å benytte ulike responsmodaliteter har videre blitt vist å ikke påvirke testresultatene hos barn med typisk utvikling (Spillane, Ross, & Vasa, 1996; Casey, Tonsing, & Alant, 2007; Kurmanaviciute & Stadskleiv, 2017). Selv om C-BiLLT er tilpasset barn med store tale- og bevegelsesvansker, kan den like godt brukes for barn uten tale- og bevegelsesvansker. C-BiLLT kartlegger både forståelse av ord og setninger, er enkel å administrere, på et nettbrett eller en datamaskin, og har normer for nederlandske barn i alderen 1;6 til 7;0 år. C-BiLLT har dermed potensiale til å dekke behovet både for en test som kan brukes til å kartlegge barn med store tale- og bevegelsesvansker, og for en test som kan kartlegge språkforståelse hos yngre barn fra 1;6 år for å realisere tidlige og tilpassede intervensjoner. Ullevål sykehus seksjon for nevrohabilitering – barn har derfor oversatt C-BiLLT til norsk. Norsk blir det første språket C-BiLLT er oversatt til, men det pågår også et oversettelsesarbeid for en engelskspråklig versjon. Før denne norske versjonen av C-BiLLT kan begynne å brukes er det viktig at den går gjennom en valideringsprosess.

4 Metode

4.1 Design

Formålet med denne studien er å undersøke om C-BiLLT er egnet til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år. For å undersøke dette er det ønskelig å statistisk analysere 1;6 til 7;6 år gamle barns skårer på C-BiLLT samt å sammenligne disse skårene med skårer på andre tester av språkforståelse og skårer på tester av ikke-verbal resonneringsevne. En kvantitativ tilnærming med et kryss-seksjonelt design er derfor benyttet, da dette gjør det mulig å samle inn testskårer fra barn i ulike aldre. Da det ikke er måleinstrumentet, men bruken av måleinstrumentet som valideres, er det viktig at valg av deltakere er grunnet i hvordan testen vil brukes og populasjonen den er intendert for, som i dette tilfelle er barn i alderen 1;6 til 7;6 år. For å kunne oppdage en språkforståelsesvanske er det nødvendig å vite hva som er typisk oppnådde skårer på C-BiLLT i ulike aldersgrupper. I denne studien ble derfor populasjonen utvalget er trukket fra definert som barn med typisk utvikling i alderen 1;6 til 7;6 år.

En tests reliabilitet er en forutsetning for vurdering av validitet. Undersøkelsen av reliabilitet blir behandlet i Haddelands (2018) masteroppgave, men reliabilitetsundersøkelsen vil kort redegjøres for da hovedfunnene i denne undersøkelsen vil presenteres før diskusjonen av resultatene fra validitetsundersøkelsen begynner. For å kunne undersøke test-retest-reliabilitet ble et utvalg av de eldste barna retestet med C-BiLLT to til fire uker etter førstegangstesting. Et utvalg av disse barna ble videre retestet med en annen testadministrator enn ved førstegangstesting for å undersøke inter-rater-reliabilitet. Barna som ble retestet ble også delt i tre grupper utfra hvilke responsmodaliteter de brukte til å besvare C-BiLLT med, for å undersøke om testbesvarelsesmåten barnet benyttet hadde en innvirkning på barnets C-BiLLT-skåre. Den ene gruppen besvarte C-BiLLT med blikkpeking ved førstegangstesting og fingerpeking ved retest, den andre gruppen besvarte C-BiLLT med fingerpeking ved førstegangstesting og blikkpeking ved retest, mens den tredje gruppen besvarte C-BiLLT med fingerpeking ved begge testdager.

4.2 Deltakere

For å sikre at utvalget representerer populasjonen hva gjelder typisk utvikling og alder ble følgende inklusjonskriterier satt for utvalget:

- 1) Alder mellom 1;6 og 7;6 år
- 2) Har ikke hatt vesentlig forsinket språk- eller taleutvikling
- 3) Har normalt syn og normal hørsel (eventuelt korrigert med briller/høreapparat)
- 4) Har ikke kroniske sykdommer eller generell forsinket utvikling
- 5) Norsk er rapportert som barnets morsmål (eventuelt i tillegg til andre språk)

Målet var å rekruttere et utvalg på 80 da dette regnes som stort nok for en pilotstudie (Johanson & Brooks, 2010). Målet var videre å rekruttere fem barn per seks måneders aldersgruppe i den yngste aldersgruppen (1;6 til 5;6 år) og ti barn per seks måneders aldersgruppe i den eldste aldersgruppen (5;6 til 8;0). Det var ønsket at den eldste aldersgruppen skulle være størst fordi det var i denne gruppen endret responsmodalitet skulle undersøkes og det var regnet med at de eldste barna forstod instruksjonene tilknyttet blikkpeking best. Under oversettelsen av C-BiLLT ble noen av setningene vurdert som så komplekse at det også var ønskelig å undersøke om C-BiLLT ville differensiere mellom barn i alderen 7;6 år. Dette er grunnen til at også barn mellom 7;0 og 8;0 år er inkludert i denne studien, selv om normene i den nederlandske versjonen av C-BiLLT kun går opp til 7;0 år.

Hovedvekten av rekrutteringen ble gjort av meg selv og masterstudent Haddeland i perioden oktober til desember 2017. Totalt ble 580 forespørsler delt ut til foresatte i syv ($n = 373$) barnehager og seks ($n = 207$) skoler i Oslo. Skolene og barnehagene lå øst og vest i Oslo samt i indre og ytre by. Av de 580 samtykket 204 til å delta i studien (35.2%), 106 fra barnehager og 98 fra skoler. Av disse var det seks barn som ikke ville eller turte å bli testet dagen testingen foregikk og ett barn i den yngste aldersgruppen (1;6-2;0) som ikke bestod C-BiLLTs pretest og dermed ble ekskludert fra studien. I tillegg var det ti barn som ikke ble testet fordi de enten var borte de dagene vi hadde avtalt å være på skolen eller barnehagen eller fordi vi ikke mottok samtykkeskjemaet før etter vi hadde vært på den aktuelle skolen eller barnehagen. Av disse ti hadde åtte oppgitt en kontaktadresse og disse ble invitert til en senere testing på Ullevål sykehus lørdag 17. mars 2018. Ingen av de foresatte ønsket imidlertid dette. Totalt antall testet ble dermed 187 av 204 (91.7%).

Tabell 1 gir en oversikt over fordelingen barn per seks måneders aldersgrupper. Den yngste deltakeren var 20 måneder og den eldste var 95 måneder ($M = 64.03$, $SD = 20.83$).

Tabell 1. Fordeling av barn i ulike aldersgrupper

Aldersgruppe	Antall	%
1;6 – 2;0	4	2.1
2;1 – 2;6	7	3.7
2;7 – 3;0	19	10.2
3;1 – 3;6	11	5.9
3;7 – 4;0	14	7.5
4;1 – 4;6	11	5.9
4;7 – 5;0	6	3.2
5;1 – 5;6	17	9.1
5;7 – 6;0	5	2.7
6;1 – 6;6	32	17.1
6;7 – 7;0	28	15.0
7;1 – 7;6	28	15.0
7;7 – 8;0	5	2.7
Totalt	187	100

Av deltakerne var 96 gutter (51.3%) og 91 jenter (48.7%). Tabell 2 gir en oversikt over foreldrerapporterte tilstander og sykdommer hos deltakerne. Disse barna ble inkludert i studien da tilstandene vanligvis ikke er assosiert med nedsatt språkforståelse. Alle barn med synsvansker korrigert av briller ($n = 4$, 2.1%) ble også inkludert i studien. For ett barn (0.5%) var det oppgitt forsinket taleutvikling i form av uttalevansker. Dette barnet ble også inkludert i studien fordi vansken ikke var en språkforståelsesvanske.

Tabell 2. Oversikt over foreldrerapporterte tilstander

	Type	Antall	%
Kronisk sykdom	Astma/allergi	3	1.6
	Eksem	1	0.5
Annet	Leppe-kjeve-ganespalte	1	0.5
Totalt		5	2.7

Det var 19 (10.2%) av de deltagende barna som hadde et annet morsmål i tillegg til norsk. En oversikt over disse barnas andre morsmål gis i Tabell 3. De tospråklige barna ble inkludert i studien da en uavhengig *t*-test avdekket at det ikke var signifikante forskjeller ($p < .05$) i gjennomsnittlig alderskontrollerte skårer mellom gruppen enspråklige og gruppen tospråklige barn.

Tabell 3. Oversikt over andre morsmål hos studiens deltakere

	Språk	Antall	%
Europeiske språk	Islandsk	1	0.5
	Svensk	2	1.1
	Engelsk	4	2.1
	Tysk	1	0.5
	Spansk	2	1.1
	Portugisisk	3	1.6
	Serbisk	1	0.5
Afrikanske språk	Berbisk	2	1.1
	Swahili	1	0.5
Asiatiske språk	Thai	1	0.5
	Kinesisk	1	0.5
Totalt		19	10.2

4.3 Kartleggingsinstrumenter

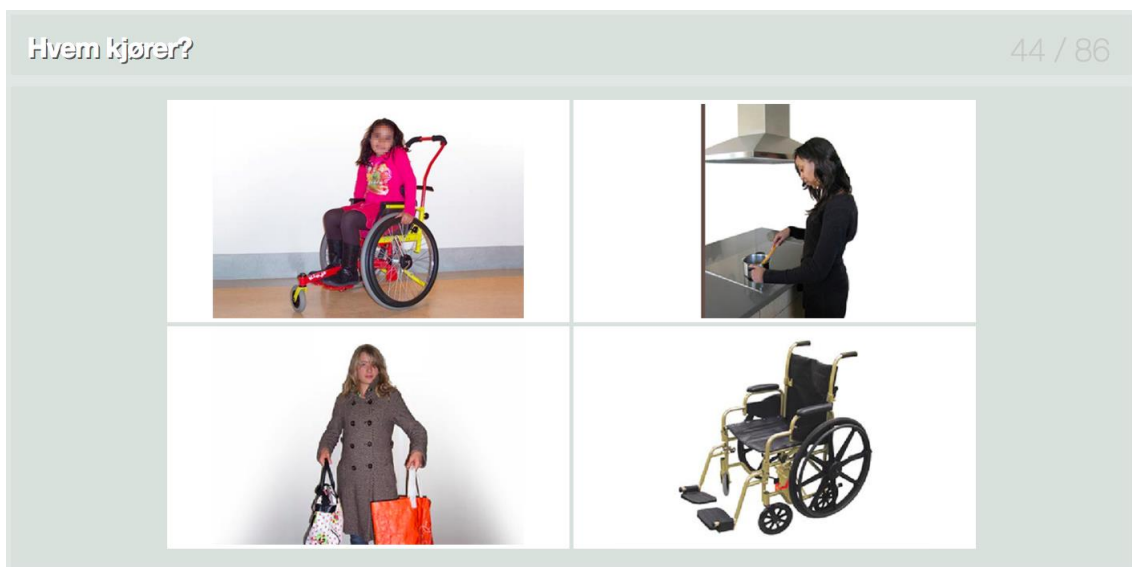
Valg av tester ble gjort på bakgrunn av den nederlandske undersøkelsen av reliabilitet og validitet (Geytenbeek, Morkink, Knol, Vermeulen, & Oostrom, 2014) samt på bakgrunn av at de har standardisert administrering og skåring. Alle studiens deltakere ble testet med C-BiLLT. Språkforståelse ble også undersøkt med forståelsesdelen av Reynells språktest, Ordgjenkjenning fra WPPSI-IV og TROG-2. I den nederlandske undersøkelsen av C-BiLLT ble Peabody Picture Vocabulary Test (PPVT) benyttet istedenfor Ordgjenkjenning og TROG-2. Den nyeste versjonen av BPVS, som er den britiske versjonen av PPVT, mangler norske normer og ble derfor ikke benyttet i denne studien. Ettersom det var ønskelig å sammenligne skårer på C-BiLLT med skårer på en test som undersøker forståelse av setninger ble derfor også TROG-2 brukt i denne studien. Ikke-verbal resonneringsevne ble i denne studien

undersøkt med CPM og Terningmønster fra WPPSI-IV. I den nederlandske studien av C-BiLLT ble kun CPM benyttet. Terningmønster er inkludert i denne studien da denne har normer ned til 2;6 år. Da det er validiteten til slutningene som kan trekkes på bakgrunn av den norske versjonen av C-BiLLT som undersøkes i denne oppgaven, vil redegjørelsen for C-BiLLT være mer omfattende enn redegjørelsen for de andre testene.

4.3.1 C-BiLLT

Bakgrunn for utvelging av innhold og bilder

C-BiLLT undersøker forståelse av talespråk og er normert på nederlandske barn i alderen 1;6 til 7;0 år. C-BiLLT bygger på den nederlandske versjonen av Reynells språktest i sin oppbygning av lingvistisk kompleksitet. Utvelgingen av testledd er videre motivert av funn på at barn med fysiske funksjonsnedsettelse kan ha begrensede livserfaring grunnet begrensede muligheter til å utforske sitt miljø (Rosenbaum et al., 2007; Geytenbeek et al., 2010a). Det er dermed bestrebet å inkludere testledd som også er gjenkjennelige for barn med store tale- og bevegelsesvansker. For eksempel spør noen testledd etter mennesker som kjører rullestol, som Testledd 44: *Hvem kjører?* (Figur 3).



Figur 3. Testledd 44 fra C-BiLLT

Bildene i C-BiLLT er tilpasset med tanke på at mange barn med CP har perseptuelle vansker som hjernesynshemming (Eng. cerebral visual impairment) som blant annet kan gjøre det vanskelig å skille figur fra bakgrunn (Fazzi et al., 2012). Bildene er derfor klare og har en hvit bakgrunn som er kontrasterende til objektene, personene og dyrene det spørres om. Ved

at bildene barna skal peke på er store, stilles det også mindre krav til at barnet må peke presist.

Oppbygning

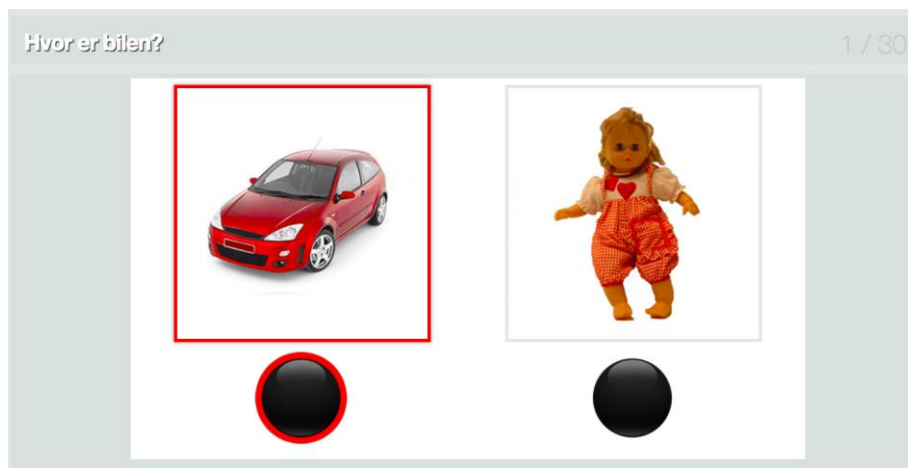
C-BiLLT utgjøres av to pretester, en læringsmodul og to datatester. De to pretestene er *pretest gjenstander* og *pretest foto*. I *pretest gjenstander* benyttes 8 kjente objekter fra barnets hjemmemiljø, mens i *pretest foto* benyttes 8 fotografier av gjenstander printet ut i A5-format. Pretestene anvendes for å vurdere om barnet kan kommunisere et valg mellom to konkrete objekter eller digitale fotografier av objekter når de hører objektets navn. Under pretesten holder testleder opp to og to bilder eller objekter, én i hver hånd, i en bestemt rekkefølge samtidig som hun spør ”Hvor er x?”. Barnet kan kommunisere sitt valg ved peking, blikkpeking eller å strekke seg etter riktig objekt eller bilde. Dersom barnet er i stand til å identifisere minst fem av de åtte objektene eller fotografiene kan barnet fortsette til den digitale delen av C-BiLLT.

Læringsmodulen er den første digitale delen av C-BiLLT. Denne brukes for at barnet skal få mulighet på å øve seg og bli kjent med sin anvendte responsmodalitet og for at testleder kan vurdere om det bør foretas noen justeringer i responsmodalitet. Barnet får blant annet se to tomme ruter på dataskjermen, og når barnet med sin foretrukne responsmodalitet velger en av disse rutene, for eksempel ved å trykke på den med fingeren, flyr en ballong opp i ruten (Figur 4). Læringsmodulen kan også anvendes til å bevisstgjøre barnet på at dets respons vil vises visuelt som en rød firkant rundt det valgte fotografiet.



Figur 4. Læringsmodulen fra C-BiLLT

Datatesten består av 86 testledd fordelt på to deler. Del 1 av *Datatesten* kartlegger vokabularforståelse og består av tre primærseksjoner med ti testledd hver. I to seksjoner skal barnet velge riktig bilde utfra talte substantivfraser som refererer til dyr, objekter og personer. I en seksjon skal barnet velge riktig bilde utfra talte verbfraser. For hvert testledd presenteres to bilder horisontalt på skjermen; ett som tilsvare den talte setningen og ett som ikke tilsvare den talte setningen. For å kontrollere for 50% sjansoeffekter har hver primærseksjon en parallellidel som administreres hvis barnet får én eller flere feil på en av primærseksjonene. I parallellidelen presenteres bildene i en annen rekkefølge og i kombinasjon med et annet bilde (se Figur 5 og Figur 6).



Figur 5. Testledd 1 fra Del 1 av C-BiLLT

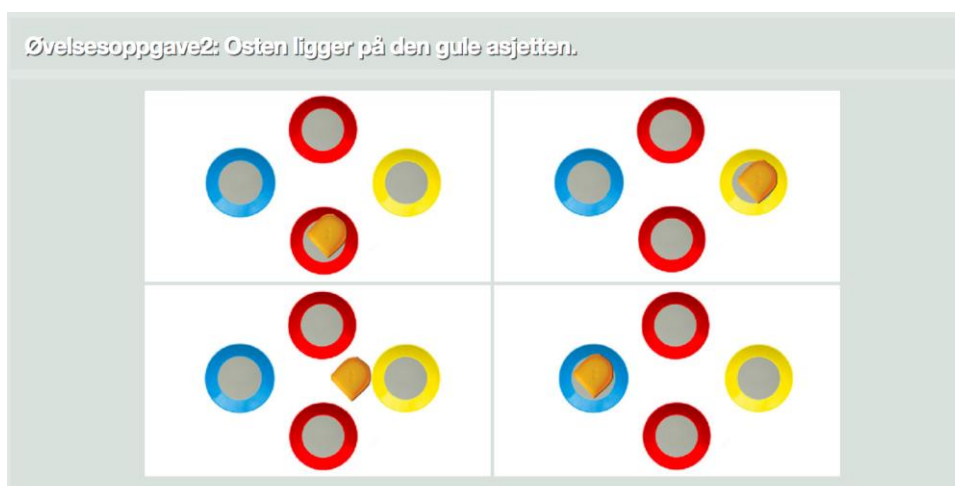


Figur 6. Testledd P9 fra parallellseksjonen av C-BiLLT

Del 2 av *Datatesten* består av 56 testledd. Barnet skal i denne delen velge riktig bilde utfra talte setninger som øker i kompleksitet av grammatisk struktur. De fire første testleddene (oppgave 31-34) kartlegger vokabularforståelse, mens de resterende testleddene (oppgave 35-86) kartlegger setningsforståelse. Del 2 er organisert i ni seksjoner. For hvert testledd presenteres nå fire bilder i en 2 x 2 matrise på skjermen, ett bilde tilsvare den talte setningen og tre bilder passer ikke til den talte setningen (Figur 7). Del 2 inneholder videre tre testledd, ett i begynnelsen av Seksjon 4 og to i begynnelsen av Seksjon 11, som er øvelsesoppgaver (Figur 8). Svarer barnet feil på disse oppgavene kan testleder korrigere barnet for å forsikre seg om at barnet forstår instruksjonene.



Figur 7. Testledd 38 fra Del 2 av C-BiLLT



Figur 8. Øvelsesoppgave 2 fra Del 2 av C-BiLLT

Tabell 4 gir en oversikt over de ulike delene og tilhørende seksjoner fra C-BiLLTs datatest. Det er også oppgitt eksempler på testledd fra hver seksjon.

Tabell 4. Seksjoner fra C-BiLLTs datatest med eksempler på testledd. Etter Geytenbeek et al., 2014, s. 7

	Antall testledd	Identifikasjon av	Eksempler
Del 1			
Seksjon 1	10	Substantiv	<i>Hvor er bilen?</i>
Seksjon 2	10	Verb	<i>Hvem sover?</i>
Seksjon 3	10	Dyr, objekter og personer	<i>Hvor er babyen?</i> <i>Hvor er støvsugeren?</i>
Del 2			
Seksjon 4	4	Vanskeligere substantiv	<i>Hvor er paraplyen?</i>
Seksjon 5	5	Enkle setninger om gjenstander i kombinasjon med verb og preposisjoner	<i>Hva kan du sove i?</i> <i>Hunden ligger i kurven</i>
Seksjon 6	5	”Hvem”-spørsmål om personer som utfører handlinger	<i>Hvem skal lufte hunden?</i>
Seksjon 7	4	Passive setninger med objekter og preposisjoner kombinert med hendelser	<i>Gutten blir dyttet av jenta</i>
Seksjon 8	9	Setninger med mer grammatisk og syntaktisk kompleksitet	<i>En av tannbørstene som er satt på plass i koppen, er lang</i>
Seksjon 9	6	Enkle aktive setninger som refererer til ikke-observerbare situasjoner med fire personer	<i>Jakob og Emma skal leke ute, hvem blir hos mamma?</i>
Seksjon 10	9	Komplekse setninger med to eller flere begreper	<i>Ett lite syltetøyglass står ved siden av det røde syltetøyglasset</i>
Seksjon 11	4	Sammensatte, komplekse setninger	<i>Først var det et eple på bordet men nå er det en banan der i stedet</i>
Seksjon 12	10	Sammensatte, komplekse setninger - fortsettelse	<i>Alt pålegget er lagt på asjettene, men det er bare på de asjettene med samme farge at det ligger ost</i>

Administrering og skåring

C-BiLLT har standardisert administrering. Det er vektlagt at testen skal administreres i rolige omgivelser. Barnets plassering foran skjermen bør videre tilpasses hvordan barnet skal besvare testen. Skal barnet bruke direkte peking bør barnet for eksempel plasseres i en nær nok avstand fra berøringsskjermen til at det kan nå frem til den samtidig som det bør

plasseres langt nok unna til at det kan få en oversikt over alle bildene. Skal barnet bruke blikkpeking bør barnet plasseres slik at øyesporeren kan fange opp hvor barnet ser på skjermen. Avstanden mellom skjermen og barnets øyne bør da tilsvare den for en person som sitter foran en vanlig dataskjerm (Tobii Technologies, 2018). *Læringsmodulen* kan her benyttes for å vurdere barnets posisjon foran skjermen. Testleder bør sitte ved barnets høyre side slik at hun kan oppnå kontakt med barnet samtidig som hun kan følge med på skjermen. Testleder skal lese opp instruksjoner og oppgavespørsmål nøyaktig slik det står skrevet i manualen. Oppgavesetningene bør videre leses opp med naturlig prosodi og hastighet. I *Datatesten* står oppgavesetningen skrevet øverst i venstre hjørne på skjermen for hvert testledd slik at testleder kan lese denne opp direkte fra skjermen. Dette kan deaktiveres for de eldre barna som kan lese. Oppgavesetningen skal som hovedregel kun leses opp én gang, men kan gjentas én gang hvis barnet ikke får setningen med seg grunnet støy eller andre distraksjoner. Det er testleder som ved å trykke på tastaturets piltaster bytter videre til neste bilde i testen etter barnet har gitt sin besvarelse. C-BiLLT skal videre avbrytes etter 8 påfølgende nullpoengsvar. C-BiLLT har ingen tidsbegrensing og det fremheves at testleder bør gi barnet god tid til å velge ett av svaralternativene. Å fjerne tidsbegrensning på tester har i denne sammenhengen vist å gi et riktigere bilde av forståelsesnivået til barn med motoriske vansker (Sherwell et al., 2014).

C-BiLLT har også standardisert skåring. Barnet får ett poeng for *pretest gjenstander* hvis de identifiserer minst fem av gjenstandene, og ett poeng for *pretest foto* hvis de identifiserer minst fem av fotografiene, slik at høyest oppnådd pretestskåre er to poeng. *Læringsmodulen* gir ingen poeng. I *Datatesten* Del 1 får barnet ett poeng riktig per oppgave slik at maksimum sumskåre for Del 1 er 30. Dersom parallellseksjonen administreres får barnet kun poeng dersom samme testledd er korrekt både i primærseksjonen og parallellseksjonen. I *Datatesten* Del 2 får barnet også ett poeng riktig per oppgave slik at maksimum sumskåre Del 2 er 56. Ingen poeng gis for øvingsoppgavene i Del 2. Maksimum oppnådde sumskåre for C-BiLLT (Pretest, Del 1 og Del 2) er til sammen 88.

Reliabilitet og validitet

Reliabiliteten og validiteten til slutningene og bruken av den nederlandske versjonen av C-BiLLT er god. Tabell 5 gir en oppsummering over konvergerende evidens, i form av høye korrelasjonen mellom skårer på C-BiLLT og Reynell og C-BiLLT og PPVT, og diskriminerende evidens, i form av lavere korrelasjoner mellom C-BiLLT og CPM, som til

sammen støtter opp under validiteten. Validiteten ble videre støttet ved en eksplorerende faktoranalyse i Geytenbeek og kollegaer (2014) sin studie, som resulterte i en endimensjonal faktorløsning karakterisert som en generell språkforståelsesevne. Faktorladningene var høye og den endimensjonale strukturen forklarte 76% av variansen i dataen (se Vedlegg 4). I Geytenbeek og kollegaer (2014) sin studie ble også den teoretiske forventninger om signifikante forskjeller i gjennomsnittlig sumskåre på C-BiLLT for ulike helårsgrupper innfridd (Vedlegg 4). Tabell 5 gir videre en oversikt over reliabilitetsstøtte i form av høye tall på inter-rater-, intra-rater- og test-retest-reliabilitet. Reliabiliteten ble ytterligere støttet i Geytenbeek og kollegaer (2014) sin studie som viste at C-BiLLT hadde god indre konsistens for barn med typisk utvikling, Chronbach's alpha = .88, og barn med CP, Chronbach's alpha = .91.

Tabell 5. Reliabilitet og validitet tilknyttet den nederlandske versjonen av C-BiLLT

År	N*	Aldersspenn	Test**	Analyse	Utfall***
2008	166	4;0-6;3	C-BiLLT/Reynell	Konvergerende evidens	$r = .93$
2009	212	1;8-4;0	C-BiLLT/Reynell	Konvergerende evidens	$r = .93$
2009	111	4;0-6;3	C-BiLLT/C-BiLLT	Test-retest	ICC: 0.97
2009	47	1;8-6;3	C-BiLLT/C-BiLLT/ Reynell	Inter-rater	ICC: 0.96
2010	113	2;3-7;0	C-BiLLT/PPVT-III	Konvergerende evidens	$r = .88$
2010	103	4;0-5;11	C-BiLLT/CPM	Diskriminerende evidens	$r = .43$
2011	54	1;8-4;0	C-BiLLT/C-BiLLT	Test-retest	ICC = 0.97
Total	806 TU				
2011	92	1;7-11;11	C-BiLLT/PPVT-III	Konvergerende evidens	$r = .87$
2012	32	1;11-11;11	C-BiLLT/C-BiLLT	Intra-rater	ICC: 0.97
2012	35	1;7-11;4	C-BiLLT/C-BiLLT	Inter-rater	ICC: 0.97
Total	159 CP				

* TU = barn med typisk utvikling, CP = Barn med Cerebral parese

** C-BiLLT = Computer-Based Instrument for Low motor Language Testing; Reynell = språkforståelsesdelen av Reynells språktest; PPVT-III = Peabody Picture Vocabulary Test third edition; CPM = Ravens Coloured Progressive Matrices

*** r = Pearsons produkt moment korrelasjonskoeffisient, ICC = Intraclass Correaltion Coefficient

4.3.2 Språkforståelsesdelen av Reynell

Reynells språktest er en språktest for barn i alderen 1;6 til 6;11 år (Hagtvat & Lillestølen, 1990). Testen undersøker språkproduksjon og språkforståelse som to ulike indekser, og i denne studien er kun språkforståelsesdelen benyttet. Hagtvat og Lillestølen (1990) beskriver språkforståelsesdelen som å ha et utviklingspsykologisk grunnlag der hvert testavsnitt er ment å reflektere et utviklingsstadium i barns typiske språkutvikling. Reynell kartlegger både forståelse av ord og setninger. Avsnittene er delt opp og oppbygd etter *Situasjonsavhengig forståelse*, *Ordforståelse*, *Relasjonsforståelse* og *Setningsforståelse* og oppgavespørsmålene blir mer situasjonsuavhengige utover i testen. Språkforståelsesdelen av Reynell utgjøres av til sammen ti avsnitt der hvert avsnitt har et tilhørende sett av konkrete. Testleder stiller konkretene opp foran barnet og leser opp ulike spørsmål eller handlingsoppfordringer tilknyttet objektene, for eksempel ”Hvor er ballen?” eller ”Sett en av de små grisene ved siden av den svarte grisen!”. Barnets oppgave er å svare på testleders utsagn ved å peke på riktig konkret eller utføre riktig handling med konkretene. Testen består av totalt 67 oppgaver og barnet kan få ett poeng riktig per oppgave slik at maksimumskåre er 67. Testen har ingen stoppkriterier og det anbefales å forsøke å gjennomføre hele testen med alle barn.

Reliabiliteten tilknyttet den norske versjonen av språkforståelsesdelen i Reynell er undersøkt ved split-half-reliabilitet, og test-retest reliabilitet (Hagtvat & Lillestølen, 1990).

Reliabilitetskoeffisientene for split-half-reliabilitet var svært akseptable i aldersgruppene 1;6-5;0, og noe lavere for barn i seksårsalderen. Korrelasjonen mellom test og retest var videre høy for språkforståelsesdelen av Reynell. Hva gjelder validitet er korrelasjonen mellom språkforståelse og talespråk undersøkt. Da korrelasjonskoeffisienten ikke var høy i noen alderstrinn antas det å gi støtte for at språkforståelses- og talespråkdelen i Reynell måler relaterte, men distinkte språkfunksjoner. Det er ikke oppgitt noen andre mål på validitet og det kan derfor knyttes usikkerhet rundt validiteten til de slutningene som kan trekkes på bakgrunn av Reynell. Ettersom dette er den eneste testen som dekker aldersspennet 1;6-6;11 år, og det er den nederlandske versjonen av Reynell som C-BiLLT bygger på, er denne testen likevel benyttet i denne studien.

4.3.3 Ordgjenkjenning fra WPPSI-IV

Ordgjenkjenning er en deltest hentet fra WPPSI-IV for barn i alderen 2;6-7;7 år (Wechsler, 2015). WPPSI-IV kartlegger generell intelligens gjennom fullskala-IQ og ulike kognitive dimensjoner fullskala-IQ antas å være utgjort av. Deltesten Ordgjenkjenning bidrar til å måle den kognitive dimensjonen Verbal Forståelsesindeks. Ordgjenkjenning måler mer spesifikt reseptivt ordforråd. Denne deltesten består av 31 oppgaver der barnet i hver oppgave får se fire bilder presentert i en 2 x 2 matrise i en stimulusbok. Testleder leser så opp en setning, for eksempel ”Pek på den som klamrer seg fast”, og barnet skal peke på det bildet som tilsvarer det testleder sier. Barnet får ett poeng per korrekte svar slik at maksimumskåre er 31. Testen avbrytes etter tre påfølgende nullpoengsvar.

Indre konsistens for den norske utgaven av Ordgjenkjenning er god og den gjennomsnittlige reliabilitetskoeffisienten innenfor hver aldersgruppe er høy, noe som indikerer god reliabilitet. Ordgjenkjenning har også påvist god validitet (Wechsler, 2015).

4.3.4 TROG-2

TROG-2 er en språktest som anvendes for å vurdere forståelse av setninger, eller reseptiv grammatikk, hos barn og ungdommer i alderen 4;0 til 16;5 år (Bishop, 2009). TROG-2 består av 80 oppgaver fordelt på 20 blokker. I hver oppgave får barnet se fire bilder presentert i en 2 x 2 matrise i en stimulusbok. Testleder leser så opp en setning, for eksempel ”Boka som er rød, ligger på blyanten”, og barnet skal peke på det av de fire bildene som tilsvarer setningens betydningsinnhold. Dersom barnet ikke svarer korrekt på alle fire oppgaver i fem etterfølgende blokker, eller ikke svarer på noen av oppgavene i to etterfølgende blokker, avbrytes testen. Barnets skåre regnes ut fra totalt antall mestrede blokker.

Den norske versjonen av TROG-2 har oppgitt høy indre konsistens, noe som indikerer god reliabilitet. Det er videre påvist konvergerende støtte til validitet ved signifikant korrelasjon mellom den norske versjonen av TROG-2 og BPVS i alle aldersgrupper.

4.3.5 CPM

CPM er en test for barn i alderen 4;0 til 11;11 år (Raven, 2008). Testen er en del av *The Raven Progressive Matrices* som er et sett ikke-verbale evnetester brukt for å måle generell kognitiv evne og mer spesifikt evnen til å utlede mening fra komplekse, ikke-verbale situasjoner. CPM består av tre oppgavesett der hvert sett inneholder 12 matriser hver. I hver oppgave får barnet se en ufullstendig matrise i en stimulusbok. Under matrisen er det presentert seks ulike deler som har lik form og størrelse til å utfylle matrisen, men der kun én del har et mønster som vil gjøre at det dannes en visuell helhet hvis den plasseres i matrisen. Barnets oppgave er å velge hvilken av de seks delene som passer inn i matrisen slik at det dannes en visuell helhet. Barnet får ett poeng per korrekt svar slik at maksimumskåre er 36. CPM har ikke stoppkriterier og det anbefales å forsøke å gjennomføre hele testen. CPM har påvist god validitet og reliabilitet ved flere kryss-kulturelle studier (Raven, 2008).

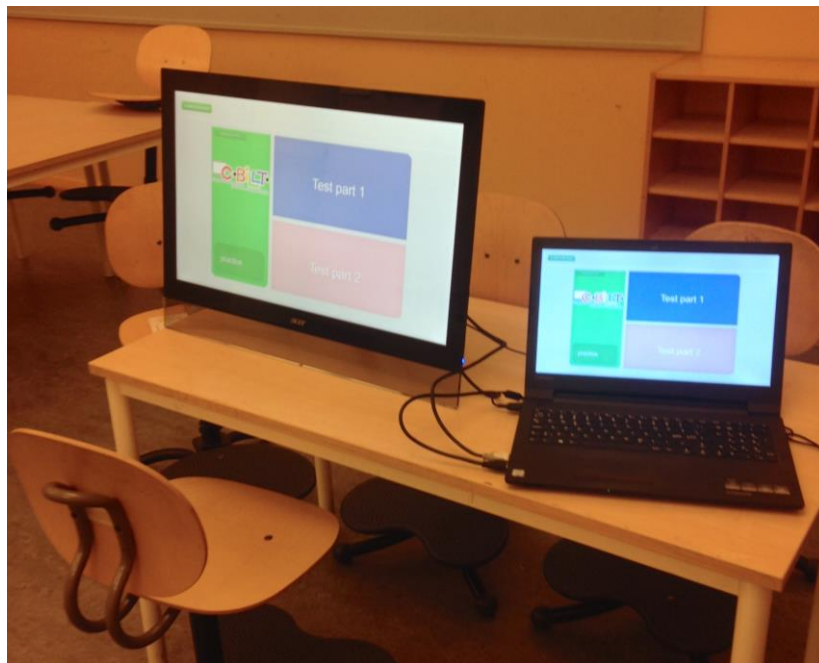
4.3.6 Terningmønster fra WPPSI-IV

Terningmønster er en ikke-verbal deltest fra WPPSI-IV for barn i alderen 2;6 til 7;7 år (Wechsler, 2015). Terningmønster er en av deltestene som bidrar til å måle visuospatial indeks. Denne indeksen omhandler ulike former for ikke-verbal resonnering som evnen til integrering av forholdet mellom deler og helhet, spatial prosessering og visuomotoriske evner. Terningmønster består av til sammen 17 oppgaver fordelt på to deler. I Del A (åtte oppgaver) benyttes ensfargede klosser som barnet skal sette sammen til ulike tårn og mønstre utfra testleders modellering. I Del B (ni oppgaver) benyttes rødhvite klosser som barnet skal sette sammen til ulike mønstre først etter testleders modellering i kombinasjon med et bilde i en stimulusbok og etter hvert kun etter bildet i stimulusboken. Barnet kan få maksimum to poeng per oppgave slik at den høyeste skåren er 34 poeng. Alle oppgaver har en tidsgrense og hvis barnet ikke fullfører oppgaven innenfor tidsgrensen gis null poeng. Testen avbrytes etter to påfølgende nullpoengsvar.

Indre konsistens for den norske utgaven av Terningmønster er god og den gjennomsnittlige reliabilitetskoeffisienten innenfor hver aldersgruppe er høy, noe som til sammen indikerer god reliabilitet. Terningmønster har også påvist god validitet (Wechsler, 2015).

4.4 Digitale verktøy

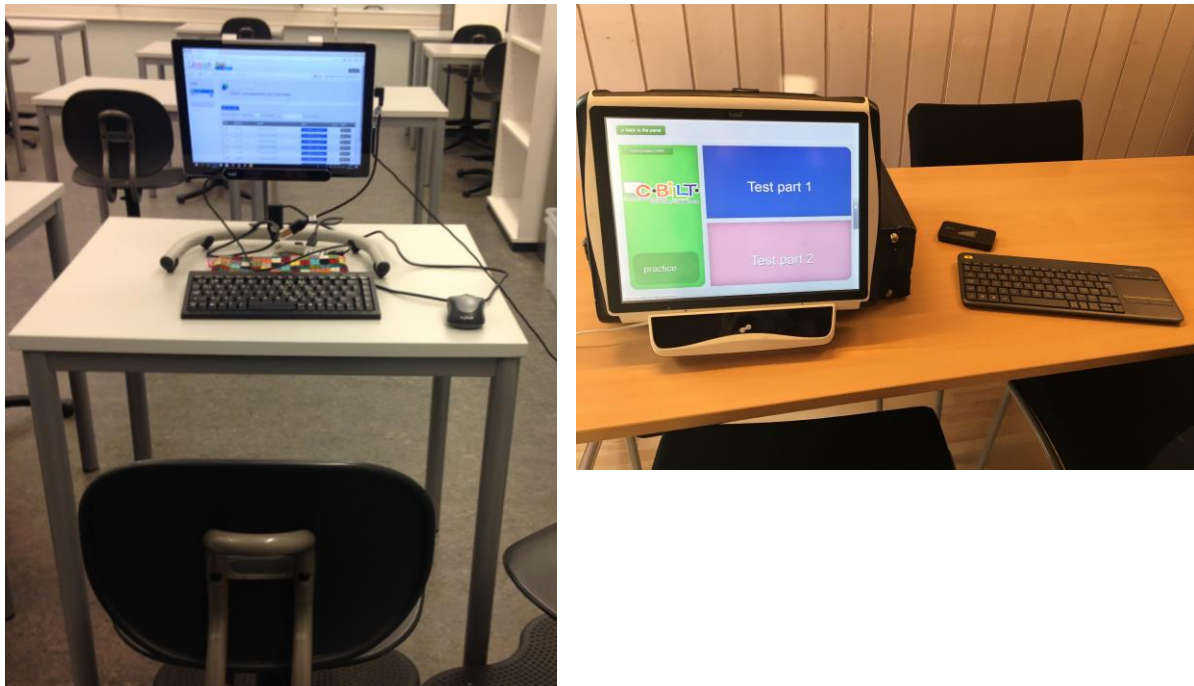
C-BiLLT ble i denne studien besvart gjennom to ulike responsmodaliteter; fingerpeking og blikkpeking. Til sammen ble det derfor benyttet tre ulike testoppsett i administreringen av C-BiLLT. Testoppsett 1 ble benyttet for fingerpeking. Her ble C-BiLLT administrert på en 19-tommers berøringsskjerm tilkoblet en PC (Figur 9). C-BiLLT ble her besvart ved at barnet trykket på det riktige svaralternativet med fingeren på berøringsskjermen, og en rød ramme dukket opp og signaliserte hvilket bilde som var valgt.



Figur 9. Testoppsett 1

Ved besvarelse gjennom blikkpeking ble to varianter av MyTobii øyestylingssystem benyttet (Figur 10). Ved Testoppsett 2 ble C-BiLLT administrert på et nettbrett med tastatur stående i et stativ. Tilkoblet nettbrettet var en øyesporer av typen Tobii PCEye Mini. Ved Testoppsett 3 ble C-BiLLT administrert på en MyTobii med berøringsskjerm og integrert blikkpekingssystem. Begge øyestylingssystemene fungerer ved at øyesporeren festet på eller integrert i skjermens ramme sender ut infrarødt lys som reflekteres i barnets øyne. I et kamera i øyesporeren blir det så tatt bilder av øynene og refleksjonene fra det infrarøde lyset, som gjør det mulig for øyestylingssystemet å regne ut øynenes posisjon og hvor brukeren ser på dataskjermen (Tobii Technologies, 2018). Hvis barnet holder blikket på et punkt på skjermen vil dette registreres som et museklikk. Barna som benyttet blikkpeking besvarte dermed C-BiLLT ved å holde blikket på bildet som tilsvarte den talte setningen og det valgte bildet

ble synlig ved at en rød ramme dukket opp rundt bildet. Ettersom begge skjermene benyttet i Testoppsett 2 og Testoppsett 3 var berørings skjermer ble disse testoppsettene også benyttet for besvarelse ved fingerpeking.



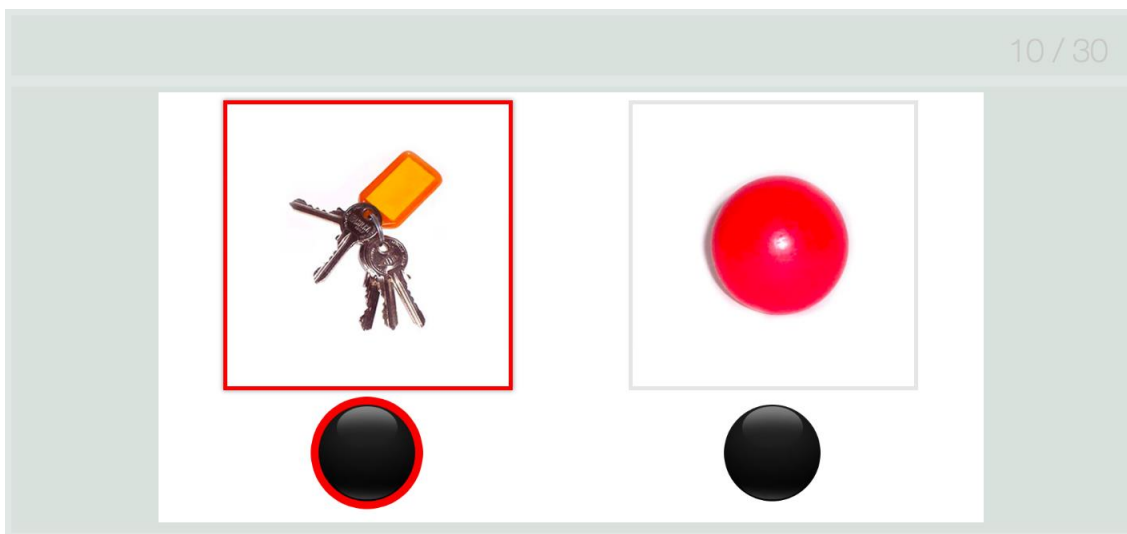
Figur 10. Testoppsett 2 (t.v.) og Testoppsett 3 (t.h.)

4.5 Prosedyre

4.5.1 Testoversettelsesprosessen

Oversettelsen av C-BiLLT fra nederlandsk til norsk bestod av tre delprosesser. I den første delprosessen gjennomgikk to uavhengige fagpersoner et oversettelsesutkast. Fagpersonene var en lege med norsk som førstespråk og nederlandsk som andrespråk og en logoped med nederlandsk som førstespråk og norsk som andrespråk. Begge har erfaring fra å jobbe med barn med tale-, språk- og bevegelsesvansker. Legens oversettelse var lik som førsteutkastet for 71 av de 86 oppgavene (82.6%), mens logopedens oversettelse var lik som førsteutkastet for 70 (81.4%) av oppgavene. For to av oppgavene skyldtes dette feilstavelser eller unøyaktigheter i det første oversettelsesutkastet. Vurderingen til legen og logopedene korrelerte signifikant, $r(85) = .67, p < .001$.

I den andre delprosessen ble testleddene oversatt tilbake fra norsk til nederlandsk av en profesjonell tolk for å forsikre at ordlyden ikke var vesentlig endret i forhold til opprinnelig versjon. Tolkens oversettelse var noe annerledes enn oversettelsen fra legen og logopeden for 20 av de 86 oppgavene (23.3%). For tre av oppgavene skyldtes det at den nederlandske originalversjonen ikke var helt overensstemmende med bildene, og dermed noe mer unøyaktig enn den norske. For eksempel var det nederlandske spørsmålet tilknyttet testledd 10 (Figur 11) ”Waar is de sleutel?” som direkte oversatt til norsk er ”Hvor er nøkkelen?”, mens det tilhørende testleddbildet avbildet flere nøkler.

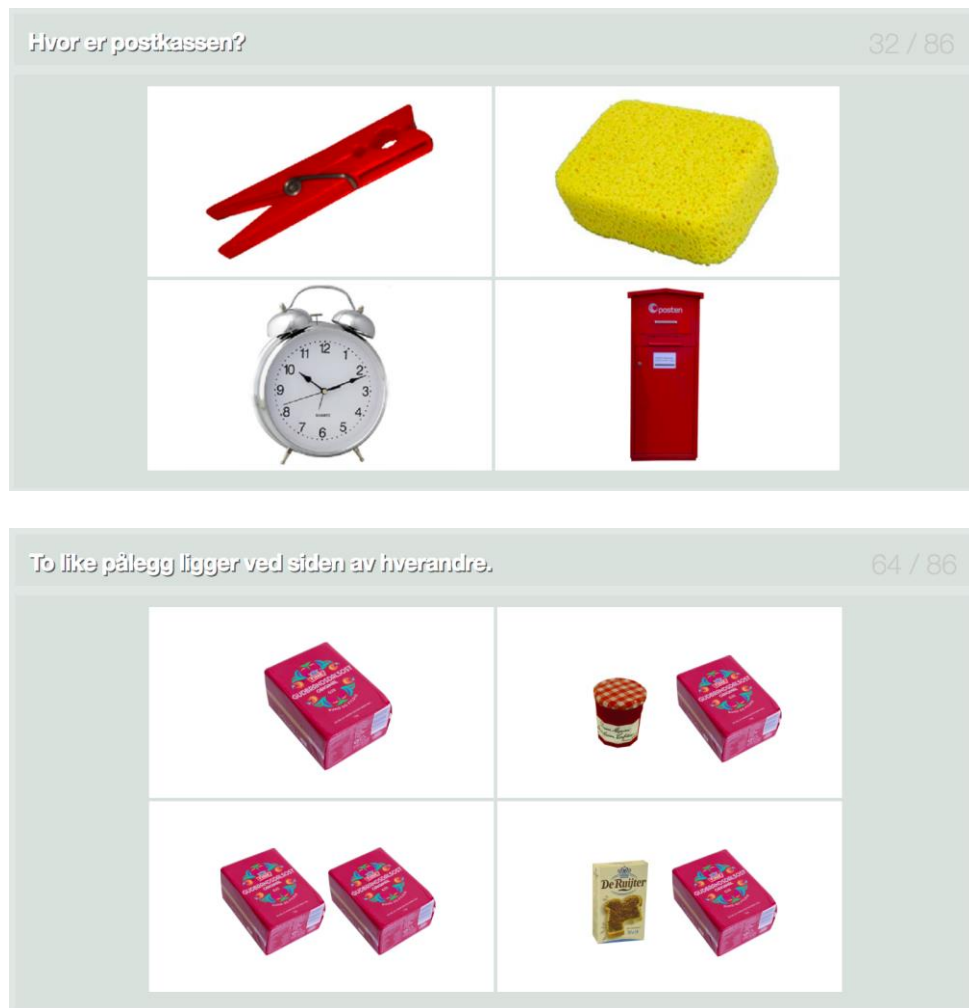


Figur 11. Testledd 10 fra C-BiLLT

I den tredje delprosessen ble det gjort en siste gjennomgang av alle testledd på et konsensumøte. Der deltok to av testens nederlandske utviklere, Geytenbeek og Bootsma, fagpersoner fra Oslo Universitetssykehus som har erfaring med utredning av språk, psykologene Stadskleiv og Heen Andersen og logoped Vøyne, samt masterstudent Haddeland og meg selv. For 53 av 86 oppgaver (61.6%) var det enighet om eksakt ordlyd blant oversettere. For de resterende 33 oppgavene ble ordlyden drøftet blant møtedeltakerne til full enighet ble nådd om en endelig versjon. For to oppgaver ble løsningen å angi at ordvalget var fritt. Dette gjald ulike ord for huske/disse/ronse, samt valgfrihet med hensyn til bruk av ordene asjett og tallerken.

Under oversettelsesprosessen ble det også besluttet å bytte ut testleddbildene som viste en nederlandsk postkasse og et nederlandsk sjokoladepålegg på boks med henholdsvis en norsk postkasse og brunost da dette er mer kjente objekter for barn vokst opp i norsk kultur (Figur

12). Fotografiene av postkasse og brunost ble brukt etter muntlig samtykke fra Posten Norge og Tine. De resterende originale bildene i C-BiLLT kunne beholdes.



Figur 12. Nye bilder av en postkasse (øverst) og brunost (nederst) i den norske versjonen av C-BiLLT

4.5.2 Testing

Testingen ble gjennomført i perioden januar til februar 2018 av meg selv ($n = 71$), masterstudent Haddeland ($n = 90$) og prosjektleder Stadskleiv ($n = 26$). I forkant av testingen deltok testerne på et todagers kurs om administreringen av C-BiLLT holdt av to av C-BiLLTS testutviklere, Geytenbeek og Bootsma. Testingen ble gjennomført i grupperom på barnets skole eller barnehage der testleder og barnet kunne sitte relativt uforstyrret. På skolene foregikk testingen både i skoletiden og i aktivitetsskoletiden og i barnehagen fra rundt 08:30 til 15:00. Noen av de minste barna hadde en voksen fra barnehagepersonalet med seg for å føle seg trygg.

Barna ble tilfeldig fordelt i to testrekkefølger A og B (Tabell 6). Testrekkefølge A ble administrert på 95 (50.8 %) barn, mens testrekkefølge B ble administrert på 92 (49.2%) barn. For alle barn ble C-BiLLT administrert først for å unngå at læringseffekter skulle påvirke testresultatene. En uavhengig *t*-test viste videre at resultatet på testene ikke ble påvirket av denne testrekkefølgen, ettersom det ikke var signifikante forskjeller ($p > .05$) verken i gjennomsnittlig sumskåre på C-BiLLT eller gjennomsnittlig *z*-skårer for de andre testene, mellom de som hadde fått testrekkefølge A og de som hadde fått testrekkefølge B.

Tabell 6. Testrekkefølge A og B

	Testrekkefølge A	Testrekkefølge B
Kartleggingsøkt 1	C-BiLLT	C-BiLLT
	Reynell	Terningmønster*
	CPM**	TROG-2**
Kartleggingsøkt 2	Ordgjenkjenning*	Reynell
	Terningmønster*	Ordgjenkjenning*
	TROG-2**	CPM**

* Terningmønster og Ordgjenkjenning ble kun administrert for barn over 2;7 år

** CPM og TROG-2 ble kun administrert for barn over 4;1 år

Testingens totale varighet varierte både individuelt mellom barna, samt mellom ulike aldersgrupper. For barna mellom 1;6 og 2;6 år tok testingen rundt 30 minutter, mens for barna mellom 2;7 og 4;0 år tok testingen fra 40 minutter til en time. For de eldste barna over 4;1 år tok testingen rundt en time og 30 minutter til to timer. De eldste barna ble tilbudt en pause halvveis i testingen og for de som ønsket dette ble testingen gjennomført i to økter. Ettersom alle testene som er benyttet i denne studien er standardiserte ble barna gitt identiske instruksjoner fra de ulike testlederne.

I denne studien var det noen forskjeller fra den nederlandske utprøvingen av C-BiLLT hva gjelder administreringen og skåring. *Pretest gjenstander* ble ikke administrert i denne studien da det ville vært praktisk utfordrende å be foreldre om å ta med kjente gjenstander fra barnets hjemmemiljø den dagen barnet skulle kartlegges. *Pretest foto* ble videre kun gjennomført på de yngste barna mellom 1;6 og 2;0 år da det ble vurdert at det kun var i denne gruppen det var behov for å undersøke om barnet var i stand til å kommunisere et valg mellom to fotografier. Det ble også bestemt at Del 1 kun skulle administreres på barna mellom 1;6 og 4;6. Dette fordi de 30 første barna svarte korrekt på alle testledd i Del 1 og fordi det ville

korte ned på testingens varighet for de eldste. De nederlandske stoppkriteriene ble ikke benyttet da vi ikke kunne være sikre på om oppbygningen av vanskelighetsgrad var lik i den norske og nederlandske versjonen av C-BiLLT. For alle barn ble det derfor forsøkt å fortsette ut den påbegynte seksjonen selv om antall feil overskred åtte. Pretesten ble ikke regnet med i sum av råskårer på C-BiLLT, slik at maksimum oppnådde sumskåre på C-BiLLT i denne studien er 86.

4.5.3 Retesting

Undersøkelsen av reliabilitet behandles utfyllende i Haddelands (2018) masteroppgave. For å undersøke test-retest-reliabilitet ble 53 av skolebarna retestet med C-BiLLT to til fire uker etter den første testingen. For å undersøke inter-rater-reliabilitet, ble videre 28 av disse barna testet av en annen testadministrator enn ved førstegangstesting. De 53 skolebarna ble også delt i tre grupper utfra hvilken responsmodalitet de benyttet til å besvare C-BiLLT med ved førstegangstesting og retesting. Dette ble gjort for å undersøke om måten barnet besvarte C-BiLLT på ville påvirke barnets skårer på C-BiLLT. Av de 53 barna brukte 20 barn blikkpeking ved førstegangstesting og fingerpeking ved retesting (BPD1), 17 barn brukte fingerpeking ved førstegangstesting og blikkpeking ved retest (BPD2), mens 16 barn brukte fingerpeking begge gangene (Kun FP). Tabell 7 gir en oversikt over de 53 barna som ble retestet fordelt på responsmodaliteter og testadministratører.

Tabell 7. Fordeling av testadministratører og responsmodaliteter

Testadministrator dag 1	Testadministrator ved retest	Responsmodalitet	Antall
KS	KS	BPD1	11
SF	SF	BPD2	7
		Kun FP	7
		ALH	BPD1
ALH	SF	BPD2	5
		Kun FP	5
		Kun FP	4
Totalt:			53

Note. KS = Kristine Stadskleiv; SF = Sara Fiske; ANH = Anne Lise Haddeland

4.6 Statistikk

4.6.1 Inspisering av data og normalitet

Dataen ble i denne studien analysert ved bruk av den statistiske programvarepakken IBM SPSS versjon 25.0. Et signifikansnivå på .05 ble benyttet da dette er vanlig i samfunnsvitenskapelig forskning (Field, 2016). I forkant av de statistiske analysene ble dataen som var plottet inn i SPSS gjennomgått og sjekket for feil.

En forutsetning for å bruke parametrisk statistikk er en vurdering av normalitet, da store avvik fra normalitet kan skape bias i bruken av parametrisk statistikk. Skjevhet, at skårer hopper seg opp i en av fordelings ender, og kurtosis, at mange eller få skårer befinner seg i fordelings haler, er to måter en fordeling kan avvike fra en normalfordeling. I en perfekt normalfordeling er verdier av skjevhet og kurtosis 0, mens skjevhets- og kurtosisverdier mellom -1 til 1 indikerer små avvik fra en normalfordeling, verdier mellom -2 til 2 indikerer moderate avvik og verdier mellom -3 til 3 indikerer at det er store avvik fra en normalfordeling (Field, 2016). Da det ikke eksisterer standardskårer for den norske versjonen av C-BiLLT ble summen av råskårer undersøkt. Den gjennomsnittlige sumskåren på C-BiLLT ($N = 187$) var 66.4 ($SD = 13.2$). Råskårer kontrollerer ikke for alder, og da flere eldre enn yngre barn deltok i denne studien, er det naturlig å forvente en del spredning, samt opphopning av høye skårer i fordelings høyre ende reflektert som en negativ skjevhetsverdi. Fordelingen sumskårer viste i tråd med dette et lite avvik hva gjelder skjevhet, -1.16 ($SE = 0.12$), og kurtosis, 0.61 ($SE = 0.35$). På grunn av en noe skjev aldersfordeling i denne studiens utvalg ble derfor skjevhet og kurtosis undersøkt per helårsgruppe (Tabell 8).

Tabell 8. Deskriptiv statistikk og skjevhets- og kurtosisverdier for C-BiLLT sumskårer fordelt på helårsgrupper

Aldersgruppe	<i>n</i>	<i>M</i>	<i>SD</i>	Skjevhet	Kurtosis	Sig.*
1;1 – 2;0	4	28.3	0.95	- 0.85	-1.29	.272
2;1 – 3;0	26	45.7	8.45	0.31	- 0.69	.452
3;1 – 4;0	25	57.2	5.69	0.20	0.98	.911
4;1 – 5;0	17	64.1	5.46	- 1.01	2.65	.041*
5;1 – 6;0	22	69.6	3.55	- 0.40	- 0.31	.635
6;1 – 7;0	60	75.1	3.60	- 0.34	- 0.48	.173
7;1 – 8;0	33	77.7	2.90	- 0.60	- 0.76	.134

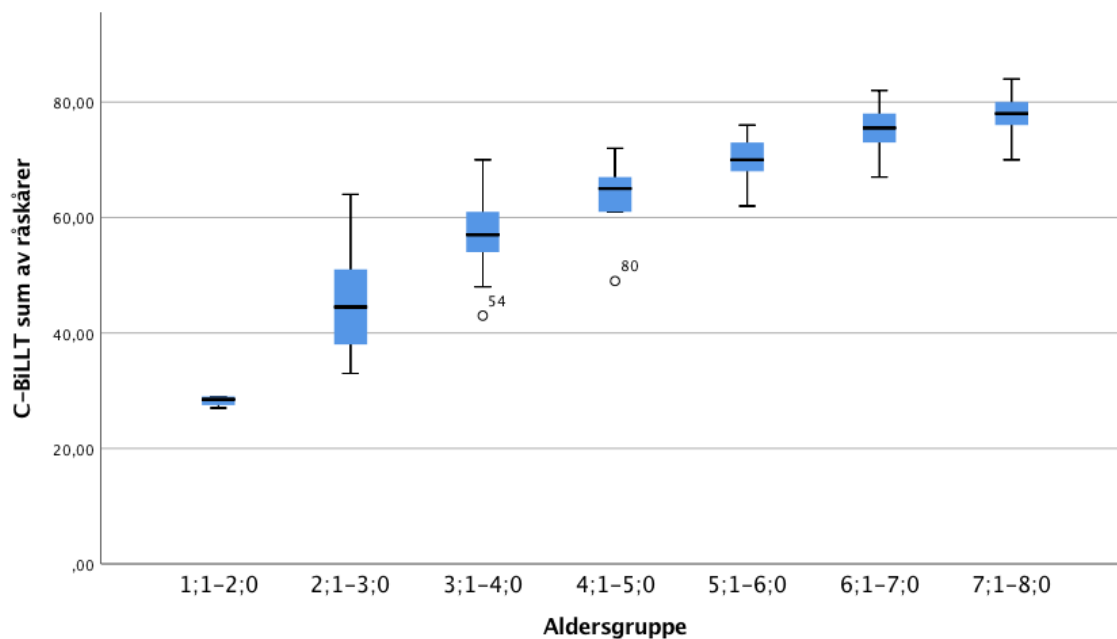
* Shapiro Wilk test

Alle aldersgruppene viste små avvik fra en normalfordeling hva gjelder skjevhet. Kun i helårsgruppen 4;1-5;0 var kurtosisverdien moderat og positiv og fordelingen signifikant forskjell fra en normalfordeling. Denne tabellen kan suppleres av Tabell 9 som gir en oversikt over fordelingen alderskontrollerte z -skårer fra de andre testene som er benyttet i denne studien. Denne tabellen viser også små avvik hva gjelder skjevhet og kurtosis. Inspeksjon av disse tabellene støtter dermed antagelsen om at utvalget ikke er grunnleggende skjevt, annet enn hva gjelder alder, noe som støtter bruken av parametrisk statistikk. Å bruke parametrisk statistikk kan også støttes i størrelsen på utvalget i denne studien. Ifølge sentralgrenseteoremet vil det være hensiktsmessig å bruke parametrisk statistikk fremfor ikke-parametrisk statistikk om utvalgsstørrelsen er >30 ettersom utvalgsfordelingen da vil ha en tilnærmet normalfordeling uavhengig av hvordan populasjonen utvalget er trukket fra er fordelt (Field, 2016).

Tabell 9. Deskriptiv statistikk og skjevhets- og kurtosisverdier for z -skårer

Test	n	M	SD	Skjevhet	Kurtosis
Reynell	189	0.48	0.87	- 0.24	- 0.36
TROG-2	122	0.59	1.07	- 0.60	0.49
Ordgjenkjenning	176	0.18	0.98	0.25	0.45
Terningmønster	176	0.34	0.84	0.74	1.21
CPM	122	0.25	0.94	0.39	- 0.50

Utvalget ble også undersøkt for uteliggere ved visuell inspeksjon av boxplot, da ekstreme verdier kan skape bias i resultatene når parametrisk statistikk benyttes (Field, 2016). Det ble identifisert to uteliggere i sumskåre på C-BiLLT i helårsgruppene. Fjerning av uteliggere krever nøye overveielser som grunner i annet enn at skårene er høyere eller lavere enn gjennomsnittet (De Vaus, 2014). Det er videre rimelig å forvente en del språklig variasjon i utvalget og ingen av skårene var ekstremt avvikende. Fordi ingen av skårene heller viste seg å være feil ved gjennomgang av dataen, ble disse informantene beholdt i analysen. Figur 13 gir en visuell fremstilling av fordelingen sumskårer i halvårsgrupper, med to synlige uteliggere.



Figur 13. Boxplot over C-BiLLT sum av råskårer i ulike helårsgrupper

4.6.2 Valg av hovedanalyser

For å kunne undersøke relasjonen mellom C-BiLLT og andre standardiserte tester ble korrelasjonen mellom sumskåre på C-BiLLT og de andre testene av språkforståelse (Reynell, Ordgjenkjenning og TROG-2) og testene av ikke-verbal resonneringsevne (CPM og Terningmønster) undersøkt. I korrelasjonsanalysen ble råskårer benyttet fordi C-BiLLT enda ikke har z -skårer. For TROG-2 ble både totalt antall blokker med fire rette og totalt antall blokker med tre rette korrelert med sumskåren på C-BiLLT ettersom det ble vurdert at antall blokker med fire rette i større grad kunne være påvirket av tilfeldige feil. Pearsons produkt moment korrelasjonskoeffisient ble benyttet i korrelasjonsanalysen. En korrelasjonskoeffisient gjør det mulig å si noe om relasjonen mellom variabler ved å gi et tall på styrken og retningen på felles variasjon, kovariasjon, mellom variablene. Pearsons produkt moment korrelasjonskoeffisient er et standardisert uttrykk for denne kovariasjonen og gir en verdi mellom -1 og +1, der -1 uttrykker perfekt negativ korrelasjon, +1 uttrykker perfekt positiv korrelasjon og 0 uttrykker ingen korrelasjon mellom de to variablene (Field, 2016). Ifølge Pearson og Hartley (1954) regnes korrelasjoner mellom 0 – .30 som svake, .30 – .60 som moderate, mens korrelasjoner $>.60$ regnes som høye.

For å kunne vurdere C-BiLLTs indre struktur og dimensjonalitet ble en faktoranalyse benyttet. En faktoranalyse bygger på korrelasjoner mellom testledd, men ettersom testleddene

i C-BiLLT er dikotome variabler og noen av testleddene ble mestret av alle deltakerne vil kovariasjonen mellom slike testledd og i neste rekke resultatene av faktoranalysen være villedende og vanskelig å tolke. Fabrigar og Wegener (2012) anbefaler å slå sammen testledd til større sammensatte enheter som en løsning på problemet med dikotome variabler. Derfor ble det bestemt å utføre en faktoranalyse på de 12 seksjonene i C-BiLLT fremfor de 86 testleddene. Det ble videre bestemt å benytte en eksplorerende faktoranalyse fordi den er åpen for hvilken faktorløsning for dataen som er best og dermed egner seg i testutvikling. I forkant av faktoranalysen ble Kaiser-Meyer-Olkin test og Bartlette's test of sphericity utført for å undersøke om mønsteret av korrelasjoner i dataen egnet seg for en faktoranalyse. Bartletts test bør være signifikant og KMO verdien bør være akseptable (.70-.80) eller svært bra (>.90) for å fortsette med faktoranalysen (Field, 2016). Resultatene bekreftet at dataen var egnet for en faktoranalyse og at det var akseptabelt å fortsette analysen, $KMO = .91$ og $X^2(66) = 1711.94, p < .001$. Field (2016) anbefaler videre at korrelasjoner mellom seksjoner bør være mellom .30 og .90 for å unngå at multikollinearitet, at to variabler er lineært og svært nært relatert, påvirker resultatene i faktoranalysen. Da inspeksjon av korrelasjoner mellom seksjoner (Vedlegg 5) innfridde dette ble valget om å gjennomføre faktoranalysen på C-BiLLTs seksjoner ytterligere støttet.

Etter undersøkelse av skjevhets- og kurtosisverdier i alle seksjoner ble det besluttet å bruke Principle Axis Factoring som eksplorerende faktoranalyse-metode ettersom denne er robust mot skjevhets- og kurtosisverdier som overskrider den anbefalte verdien på +/- 2 (Fabrigar & Wegener, 2012). Faktorer ble ekstrahert etter Kaisers kriterie om en eigenvalue >1 i kombinasjon med visuell inspeksjon av scree plot. En Oblique rotasjon (Promax) ble benyttet fordi denne rotasjonsmåten åpner for korrelasjoner mellom faktorer, noe som er teoretisk å forvente i studiet av mentale evner (Fabrigar & Wegener, 2012). Mønstermatrisen vil rapporteres for å gi en oversikt over hver faktors unike påvirkning på en C-BiLLT-seksjon, kontrollert for påvirkningen av andre faktorer. Her vil også ekstraherte kommunaliteter rapporteres. Ettersom faktorene er forventet å korrelere vil også strukturmatrisen rapporteres ettersom denne gir en oversikt over faktorladningene uten å kontrollere for påvirkningen av andre faktorer. Fabrigar og Wegener (2012) rapporterer at kommunaliteter < .40 er lave, mellom .40 og .70 er moderate og >.70 er høye, og at faktorladninger regnes som høye når de er >.50. Når antall deltakere er mellom 100-200 er faktorladninger videre regnet som å gi et meningsfullt bidrag til den overordnede faktorløsningen når de er over .30 (Field, 2016). I denne studien vil derfor faktorladninger <.30 ikke rapporteres. Det er videre anbefalt at hver

faktor lades høyt av minst 3 seksjoner og at differansen mellom en seksjons ladning på flere faktorer er over .20 (Costello & Osborne, 2005).

For å undersøke om C-BiLLT er tilstrekkelig sensitiv til å fange opp forskjeller i språkforståelsesnivå i ulike aldersgrupper ble det foretatt en enveis ANOVA på gjennomsnittlige C-BiLLT sumskårer i halvårsgrupper og helårsgrupper. Selv om ANOVA regnes som en robust test, kan homogen varians skape bias i analysen dersom gruppestørrelsene er ulike, noe som er tilfelle i denne studien. Levenes test viste at variansen i de ulike gruppene var signifikant forskjellig i halvårsgrupper, $F(12, 174) = 2.834, p < .001$, og helårsgrupper, $F(6, 180) = 9.132, p < .001$, og at antagelsen om homogen varians var brutt. Derfor ble en korrigerert versjon av F -ratio, Welch's F , og Games-Howell post hoc analyse, benyttet ettersom disse er robuste selv når gruppestørrelser er ulike og antagelsen om homogen varians er brutt (Field, 2016). For å kunne vurdere tilstedeværelsen av gulv- og takeffekter ble det utformet en tabell med utregninger av deskriptiv statistikk for C-BiLLT-skårer i ulike halvårsgrupper, ettersom gulv- og takeffekter er tenkt å reflekteres i tilstedeværelsen av gjennomsnittlige sumskårer nær henholdsvis minimumskåre eller maksimumskåre samt lave standardavvik som indikerer lite spredning. Tilsvarende deskriptiv statistikk for skårer på Reynell og skårer fra den nederlandske utprøvingen av C-BiLLT vil også rapporteres for å henholdsvis kunne sammenligne skårer fra C-BiLLT med en annen test og med nederlandske C-BiLLT-skårer som kan regnes som mer robuste da antall deltakere er betraktelig flere enn i denne studien.

4.7 Etikk

Et sentralt krav i de forskningsetiske retningslinjene (NESH, 2016) og er at datainnsamlingen ikke settes i gang før deltakerne har gitt sitt frivillige, informerte og uttrykkelige samtykke. I denne studien ble dette kravet overholdt ved at samtykket ble innsamlet som en skriftlig signatur på et samtykkeskjema. På samtykkeskjemaet og tilhørende informasjonsskriv (Vedlegg 1 og 2) var det informasjon om rasjonale, prosedyre, oppbevaring og bruk av data og fordeler og ulemper ved studien samt at samtykket når som helst kunne trekkes tilbake uten å oppgi noen grunn. Deltakerne i denne studien er barn under 15 år og i henhold til NESH (2016) ble samtykket derfor innhentet fra barnas foresatte. Barn omfattes av FNs barnekonvensjon og NESH (2016) fremhever i lys av det at barn som deltar i forskning har et særlig krav på beskyttelse ved at deres interesser og behov ivaretas på andre måter enn hos voksne

forskningsdeltakere. Dette innebærer blant annet at barnet har rett til å bli informert og si sin mening om saker som angår dem (NESH, 2016; Barnekonvensjonen, 2003). Barna som deltok i denne studien ble derfor gitt alderstilpasset informasjon om studien dagen testingen foregikk. De fleste barna var nysgjerrige og spente på hva de skulle få bli med på, mens noen var mer reserverte og trengte mer informasjon og trygghet. Barna som ikke ville eller turte å bli testet ble ikke tvunget til å delta i studien. Under testingen ble det videre tatt hensyn til barnets tilstand og barnet ble informert om at de når som helst kunne si ifra hvis de ble slitne eller måtte på do slik at de kunne få en pause. Foreldrene ble også informert om at barnet kunne ha en voksen med seg under testingen for å føle seg trygg. Som takk for deltakelsen fikk alle barn et takkekort og etter hver test fikk de en liten pause der de kunne velge ett klistermerke å pynte takkekortet med. Foresatte til barna som ikke ble testet fordi de var borte dagene vi hadde avtalt å være på skolen eller barnehagen eller fordi samtykkeskjemaet ble mottatt etter vi hadde vært på den aktuelle skolen eller barnehagen, fikk tilbud om en ny kartlegging slik at ingen skulle oppleve å ha gitt sitt samtykke uten å bli kontaktet.

I denne studien ble det samlet inn personopplysninger og det stilles da særlige krav til oppbevaring og formidling av data (NESH, 2016; Personopplysningsloven, 2001, §13). Studien ble godkjent av Norsk Senter for Forskningsdata før rekrutteringen begynte (Vedlegg 3). Kravet om konfidensiell og forsvarlig oppbevaring ble overholdt ved at samtykkeskjemaene, der foreldrene hadde oppgitt informasjon om adresse, barnets navn, fødselsdato og helsemessige forhold, som eventuelle koniske sykdommer eller språklige, auditive eller visuelle problemer, ble lagret i en safe på Ullevål sykehus og dataen fra kartleggingen ble lagret på sykehusets forskningsserver. Dataen som kunne identifisere barna ble videre oppbevart adskilt fra testresultatene, og barna ble gitt et ID-nummer som knyttet dem til opplysningene gjennom en navneliste. Kun autorisert personell knyttet til prosjektet hadde tilgang til navnelisten og kunne finne tilbake til barnet. Dataen fra denne studien formidles som sumskårer og det er derfor ikke mulig å identifisere enkeltbarn i resultatene som publiseres. Foreldrene ble i samtykkeskjemaet opplyst om at de hadde rett til innsyn om personopplysninger lagret om dem og at opplysningene vil bli slettet senest 2020.

Ettersom C-BiLLT er databesvart lagres barnets oppnådde skårer digitalt og dette medfører en særlig forsiktighet med tanke på faren for spredning av personlig informasjon på nett. Kun de med sertifisering i C-BiLLT får utdelt et brukernavn og passord med tilgang til en egen

C-BiLLT-brukerprofil. All informasjon på denne brukerprofilen er lagret på en sentralserver i Nederland, og testleder har kun tilgang til informasjonen som er lagret på sin egen brukerprofil. Når et barn skal testes oppretter testleder en testprofil til barnet på sin egen brukerprofil. For å opprette denne skriver testleder inn barnets fødselsdato, som så automatisk omregnes til antall år og måneder ved testdato slik at det ikke er mulig å regne seg tilbake til barnets eksakte fødselsdato. Barnet får også utdelt et syvsifret ID-nummer som blir testleders kodenøkkel. På denne måten er det kun aidentifisert informasjon som lagres digitalt.

I samtykkeskjemaet ble det presisert at foresatte ville få tilbakemelding dersom testresultatene til barnet var i lavområdet, definert som 2 standardavvik under gjennomsnittet, og dermed ga mistanke om at barnet hadde vansker og kunne trenge ekstra oppfølging. Testresultatene til ti barn ga mistanke om mulige vansker. For hver av disse barna ble det skrevet rapporter med informasjon om barnets testresultater. Rapporten ble sendt til barnas foresatte og disse ble også invitert til en konsultasjon med prosjektleder for mer informasjon om barnets skårer før eventuell videre henvisning. Det var imidlertid ingen foreldre som tok kontakt. Det ble også besluttet at det skulle skrives en rapport for barnet dersom barnet oppnådde skårer i høyområdet, definert som 2 standardavvik over gjennomsnittet, ettersom disse barna også kan ha behov for ekstra tiltak. To barn oppnådde slike testresultater, men det ble kun skrevet rapport som ble sendt til foresatte for ett av disse barna ettersom det ikke var oppgitt noen kontaktopplysninger for det andre barnet.

5 Resultater

5.1 Undersøkelse av korrelasjon mellom tester

Korrelasjonsanalysen (Tabell 10) viste at korrelasjonen var høyest mellom C-BiLLT og Reynell. Korrelasjonen var lavere men fortsatt høy mellom C-BiLLT og TROG-2, Ordgjenkjenning, og Terningmønster. Korrelasjonen mellom C-BiLLT og CPM var moderat til høy. Alle korrelasjonene var signifikante ($p < .001$). Korrelasjonen mellom C-BiLLT og antall blokker med fire riktige fra TROG-2 var noe lavere, $r(121) = .67, p < .001$, enn korrelasjonen mellom C-BiLLT og antall blokker med tre rette fra TROG-2.

Tabell 10. Oversikt over korrelasjoner mellom råskårer på ulike tester

		C-BiLLT	Reynell	TROG-2**	Ord- gjenkjenning	Terning- mønster	CPM
Reynell	<i>r</i>	.91*	-				
	<i>n</i>	187	-				
TROG-2**	<i>r</i>	.73*	.71*	-			
	<i>n</i>	122	122	-			
Ord- gjenkjenning	<i>r</i>	.84*	.86*	.57*	-		
	<i>n</i>	176	176	122	-		
Terning- mønster	<i>r</i>	.84*	.79*	.52*	.77*	-	
	<i>n</i>	176	176	122	176	-	
CPM	<i>r</i>	.65*	.60*	.61*	.50*	.68*	-
	<i>n</i>	122	122	122	122	122	-

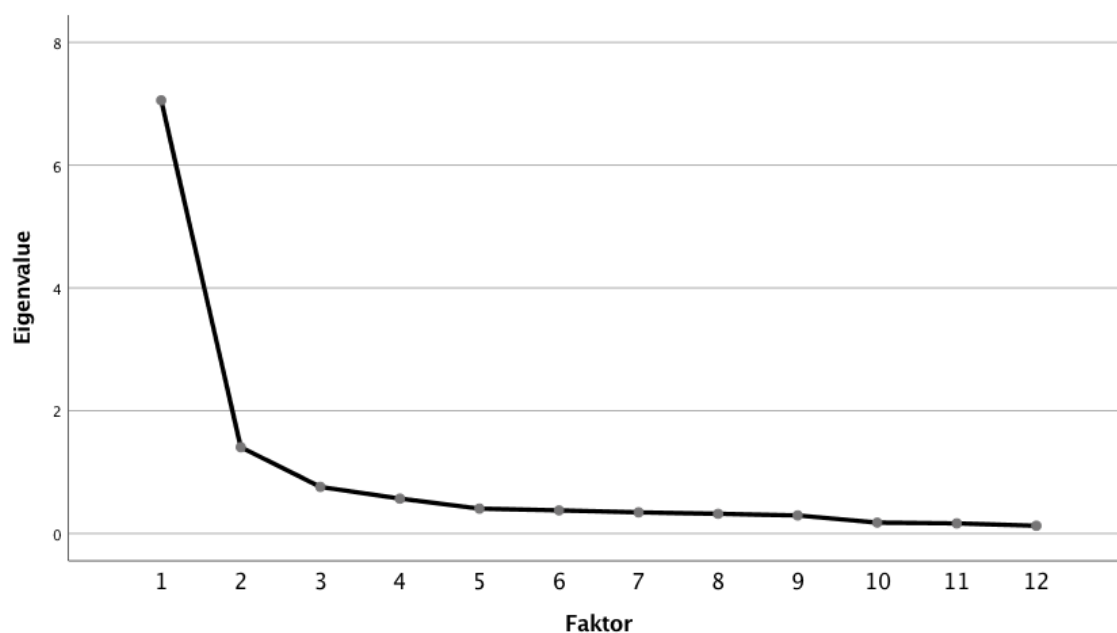
* $p < .001$

** For TROG-2 er sum av antall blokker med tre rette benyttet

Da korrelasjonen mellom sumskårer på C-BiLLT og Terningmønster var uventet høy ble også korrelasjonen mellom testene undersøkt isolert hos barna over 4;1 år. Korrelasjonen mellom sumskårer på C-BiLLT og Terningmønster hos barn mellom 4;1 til 8;0 år var moderat til høy, $r(121) = .64, p < .001$.

5.2 Utforskning av faktorstrukturen til C-BiLLT

Den eksplorative faktoranalyse på seksjonene i C-BiLLT resulterte i at to faktorer med eigenvalues større enn Kaisers kriterie ble ekstrahert. Dette sammenfalt også med visuell analyse av scree plot (Figur 14). Den første faktoren hadde en høy initial eigenvalue på 7.06 og forklarer 58.8% av variansen i dataen. Den andre faktoren hadde en initial eigenvalue på 1.41 og forklarer ytterligere 11.7% av variansen i dataen. Til sammen forklarte faktorene 70.5% av variansen i data.



Figur 14. Scree plot fra den eksplorative faktoranalysen

Tabell 11 og 12 gir en oversikt over faktorladningene etter Oblique rotasjon (Promax). Mønstermatrisen gir en oversikt over hver faktors unike påvirkning på seksjonene i C-BiLLT, kontrollert for påvirkningen av andre faktorer (Tabell 11). For hver faktor var faktorladningene høye for minst 3 seksjoner, slik at faktorene kan tolkes som solide og meningsfulle faktorer. Seksjon 1 til 4 lader kun på Faktor 2 mens Seksjon 7 til 12 lader kun på Faktor 1. Seksjon 5 og 6 lader både på Faktor 1 og Faktor 2 og differansen mellom ladningene er under .20 for begge seksjoner, henholdsvis .09 og .12. Rotasjonen avslørte at korrelasjonen mellom Faktor 1 og Faktor 2 var høy, .69. Strukturmatrisen (Tabell 12) viser videre faktorladningene uten å kontrollere for påvirkningen av andre faktorer. I denne matrisen har alle seksjoner en faktorladning $>.40$ på begge faktorene.

Tabell 11. Mønstermatrise fra den eksplorerende faktoranalysen

C-BiLLT seksjon	Antall testledd	Faktorer		Kommunaliteter
		Faktor 1	Faktor 2	
Seksjon 1 – Substantiv	10	-	.77	.55
Seksjon 2 – Verb	10	-	.85	.66
Seksjon 3 – Dyr, objekter og personer	10	-	.60	.38
Seksjon 4 – Vanskeligere substantiv	4	-	.86	.60
Seksjon 5 – Enkle setninger om gjenstander i kombinasjon med verb og preposisjoner	5	.40	.48	.65
Seksjon 6 – ”Hvem”-spørsmål om personer som utfører handlinger	5	.50	.38	.65
Seksjon 7 – Passive setninger med objekter og preposisjoner kombinert med hendelser	4	.73	-	.56
Seksjon 8 – Setninger med mer grammatisk og syntaktisk kompleksitet	9	.73	-	.84
Seksjon 9 – Enkle aktive setninger som refererer til ikke-observerbare situasjoner med fire personer	6	.84	-	.69
Seksjon 10 – Komplekse setninger med to eller flere begreper	9	.95	-	.85
Seksjon 11 – Sammensatte, komplekse setninger	4	.84	-	.69
Seksjon 12 – sammensatte, komplekse setninger - fortsettelse	10	.91	-	.64
Eigenvalues		7.06	1.42	
% varians forklart		58.8	11.7	

Note. Faktorladninger <.30 er ikke rapportert

Tabell 12. Strukturmatrisen fra den eksplorerende faktoranalysen

C-BiLLT seksjon	Antall testledd	Faktor	
		Faktor 1	Faktor 2
Seksjon 1 – Substantiv	10	.48	.74
Seksjon 2 – Verb	10	.53	.81
Seksjon 3 – Dyr, objekter og personer	10	.44	.61
Seksjon 4 – Vanskeligere substantiv	4	.49	.73
Seksjon 5 – Enkle setninger om gjenstander i kombinasjon med verb og preposisjoner	5	.73	.75
Seksjon 6 – ”Hvem”-spørsmål om personer som utfører handlinger	5	.76	.72
Seksjon 7 – Passive setninger med objekter og preposisjoner kombinert med hendelser	4	.75	.53
Seksjon 8 – Setninger med mer grammatisk og syntaktisk kompleksitet	9	.90	.75
Seksjon 9 – Enkle aktive setninger som refererer til ikke-observerbare situasjoner med fire personer	6	.83	.56
Seksjon 10 – Komplekse setninger med to eller flere begreper	9	.92	.61
Seksjon 11 – Sammensatte, komplekse setninger	4	.83	.56
Seksjon 12 – sammensatte, komplekse setninger - fortsettelse	10	.79	.45

5.3 Testskårer ved ulike aldersgrupper

Resultatet fra enveis ANOVA viste en statistisk signifikant forskjell mellom gjennomsnittlig sum av raskårer på C-BiLLT ved ulike halvårsgrupper, $F(12, 38.68) = 501.72, p < .001$, og helårsgrupper, $F(6, 42.14) = 1075.36, p < .001$. Post hoc analyse med Games-Howell test viste videre en signifikant forskjell i gjennomsnittlig C-BiLLT sum av raskårer mellom de tre første halvårsgruppene (Tabell 13) og mellom alle helårsgruppene (Tabell 14).

Tabell 13. Oversikt over forskjeller i gjennomsnittlig C-BiLLT sumskåre mellom halvårsgrupper

Aldersgrupper	<i>M</i> forskjell*	Sig.	95% konfidensintervall	
			Nedre grense	Øvre grense
1;6 – 2;0 og 2;1 – 2;6	- 9.75	.012**	- 17.46	- 2.30
2;1 – 2;6 og 2;7 – 3;0	- 10.47	.012**	-19.32	- 1.62
2;7 – 3;0 og 3;1 – 3;6	- 7.34	.159	- 16.01	1.32
3;1 – 3;6 og 3;7 – 4;0	- 2.47	.994	- 10.72	5.79
3;7 – 4;0 og 4;1 – 4;6	- 6.17	.176	-13.64	1.30
4;1 – 4;6 og 4;7 – 5;0	0.95	1.000	- 15.26	17.17
4;7 – 5;0 og 5;1 – 5;6	- 5.56	.849	- 21.99	10.88
5;1 – 5;6 og 5;7 – 6;0	- 2.34	.944	- 10.03	5.35
5;7 – 6;0 og 6;1 – 6;6	-3.66	.560	- 11.51	4.18
6;1 – 6;6 og 6;7 – 7;0	- 0.03	1.000	- 3.21	3.27
6;7 – 7;0 og 7;1 – 7;6	- 2.53	.215	- 5.63	0.56
7;1 – 7;6 og 7;7 – 8;0	- 0.83	1.000	- 6.63	4.97

* *M* forskjell = gjennomsnittlig forskjell i skårer mellom halvårsgrupper

** $p < .05$

Tabell 14. Oversikt over forskjeller i gjennomsnittlig C-BiLLT sumskåre mellom helårsgrupper

Aldersgrupper	<i>M</i> forskjell*	Sig.	95% konfidensintervall	
			Nedre grense	Øvre grense
1;1 – 2;0 og 2;1 – 3;0	- 17.40	> .001**	- 25.26	- 9.55
2;1 – 3;0 og 3;1 – 4;0	- 11.55	> .001**	- 15.64	- 7.45
3;1 – 4;0 og 4;1 – 5;0	- 6.91	.006**	- 11.51	- 2.32
4;1 – 5;0 og 5;1 – 6;0	- 5.47	.020**	- 10.19	- 0.75
5;1 – 6;0 og 6;1 – 7;0	- 5.46	> .001**	- 9.10	- 1.82
6;1 – 7;0 og 7;1 – 8;0	- 2.65	.004**	- 5.81	0.52

* *M* forskjell = gjennomsnittlig forskjell i skårer mellom helårsgrupper

** $p < .05$

Tabell 15 viser at gjennomsnittlig sumskåre på C-BiLLT øker med økende halvårsgruppe samt at spredningen i skårer, vist ved standardavvik, blir mindre med alder. Den laveste oppnådde råskåren på C-BiLLT var 27 og den høyeste oppnådde råskåren var 84, og det var dermed ingen som fikk minimumskåre (0) eller maksimumskåre (86). Tabell 16 gir en oversikt over tilsvarende nederlandske data. De eldste halvårsgruppene 7;1-7;6 og 7;7-8;0 er her oppgitt som en helårsgruppe 7;1-8;0, og denne helårsgruppen er derfor også oppgitt i Tabell 15 for å gjøre sammenligningen mellom nederlandsk og norsk data lettere.

Tabell 15. Økning i gjennomsnittlig sumskåre på C-BiLLT i halvårsgrupper

Aldersgruppe	<i>n</i>	<i>M</i>	<i>SD</i>	Minimum	Maksimum
1;7 – 2;0	4	28.3	0.96	27	29
2;1 – 2;6	7	38.0	4.08	33	45
2;7 – 3;0	19	48.5	7.90	34	64
3;1 – 3;6	11	55.8	5.19	48	66
3;7 – 4;0	14	58.3	6.00	43	70
4;1 – 4;6	11	64.5	4.11	61	71
4;7 – 5;0	6	63.5	7.82	49	72
5;1 – 5;6	17	69.1	3.56	62	74
5;7 – 6;0	5	71.4	3.21	68	76
6;1 – 6;6	32	75.1	3.54	68	81
6;7 – 7;0	28	75.0	3.69	67	82
7;1 – 7;6	28	77.6	2.95	70	84
7;7 – 8;0	5	78.4	2.41	75	81
7;1 – 8;0	33	77.7	2.85	70	84

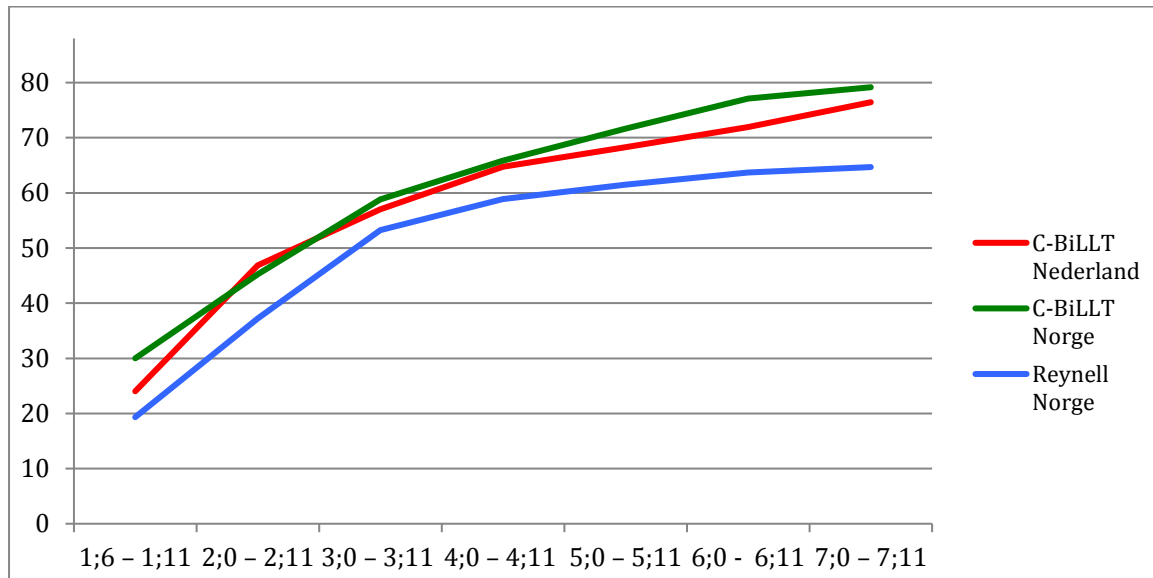
Note. Maksimum oppnådde råskåre = 86

Tabell 16. Økning i gjennomsnittlig sumskåre på C-BiLLT fra den nederlandske utprøvingen av C-BiLLT

Aldersgruppe	<i>n</i>	<i>M</i>	<i>SD</i>	Minimum	Maksimum
1;6 – 1;11	48	24.0	11.14	2	46
2;0 – 2;6	74	43.1	10.17	11	61
2;7 – 3;0	100	49.6	6.21	33	64
3;1 – 3;6	100	55.9	5.80	39	67
3;7 – 4;0	65	60.3	4.90	48	71
4;1 – 4;6	103	63.4	4.40	50	72
4;7 – 5;0	117	65.9	4.72	58	78
5;1 – 5;6	132	67.3	4.75	55	78
5;7 – 6;0	116	69.4	4.08	57	77
6;1 – 6;6	84	71.1	3.89	57	80
6;7 – 7;0	46	73.4	3.82	64	82
7;1 – 8;0	50	76.4	4.06	69	86

Note. Maksimum oppnådde råskåre = 88

Figur 15 gir en visuell fremstilling over tall fra denne studien på økning i gjennomsnittlig sumskåre på C-BiLLT og Reynell, og tall på økning i gjennomsnittlig sumskåre fra den nederlandske utprøvingen av C-BiLLT. For å unngå at de norske og nederlandske C-BiLLT-sumskårene er systematisk forskjellig, grunnet at det i denne studien ikke har blitt regnet med de to poengene fra pretesten i sumskåren, er de to poengene fra pretesten regnet med i hver deltakers sumskåre også for den norske C-BiLLT-sumskåren i dette linjediagrammet.



Figur 15. Linjediagram over fordeling av gjennomsnittlig sumskårer i helårsgrupper

Note. Maksimum oppnådd råskåre på C-BiLLT Nederland = 88; Maksimum oppnådd råskåre på C-BiLLT Norge = 88; Maksimum oppnådd råskåre på Reynell = 67.

6 Diskusjon

Barn med store tale- og bevegelsesvansker utgjør en heterogen gruppe med variert kognitivt og kommunikativt funksjonsnivå. Alle har behov for tilpassede intervensjoner og ASK tidlig i sin utvikling for å få tilgang til læring, utvikling og bli aktive deltakere i sitt miljø (Light et al., 2002; Murray & Goldbart, 2009; Batorowicz et al., 2016). For at tiltak skal tilpasses den enkeltes behov bør de bygge på en kartlegging av blant annet barnets språkforståelse. At gjennomgang av eksisterende normerte språkforståelsestester indikerer at det er en mangel på norske språkforståelsestester som dekker aldersspennet 1;6 til 7;6 år og som er tilpasset barn med store tale- og bevegelsesvansker, fremmer behovet for en ny språkforståelsestest. Om C-BiLLT kan dekke dette behovet avhenger av hvorvidt teori og empiri støtter den intenderte tolkningen og bruken av C-BiLLT-skårer (AERA et al., 2014). I dette kapittelet vil problemstillingen *Hvor egnet er den norske versjonen av C-BiLLT til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år?* diskuteres i lys av teori og empiriske funn. De tre forventningene beskrevet innledningsvis vil danne grunnlaget for diskusjonen. Kapittelet vil avsluttes med en diskusjon av styrker og svakheter ved denne studien og funnenes kliniske implikasjoner. Først vil imidlertid en presentasjon av hovedfunnene fra reliabilitetsundersøkelsen presenteres, da en tests reliabilitet er en forutsetning for validitet (AERA et al., 2014).

6.1 Hovedfunnene fra reliabilitetsundersøkelsen

Haddelands (2018) undersøkelse av reliabilitet ga gode resultater. Indre konsistens for C-BiLLT var høy og ble av Haddeland drøftet opp mot gjennomsnittlig riktig svarprosent for hvert testledd. Oversikten over C-BiLLTs testledd med tilhørende gjennomsnittlig riktig svarprosent er i denne oppgaven lagt ved som Vedlegg 6. Det ble videre påvist at C-BiLLT har god inter-rater-reliabilitet, men at korrelasjonen mellom test og retest var noe lav. Dette grunnet mest sannsynlig i at tidsintervallet mellom test og retest for noen barn ble for kort og at det dermed oppstod en læringseffekt. Resultatene viser i denne sammenheng at korrelasjonen mellom test og retest var større hos de barna som ble retestet en måned etter førstegangstesting enn de barna som ble retestet to uker etter førstegangstesting, noe som gir evidens for at test-retest-reliabiliteten til C-BiLLT er god dersom tidsintervallet mellom test og retest økes. Resultatene fra Haddelands (2018) undersøkelse viste videre ingen signifikant forskjell mellom oppnådd sumskåre på C-BiLLT ved bruk av de ulike responsmodalitetene

fingerpeking eller blikkpeking. Dette indikerer at C-BiLLT kan besvares med ulike responsmodaliteter uten at dette vil ha en innvirkning på barnets skåre. Funnene fra Haddelands (2018) undersøkelse av reliabilitet gir dermed belegg for å begynne diskusjonen av validiteten til de slutningene som kan trekkes på bakgrunn av den norske versjonen av C-BiLLT.

6.2 Evidens basert på relasjoner til andre variabler

6.2.1 Konvergerende evidens

Når et konstrukt som språkforståelse defineres fremhever som nevnt Cronbach og Meehl (1955) at konstruktet plasseres i et nomologisk nettverk der forventede relasjoner mellom konstruktet og andre konstrukt beskrives. Videre presiserer de at en valideringsprosess er mulig når slike påstander angående det nomologiske nettverket fører til forventede relasjoner mellom variabler som er observerbare. AERA, APA og NCME (2014) beskriver i tråd med dette at en viktig støtte til tolkningene og bruken av en test kommer fra empirisk evidens på testens relasjon til variabler som er eksterne for testen. Den første forventningen i innledningen beskriver i denne sammenhengen forventningen om at den observerbare relasjonen mellom variabelen *skårer på C-BiLLT* og de eksterne variablene *skårer på tester av språkforståelse* og *skårer på tester av ikke-verbal resonnering* oppfører seg på teoretisk forventede måter:

- 1) Det er forventet en høyere korrelasjon mellom skårene på C-BiLLT og skårer på andre tester av språkforståelse enn mellom skårene på C-BiLLT og skårer på tester av ikke-verbal resonnering.

C-BiLLT, Reynell, Ordgjenkjenning og TROG-2 er alle utformet for å måle samme konstrukt, språkforståelse. At alle korrelasjonskoeffisientene mellom sumskårer på C-BiLLT og de andre testene av språkforståelse var høye, indikerer sterke positive relasjoner mellom skårene og innfrir den første delen av den ovennevnte forventningen. Slike korrelasjoner kan tolkes som at alle testene måler ferdigheter som testes med alle tester, noe som støtter antagelsen om at C-BiLLT måler språkforståelse. Alle korrelasjonene var videre signifikante, noe som indikerer at det er usannsynlig at korrelasjonene hadde blitt så store i utvalget dersom det ikke skulle være en tilsvarende effekt i populasjonen, og at man dermed kan ha tillit til at de observerte relasjonene er statistisk meningsfulle (Field, 2016). Validiteten til

tolkningen og bruken av C-BiLLT-skårer støttes dermed av tolkningen av konvergerende evidens. Disse funnene er videre også i tråd med studier som viser at språklige komponenter som grammatisk forståelse og forståelse av ords semantiske innhold, eller vokabularforståelse, er positivt korrelert og dermed nært sammenhengende ferdigheter hos barn (Colledge et al., 2002; Tomblin & Zhang, 2006; LARRC et al., 2018). Selv om alle korrelasjonen mellom sumskåre på C-BiLLT og de andre testene av språkforståelse var høy, var det imidlertid også forskjeller i styrken på korrelasjonskoeffisienten mellom C-BiLLT og de ulike språkforståelsestestene. AERA, APA og NCME (2014) fremhever at slike forskjeller kan være viktig å undersøke for å nyansere forståelsen av relasjonen mellom de underliggende konstruktene som måles av testene og tolkningen av testskårer, samt relasjonen mellom innholdet i testen og konstruktet testen er utformet for å måle.

En mulig årsak til forskjeller i korrelasjoner kan være at alle testene tapper samme overordnede konstrukt, språkforståelse, men at testene vektlegger ulike fasetter ved dette konstruktet i operasjonaliseringsprosessen. Utvelging av attributtene som skal dekke innholdsdomenet testene er utformet for å måle, prosessen Benson (1998) beskriver som å trekke forbindelseslinjer mellom det teoretiske og empiriske plan, vil dermed se noe annerledes ut for de ulike testene og dermed resultere i en noe ulik måling av samme konstrukt. Den høyeste korrelasjonen var mellom skårene på C-BiLLT og skårene på Reynell, noe som også var tilfelle i den nederlandske utprøvingen av C-BiLLT. Reynell er i likhet med C-BiLLT utformet for å kartlegge språkforståelse både gjennom vokabularforståelse og forståelse av grammatikk. Ordgjenkjenning og TROG-2 er på den andre siden utformet for å kartlegge henholdsvis kun reseptivt vokabular og kun reseptiv grammatikk (Wechsler, 2015; Bishop, 2009). At skårene på C-BiLLT samvarierer mer med skårene på Reynell enn med skårene på Ordgjenkjenning og TROG-2 er i denne sammenhengen ikke uventet ettersom C-BiLLT og Reynell deler flere av de samme attributtene og dermed dekker mer sammenfallende sider av det teoretiske konstruktet.

Det må imidlertid også tas med i betraktning at den høye korrelasjonen mellom C-BiLLT og Reynell er naturlig i lys av at C-BiLLT bygger på Reynell i sin oppbygning av lingvistisk kompleksitet. Presisering av innholdsdomenet og utvelging av attributter til C-BiLLT er dermed sammenfallende med Reynell ikke bare fordi de er utformet for å måle det samme, men også fordi C-BiLLT bygger på Reynell og dermed vil ha noe lignende testledd. For eksempel er testledd 37: *Hvor er den lengste røde blyanten?* og testledd 67: *Hvem skal gå på*

skolen senere, men gjør det ikke nå? fra Reynell lignende testledd 51: *Hvor er det lengste røde toget?* og testledd 60: *Hvem kommer til å leke ute senere, men gjør det ikke ennå?* Fra C-BiLLT. Price (2017) fremhever i denne sammenhengen at kriteriesmitte oppstår når målet som brukes som kriterie består av noen av de samme testleddene som testen som studeres. Dette kan være problematisk i valideringsprosessen da de to målene ikke er uavhengige av hverandre. Price (2017) fremhever videre at en test som brukes som kriterie bør ha påvist god validitet og reliabilitet. Som nevnt kan det knyttes usikkerhet rundt validiteten til de slutningene som kan trekkes på bakgrunn av den norske versjonen av Reynell grunnet få nye undersøkelser av validitet. Tilstedeværelsen av konstruktunderrepresentasjon eller konstruktirrelevant varians i Reynell er dermed usikker. I tillegg strekker normene til Reynell seg kun til seksårsalder og dekker dermed ikke hele aldersspennet i denne studien, noe som gjør det rimelig å forvente takeffekter for de eldste aldersgruppene og at variasjon i språkforståelse ikke fanges opp like godt av Reynell for denne gruppen. Det kan dermed være problematisk å bruke Reynell som et kriterie i undersøkelsen av konvergerende evidens. At skårer på C-BiLLT også korrelerer høyt med de andre testene av språkforståelse er dermed en viktig validitetsstøtte.

En annen mulighet er at forskjellen på korrelasjonene mellom de ulike språkforståelsestestene og C-BiLLT er et bilde på aldersforskjellen mellom barna som er testet med de ulike testene. Som nevnt indikerer flere studier at flerdimensjonalitet i språkevnen er en fremkommende snarere enn medfødt kvalitet ved språkssystemet, noe som reflekteres i en nedgang i korrelasjonen mellom vokabular og grammatikk etter hvert som barnet blir eldre (Tomblin & Zhang, 2006; Lyster et al., 2010; LARRC et al., 2018). TROG-2 har kun normer fra 4;0 år og ble dermed ikke administrert på barna under 4;0 år, noe som betyr at Ordgjenkjenning i tillegg ble administrert på de 44 barna fra 2;6 til 4;1 år, og C-BiLLT og Reynell i tillegg ble administrert på de 55 barna fra 1;6 til 4;1 år. Grunnen til at korrelasjonen mellom skårer på C-BiLLT og skårer på Reynell og Ordgjenkjenning er høyere enn korrelasjonen mellom skårer på C-BiLLT og skårer på TROG-2, kan dermed skyldes at C-BiLLT, Reynell og Ordgjenkjenning deler mer av variasjonen i en helhetlig språklig evne hos de yngste barna.

I tillegg til utvelging av testledds innhold, hva det skal spørres etter i testen, innebærer operasjonaliseringen av et teoretisk konstrukt også utvelging av hvordan attributtene skal måles for eksempel gjennom ordlegging, testleddformat og svarmåte (Nunnally & Bernstein, 1994). Campell og Fiske (1959) fremhever i sammenheng med dette at det i tolkningen av

konvergerende evidens også må tas i betraktning at systematisk variasjon i testskårer kan skyldes responser til hvordan attributtene måles i tillegg til responser til testinnhold.

C-BiLLT, Ordgjenkjenning og TROG-2 måler språkforståelse ved at barnet skal velge ett av bildene i en bildematrise som tilsvarer det testleder sier. Reynell skiller seg ut da testen måler språkforståelse ved at barnet skal utføre en handling med ulike objekter som tilsvarer det testleder sier. At korrelasjonen mellom skårer på C-BiLLT og Reynell likevel er høyest kan dermed tolkes som en støtte til at variasjonen i skårer i liten grad påvirkes av svarmåte.

På den andre siden kan TROG-2 beskrives som å skille seg ut fra de andre testene ved at denne er mer komplisert, da den til forskjell fra de andre testene er utformet for eldre barn fra 4;0 til 16;5 år. En god del barn uttrykte også at de opplevde denne testen som krevende, kjedelig og lang, og flere måtte minnes på å høre ferdig hele setningen før de pekte på et bilde. Til forskjell fra dette kan det hende at C-BiLLT og Reynell oppleves mer oppmerksomhetsfangende og motiverende ved at de utnytter elementer barn ofte assosierer med lek og spill, som berøringsskjermen i C-BiLLT og forskjellige konkreter i Reynell. Korrelasjonskoeffisienten mellom skårer på TROG-2 og Ordgjenkjenning og TROG-2 og Reynell var videre enda lavere enn korrelasjonen med skårer på C-BiLLT. En mulig årsak til at korrelasjonen mellom skårene på C-BiLLT og TROG-2 var den laveste av språkforståelsestestene kan dermed også skyldes at variasjon i skårer på TROG-2 i større grad enn C-BiLLT, Reynell og Ordgjenkjenning er påvirket av lav motivasjon og varierende oppmerksomhet. Dette kan også være en mulig forklaring på at korrelasjonen mellom sumskårer på C-BiLLT og antall blokker med fire rette på TROG-2 var lavere enn mellom sumskårer på C-BiLLT og antall blokker med tre rette på TROG-2. Dersom skårene på TROG-2 ble påvirket av varierende oppmerksomhet vil dette få større konsekvenser for antall blokker med fire rette, der barnet må ha alle fire spørsmål korrekt for å få ett poeng, enn antall blokker med tre rette, der barnet kan ha ett spørsmål feil men likevel få ett poeng for de tre resterende mestrede spørsmålene i blokken. En mulighet er dermed at totalt antall mestrede blokker med tre rette på TROG-2 i denne studien gir et riktigere bilde på barnets språkforståelsesnivå enn totalt antall mestrede blokker med fire rette.

6.2.2 Diskriminerende evidens

Selv om korrelasjonene mellom C-BiLLT og de andre testene av språkforståelse indikerte at testene tapper det samme konstruktet, vil en slik korrelasjon si mindre om hvorvidt variasjon

i skårer på C-BiLLT også kan skyldes variasjon i andre evner enn språkforståelse. I definisjonen av teoretiske forventninger om relasjonen mellom testen og andre eksterne variabler er det også forventninger om hvilke eksterne variabler som ikke vil være relatert til testen. Campbell og Fiske (1959) fremhever videre at konvergerende evidens alltid må suppleres av diskriminerende evidens. I denne sammenhengen forventes det at skårene på C-BiLLT korrelerer mindre med skårene på ikke-verbale tester, enn med skårene på tester av språkforståelse, ettersom de ikke-verbale testene er utformet for å måle konstrukter som teoretisk er forventet å være mindre relatert til konstruktet språkforståelse. At korrelasjonen mellom skårer på C-BiLLT og CPM var mer moderat enn korrelasjonen mellom skårer på C-BiLLT og testene av språkforståelse gir dermed støtte til denne forventningen og at C-BiLLT måler språkforståelse og ikke ikke-verbale resonneringsevner. Korrelasjonen mellom skårer på C-BiLLT og CPM var høyere enn korrelasjonen mellom den nederlandske versjonen av C-BiLLT og CPM. Noe korrelasjon mellom tester av språk og tester av ikke-verbal resonnering er imidlertid å forvente utfra synet om at kognitive ferdigheter ikke er urelaterte modulære evner (Karmiloff-Smith, 2009), og at prestasjoner på språkforståelsesoppgaver også vil påvirkes av generelle kognitive prosesser som oppmerksomhet og arbeidsminne (Bishop, 2014; Norbury et al., 2002).

Det var imidlertid uventet at korrelasjonen mellom skårene på C-BiLLT og Terningmønster var såpass høy, og dette kan ikke kun forklares med at språk og ikke-verbal resonnering er relaterte kognitive ferdigheter. Det er heller ikke mulig å sammenligne denne korrelasjonen med nederlandske data da korrelasjonen mellom andre tester av ikke-verbal resonneringsevne enn CPM ikke ble undersøkt for den nederlandske utprøvingen av C-BiLLT. Dersom variasjonen i skårer på C-BiLLT også er påvirket av barns ikke-verbale resonneringsevne kan dette bidra til konstruktirrelevant variasjon, noe som svekker validiteten til tolkningene og bruken av testskårer som et rent mål på barns språkforståelse. Et viktig skille mellom Terningmønster og CPM er imidlertid at CPM kun ble administrert på barna over 4;1 år, mens Terningmønster i tillegg ble administrert på barna fra 2;7 år. Den høye korrelasjonen mellom C-BiLLT og Terningmønster kan derfor forklares med at kognitive ferdigheter er mindre differensiert hos yngre barn og at en språklig spesialisering er en fremkommende snarere enn medfødt egenskap (Karmiloff-Smith, 2009; Tideman & Gustafsson, 2011). I lys av dette er det å forvente at variasjonen i skårene hos de yngre i mindre grad enn hos de eldre barna vil skille seg fra hverandre, noe som vil reflekteres i en høyere korrelasjon mellom skårene på tester av språk og ikke-verbal resonneringsevne. En slik antagelse kan også støttes

ved en nærmere inspeksjon av korrelasjonsmatrisen (Tabell 10) som viser at korrelasjonen mellom skårer på Terningmønster og språkforståelsestestene som også ble administrert på de minste barna er høy, mens korrelasjonen mellom skårer på Terningmønster og TROG-2, som kun ble administrert på de eldste barna, er moderat. Når barna under 4;1 år utelates fra korrelasjonsanalysen blir videre korrelasjonen mellom skårer på C-BiLLT og Terningmønster tilsvarende den mellom C-BiLLT og CPM. I lys av dette er en rimelig konklusjon at den høye korrelasjonen mellom skårene på C-BiLLT og Terningmønster ikke skyldes at C-BiLLT også måler ikke-verbal resonneringsevne, men er en refleksjon av at språk og andre kognitive ferdigheter er mindre differensiert hos yngre barn.

6.3 Evidens basert på indre struktur

6.3.1 Endimensjonalitet eller flerdimensjonalitet?

Selv om konvergerende og diskriminerende evidens støtter den første forventningen og tolkningen av C-BiLLT-skårer som et mål på barns språkforståelse, fremhever AERA, APA og NCME (2014) at ytterligere støtte til validitet komme fra at testens indre struktur og faktorløsning er i tråd med teoretiske forventninger om konstruktet testen er utformet for å måle. I denne sammenhengen er følgende forventning beskrevet:

- 2) Det er forventet at en eksplorerende faktoranalyse på C-BiLLT resulterer i samme faktorløsning som for den nederlandske versjonen av C-BiLLT; en endimensjonal faktorløsning som representerer en overordnet språkforståelsesevne.

Flere studier indikerer at språkevnen hos barn kan representeres av én latent variabel eller en endimensjonal faktorløsning (Colledge et al., 2002; Tomblin & Zhang, 2006; Anthony et al., 2014; Klem et al., 2014; LARRC, 2015). I lys av dette var det uventet at resultatet fra den eksplorerende faktoranalysen i denne studien resulterte i to ekstraherte faktorer, og ikke én. En tests dimensjonalitet har stor innvirkning på tolkningen av testskårer. Når seksjoner i en test summeres til én sumskåre, som i C-BiLLT, er en implisitt antagelse at en slik test måler én dimensjon, ved at det er én dominant latent variabel som forklarer korrelasjonene i testledd, og at denne dimensjonen måles langs et kontinuum (Slocum-Gori & Zumbo, 2011). Dersom testen i realiteten er flerdimensjonal, men påtvinges en endimensjonal form og tolkes på en endimensjonal måte, kan dette føre til konstruktunderrepresentasjon, ved at sider ved konstruktet ikke måles, noe som truer validiteten til tolkningene og bruken av testskårer

(Messic, 1995). Videre vil tolkningen av en slik sumskåre rent praktisk kompliseres ettersom sumskåren da vil påvirkes på uvis måte av variasjoner i to konstrukter og ikke ett.

Et viktig poeng er imidlertid at selv om en endimensjonal faktorløsning generelt foretrekkes, er det i virkeligheten sjeldent at det konstrueres rene endimensjonale tester ettersom det som måles i samfunnsvitenskapen ofte er komplekse og sammensatte fenomener som ikke lar seg måle som én attributt (Nunnally & Bernstein, 1994; Friborg, 2010). Tester som har ett dominerende latenstrekk supplert av en annen mindre faktor er dermed akseptert, og betegnes som essensiell endimensjonalitet (Slocum-Gori & Zumbo, 2010). Det viktigste blir i denne sammenhengen at faktorene som ekstraheres, og den sammensatte faktorløsningen, er tolkbar. Det blir dermed viktig å diskutere hvordan dimensjonaliteten og faktorene ekstrahert fra C-BiLLT kan forstås.

Å undersøke hva som kjennetegner og skiller de seksjonene som bidrar til Faktor 1 og de seksjonene som bidrar til Faktor 2 i mønstermatrisen (Tabell 11) kan gi en indikasjon på hvilke evner de to ekstraherte faktorene representerer. De seksjonene som kun lader på Faktor 2 er seksjonene 1 til 4. Disse seksjonene er tidligere beskrevet som å kartlegge vokabularforståelse, og barnet kan dermed mestre oppgavene i disse seksjonene kun ved å reagere på enkeltord i setningen. Barnet får i disse seksjonene høre substantivfraser og verbfraser og skal matche disse til bilder av det tilsvarende objektet, dyret, mennesket eller handlingen. Slike oppgaver vil i henhold til Bishops (2014) modell stille krav til barnets evne til matche sekvenser av lyder i talestrømmen med mentale representasjoner av ords fonologiske form og semantiske innhold i barnets mentale leksikon. En tolkning av Faktor 2 kan dermed være at denne representerer barns forståelse av ord; deres vokabularforståelse. De seksjonene som kun lader på Faktor 1 er Seksjon 7 til 12. Disse skiller seg fra Seksjon 1 til 4 ettersom de er utformet for å kartlegge forståelse av mer syntaktisk komplekse setninger. For eksempel i testledd 48 i Seksjon 7 er setningen *Jenta blir dyttet av mannen*. Her kan barnet verken basere seg på forståelse av enkeltord i setningen, på kunnskap om hva det mest sannsynlige scenarioet er eller den vanlige subjekt-verbal-objekt-strukturen i norsk språk, for å forstå setningens innhold og matche dette til riktig bilde. Mestringen av slike oppgaver stiller større krav til integrering av de ulike nivåene i språkforståelse ved at barnet i tillegg til forståelsen av enkeltord må utlede meningen som ligger i kombinasjonen og rekkefølgen av setningsledd (Bishop, 2014; Lust, 2006). En tolkning av Faktor 1 kan dermed være at denne representerer barns forståelse av setninger; deres grammatiske forståelse.

Selv om Faktor 2 reflekterer vokabularforståelse og Faktor 1 representerer grammatisk forståelse er dette fortsatt ingen forklaring på hvorfor faktoranalysen resulterte i to ekstraherte faktorer og ikke én generell faktor som kan betegnes som å representere barns generelle språkforståelsesevne. En mulig tolkning kan være at en slik faktorløsning representerer hvordan ulike språkdomener reflekterer ulike underliggende systemer. Ut fra nativistiske teorier vil en slik faktorløsning være å forvente ettersom vokabular og grammatikk representerer separate modulære systemer hvis utvikling og funksjon styres av ulike underliggende mentale mekanismer (Pinker, 1998; Chomsky, 1959). Variasjonen i barns forståelse av ord, og korrelasjonen mellom seksjoner utformet for å måle vokabularforståelse, vil dermed være uavhengig av variasjonen i barns forståelse av grammatikk, og korrelasjonen mellom seksjoner utformet for å måle grammatisk forståelse. Dersom en slik tolkning aksepteres kan det være problematisk å kombinere skårer fra C-BiLLT-seksjonene som lader på ulike faktorer til ett mål, da det i realiteten er to separate evner som måles. Før en slik tolkning aksepteres må imidlertid faktorladningenes egenskaper i mønstermatrisen tas med i betraktning. Faktorladningene i mønstermatrisen representerer regresjonskoeffisienter som belyser hvordan endringer i avhengig variabel bidrar til endringer i uavhengig variabel, mens andre uavhengige variabler holdes konstant (Field, 2016). Faktorladningene i mønstermatrisen gir dermed et bilde på hvordan endringer i én faktor bidrar til endringer i seksjoner, mens det kontrolleres for påvirkningen av andre faktorer. Det blir da mulig å se et klart og usmittet bilde av relasjonene mellom to variabler; en seksjon og en faktor, og danne seg en oversikt over unike ladninger av en gitt faktor i en seksjon. En mulig bakside med et slikt mål er at faktorene som ekstraheres kan fremstå som mer separate og uavhengige enn de egentlig er, samt at en faktors ladning i en gitt seksjon fremstår som mer unik enn den egentlig er. I samfunnsvitenskapen er det videre generelt akseptert at ekstraherte faktorer kan korrelere med hverandre ettersom evner er forventet å påvirkes av flere faktorer (Costello & Osborne, 2015). Ut fra synet på at språkforståelse avhenger av en integrering av flere språklige nivåer (Bishop, 2014) vil det i denne sammenhengen være teoretisk å forvente at Faktor 1; grammatisk forståelse, og Faktor 2; vokabularforståelse, korrelerer med hverandre og samtidig påvirker variasjonen i alle seksjoner. Det blir dermed nødvendig å også diskutere dimensjonalitet ut fra strukturmatrisen.

I strukturmatrisen representerer faktorladningene bivarierte korrelasjonskoeffisienter (Field, 2016). Slike mål er nyttige for å etablere styrken og retningen på relasjonen mellom to

variabler. Til forskjell fra en regresjonskoeffisient kontrollerer ikke korrelasjonskoeffisienter for påvirkningen av mulige tredjevariabler, noe som gjør det mulig å si noe om relasjonen mellom en seksjon og en faktor uten at påvirkningen fra andre faktorer kontrolleres for. Ved visuell inspeksjon av strukturmatrisen (Tabell 12) fremgår det at alle seksjoner gir et meningsfullt bidrag til begge faktorene noe som kan sees i sammenheng med at de ekstraherte faktorene korrelerte høyt med hverandre. Faktor 1; grammatisk forståelse, og Faktor 2; vokabularforståelse, kan dermed ikke tolkes som urelaterte latente variabler, men som relaterte evner som begge påvirker variasjonen i alle seksjoner. Fabrigar og Wegener (2012) fremhever videre at når ekstraherte faktorer korrelerer høyt med hverandre kan grunnen til denne relasjonen være at begge faktorer påvirkes av en høyere-ordens faktor. I lys av dette kan den høye korrelasjonen mellom Faktor 1 og Faktor 2 tolkes som at begge faktorene i denne studien representerer distinkte aspekter ved et bredere språkforståelseskonstrukt og at en enhetlig språklig evne bidrar til variasjon i alle seksjoner. Et slikt resultat trenger dermed ikke å tolkes som eksistensen av modulære systemer og separate språkforståelsesevner. Selv om flere studier av språkevns dimensjonalitet konkluderer med at språkevnen til barn under syv år best kan representeres med en endimensjonal modell, beskrev noen av de samme studiene hvordan en endimensjonal modell også kom til uttrykk ved at en tofaktormodell med to høyt korrelerte faktorer som representerer vokabular og grammatikk også passet dataen for alle aldersgrupper (Tomblin & Zhang, 2006) og at variasjonen i data kunne forklares av en flerdimensjonal modell på det underordnede nivået og endimensjonal modell på et overordnet nivå (Klem et al., 2014). I lys av dette vil det derfor ikke være problematisk å tolke sumskåren på C-BiLLT som å reflektere et mål på barnets enhetlige språkforståelsesevne, en evne som både påvirkes av variasjoner i grammatisk forståelse og vokabularforståelse, noe som videre støtter den andre forventningen beskrevet innledningsvis.

Selv om det at to faktorer ble ekstrahert kan begrunnes i at disse to faktorene påvirkes av en overordnet faktor betegnet som en overordnet språkforståelsesevne forklarer ikke dette hvorfor resultatet av den eksplorerende faktoranalysen i denne studien resulterte i to ekstraherte faktorer, mens resultatet av faktoranalysen i den nederlandske studien (Geytenbeek et al., 2014) resulterte i en endimensjonal faktorløsning. En mulig årsak til dette kan være at det norske utvalget har en noe skjev aldersfordeling ved at størstedelen av utvalget er over fem år, mens det nederlandske utvalget har en mer jevn aldersfordeling. I lys av studier som indikerer at språkevnen hos barn vil bli mer differensiert med alder, reflektert

i lavere korrelasjoner mellom mål på grammatikk og vokabular (Hayiou-Thomas et al., 2006; Lyster et al., 2010; LARRC et al., 2018), kan en mulig grunn til forskjellen mellom ekstraherte faktorer i denne studien og den nederlandske studien være at utvalget i denne studien er gjennomsnittlig eldre og at vokabular- og grammatikkferdigheter dermed vil være mer differensiert. En annen forskjell mellom det nederlandske og norske utvalget kan imidlertid også være at de norske deltakerne var mer homogene enn de nederlandske, både fordi det var en skjevere aldersfordeling og fordi det utfra inspeksjon fra gjennomsnittlige z-skårer (Tabell 9) kan virke som om de norske deltakerne lå noe over gjennomsnittet både hva gjelder språklige og ikke-verbale evner. Dersom utvalget i denne studien er homogent kan dette ha ført til lite variasjon i skårer noe som igjen kan ha ført til noe lavere korrelasjoner mellom seksjoner (Fabrigar & Wegener, 2012). En mulighet er dermed at den ekstraherte faktorløsningen i denne studien også er påvirket av utvalget.

6.3.2 Er kryssladninger et problem for testtolkning?

I en faktoranalyse er det generelt ikke ønskelig med kryssladninger i mønstermatrisen, det vil si variabler som lader på mer enn én faktor, ettersom en slik variabel ikke kan sees på som et rent mål på det underliggende konstruktet (Costello & Osborne, 2005). Dette vil imidlertid avhenge av teoretiske forventninger til det underliggende konstruktet. En seksjon som krysslader, også kalt kompleks variabel, kan dermed beholdes dersom det kan argumenteres for at dette er i tråd med det underliggende konstruktets natur (Fabrigar & Wegener, 2012). Resultatet fra faktoranalysen i denne studien viste at to seksjoner i C-BiLLT, Seksjon 5 og Seksjon 6, ladet på begge ekstraherte faktorer og at forskjellen mellom ladningene på faktorene var under den foretrukne grensen på .20 for begge seksjoner. Disse seksjonene skilte seg også ut ved å ha lavere faktorladninger enn de andre seksjonene. Ved å se nærmere på hva som kjennetegner testleddene i disse seksjonene kan en slike kryssladninger og faktorladninger forstås. Nedenfor presenteres et utklipp fra Vedlegg 6 (Tabell 17) med en oversikt over testleddene i Seksjon 5 og 6.

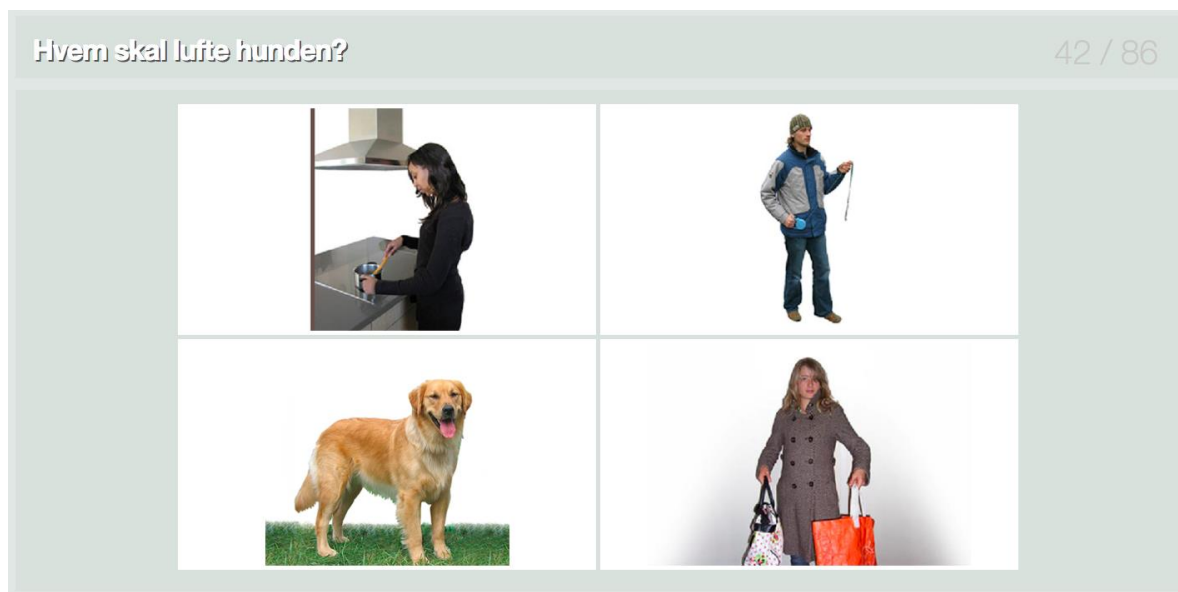
Tabell 17. Utklipp fra Vedlegg 6 – Testledd 35 til 44 fra Seksjon 5 og 6

Seksjon	Testledd	<i>N</i>	<i>M</i> %*	<i>SD</i>
Seksjon 5	35: <i>Hva kan du sove i?</i>	187	94	24.6
	36: <i>Hva kan du kjøre i?</i>	187	91	28.0
	37: <i>Hva kan du tørke med?</i>	187	91	28.0
	38: <i>Hunden ligger i kurven</i>	187	84	36.0
	39: <i>Boka står på asjetten</i>	187	79	40.7
Seksjon 6	40: <i>Hvem spiller musikk?</i>	187	92	27.2
	41: <i>Hvem bærer noe?</i>	187	93	25.5
	42: <i>Hvem skal lufte hunden?</i>	187	69	46.4
	43: <i>Hvem har handlet?</i>	187	83	37.3
	44: <i>Hvem kjører?</i>	187	91	28.0

* *M* % = Gjennomsnittlig riktig svarprosent

Seksjon 5 utgjøres av fem testledd bestående av enkle setninger om gjenstander i kombinasjon med verb og preposisjoner, mens Seksjon 6 utgjøres av fem testledd bestående av "Hvem"-spørsmål om personer som utfører handlinger. Begge disse seksjonene skiller seg fra seksjonene som kun ladet på Faktor 1; grammatisk forståelse, ved at de fleste testleddene kan mestres ved å reagere på enkeltord i setningen, noe som er en vanlig strategi hos yngre barn (Strohner og Nelson, 1974). For eksempel er det mulig for barn som kun mestrer verbene *sove* og *bærer* å velge riktig bilde, henholdsvis en seng og en jente som bærer handleposer, i Seksjon 5 testledd 35: *Hva kan du sove i?* og Seksjon 6 testledd 41: *Hvem bærer noe?*. Årsaken til at barn mestrer slike setninger kan dermed være at de baserer seg på enkeltordene i setningen, sin vokabularforståelse. Variasjon i mestring av disse seksjonene vil dermed delvis skyldes variasjon i Faktor 2; vokabularforståelse. Dette kan også illustreres ved at gjennomsnittlig prosent som har mestret disse setningene er høy og standardavviket lavt. Seksjonene skiller seg videre også fra seksjonene som kun lader på Faktor 2; vokabularforståelse, ved at mestringen av noen testledd i større grad avhenger av syntaktisk forståelse. For eksempel er mestring av testledd 42: *Hvem skal lufte hunden?* i Seksjon 6 i større grad avhengig av mestringen av ordene i kombinasjon, snarere enn enkeltordet *hunden*, for å velge riktig bilde (Figur 16). Videre vil mestring av testledd 39: *Boka står på asjetten* i Seksjon 5 avhenge av at barnet forstår at setningsoppbygningen av ordene *boka* og *asjetten* rundt verbet og preposisjonen *står på* indikerer at det er boka som står på asjetten og ikke omvendt (Figur 17). Gjennomsnittlig prosent som har mestret disse setningene er også lavere og standardavviket høyere sammenlignet med testledd 35 og 41. Det eksisterer dermed en

overlapp mellom vokabularforståelse og grammatisk forståelse i disse seksjonene, noe som kan forklare at de lader på begge faktorer. At disse seksjonene hadde noe lavere faktorladning enn de andre seksjonene i mønstermatrisen kan videre forklares ved at variansen i dem påvirkes av begge seksjoner, noe som også reflekteres i at faktorladningene øker betraktelig i strukturmatrisen.



Figur 16. Testledd 42 fra Seksjon 6 av C-BiLLT



Figur 17. Testledd 39 fra Seksjon 5 av C-BiLLT

Validiteten til slutninger som kan trekkes på bakgrunn av en test må alltid sees i sammenheng med bruken av testen (AERA et al., 2014). Dersom C-BiLLT var utformet for å danne

separate indekser for to relaterte konstrukter som ga informasjon om barns vokabularforståelse og grammatisk forståelse hver for seg, ville slike seksjoner være problematisk for validiteten til slutningene som kan trekkes på bakgrunn av dem ettersom variasjonen i disse seksjonene er avhengig både av Faktor 1; grammatisk forståelse, og Faktor 2; vokabularforståelse. Men ettersom hensikten med C-BiLLT er å si noe om reseptivt språklig funksjonsnivå, ikke vokabular og grammatikk separat, er en slik kryssladning mindre problematisk for validiteten tilknyttet tolkningen og bruken av sumskåren på C-BiLLT. De kan snarere ytterligere belyse og bekrefte interaksjonen mellom vokabularforståelse og grammatisk forståelse i seksjonene i C-BiLLT.

6.3.3 Hva reflekterer lav kommunalitet?

Å undersøke kommunaliteter kan gi en viktig indikator på om variansen i noen seksjoner skyldes andre faktorer enn de som er ekstrahert, og om seksjonene dermed ikke er en del av det samme innholdsdomenet som de andre seksjonene i testen (Fabrigar & Wegener, 2012). Kommunaliteten for Seksjon 3 i Del 1 skilte seg i denne sammenhengen ut ved å ha lav kommunalitet noe som indikerer at unike faktorer forklarer mer av variasjonen i denne seksjonen enn de ekstraherte faktorene. Fabrigar og Wegener (2012) skiller mellom to typer unike faktorer, og det er av stor betydning å vurdere hvorvidt den lave kommunaliteten skyldes det de kaller spesifikke unike faktorer, at testleddene i denne seksjonen tapper en annen faktor enn de to ekstraherte faktorene, eller om det skyldes høye nivåer av tilfeldige målefeil, såkalte feil-faktorer. Hvis førstnevnte er tilfelle kan dette indikere at konstruktet språkforståelse feilrepresenteres med denne seksjonen, noe som kan føre til konstruktirrelevant variasjon som svekker validiteten til de slutningene som kan gjøres på bakgrunn av sumskåren på C-BiLLT. Hva gjelder spørsmålsformuleringen har imidlertid både Seksjon 1 og Seksjon 3 lik setningsoppbygging av typen *Hvor er x* og det er dermed usannsynlig at det er noe annerledes ved spørsmålsformuleringen i Seksjon 3 som fører til lav kommunalitet. I både Seksjon 1 og Seksjon 3 spørres det også om substantiv, og Seksjon 3 skiller seg dermed heller ikke stort ut hva gjelder innholdet i ordene det spørres om.

Seksjon 3 skiller seg imidlertid ut fra de øvrige seksjonene i Del 1 ved at dette er den siste seksjonen i Del 1. Det er dermed mulig at barnas oppmerksomhet avtok utover i Del 1 fordi de begynte å kjede seg ettersom oppgavene i seksjonene 1 til 3 er like i ordlyden. Variasjonen i Seksjon 3 kan derfor skyldes unike faktorer i form av usystematisk målefeil, feil-faktorer,

ettersom denne seksjonen kommer etter Seksjon 1 og 2. Ved visuell inspeksjon av tabellen over riktig svarprosent for testleddene i C-BiLLT (Vedlegg 6) ser det imidlertid ikke ut til at gjennomsnittlig prosent som har mestret testleddene i Seksjon 3 skiller seg ut fra Seksjon 1 og 2. En annen mulighet er dermed at den lave kommunaliteten er en konsekvens av at homogenitet i utvalget påvirker variansen i seksjoner og dermed korrelasjonen mellom seksjoner, noe som gjør resultatene villedende og vanskelig å tolke (Fabrigar & Wegener, 2012). Selv om faktorladningen til Seksjon 3 er noe lavere enn de øvrige seksjonene i Del 1 er den likevel høy, og det er dermed støtte for å inkludere denne seksjonen i testen selv om den hadde noe lav kommunalitet.

6.4 Evidens basert på forskjeller mellom og innad i aldersgrupper

Det er ikke måleinstrumentet i seg selv som valideres, men bruken av måleinstrumentet (Nunnally & Bernstein, 1994). Selv om resultatene fra korrelasjonsanalysen og faktoranalysen indikerer at C-BiLLT er et egnet kartleggingsinstrument til å trekke slutninger om en generell språkforståelsesevne, gjenstår diskusjonen om for hvem slutningene kan trekkes. Dersom C-BiLLT-skårer skal brukes til å utforme tilpassede tiltak og intervensjoner for barn i alderen 1;6 til 7;6 år er det viktig at det kan trekkes valide slutninger om språkforståelsesnivået til barn i alle aldersgrupper. Dersom C-BiLLT ikke kan differensiere mellom språkforståelsesnivået til barn i ulike aldersgrupper som er forventet å ha ulikt språkforståelsesnivå, kan dette indikere at de utvalgte attributtene i C-BiLLT ikke dekker innholdsdomenet i språkforståelse for alle aldersgrupper. En vurdering av hvorvidt positive konsekvenser av testing vil realiseres for alle aldersgrupper er dermed viktig i valideringsprosessen (AERA et al., 2014). Den andre forventningen beskrevet innledningsvis er i tråd med dette:

- 3) Det er forventet at C-BiLLT er sensitiv nok til å fange opp forskjeller i språkforståelsesnivå mellom barn i ulike aldersgrupper og barn i samme aldersgruppe. Dette bør reflekteres i signifikante forskjeller i gjennomsnittlig sumskåre mellom ulike aldersgrupper og fravær av gulveffekter og takeffekter innad i ulike aldersgrupper.

Selv om det er forventet at forståelse av komplekse setninger er godt etablert rundt seksårsalder, vil barn fortsette å utvikle sin språkforståelse også etter seksårsalder blant annet i form av flere ord lagret i mentalt leksikon og økt mestring og forståelse for mer komplekse grammatiske strukturer (Tomblin & Zhang, 2006; Lyster et al., 2010; LARRC et al., 2018). Det kan dermed forventes å eksistere forskjeller mellom språkforståelsesnivået til seksåringer, sjuåringer og åtteåringer. At resultatet fra variansanalysen viste signifikante forskjeller mellom gjennomsnittlige skårer på C-BiLLT i alle helårsgrupper mellom 1;6 til 8;0 år er dermed en støtte til forventningen om at de utvalgte attributter som utgjør C-BiLLT kan brukes til å differensiere mellom ulike språkforståelsesnivåer i ulike aldersgrupper. Dette resultatet skilte seg fra den nederlandske variansanalysen (Geytenbeek et al., 2014) der det ikke var signifikante forskjeller i gjennomsnittlig sumskåre på C-BiLLT mellom seks og syv år gamle barn (Vedlegg 4). Dette må imidlertid sees i sammenheng med at den nederlandske variansanalysen fra 2014 ble foretatt før Seksjon 12 *Sammensatte, komplekse setninger – fortsettelse* ble utformet og innført som den siste seksjonen i C-BiLLT. Dette kan indikere at de ti testleddene i Seksjon 12 er nødvendige attributter for å dekke innholdsdomenet i konstruert språkforståelse for barn over seks år.

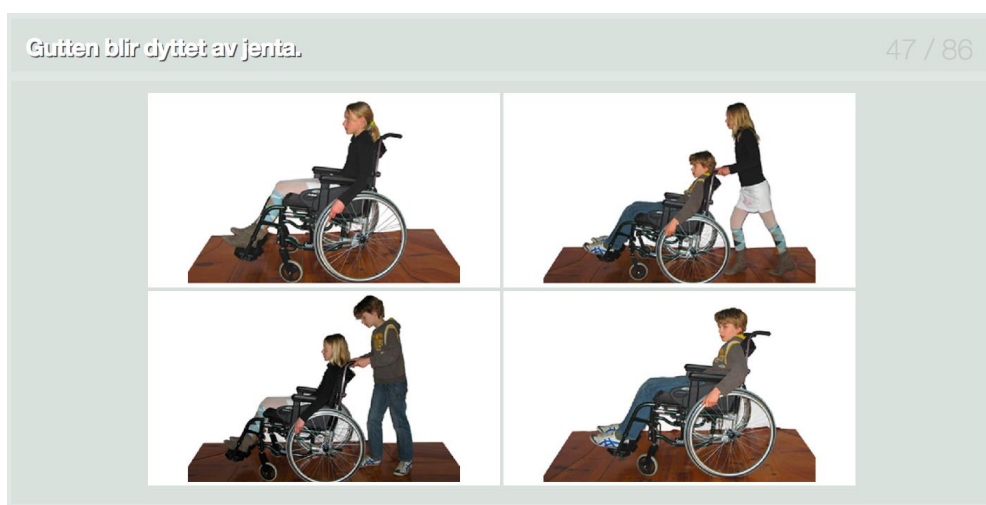
At det i denne studien også var signifikante forskjeller i gjennomsnittlige C-BiLLT-skårer mellom de tre halvårsgruppene mellom 1;6 og 3;0 år kan tolkes som en støtte til at C-BiLLT er særlig sensitiv til å fange opp forskjeller i språkforståelsesnivå hos denne yngre barnegruppen. Det var ikke signifikante forskjeller i gjennomsnittlig skåre på C-BiLLT mellom halvårsgruppene over tre år. Dette kan på den ene siden skyldes at utvalget attributter i C-BiLLT ikke er sensitivt eller dekkende nok til å differensiere mellom disse halvårsgruppene. På den andre siden kan dette også være et statistisk fenomen. Til forskjell fra i korrelasjonsanalysen er utvalget i variansanalysen delt opp i grupper, noe som gjør utvalget i hver gruppe betraktelig mindre. Utvalgsstørrelsen har videre stor innvirkning på muligheten for å oppdage signifikante forskjeller som kan slutes til å også gjelde i populasjonen. Dette er fordi det ved små utvalg er heftet større usikkerhet til om parametere estimert fra utvalget også gjelder i populasjonen, og denne usikkerheten reflekteres som høye tall på standardfeil. Standardfeil er videre med i beregningen av konfidensintervall som i en variansanalyse oppgir et intervall for hvor den gjennomsnittlige aldersgruppeforskjellen i populasjonen sannsynligvis vil falle (Field, 2016). Dersom dette intervallet krysser null indikerer dette at den sanne forskjellen mellom aldersgrupper i populasjonen, estimert fra den observerte forskjellen mellom aldersgrupper i utvalget, kan være null og dermed ikke-

eksisterende. Når utvalget er lite og standardfeilen er høy, vil konfidensintervallet bli større, noe som øker sannsynligheten for at dette intervallet krysser null og at en gjennomsnittlig forskjell mellom aldersgrupper defineres som ikke-signifikant. I små utvalg kan dermed en stor forskjell i gjennomsnitt mellom to aldersgrupper fortsatt defineres som ikke-signifikant på grunn av den store usikkerheten knyttet til denne forskjellen. Field (2016) beskriver at dette øker sannsynligheten for å gjøre Type-II-feil, som i dette tilfelle er å avvise at det er en reell forskjell i gjennomsnittlig skårer på C-BiLLT mellom halvårsgrupper over tre år i populasjonen, selv om denne forskjellen egentlig eksisterer. Dette kan også sees i sammenheng med longitudinelle studier som viser at perioden fra 1;6 til 3;0 årsalder kjennetegnes av en akselererende språklig vekst (Rowe et al., 2012; LARRC et al., 2018) og at det dermed vil eksistere større forskjeller mellom språkforståelsen mellom barn i yngre halvårsgrupper enn barn i eldre halvårsgrupper. Ved visuell inspeksjon av variansanalysen for halvårsgrupper ser det i tråd med dette ut til at forskjellene mellom gjennomsnitt var høyest for halvårsgruppene <3;0. At det kun var signifikante forskjeller mellom de tre halvårsgruppene mellom 1;6 og 3;0 kan dermed skyldes at denne forskjellen var stor nok til å bli definert som signifikant selv om gruppen hadde et lite antall, og behøver dermed ikke å skyldes at C-BiLLT ikke er sensitiv nok til å oppdage forskjeller mellom halvårsgruppene over tre år.

Selv om resultatet fra variansanalysen støtter at C-BiLLT kan differensiere mellom språkforståelsesnivået i ulike aldersgrupper er et annet spørsmål om C-BiLLT kan brukes til å differensiere mellom barn innenfor samme aldersgruppe. Spredning i typisk utviklede barns språkevne er naturlig å forvente (Simonsen et al., 2014) og dette bør også reflekteres som spredning i oppnådde C-BiLLT-skårer innenfor ulike aldersgrupper. Om det er lite spredning og gjennomsnitt i aldersgrupper er nære minimumskåre eller maksimumskåre kan dette være en indikasjon på henholdsvis gulv- eller takeffekter og at testen ikke er egnet til å differensiere mellom språkforståelsesnivået hos de yngste eller språklig lavtfungerende barna og eldste eller språklig høytfungerende barna (Price, 2017). Ved visuell inspeksjon av oversikten over norsk deskriptiv statistikk (Tabell 15) er det i denne sammenhengen positivt at gjennomsnittet i aldersgruppene under fem år er et stykke unna minimumskåre og at det er en del spredning, reflektert som høye standardavvik, fra dette gjennomsnittet innenfor hver aldersgruppe. Dette indikerer fravær av gulveffekter og støtter bruken av C-BiLLT til å differensiere mellom språkforståelsesnivået til yngre barn så vel som lavtfungerende barn. Dette støttes også ved at lignende deskriptiv statistikk er å observere i oversikten over

nederlandsk data (Tabell 16), da denne dataen kan regnes som mer solid grunnet et betraktelig høyere antall innenfor hver aldersgruppe.

Spredningen innenfor aldersgrupper ser videre ut til å synke med økende halvårsgruppe både for barna i denne studien og barna i den nederlandske studien. Å basere seg på kontekst og forståelse av innholdet i enkeltord for å tolke mer grammatisk komplekse språklige ytringer er som nevnt en vanlig strategi hos yngre barn (Strohner & Nelson, 1974). I språktester, som C-BiLLT, kan imidlertid ikke barnet basere seg på kontekst for å velge riktig svaralternativ, men må basere seg på den språklige ytringen i seg selv. I de mer grammatisk komplekse setningene i Del 2 av C-BiLLT, som testledd 47 *Gutten blir dyttet av jenta* (Figur 18), er det ikke mulig for barnet å kun basere seg på forståelsen av enkeltord i setningen, da ordenes plassering i forhold til hverandre også er avgjørende for setningens betydning. For de yngre barna er det dermed naturlig å forvente mer spredning ettersom flere svaralternativer sannsynligvis velges basert på delvis forståelse. Dersom det kan forventes at barn over fem år mer stabilt kan basere seg på den språklige ytringen i seg selv for å velge bildet som tilsvarende setningens betydningsinnhold, kan det dermed også forventes at det vil være mindre spredning i skårer for barna over fem år. En forskjell mellom norsk og nederlandsk data var at standardavvikene i de to yngste nederlandske halvårsgruppene indikerte at det var mest spredning i disse gruppene, mens standardavvikene for tilsvarende norske halvårsgrupper indikerte lite spredning i disse gruppene. De to norske halvårsgruppene har imidlertid et betraktelig lavere antall enn de tilsvarende nederlandske halvårsgruppene. At det var mindre spredning enn forventet i de to første norske halvårsgruppene vil dermed mest sannsynlig være et resultat av at de barna som ble testet tilfeldigvis var på samme språkforståelsesnivå.



Figur 18. Testledd 47 fra C-BiLLT

Forventningen om fravær av takeffekter ble støttet av at ingen barn i denne studien oppnådde maksimumskåre på C-BiLLT. Reynell har kun normer opp til seksårsalder, noe som gjør det naturlig å forvente takeffekter og små forskjeller mellom de eldste aldersgruppene. Ved visuell inspeksjon av linjediagrammet (Figur 15) er dette synlig som at økningen i gjennomsnittlig råskåre på Reynell viser en tydelig nedflatning ved fireårsalder. Skårene for C-BiLLT viser til sammenligning en brattere økning hos de eldste aldersgruppene, noe som også støtter fravær av takeffekter for de eldste barna. Dersom forskjellen mellom barns språkevnene er relativt stabil (Tomblin et al., 2014; Bornstein et al., 2013; 2016) bør det videre også forventes variasjon i språkforståelses hos barna innen de eldste aldersgruppene, og ytterligere støtte til at C-BiLLT kan brukes til å differensiere mellom språkforståelsesnivået til de eldste barna kan komme fra en inspeksjon av spredningen i de eldste aldersgruppene. Sammenlignet med nederlandske tall var imidlertid standardavvikene i de eldste aldersgruppene fra denne studien noe lave. Særlig standardavvikene i aldersgruppene over 7;1 år skilte seg i denne studien ut som lave sammenlignet med de andre aldersgruppene. Ved visuell inspeksjon av linjediagrammet blir det også tydelig at gjennomsnittlig sumskårer hos de norske barna over 5;1 år er jevnt over høyere enn i den nederlandske studien, og at det er en nedflatning i økning i gjennomsnittlig sumskåre på C-BiLLT i aldersgruppene over 6;1 år. Et spørsmål i denne sammenhengen er om grunnen til lave standardavvik i de eldste aldersgruppene er at det er få testledd som tapper mer avansert språkforståelse forventet hos barna over 6;1 år i utvalget, eller at noe av den grammatiske kompleksiteten har gått tapt i oversettelsen fra nederlandsk til norsk. Dette vil i så fall være et eksempel konstruktunderrepresentasjon, at de utvalgte attributtene i C-BiLLT ikke dekker hele innholdsdomenet i konstruktet språkforståelse for de eldste barna. Normene som utledes fra slike skårer vil da få begrenset nytte på enkeltindividsnivå ettersom de ikke vil gjenspeile faktisk språklig variasjon (Sattler, 2008). For eksempel vil et standardavvik på 2.41, som i aldersgruppen 7;7 til 8;0, bety at det kun skal fire feil til for å havne nesten to standardavvik under gjennomsnittet og dermed i lavområdet. Dette kan svekke bruken av slike skårer som et mål på barnets språkforståelsesnivå ettersom det blir vanskelig å presist differensiere mellom språkforståelsesnivået hos de eldste barna, noe som gjør at forventningen skissert over ikke fullt ut innfris.

Det er imidlertid viktig å presisere at diskusjonen om standardavvik i ulike aldersgrupper er en diskusjon basert på deskriptiv statistikk som ikke oppgir sannsynligheten for om de oppnådde tallene er usikre tall som ikke kan sluttes til populasjonen. Selv om

standardavvikene i de eldste aldersgruppene i dette utvalget var noe lave, er det ikke nødvendigvis sikkert at disse tallene også gjelder for tilsvarende aldersgrupper i populasjonen. Som nevnt er det heftet større usikkerhet ved slutninger som trekkes på bakgrunn av små utvalg, noe som er grunnen til at det optimalt bør være minst hundre personer innenfor hver aldersgruppe når en test skal normeres (Sattler, 2008). Nederlandske standardavvik kan på så måte regnes som mer robuste, da antall deltakere innen hver aldersgruppe er betydelig større. At standardavvikene for de eldste barna over fem år i denne studien var lavere enn tilsvarende nederlandske standardavvik kan dermed også indikere at det er noe med deltakerne i denne norske studien, og dermed ikke nødvendigvis at testleddene i den norske versjonen av C-BiLLT ikke differensierer mellom de eldste barna. Selv om det i denne studien er bestrebet å rekruttere barn fra øst og vest i Oslo og i indre og ytre by, var det kun 35.2% av utvalget som takket ja til å delta i studien. Ifølge De Vaus (2014) er en mulighet at de som takker ja til å delta skiller seg fra de som takker nei og at studiens deltakere dermed skiller seg systematisk fra totalt antall forespurte. Som nevnt kan det ved visuell inspeksjon av gjennomsnittlige z -skårer (Tabell 9) virke som om barna i dette utvalget skiller seg noe fra populasjonen barn med typisk utvikling ved at de har noe sterkere språklige og ikke-verbale evner enn gjennomsnittet. En annen mulig forklaring på at standardavvikene for de eldste aldersgruppene i det norske utvalget var noe lavere enn tilsvarende aldersgrupper i det nederlandske utvalget kan dermed være at utvalget i denne studien er mer homogent og jevnt over språklig sterkt noe som gjør standardavvikene unaturlig lave.

6.5 Sammenstilling av evidens

Ifølge AERA, APA og NCME (2014) handler ikke validitet kun om å samle evidens, men å integrere disse evidensformene til en koherent beskrivelse om i hvilken grad eksisterende evidens og teori støtter den intenderte tolkningen av testskårer for spesifikk bruk. Dette henger sammen med synet på validitet som et mangefasettert begrep best definert som "the degree to which evidence and theory supports the interpretations of test scores for proposed uses of tests" (AERA et al., 2014, s. 11).

Undersøkelsen av relasjonen mellom C-BiLLT og eksterne variabler i form av standardiserte tester tyder på at konstruktet som ligger under variasjoner i skårer på C-BiLLT i stor grad er relatert til konstruktene som ligger under variasjoner i skårer på Reynell, Ordgjenkjenning og

TROG-2. Dette støtter antagelsen om at det som måles med C-BiLLT også måles med disse testene, og at det som måles med C-BiLLT bør beskrives som språkforståelse. Bruken av C-BiLLT-skårer til å trekke slutninger om barns generelle språkforståelsesevne støttes også av undersøkelsen av C-BiLLTs indre struktur. Selv om variasjonene i seksjoner i faktoranalysen ble forklart av to faktorer antatt å representere vokabularforståelse og grammatisk forståelse, kan den høye korrelasjonen mellom disse faktorene tolkes som at de begge representerer aspekter ved et bredere språkforståelseskonstrukt og at en enhetlig språklig evne bidrar til variasjoner i alle seksjoner. Dette er også i tråd med synet på barns forståelse av vokabular og grammatikk som å representere spesialiserte men nært sammenhengende og interagerende ferdigheter, snarere enn ferdigheter styrt av separate og modulære systemer som kan være dissosierte. Det er dermed støtte for å kombinere skårer på seksjoner i C-BiLLT til én sumskåre tenkt å representere variasjoner i én dimensjon; variasjoner i en generell språkforståelsesevne.

At skårene på C-BiLLT kan brukes til å trekke slutninger om barns generelle språkforståelsesevne støttes også av fravær av konstruktirrelevant variasjon. I analysen av C-BiLLTs indre struktur var det kun én seksjon som hadde lav kommunalitet. Ved nærmere inspeksjon av denne lave kommunaliteten virket det videre usannsynlig at den skyldtes at testleddene i denne seksjonen tapper en faktor urelatert til språkforståelse, men at den snarere var en konsekvens av lite variasjon i testskårer grunnet at deltakerne muligens var noe homogene hva gjelder alder og språkforståelsesnivå. Videre viste undersøkelsen av relasjonen mellom C-BiLLT og eksterne variabler at C-BiLLT korrelerte mindre med CPM, og mindre med Terningmønster for barna over fire år, enn med andre tester av språkforståelse. Til sammen støtter dette at utvalgte attributtene i C-BiLLT i liten grad bidrar til konstruktirrelevant varians og at C-BiLLT dermed kan regnes som et rent mål på barns språkforståelse.

Variansanalysen viste videre at det var signifikante forskjeller i gjennomsnittsskåre på C-BiLLT mellom alle helårsgrupper og mellom de tre første halvårsgruppene. Dette indikerer at attributtene som utgjør C-BiLLT dekker innholdsdomenet i konstruktet språkforståelse for barn i ulike aldersgrupper, samt at C-BiLLT er egnet til å differensiere mellom språkforståelsesnivået til barn i ulike aldersgrupper. Inspeksjon av både norske og nederlandske gjennomsnitt og standardavvik i ulike halvårsgrupper indikerer videre fravær av gulveffekter og at den norske versjonen av C-BiLLT dermed kan egne seg til å differensiere

mellom språkforståelsesnivået til de yngste barna og språklig lavtfungerende barn. Ingen barn oppnådde maksimumskåre på C-BiLLT, noe som videre indikerer fravær av takeffekter. Det var imidlertid noe mindre spredning enn forventet i C-BiLLT-skårer innad i de eldste aldersgruppene over fem år. Om dette reflekterer konstruktunderrepresentasjon, at attributtene som utgjør C-BiLLT er mindre egnet til å differensiere mellom de eldste eller de språklig høytfungerende barna, eller om det kun er en refleksjon av at deltakerne i denne studien er mer homogene og språklig sterke og dermed ikke varierer stort fra hverandre hva gjelder språkforståelsesnivå, er imidlertid usikkert. At disse tallene avviker fra tilsvarende nederlandske tall som kan regnes som mer robuste da antall deltakere innenfor hver aldersgruppe er betraktelig større, kan imidlertid indikere at sistnevnte forklaring er mest sannsynlig. Dette fordrer at C-BiLLT utprøves på et større og mer heterogent utvalg før det kan gjøres en endelig vurdering om hvorvidt C-BiLLT egner seg til å differensiere mellom de eldste og de språklig høytfungerende barna.

Til sammen støtter dette at den norske versjonen av C-BiLLT er godt egnet til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år. Sammenstilling av evidens tyder på at attributtene C-BiLLT utgjøres av representere innholdsdomenet i konstruktet språkforståelse, samt at disse attributtene dekker innholdsdomenet for språkforståelseevnen til barn i ulike aldersgrupper og innad i ulike aldersgrupper. Resultatene fra denne studien kan imidlertid tyde på at det må heftes en noe større usikkerhet til nøyaktigheten i slutningene som kan trekkes om språkforståelse hos barna over fem år på bakgrunn av skårer på C-BiLLT, før C-BiLLT er prøvd ut på et større og mer heterogent utvalg.

6.6 Styrker og begrensninger

Når kvantitative design benyttes er en utfordring om antallet deltakere er stort nok til å kunne slutte at resultatene fra studien også er gjeldende for populasjonen (De Vaus, 2014). En styrke ved denne studien er derfor at antall deltakere var relativt stort, noe som støtter generaliserbarheten av funnene i denne studien. At det var mulig å benytte parametriske statistikk er også en fordel, da parametriske statistikk er forbundet med større statistisk styrke enn ikke-parametriske statistikk (Field, 2016).

En annen styrke er at C-BiLLT i denne studien er validert opp mot flere standardiserte og anerkjente tester av både språkforståelse og ikke-verbal resonneringsevne, og at det i denne

studien er inkludert flere tester enn i den nederlandske utprøvingen av C-BiLLT. Dette gjør det mulig å nyansere forståelsen av hva C-BiLLT måler. En mulig begrensning med dette er imidlertid at kartleggingsøkten ble noe lang, særlig for de eldste barna over 4;1 år. Det er dermed mulig at skårene på testene i denne studien er påvirket av tilfeldige feil grunnet barnas manglende motivasjon eller varierende oppmerksomhet. Det er i denne sammenhengen en styrke at testene ble administrert i to ulike rekkefølger og at det dermed var mulig å undersøke og fastslå at det ikke var signifikante forskjeller i testskårer på noen av testene mellom barna som hadde fått ulik testrekkefølge.

Til tross for at antall deltakere i denne studien var stort, ble antallet deltakere fortsatt noe lavt når det kom til sammenligninger gjort mellom aldersgrupper og innad i aldersgrupper. Små utvalgsstørrelser er heftet med større usikkerhet og kan dermed redusere styrken parametriske tester har til å oppdage og avgjøre om effekter er signifikante og om effekter er ikke-signifikante, noe som øker sannsynligheten for å gjøre henholdsvis Type-II- og Type-I-feil (Field, 2016). Det er dermed heftet større usikkerhet til slutningene fra variansanalysen og konklusjonen angående gulveffekter og takeffekter i denne studien. En styrke er likevel her at det var mulig å sammenligne funnene i denne studien med tilsvarende nederlandsk data som ikke er forventet å avvike i stor grad fra norsk data og som har betraktelig flere deltakere.

Selv om utvalget i denne studien er stort garanterer dette ikke fraværet av systematiske forskjeller fra populasjonen *barn i alderen 1;6 til 7;6 år med typisk utvikling*. De Vaus (2014) fremhever i denne sammenhengen at prinsippet om fritt samtykke kan forhindre deltakernes representativitet ettersom de som velger å delta i studien kan skille seg systematisk fra de som velger å ikke delta. I denne studien kan det virke som om deltakerne er noe over gjennomsnittet både hva gjelder språklige og ikke-verbale evner, og dette vil i så fall svekke generaliserbarheten av funnene i denne studien. Deltakerne har videre en skjev aldersfordeling, med flere eldre enn yngre deltakere. I denne sammenhengen er en begrensning ved denne studien at det kun var mulig å utføre faktoranalysen på seksjonene i C-BiLLT og ikke på testleddene. Dette skyldes sannsynligvis at deltakerne var noe homogene hva gjelder språkforståelsesnivå, og at det dermed ble for lite variasjon i skårer på testledd slik at korrelasjonene mellom testledd ikke var meningsfulle eller tolkbare. Selv om resultatene fra faktoranalysen viste at alle seksjoner ladet på minst en av de ekstraherte faktorene, var det dermed ikke mulig å vurdere hvorvidt noen spesifikke testledd skiller seg

ut ved å ikke lade på noen av de ekstraherte faktorene og dermed bør endres fordi de i liten grad tapper konstruert språkforståelse.

6.7 Kliniske implikasjoner

I denne oppgaven har det blitt skissert et behov for en standardisert språkforståelsestest som kan benyttes for å kartlegge språkforståelse hos barn med store tale- og bevegelsesvansker, ettersom eksisterende standardiserte tester av språkforståelse krever besvarelse gjennom motorisk eller verbal respons og dermed er lite egnet for denne gruppen barn. I sammenheng med dette rapporteres det at få barn med store tale- og bevegelsesvansker kartlegges (Andersen, et al., 2008; Sigurdadottir et al., 2008; Sherwell et al., 2014; Stadskleiv et al., 2015). Dette kan tyde på at tiltak for mange av barna baseres på antakelser om språklige funksjonsnivå. Da det ikke eksisterer et en-til-en-forhold mellom hva barn med store tale- og bevegelsesvansker kan uttrykke av språk og hva de forstår av språk, og at flere ikke-talende barn med CP har aldersadekvat reseptivt språk (Bishop et al., 1990; Frampton et al., 1998; Straub & Obrzut, 2009; Geytenbeek et al., 2015) er det grunn til å betvile gyldigheten til slike antagelser og tiltakene som baseres på dem.

At resultatene fra denne studien og reliabilitetsstudien støtter bruken av C-BiLLT-skårer til å trekke valide og reliable slutninger om barns språkforståelse, samt at bruk av fingerpeking og blikkpeking som besvarelsesmåte ikke hadde noen innvirkning på barnas skårer, indikerer at C-BiLLT har potensiale til å fylle dette behovet. Dette kan ha store kliniske implikasjoner.

Eksistensen av en test som C-BiLLT kan bidra til at flere barn med store tale- og bevegelsesvansker kartlegges enn det som rapporteres i dag. Dette kan igjen medføre at tiltak og ASK blir mer tilpasset barnets behov da utformingen av dem kan baseres på valide og reliable slutninger om barnets språkforståelsesnivå snarere enn antagelser. At barn med store tale- og bevegelsesvansker får tilgang til tilpassede intervensjoner og ASK er videre fundamentalt ikke bare for språklig utvikling, men også for deres muligheter til sosial deltakelse, læring og generelle kognitiv utvikling (Light et al., 2002; Tetzchner & Martinsen, 2002; Murray & Goldbart, 2009; Batorowicz et al., 2016). Kunnskap om barnets språkforståelsesnivå gir videre verdifull informasjon om hvordan barnets omsorgspersoner, lærere og andre bør interagere med barnet på en måte som verken er for språklig kompleks eller for enkelt, men tilpasset barnets faktiske språkforståelsesfunksjonsnivå. Mei og kollegaer (2016) beskriver videre at det er en mangel på kunnskap om språkforståelsesnivå

hos barn med CP blant annet grunnet mangel på standardiserte tester. En annen viktig konsekvens av en test som C-BiLLT kan dermed være økt kunnskap på forsknings- og praksisfeltet om språkforståelse hos barn med CP, og et generelt økt fokus på intervensjoner som minker avstanden som kan eksistere mellom deres ekspressive tale og reseptive språk.

At det ikke var noen forskjell på barnets skårer utfra om de besvarte C-BiLLT ved fingerpeking eller blikkpeking er videre i tråd med funn fra tidligere studier som undersøker testbesvarelse gjennom ulike responsmodaliteter (Spillane et al., 1996; Casey et al., 2007; Kurmanaviciute & Stadskleiv, 2017). Slike funn kan dermed støtte bruken av ulike responsmodaliteter for å besvare andre typer oppgaver og tester som har flervalgsalternativer, noe som videre gir støtte til bruk av ulike responsmodaliteter for å tilrettelegge kartlegging av for eksempel ikke-verbal resonneringsevne hos barn med store tale- og bevegelsesvansker.

C-BiLLT kan også benyttes i kartlegging av språkforståelse hos barn uten tale- og bevegelsesvansker. Testen er enkel og administrere og skåre, og at den er databesvart kan også gjøre at den fremstår som attraktiv for barn. I denne studien viste for eksempel ingen av barna motvilje mot å bli retestet med C-BiLLT, noe som kan indikere at de fant testen spennende og underholdende. Mistanke om forsinket språkutvikling er videre en vanlig henvisningsgrunn og det er dermed mange barn som har behov for kartlegging av språkforståelse. At flere studier indikerer at barns rangordnede språkforståelsesevne blir mer stabil over tid (Bornstein, et al., 2014; Tomblin et al., 2014; Bornstein et al., 2016) fordrer at en kartlegging gjøres tidlig slik at tilpassede tiltak og intervensjoner kan settes inn tidlig i barnets utvikling. Tidlig innsats i barnehagen er i sammenheng med dette et vektlagt fokusområde fra politisk hold (Utdanningsdirektoratet, 2017). Ettersom C-BiLLT dekker aldersspennet 1;6 til 7;6 år har den dermed også potensiale til å bli et mye brukt verktøy i kartlegging av språkforståelse hos yngre barn uten store tale- og bevegelsesvansker.

7 Avslutning

Formålet med denne studien har vært å undersøke hvor egnet den norske versjonen av C-BiLLT er til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 til 7;6 år. Barn med store tale- og bevegelsesvansker har et behov for tilpassede intervensjoner og ASK som bygger på en kartlegging av blant annet språkforståelse. Denne studiens formål grunner i behovet for en språkforståelsestest tilpasset denne gruppen, samt et behov for en ny standardisert og normert norsk språkforståelsestest som dekker aldersspennet 1;6 til 7;6 år og dermed kan benyttes for å realisere tidlig innsats. For å undersøke dette har 187 barn i alderen 20 til 95 måneder blitt kartlagt med C-BiLLT og andre tester av språkforståelse og tester av ikke-verbal resonneringsevne. Barnas skårer har så blitt analysert statistisk og tre forventninger til validitet er diskutert i lys av de empiriske resultatene og teori.

Den første forventningen var at korrelasjonsanalysen skulle vise høyere korrelasjoner mellom skårene på C-BiLLT og skårer på andre tester av språkforståelse enn mellom skårene på C-BiLLT og skårer på tester av ikke-verbal resonneringsevne. Denne forventningen ble innfridd ved at C-BiLLT-skårer korrelerte høyere med skårer på Reynell, Ordgjenkjenning og TROG-2, enn med den ikke-verbale resonneringsevnetesten CPM. Korrelasjonen mellom C-BiLLT og den ikke-verbale resonneringsevnetesten Terningmønster var høy, men ble mer moderat da barna under 4;1 år ble fjernet fra korrelasjonsanalysen. Dette ble tolket som å være i tråd med studier som viser at kognitive evner er mindre differensiert hos yngre barn (Colledge et al., 2002; Tideman & Gustafsson, 2011), snarere enn at C-BiLLT måler ikke-verbal resonneringsevne. Til sammen støttet dette antagelsen om at det som måles med C-BiLLT også måles med andre tester av språkforståelse, men ikke av tester av ikke-verbal resonneringsevne, og at det som måles med C-BiLLT dermed kan beskrives som et rent mål på språkforståelse.

Den andre forventningen var at en eksplorerende faktoranalyse på C-BiLLT skulle resultere i samme faktorløsning som for den nederlandske versjonen av C-BiLLT; en endimensjonal faktorløsning som representerer en overordnet språkforståelsesevne. Faktoranalysen i denne studien resulterte i at to faktorer, tolket som å representerte vokabularforståelse og grammatisk forståelse, ble ekstrahert. At deltakerne i denne studien hadde en skjev aldersfordeling, med flere eldre barn enn deltakerne i den nederlandske undersøkelsen, ble videre trukket frem som en mulig grunn til denne faktorløsningen da flere studier indikerer at

språkevnene hos barn vil bli mer differensiert med alder, reflektert i lavere korrelasjoner mellom mål på grammatikk og vokabular (Hayiou-Thomas et al., 2006; Lyster et al., 2010; LARRC et al., 2018). At det i denne studien likevel var en høy korrelasjon mellom de to ekstraherte faktorene, og at begge faktorer forklarte meningsfulle variasjoner i alle seksjoner, ble videre tolket som at de begge representerer aspekter ved et bredere språkforståelseskonstrukt og at en enhetlig språklig evne bidrar til variasjoner i alle seksjoner. Dette støttet dermed den andre forventningen og at skårer på seksjoner i C-BiLLT kan kombineres til én sumskåre tenkt å representere variasjoner i én dimensjon; variasjoner i en generell språkforståelsesevne.

Den tredje og siste forventningen var at C-BiLLT skulle være sensitiv nok til å fange opp forskjeller i språkforståelsesnivå mellom barn i ulike aldersgrupper og innad i samme aldersgruppe. Denne forventningen ble delvis innfridd. Resultatene fra enveis ANOVA viste signifikante forskjeller mellom barn i alle helårsgruppene, noe som ble tolket som at C-BiLLT er egnet til å differensiere mellom barn i ulike aldersgrupper. Antagelsen om at C-BiLLT egner seg til å differensiere mellom språkforståelsesnivået til barna i de yngste aldersgruppene og språklig lavtfungerende barn ble støttet ved inspeksjon av deskriptiv statistikk som indikerte fravær av gulveffekter og tilstedeværelsen av en del spredning innad i aldersgruppene under fem år. Dette ble også tolket som at C-BiLLT dekker innholdsdomenet ved konstruktet språkforståelse for barna i disse aldersgruppene. Det var videre fravær av takeffekter men noe lavere spredning i de eldste aldersgruppene enn i tilsvarende aldersgrupper fra den nederlandske undersøkelsen. Dette ble tolket som å enten representere at den norske versjonen av C-BiLLT er mindre egnet til å differensiere mellom språkforståelsesnivået til de eldste og språklig høytfungerende barna, eller at deltakerne i denne studien var mer homogene enn i den nederlandske undersøkelsen og at standardavvikene i de eldste aldersgruppene i denne studien dermed ble unaturlig lave.

Til sammen støtter evidens gjennomgått i denne oppgaven at den norske versjonen av C-BiLLT er godt egnet til å trekke valide slutninger om språkforståelse hos barn mellom 1;6 og 7;6 år. Sammenstilling av resultatene indikerer at attributtene C-BiLLT utgjøres av representere innholdsdomenet i konstruktet språkforståelse, samt at disse attributtene dekker innholdsdomenet for språkforståelsesevnen både til barn i ulike aldersgrupper og barn innad i samme aldersgrupper. Resultatene gjennomgått i denne studien kan imidlertid tyde på at det må heftes en noe større usikkerhet til presisjonen i slutningene som kan trekkes om

språkforståelse hos barna over fem år på bakgrunn av skårer på C-BiLLT, før C-BiLLT er undersøkt på et større og mer heterogent utvalg.

7.1 Veien videre

En valideringsprosess er en prosess uten ende som innebærer pågående akkumulering av evidens (Price, 2017). I tråd med vektleggingen av at det er bruken av måleinstrumentet som valideres, og ikke måleinstrumentet i seg selv, er det viktig at C-BiLLT prøves ut også på barn med CP og barn med store tale- og bevegelsesvansker slik at det kan heftes sikkerhet vet at funnene i denne studien også kan generaliseres til å gjelde disse barna. Det er videre også ønskelig at C-BiLLT prøves ut på et større antall deltakere som er mer heterogene med tanke på språkforståelsesnivå og alder. Dette kan for det første muliggjøre at faktoranalysen kan utføres på testleddene i C-BiLLT fremfor seksjonene. Det vil da bli mulig å undersøke om noen spesifikke testledd skiller seg ut og bør forandres for å sikre at alle testledd tapper aspekter ved språkforståelse. Å utprøve C-BiLLT på et større og mer heterogent utvalg kan for det andre bidra til at konklusjonen angående om C-BiLLT er egnet til å differensiere mellom barn innad i ulike aldersgrupper kan trekkes på et sikrere grunnlag. For å etablere ytterligere evidens for dette kan det videre være ønskelig å undersøke om C-BiLLT er egnet til å differensiere mellom grupper som er forventet å ha signifikant forskjellig språkforståelsesnivå, som barn med typisk språkutvikling og barn med påviste språkvansker. Kartlegging av flere barn i alle aldersgrupper er også en forutsetning for å etablere normer.

Litteraturliste

- Abend, O., Kwiatowski, T., Smith, J. N., Goldwater, S., & Steedman, M. (2017). Bootstrapping language acquisition. *Cognition*, *164*, 116-143. doi:10.1016/j.cognition.2017.02.009
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing* (3. utg.). Washington, DC: Authors.
- Andersen, G., Irgens, L. M., Haagaas, I., Skranes, J. S., Meberg, A. E., & Vik, T. (2008). Cerebral palsy in Norway: Prevalence, subtypes and severity. *European Journal of Paediatric Neurology*, *12*(1), 4-13. doi:10.1016/j.ejpn.2007.05.001
- Andersen, G. L., Mjøen T., & Vik, T. (2010). Prevalence of speech problems and the use of augmentative and alternative communication in children with cerebral palsy: a registry-based study in Norway. *Perspectives on Augmentative and Alternative Communication*, *19*(1), 12-20. doi:10.1044/aac19.1.12
- Anthony, J. L., Davis, C., Williams, J. M., & Anthony, T. I. (2014). Preschoolers' oral language abilities: A multilevel examination of dimensionality. *Learning and individual differences*, *35*, 56-61. doi:10.1016/j.lindif.2014.07.004
- Arciuli, J. & Torkildsen, J. V. K. (2012). Advancing our understanding of the link between statistical learning and language acquisition: the need for longitudinal data. *Frontiers in Psychology*, *3*, 1-9. doi:10.3389/fpsyg.2012.00324
- Barnekonvensjonen. (2003). *FNs konvensjon om barnets rettigheter: Vedtatt av De Forente Nasjoner den 20. november 1989, ratifisert av Norge den 8. januar 1991: Revidert oversettelse mars 2003 med tilleggsprotokoller*. Oslo: Barne- og familiedepartementet.
- Batorowicz, B., Stadskleiv, K., von Tetzchner, S., & Missiuna, C. (2016). Children who use communication aids instructing peer and adult partners during play-based activity. *Augmentative and Alternative Communication*, *32*(2), 105-119. doi:10.3109/07434618.2016.1160150
- Benson, J. (1998). Developing a strong program of construct validation: a test anxiety example. *Educational Measurement: Issues and Practice*, *17*(1), 10-17. doi:10.1111/j.1745-3992.1998.tb00616.x

- Bishop, D. V. M., Brown B. B., & Robson, J. (1990). The relationship between phoneme discrimination, speech production, and language comprehension in cerebral-palsied individuals. *Journal of Speech and Hearing Research*, 33(2), 210-209.
doi:10.1044/jshr.3302.210
- Bishop, D. V. M. (1997). Cognitive Neuropsychology and developmental disorders: Uncomfortable Bedfellows. *The Quarterly Journal of Experimental Psychology*, 50(4), 899-923. doi:10.1080/713755740
- Bishop, D. V. M. (2009). *Test for reception of Grammar, second edition, Norwegian version*. Bromma, Sweden: Pearson Assessment.
- Bishop, D. V. M. (2014). *Uncommon Understanding. Development and Disorders of Language Comprehension in Children*. London: Psychology Press Ltd.
- Bjerkan, M. K. (2011). Fonologi. I K. E. Kristoffersen, H. G. Simonsen, A. Sveen (Red.), *Språk; en grunnbok*. (s. 198-221). Oslo: Universitetsforlaget.
- Bornstein, M. H., Hahn, C. S., Putnick, D. L., & Suwalsky, J. T. (2014). Stability of core language skill from early childhood to adolescence: a latent variable approach. *Child Development*, 85(4), 1346-1356. doi:10.1111/cdev.12192
- Bornstein, M. H., Hahn, C. S., & Putnick, D. L. (2016). Long-term stability of core language skills in children with contrasting language skills. *Developmental Psychology*, 52(5), 704-716. doi:10.1037/dev0000111
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
doi:10.1037/h0046016
- Cans, C., De-la-Cruz, J., & Mermet, M. A. (2008). Epidemiology of cerebral palsy. *Paediatrics and Child Health*, 18(9), 393-398. doi:10.1016/j.paed.2008.05.015
- Casey, M., Tonsing, K. M., & Alant, E. (2007). Comparison of a non-spoken response mode and a spoken response mode in a test of phonological awareness. *South African Journal of Occupational Therapy*, Nov., 25-28. Hentet fra <http://repository.up.ac.za/handle/2263/6279>
- Choi, J. Y., Choi, Y. S., & Park E. S. (2017). Development and brain magnetic resonance imaging characteristics in preschool children with cerebral palsy. *Journal of Speech, Language and Hearing Research*, 60(5), 1330-1338.
doi:10.1044/2016_JSLHR-L-16-0281
- Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control*, 2(2), 137-167. doi:10.1016/S0019-9958(59)90362-6

- Colledge, E., Bishop, D. V. M., Koeppen-Schomerus, G., Price, T. S., Happé, F. G. E., Eley, T. C., . . . Plomin, R. (2002). The structure of language abilities at 4 years: A twin study. *Developmental Psychology*, *38*(5), 749-757. doi:10.1037//0012-1649.38.5.749
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, *10*(7), 1-9. Hentet fra: <http://pareonline.net/pdf/v10n7.pdf>
- Cronbach, L. J. & Meehl, P. E. (1995). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302. doi:10.1037/h0040957
- De Vaus, D. A. (2014). *Surveys in Social Research. An Introduction* (6. utg.). London: Routledge.
- Eliasson, A C., Rosblad, B., Krumlinde-Sundholm, L., Beckung, E., Arner, M., Ohrwall, A-M., & Rosenbaum, P. (2006). Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Developmental Medicine & Child Neurology*, *48*(7), 549-554. doi:10.1111/j.1469-8749.2006.tb01313.x
- Fabrigar, L. R. & Wegener, D. T. (2012). *Exploratory Factor Analysis*. New York: Oxford University Press.
- Fazzi, E., Signorini, S. G., Piana, R. L., Bertone, C., Misefari, W., Galli, J. . . . Bianchi, P. E. (2012). Neuro-ophthalmological disorders in cerebral palsy: ophthalmological, oculomotor and visual aspects. *Developmental Medicine & Child Neurology*, *54*(8), 730-736. doi:10.1111/j.1469-8749.2012.04324.x
- Field, A. (2016). *Discovering Statistics Using IBM SPSS Statistics*. London: Sage Publications Ltd.
- Frampton, I., Yude, C., & Goodman, R. (1998). The prevalence and correlates of specific difficulties in a representative sample of children with hemiplegia. *British Journal of Educational Psychology*, *68*(1), 39-51. doi: 10.1111/j.2044-8279.1998.tb01273.x
- Friborg, O. (2010). Klassisk testteori og utvikling av spørreinstrumenter. I M. Martinussen (Red). *Kvantitativ Forskningsmetodologi i Samfunns og Helsefag* (s. 15–54). Bergen: Fagbokforlaget.
- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*(5), 887-898. doi:10.1037/0012-1649.28.5.887

- Geytenbeek, J. J., Heim, M. M. J., Vermeulen, J. R., & Oostrom, K. J. (2010a). Assessing Comprehension og spoken language in children with cerebral palsy: Application of a newly developed computer-based instrument. *Augmentative and Alternative Communication*, 26(2), 97-107. doi:10.3109/07434618.2010.482445
- Geytenbeek, J. J., Harlaar, L., Stam, M., Ket, H., Bevher, J. G., Oostrom, K., & Vermeulen, R. J. (2010b). Utility of language comprehension tests for unintelligible or non speaking children with cerebral palsy: a systematic review. *Developmental Medicine & Child Neurology*, 52(12), 267-277. doi:10.1111/j.1469-8749.2010.03807.x
- Geytenbeek, J. J., Mokkink, L. B., Knol, D. L., Vermeulen, R. J., & Oostrom, K. J. (2014). Reliability and validity of the C-BiLLT: A new instrument to assess comprehension of spoken language in young children with cerebral palsy and complex communication needs. *Augmentative and Alternative Communication*, 3, 1-15. doi:10.3109/07434618.2010.482445.
- Geytenbeek, J. J. M., Heim, M. J. M., Knol, D. J., Vermeulen, R. J. & Oostrom, K. J. (2015). Spoken language comprehension of phrases, simple and compound-active sentences in non-speaking children with severe cerebral palsy. *International Journal of Language & Communication Disorders*, 50(4), 499-515. doi:10.1111/1460-6984.12151
- Gleitman, L. R. (1994). Words, Words, Words. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 346(1315), 71-77. doi:10.1098/rstb.1994.0130
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14(1), 23-45. doi:10.1017/S030500090001271X
- Hayiou-Thomas, M. E., Kovas, Y., Harlaar, N., Plomin R., Bishop, D. V. M., & Dale, P. S. (2006). Common aetiology for diverse language skills in 4 ½ year old twins. *Journal of Child Lanugage*, 33(2), 339-368. doi:10.1017/S0305000906007331
- 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.
- Hagtvet, B. & Lillestølen, R. (1990). *Reynells Språkttest*. Oslo: Gyldendal Akademisk.

- Hidecker, M. J., Paneth, N., Rosenbaum, P. L., Kent, R. D., Lillie, J., Eulenberg, J. B., ... Taylor, K. (2011). Developing and validating the Communication Function Classification System for individuals with cerebral palsy. *Developmental Medicine & Child Neurology*, 53(8), 704-710. doi:10.1111/j.1469-8749.2011.03996.x.
- Himmelman, K., Beckung, E., Hagberg, G., & Uvebrant, P. (2006). Gross and fine motor function and accompanying impairments in cerebral palsy. *Developmental Medicine & Child Neurology*, 48(6), 417-423. doi:10.1111/j.1469-8749.2006.tb01289.x
- Himmelman, K., Hagberg, U., & Uvebrandt, P. (2010). The changing panorama of cerebral palsy in Sweden. Prevalence and origin in the birth-year period 1999-2002. *Acta Paediatrica*, 99(9), 1337-1343. doi:10.1111/j.1651-2227.2010.01819.x
- Himmelman, K., Lindh, K. & Hidecker, M. J. (2012). Communication ability in cerebral palsy: a study from the CP register of western Sweden. *European Journal of Paediatric Neurology*, 17(6), 568-574. doi:10.1016/j.ejpn.2013.04.005
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368-1378. doi:10.1111/1467-8624.00612
- Huttenlocher, J., Haight, W., Bryk, A., Selzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236-248. doi:10.1037/0012-1649.27.2.236
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61(4), 343-365. doi:10.1016/j.cogpsych.2010.08.002
- Johanson, G. A. & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement*, 70(3), 394-400. doi:10.1177/0013164409355692
- Jusczyk, P. W. & Luce, P. A. (2002). Speech perception and spoken word recognition: past and present. *Ear & Hearing*, 23(1), 2-40. doi:10.1097/00003446-200202000-00002
- Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: rethinking the study of developmental disorders. *Developmental Psychology*, 45(1), 56-63. doi:10.1037/a0014506.
- Kid, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, 48(1), 171-84. doi:10.1037/a0025405

- Klem, M., Gustafsson, J-E., & Hagtvet, B. (2014). Dimensionality of language ability in four-year-olds: Construct validation of a language screening tool. *Scandinavian Journal of Educational Research*, 59(2), 195-213.
doi:10.1080/00313831.2014.904416
- Kleven, T. A. (2008). Validity and validation in qualitative and quantitative research. *Nordisk Pedagogikk*, 28(3), 219-233. Hentet fra https://www.idunn.no/np/2008/03/validity_and_validation_in_qualitative_and_quantitative_research
- Kristoffersen, K. E. (2011). Hva er språk?. I K. E. Kristoffersen, H. G. Simonsen, A. Sveen (Red.), *Språk; en grunnbok*. (s. 17-35). Oslo: Universitetsforlaget.
- Kurmanacivute, R. & Stadskleiv, K. (2017). Assessment of verbal comprehension and cognitive reasoning when standard response mode is challenging: a comparison of different response modes and an exploration of their clinical usefulness. *Cogent Psychology*, 4(1), 1-17. doi:10.1080/23311908.2016.1275416
- Lahey, M. (1988). *Language disorders and language development*. Needham, MA: Macmillan.
- Language and Reading Research Consortium. (2015). The dimensionality of language ability in young children, *Child Development*, 86(6), 1948-1965.
doi: 10.1111/cdev.12450
- Language and Reading Research Consortium, Jiang, H., Logan, J. A., & Jia, R. (2018). Modeling the nature of grammar and vocabulary trajectories from prekindergarten to third grade. *Journal of Speech, Language, and Hearing Research*, 61(4), 910-923.
doi:1044/2018_JSLHR-L-17-0090.
- Light, J. C., Parsons, A. R. & Drager, K. (2002). "There is more to life than cookies": developing interactions for social closeness with beginning communicators who use AAC. I J. Reichle, D. R. Beukelman & J. C. Light (Red.), *Exemplary practices for beginning communicators. Implications for AAC* (s. 187-214). Baltimore: Paul H Brookes Publishing Co.
- Lust, B. C. (2006). *Child Language. Acquisition and growth*. New York: Cambridge University Press.

- 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. Hentet fra: <https://www.utdanningsnytt.no/globalassets/filer/pdf-av-spesialpedagogikk/2010/spesialpedagogikk-9-2010.pdf>
- Marchman, V. A. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21(2), 339-366. doi:10.1017/S0305000900009302
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57-77. doi:10.1207/s15516709cog1401_4
- Mei, C., Reilly, S., Reddihough, D., Mensah, F., Pennington, L., & Morgan, A. (2016). Language outcomes of children with cerebral palsy aged 5 years and 6 years: a population-based study. *Developmental Medicine & Child Neurology*, 58(6), 605-611. doi:10.1111/dmcn.12957
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. doi:10.1037/0003-066X.50.9.741
- Miller, J. F. & Paul, R. (1995). *The clinical assessment of language comprehension*. Baltimore: Paul H. Brooks Publishing Co.
- Moyle, J. M., Weismer, S. E., Evans, J. L., & Lindstrom, M. J. (2006). Longitudinal Relationship between lexical and grammatical development in typical and late talking children. *Journal of Speech, Language, and Hearing Research*, 50(2), 508-528. doi:10.1044/1092-4388(2007/035)
- Murray, J., & Goldbart, J. (2009). Cognitive and language acquisition in typical and aided language learning: A review of recent evidence from an aided communication perspective. *Child Language Teaching and Therapy*, 25(1), 31-58. doi:10.1177/0265659008098660
- NESH (2016). *Forskningsetiske Retningslinjer for Samfunnsvitenskap, Humaniora, Juss og Teologi* (4. utg.). Hentet fra https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf
- Norbury, C. F., Bishop, D. M. V., & Briscoe, J. (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module? *Applied Psycholinguistics*, 23(2), 247-268. doi:10.1017.S0142716402002059

- Nordberg, A., Miniscalco, C., Lohmander, A., & Himmelmann, K. (2013). Speech problems affect more than one in two children with cerebral palsy: Swedish population-based study. *Acta paediatrica*, *102*, 161-166. doi:10.1111/apa.12076
- Nunnally, J. C. & Berinstein, I. H. (1994) *Psychometric Theory*. New York: McGraw-Hill.
- Odding, E., Roebroek, M. E., & Stam, H. J. (2006). The epidemiology of cerebral palsy: incidence, impairments and risk factors. *Disability and Rehabilitation*, *28*(4), 183-191. doi:10.1080/09638280500158422
- Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E. & Galuppi, B. (1997). Development of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine & Child Neurology*, *39*(4), 214-223. doi:10.1111/j.1469-8749.1997.tb07414.x
- Parkes, J. Hill, N., Platt, M. J., & Donnelly, C. (2010). Oromotor dysfunction and communication impairments in children with cerebral palsy: a register study. *Developmental Medicine & Child Neurology*, *52*(12), 1113-1119. doi:10.1111/j.1469-8749.2010.03765.x.
- Pearson, E. S., & Hartley, H. O. (1954). *Biometrika Tables for Statisticians*. New York: Cambridge University Press.
- Pennington, L. & McConachie, H. (2001). Predicting patterns of interaction between children with cerebral palsy and their mothers. *Developmental Medicine & Child Neurology*, *43*(2), 83-90. doi:10.1111/j.1469-8749.2001.tb00720.x
- Personopplysningsloven. (2001). Lov om behandling av personopplysninger m.v. av 1. januar 2001. Hentet fra: https://lovdata.no/dokument/NL/lov/2000-04-14-31/KAPITTEL_2#%C2%A712
- Petscher, Y., Connor, C. M. D., & Otaiba, S. A. (2012). Psychometric analysis of the diagnostic evaluation of language variation assessment. *Assessment for Effective Intervention*, *37*(4), 243-250. doi:10.1177/1534508411413760
- Pinker, s. (1998). Words and Rules. *Lingua*, *106*, 219-242. Hentet fra <https://stevenpinker.com/publications/words-and-rules-paper>
- Pirila, S., van der Meere, J., Pentikainen, T., Ruusu-Niemi, P., Korpela, R., Kilpinen, J., ... Kilpinen, R. (2007). Language and motor speech skills in children with cerebral palsy. *Journal of Communication Disorders*, *40*(2), 116-128. doi:10.1016/j.jcomdis.2006.06.002

- Plunkett, K. & Marchman, V. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48(1), 21-69.
doi:10.1016/0010-0277(93)90057-3
- Price, L. R. (2017). *Psychometric methods. Theory into practice*. New York: The Guilford Press.
- Raven, J. (2008). *Colored progressive matrices and Crichton vocabulary scale manual*. Bromma, Sweden: Pearson Assessment.
- Romberg, A. R. & Saffran, J. R. (2010). Statistical learning and acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 906-914. doi:10.1002/wcs.78
- Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., ... Jacobsson, B. (2007). A Report: the definition and classification of cerebral palsy april 2006. *Developmental Medicine & Child Neurology*, 49(6), 8-14.
doi:10.1111/j.1469-8749.2007.tb12610.x
- Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, 83(2), 508-525.
doi:10.1111/j.1467-8624.2011.01710.x
- Sabbadini, M., Bonanni, R., Carlesimo, G. A., & Caltagirone, C. (2001). Neuropsychological assessment of patients with severe neuromotor and verbal disabilities. *Journal of Intellectual Disability Research*, 44(2), 169-179.
doi:10.1046/j.1365-2788.2001.00301.x
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old Infants. *Science*, 247(5294), 1926–1928. doi:10.1126/science.274.5294.1926
- Sankar, C., & Mundkur, N. (2005). Cerebral palsy – definition, classification, etiology and early diagnosis. *Indian Journal of Pediatrics*, 72(10), 865-868.
doi:10.1007/BF02731117.
- Sattler, J. M. (2008). *Assessment of children: cognitive foundations* (5. utg.). San Diego: Sattler Publishing Inc.
- Semel, E., Wiig, E. H., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals, fourth edition, Norwegian version*. Bromma, Sweden: Pearson Assessment.
- Sherwell, S., Reid, S. M., Reddihough, D. S., Wrennal, J., Ong, B., & Stargatt, R. (2014). Measuring intellectual ability in children with cerebral palsy: can we do better? *Research in Developmental Disabilities*, 35(10), 2558-2567.
doi:10.1016/j.ridd.2014.06.019

- Siegel, E. & Cress, C. J. (2002). Overview of the emergence of early AAC behaviors: progression from communication to symbolic skills. I J. Reichle, D. R. Beukelman & J. C. Light (Red.), *Exemplary practices for beginning communicators. Implications for AAC* (s. 25-52) . Baltimore: Paul H Brookes Publishing Co.
- Sigurdardottir, S., Eiriksdottir, A., Gunnarsdottir, E., Meintema, M., Arnadottir, U., & Vik, T. (2008). Cognitive profile in young Icelandic children with cerebral palsy. *Developmental Medicine & Child Neurology*, 50(5), 357-362. doi:10.1111/j.1469-8749.2008.02046.x
- Simonsen, H. G. & Theil, R. (2011). Morfologi. I K. E. Kristoffersen, H. G. Simonsen, A. Sveen (Red.), *Språk; en grunnbok*. (s. 249-294). Oslo: Universitetsforlaget
- Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, 34(1), 3-23. doi: 10.1177/01427237135109
- Slocum-Gori, S. L. & Zumbo, B. D. (2011). Assessing the unidimensionality of psychological scales: Using multiple criteria from factor analysis. *Social Indicators Research*, 102(3), 443-461. doi:10.1007/s11205-010-9682-8
- Spillane, M. M., Ross, K. K., & Vasa, S. F. (1996). A comparison of eye – gaze and standard response mode on the PPVT – R. *Psychology in the Schools*, 33(4), 265-271. doi:10.1002/(sici)1520-6807(199610)33:4<265::aid-pits1>3.3.co;2-h
- Stadskleiv, K., Schiørbeck, H., Seeland, L., & Lillehaug, H. A. (2014). *Assessing the “non assessable” children – a case study*. Paper presentert på The 16th Biennial Conference of the International Society for Augmentative and Alternative Communication, Lisboa, Portugal.
- Stadskleiv, K. (2015). Kartlegging. I K-A. Bottegaard Næss & A. V. Karlsen (Red.), *God kommunikasjon med ASK-brukere* (s. 73-118). Bergen: Fagbokforlaget.
- Stadskleiv, K., Vik, T., Andersen, G. L. & Lien, E. (2015). Is cognition assessed appropriately in children with cerebral palsy? *Developmental Medicine & Child Neurology*, 57(4), 23–23. doi:10.1111/dmcn.12778_52
- Straub, K. & Obrzut, J. E. (2009). Effects of Cerebral Palsy on Neuropsychological function. *Journal of Developmental and Physical Disabilities*, 21(2), 153-167. doi:10.1007/s10882-009-9130-3

- Strohner, H. & Nelson, K. (1974). The young child's development of sentence comprehension: Influence of event probability, nonverbal context, syntactic form, and strategies. *Child Development*, 45(3), 567-576. doi:10.2307/1127821
- Surveillance of Cerebral Palsy in Europe. (2000). Surveillance of Cerebral Palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Developmental Medicine and Child Neurology*, 42(12), 816-824. doi:10.1111/j.1469-8749.2000.tb00695.x
- Surveillance of Cerebral Palsy in Europe. (2002). Prevalence and characteristics of children with cerebral palsy in Europe. *Developmental Medicine & Child Neurology*, 44(9), 633-640. doi:1469-8749.2002.tb00848.x
- Tetzchner, S. von & Martinsen, H. (2002). *Alternativ og supplerende kommunikasjon: Innføring i tegnopplæring og bruk av kommunikasjons hjelpemidler for mennesker med språk- og kommunikasjonsvansker*. Oslo: Gyldendal Akademisk.
- Tideman, E., & Gustafsson, J. E. (2004). Age-related differentiation of cognitive abilities in ages 3-7. *Personality and Individual Differences*, 36(8), 1965-1974. doi:10.1016/j.paid.2003.09.004
- Tobii Technologies. (2018). *Slik fungerer øyestyring*. Hentet fra <http://www.tobiidynavox.no/slik-fungerer-oyestyring/>
- Tomasello, M. (2003). *Constructing a Language. A usage-based theory of language acquisition*. London: Harvard University Press.
- Tomblin, J. B. (2008). Validating diagnostic standards for specific language impairment using adolescent outcomes. I C. F. Norbury, J. B. Tomblin, & D. V. M. Bishop (Red.), *Understanding developmental language disorders* (s. 93-114). Hove: Psychology Press.
- 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
- Tomblin, J. B., Nippold, M. A., Fey, M. E., & Zhang, X. (2014). The character and course of individual differences in spoken language. I J. B. Tomblin & M. A. Nippold (Red.), *Understanding individual differences in language development across the school years* (s. 47-78). New York: Psychology Press.
- Utdanningsdirektoratet. (2017) *Rammeplan for barnehagen: Innhold og oppgaver*. Hentet fra: <https://www.udir.no/laring-og-trivsel/rammeplan/pedagogisk-virksomhet/ekstra-stotte/>

- van der Lely, H. K. J., Rosen, S. & McClelland, A. (1997). Evidence for a grammar-specific deficit in children. *Current Biology*, *19*(23), 1253-1258.
doi:10.1016/S0960-9822(07)00534-9
- van der Lely, H. K. J. (2005). Domain-specific cognitive systems: insight from grammatical SLI. *Trends in Cognitive Sciences*, *9*(2), 53-59. doi: 10.1016/j.tics.2004.12.002
- Voorman, J. M., Dallmeijer, A. J., Van Eck, M., Schuengel, C., & Becher, J. G. (2010). Social functioning and communication in children with cerebral palsy: association with disease and characteristics and personal and environmental factors. *Developmental Medicine & Child Neurology*, *52*(5), 441-447.
doi:10.1111/j.1469-8749.2009.03399.x
- Vos, R. C., Dallmeijer, A., Verhoef, M., Van Schie, P. E., Voorman, J. M., Wiegerink, D. J., ... Becher, J. G. (2014). Developmental trajectories of receptive and expressive communication in children and young adults with cerebral palsy. *Developmental Medicine & Child Neurology*, *56*(10), 951-959. doi:10.1111/dmcn.12473
- Wechsler, D. (2015). *Wechsler Preschool and Primary Scale of Intelligence, fourth edition, Norwegian version*. Bromma, Sweden: Pearson Assessment.

Vedlegg 1: Samtykkeskjema



Forespørsel om deltakelse i forskningsprosjektet

"C-BiLLT prosjektet"

Bakgrunn

Dette er et spørsmål til dere om å la deres barn delta i en forskningsstudie der formålet er å utvikle en norskspråklig versjon av språkforståelsestesten C-BiLLT (The Computer-Based Instrument for Low Motor Language Testing). Språkforståelsestesten administreres på en PC, og er beregnet på barn i alderen 1,5 til 7,5 år.

Prosjektleder er psykologspesialist og ph.d. Kristine Stadsleiv. Studien gjennomføres som et samarbeid mellom Seksjon for nevrohabilitering –barn ved Oslo Universitetssykehus og Institutt for spesialpedagogikk, Universitetet i Oslo.

Hva innebærer studien?

Studien innebærer at deres barn testes i barnehagen/skolen over 1-2 timer. Testingen vil foregå i januar og februar 2018. Testingen vil bli gjennomført av masterstudentene Sara Fiske eller Anne-Lise Haddeland, eller av prosjektleder.

Testingen vil foregå i et rolig rom, og en voksen som kjenner barnet og som barnet er trygg på kan være til stede i undersøkelsesrommet hele tiden.

Barna vil undersøkes med språkforståelsestesten C-BiLLT, samt med andre tester av språk og av ikkespråklig forståelse. Årsaken til at også andre tester brukes er at det vil gi oss informasjon om C-BiLLT måler det den skal, nemlig språkforståelse, og at den måler noe annet enn det vi får informasjon om på ikke-språklige tester. Det vil benyttes anerkjente instrumenter som allerede er normert på norsk. Avhengig av barnets alder vil språktestene Reynell, TROG-2 (Test for Reception of Grammar) og Ordgjenkjenningssoppgaven fra WPPSI-IV benyttes, samt de ikke-språklige oppgavene Terningmønster fra WPPSI-IV og Ravens fargematriser.

For barn i alderen 5,5 til 7,5 år vil vi teste barna med C-BiLLT to ganger, med to ukers mellomrom. Den ene gangen skal de besvare oppgavene med å peke med hånden direkte på PC-skjermen, mens de den andre gangen skal bruke databasert blikkepekestyr til å besvare oppgaven med. Utstyret fanger opp hvor på dataskjermen et barn ser, og brukes vanligvis av barn som ikke har mulighet til å bruke hendene til å peke med på grunn av store motoriske vansker.

For å undersøke egenskapene til testen C-BiLLT er det nødvendig å teste den ut på barn som ikke har noen kjente vansker. Det betyr at barna som deltar ikke kan ha vesentlige motoriske vansker (de må kunne peke med en hånd), ikke større syns- eller hørselsvansker enn at det kan korrigeres med briller/høreapparat, ikke ha hatt forsinket utvikling generelt eller forsinket språkutvikling, samt ha norsk som morsmål.

Mulige fordeler og ulemper

Deres barn vil ikke ha noen spesielle fordeler av studien, men erfaringer fra studien vil senere kunne hjelpe barn med språkvansker, samt barn med store tale- og bevegelsesvansker.

Dere vil ikke få noen tilbakemelding på testresultatene så fremt de ligger innenfor det normale variasjonsområdet sett i forhold til barnets alder. Det normale variasjonsområdet er definert som resultater over 16. persentil. Dersom barnet får resultater under dette området vil dere få tilbud om en samtale med prosjektleder.

Hva skjer med prøvene og informasjonen om ditt barn?

Opplysninger fra testene skal kun brukes slik som beskrevet i hensikten med studien. Alle opplysningene vil bli behandlet uten navn og fødselsnummer. En kode knytter barnet til opplysningene gjennom en navneliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til navnelisten og som kan finne tilbake til barnet. Det vil ikke være mulig å identifisere barnet i resultatene av studien når disse publiseres.

Hvis dere sier ja til å delta i studien, har dere rett til å få innsyn i hvilke opplysninger som er registrert om dere. Dere har videre rett til å få korrigert eventuelle feil i de opplysningene vi har registrert. Dersom dere trekker barnet fra studien, kan dere kreve å få slettet innsamlede opplysninger. Opplysningene blir senest slettet 2020.

Frivillig deltakelse

Det er frivillig å delta i studien. Dersom dere ikke ønsker at deres barn skal delta, trenger dere ikke å oppgi noen grunn.

Dersom dere ønsker at deres barn skal delta, undertegner dere samtykkeerklæringen på neste side. Om dere nå sier ja til deltakelse på vegne av barnet, kan dere senere trekke tilbake samtykket uten å oppgi noen grunn. Dersom dere senere ønsker å trekke tilbake samtykket, kan psykolog Kristine Stadskleiv, tlf 22 11 76 35, kontakte s.

Samtykke for deltakelse i studien «C-BiLLT prosjektet»

Gjelder _____, f. _____

i avdeling/klasse _____ i/ ved _____ barnehage/skole.

Opplysninger om barnet:

Barnets fødselsdato (dd.mm.åååå) _____

	Stemmer	Stemmer ikke
Normalt syn uten bruk av briller	<input type="checkbox"/>	<input type="checkbox"/>
Ser godt ved bruk av briller	<input type="checkbox"/>	<input type="checkbox"/>
Normal hørsel uten bruk av høreapparat	<input type="checkbox"/>	<input type="checkbox"/>
Hører godt ved bruk av høreapparat	<input type="checkbox"/>	<input type="checkbox"/>
Barnet har ikke større motoriske vansker enn at det kan peke (<i>barnet må kunne peke for å svare på testene</i>)	<input type="checkbox"/>	<input type="checkbox"/>
Har ikke hatt vesentlig forsinket utvikling på noen områder	<input type="checkbox"/>	<input type="checkbox"/>
Har ikke hatt vesentlig forsinket språkutvikling	<input type="checkbox"/>	<input type="checkbox"/>
Har en kronisk sykdom. Hvis ja, spesifiser gjerne: _____	<input type="checkbox"/>	<input type="checkbox"/>
Har en nevrologisk tilstand/sykdom	<input type="checkbox"/>	<input type="checkbox"/>
Har norsk som morsmål	<input type="checkbox"/>	<input type="checkbox"/>
Har andre morsmål i tillegg. Hvis ja, spesifiser: _____	<input type="checkbox"/>	<input type="checkbox"/>

Vi er innforstått med at det ikke rutinemessig gis tilbakemelding på testresultatene, men at vi kontaktes og tilbys en samtale dersom barnet får noe lavere skårer enn forventet for alderen. I så fall ønsker vi at invitasjon til samtale skal sendes til følgende adresse:

Jeg er villig til å la mitt barn delta i studien og at barnehagen/skolen kan kontaktes for å avtale tid for testing

 (Signert av pårørende, på vegne av prosjektdeltaker, dato)

Samtykket kan leveres i barnehage/skolen eller returneres til:

 Oslo Universitetssykehus
 Seksjon for nevrohabilitering – barn/Ullevål
 Ved psykologspesialist Kristine Stadskleiv
 Postboks 4956, Nydalen
 0424 Oslo

Vedlegg 2: Informasjonsskriv



Oslo universitetssykehus HF
Ullevål sykehus
Postboks 4956 Nydalen
0424 Oslo

Sentralbord: 02770

Til foresatte til barn i alderen 1,5 til 7,5 år

Barne- og ungdomsklinikken
Barneavdeling for nevrofag
Seksjon for nevrohabilitering - barn US

Vår ref.: C-BiLLT prosjektet
Oppgis ved all henvendelse
Deres ref.:
Saksbeh.: Kristine Stadskleiv
Dato: 31.8.2017

Seksjon for nevrohabilitering – barn (barnehabiliteringen) ved Oslo Universitetssykehus har sammen med Institutt for spesialpedagogikk, Universitet i Oslo, et forskningsprosjekt der formålet er å utvikle en norskspråklig versjon av språkforståelsestesten C-BiLLT (The Computer-Based Instrument for Low Motor Language Testing).

Språkforståelsestesten administreres på en PC, og er beregnet på barn i alderen 1,5 til 7,5 år. Testen er særlig velegnet for å undersøke forståelse for ord og setninger hos barn som ikke har talespråk eller som har store motoriske vansker. Dette er en gruppe barn vi følger opp ved vår seksjon, og er årsaken til at sykehuset er involvert i testutviklingen.

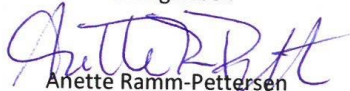
C-BiLLT testen er opprinnelig utviklet i Nederland, og blir nå oversatt til norsk. For å kunne bruke testen i klinisk sammenheng er vi avhengige av å vite at den måler det den skal på en sikker måte. For å finne ut dette trenger vi å undersøke barn som ikke har noen kjente vansker og som har norsk som morsmål. Dette er bakgrunnen for at dere har fått denne henvendelse gjennom barnets barnehage/skole.

Studien innebærer at deres barn testes i barnehagen/skolen over 1-2 timer. Testingen vil foregå i januar og februar 2018. Testingen vil bli gjennomført av masterstudentene Sara Fiske og Anne-Lise Haddeland, eller av prosjektleder. Prosedyren for testingen er beskrevet i vedlagte informasjonsskriv.

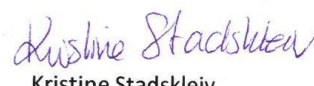
Dataene om barnet vil oppbevares på sykehusets forskningsserver. Data som kan identifisere barnet vil bli oppbevart adskilt fra testresultatene. Testresultatene vil kun benyttes til det som er beskrevet i informasjonsskrivet.

Prosjektleder er psykologspesialist og ph.d. Kristine Stadskleiv. Om det er spørsmål eller kommentarer kan hun kontaktes på tlf 22 11 76 35 eller mail kstadskl@ous-hf.no.

Med vennlig hilsen


Anette Ramm-Pettersen
Avdelingsleder


Petra Aden
Seksjonsleder


Kristine Stadskleiv
Prosjektleder



Vedlegg 3: Godkjenning fra NSD



Kristine Stadsleiv
Postboks 1140 Blindern
0318 OSLO

Vår dato: 02.10.2017

Vår ref: 55604 / 3 / AMS

Deres dato:

Deres ref:

Tilbakemelding på melding om behandling av personopplysninger

Vi viser til melding om behandling av personopplysninger, mottatt 31.08.2017.

Meldingen gjelder prosjektet:

55604	<i>Innovative assessment of language comprehension in children with severe motor impairments: development of the Norwegian version of the C-BiLLT test</i>
Behandlingsansvarlig	Universitetet i Oslo, ved institusjonens øverste leder
Daglig ansvarlig	Kristine Stadsleiv
Student	Sara Fiske

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget [skjema](#). Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en [offentlig database](#).

Personvernombudet vil ved prosjektets avslutning, 31.12.2018, rette en henvendelse angående status for behandlingen av personopplysninger.

Dersom noe er uklart ta gjerne kontakt over telefon.

Vennlig hilsen

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

Katrine Utaaker Segadal

Anne-Mette Somby

Kontaktperson: Anne-Mette Somby tlf: 55 58 24 10 / anne-mette.somby@nsd.no

Vedlegg: Prosjektvurdering

Kopi: Sara Fiske, sara.i.fiske@gmail.com

Vedlegg 4: Nederlandsk data

Tabell 18. Resultater fra den nederlandske faktoranalysen. Hentet fra Geytenbeek et al., 2014, s. 10

Seksjon*	Antall testledd	Identifikasjon av	faktorladning
1	10	Substantiv	.91 – .97
2	10	Verb	.93 – .97
3	10	Dyr, objekter og personer	.96 – .98
4	4	Vanskeligere substantiv	.93 – .96
5	5	Enkle setninger om gjenstander i kombinasjon med verb og preposisjoner	.90 – .96
6	5	”Hvem”-spørsmål om personer som utfører handlinger	.75 – .94
7	4	Passive setninger med objekter og preposisjoner kombinert med hendelser	.43** – .75
8	9	Setninger med mer grammatisk og syntaktisk kompleksitet	.56 – .97
9	6	Enkle aktive setninger som refererer til ikke-observerbare situasjoner med fire personer	.75 – .88
10	9	Komplekse setninger med to eller flere begreper	.42** – .94
11	4	Sammensatte, komplekse setninger	.63 – .89

* Faktoranalysen ble utført før Seksjon 12 var utformet og lagt til i C-BiLLT

** Kun testledd 45 i Seksjon 7 og testledd 70 i Seksjon 10 hadde en faktorladning <.50

Tabell 19. Resultater fra den nederlandske variansanalysen. Hentet fra Geytenbeek et al., 2014, s. 10

Alder	M forskjell	95% konfidensintervall
2 - 1	22.41*	19.31 – 25.51
3 - 2	11.66*	9.46 – 13.85
4 - 3	6.43*	4.43 – 8.44
5 - 4	3.20*	1.40 – 5.01
6 - 5	3.25*	1.08 – 5.43
7 - 6	1.54	4.37 – 7.46

* M forskjell = gjennomsnittlig forskjell i skårer mellom helårsgrupper

** $p < .001$

Vedlegg 5: Korrelasjoner mellom seksjoner

Tabell 20. Korrelasjoner mellom de 12 seksjonene i C-BiLLT (N = 187)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
S1	-											
S2	.68*	-										
S3	.41*	.42*	-									
S4	.51*	.59*	.63*	-								
S5	.62*	.55*	.44*	.55*	-							
S6	.42*	.67*	.39*	.53*	.67*	-						
S7	.40*	.37*	.39*	.39*	.55*	.53*	-					
S8	.54*	.61*	.45*	.51*	.71*	.79*	.70*	-				
S9	.38*	.45*	.34*	.40*	.58*	.67*	.58*	.75*	-			
S10	.44*	.47*	.39*	.42*	.66*	.67*	.69*	.81*	.79*	-		
S11	.38*	.41*	.38*	.41*	.62*	.62*	.63*	.73*	.66*	.78*	-	
S12	.31*	.34*	.31*	.32*	.52*	.54*	.63*	.67*	.66*	.71*	.67*	-

Note. S1 = Seksjon 1, S2 = Seksjon 2, S3 = Seksjon 3, ..., S12 = Seksjon 12

* $p < .001$

Vedlegg 6: Gjennomsnittlig riktig svarprosent for alle testledd i C-BiLLT

Seksjon	Testledd	<i>N</i>	<i>M</i> %	<i>SD</i>
Del 1				
Seksjon 1	1: <i>Hvor er bilen?</i>	187	99	7.3
	2: <i>Hvor er ballen?</i>	187	99	7.3
	3: <i>Hvor er tåteflasken?</i>	187	99	7.3
	4: <i>Hvor er dukken?</i>	187	99	7.3
	5: <i>Hvor er skjeen?</i>	187	100	0.0
	6: <i>Hvor er TV'en?</i>	187	98	14.5
	7: <i>Hvor er lampen?</i>	187	99	7.3
	8: <i>Hvor er skoen?</i>	187	100	0.0
	9: <i>Hvor er gryta?</i>	187	97	16.2
	10: <i>Hvor er nøklene?</i>	187	98	12.6
Seksjon 2	11: <i>Hvem spiser?</i>	187	99	10.3
	12: <i>Hvem drikker?</i>	187	100	0.0
	13: <i>Hvem sover?</i>	187	98	12.6
	14: <i>Hvem vasker seg?</i>	187	98	12.6
	15: <i>Hvem klemmer?</i>	187	96	19.0
	16: <i>Hvem leker?</i>	187	99	10.3
	17: <i>Hvem skjærer?</i>	187	95	22.6
	18: <i>Hvem husker?</i>	187	99	10.3
	19: <i>Hvem kler på seg?</i>	187	99	10.3
	20: <i>Hvem tegner?</i>	187	98	12.6
Seksjon 3	21: <i>Hvor er stolen?</i>	187	99	7.3
	22: <i>Hvor er bordet?</i>	187	99	10.3
	23: <i>Hvor er sengen?</i>	187	100	0.0
	24: <i>Hvor er hunden?</i>	187	100	0.0
	25: <i>Hvor er anda?</i>	187	99	10.3
	26: <i>Hvor er støvsugeren?</i>	187	99	7.3
	27: <i>Hvor er telefonen?</i>	187	99	10.3
	28: <i>Hvor er babyen?</i>	187	100	0.0
	29: <i>Hvor er krana?</i>	187	94	23.6
	30: <i>Hvor er gutten?</i>	187	99	7.3

Del 2				
Seksjon 4	31: <i>Hvor er trommen?</i>	187	96	19.0
	32: <i>Hvor er postkassen?</i>	187	98	14.5
	33: <i>Hvor er ryggsekken?</i>	187	98	12.6
	34: <i>Hvor er paraplyen?</i>	187	97	17.7
Seksjon 5	35: <i>Hva kan du sove i?</i>	187	94	24.6
	36: <i>Hva kan du kjøre i?</i>	187	91	28.0
	37: <i>Hva kan du tørke med?</i>	187	91	28.0
	38: <i>Hunden ligger i kurven.</i>	187	84	36.0
	39: <i>Boka står på asjetten.</i>	187	79	40.7
Seksjon 6	40: <i>Hvem spiller musikk?</i>	187	92	27.2
	41: <i>Hvem bærer noe?</i>	187	93	25.5
	42: <i>Hvem skal lufte hunden?</i>	187	69	46.4
	43: <i>Hvem har handlet?</i>	187	83	37.3
	44: <i>Hvem kjører?</i>	187	91	28.0
Seksjon 7	45: <i>Telefonen er lagt foran kurven.</i>	187	42	49.4
	46: <i>Skoene er satt på plass i hylla.</i>	187	78	41.5
	47: <i>Gutten blir dyttet av jenta.</i>	187	68	46.8
	48: <i>Jenta blir dyttet av mannen.</i>	187	83	37.3
Seksjon 8	49: <i>Hvilken ball er minst?</i>	187	89	31.7
	50: <i>Hvor er den gule ballen?</i>	187	95	21.5
	51: <i>Hvor er det lengste røde toget?</i>	187	85	35.8
	52: <i>Alle de røde nøklene er i koppen.</i>	187	74	44.1
	53: <i>Den svarte nøkkelen ligger under koppen.</i>	187	88	32.3
	54: <i>Den lille røde tannbørsten står i midten.</i>	187	78	41.8
	55: <i>En av tannbørstene som er satt på plass i koppen, er lang.</i>	187	68	46.6
	56: <i>Av de tre røde nøklene, er to tatt ut av koppen</i>	187	68	46.8
	57: <i>Hvilke tannbørster er ikke satt på plass i koppen?</i>	187	55	49.9
Seksjon 9	58: <i>Jakob og Emma skal leke ute, hvem blir hos mamma?</i>	187	70	46.2
	59: <i>Hvem pleide å leke før, men gjør det ikke nå lenger?</i>	187	43	49.6
	60: <i>Hvem kommer til å leke ute senere, men gjør det ikke ennå?</i>	187	48	50.1
	61: <i>Hvem skal leke ute sammen med Emma?</i>	187	75	43.5

	62: <i>Mamma spiser Emmas kjeks, hvem blir sint?</i>	187	70	45.9
	63: <i>Emma og mamma skifter bleie på babyen. Hvem er det som hjelper mamma?</i>	187	79	41.1
Seksjon 10	64: <i>To like pålegg ligge ved siden av hverandre.</i>	187	81	39.5
	65: <i>Halvparten av brunosten er spist opp.</i>	187	39	48.8
	66: <i>På det største syltetøyglasset kan du se forsiden.</i>	187	69	46.9
	67: <i>Bak kluten står det ingen syltetøyglass.</i>	187	68	46.8
	68: <i>Et lite syltetøyglass står ved siden av det røde syltetøyglasset.</i>	187	59	49.3
	69: <i>Kluten ligger, og <u>ett</u> av syltetøyglassene står i kurven.</i>	187	68	46.6
	70: <i><u>Også</u> det gule syltetøyglasset står foran brunosten.</i>	187	28	44.9
	71: <i>Alle syltetøyglassene står rundt kurven.</i>	187	66	47.4
	72: <i>Alt pålegget bortsett fra det gule syltetøyglasset, er borte.</i>	187	55	49.9
Seksjon 11	73: <i>Ikke på den gule, men på den blå asjetten ligger det et stykke ost og et eple.</i>	187	76	42.5
	74: <i>Brunosten står på den blå asjetten og syltetøyglasset står ved siden av den gule asjetten.</i>	187	52	50.1
	75: <i>Først var det et eple på bordet, men nå er det en banan der i stedet.</i>	187	61	48.8
	76: <i>Bordet er nesten ferdig dekket, men syltetøyglasset og osten mangler fortsatt.</i>	187	45	49.9
Seksjon 12	77: <i>All frukten er lagt på asjettene. Bare på de <u>røde</u> asjettene ligger de samme sortene frukt, mens disse sortene ikke ligger på den blå eller den gule asjetten.</i>	187	38	48.7
	78: <i>Med unntak av én asjett, så er det servert mat på alle asjettene.</i>	187	28	44.9
	79: <i>Alt pålegget er lagt på asjettene, men det er bare på de asjettene med samme farge at det ligger ost.</i>	187	22	41.8
	80: <i>Alle er ferdig med å spise, men på to asjetter som har forskjellig farge ligger det fortsatt avfall.</i>	187	43	49.6

81: <i>Flertallet av menneskene ved bordet har fått druer, men klasen ble ikke likt fordelt da personen som spiser fra den blå asjetten bare har fått en liten del av klasen.</i>	187	10	29.6
82: <i>Eplene ligger på asjetten som står til venstre for de røde asjettene.</i>	187	29	45.4
83: <i>Jakob har allerede kiwiskiver på sin asjett, mens Emma ennå ikke har skrelt sin.</i>	187	48	50.1
84: <i>Hvis Jakob har lagt kiwiskiver på Emmas asjett, men Emma har lagt dem og sin egen frukt på asjetten til høyre for seg. På hvilken asjett har Emma lagt frukten da?</i>	187	12	32.2
85: <i>Jakob, som tok asjetten til Emma, har lagt på to bananer og en drueklase.</i>	187	40	49.1
86: <i>Fordi nesten alle asjettene allerede er stablet, har Jakob ikke lagt eplet som er skrelt eller bananene på stabelen.</i>	187	30	46.2
