



UiT Norges arktiske universitet

Institutt for lærerutdanning og pedagogikk

Forholdet mellom vokabular og nonverbal resonnering

Er det en gjensidig påvirkning mellom dem over tid?

Anja Sofie Richardsen

Masteroppgave i spesialpedagogikk, PED-3901, mai 2024

Sammendrag

Gjensidighetsteorien er en av flere teorier som prøver å forklare kognitiv utvikling. Den foreslår at ulike kognitive ferdigheter, som eksempelvis vokabular og nonverbal resonnering, gjensidig påvirker hverandres utvikling. Tidligere studier, noen med et tydelig spesialpedagogisk fokus, peker i ulike retninger om hvorvidt de støtter gjensidighetsteorien. Det er ikke tidligere gjort studier som undersøker teorien i et norsk utvalg.

I denne masteroppgaven blir gjensidighetsteorien testet i et utvalg totalt bestående av 215 norske barn. Data fra da barna var 4 og 6 år brukes for å undersøke hvorvidt det er gjensidige forhold mellom utviklingen av vokabular (BPVS) og nonverbal resonnering (Terningmønster). Analysemetodene i undersøkelsen er *Cross-Lagged Panel Models* (CLPM) og *Latent Change Score Models* (LCSM). Ved å dele opp utvalget etter hvordan deltakerne skårer på BPVS og Terningmønster, vil undersøkelsen både teste gjensidighetsteorien i det totale utvalget (CLPM og LCSM) og hos barna med utfordringer knyttet til vokabular og nonverbal resonnering (CLPM).

Gjennom analysene i denne masteroppgaven er det ikke funnet støtte for gjensidighetsteorien, hverken i analysene av hele utvalget eller i analysene av barna med lave skårer innenfor vokabular eller nonverbal resonnering. Det er likevel identifisert noen svakheter i analysene som gjør at særlig resultatene knyttet til deltakerne med lave skårer må tolkes med noe forsiktighet. Funnene i masteroppgaven har likevel viktige teoretiske og praktiske implikasjoner som vil bli diskutert nedenfor.

Innholdsfortegnelse

1	Innledning.....	1
1.1	Begrepsintroduksjon.....	1
1.2	Begrunnelse for valg av tema.....	2
1.3	Hensikt og problemstillinger.....	3
1.4	Masteroppgavens struktur.....	3
2	Teoretisk og empirisk bakgrunn.....	5
2.1	Vokabular.....	5
2.2	Nonverbal resonnering.....	7
2.3	Vansker tilknyttet vokabular og nonverbal resonnering.....	9
2.3.1	Språkforstyrrelser.....	9
2.3.2	Utviklingshemming og generelle lærevansker.....	12
2.4	Teorier om positive manifold og kognitiv utvikling.....	13
2.4.1	Gjensidighetsteorien.....	14
2.4.2	<i>G</i> -faktor-teorien.....	17
2.4.3	Investeringsteorien.....	18
2.4.4	Samplingteorien.....	20
2.4.5	Samlet oversikt over de ulike teoriene.....	21
2.4.6	Oversikt over kausalitet i teoriene.....	22
2.5	Oppsummering.....	23
3	Metode.....	25
3.1	Kvantitativ metode og vitenskapsteoretisk forankring.....	25
3.2	Forskningsetikk.....	27
3.3	Utvalg.....	27
3.4	Studiedesign.....	28
3.5	Datainnsamling.....	29
3.6	Testinstrumenter.....	29

3.6.1	Vokabular	29
3.6.2	Nonverbal resonnering	29
3.7	Definering og identifisering av «vanskegrupper»	30
3.8	Korrelasjon og lineær regresjon	31
3.9	Cross-Lagged Panel Models.....	32
3.10	Latent Change Score Models	35
3.11	Validitet.....	36
4	Resultater.....	39
4.1	Deskriptiv statistikk og forberedende analyser	39
4.1.1	Hele utvalget	39
4.1.2	Vanskegrupper	41
4.2	Lar gjensidighetsteorien seg replikere på et norsk utvalg?	42
4.2.1	CLPM hele utvalget	42
4.2.2	LCSM hele utvalget	44
4.2.3	Sammenligning av modeller.....	45
4.3	Er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?	46
4.3.1	CLPM vanskegruppe vokabular.....	46
4.3.2	CLPM vanskegruppe nonverbal resonnering.....	47
5	Diskusjon.....	49
5.1	Vurdering av validitet.....	49
5.1.1	Statistisk validitet	49
5.1.2	Indre validitet	51
5.1.3	Begrepsvaliditet.....	53
5.1.4	Ytre validitet.....	55
5.2	Teoretiske implikasjoner	55
5.3	Praktiske implikasjoner	58

5.4	Mulige forklaringer på at funnene til Kievit og kolleger ikke har latt seg replikere	60
5.5	Videre forskning	61
5.6	Oppsummering	63
	Referanseliste	65

Tabelliste

Tabell 1:	Oversikt over teorier om kognitiv utvikling og positive manifold	22
Tabell 2:	Deskriptiv statistikk og psykometriske egenskaper for BPVS og Terningmønster (hele utvalget)	40
Tabell 3:	Korrelasjonstabell med BPVS og Terningmønster	40
Tabell 4:	Variabelverdier på 15. persentil	41
Tabell 5:	Deskriptiv statistikk og psykometriske egenskaper for BPVS og Terningmønster (vanskegrupper)	42
Tabell 6:	CLPM BPVS og Terningmønster, resultater fra hele utvalget	43
Tabell 7:	LCSM BPVS og Terningmønster, resultater fra hele utvalget	44
Tabell 8:	Sammenstilling av resultater fra LCSM og CLPM	45
Tabell 9:	CLPM BPVS og Terningmønster, resultater fra vanskegruppe vokabular	46
Tabell 10:	CLPM BPVS og Terningmønster, resultater fra vanskegruppe nonverbal resonnering	47

Figurliste

Figur 1:	Terminologi om språkforstyrrelser, omarbeidet etter CATALISE-studien (Bishop et al., 2017)	10
Figur 2:	Kausalitet i de ulike teoriene	23
Figur 3:	Cross-Lagged Panel Model (CLPM)	33
Figur 4:	Vokabular og nonverbal resonnering i CLPM	34
Figur 5:	Latent Change Score Models	36
Figur 6:	CLPM BPVS og Terningmønster, resultater fra hele utvalget	43
Figur 7:	LCSM BPVS og Terningmønster, resultater fra hele utvalget	45
Figur 8:	CLPM BPVS og Terningmønster, resultater fra vanskegruppe vokabular	47

Figur 9: CLPM BPVS og Terningmønster, resultater fra vanskegruppe nonverbal resonnering
..... 48

Forord

For meg markerer denne masteroppgaven slutten på mange år som student innenfor flere forskjellige fagfelt. Når jeg nå omsider er ferdig, er det flere jeg vil takke.

Jeg vil først takke UiT Norges arktiske universitet for muligheten jeg har fått til å lære veldig mye forskjellig i løpet av min studietid.

Til min masteroppgave har jeg fått bruke et datasett som kommer fra Institutt for spesialpedagogikk ved Universitetet i Oslo. Det er jeg veldig glad for, og jeg er svært takknemlig for alt arbeidet som har blitt lagt i innsamlingen av dette datamaterialet. Men ikke minst er jeg takknemlig for at alle barna (nå unge voksne) i studien faktisk har deltatt i dette prosjektet gjennom store deler av sin barndom.

Jeg vil også gi min varmeste takk til mine veiledere, Monica Melby-Lervåg og Trude Nergård Nilssen. Dere har hjulpet meg med å velge et utfordrende og spennende tema for masteroppgaven, og delt av deres enorme kunnskap og erfaring. Samtidig har dere vært veldig oppmuntrende og støttende hele veien, noe jeg har satt stor pris på. Jeg kunne ikke bedt om bedre veiledning. Tusen, tusen takk!

1 Innledning

Mange ulike teorier har gjennom lang tid forsøkt å forklare barns kognitive utvikling. Innenfor spesialpedagogikk er slike teorier interessante blant annet fordi de kan bidra til økt forståelse av lærevansker og tiltak. I dette prosjektet utforskes en slik teori; masteroppgaven undersøker hvorvidt det er en gjensidig påvirkning mellom vokabular og nonverbal resonnering hos barn over tid. Dette spørsmålet har både teoretiske og praktiske implikasjoner, og har engasjert forskere innenfor psykologi og læring i en årrekke. I denne masteroppgaven skal spørsmålet utforskes kvantitativt med et datasett som kommer fra en prospektiv longitudinell studie ved Institutt for spesialpedagogikk ved Universitetet i Oslo (Hjetland et al., 2019). Aller først kreves imidlertid en begrepsintroduksjon.

1.1 Begrepsintroduksjon

Teorien om at det eksisterer et gjensidig forhold mellom blant annet vokabular og nonverbal resonnering har på engelsk det etablerte begrepet «mutualism theory». Heretter vil jeg kalle det «gjensidighetsteorien». Et sentralt element innenfor denne teorien er at alle kognitive ferdigheter viser en viss grad av sammenheng med hverandre, selv om de tilhører ulike domener. Dette er et fenomen som ble oppdaget av Charles Spearman (1904), og det har siden blitt underbygget gjentatte ganger i flere ulike undersøkelser (Kievit et al., 2017). På engelsk er et etablert begrep for fenomenet «positive manifold». Ordet manifold betyr «grenrør», også på norsk, og på grunn av den språklige likheten vil det engelske begrepet brukes også her. Positive manifold sier altså at det er en positiv korrelasjon mellom ulike kognitive ferdigheter hos et individ. I henhold til gjensidighetsteorien er årsaken til korrelasjonen at ulike kognitive ferdigheter påvirker hverandres utvikling (van der Maas et al., 2006).

Man kan studere gjensidighetsteorien med utgangspunkt i ulike typer kognitive ferdigheter. De kognitive ferdighetene det er fokus på her, er vokabular og nonverbal resonnering. I denne masteroppgaven vil begrepet «kognitive ferdigheter» brukes som fellesbetegnelse for ulike ferdigheter som vokabular og nonverbal resonnering. «Intelligens» er dessuten et nært relatert begrep. Dette begrepet vil brukes når det er snakk om målt IQ-skår, eller når det inngår som en del av andre etablerte begreper (for eksempel flytende intelligens og krystallisert intelligens).

1.2 Begrunnelse for valg av tema

De siste årene har gjensidighetsteorien vært mye omdiskutert i internasjonal forskning. Det skyldes i hovedsak en serie undersøkelser gjennomført av Kievit og kolleger (Griffiths et al., 2022; Kievit et al., 2019; Kievit et al., 2017). Gjennom kartlegging av vokabular og nonverbal resonnering hos ungdommer og unge voksne, konkluderer Kievit et al. (2017) med at gjensidighetsteorien forklarer positive manifold bedre enn andre teorier. Senere ble resultatene replikert med yngre barn (6–8 år) som utvalg (Kievit et al., 2019). I tillegg viser en nyere studie at gjensidighetsteorien beskriver utviklingen av vokabular og nonverbal resonnering hos barn både med og uten språkforstyrrelser (Griffiths et al., 2022). Andre studier har sett på gjensidighetsteorien med andre typer kognitive ferdigheter som utgangspunkt (se for eksempel Peng et al., 2019). Per i dag har det likevel ikke blitt gjort studier som undersøker om gjensidighetsteorien kan forklare utviklingen av vokabular og nonverbal resonnering, eller andre typer kognitive ferdigheter, i et norsk utvalg.

Samtidig peker undersøkelser i ulike retninger når det gjelder hvorvidt de støtter gjensidighetsteorien. Det finnes blant annet forskning som ikke finner støtte for denne forklaringsmodellen, eller som støtter andre forklaringer på positive manifold og kognitiv utvikling, for eksempel Gignac (2014, 2016). Siden resultatene knyttet til gjensidighetsteorien altså ikke er entydige, er det et behov for ytterligere forskning.

Siden fokuset i denne masteroppgaven er spesialpedagogikk, er det også ønskelig å teste hvorvidt teorien er gyldig for barn som i utgangspunktet har utfordringer med språk og læring. Som nevnt har andre studier allerede sett på hvorvidt gjensidighetsteorien er holdbar for barn med slike utfordringer (Griffiths et al., 2022), men de fleste testene av gjensidighetsteorien har likevel ikke hatt dette fokuset. Det er heller ikke utforsket blant norske barn med utfordringer med språk og nonverbal resonnering.

Masteroppgaven vil derfor både ha teoretisk og praktisk betydning. Teoretisk fordi den vil bidra til forståelsen av kognitiv utvikling, språkutvikling inkludert. Praktisk fordi en slik forståelse kan gi føringer for hvordan man kan arbeide med å fremme denne utviklingen hos barn og unge, også i spesialpedagogiske sammenhenger. Med andre ord kan masteroppgaven bidra til økt kunnskap om spesialpedagogiske tiltak. Siden gjensidighetsteorien omfatter utviklingen av alle typer kognitive ferdigheter, vil masteroppgavens teoretiske og praktiske betydning dessuten strekke seg utover vokabular og nonverbal resonnering, selv om det er disse ferdighetene det i hovedsak skal handle om her.

1.3 Hensikt og problemstillinger

Hensikten med denne masteroppgaven er å altså undersøke om funnene internasjonalt som støtter gjensidighetsteorien for kognitiv utvikling lar seg replikere, og utvide, til et norsk utvalg. I tillegg vil det også bli sett på om funnene er overførbare til de som har utfordringer med språk og læring.

Undersøkelsen har følgende problemstillinger:

1. *I hvilken grad lar gjensidighetsteorien seg replikere på et norsk utvalg?*
2. *I hvilken grad er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?*

1.4 Masteroppgavens struktur

Her i kapittel 1 har temaet for masteroppgaven blitt introdusert og begrunnet. Kapittelet består også av en begrepsavklaring av de mest sentrale begrepene i oppgaven. I tillegg har prosjektets hensikt og problemstillinger blitt presentert.

Kapittel 2 består av masteroppgavens teoretiske og empiriske bakgrunn. Der vil fokuset først ligge på å få frem det som er relevant for masteroppgaven når det gjelder vokabular og nonverbal resonnering. I tillegg følger en redegjørelse for spesialpedagogiske vansker som er knyttet til disse ferdighetene. Kapittelet består også av en gjennomgang av gjensidighetsteorien og teorier som er mulige alternativer til denne.

Metoden for prosjektet blir presentert i kapittel 3. Der inngår beskrivelser av datasettet som er brukt, herunder beskrivelser av utvalget, studiedesignet, datainnsamlingsprosedyren og testinstrumenter. I tillegg omtales viktige momenter tilknyttet masterprosjektets statistiske analyser, og analyseteknikkene *Cross-Lagged Panel Models* og *Latent Change Score Models* vil særlig bli gjort rede for. Kapittelet omfatter også en redegjørelse for vitenskapsteori i kvantitativ metode, forskningsetikk og validitet.

Resultatene av prosjektets statistiske analyser presenteres i kapittel 4. Først presenteres deskriptiv statistikk og forberedende analyser, før resultatene av analysene knyttet direkte opp mot problemstillingene blir gått gjennom.

I kapittel 5 diskuteres de teoretiske og praktiske implikasjonene av funnene i masteroppgaven, i tillegg til at det pekes på forslag til videre forskning. Det vil da særlig legges vekt på

implikasjoner og videre forskning relatert til spesialpedagogikk. I tillegg diskuteres resultatene i denne masteroppgaven opp mot svært sentrale tidligere studier av gjensidighetsteorien. I kapittelet inngår også en diskusjon av validitet i prosjektet. Diskusjonen av validiteten vil foregå med utgangspunkt i et validitetssystem (Shadish et al., 2002) som først vil presenteres i kapittel 3.

Det gjøres oppmerksom på at masteroppgaven er skrevet med utgangspunkt i en projektskisse som også ble levert som eksamensbesvarelse i PED-3055 våren 2023. Dette innebærer at det vil være noen likheter mellom masteroppgaven og en eksamensbesvarelse jeg har levert tidligere. Under arbeidet med projektskissen og masteroppgaven har universitetet gjort det klart at slike likheter er tillatt.

2 Teoretisk og empirisk bakgrunn

I dette kapittelet vil det redegjøres for teoretisk og empirisk bakgrunn for masteroppgaven. Først kommer definisjoner og en nærmere redegjørelse for de to kognitive ferdighetene som er i fokus: vokabular og nonverbal resonnering. Noen barn og unge har utfordringer innen disse områdene, og siden problemstilling nummer to handler om disse barna, vil det deretter redegjøres for hva vi i dag vet om slike vansker. Etter dette rettes fokuset mot kognitiv utvikling – altså mot utvikling av vokabular og nonverbal resonnering – og sentrale teorier på feltet belyses. Mer bestemt vil gjensidighetsteorien, g-faktor-teorien, investeringsteorien og samplingteorien da bli diskutert.

2.1 Vokabular

Begrepet «vokabular» viser til en persons ordforråd og dermed ordene personen behersker. Å beherske et ord kan involvere en rekke ulike kompetanser, som eksempelvis kunnskap om ordets betydning, uttale, skrivemåte, grammatiske funksjon og bruksområder (Nation, 2013, s. 49). I stor grad dreier likevel vokabularferdigheter seg om semantisk språk, altså språkets innhold (Rygvoold, 2017). Nært relatert til begrepet «vokabular», er begrepet «mentalt leksikon». Uttrykket «mentalt leksikon» viser til at ord lagres i minnet etter hvert som vokabularet utvikles (Hoff, 2014).

Av mange blir vokabularet ansett som den mest sentrale av de ulike komponentene som utgjør generelle språkferdigheter (Vermeer, 2001). Likevel har det i senere tid vært betydelig forskningsinteresse knyttet til hvorvidt ulike språklige komponenter fungerer og utvikles modulært eller mer enhetlig. Tidligere hevdet forskere, blant annet med utgangspunkt i Chomskys teorier, at språk ble utviklet i mer separate enheter eller moduler for eksempel spesialisert mot ordforråd eller grammatikk (Helland, 2019). Nyere forskning antyder imidlertid at utviklingen av de ulike språkferdighetene (som grammatikk og vokabular) går hånd i hånd og danner en latent ferdighet hvor rangordningen mellom barna er svært stabil over tid (Bornstein et al., 2014). Dette peker i retning av at vokabular og andre språkferdigheter er nært relatert til hverandre og at de fremstår enhetlige, også i utviklingen. Derfor er vokabularet en språkferdighet som fungerer sammen med, og ikke adskilt fra, andre språkferdigheter.

Likevel sees vokabular ofte som grunnlag og basis for andre språkferdigheter, både språkforståelse og -produksjon (Vermeer, 2001). Ikke minst har vokabular også en sentral

betydning for skriftspråklige ferdigheter, og da særlig for leseforståelse (Snowling & Hulme, 2021; Vermeer, 2001). Hagtvet et al. (2011) kommer i sin metaundersøkelse frem til at små barns vokabular er relatert til senere leseferdigheter, i form av både ordavkodning og leseforståelse. Videre har det blitt vist at mens avkodingsferdigheter er avgjørende for leseforståelsen for barn som er i begynnelsen av sin leseutvikling, så blir språkforståelse og vokabular viktigere når avkodingsferdighetene stabiliserer seg (Hjetland et al., 2019). I tillegg har man også sett at et begrenset ordforråd er en tilleggsfaktor for å utvikle dysleksi for barn med familiær sårbarhet (van Viersen et al., 2017). På den andre siden er det også god forskningsmessig støtte for at lesing og leseferdigheter bidrar til og er viktig for ordforrådets utvikling (Duff et al., 2015; Nagy et al., 1985).

Utover muntlige og skriftlige språkferdigheter, har tidligere forskning dessuten vist at vokabularet har betydning for en rekke andre egenskaper og ferdigheter. Dette gjelder blant annet sosiale relasjoner (Gertner et al., 1994) samt skolefaglige og akademiske prestasjoner (Dollinger et al., 2008; Schuth et al., 2017). Det vil si at vokabularferdigheter er sentralt for gode språklige og skriftspråklige ferdigheter, men også for andre typer ferdigheter.

Forskningslitteraturen skiller ofte mellom breddevokabular og dybdevokabular.

Breddevokabular viser til størrelsen på ordforrådet, altså det totale antallet ord man behersker, mens dybdevokabularet sier noe om kvaliteten på kunnskapen om de ulike ordene (Næss et al., 2017). God dybdekunnskap blir gjerne kjennetegnet av at man har en fyldig definisjon av ulike ord, og at man kan kategorisere et ord og assosiere det til andre ord og begreper (Næss et al., 2017; Schmitt, 2014). En persons breddevokabular måles ved at man teller hvor mange av et gitt antall ord personen mestrer, mens mål på dybdevokabular er ute etter å nærmere beskrive personens kunnskap om enkeltord (Schmitt, 2010). Mål av dybdevokabular er imidlertid særlig utfordrende fordi definisjonen av konseptet varierer mellom ulike kilder, og det finnes i dag ikke en standard for hvordan dybdevokabular bør måles (Schmitt, 2014). Samtidig er det foreslått at breddevokabular og dybdevokabular er to konsepter som i hvert fall til en viss grad måler det samme, fordi de er avhengige av hverandre og har en sterk sammenheng med hverandre (Vermeer, 2001).

I tillegg er det vanlig å skille mellom impressivt og ekspressivt vokabular, også kalt reseptivt og produktivt vokabular. Det impressive eller reseptive vokabularet er ordene man forstår når man hører eller leser dem, mens det ekspressive eller produktive vokabularet er ordene man selv kan bruke muntlig eller skriftlig (Schmitt, 2019). En persons impressive vokabular er

større enn det ekspressive, ifølge noen kilder omtrent dobbelt så stort (Brysbaert et al., 2016). Det impressive vokabularet kan måles med øvelser som tester hvorvidt man gjenkjenner et ord og kan indikere hva det betyr, mens ekspressivt vokabular måles ved ulike øvelser som tester hvorvidt man selv klarer å hente frem riktig ord fra minnet (Schmitt, 2010). Det er imidlertid en sterk relasjon mellom impressivt og ekspressivt vokabular. Når en persons vokabular over tid øker i størrelse, går utviklingen nemlig fra ingen kunnskap om et ord, gjennom impressiv kunnskap og endelig til ekspressiv kunnskap (Schmitt, 2019).

For å oppsummere er vokabular en viktig språkferdighet, som har stor betydning for andre språklige og skriftspråklige ferdigheter. Nyere forskning viser at vokabular og andre språkferdigheter er relatert til hverandre, på en måte som gjør at de fungerer mer eller mindre helhetlig (Bornstein et al., 2014). Derfor er det vanskelig å skille vokabularferdigheter fra andre språkferdigheter. Videre har det i denne gjennomgangen kommet frem at begrepsparene «breddevokabular og dybdevokabular» og «impressivt og ekspressivt vokabular» er mye brukt, men at det også er vanskelig å skille breddevokabular fra dybdevokabular, og impressivt fra ekspressivt vokabular, fordi disse ferdighetene også er svært nært relaterte.

2.2 Nonverbal resonnering

Mens vokabularferdigheter altså både henger nært sammen med mange ulike språkferdigheter, og er viktig for mange ulike språkferdigheter, er nonverbal resonnering på sin side en såkalt ikke-språklig kognitiv ferdighet. Mer presist kjennetegnes nonverbal resonnering av problemløsning som er mer eller mindre uavhengig av språklig kompetanse, herunder både språkproduksjon og språkforståelse (Kuschner, 2013). I tillegg er oppgaver som krever nonverbal resonnering lite avhengige av kultur og tidligere erfaring eller læring (Peng et al., 2019). Evnen til logisk og abstrakt tenkning er derimot sentral (Ferrer et al., 2009). Nonverbal resonnering kan være induktiv og/eller deduktiv (Peng et al., 2019). Induktiv resonnering vil si at man oppdager generelle regler basert på spesifikke eksempler, mens deduktiv resonnering vil si at man danner korrekte prediksjoner basert på generelle regler (Sternberg et al., 2008, s. 190–191).

I flere publikasjoner omtales nonverbal resonnering som flytende intelligens (Peng et al., 2019; Primi, 2002). Flytende intelligens er et begrep som sammen med begrepet krystallisert intelligens ble introdusert av Cattell (1943, 1963). Ofte defineres flytende intelligens som evnen til resonnering og problemløsning når man står overfor nye og ukjente kognitive oppgaver (Salthouse et al., 2008). Det skiller seg altså fra krystallisert intelligens, som

innebærer kognitive ferdigheter som man har lært seg (Sternberg et al., 2008, s. 14). Språkferdigheter og vokabular er blant de ferdighetene som ofte betegnes som krystallisert intelligens (Peng et al., 2019). I noen publikasjoner gjøres det likevel et poeng av at nonverbal resonnering og flytende intelligens ikke er identiske begreper, fordi nonverbal resonnering først og fremst kjennetegnes av ikke-språklig problemløsning, mens flytende intelligens kjennetegnes av hvordan man løser nye og ukjente problemer (Nutley et al., 2011; Peng et al., 2019).

Kartlegging av nonverbal resonnering foregår gjennom problemløsningsoppgaver som ofte innebærer at man skal manipulere ulike former for visuell informasjon (Kuschner, 2013). Særlig brukes oppgaver der man skal bruke logikk for å lage eller fullføre figurer eller serier med figurer (Ferrer et al., 2009; Peng et al., 2019). I nonverbale kartlegginger finnes det noen tester som hverken krever språkforståelse eller språkproduksjon, mens andre tester har verbale instruksjoner som skal gis av en testadministrator (McCallum, 2017). Det vil si at noen nonverbale tester likevel stiller visse krav til språkferdigheter hos personen som skal testes. Tester som skal måle nonverbal resonnering kan også brukes som kartleggingsverktøy for å måle intelligens som helhet (Helland-Riise & Martinussen, 2017), og derfor er det en nær relasjon mellom begrepene «nonverbal resonnering» og «intelligens».

Som fenomenet positive manifold også forteller oss, henger evnen til nonverbal resonnering sammen med ulike andre kognitive ferdigheter, både språklige og ikke-språklige, som er sentrale i hverdagslige og skolefaglige sammenhenger. Det har blant annet vist seg å være en tydelig positiv sammenheng mellom nonverbal resonnering og arbeidsminne, altså evnen til å holde på og bearbeide informasjon mens man løser en oppgave (Kane et al., 2005). Det er også en vanlig oppfatning innen forskning at evnen til nonverbal resonnering ligger til grunn for elevers skoleprestasjoner (Peng & Kievit, 2020). I tillegg kommer det frem i en større metastudie at barn og unge som strever med språket også strever mer enn sine jevnaldrende med nonverbal resonnering (Gallinat & Spaulding, 2014). Videre er det ifølge Saar et al. (2023) slik at barn med utviklingsmessige språkforstyrrelser har større språklige utfordringer jo mer alvorlige nonverbale vansker de samtidig har. Det skal imidlertid nevnes at det også finnes studier der barn med språkforstyrrelser skårer gjennomsnittlig innen nonverbal resonnering (Calder et al., 2023). Begrepet utviklingsmessige språkforstyrrelser, og forholdet mellom slike vansker og nonverbal resonnering, utdypes nærmere nedenfor i kapittel 2.3.

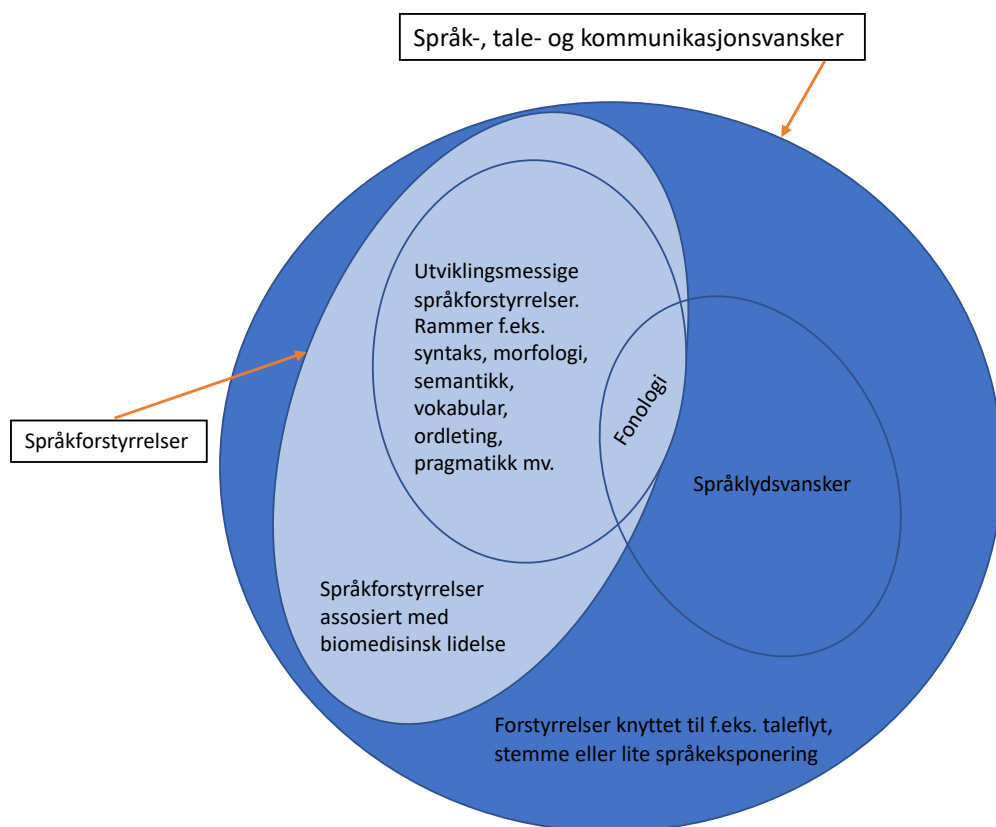
Oppsummeringsvis er nonverbal resonnering en betegnelse på kognitive ferdigheter som i liten eller ingen grad er avhengige av språk, kultur eller tidligere læring og erfaring. Ferdigheter knyttet til nonverbal resonnering er derimot kjennetegnet av evnen til logisk og abstrakt tenkning og problemløsning. Dette testes ofte med problemløsningsoppgaver som går ut på at man skal bearbeide og manipulere visuell informasjon. Videre har denne gjennomgangen pekt på at evnen til nonverbal resonnering henger sammen med en rekke andre typer ferdigheter, både språklige og andre ikke-språklige ferdigheter.

2.3 Vansker tilknyttet vokabular og nonverbal resonnering

Noen barn og unge vil imidlertid streve med vokabular og språk, eller med nonverbal resonnering, og disse to gruppene skal vi se nærmere på her. Sentrale begreper som vil defineres og gjennomgås her er «språkforstyrrelser», «utviklingshemming» og «generelle lærevansker».

2.3.1 Språkforstyrrelser

Når man snakker om vansker knyttet til språk, bruker man i dag terminologi utarbeidet gjennom CATALISE-studien (Bishop et al., 2017). I denne studien deltok engelskspråklige eksperter fra flere land. I ettertid er det også gjort en tilsvarende studie der norske eksperter har deltatt (Kristoffersen et al., 2021). Som følge av disse konsensusstudiene kalles vansker med språk i dag for «språkforstyrrelser». Når språkforstyrrelsen har sin årsak i en biomedisinsk lidelse, skal det kalles «språkforstyrrelse assosiert med den biomedisinske lidelsen». Eksempelvis kalles språkforstyrrelser som skyldes autisme «språkforstyrrelser assosiert med autisme». Andre eksempler på slike biomedisinske lidelser er hjerneskader, utviklingshemming og Downs syndrom. Når språkforstyrrelsen imidlertid ikke skyldes slike biomedisinske tilstander, skal det kalles «utviklingsmessige språkforstyrrelser». Dette forkortes til DLD fordi den engelske terminologien er «developmental language disorders», og man bruker den samme forkortelsen på norsk (Kristoffersen et al., 2021). Videre skilles språkforstyrrelser fra forstyrrelser som kun er knyttet til språklyder, samt fra vansker som skyldes lite eksponering for språket, stemmevansker eller taleflytvansker. For å tydeliggjøre terminologien, viser figur 1 et venndiagram som er omarbeidet etter den opprinnelige CATALISE-studiens figur over språk-, tale og kommunikasjonsvansker (Bishop et al., 2017). Den nevnte norske konsensusstudien har også utarbeidet en lignende oversettelse av den samme figuren.



Figur 1: Terminologi om språkforstyrrelser, omarbeidet etter CATALISE-studien (Bishop et al., 2017).

Utviklingsmessige språkforstyrrelser kan, som figur 1 også viser, føre til ulike uttrykk, og det er en heterogen vanskegruppe (Bishop et al., 2017). Denne typen heterogenitet ser man også innenfor andre lærevansker som dysleksi og dyskalkuli (Peters & Ansari, 2019). Likevel peker Bishop et al. (2017) på at det finnes vanskeområder som går igjen hos barn og unge med språkforstyrrelser. Det dreier seg blant annet om problemer med språklyder, setningsoppbygging, pragmatikk, ordleting og/eller vokabular. Fokuset i denne masteroppgaven er særlig vokabular. Angående dette har tidligere forskning pekt i retning av at barn med utviklingsmessige språkforstyrrelser har et mindre bredde- og dybdevokabular enn sine jevnaldrende (McGregor et al., 2013). Likevel viser derimot nyere studier at barn med utviklingsmessige språkforstyrrelser presterer godt innen gjennomsnittsområdet på standardiserte tester som måler reseptivt vokabular (Calder et al., 2023). Derfor er det ikke anbefalt å bruke tester av reseptivt vokabular i diagnostisering av språkforstyrrelser, fordi det kan føre til underdiagnostisering.

Vansker med språk og vokabular er likevel gjerne kjennetegnet av kvantitative, men ikke kvalitative, forskjeller fra den øvrige befolkningen. Det er fordi disse ferdighetene har en normalfordeling i befolkningen (Hagen et al., 2014). Derfor er også grensen for hvor store utfordringer som må til for at noe skal defineres som en vanske utydelig. Dette er også kjent fra andre typer lærevansker, og medfører at cut-off-punktet mellom barn med og uten vansker vil variere fra situasjon til situasjon (Peters & Ansari, 2019). Innen forskning er det likevel gjerne slik at deltakere inkluderes i en språkforstyrrelsesgruppe dersom man har skårer innen språk som ligger minst -1 standardavvik under gjennomsnittskårene for aldersgruppa (Snowling, Hayiou-Thomas, et al., 2020). Også i psykometri brukes ofte en grense på -1 standardavvik, altså skårer på omtrent 15. persentil eller lavere, for å identifisere avvik som kan være av betydning i tester eller målinger av enkeltindivider (Pearson Clinical Assessment's Scientific Council, 2023). Dette fordi skårer på dette nivået indikerer en signifikant lavere skår eller verdi sammenlignet med resten av populasjonen.

Det man i dag kjenner som utviklingsmessige språkforstyrrelser, ble tidligere kalt spesifikke språkvansker. For å omfattes av dette begrepet var det et krav om at språkvanskene måtte være uventet sammenlignet med øvrige kognitive ferdigheter som nonverbal resonnering (Rygvoold, 2017). I CATALISE-studien kom man til enighet om at dette kravet måtte ses bort fra (Bishop et al., 2017). Det innebærer at man i dag kan få diagnosen utviklingsmessige språkforstyrrelser selv om man også strever med nonverbal resonnering. Denne nye forståelsen har også empirisk støtte fordi det viser seg at barn som primært strever med språk også skårer lavere enn jevnaldrende når det gjelder nonverbal resonnering (Gallinat & Spaulding, 2014).

Det finnes ulike intervensjoner med dokumenterte effekter som kan avhjelpe språkforstyrrelser (Hulme et al., 2020). Forskjellige intervensjonsprogrammer der man jobber med forskjellige språkferdigheter som barnet strever med, kan blant annet ha effekt (Rogde et al., 2019). Forskning peker imidlertid i retning av at trening av arbeidsminne eller andre typer generelle kognitive ferdigheter ikke gir effekt på verbale ferdigheter (Hagen et al., 2014; Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016; Strong et al., 2011). Undersøkelser indikerer også at personer med språkforstyrrelser responderer nokså likt på tiltak, uavhengig av om de skårer gjennomsnittlig eller lavt gjennomsnittlig på nonverbal resonnering (Bishop et al., 2016).

2.3.2 Utviklingshemming og generelle lærevansker

Vansker med nonverbal resonnering er knyttet til intellektuelle evner og består av vansker med problemløsning, planlegging, kompleksitet, abstrakt og logisk tenkning, samt med akademisk og erfaringsbasert læring (Harris, 2013). Nedenfor gjennomgås noen begreper som er tilknyttet slike vansker.

Spesielt uttalte problemer med nonverbal resonnering er et hovedkjennetegn og kriterium for diagnosen psykisk utviklingshemming (Gallinat & Spaulding, 2014). I medisinsk sammenheng brukes termen «psykisk utviklingshemming», men i spesialpedagogisk sammenheng bruker man ofte heller «utviklingshemming» (Sigstad, 2017). Derfor brukes «utviklingshemming» også her. Adskilt etter alvorlighetsgrad, er dette en gruppe diagnoser som innebærer at man har en målt IQ-skår lavere enn omtrent 70, samt betydelige problemer med adaptive ferdigheter (Verdens helseorganisasjon, 2019). I likhet med språkferdigheter er IQ-skår og også evnen til nonverbal resonnering normalfordelt i befolkningen (Gallinat & Spaulding, 2014). I DSM-5 og ICD-11, de nyeste utgavene av diagnosemanualene, legges det imidlertid mindre vekt enn tidligere på IQ-skår i diagnostisering av utviklingshemming. Tendensen er med andre ord at vurdering av adaptive ferdigheter har blitt mer sentral i diagnostiseringssammenheng (Harris, 2013). Adaptive ferdigheter handler om hverdagslig fungering og kan deles inn i sosiale, praktiske og begrepsmessige ferdigheter (Harris, 2013; Sigstad, 2017).

Mindre uttalte vansker med nonverbal resonnering kan omfattes av begrepet «generelle lærevansker». Dette innebærer at man har lærevansker som ikke er avgrenset til bare ett område, og videre at man har en IQ-skår som ligger noe under gjennomsnittet (Krämer et al., 2021). «Generelle lærevansker» brukes ofte for å beskrive lærevanskene til personer som ligger i gråsonen mellom utviklingshemming og gjennomsnittlig fungering (Eckhoff, 1997). Noen inkluderer også personer med lett utviklingshemming i begrepet «generelle lærevansker» (Haugen, 2016; Krämer et al., 2021).

Selv om personer med utviklingshemming og generelle lærevansker kjennetegnes av problemer med nonverbal resonnering, er også vansker knyttet til språket en utfordring for disse gruppene. En av årsakene til dette er som nevnt tidligere at de ulike kognitive ferdighetene har sammenheng med hverandre. Språkforstyrrelsene til personer med utviklingshemming kalles som nevnt «språkforstyrrelser assosiert med utviklingshemming». Derimot kan språkforstyrrelsene til personer som har generelle lærevansker uten

utviklingshemming plasseres i kategorien «utviklingsmessige språkforstyrrelser» (Bishop et al., 2017). Tidligere forskning viser at språkprofilen til barn med utviklingsmessige språkforstyrrelser ikke er avhengig av om de har generelle lærevansker eller ikke, mens barn med språkforstyrrelser assosiert med utviklingshemming har mer alvorlige språklige utfordringer (Norbury et al., 2016).

Det er uklart om vansker med nonverbal resonnering, som man særlig ser hos barn med utviklingshemming eller generelle lærevansker, er trenbare. I utgangspunktet var det antatt at flytende intelligens ikke kunne læres eller trenes (Horn & Cattell, 1966). Man kan likevel forbedre sin skår på nonverbale tester ved å lære hvordan man løser oppgavene (Bors & Vigneau, 2003). Forskning tyder imidlertid på at dette ikke egentlig reflekterer økte evner til nonverbal resonnering (te Nijenhuis et al., 2007). Derfor har det vært interesse for å se hvorvidt trening av andre kognitive ferdigheter, særlig arbeidsminne, kunne føre til forbedringer av nonverbal resonnering (Jaeggi et al., 2008). Resultatene til Nutley et al. (2011) viser likevel at evnen til nonverbal resonnering faktisk kan forbedres gjennom treningsprogrammer som fokuserer på nettopp nonverbal resonnering, og at trening av arbeidsminne ikke er effektivt. Riktignok finnes det studier som støtter at arbeidsminnetrening kan forbedre nonverbal resonnering (Au et al., 2014; Jaeggi et al., 2008; Karbach & Verhaeghen, 2014). Likevel setter en større metaundersøkelse spørsmålstegn ved slike funn, og viser at trening av arbeidsminnet trolig ikke er hensiktsmessig for å forbedre andre kognitive ferdigheter, herunder nonverbal resonnering (Melby-Lervåg et al., 2016).

2.4 Teorier om positive manifold og kognitiv utvikling

Spørsmål om kognitiv utvikling engasjerer som nevnt forskere innen forskjellige disipliner. Forskning på dette har lange tradisjoner, og følgelig har det gjennom tidene blitt lansert mange ulike teorier som forsøker å forklare utviklingen av kognitive ferdigheter. Nedenfor beskrives noen av disse teoriene og forskningen som har blitt gjort om dem. Teoriene det fokuseres på har til felles at de alle berører utviklingen av flere ulike typer kognitive ferdigheter, blant annet vokabular og nonverbal resonnering.

Videre har teoriene nedenfor til felles at de alle representerer alternative forklaringer på positive manifold. Spearman (1904) oppdaget som nevnt at ulike kognitive ferdigheter korrelerte med hverandre, og kalte fenomenet «positive manifold». Det innebærer for eksempel at personer som har gode anlegg for én type kognitiv ferdighet, gjerne også har gode anlegg for andre typer kognitive ferdigheter (og motsatt). Selv om dette er et fenomen

som har stor aksept (van der Maas et al., 2006), er det som vi skal se usikkerhet knyttet til hvorfor ferdighetene korrelerer.

I diskusjonen om hvorfor de ulike ferdighetene har sammenheng med hverandre og korrelerer, er det viktig å huske på forskjellen mellom korrelasjon og kausalitet. To ferdigheter kan vise en sammenheng med hverandre uten at de på noen måte påvirker hverandre. Sammenhengen kan for eksempel ha oppstått fordi en felles underliggende faktor påvirker begge ferdighetene. Sammenhengen kan imidlertid også ha oppstått fordi de påvirker hverandre i løpet av utviklingen, enten ved at den ene påvirker den andre, eller ved at det er et gjensidig forhold mellom dem. Teoriene vi skal se på her forsøker å forklare sammenhengen på ulike måter: *G*-faktor-teorien og samplingteorien tar utgangspunkt i to forskjellige underliggende tredjevariabler. Gjensidighetsteorien tar utgangspunkt i at det er en direkte og gjensidig årsakssammenheng mellom ulike kognitive ferdigheter. Investerings-teorien hevder at noen typer kognitive ferdigheter ensidig påvirker utviklingen av andre typer kognitive ferdigheter.

2.4.1 Gjensidighetsteorien

Hovedfokuset i denne oppgaven er gjensidighetsteorien. Dette er en teori som først ble lansert av van der Maas et al. (2006). Hovedpoenget i denne teorien er at ulike kognitive ferdigheter påvirker og understøtter hverandre i utviklingen, noe som resulterer i at det i løpet av utviklingsperioden vokser frem en positiv korrelasjon – positive manifold – mellom de ulike ferdighetene hos et individ. En annen sentralt antakelse i teorien er at positive manifold først oppstår etter at de ulike kognitive ferdighetene har begynt å etablere seg. Det kausale forholdet teorien beskriver er altså at de ulike kognitive ferdighetene har en gjensidig årsakssammenheng med hverandre, slik at én kognitiv ferdighet som har begynt å etablere seg påvirker utviklingen av andre kognitive ferdigheter, og motsatt.

Det er særlig to banebrytende undersøkelser, utført av Kievit og kolleger, som har motivert denne masterstudien (Kievit et al., 2019; Kievit et al., 2017). Disse to undersøkelsene fokuserer begge på vokabular og nonverbal resonnering. Begge undersøkelsene er sentrale for hvorfor gjensidighetsteorien i dag er aktuell i faglige diskusjoner om kognitiv utvikling og særlig utviklingen av vokabular og nonverbal resonnering. Ettersom disse studiene til Kievit og kolleger er spesielt aktuelle og viktige for denne masteroppgaven, vil det sees nærmere på hver av dem nedenfor.

I undersøkelsen fra 2017 (Kievit et al., 2017) var målet å undersøke hvilken teori som best forklarte utviklingen av vokabular og nonverbal resonnering hos ungdommer og unge voksne i England. Teoriene det ble sett på var gjensidighetsteorien, *g*-faktor-teorien og investeringsteorien. Utviklingen av vokabular og nonverbal resonnering ble undersøkt ved at man på to ulike måletidspunkter (med i gjennomsnitt 1,48 år mellom hvert måletidspunkt) gjorde tester av vokabular (Ordforståelse fra WAIS-IV) og nonverbal resonnering (Matriser fra WAIS-IV). Totalt deltok 784 deltakere i første testrunde, og 563 av deltakerne ble igjen kartlagt i andre runde. I første testrunde var deltakerne 14–25 år gamle. Dataene fra testingen ble analysert ved at man brukte *Latent Change Score Models*, en analysemetode som vil bli nærmere beskrevet under metodekapittelet i denne masteroppgaven. Resultatene fra 2017-studien til Kievit og kolleger viste at gjensidighetsteorien forklarte utviklingen av vokabular og nonverbal resonnering bedre enn både investeringsteorien og *g*-faktor-teorien. Dette fordi gjensidighetsteorien hadde bedre «model fit» enn de øvrige modellene, og fordi man så at vokabular påvirket utviklingen av nonverbal resonnering ($\gamma = 0,051$) samtidig som nonverbal resonnering påvirket utviklingen av vokabular ($\gamma = 0,163$). Dermed ble det antatt at disse to ferdighetene påvirket hverandres utvikling gjensidig.

I den påfølgende undersøkelsen var målet å undersøke om resultatene fra 2017-undersøkelsen lot seg replikere på et utvalg med yngre deltakere (Kievit et al., 2019). I denne undersøkelsen deltok 227 barn i alderen 6–8 år. Igjen ble det sett på om det var gjensidighetsteorien, *g*-faktor-teorien eller investeringsteorien som best forklarte utviklingen av vokabular og nonverbal resonnering. I analysene ble det også i denne undersøkelsen brukt *Latent Change Score Models*, men her brukte man data fra tre i stedet for to forskjellige måletidspunkter i analysene. Vokabular ble målt med deltesten Ordforståelse fra WASI, mens nonverbal resonnering ble målt med Matriser fra WASI. I likhet med undersøkelsen fra 2017, viste denne studien at gjensidighetsteorien bedre enn både *g*-faktor-teorien og investeringsteorien forklarte utviklingen av vokabular og nonverbal resonnering. Også her så man at gjensidighetsteorien hadde best «model fit». I tillegg så man at vokabular påvirket utviklingen av nonverbal resonnering ($\gamma = 0,273$) samtidig som nonverbal resonnering påvirket utviklingen av vokabular ($\gamma = 0,272$), noe som gjaldt mellom samtlige måletidspunkter. Derfor ble det også i denne studien påstått at det er et gjensidig forhold i utviklingen av ulike kognitive ferdigheter.

Det finnes i tillegg en rekke studier som støtter gjensidighetsteorien, men som har fokusert på andre typer kognitive ferdigheter. Ferrer et al. (2007) fant blant annet positive gjensidige

forhold mellom lesing og ulike andre kognitive ferdigheter. I en annen studie er det også vist at tidlige språkferdigheter predikerer flere kognitive ferdigheter senere i barndommen (Marchman & Fernald, 2008). I tillegg kommer Peng et al. (2019) frem til at gjensidighetsteorien kan bidra til å forklare forholdet mellom nonverbal resonnering og både lesing og matematikkferdigheter. Også Peng og Kievit (2020) viser til at det kan være gjensidige forhold mellom akademiske ferdigheter som lesing og matematikk og ulike andre kognitive ferdigheter. Forskjellige typer matematikkferdigheter virker også å påvirke hverandre gjensidig (Hofman et al., 2018). En annen studie (Schmitt et al., 2017) har utforsket det longitudinelle forholdet mellom barns matematikkferdigheter, lesing og eksekutive funksjoner (bestående av arbeidsminne, impuls kontroll og kognitiv fleksibilitet). I denne studien ble det gjort funn som delvis støttet gjensidighetsteorien, fordi hvorvidt det var gjensidige forhold varierte avhengig av hvilke ferdigheter man målte og mellom ulike tidspunkter i barnas utviklingsperiode.

Siden masteroppgavens fokus er spesialpedagogikk, er det spesielt interessant å se hva tidligere spesialpedagogisk forskning sier om gjensidighetsteorien. Som nevnt konkluderte Griffiths et al. (2022) med at gjensidighetsteorien var holdbar for barn både med og uten språkforstyrrelser. De fant altså ingen forskjeller mellom barn med og uten språkforstyrrelser på dette punktet. Dette skiller seg fra studien til Ferrer et al. (2010), som omhandlet barn med og uten dysleksi. De fant gjensidige forhold mellom intelligens og lesing hos barna uten dysleksi, men det samme gjaldt ikke for barna med dysleksi. En mulig årsak til at gjensidighetsteorien i disse studiene er holdbar for vokabular og nonverbal resonnering hos barn med språkforstyrrelser, men ikke for lesing og intelligens hos barn med dysleksi, er at barn med dysleksi ofte unngår lesing og på den måten «bryter» det gjensidige forholdet mellom lesing og intelligens, mens barn med språkforstyrrelser ikke har anledning til å unngå språk (Ferrer et al., 2010; Griffiths et al., 2022).

Selv om det finnes en del forskning som støtter gjensidighetsteorien eller deler av den, finnes det også motstridende funn. For eksempel fant Quinn et al. (2015) i sin studie at vokabular påvirket utviklingen av leseforståelse, men at forholdet ikke var gjensidig. Videre kommer Blaga et al. (2009) frem til at ulike kognitive ferdigheter har stor grad av stabilitet i sine utviklingsmønstre, og at utviklingsmønsteret ikke ser ut til å være påvirket av vokabularferdigheter. Som vi skal utforske nærmere nedenfor, er det dessuten usikkerhet knyttet til gjensidighetsteorien fordi det også finnes empirisk støtte for andre teorier, blant

annet *g*-faktor-teorien (Gignac, 2014, 2016), investeringsteorien (Kvist & Gustafsson, 2008) og samplingteorien (Bartholomew et al., 2013; Bartholomew et al., 2009).

2.4.2 G-faktor-teorien

Den mest tradisjonelle forklaringen på positive manifold er *g*-faktor-teorien (Spearman, 1927). Den går ut på at det finnes en underliggende prosess eller mentalkapasitet hos individet, som påvirker evnen til å utføre mange forskjellige kognitive oppgaver. Denne mentalkapasiteten kalles *g*-faktoren, som står for generell intelligens. Innenfor *g*-faktor-teorien antar man at ulike kognitive ferdigheter er relatert til hverandre bare fordi de på ulike måter er avhengige av *g*-faktoren. I motsetning til gjensidighetsteorien antar man altså at de ulike kognitive ferdighetene ikke påvirker hverandre, og at det ikke finnes direkte årsakssammenhenger på tvers av ulike kognitive egenskaper (Kievit et al., 2017). Blant annet vil dette si at *g*-faktor-teorien hevder at vokabular og nonverbal resonnering korrelerer med hverandre fordi begge ferdighetene avhenger av *g*-faktoren. Det kausale i denne teorien er altså at det er en tredjevariabel som ligger til grunn for begge ferdighetene og forklarer sammenhengen mellom dem.

Selv om det ikke er funnet noen hjerneanatomiske strukturer som knyttes til *g*-faktoren (Haier et al., 2009), blir det innenfor denne teorien antatt at *g*-faktoren er en biologisk og arvelig mentalkapasitet (Jensen, 1998). Den kommer på plass svært tidlig i barndommen, og deretter vil man ut fra denne etablere og forbedre ulike kognitive ferdigheter som vil korrelere med hverandre (Kievit et al., 2017). Det vil si at *g*-faktor-teorien antar at også utviklingen av de ulike kognitive ferdighetene er avhengig av *g*-faktoren. Likevel er utviklingsperspektivet noe underkommunisert innenfor *g*-faktor-teorien sammenlignet med for eksempel gjensidighetsteorien (van der Maas et al., 2006). Altså er det mer fokus på at *g*-faktoren fører til positive manifold, og mindre fokus på hvordan dette forekommer.

En sentral forskjell mellom *g*-faktor-teorien og gjensidighetsteorien er at gjensidighetsteorien som nevnt antar at positive manifold ikke er til stede innledningsvis i barns kognitive utviklingsperiode, men at det oppstår og forsterkes over tid (van der Maas et al., 2006). Ifølge *g*-faktor-teorien er derimot positive manifold mer stabilt og også til stede hos de minste barna, fordi man antar at *g*-faktoren etableres tidlig. Da Gignac (2014) testet dette, fant han støtte for at positive manifold er et relativt stabilt fenomen, som også finnes hos de minste barna. Derfor var hans funn mest i samsvar med *g*-faktor-teorien. Også Shahabi et al. (2018) kom frem til dette da de testet noe av det samme.

G-faktor-teorien blir også støttet og underbygget i andre publikasjoner (Gignac, 2016; Jensen, 1998). Hvorvidt det i realiteten finnes en underliggende mentalkapasitet som *g*-faktoren er imidlertid debattert, med flere synspunkter både for og mot (Grigorenko & Sternberg, 2002). McArdle et al. (2002) konkluderer på sin side med at teorier som antar at det finnes en enkelt enhetlig *g*-faktor blir for unyanserte, fordi utviklingsløpene til ulike kognitive ferdigheter er for ulike. Investeringsteorien som omtales nedenfor, bygger imidlertid på en mer kompleks forståelse av *g*-faktoren.

2.4.3 Investeringsteorien

Cattell (1971, 1987) studerte *g*-faktoren videre, og formulerte over tid en ny teori om hvordan man kan forklare positive manifold og kognitiv utvikling. Denne har fått navnet «investment theory». Heretter vil jeg bruke en norsk oversettelse og kalle den for «investeringsteorien».

Bakgrunnen for investeringsteorien er en videreutvikling av Spearman's (1927) *g*-faktor-teori, nemlig en teori som blant annet kalles *Gf–Gc*-teorien (Cattell, 1943, 1963; Horn, 1968; Horn & Cattell, 1966). Den innebærer at *g*-faktoren er delt inn i flytende intelligens og krystallisert intelligens. I internasjonale publikasjoner er disse delene ofte forkortet til henholdsvis *Gf* og *Gc*. Som tidligere beskrevet, er flytende intelligens evnen til å løse nye kognitive oppgaver og problemer, mens krystallisert intelligens er kognitive ferdigheter man har tilegnet seg, eksempelvis lesing, matematikk og vokabular. Flytende intelligens er i stor grad medfødt, og knyttes til arvelighet og biologi, mens krystallisert intelligens knyttes til læring, utdanning og miljøpåvirkning (Gustafsson, 1984; McArdle et al., 2002). Blant annet fordi den i ettertid har blitt slått sammen med Carroll (1993) sin teori om kognitive evner, har *Gf–Gc*-teorien fortsatt anerkjennelse i dag (McGrew, 2009).

Mens *Gf–Gc*-teorien skal forklare hvordan kognitive ferdigheter er strukturert, bygger investeringsteorien på det samme grunnlaget, men handler spesifikt om hvordan kognitive ferdigheter utvikles og hvorfor positive manifold eksisterer. I korte trekk går investeringsteorien ut på at krystallisert intelligens er noe man tilegner seg ved at man «investerer» sin flytende intelligens mens man lærer, derav navnet «investeringsteorien» (Kievit et al., 2017). Med andre ord er det en medfødt flytende intelligens som skaper grunnlaget for utviklingen av krystallisert intelligens. Ifølge investeringsteorien forklares derfor sammenhengen mellom ulike kognitive ferdigheter, altså positive manifold, med at utviklingen av alle former for krystallisert intelligens er avhengig av flytende intelligens.

Kausaliteten som beskrives i denne teorien går altså fra flytende intelligens til krystallisert intelligens.

Videre vil dette si at investeringsteorien hevder at utviklingen av vokabular er avhengig av nonverbal resonnering, men ikke motsatt. Det er fordi vokabular er en form for krystallisert intelligens, mens nonverbal resonnering er tilnærmet det samme som flytende intelligens. På den måten skiller investeringsteorien seg fra gjensidighetsteorien, fordi sistnevnte hevder at det er en gjensidig påvirkning i utviklingen av vokabular og nonverbal resonnering.

En forskjell mellom investeringsteorien og gjensidighetsteorien som Peng et al. (2019) påpeker, er dessuten at investeringsteorien hevder at relasjonen mellom flytende og krystallisert intelligens er spesielt sterk i tidlig kognitiv utvikling. Før de har etablert seg fullstendig, skal oppgaver som krever krystallisert intelligens nemlig være særlig avhengige av flytende intelligens. Det vil si at investeringsteorien, som *g*-faktor-teorien, forventer en tydelig sammenheng mellom eksempelvis vokabular og nonverbal resonnering blant yngre barn. På den andre siden hevder gjensidighetsteorien at sammenhengen mellom ulike kognitive ferdigheter, som for eksempel vokabular og nonverbal resonnering, blir sterkere over tid (van der Maas et al., 2006).

Forskning på investeringsteorien har hatt litt blandede resultater. Schweizer og Koch (2002) fant støtte for investeringsteorien da de undersøkte om læring fungerte som mediator mellom flytende og krystallisert intelligens, altså om flytende intelligens førte til læring som igjen førte til krystallisert intelligens. Resultatene til Kvist og Gustafsson (2008), som så på forholdet mellom *g*-faktoren og flytende intelligens, taler også for investeringsteorien. Peng et al. (2019) fant delvis støtte for investeringsteorien, men sterkere støtte for andre typer teorier. Rindermann et al. (2010) undersøkte investeringsteorien i ulike utvalg, men viste at det over tid var gjensidige forhold mellom flytende og krystallisert intelligens, noe som er mer i tråd med gjensidighetsteorien. Likevel viste de at krystallisert intelligens i større grad enn flytende intelligens påvirkes av miljøfaktorer, slik *Gf–Gc*-teorien og investeringsteorien går ut fra. Også Ferrer og McArdle (2004) fant noe støtte for investeringsteorien, men foreslo at den ikke er kompleks nok til å beskrive kognitiv utvikling. Det finnes også studier som ikke har funnet noe som kunne underbygge investeringsteorien (Christensen et al., 2013).

2.4.4 Samplingteorien

Både g-faktor-teorien og investeringsteorien antar altså at det finnes en form for medfødt mentalkapasitet som er ansvarlig for hvordan man utfører ulike kognitive oppgaver, selv om investeringsteorien mer bestemt innebærer at det finnes noen typer kognitive ferdigheter (flytende intelligens) som har en direkte påvirkning på andre typer kognitive ferdigheter (krystallisert intelligens). I likhet med gjensidighetsteorien inneholder imidlertid ikke Thomsons (1951) samplingteori noen form for medfødt underliggende mentalkapasitet. Som vi skal se antar likevel samplingteorien at korrelasjonen mellom ulike kognitive ferdigheter skyldes en tredjevariabel, og ikke at ulike kognitive ferdigheter påvirker hverandre.

Biologisk bygger samplingteorien (Thomson, 1951) på at menneskets kognitive egenskaper fungerer ved at hjernen hele tiden danner mange ulike nevrologiske forbindelser (Bartholomew et al., 2009). En person med sterke kognitive ferdigheter kjennetegnes derfor av evnen til å danne flere slike forbindelser enn en person med svakere kognitive ferdigheter. Videre kjennetegnes også kognitiv utvikling av at man i utviklingsperioden klarer å danne flere og flere ulike forbindelser.

Innenfor samplingteorien går man ut fra at når man utfører én type kognitiv oppgave, iverksettes ett utvalg av disse kognitive forbindelsene, og når man utfører en annen type kognitiv oppgave, iverksettes et annet utvalg av forbindelser (Bartholomew et al., 2009). Man antar altså at ulike kognitive ferdigheter fungerer mer eller mindre separat fra hverandre (Egeland, 2022). Likevel, når man utfører en kognitiv test, vil det i tråd med samplingteorien også bli aktivert hjerneforbindelser som er assosiert med andre typer kognitive oppgaver, og det er dette som i denne teorien gir oppgav til positive manifold (Bartholomew et al., 2009). Det vil si at korrelasjonen mellom kognitive ferdigheter forklares av hjernens måte å lage koblinger på, og dette blir en underliggende tredjefaktor. Dette betyr at dersom man kontrollerer for at hjernen danner koblinger på denne måten, så vil ikke de ulike kognitive ferdighetene lenger vise en korrelasjon med hverandre.

Den korrelasjonen man finner mellom forskjellige kognitive ferdigheter skyldes derfor ifølge samplingteorien i stor grad at testinstrumentene man bruker til å måle de forskjellige ferdighetene ikke er nøyaktige nok, eller viser en viss overlapp med hverandre, slik at testinstrumenter som er utviklet for å måle en type kognitiv ferdighet også aktiverer hjerneforbindelser som er assosiert med andre typer kognitive oppgaver (Bartholomew et al., 2009). Et eksempel på denne overlappen kan i så fall være hvordan mål på nonverbal

resonnering som nevnt noen ganger har verbale instruksjoner og derfor også er noe språkavhengige (McCallum, 2017).

Tidligere forskning viser at samplingteorien rent statistisk forklarer positive manifold like godt som g -faktor-teorien, og dessuten at den biologiske begrunnelsen for samplingteorien er holdbar (Bartholomew et al., 2009). I tillegg har det blant forskere vært fornyet interesse rundt teorien i nyere tid. Gjennom ulike publikasjoner har Thomsons opprinnelige samplingteori blitt testet og arbeidet videre med (Bartholomew et al., 2013; Bartholomew et al., 2009). I tillegg har en ny variant av samplingteorien også blitt presentert (Kovacs & Conway, 2019). Den går ut på at utførelsen av en kognitiv oppgave krever flere både domenegenerelle og domenespesifikke kognitive prosesser, og at positive manifold oppstår fordi noen av disse prosessene, særlig de domenegenerelle, går igjen i utførelsen av ulike kognitive oppgaver.

En grunn til at samplingteorien i dag igjen tas opp i forskning kan være at den har et sentralt likhetstrekk med den svært aktuelle gjensidighetsteorien: at positive manifold ikke skyldes en underliggende mentalkapasitet (g -faktoren), og at denne underliggende mentalkapasiteten ikke eksisterer (Bartholomew et al., 2013). Det er samtidig verdt å nevne at både samplingteorien og gjensidighetsteorien likevel anerkjenner at psykometriske mål av g -faktoren fortsatt er relevante, ettersom de fanger opp interindividuelle forskjeller i intelligens, og har en viktig prediktiv funksjon for mange ulike utfall (Bartholomew et al., 2013; Kovacs & Conway, 2019; van der Maas et al., 2006). Mål for g -faktoren predikerer blant annet personens skoleprestasjoner, yrkesstatus, generelle helse og levetid, og er derfor sentralt og ofte inkludert både i forskning og klinisk arbeid, blant annet under utredning av lærevansker (Jensen, 2002).

2.4.5 Samlet oversikt over de ulike teoriene

For å sammenfatte de ulike teoriene om kognitiv utvikling og positive manifold, er viktige punkter som kjennetegner dem, og skiller dem fra hverandre, oppsummert i oversikten i tabell 1.

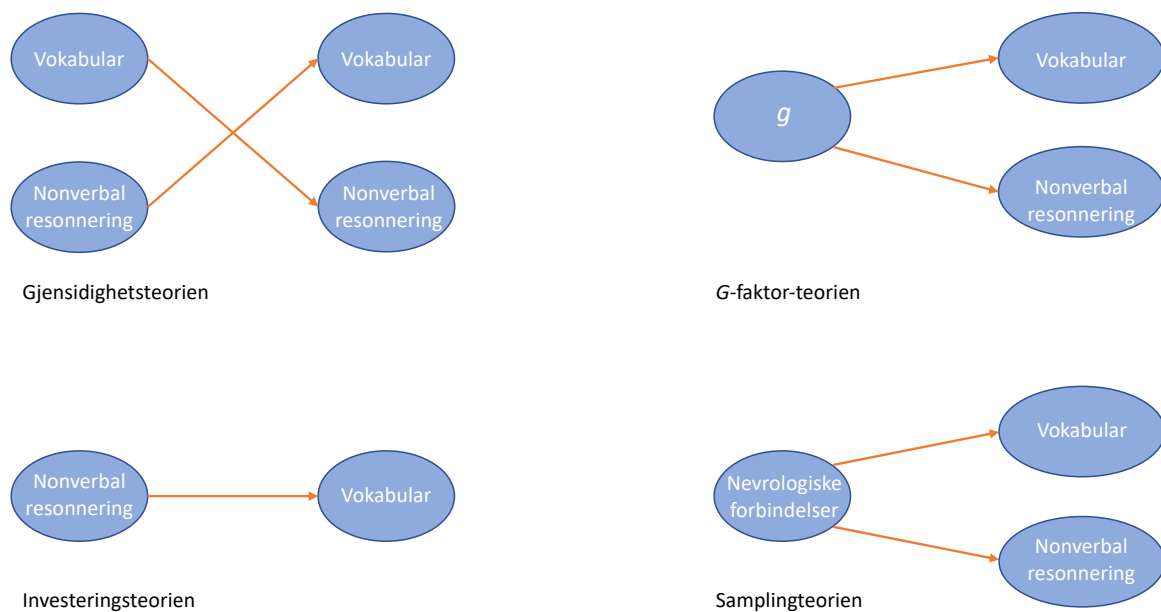
Tabell 1: Oversikt over teorier om kognitiv utvikling og positive manifold

	1	2	3	4
Positive manifold er et akseptert statistisk fenomen	✓	✓	✓	✓
Positive manifold oppstår over tid	✓			
Positive manifold er til stede også hos de yngste barna		✓	✓	✓
Positive manifold skyldes i stor grad at kognitive tester er noe upresise				✓
G-faktoren er et viktig statistisk fenomen	✓	✓	✓	✓
G-faktoren er et reelt psykologisk konstrukt		✓	✓	
G-faktoren er et helhetlig konstrukt som alle kognitive evner er avhengig av		✓		
G-faktoren er delt i flytende (nonverbal resonnering) og krystallisert intelligens (f.eks. vokabular), der førstnevnte påvirker utviklingen av sistnevnte			✓	
Alle kognitive ferdigheter påvirker hverandres utvikling	✓			
Merknad: 1 = Gjensidighetsteorien, 2 = G-faktor-teorien, 3 = Investeringsteorien, 4 = Samplingteorien				

2.4.6 Oversikt over kausalitet i teoriene

Spørsmålet om kausalitet er et særlig viktig område som skiller de overnevnte teoriene fra hverandre. I tillegg er kausalitet også et sentralt tema i denne masteroppgaven, ettersom det forsøkes å klargjøre om det er en gjensidig årsakssammenheng mellom vokabular og nonverbal resonnering. Derfor viser figur 2 en oversikt som nærmere tydeliggjør hvordan de ulike teoriene antar at de kausale forholdene mellom ulike kognitive ferdigheter er. Det fremgår i figur 2 at 1) gjensidighetsteorien antar at det er gjensidige kausale forhold mellom

vokabular og nonverbal resonnering, 2) *g*-faktor-teorien antar at *g*-faktoren er en underliggende tredjefaktor som påvirker både vokabular og nonverbal resonnering, 3) investeringsteorien antar at flytende intelligens (nonverbal resonnering) påvirker krystallisert intelligens (vokabular) og 4) samplingteorien antar at måten hjernen danner ulike neurologiske forbindelser på når man utfører kognitive oppgaver er årsaken til at vokabular og nonverbal resonnering korrelerer og fremtrer slik de gjør.



Figur 2: Kausalitet i de ulike teoriene

2.5 Oppsummering

I dette kapittelet er det gjort rede for begrepene «vokabular» og «nonverbal resonnering», og videre for vansker som knytter seg til disse begrepene: språkforstyrrelser, generelle lærevansker og utviklingshemming. Deretter ble ulike sentrale teorier om kognitiv utvikling, herunder utviklingen av vokabular og nonverbal resonnering, presentert. De ulike teoriene representerer også ulike forklaringer på fenomenet positive manifold. I gjennomgangen ovenfor kom det frem at det finnes noen grad av empirisk støtte for alle de forskjellige teoriene som ble presentert (gjensidighetsteorien, *g*-faktor-teorien, investeringsteorien og samplingteorien), selv om teoriene står for vidt ulike forklaringer på kognitiv utvikling og positive manifold.

Basert på denne gjennomgangen er det viktig å klargjøre mekanismene som ligger bak positive manifold og kognitiv utvikling, herunder også utviklingen av vokabular og nonverbal

resonnering. Særlig er det interessant å undersøke hvorvidt den for tiden svært aktuelle gjensidighetsteorien er en holdbar forklaringsmodell. Dette er også interessant fra et spesialpedagogisk perspektiv, ettersom det kan legge noen føringer for spesialpedagogisk forståelse og praksis. Formålet med dette masterprosjektet er som nevnt å bidra til å få mer klarhet rundt gjensidighetsteorien gjennom følgende problemstillinger:

- 1. I hvilken grad lar gjensidighetsteorien seg replikere på et norsk utvalg?*
- 2. I hvilken grad er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?*

3 Metode

I dette kapittelet følger en beskrivelse av metoden som har blitt brukt for å besvare forskningsspørsmålene redegjort for ovenfor.

Metodekapittelet har først en generell redegjørelse for kvantitativ metode og det vitenskapsteoretiske grunnlaget for kvantitativ forskning bygger på. Deretter følger en redegjørelse for forskningsetiske betraktninger, utvalg, studiedesign, datainnsamling, testinstrumenter, statistiske analyser og validitet.

Datasettet som brukes i dette masterprosjektet er hentet fra en tidligere undersøkelse utført ved Institutt for spesialpedagogikk ved Universitetet i Oslo, der en større kohort barn ble fulgt over tid og kartlagt med ulike testinstrumenter innen blant annet språk og nonverbal resonnering (Hjetland et al., 2019). Flere av de metodologiske valgene som er relevante for masteroppgaven ble tatt under den opprinnelige undersøkelsen. De metodologiske beslutningene knyttet til hvordan datamaterialet kan brukes for å svare på masteroppgavens problemstillinger, er likevel tatt av meg i samråd med mine veiledere.

3.1 Kvantitativ metode og vitenskapsteoretisk forankring

Denne masteroppgaven er en kvantitativ undersøkelse. Slike undersøkelser kjennetegnes primært av at man bruker talldata, variabler og statistikk for å besvare sine problemstillinger, i tillegg til at både fremgangsmåten og den vitenskapsteoretiske forankringen i kvantitativ forskning tilstreber å være mer objektiv enn i kvalitativ forskning (Creswell & Guetterman, 2021; Ringdal, 2018). I kvantitative forskningsartikler er det imidlertid ikke noen tradisjon for å gjøre rede for sitt vitenskapsteoretiske grunnlag (Ringdal, 2018, s. 54). Siden dette er en masteroppgave, vil jeg likevel kort gjøre dette nedenfor.

Tradisjonelt bygger vitenskapsteorien innen kvantitativ metode på positivismen eller en videreføring av denne (Nyeng, 2012; Ringdal, 2018, s. 54). Positivism er et kunnskapssyn med en lang historie som bygger på en målsetting om å gjøre nøytrale og objektive observasjoner, og på den måten å skaffe seg verifisert og sikker kunnskap gjennom forskning (Ringdal, 2018, s. 38). Likevel vil de fleste forskere som arbeider innenfor kvantitativ metode i samfunnsfagene i dag nok også anerkjenne at observasjoner ikke er objektive, at andre metoder kan bidra med kunnskap, og at forskningskunnskap ikke kan verifiseres en gang for alle (Ringdal, 2018, s. 39). Den delen av positivismen som likevel særlig består i kvantitativ

forskning, er ønsket om å være så upartisk og presis i forskningen som mulig (Nyeng, 2012). Dette etterstrebes også i dette masterprosjektet.

Moderne kvantitativ forskning bygger imidlertid i stor grad heller på kritisk realisme enn positivisme (Jakobsen, 2021). Kritisk realisme kjennetegnes først og fremst av antakelsen om at verden eksisterer uavhengig av menneskets kunnskap om den, og derfor at vår kunnskap om verden kan være feilaktig, og dessuten at det som eksisterer i verden overgår det som mennesket direkte kan observere (Jakobsen, 2021, s. 193). Derfor hevder man innen kritisk realisme at vitenskapelig kunnskap ikke utelukkende består av det man har observert objektivt, noe som skiller dette kunnskapssynet fra den mer tradisjonelle positivismen (Jakobsen, 2021, s. 194; Ringdal, 2018, s. 38). Metodologisk er kritisk realisme forbundet med mindre rigiditet enn positivismen, ved at man ikke favoriserer en enkelt type metode (som for eksempel observasjon), men heller søker en tverrfaglig og mangfoldig metodebruk som er tilpasset de forholdene man ønsker å se nærmere på i forskningen (Jakobsen, 2021, s. 255–256).

Nært beslektet med både positivisme og kritisk realisme er dessuten hypotetisk-deduktiv metode, hvor man har en hypotese og tester den med en undersøkelse som følger av hypotesen (Ringdal, 2018, s. 45). Dette er også relevant vitenskapsteori for denne masteroppgaven fordi det skal testes om gjensidighetsteorien kan styrkes, ved å se på mulige gjensidige forhold mellom vokabular og nonverbal resonnering.

Det er imidlertid viktig å merke seg at kvantitative prosjekter ikke er fullstendig løsrevet fra mer kvalitative vitenskapsteoretiske aspekter. Et eksempel på det er at Nyeng (2012, s. 79) hevder at ethvert kvantitativt prosjekt innledningsvis inneholder et kvalitativt element, fordi fenomenene som skal tallfestes, eksempelvis vokabular og nonverbal resonnering, først må defineres og tillegges en mening. Denne meningen er gjerne konstruert og fortolket av oss mennesker. I kvalitative metoder mener man nettopp at verden konstrueres (sosialkonstruktivisme) og fortolkes (hermeneutikk) av mennesker (Gilje, 2019; Thagaard, 2018). I Shadish, Cook og Campbells validitetssystem for eksperimentelle kvantitative undersøkelser (Shadish et al., 2002), som blir nærmere omtalt og brukt nedenfor, understrekes det også at mange av slutningene i en kvantitativ undersøkelse er av kvalitativ art, for eksempel slutninger knyttet til operasjonaliseringer av begreper og generaliseringer. Dette masterprosjektet har altså ikke utelukkende en vitenskapsteoretisk forankring i positivismen,

kritisk realisme eller hypotetisk-deduktiv metode, men fokuserer også på betydningen av kvalitative slutninger i en kvantitativ undersøkelse.

3.2 Forskningsetikk

Før datainnsamlingen i sin tid gikk i gang, ble prosjektet meldt inn til Norsk senter for forskningsdata (NSD) (Hjetland et al., 2019). I tillegg innhentet forskerne informert samtykke fra de foresatte til barna som skulle delta, slik det følger av retningslinjene til Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH, 2021). Datasettet ble senere anonymisert. Siden disse etiske reglene er ivaretatt, er det ikke behov for melding til Sikt (tidligere NSD) eller ytterligere informert samtykke i forbindelse med denne masteroppgaven. Samtidig har andre forskningsetiske hensyn vært av stor betydning i løpet av masterprosjektet. Blant annet har valg av metode og senere tolkning og vurdering av resultater blitt gjort med stor omhu. Det har vært viktig særlig fordi feilaktige eller upresise analysemetoder eller resultater kan føre til upålitelige implikasjoner for praksis og videre forskning.

3.3 Utvalg

Rekrutteringen til prosjektet foregikk i en kommune på Østlandet ved at alle barn i en alderskohort på i overkant av ett år ble invitert til å delta (ca. 250 barn). Kriteriene for invitasjonen var likevel norsk morsmål (siden undersøkelsens opprinnelige formål var språkutvikling hos typisk utviklede norskspråklige barn), i tillegg til at man ikke var tilmeldt PPT eller hadde kjente vansker som for eksempel språkforstyrrelser, utviklingshemming eller andre lærevansker (Hjetland & Johansen, 2009; Hjetland et al., 2019). Totalt ble 215 deltakere tatt inn i studien, derav 109 gutter og 106 jenter (Hjetland et al., 2019).

Den opprinnelige studien var et flerårig prosjekt, og som vi skal se nærmere på nedenfor ble det samme utvalget fulgt opp med gjentatte kartlegginger. Likevel var frafallet forholdsvis lite. Fra første måletidspunkt da barna var 4 år, og frem til siste måletidspunkt da barna var 9 år, var frafallet 14,9 %. Det meste av frafallet (rundt to tredeler) kom imidlertid av at deltakere ble tatt ut av studien fordi de startet på skolen senere enn størstedelen av utvalget, mens det resterende frafallet skyldtes fraflytting fra kommunen (Hjetland et al., 2019). Ved kartleggingstidspunktene som er relevante for denne masteroppgaven var barna henholdsvis 4,3 og 6,3 år gamle, og utvalgsstørrelsen på de to tidspunktene var henholdsvis 201 og 200.

Generaliserbarhet vil si i hvilken grad utvalget representerer den populasjonen det er ment å reflektere. Her er målet at utvalget representerer norske barn og unge i alderen 4 til 6 år. En viktig forutsetning for generaliserbarhet er at utvalget er tilfeldig trukket fra populasjonen (Ringdal, 2018, s. 208). Det betyr at man ikke med sikkerhet kan si at resultatene fra masteroppgaven kan generaliseres til alle norske barn i den aktuelle alderskohorten, ettersom deltakerne ble selektert fordi de kom fra samme geografiske område. Likevel ble den aktuelle østlandskommunen valgt ut fordi den er en gjennomsnittlig norsk kommune når det gjelder sosioøkonomisk status (Hjetland et al., 2019). Derfor kan det argumenteres for at utvalget er relativt representativt for norske barn for øvrig. Likevel bør man merke seg at det ikke er et randomisert trukket utvalg, så det kan være variabler som ikke har blitt målt her, hvor utvalget skiller seg fra populasjonen. Men siden vi her snakker om forholdet mellom kognitive ferdigheter, antas resultatene å være rimelig generaliserbare. Det er fordi variasjoner i slike ferdigheter vil finnes overalt i landet, og fordi sosioøkonomisk status ser ut til å være relativt representativt.

3.4 Studiedesign

Studiedesignet til masteroppgaven er et prospektivt longitudinelt design. I en longitudinell studie undersøker man trender eller utvikling ved at man innhenter data på ulike måletidspunkter (Creswell & Guetterman, 2021, s. 432). Det finnes ulike måter å utføre en longitudinell studie på. Prospektive longitudinelle studier kjennetegnes av at man samler inn data om de samme informantene fra rekrutteringstidspunktet og fremover i tid (Creswell & Guetterman, 2021, s. 433; Ringdal, 2018, s. 158). Siden studiedesignet innebærer at det finnes målinger om de samme informantene på ulike tidspunkter, er det mulig å følge utviklingen av blant annet vokabular og nonverbal resonnering over tid. Det blir også mulig å se på hvordan de ulike variablene kan påvirke hverandre over tid.

Masteroppgaven er dessuten en observasjonell studie, det vil si ikke en eksperimentell studie. Eksperimenter kjennetegnes av at man gir deltakerne en intervensjon og måler effekten, og helst er deltakerne randomisert i en eksperimentgruppe og en kontrollgruppe (Ringdal, 2018, s. 132). I observasjonelle studier innfører man derimot ikke noen form for intervensjon eller tiltak. Eksperimenter er kjent som den klart sikreste måten å studere kausalitet på, fordi man under kontrollerte forhold har muligheten til å manipulere «årsaksvariabelen» ved å introdusere en intervensjon, for så å måle effekten på «virkningsvariabelen». Selv om masteroppgaven er en observasjonsstudie, er likevel målet også her å se på mulige kausale

sammenhenger. Dette krever imidlertid at det man må forsøke å kompensere for det man mister som følge av at designet ikke er eksperimentelt. Dette vil gjøres ved hjelp av statistiske modeller.

3.5 Datainnsamling

Datainnsamlingen startet i desember 2007, da barna var rundt 4 år. Deretter ble de samme barna fulgt opp og kartlagt årlig. Totalt ble de kartlagt i seks runder, med omtrent ett år mellom hver runde. Her brukes som nevnt data fra målepunktene ved 4 og 6 år.

Datainnsamlingsprosedyren foregikk som individuelle kartlegginger i barnets egen barnehage eller skole. Testene ble administrert av forskningsassistenter som hadde fått nødvendig opplæring. For å sikre mest mulig likhet mellom testsituasjonene til alle barna, ble testene utført i en bestemt rekkefølge. I noen av rundene med datainnsamling var det behov for flere økter med testing per barn. Hver enkelt økt med testing varte i 30–45 minutter da barna var små, og 60 minutter etter hvert som de ble litt eldre. Det ble brukt ulike godt dokumenterte og etablerte tester innen vokabular, lesing, skriving, lytteforståelse og kognitive evner. I de neste avsnittene beskrives de av testinstrumentene som er brukt i dette masterprosjektet.

3.6 Testinstrumenter

3.6.1 Vokabular

Barnas vokabular ble kartlagt med den norske oversettelsen (Lyster et al., 2010) av British Picture Vocabulary Scale, BPVS II (Dunn et al., 1997). Heretter kalles denne testen BPVS. BPVS måler vokabularet til barnet ved at en testleder sier ett og ett ord, hvorpå barnet skal velge ett og ett bilde som innholdsmessig passer best til ordet som testlederen uttaler. Testen er tilpasset barn i aldersgruppen 3–16 år. I den norske utgaven er det 12 deler, og 12 oppgaver i hver del. Som det går frem av testprosedyren, er det særlig reseptivt vokabular, altså evnen til å forstå ord uten å selv måtte uttrykke ordene, som måles med BPVS. BPVS blir regnet for å ha god reliabilitet og validitet når man måler barns vokabular. Det er blant annet fordi de britiske og amerikanske variantene har god validitet, og fordi det er en korrelasjon mellom norske BPVS og andre språktester (Lyster et al., 2010). En nærmere diskusjon av reliabilitet og validitet knyttet til BPVS i dette prosjektet følger i kapittel 5.

3.6.2 Nonverbal resonnering

Terningmønster, eller «Block Design», er mye brukt i tidligere forskning, og inngår i dette prosjektet som mål på nonverbal resonnering (Hjetland et al., 2019). Det er en test som er

spesielt utformet for å vurdere kognitive evner uten å være avhengig av språk (Kohs, 1920). Testen inngår som en deltest i de mye brukte Wechsler-testene som måler intelligens (se for eksempel Wechsler, 2009), og den kan dermed anses som et anerkjent måleverktøy.

Terningmønster går ut på at man ved hjelp av en- eller tofargede terninger skal kopiere ulike mønstre man får presentert. Det betyr at testpersonen får løse et problem uten å bruke språket, men det gis verbale testinstruksjoner. Som flere andre tester av nonverbal resonnering, krever Terningmønster både induktiv og deduktiv resonnering (Peng et al., 2019). Det betyr at man både skal oppdage regler i et mønster (induktivt) og anvende disse reglene for sette sammen et tilsvarende mønster (deduktivt) (Sternberg et al., 2008, s. 190–191). Reliabilitet og validitet knyttet til Terningmønster i dette prosjektet vil bli drøftet i kapittel 5 i masteroppgaven.

3.7 Definerings og identifisering av «vanskegrupper»

For å besvare problemstilling nummer 2 «I hvilken grad er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?» må deltakerne som strever med vokabular og/eller nonverbal resonnering identifiseres. Som nevnt har ingen av deltakerne utviklingshemming, språkforstyrrelser eller andre kjente, diagnostiserte lærevansker. I dette masterprosjektet er vanskegruppene operasjonalisert ved hjelp av to undergrupper, «vanskegruppe vokabular» og «vanskegruppe nonverbal resonnering», med de av deltakerne som skårer lavere enn 31 på BPVS og 20 på Terningmønster ved testing i 4-årsalderen.

Cut-off-punktet for å havne i en av vanskegruppene i denne undersøkelsen settes til 15 %, noe som tilsvarer omtrent –1 standardavvik under gjennomsnittet for utvalget. Som redegjort for i kapittel 2.3.1 er dette vanlig i forskningssammenheng og i psykometri. Det vil si at barn som hadde en skår på BPVS og/eller Terningmønster som lå på 15. persentil eller lavere, blir plassert i henholdsvis «vanskegruppe vokabular» og/eller «vanskegruppe nonverbal resonnering». 15 % er satt som cut-off fordi det antas at vanskegruppene da vil utgjøre en passe stor andel av det totale utvalget, samt at barna som havner i disse gruppene vil ha en del utfordringer knyttet til vokabular eller nonverbal resonnering. Likevel er det flytende overganger mellom «normalfungering» og «vansker» både når det gjelder vokabular og nonverbal resonnering, og det finnes ikke en fasit for hvor man skal sette cut-off (Peters & Ansari, 2019).

Når det gjelder begrepsbruk, kalles disse gruppene vanskegrupper fordi betegnelse «språkforstyrrelser» eller «utviklingshemming» ikke var aktuelle, ettersom dette var eksklusjonskriterier i studien. Benevnelsen «generelle lærevansker» var aktuell som alternativ til «vanskegruppe nonverbal resonnering», men ble valgt bort fordi barn med generelle lærevansker, som navnet tilsier, har mer generelle utfordringer, mens det her er nettopp nonverbal resonnering som er i fokus.

Selv om det både her og generelt innenfor spesialpedagogisk forskning ofte deles inn i vanskegrupper på denne måten, bør det nevnes at dette er forbundet med noen betydelige metodologiske utfordringer. Spesielt er det problematisk at siden utvalgsstørrelsen ofte blir lav, så vil testene man utfører få redusert statistisk styrke (Toffalini et al., 2021). Statistisk styrke betyr her sannsynligheten for at man konkluderer med at man har funnet en sammenheng som faktisk eksisterer (Creswell & Guetterman, 2021, s. 227). Det vil si at lav statistisk styrke fører til at analysene i mindre grad vil være i stand til å avdekke reelle sammenhenger mellom variablene i undersøkelsen. Som vi skal se nærmere på nedenfor i resultatdelen, medfører dessuten oppdelingen av utvalget i denne undersøkelsen at reliabiliteten til særlig noen av variablene i vanskegruppene blir lav, noe som igjen fører til at analysene med disse variablene må tolkes med en viss forsiktighet.

3.8 Korrelasjon og lineær regresjon

Analysemetoden i denne masteroppgaven er i hovedsak *Cross-Lagged Panel Models*, i tillegg til at det blir anvendt en *Latent Change Score Model*. De to ulike modellene blir nærmere forklart nedenfor. Fordi korrelasjon og lineær regresjon er enklere statistiske analyser som begge modellene bygger på, vil dette kort bli omtalt først.

Korrelasjon er sammenhengen mellom variabler (Ringdal, 2018, s. 313), og er et mål som gir informasjon om hvorvidt verdiene til to variabler henger sammen med hverandre. Korrelasjon er ikke ensbetydende med kausalitet, altså at den ene variabelen er årsaken til den andre. I tillegg til korrelasjon mellom variablene, har kausalitet også andre kriterier. Blant annet må «årsaksvariabelen» tidsmessig inntreffe før «virkningsvariabelen», og andre mulige forklaringer på sammenhengen mellom dem må være utelukket (Ringdal, 2018, s. 50). Når data er samlet på flere tidspunkter, for eksempel i et longitudinelt design, kan man sikre at «årsaksvariabelen» inntreffer tidligere enn «virkningsvariabelen». Derfor var det tidligere vanlig å tolke seg frem til om det var gjensidige forhold mellom variabler ved å se på rene korrelasjonsmål på tvers av minst to variabler som var målt på ulike tidspunkter (Kearney,

2017). Som vi skal se nærmere på nedenfor, er denne metoden mer upresis enn blant annet *Cross-Lagged Panel Models*, fordi man i mindre grad tar høyde for at korrelasjonen kan skyldes noe annet enn kausalitet.

I likhet med korrelasjon gir lineær regresjon et tallmessig estimat på styrken på sammenhengen mellom variabler. Gjennom regresjon kan sammenhengen uttrykkes ved hjelp av en ligning, noe som gjør at regresjonsanalyser kan brukes til å predikere verdien på én variabel, basert på verdien til en eller flere andre variabler (Ringdal, 2018, s. 405).

Tradisjonelt har regresjonsanalyse vært mye brukt innenfor kvantitativ metode. Vanlige regresjonsmodeller har imidlertid noen begrensninger, blant annet at de ofte er basert på at man predikerer ett utfallsmål om gangen. Dette er uheldig i dette prosjektet fordi man ikke direkte kan kombinere resultatene fra en regresjonsanalyse med ett utfallsmål (f.eks. vokabular) med en regresjonsanalyse med et annet utfallsmål (f.eks. nonverbal resonnering). Derfor vil ikke en vanlig regresjonsmodell kunne svare på hvorvidt det er et gjensidig forhold mellom vokabular og nonverbal resonnering.

3.9 Cross-Lagged Panel Models

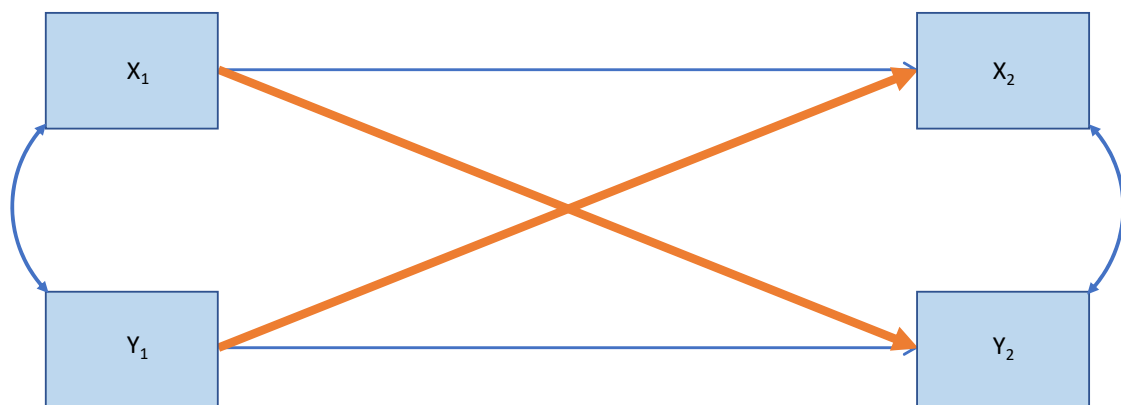
Cross-Lagged Panel Models, heretter også omtalt som CLPM, tilhører en gruppe analyser som kalles «*Structural Equation Modeling*» («SEM»). Hovedpoenget med SEM er at man lager ulike statistiske modeller som gir informasjon om korrelasjonen og korrelasjonsretningen mellom variabler (Creswell & Guetterman, 2021, s. 403). Statistisk bygger SEM på regresjonsanalyse. Fordelen med SEM sammenlignet med en vanlig lineær regresjonsmodell som beskrevet over, er likevel at man kan se på mer komplekse sammenhenger, og på flere utfallsmål samtidig.

CLPM undersøker hvordan minst to ulike variabler, som er målt på minst to tidspunkter, gjensidig påvirker hverandre over tid (Kearney, 2017). Metoden går ut på at man ser på hvordan variabler i et datasett forholder seg til hverandre, ved å plassere dem inn i analysemodellen. I praksis gjøres dette ved hjelp av analyseprogrammer. I denne masteroppgaven brukes Stata til å utføre CLPM. Her brukes CLPM med to ulike variabler er målt på to ulike tidspunkter, men modellen kan også brukes hvis man har flere variabler og/eller flere måletidspunkter.

CLPM er et sett med regresjonsligninger som inneholder minst to variabler målt på minst to tidspunkter (Selig & Little, 2012). Altså har man variablene X_1 , X_2 , Y_1 og Y_2 . I likhet med

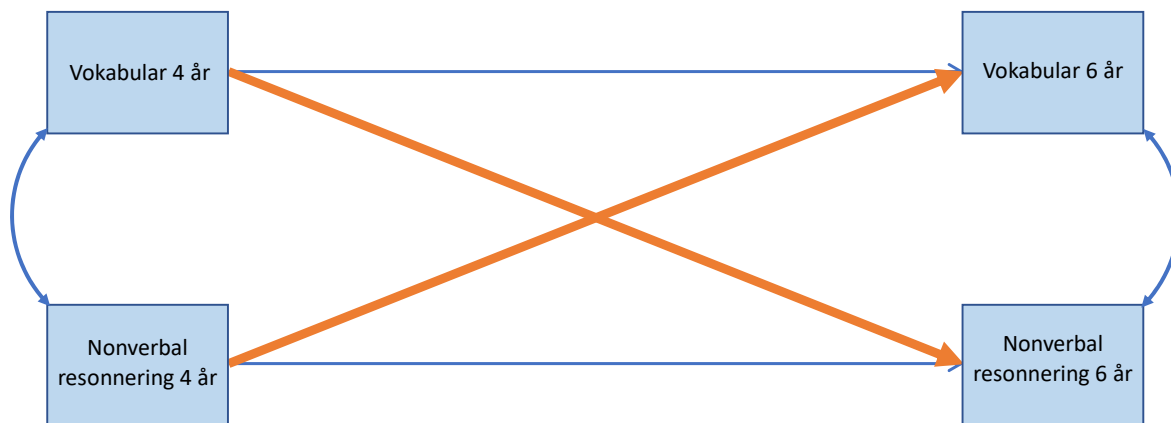
vanlig lineær regresjon, er det i CLPM regresjonskoeffisienter som reflekterer den statistiske sammenhengen mellom 1) Y_2 og X_1 og 2) X_2 og Y_1 . Samtidig kontrollerer regresjonsligningene for at hver enkelt av variablene kan ha sammenheng med sin egen verdi over tid, altså kontrolleres det for stabiliteten til variablene X og Y. Det kontrolleres også for at variablene X og Y kan korrelere med hverandre også når de er målt samtidig (Kearney, 2017). Dette gjør at CLPM blir en mer presis metode når man ser på gjensidige forhold mellom variabler, sammenlignet med for eksempel regresjonsmodeller med flere prediktorer og ett utfallsmål.

Figur 3 illustrerer CLPM. De diagonale oransje pilene er en illustrering av forholdet mellom 1) Y_2 og X_1 og 2) X_2 og Y_1 . Samtidig viser figuren at CLPM kontrollerer for korrelasjonen mellom X og Y når begge er målt på samme tidspunkt, og for stabiliteten til både X og Y over tid.



Figur 3: Cross-Lagged Panel Model (CLPM)

Figur 4 viser hvordan CLPM anvendes i denne masteroppgaven. Det som blir sett på her er utvikling i vokabular og nonverbal resonnering. Modellen kan besvare spørsmålet om hvorvidt nonverbale ferdigheter påvirker vokabularutvikling mellom 4 og 6 år, og tilsvarende for nonverbale ferdigheter, hvorvidt vokabular forklarer utviklingen av nonverbal resonnering mellom 4 og 6 år. Målingene av vokabular gjøres som nevnt med BPVS, mens målingene av nonverbal resonnering gjøres med Terningmønster.



Figur 4: Vokabular og nonverbal resonnering i CLPM

Det er imidlertid også viktig å påpeke at CLPM har begrensninger. Den viktigste begrensningen er at CLPM brukes til å analysere data i observasjonelle studier, og derfor at det alltid vil være en viss tvil knyttet til kausale slutninger basert på CLPM (Rogosa, 1980). Det er fordi et randomisert eksperiment som tidligere nevnt er den eneste sikre måten å teste en kausal sammenheng på. Av ulike grunner kan det likevel være at eksperimentelle studier, som her, ikke er gjennomførbare, og følgelig må kausale slutninger i forskning ofte basere seg på observasjonsstudier (Rubin, 2007). Selv om man uten eksperimenter ikke kan trekke sikre konklusjoner om kausalitet, kan man fortsatt forsøke å komme så nært som mulig, blant annet gjennom valg av statistiske analysemetoder (Rohrer, 2018). CLPM (og LCSM som omtales nedenfor) er forsøk på dette.

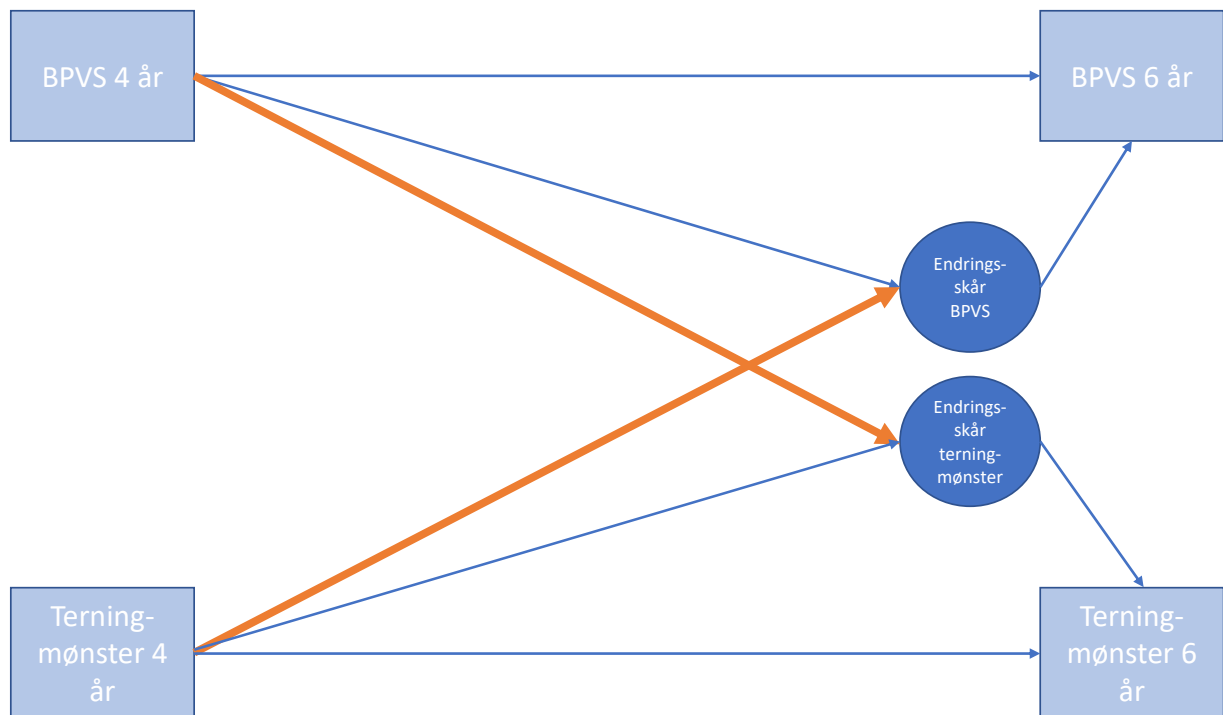
Utover utfordringen med kausale slutninger i observasjonelle studier, har CLPM også noen andre begrensninger. Blant annet er CLPM en analysemodell som ikke håndterer målefeil i variablene så godt, slik at målefeil kan påvirke konklusjonene i analysen (Rogosa, 1980). I tillegg hviler CLPM på en rekke antakelser som kanskje ikke i realiteten kan oppfylles, som for eksempel at alle mulige prediktorer er inkludert i modellen, at variablene på hvert måletidspunkt er målt samtidig og at innholdet i begrepene som måles er uendret over tid (Kearney, 2017). Likevel er det anerkjent at CLPM kan gi gode forskningsbidrag om blant annet kommunikasjon og kognitiv utvikling (Kearney, 2017; Selig & Little, 2012). Derfor er CLPM valgt som en analysemetode i masterprosjektet.

3.10 Latent Change Score Models

For å teste holdbarheten til resultatene med CLPM som analysemetode, og for å se nærmere på forholdet mellom vokabular og nonverbal resonnering, vil det i dette prosjektet også bli brukt en analysemetode kalt «*Latent Change Score Models*» eller «LCSM». Årsaken til at det er ønskelig å bruke LCSM som supplement til CLPM er at det vil gi en viss kontroll på målefeil som CLPM ikke håndterer. LCSM er også en type analyse som er mye brukt tidligere for å studere gjensidighetsteorien (Griffiths et al., 2022; Kievit et al., 2019; Kievit et al., 2017). På grunn av at utvalgsstørrelsen blir noe lav når bare vanskegruppene inngår i analysene, og fordi det av den grunn antas at det kan bli vanskelig å få statistisk signifikante resultater i vanskegruppene, er det valgt å bare bruke LCSM i analyser av hele utvalget.

I likhet med CLPM er LCSM også en type SEM som bygger på regresjonsanalyser. Det finnes imidlertid mange ulike varianter av LCSM. I varianten som brukes her er det to ulike observerte variabler målt på to ulike tidspunkter, altså BPVS 4 og 6 år og Terningmønster 4 og 6 år. I tillegg til disse variablene inngår det to latente variabel i modellen, som angir endringen i verdi på BPVS og endringen i verdi på Terningmønster. At variablene er latente vil si at de ikke er direkte målt i datainnsamlingen, men at verdien finnes gjennom matematisk manipulering av andre variabler i datasettet. Verdiene på de latente variablene i denne modellen blir «BPVS 6 år – BPVS 4 år» og «Terningmønster 6 år – Terningmønster 4 år». På engelsk har disse latente variablene navnene «change score», derav navnet til analysemodellen. Her brukes «endringsskår» som en norsk oversettelse.

Det interessante i denne analysen er sammenhengen mellom 1) endringsskår Terningmønster og BPVS ved 4 år og 2) endringsskår BPVS og Terningmønster ved 4 år. Med andre ord er målet med disse analysene å finne ut om vokabular kan påvirke endring i nonverbal resonnering over tid, og omvendt om nonverbal resonnering kan påvirke endring i vokabular over tid. Figur 5 er en forenklet illustrasjon av LCS-modellen som brukes her, hvor igjen de oransje pilene er de statistiske forholdene som er i fokus.



Figur 5: Latent Change Score Models

Til denne analysen brukes RStudio som analyseprogram. Den aktuelle fremgangsmåten og analysekoden er presentert av Kievit et al. (2018). Det er analysekoden for «Bivariate Latent Change Score Model» med R som vil bli brukt, men med tilpasninger til masteroppgavens datasett.

3.11 Validitet

Validitet handler om hvorvidt konklusjonene man trekker i prosjektet er korrekte og sanne (Creswell & Guetterman, 2021, s. 345). Her følger en kort presentasjon av validitetstypologien til Shadish et al. (2002), som senere i kapittel 5 skal brukes til å vurdere validiteten i masterprosjektet. Typologien er i utgangspunktet utviklet for eksperimentelle studier. Selv om masteroppgaven er en observasjonell studie, har den likevel til felles med eksperimentelle studier at den handler om kausalitet. Derfor er typologien til Shadish og kolleger et relevant og interessant rammeverk for å vurdere validiteten også her.

Validitetstypologien består av fire typer validitet: statistisk validitet, indre validitet, begrepsvaliditet og ytre validitet.

Statistisk validitet er validiteten til slutningene man trekker om korrelasjonen mellom variablene i prosjektet (Shadish et al., 2002). Litt forenklet handler dette om å vurdere om det

er en sammenheng mellom variablene, og hvor sterk denne sammenhengen er. Sentrale begreper her er for det første «effektstørrelse» eller «forklart varians», noe som viser til hvor stor forklaringsgrad en eller flere variabler uavhengige har på en avhengig variabel (Ringdal, 2018, s. 409). Et annet sentralt begrep her er «statistisk signifikans». Dette begrepet knytter seg til nullhypotesetesting, og brukes når nullhypotesen i en undersøkelse skal forkastes. Nullhypotesen er et utsagn om populasjonen, og innebærer som regel at det man undersøker i en studie *ikke* stemmer (Ringdal, 2018, s. 285). Motsatsen er den alternative hypotesen, som innebærer det motsatte. Nullhypotesen forkastes oftest når man i utvalget gjør et funn som det er maksimalt 5 % sjans for å finne hvis nullhypotesen faktisk er sann. At man setter 5 % som grense betyr at signifikansnivået er satt til 0,05, og man tolererer bare p-verdier på 0,05 eller lavere for å kunne si at man har et statistisk signifikant funn. 0,05 er det vanligste signifikansnivået (Ringdal, 2018, s. 286).

Indre validitet handler om kausaliteten mellom variablene, og hvorvidt slutninger man trekker angående kausalitet innad i utvalget er gyldige (Shadish et al., 2002). Siden masteroppgaven som nevnt ikke er en eksperimentell studie, er dette særlig utfordrende. I vurderingen av prosjektets indre validitet, vil det derfor særlig bli lagt vekt på hvordan det i dette prosjektet er forsøkt å kompensere for de truslene mot indre validitet som oppstår i observasjonsstudier hvor man mangler eksperimentell kontroll.

Begrepsvaliditet viser til hvorvidt variabler i et datasett faktisk representerer de begrepene de er ment å måle (Shadish et al., 2002). Med andre ord handler det om hvorvidt man har klart å operasjonalisere begrepene i en studie med måleverktøyene man har brukt. I dette prosjektet er det derfor spesielt sentralt å vurdere om BPVS og Terningmønster er tilstrekkelige og rimelige operasjonaliseringer av «vokabular» og «nonverbal resonnering». Under begrepsvaliditet i dette validitetssystemet inngår også begrepet «reliabilitet». Reliabilitet viser til hvorvidt målinger er stabile og pålitelige, og med lite tilfeldige målefeil, slik at en gjentatt måling vil gi det samme resultatet (Ringdal, 2018, s. 103). Årsaken til at reliabilitet går inn under begrepsvaliditet i validitetssystemet som brukes her, er at et mål med dårlig reliabilitet aldri heller kan ha god begrepsvaliditet, fordi det vil måle mye støy og dermed ikke gi en god representasjon av begrepet man skal måle. I dette prosjektet brukes Cronbachs alpha (α) for å avgjøre variablenes reliabilitet. Dette er et mål på intern konsistens mellom «items» som sammen utgjør de ferdige variablene BPVS og Terningmønster. Man ønsker en Cronbachs alpha på minst 0,70 (Ringdal, 2018, s. 367).

Ytre validitet er den siste typen validitet i typologien til Shadish et al. (2002). I likhet med indre validitet handler også ytre validitet om årsakssammenhenger. Her er man imidlertid interessert i å avgjøre hvorvidt slutninger om kausalitet i prosjektets utvalg også kan generaliseres og gjelde utenom utvalget. I dette prosjektet handler ytre validitet altså om hvorvidt resultatene i utvalget også kan sies å gjelde for norske barn generelt i alderen 4 til 6 år.

4 Resultater

I dette kapittelet presenteres funnene fra undersøkelsen. Først presenteres deskriptiv statistikk og forberedende analyser som ble gjort for å skaffe en oversikt over variablene i prosjektet, samt for å klargjøre for videre analyser. Deretter følger resultatene av analysene som brukes direkte for å svare på problemstillingene. Først presenteres da resultatene fra CLPM og LCSM for hele utvalget. Disse analysene er særlig knyttet opp mot problemstilling nummer 1:

1. *I hvilken grad lar gjensidighetsteorien seg replikere på et norsk utvalg?*

Deretter presenteres resultatene fra CLPM med henholdsvis bare vanskegruppe vokabular og vanskegruppe nonverbal resonnering. Disse analysene er særlig knyttet opp mot problemstilling nummer 2:

2. *I hvilken grad er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?*

4.1 Deskriptiv statistikk og forberedende analyser

4.1.1 Hele utvalget

Tabell 2 er en oversikt over deskriptiv statistikk og psykometriske egenskaper (utvalgsstørrelse for målet, laveste og høyeste skår, gjennomsnitt, standardavvik og Cronbachs alpha) for testing med BPVS og Terningmønster målt ved 4 og 6 år. Det som særlig står ut her er at reliabilitetsmålene for BPVS er stabilt høye, men at reliabiliteten til Terningmønster er lavere, særlig ved testing ved 4 år. I tillegg kommer det frem i tabellen at spredningen i skårer er større for BPVS enn for Terningmønster. Dette henger sammen med at BPVS er et mer fingradert måleverktøy enn Terningmønster. Tabellen viser også at utvalgsstørrelsen varierer litt på de ulike måletidspunktene. I analysene er «missing values» håndtert ved at det er brukt «listwise deletion», det vil si at når en verdi mangler på et barn på en test, slettes alle data på dette barnet.

Tabell 2: Deskriptiv statistikk og psykometriske egenskaper for BPVS og Terningmønster (hele utvalget)

	N	Min.	Max.	Gj.snitt	Std.avvik	α
BPVS 4 år	195	15	75	42.28	11.212	0.911
BPVS 6 år	190	43	104	73.39	11.945	0.913
Terningmønster 4 år	196	12	34	22.05	3.232	0.560
Terningmønster 6 år	190	19	40	28.94	4.186	0.703

I tabell 3 presenteres korrelasjoner (Pearsons r) mellom de totalt fire variablene. Det var svake til moderate positive korrelasjoner mellom alle de ulike variablene. Alle korrelasjonene var statistisk signifikante. Det betyr at p-verdiene i tabellen maksimalt er 0,05. Korrelasjonene mellom variablene som ble målt på samme tid var 0,293 (BPVS 4 år og Terningmønster 4 år) og 0,213 (BPVS 6 år og Terningmønster 6 år). Korrelasjonene mellom samme variabel målt på ulike tidspunkter var 0,447 for BPVS og 0,455 for Terningmønster. Dette viser til variablenes stabilitet over tid. Longitudinelle korrelasjoner på tvers av type variabel var 0,160 (mellom Terningmønster 6 år og BPVS 4 år) og 0,246 (mellom BPVS 6 år Terningmønster 4 år).

Tabell 3: Korrelasjonstabell med BPVS og Terningmønster

	1	2	3	4
1. BPVS 4 år	–			
2. BPVS 6 år	0.447**	–		
3. Terningmønster 4 år	0.293**	0.246**	–	
4. Terningmønster 6 år	0.160*	0.213**	0.455**	–

Korrelasjonsmål er Pearsons r

** Signifikant på 0.01-nivå

* Signifikant på 0.05-nivå

4.1.2 Vanskegrupper

Cut-off-punktet for å havne i en av vanskegruppene i denne undersøkelsen var 15 % lavest presterende. Tabell 4 viser at 15. persentilene for BPVS og Terningmønster målt ved 4 år var henholdsvis 31 og 20. De av deltakerne som hadde en skår på 31 eller lavere defineres altså her som vanskegruppe vokabular, mens de som hadde en skår på 20 eller lavere defineres som vanskegruppe nonverbal resonnering. Rent praktisk ble dette regnet ut ved hjelp av SPSS.

Tabell 4: Variabelverdier på 15. persentil

	15. persentil
BPVS 4 år	31 (n = 33)
Terningmønster 4 år	20 (n = 64)

Tabell 5 inneholder deskriptiv statistikk og psykometriske egenskaper for variablene BPVS og Terningmønster når det bare er sett på henholdsvis vanskegruppe vokabular og vanskegruppe nonverbal resonnering. Her er det spesielt viktig å påpeke at inndelingen i vanskegrupper gjør at reliabilitetsmålene for variablene i fireårsalderen har gått betraktelig ned. Særlig gjelder dette Terningmønster ved 4 år. Dette er noe som vil diskuteres nærmere i kapittel 5.

Tabell 5: Deskriptiv statistikk og psykometriske egenskaper for BPVS og Terningmønster (vanskegrupper)

	N	Min.	Max.	Gj.snitt	Std.avvik	α
Vanskegruppe vokabular						
BPVS 4 år	33	15	31	26.79	4.574	0.642
BPVS 6 år	31	43	87	64.94	11.275	0.910
Terningmønster 4 år	33	13	26	20.85	3.022	0.441
Terningmønster 6 år	31	20	38	28.84	4.612	0.756
Vanskegruppe nonverbal resonnering						
BPVS 4 år	64	15	63	37.72	9.358	0.885
BPVS 6 år	61	46	99	69.67	10.568	0.893
Terningmønster 4 år	64	12	20	18.58	2.188	0.081
Terningmønster 6 år	61	19	36	26.98	3.263	0.565

4.2 Lar gjensidighetsteorien seg replikere på et norsk utvalg?

4.2.1 CLPM hele utvalget

SEM-builder i Stata ble brukt for å undersøke hvordan dataene henger sammen når de plasseres i en *Cross-Lagged Panel Model*. For å svare på problemstilling 1 «I hvilken grad lar gjensidighetsteorien seg replikere på et norsk utvalg?» er det brukt data fra hele utvalget i modellen. Tabell 6 oppsummerer det som kom frem i denne analysen.

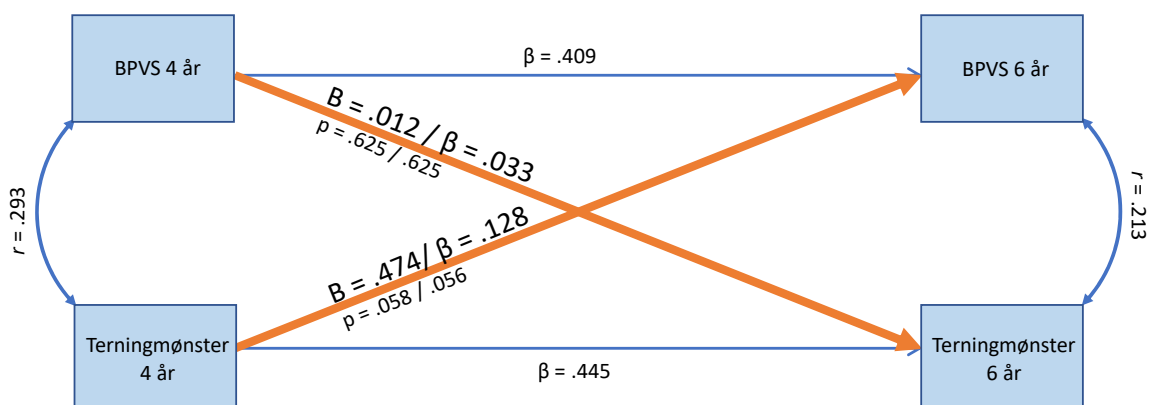
Regresjonskoeffisientene i tabell 6 presenteres både i standardisert (β) og ikke-standardisert (B) form. Begge deler sier noe om forholdet mellom variablene, men hvis man skal sammenligne forholdene mellom de ulike variablene i modellen med hverandre, må man se på

de standardiserte koeffisientene (Kearney, 2017). Som vi kan se av tabell 6 var de longitudinelle regresjonskoeffisientene mellom BPVS og Terningmønster ikke signifikante. Dette gjelder både forholdet mellom Terningmønster 4 år og BPVS 6 år, samt mellom BPVS 4 år og Terningmønster 6 år. Likevel utpekte forholdet mellom BPVS 4 år og Terningmønster 6 år seg som spesielt svakt.

Tabell 6: CLPM BPVS og Terningmønster, resultater fra hele utvalget

		B	β
BPVS 6 år	Terningmønster 4 år	0.474 (p = .058)	0.128 (p = .056)
	BPVS 4 år	0.436 (p < .001)	0.409 (p < .001)
Terningmønster 6 år	Terningmønster 4 år	0.576 (p < .001)	0.445 (p < .001)
	BPVS 4 år	0.012 (p = .625)	0.033 (p = .625)

De samme resultatene oppsummeres også visuelt i figur 6. Det gjøres her oppmerksom på at regresjonskoeffisientene mellom samme kognitive ferdighet målt på ulike tidspunkter, samt korrelasjonskoeffisientene mellom ulike ferdigheter målt på samme tidspunkt, alle var signifikante på 0.01-nivå.



Figur 6: CLPM BPVS og Terningmønster, resultater fra hele utvalget

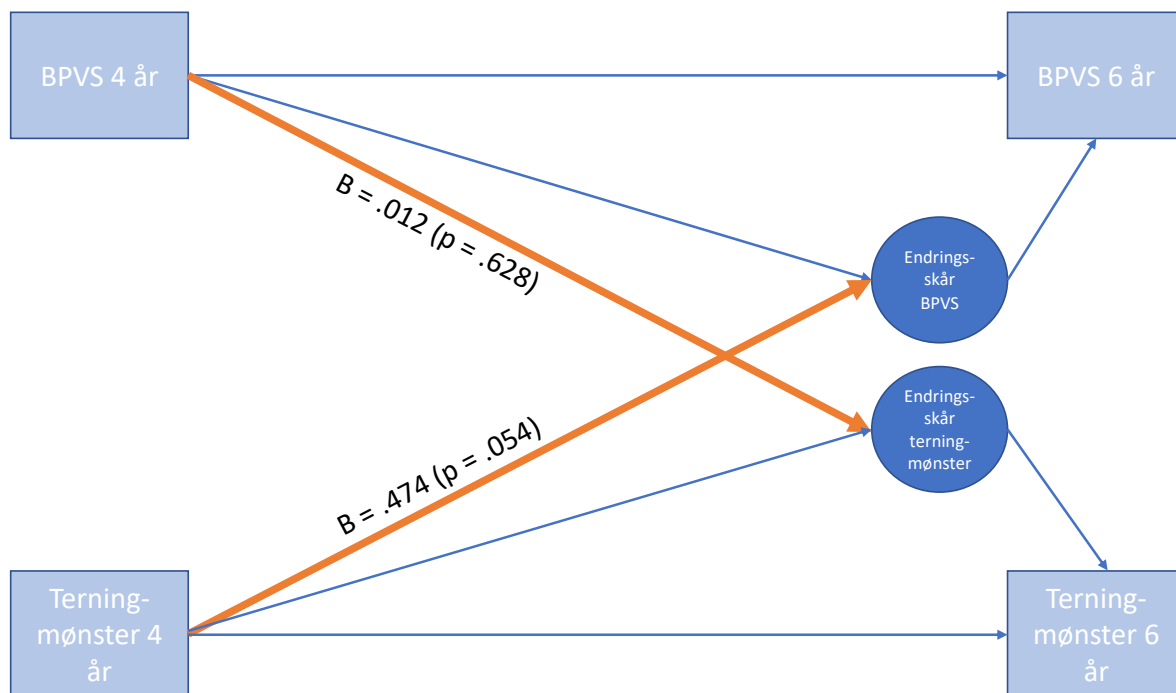
4.2.2 LCSM hele utvalget

De samme dataene ble også analysert i en LCS-modell ved hjelp av RStudio. Dette for å teste problemstillingen med en annen modell, samt for å undersøke holdbarheten til resultatene som kom frem med CLPM. Analysen med LCSM ble bare gjort med hele utvalget, slik at analysen ikke ble gjentatt for undergruppene vanskegruppe vokabular eller vanskegruppe nonverbal resonnering senere. Grunnen til det var at utvalgsstørrelsen er noe lav når bare vanskegruppene er inkludert, og fordi det derfor ble antatt at det ville bli vanskelig å få statistisk signifikante resultater i vanskegruppene. Resultatene fra LCSM vises i tabell 7. Ingen av regresjonskoeffisientene i denne analysen var signifikante, og i likhet med det vi så for CLPM, er det spesielt samvariasjonen mellom endringsskår Terningmønster og BPVS 4 år som utpeker seg som svak. Det gjøres oppmerksom på at denne modellen er basert på en ikke-standardisert løsning, altså er regresjonskoeffisientene i tabell 7 ikke standardiserte.

Tabell 7: LCSM BPVS og Terningmønster, resultater fra hele utvalget

B		
Endringsskår BPVS	Terningmønster 4 år	0.474 (p = .054)
Endringsskår Terningmønster	BPVS 4 år	0.012 (p = .628)

Visuelt presenteres de samme resultatene i figur 7. Figuren er en forenklet fremstilling av LCS-modellen, og inkluderer bare forholdet mellom 1) endringsskår Terningmønster og BPVS 4 år og 2) endringsskår BPVS og Terningmønster 4 år.



Figur 7: LCSM BPVS og Terningmønster, resultater fra hele utvalget

4.2.3 Sammenligning av modeller

Analyseprogrammet RStudio gir ikke standardiserte koeffisienter når man arbeider med LCSM slik det er gjort her. For å sammenligne resultatene fra LCSM og CLPM anvendes derfor de ikke-standardiserte koeffisientene fra begge modellene. En sammenstilling av resultatene vises i tabell 8. Det går frem av tabellen at de to ulike modellene fanget opp det samme, ettersom resultatene var påfallende like.

Tabell 8: Sammenstilling av resultater fra LCSM og CLPM

		LCSM	CLPM
Terningmønster 6 år	BPVS 4 år	0.012 (p = .628)	0.012 (p = .625)
BPVS 6 år	Terningmønster 4 år	0.474 (p = .054)	0.474 (p = .058)

Merknad: For enkelhets skyld brukes «Terningmønster 6 år» og «BPVS 6 år» som variabelnavn, men for LCSM er det mer presise navnet hhv. «Endringskår Terningmønster» og «Endringskår BPVS».

4.3 Er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?

Problemstilling 2 har blitt undersøkt ved at det har blitt brukt CLPM med kun undergruppene «vanskegruppe vokabular» og «vanskegruppe nonverbal resonnering». Resultatene av disse undersøkelsene presenteres nedenfor.

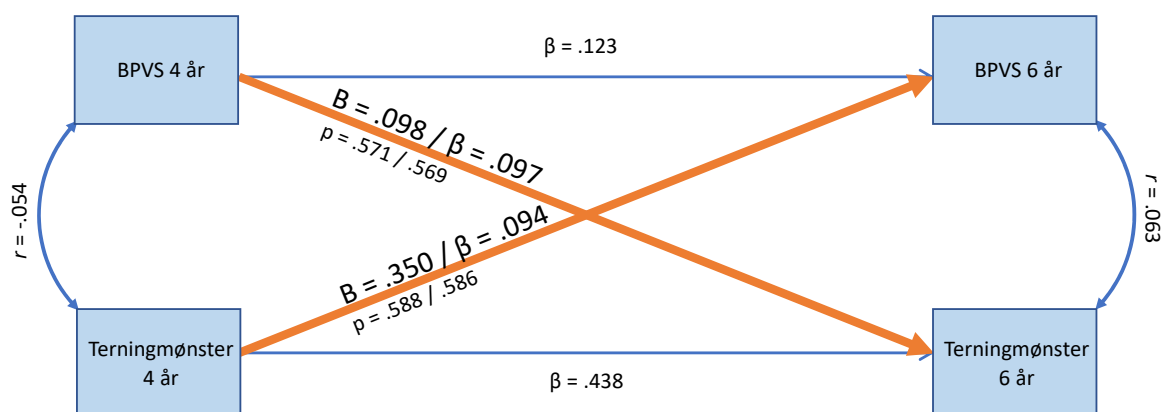
4.3.1 CLPM vanskegruppe vokabular

Ved å kun analysere data fra deltakerne med BPVS-skårer i 15. persentil, kjent her som vanskegruppe vokabular, fremkom resultatene som vises i tabell 9. Med unntak av den standardiserte regresjonskoeffisienten mellom Terningmønster 6 år og Terningmønster 4 år, var det ingen av regresjonskoeffisientene her som oppnådde statistisk signifikans.

Tabell 9: CLPM BPVS og Terningmønster, resultater fra vanskegruppe vokabular

		B	β
BPVS 6 år	Terningmønster 4 år	0.350 (p = .588)	0.094 (p = .586)
	BPVS 4 år	0.303 (p = .517)	0.123 (p = .513)
Terningmønster 6 år	Terningmønster 4 år	0.665 (p = .0052)	0.438 (p = .002)
	BPVS 4 år	0.098 (p = .571)	0.097 (p = .569)

Visuelt fremstilles de samme resultatene i figur 8. Her gjøres det oppmerksom på at korrelasjonskoeffisientene mellom ulike kognitive ferdigheter målt på samme tidspunkt ikke er signifikante. Regresjonskoeffisienten mellom BPVS 6 år og BPVS 4 år er ikke signifikant, mens regresjonskoeffisienten mellom Terningmønster 6 år og Terningmønster 4 år er signifikant.



Figur 8: CLPM BPVS og Terningmønster, resultater fra vanskegruppe vokabular

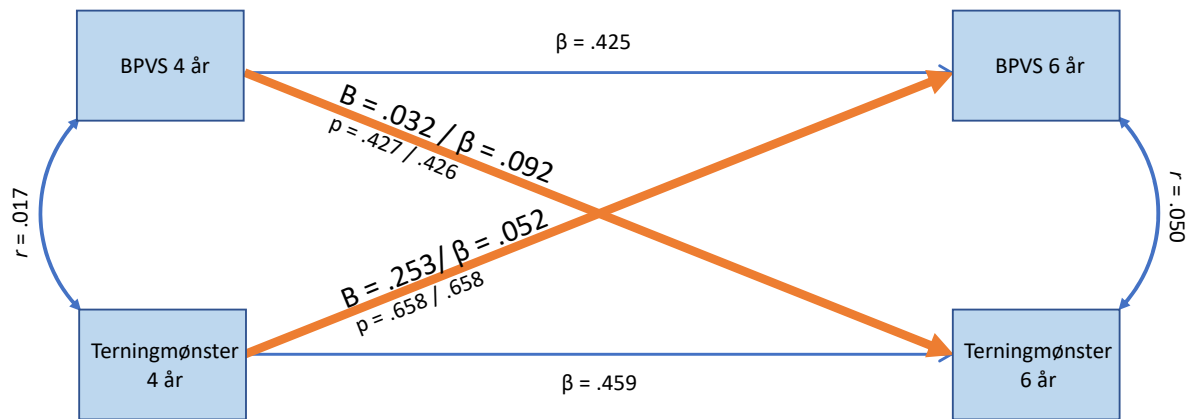
4.3.2 CLPM vanskegruppe nonverbal resonnering

Ved å kun analysere data fra deltakerne med skårer på Terningmønster i 15. persentil, kjent her som vanskegruppe nonverbal resonnering, fremkom resultatene som vises i tabell 10. Her var det statistisk signifikans for alle regresjonskoeffisientene mellom samme variabel målt på forskjellig tidspunkt (samtlige har $p < .001$), mens de øvrige regresjonskoeffisientene ikke var statistisk signifikante.

Tabell 10: CLPM BPVS og Terningmønster, resultater fra vanskegruppe nonverbal resonnering

		B	β
BPVS 6 år	Terningmønster 4 år	0.253 (p = .658)	0.052 (p = .658)
	BPVS 4 år	0.482 (p < .001)	0.425 (p < .001)
Terningmønster 6 år	Terningmønster 4 år	0.686 (p < .001)	0.459 (p < .001)
	BPVS 4 år	0.032 (p = .427)	0.092 (p = .426)

Visuelt fremstilles de samme resultatene i figur 9. Her gjøres det oppmerksom på at korrelasjonskoeffisientene mellom ulike kognitive ferdigheter målt på samme tidspunkt ikke er signifikante. Regresjonskoeffisientene mellom samme ferdighet målt på ulike tidspunkter er begge signifikante.



Figur 9: CLPM BPVS og Terningmønster, resultater fra vanskegruppe nonverbal resonnering

5 Diskusjon

Gjennom undersøkelsene som er gjort i denne masteroppgaven, er det ikke funnet støtte for at gjensidighetsteorien forklarer utviklingen av vokabular og nonverbal resonnering, hverken i analysene av hele utvalget eller i analysene av vanskegruppene. Resultatene indikerer altså at gjensidighetsteorien ikke lar seg replikere på et norsk utvalg, og videre at den ikke er holdbar for de av barna som strever med vokabular eller nonverbal resonnering.

Likevel er det identifisert noen svakheter i analysene som gjør at særlig konklusjonene knyttet til vanskegruppene er noe usikre. Dette og andre begrensninger i studien vil diskuteres nedenfor når det gjøres en vurdering av prosjektets validitet.

Undersøkelsene i denne masteroppgaven kan bidra til vår forståelse av kognitiv og språklig utvikling, også i spesialpedagogiske sammenhenger. Derfor har undersøkelsene både teoretisk og praktisk betydning. Dette vil også bli diskutert nedenfor. Siden funnene i denne masteroppgaven spriker fra studiene som har inspirert mine undersøkelser (Kievit et al., 2019; Kievit et al., 2017), vil diskusjonen dessuten inneholde noen mulige forklaringer på dette. Avslutningsvis inneholder diskusjonen noen forslag til videre forskning.

5.1 Vurdering av validitet

Her brukes som beskrevet typologien til Shadish et al. (2002) for å vurdere validiteten til slutningene som trekkes på bakgrunn av masteroppgavens resultater.

5.1.1 Statistisk validitet

Statistisk validitet handler som nevnt om å vurdere hvorvidt man trekker gyldige slutninger om korrelasjonen mellom variabler. Dette er en vurdering som baserer seg to forhold (Shadish et al., 2002). Det ene er effektstørrelse, altså hvor stor effekt en variabel har på en annen, noe som blant annet bestemmes av styrken på korrelasjonen mellom variablene. Det andre er den statistiske signifikansen til effektstørrelsen.

5.1.1.1 Effektstørrelse

En mulig trussel for statistisk validitet er reliabiliteten til variablene. Lav reliabilitet kan føre til at man feilvurderer og særlig at man undervurderer effektstørrelsen på sammenhenger i undersøkelsen (Ringdal, 2018, s. 367; Shadish et al., 2002). I dette prosjektet var reliabiliteten til BPVS svært god, med $\alpha = 0,911$ ved første måling og $\alpha = 0,913$ ved andre måling.

Reliabiliteten til Terningmønster var derimot lavere, med $\alpha = 0,560$ ved første måling og $\alpha =$

0,703 ved andre måling. Særlig har målingen av Terningmønster ved 4 år noe lav reliabilitet, noe som gjør at en viss forsiktighet bør utvises i tolkningen av resultatene.

De nevnte reliabilitetsmålene gjelder når man ser på hele utvalget under ett. Når man derimot ser kun på vanskegruppene (se tabell 5), er reliabiliteten til nesten samtlige variabler lavere. Variablene for Terningmønster, hovedsakelig ved 4 år og i vanskegruppe nonverbal resonnering, har dessverre en særlig utilfredsstillende reliabilitet. Dette kan ha flere årsaker. Muligens kan det skyldes en gulveffekt, altså at testen var for vanskelig for barna i studien. Det ble likevel valgt å inkludere analysene med disse variablene i masteroppgaven, fordi forskningsspørsmålene er interessante, og fordi det uansett var ønskelig å belyse hvilke resultater disse analysene ledet til. Selvfølgelig må det likevel påpekes at det på grunn av reliabiliteten er usikkerhet knyttet til resultatene som gjelder problemstilling 2.

For å begrense målefeil, kunne analysene i prosjektet med fordel tatt utgangspunkt i latente variabler for vokabular og nonverbal resonnering. Dette ville omfattet å sette sammen ulike tester til «samlevariabler», i stedet for å utelukkende se på mål gjort med BPVS og Terningmønster. Dette var imidlertid også grunnen til at latente variabler ikke var et alternativ i dette prosjektet, fordi datasettet ikke inneholdt flere typer tester av vokabular og nonverbal resonnering som var utført både ved 4 og 6 år. Jeg lykkes heller ikke i å finne et annet datasett som var egnet i denne oppgaven.

Noe som likevel bidrar positivt til prosjektets statistiske validitet, er at det er brukt to ulike analytiske modeller for å undersøke forholdet mellom vokabular og nonverbal resonnering. At analysene av det totale utvalget gjort med både CLPM og LCSM stort sett ga identiske resultater, gjør at det knyttes større troverdighet til korrelasjonsstyrken mellom variablene, men også til p-verdien til disse korrelasjonsmålene.

5.1.1.2 Statistisk signifikans

Statistisk validitet handler også om statistisk signifikans. Dette er begrep knyttet til nullhypotesetesting, som er beskrevet under metodekapittelet. De sentrale nullhypotesene i dette prosjektet er knyttet til analysene gjort med CLPM og LCSM. For problemstilling 1 handler nullhypotesen om at det i disse modellene ikke er korrelasjon mellom 1) Terningmønster 6 år og BPVS 4 år og 2) BPVS 6 år og Terningmønster 4 år. For problemstilling 2 gjelder de samme nullhypotesene, men bare for barn som strever med vokabular og/eller nonverbal resonnering. Siden det som presentert under resultatdelen ikke

ble funnet noen statistisk signifikante forhold mellom disse variablene i CLPM eller LCSM, er det ikke grunnlag for å forkaste de nevnte nullhypotesene i prosjektet.

I nullhypotesetesting finnes det to ulike typer feilslutninger, type I-feil og type II-feil (Shadish et al., 2002). Type I-feil er å forkaste en nullhypotese som i virkeligheten er sann, mens type II-feil vil si å beholde en nullhypotese som i virkeligheten er usann. Siden det her ikke er funnet grunnlag for å forkaste nullhypotesene, vil det altså være snakk om en type II-feil i dette prosjektet hvis nullhypotesene i virkeligheten er usanne.

For å minimere sjansen for en type II-feil, kan man øke den statistiske styrken ved for eksempel å øke utvalgsstørrelsen (Shadish et al., 2002). For problemstilling 1 som gjelder hele utvalget, er utvalgsstørrelsen allerede av en god størrelse (se tabell 2). For problemstilling 2 derimot, som bare gjelder vanskegruppe vokabular og vanskegruppe nonverbal resonnering, er det en betydelig lavere utvalgsstørrelse (se tabell 5). Dette øker sjansen for å gjøre en type II-feil. Likevel, siden det heller ikke fantes signifikante resultater i hele utvalget, er det på ingen måte sikkert at økt utvalgsstørrelse ville ført til signifikante resultater i vanskegruppene. Det er fordi forskjellen mellom vansker og «normalfungering» på disse områdene som nevnt bare er kvantitative, og ikke kvalitative.

Flere peker likevel på at klassisk nullhypotesetesting er forbundet med en for stor grad av «enten eller»-tenkning, ved at en p-verdi på for eksempel 0,49 fører til forkasting av nullhypotesen, mens en p-verdi på 0,51 fører til at nullhypotesen beholdes (Ringdal, 2018, s. 387; Shadish et al., 2002). Å rapportere p-verdi og å diskutere resultatene representerer en mindre rigid fremgangsmåte. P-verdi angir her sjansen for å finne minimum den korrelasjonen man har funnet, dersom det i virkeligheten ikke finnes en sammenheng mellom variablene. Noen av p-verdiene i modellene over er i grenseland mellom «signifikant» og «ikke signifikant» ($p = 0,054-0,058$). Det gjelder forholdet mellom BPVS 6 år og Terningmønster 4 år når hele utvalget inngikk i analysene med CLPM og LCSM. Her kunne man kanskje fått signifikante resultater ved eksempelvis å øke utvalgsstørrelsen.

5.1.2 Indre validitet

Indre validitet handler som beskrevet om hvorvidt korrelasjoner man har funnet innebærer at det er årsakssammenhenger mellom variablene i utvalget (Shadish et al., 2002). Indre validitet bygger på statistisk validitet, og siden det ikke ble funnet statistisk signifikante forhold mellom BPVS og Terningmønster i CLPM eller LCSM, er det i dette prosjektet ikke grunnlag

for å si at det er en kausal sammenheng mellom variablene. Nedenfor utforskes det nærmere hvor skikket denne observasjonsstudien er for å trekke denne slutningen om gjensidige kausale sammenhenger mellom vokabular og nonverbal resonnering i utvalget.

Som tidligere nevnt er eksperimenter det sikreste studiedesignet for å uttale seg om kausale sammenhenger mellom variabler. Det er fordi man kan kontrollere for noen viktige forhold. Dette gjelder særlig tidsaspektet, altså at årsak må inntreffe før virkning (som kontrolleres for i eksperimenter ved at årsaken innføres i forbindelse med eksperimentet), og utelukkning av en tredjevariabel (som kontrolleres for i eksperimenter ved at man randomiserer utvalget og deler inn i en tiltaksgruppe og en kontrollgruppe). Selv om det er svært utfordrende i observasjonelle studier å kontrollere for dette, særlig for mulige tredjevariabler, finnes det muligheter som gjør at man kan komme nærmere å trekke gyldige kausale slutninger (Rohrer, 2018).

I dette masterprosjektet er tidsaspektet og altså retning kontrollert for fordi studien er longitudinell, og data er innsamlet på flere tidspunkter fremover i tid. Gjennom de statistiske analysemodellene som er brukt, CLPM og LCSM, er det til en viss grad også kontrollert for tredjevariabler som kan påvirke utviklingen til vokabular og/eller nonverbal resonnering. Det er fordi analysemodellene kontrollerer for variablenes stabilitet over tid, og for at variablene som er målt samtidig kan korrelere med hverandre (Kearney, 2017). Likevel kan det med disse analysemodellene ikke fullstendig utelukkes at det finnes ukjente tredjevariabler som ligger til grunn og som både påvirker utviklingen av vokabular og nonverbal resonnering. Akkurat i studier av gjensidighetsteorien er dette litt uheldig fordi den ukjente tredjevariabelen i teorien kan være en form for mentalkapasitet eller *g*-faktor, som jo gjensidighetsteorien tar avstand fra.

Når man ser på tabell 3 ovenfor, som angir rene mål av Pearsons *r* for variablene i prosjektet, er det likevel tydelig at det er viktig å bruke modeller som CLPM og LCSM når man skal studere kausale sammenhenger basert på observasjonsdata. Det er fordi det, som det fremgår av tabell 3, er statistisk signifikante korrelasjonsmål mellom samtlige variabler i prosjektet, inkludert mellom Terningmønster 6 år og BPVS 4 år, samt BPVS 6 år og Terningmønster 4 år. Dette kunne blitt feiltolket som gjensidige kausale sammenhenger mellom variablene, særlig fordi dette både har teoristøtte og empirisk støtte (Kievit et al., 2019; Kievit et al., 2017; van der Maas et al., 2006). Når de samme variablene plasseres inn i CLPM og LCSM, som er mer presise fordi de altså kontrollerer for noen viktige forhold, finner man imidlertid

ikke holdepunkter for at det er en årsakssammenheng mellom variablene. Et sentralt poeng er også at begge modellene, altså både CLPM og LCSM, ga de samme resultatene, noe som styrker konklusjonen om at det ikke er et gjensidig kausalt forhold mellom vokabular og nonverbal resonnering i dette utvalget.

I denne masteroppgaven er formålet ikke bare å se om det er en årsakssammenheng mellom vokabular og nonverbal resonnering, men også å undersøke om det er et gjensidig årsaksforhold mellom dem. Dersom undersøkelsen hadde gitt signifikante resultater der nonverbal resonnering påvirket vokabular, som på sin side påvirket nonverbal resonnering, ville det vært grunnlag for å si at det var gjensidige forhold. Dette var ikke tilfellet. Dersom undersøkelsen hadde gitt signifikante resultater der enten vokabular eller nonverbal resonnering påvirket den andre, men ikke motsatt, kunne man konkludert med at det var et ikke-gjensidig årsaksforhold. Dette var heller ikke tilfellet i denne undersøkelsen, selv om forholdet mellom Terningmønster 4 år og BPVS 6 år som nevnt var sterkere enn forholdet mellom BPVS 4 år og Terningmønster 6 år når data fra hele utvalget ble analysert.

5.1.3 Begrepsvaliditet

Å vurdere begrepsvaliditet handler om å avgjøre om man har målt de begrepene man faktisk mente å måle, slik at man trekker slutninger om de riktige begrepene (Nyeng, 2012). Generelt påvirker tilfeldige målefeil (lav reliabilitet) begrepsvaliditeten i negativ retning. Under kapittel 5.1.1 er det beskrevet at enkelte av reliabilitetsmålene i denne undersøkelsen var lavere enn ønskelig, og der dette er tilfellet, truer det også begrepsvaliditeten til undersøkelsen (Nyeng, 2012). I tillegg vil systematiske målefeil være en mulig trussel for begrepsvaliditeten. Det betyr at begrepsvaliditet også er avhengig av at måleverktøyene man bruker ikke systematisk fanger opp forhold som ligger utenfor de tilsiktede begrepene (Nyeng, 2012; Ringdal, 2018, s. 103). Videre er god begrepsvaliditet også kjennetegnet av at man fanger opp hele bredden av begrepene man tar for seg (Ringdal, 2018, s. 105).

Nedenfor blir det sett på hvorvidt begrepene vokabular og nonverbal resonnering er operasjonalisert på en god måte i dette prosjektet ved hjelp av måleverktøyene BPVS og Terningmønster. I tillegg er det relevant å vurdere hvorvidt vanskegruppene er definert på en tilfredsstillende måte.

5.1.3.1 Vokabular

Når det gjelder BPVS, som er brukt for å måle vokabularferdighetene til barna i denne undersøkelsen, tyder tidligere publikasjoner som allerede nevnt på at begrepsvaliditeten er god (Lyster et al., 2010; Miller & Lee, 1993). Dette gir grunn til å tro at testing med BPVS i liten grad inkluderer andre begreper enn nettopp vokabular. Samtidig er BPVS et verktøy som primært måler breddevokabular og impressivt vokabular (Lyster et al., 2010). Man kan derfor stille spørsmål om hvorvidt det ville gitt andre resultater dersom man brukte et måleverktøy som i større grad stilte krav til dybdevokabular og ekspressivt vokabular. Likevel viser tidligere undersøkelser at det er en høy korrelasjon mellom dybdevokabular og breddevokabular (Vermeer, 2001). Også impressivt og ekspressivt språk er nært relatert og korrelerer med hverandre (Schmitt, 2019; van Viersen et al., 2017). Det er derfor en rimelig grunn til å anta at begrepet vokabular er tilstrekkelig operasjonalisert og at måleverktøyet BPVS fanger opp bredden i begrepet.

5.1.3.2 Nonverbal resonnering

Hovedankepunktet for begrepsvaliditeten til nonverbal resonnering i denne undersøkelsen er antakeligvis reliabiliteten som allerede er drøftet tidligere i dette kapittelet. Likevel er Terningmønster et anerkjent mål på nonverbal resonnering som ikke stiller spesielle krav til tidligere kunnskap eller erfaring, og som dessuten dekker både den induktive og den deduktive delen av nonverbal resonnering (Peng et al., 2019). Det er imidlertid ikke nødvendigvis et krav at en test av nonverbal resonnering både er induktiv og deduktiv, fordi nyere forskning peker i retning av at begge delene baserer seg på de samme kognitive prosessene (Stephens et al., 2018). Peng et al. (2019) har på den andre siden hevdet at Terningmønster ikke er et helt rent mål av nonverbal resonnering, fordi visuospatiale evner også vil påvirke testresultatet. I tillegg ble det gitt verbale instruksjoner i forbindelse med Terningmønster i studien, slik at det også finnes en viss mulighet for at verbale evner kan virke inn i testsituasjonen. Likevel vil operasjonaliseringer av begreper som eksempelvis nonverbal resonnering aldri bli helt feilfrie (Nyeng, 2012), og i alt er det i dette prosjektet vurdert at Terningmønster er et tilfredsstillende måleverktøy. Det er blant annet fordi Terningmønster er en test som også tidligere har blitt brukt for å måle nonverbal resonnering i undersøkelser av gjensidighetsteorien (Griffiths et al., 2022).

5.1.3.3 Vanskegrupper

Foruten begrepsvaliditeten til nonverbal resonnering og vokabular, er det også interessant å se på hvorvidt vanskegruppene i masterprosjektet er definert på et rimelig grunnlag. Barn med

kjente språkforstyrrelser, lærevansker eller utviklingshemming ble som kjent utelukket fra den opprinnelige studien (Hjetland et al., 2019). I denne masteroppgaven inngår derfor ikke de barna i populasjonen som muligens skårer lavest på vokabular og nonverbal resonnering. Dette kan ses som en trussel mot begrepsvaliditeten når det kommer til problemstilling nummer 2. Samtidig er det antakeligvis rimelig å anta at de barna som er inkludert i vanskegruppene her, faktisk har en del utfordringer knyttet til vokabular og nonverbal resonnering. Dette fordi kun 15 % av de andre deltakerne hadde samme eller lavere skår. Igjen må det dessuten også påpekes at det ikke er noe som tyder på at barn med språkforstyrrelser eller nonverbale lærevansker har kvalitativt forskjellige utfordringer med vokabular og nonverbal resonnering sammenlignet med andre barn, og det er en uklar grense mellom «vansker» og «ikke-vansker». Videre er det ikke på forhånd gitt at barn med diagnostiserte lærevansker nødvendigvis ville fått blant de laveste skårene på de testene som er brukt her (Calder et al., 2023).

5.1.4 Ytre validitet

Ytre validitet handler om hvorvidt man kan generalisere konklusjoner fra studien, slik at de også er gyldige for andre enn deltakerne i utvalget (Shadish et al., 2002). Siden det i masterprosjektet ikke ble funnet støtte for gjensidighetsteorien, peker det i retning av at denne teorien ikke forklarer positive manifold eller kognitiv utvikling hos norske barn og unge. I masteroppgaven er det imidlertid brukt et utvalg som skal representere norske barn, selv om deltakerne ikke er tilfeldig utvalgt fra hele populasjonen, men tilhører samme kommune (Hjetland et al., 2019). Som allerede gjort rede for under metodekapittelet, er det likevel vurdert at resultatene i dette prosjektet antakeligvis er generaliserbare til norske barn for øvrig. Dette skyldes at variablene som er brukt i masterprosjektet handler om kognitive ferdigheter, og det antas her at variasjoner i slike ferdigheter vil finnes overalt i landet. I tillegg er sosioøkonomisk status i hjemkommunen til deltakerne i undersøkelsen antatt å være representativ for øvrige norske barn også.

5.2 Teoretiske implikasjoner

I motsetning til tidligere studier (Griffiths et al., 2022; Kievit et al., 2019; Kievit et al., 2017), ble det i dette norske utvalget ikke funnet noe som underbygger at det er et gjensidig forhold mellom vokabular og nonverbal resonnering. Det indikerer at det i denne masteroppgaven ikke er funnet støtte for gjensidighetsteorien, altså for hypotesen om at ulike kognitive ferdigheter støtter hverandre sin utvikling. Disse funnene gjelder både for det totale utvalget i

masteroppgaven, og for barna i utvalget som strever med vokabular og/eller nonverbal resonnering.

I analysene som kun angikk vanskegruppene var det likevel noen utfordringer knyttet til utvalgsstørrelsen og lave reliabilitetsmål. Derfor skal man være litt forsiktig med å konkludere med at det ikke er et gjensidig forhold mellom vokabular og nonverbal resonnering i disse gruppene. Samtidig er det i utgangspunktet ingen åpenbar grunn til at forholdet mellom vokabular og nonverbal resonnering skal være annerledes i vanskegruppene enn i totalutvalget, fordi det ikke finnes et kvalitativt skille mellom barn med og uten vansker (Bishop et al., 2017; Peters & Ansari, 2019). I en tidligere studie har man også sett at forholdet mellom vokabular og nonverbal resonnering var sammenlignbart hos barn med og uten språkforstyrrelser (Griffiths et al., 2022). I motsetning til dette masterprosjektet, fant den nevnte studien imidlertid støtte for gjensidighetsteorien. Der forskning tidligere har vist at forholdet mellom ulike kognitive ferdigheter kan være ulikt hos barn med og uten lærevansker, innebærer dette at gjensidige forhold finnes mellom ferdighetene i «normalgrupper», men ikke hos barn med dysleksi (Ferrer et al., 2010). Det gir med andre ord heller ikke grunn til å anta at gjensidighetsteorien er holdbar for barna i vanskegruppene i denne masteroppgaven, fordi den uansett ikke var holdbar i totalutvalget.

At funnene i masteroppgaven ikke er i tråd med gjensidighetsteorien, åpner for at andre teorier om kognitiv utvikling og positive manifold bør vurderes. Det gjør også studier som tidligere har pekt på at andre teorier er mer plausible enn gjensidighetsteorien. Her kan det blant annet igjen nevnes studier som støtter *g*-faktor-teorien (Gignac, 2014, 2016), investeringsteorien (Kvist & Gustafsson, 2008) og samplingteorien (Bartholomew et al., 2013; Bartholomew et al., 2009).

Siden dette masterprosjektet handlet spesifikt om å se hvorvidt gjensidighetsteorien lot seg replikere, har disse andre teoriene naturligvis vært mindre i fokus i analysene. Selv om den ikke var statistisk signifikant, var det likevel en tendens i mine analyser at relasjonen mellom vokabular 6 år og nonverbal resonnering 4 år var sterkere enn den motsatte relasjonen (nonverbal resonnering 6 år og vokabular 4 år). Det er interessant at dette funnet er mer i tråd med investeringsteorien enn gjensidighetsteorien. Som vi tidligere har sett (se kapittel 2, bl.a. figur 2) er nemlig et kjennetegn ved investeringsteorien at flytende intelligens (nonverbal resonnering) innvirker på utviklingen av krystallisert intelligens (for eksempel vokabular), men at forholdet altså ikke er gjensidig. Selv om det ble observert tegn til at utviklingen av

vokabular var mer avhengig av nonverbal resonnering enn motsatt, var funnene likevel ikke tilstrekkelige til å konkludere med at undersøkelsene var fullstendig i samsvar med investeringsteorien heller. Analyser utført i denne masteroppgaven gir videre ingen grunnlag hverken for å styrke eller utfordre *g*-faktor-teorien eller samplingteorien.

Det har ikke vært en del av dette masterprosjektet å undersøke hvorvidt det er mer hensiktsmessig å forklare kognitiv utvikling eller positiv manifold ved å integrere elementer fra flere ulike teorier. Dette er imidlertid en mulighet som er pekt på i flere tidligere publikasjoner. For eksempel foreslår Peng et al. (2019) at miljøfaktorer som skolegang og sosioøkonomisk status, sammen med elementer fra både gjensidighetsteorien og investeringsteorien, kan gi et mer fullstendig bilde av barns kognitive utvikling. I sin studie diskuterer også Gignac (2014) at det er mulig at *g*-faktor-teorien og gjensidighetsteorien kan sameksistere, dersom det finnes en underliggende medfødt mentalkapasitet (*g*-faktor) som innledningsvis er bestemmende for kognitiv utvikling innen ulike domener, samtidig som ulike kognitive ferdigheter etter hvert som de blir etablert også påvirker hverandres videre utvikling. Dessuten påpeker også Kievit og kolleger (Kievit et al., 2017) at deres funn støtter at det er gjensidige forhold mellom ulike kognitive ferdigheter, men at denne gjensidigheten kan være noe som forsterker preeksisterende individuelle forskjeller i kognitivfunksjon. Samtidig finnes det også andre publikasjoner som har forsøkt å danne nye modeller for kognitiv utvikling ved å inkorporere sentrale elementer fra flere allerede etablerte teorier (se f.eks. Gustafsson, 1984). At kognitiv utvikling kan forklares av en form for ny eller «blandet» teori kan selvsagt ikke avvises i dette masterprosjektet, selv om det er et tema som ikke har blitt nærmere utforsket her.

I tillegg til implikasjoner for hvordan man kan forstå kognitiv utvikling, kan studier av gjensidighetsteorien også ha betydning for hvordan man skal forstå ulike lærevansker. Ferrer et al. (2010) konkluderer med at det ikke er et gjensidig forhold mellom leseferdigheter og intelligens hos barn med dysleksi, selv om dette gjensidige forholdet eksisterer hos barn uten dysleksi. Dette er i tråd med en diskrepansdefinisjon av dysleksi, noe som går ut på at lesevanskene er uventede sett i lys av øvrig kognitiv fungering. Denne typen definisjon har man gått bort fra i senere tid (Snowling, Hulme, et al., 2020). Særlig gjelder dette i den nye definisjonen av språkforstyrrelser, hvor termen «utviklingsmessige språkforstyrrelser» nå kan brukes selv om andre kognitive evner ligger under gjennomsnittsområdet (CATALISE-studien, Bishop et al., 2017). Litt ulikt Ferrer et al. (2010), finner Griffiths et al. (2022) støtte for gjensidighetsteorien blant barn både med og uten språkforstyrrelser. Derfor antar de at

språkferdighetene til barn med språkforstyrrelser har en negativ påvirkning på utviklingen av andre kognitive ferdigheter. De påpeker at dette funnet underbygger den nye konsensusen om språkforstyrrelser. Dessuten kan det bidra til å styrke antakelsen om at det ikke er kvalitative forskjeller mellom barn med og uten språkforstyrrelser.

I denne masteroppgaven er det ikke funnet resultater tilsvarende studien til Griffiths et al. (2022), fordi det ikke er funnet holdepunkter for gjensidighetsteorien hverken hos barn med lærevansker eller hos barna i totalutvalget. Resultater som ikke støtter gjensidighetsteorien for barn med språkforstyrrelser, er likevel ikke nødvendigvis et argument mot konsensusen i CATALISE. Det skyldes at det uansett, gjennom fenomenet positive manifold, er anerkjent at det vil være en korrelasjon mellom språkferdigheter og nonverbal resonnering på individnivå. Da er det også å forvente at mange barn som strever med språket, vil ha tilleggsutfordringer med nonverbal resonnering. Man kan argumentere for at dette vil gjelde uavhengig av hva som er årsaksmekanismene bak positive manifold. En mulighet er likevel at positive manifold ikke gjør seg gjeldende for barn med språkforstyrrelser. Dette spørsmålet faller likevel utenfor formålet med denne masteroppgaven.

5.3 Praktiske implikasjoner

Dersom man finner gjensidige forhold mellom vokabular og nonverbal resonnering, kan dette indikere at ferdigheter i nonverbal resonnering profiterer på språktiltak (Griffiths et al. 2022), og motsatt at tiltak som går på nonverbal resonnering kan forbedre barns språkferdigheter. Når det i denne masteroppgaven ikke er funnet støtte for at det er gjensidige forhold mellom vokabular og nonverbal resonnering, peker dette derimot i retning av at språkforstyrrelser bør avhjelpest med språktiltak, mens vansker med nonverbal resonnering bør avhjelpest med tiltak som går på nonverbal resonnering. Med andre ord peker funnene i masteroppgaven på at spesialpedagogiske tiltak bør være målrettede og ta for seg det aktuelle vanskeområdet, og at man ikke uten videre kan forvente en overføringsverdi ved å trene på andre typer ferdigheter.

Disse implikasjonene gjelder særlig for vansker knyttet til vokabular og nonverbal resonnering som det har vært fokus på her. Men siden gjensidighetsteorien også omfatter andre kognitive ferdigheter (van der Maas et al., 2006), er det en viss grunn til å anta at det samme kan gjelde for andre typer lærevansker. Det vil si at de praktiske implikasjonene kan ha betydning for barn med for eksempel språkforstyrrelser, utviklingshemming eller generelle lærevansker, men også for barn med andre spesialpedagogiske behov, blant annet dysleksi eller matematikkvansker.

Positive manifold forteller oss likevel at barn som strever med én kognitiv ferdighet, har økt risiko for også å streve med andre typer kognitive ferdigheter. Når spesialpedagogiske tiltak i disse tilfellene er indikert, tyder resultatene fra denne masteroppgaven på at man bør tilby intervensjoner som har dokumentert effekt mot hver av de ulike vansketypene. Med andre ord er det ikke tilstrekkelig å innføre tiltak innenfor én kognitiv ferdighet i håp om å se forbedringer også innenfor andre kognitive domener. Det kunne derimot vært en mer rimelig strategi dersom man hadde funnet støtte for gjensidighetsteorien.

I datasettet som er brukt i denne masteroppgaven, er det som nevnt en tendens som kan peke i retning av investeringsteorien, selv om denne ikke oppnår statistisk signifikans. Hvis senere undersøkelser finner sterkere støtte for investeringsteorien, kan dette være et argument for at språkforstyrrelser kan avhjelpes via oppgaver som går på nonverbal resonnering. Dette fordi nonverbal resonnering innenfor investeringsteorien skaper grunnlaget for utviklingen av blant annet språk- og vokabularferdigheter (Kievit et al., 2017). På den andre siden viser tidligere publikasjoner at øving innenfor generelle kognitive ferdigheter ikke kan anbefales som språktiltak (Hagen et al., 2014; Strong et al., 2011).

Som gjort rede for i kapittel 2.3.2, er det dessuten også en viss usikkerhet knyttet til om nonverbal resonnering er trenbart. Dette betyr at det ikke nødvendigvis er hensiktsmessig å trene nonverbal resonnering med tanke på forbedring av språkferdigheter eller vokabular, selv om det skulle vise seg å finnes støtte for eksempelvis gjensidighetsteorien eller investeringsteorien. Det betyr også at det er knyttet tvil til effekten av trening på nonverbal resonnering og om det bidrar til forbedring av disse ferdighetene i seg selv. Derfor ville det være interessant å utforske om trening i andre ferdigheter, som for eksempel vokabular, kan være nyttig for å forbedre nonverbal resonnering. Undersøkelsen i denne masteroppgaven indikerer imidlertid at dette antakeligvis ikke er hensiktsmessig, ettersom gjensidighetsteorien ikke lot seg replikere.

At trening i én spesifikk kognitiv ferdighet ikke nødvendigvis har en direkte effekt på andre kognitive ferdigheter, er velkjent i nyere forskning, spesielt innenfor studier av arbeidsminnet. Lenge var det antatt at fonologisk arbeidsminne påvirket og understøttet utviklingen av vokabularet (Gathercole et al., 1992), og derfor at arbeidsminnetrening kunne være hensiktsmessig for ordforrådet. Nyere funn viser imidlertid at dette trolig ikke er tilfellet (Melby-Lervåg et al., 2012). Funn fra metaanalyser indikerer dessuten at trening av arbeidsminne ikke er effektivt for å forbedre andre kognitive ferdigheter som for eksempel

språk eller nonverbal resonnering (Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016). Disse studiene kan antyde at spesialpedagogiske intervensjoner bør rettes konkret mot det som er den aktuelle vansken, fordi man har sett at arbeidsminnetrening trolig ikke gir effekt på andre typer ferdigheter. Dette stemmer overens med det som fremkommer i denne masteroppgaven, fordi funnene her også peker i retning av at spesialpedagogiske tiltak bør ta for seg det konkrete vanskeområdet, og ikke andre typer ferdigheter i håp om at det vil gi en overføringsverdi til det som er vanskeområdet.

5.4 Mulige forklaringer på at funnene til Kievit og kolleger ikke har latt seg replikere

Funnene i denne masteroppgaven skiller seg med andre ord fra de tidligere studiene av gjensidighetsteorien utført av Kievit og kolleger (Kievit et al., 2019; Kievit et al., 2017). Det er viktig å drøfte mulige årsaker til disse resultatforskjellene, siden Kievit og kollegers undersøkelser har dannet grunnlaget for denne masteroppgaven.

En mulig forklaring kan knytte seg til begrensninger ved henholdsvis masterprosjektet eller Kievit og kollegers studier (Kievit et al., 2019; Kievit et al., 2017). Blant annet kan de lave reliabilitetsmålene i dette masterprosjektet fordekke sammenhenger som kunne støttet gjensidighetsteorien. Samtidig kan det også være begrensninger ved studiene til Kievit og kolleger som har gjort at det gjensidige forholdet mellom vokabular og nonverbal resonnering hos dem har blitt overvurdert. Det kan ikke utelukkes at analysene i Kievit-undersøkelsene ikke kontrollerer for målefeil godt nok, blant annet fordi det er brukt bare ett måleverktøy for hver av de kognitive ferdighetene som inngår i studiene, i stedet for at det har blitt brukt latente variabler. Det må samtidig påpekes at det også i denne masteroppgaven kun er brukt ett måleverktøy per kognitive ferdighet som er målt.

En annen forklaring på resultatforskjellene kan knytte seg til forskjellene mellom masterprosjektet og Kievit-studiene (Kievit et al., 2019; Kievit et al., 2017). For det første brukte Kievit-undersøkelsene og dette masterprosjektet ulike måleverktøy for både vokabular og nonverbal resonnering (henholdsvis Ordforståelse og Matriser, og BPVS og Terningmønster). Imidlertid fant Kievit og kolleger støtte for gjensidighetsteorien i en senere studie der de brukte lignende tester som i denne masteroppgaven (Griffiths et al., 2022). For det andre er det en forskjell i alder hos deltakerne i masterprosjektet og undersøkelsene til Kievit og kolleger. Det yngste utvalget til Kievit et al. (2019) var 6 til 8 år, mens barna i denne undersøkelsen var 4–6 år. Selv om dette kan forklare at funnene våre er forskjellige,

hevder Kievit og kolleger både i 2017 og 2019 at gjensidige forhold mellom vokabular og nonverbal resonnering trolig er sterkere i perioder som er preget av rask kognitiv utvikling. Dersom denne hypotesen stemmer burde den også gjelde for barn i alderen 4 til 6 år. For det tredje er det ulike språk i de ulike undersøkelsene, med et norskspråklig utvalg i masterprosjektet og engelskspråklige utvalg i Kievit og kolleger sine studier. Språkforskjellen i utvalgene gir imidlertid ikke grunn til å forvente de observerte resultatforskjellene. Dette fordi det her er snakk om en så grunnleggende og basal ferdighet som vokabular.

5.5 Videre forskning

Siden resultatene angående gjensidighetsteorien er sprikende, er det behov for videre forskning på dette temaet. Det er viktig at fremtidig forskning søker å få mer klarhet omkring hvorvidt gjensidighetsteorien er gyldig, ikke minst fordi vår forståelse av kognitiv utvikling både har teoretiske og praktiske implikasjoner, spesielt i spesialpedagogisk sammenheng.

Tidligere i masteroppgaven har det blitt pekt på at eksperimentelle studier er den eneste sikre metoden hvis man ønsker å konkludere om kausale sammenhenger mellom variabler. Både denne studien og tidligere studier av gjensidighetsteorien er imidlertid observasjonsstudier. Derfor er det viktig å utføre eksperimentelle studier av gjensidighetsteorien før man eventuelt konkluderer med noen praktiske implikasjoner som man iverksetter i arbeidet med barn og unge, enten i spesialpedagogiske eller andre sammenhenger.

Eksperimentelle studier av gjensidighetsteorien ville innebære at man for eksempel gir intervensjoner rettet mot vokabular for å observere hvorvidt vokabulartrening påvirker nonverbal resonneringsevne, og omvendt, at man gir intervensjoner rettet mot nonverbal resonnering for å observere hvorvidt nonverbal resonneringsevne påvirker vokabular. Det er likevel noe usikkert om akkurat dette er hensiktsmessig i studier av gjensidighetsteorien. Det skyldes at det fremdeles er uvisst om nonverbal resonnering i det hele tatt er en ferdighet man kan trene på (se kapittel 2.3.2 over for en redegjørelse). Likevel er det en mulighet å studere gjensidighetsteorien med utgangspunkt i andre kognitive ferdigheter som man er mer sikre på er trenbare. Eksempelvis kunne en eksperimentell studie av gjensidighetsteorien tatt utgangspunkt i ferdigheter som vokabular, lesing og/eller matematikk. På nåværende tidspunkt er jeg ikke kjent med at dette har blitt gjort tidligere.

Det er også et behov for ytterligere forskning på gjensidighetsteorien der man fokuserer spesifikt på barn med spesialpedagogiske utfordringer. Det er flere grunner til dette. For det

første har tidligere studier med spesialpedagogisk fokus sprikende resultater angående hvorvidt gjensidighetsteorien er holdbar for barn med ulike typer lærevansker (Ferrer et al., 2010; Griffiths et al., 2022). I tillegg kan studier av gjensidighetsteorien legge noen viktige føringer for hvordan man skal utforme tiltak for barn og unge med ulike typer lærevansker. Derfor kan man også argumentere for at å ha fokus på barn med spesialpedagogiske behov er særlig viktig i videre studier av gjensidighetsteorien.

Det er imidlertid spesielt utfordrende å studere gjensidighetsteorien hos barn og unge med spesialpedagogiske vansker. Dette kommer særlig av at utvalgsstørrelsen har en tendens til å bli lav når man studerer slike grupper. Dette er et kjent problem innen spesialpedagogisk forskning, og skyldes gjerne at det er vanskelig å rekruttere et tilstrekkelig antall barn med spesialpedagogiske behov til forskning, ikke minst fordi deltakelsen ofte er tidkrevende og krever gjentatt oppfølging og testing (Toffalini et al., 2021). En løsning på dette problemet, som også er brukt i dette masterprosjektet, er at man definerer vanskegrupper ut fra hvem som har lave skårer innen ulike ferdigheter i et større datasett. Likevel førte også denne løsningen i dette masterprosjektet til at utvalgsstørrelsen i vanskegruppene ble noe lav. Av den grunn kan man påpeke at fremtidige studier av gjensidighetsteorien hos barn med lærevansker bør ha som mål å rekruttere et utvalg som er av en viss størrelse. Dette vil gi undersøkelsen bedre statistisk styrke.

Når man definerer vanskegrupper som de av deltakerne i et større utvalg som ligger under en bestemt cut-off på gitte variabler, er det også slik at disse deltakerne ikke nødvendigvis har reelle diagnostiserte lærevansker. Dette gjelder særlig i dette prosjektet fordi en av eksklusjonskriteriene i datasettets opprinnelige undersøkelse var at barna ikke hadde slike diagnostiserte vansker. Dette er i utgangspunktet ikke nødvendigvis veldig problematisk, fordi lærevansker som for eksempel generelle lærevansker eller språkforstyrrelser beskriver personer som skårer i de lavere områdene i en normalfordeling av gitte ferdigheter. Likevel er det nok fordelaktig om fremtidig forskning på gjensidighetsteorien innenfor spesialpedagogikk ser spesifikt på barn med kjente eller diagnostiserbare utfordringer. Dette er særlig fordi mye fortsatt er ukjent når det kommer til gjensidighetsteorien, slik at man ikke uten videre kan anta at barn med og uten lærevansker er sammenlignbare på dette området.

I dette masterprosjektet, samt hos Kievit et al. (2017), er gjensidige forhold mellom vokabular og nonverbal resonnering studert med utgangspunkt i to måletidspunkter. I andre studier er det brukt tre ulike måletidspunkter (Griffiths et al., 2022; Kievit et al., 2019). Et mål for

fremtidig forskning bør muligens likevel være at man bruker data som er samlet inn på enda flere tidspunkter, eksempelvis ved fire ulike måletidspunkter. Dette vil gjøre at man i større grad kan studere hvordan de ulike kognitive ferdighetene utvikler seg og forholder seg til hverandre over tid. Undersøkelser med flere enn to måletidspunkter kan også bruke mer avanserte varianter av SEM som analysemetode, eksempelvis *Random Intercept Cross-Lagged Panel Models*, som gjør at man kan kontrollere for stabile faktorer hos deltakerne, eksempelvis genetikk eller sosial bakgrunn (Hamaker et al., 2015). Dette vil kunne gi et klarere svar angående gjensidighetsteorien i fremtiden.

5.6 Oppsummering

I dette masterprosjektet har gjensidighetsteorien blitt testet gjennom følgende problemstillinger:

1. *I hvilken grad lar gjensidighetsteorien seg replikere på et norsk utvalg?*
2. *I hvilken grad er gjensidighetsteorien holdbar for barn som strever med vokabular og/eller nonverbal resonnering?*

Resultatene av undersøkelsene har vist at gjensidighetsteorien ikke har latt seg replikere på dette norske utvalget, hverken i totalutvalget eller i de nevnte vanskegruppene. Likevel, i diskusjonen av validiteten til prosjektet, har det kommet frem at resultatene må tolkes med noe forsiktighet, særlig fordi noen av reliabilitetsmålene i prosjektet er lave. Dette gjelder spesielt i analysene av vanskegruppene, og i størst grad i vanskegruppen for nonverbal resonnering.

I tillegg må man være forsiktig med å konkludere med hvorvidt gjensidighetsteorien er holdbar eller ikke fordi studiene som er gjort på teorien til nå er basert på observasjonsdata. Derfor er det i fremtidig forskning først og fremst viktig at man utfører eksperimentelle studier hvor man tester teorien. I den overnevnte diskusjonen er det også pekt på at videre forskning bør basere seg på testing som er gjort på flere måletidspunkter, eksempelvis fire ulike måletidspunkter, og at man bruker mer avanserte analysemetoder hvor man kan kontrollere for flere variabler. Samtidig er det pekt på at fremtidig forskning på gjensidighetsteorien med fordel kan ha et spesialpedagogisk fokus. Dette fordi de praktiske implikasjonene av studier av gjensidighetsteorien kan ha stor betydning for barn og unge med spesialpedagogiske behov.

Selv om det altså gjenstår mer forskning før man med sikkerhet kan konkludere med hvorvidt gjensidighetsteorien er holdbar eller ikke, har dette masterprosjektet noen sentrale teoretiske og praktiske implikasjoner. De teoretiske implikasjonene av denne undersøkelsen er at gjensidighetsteorien trolig ikke er tilstrekkelig for å forklare kognitiv utvikling hos barn og unge. Det åpner for at man heller bør vurdere andre typer forklaringsmodeller. Av dette følger også praktiske implikasjoner som går ut på at spesialpedagogiske tiltak bør rettes mot det som er den aktuelle vansken, og at man ikke uten videre kan forvente en overføringsverdi ved å trene på andre typer ferdigheter.

Referanseliste

- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuohl, M. & Jaeggi, S. M. (2014). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 22(2), 366–377. <https://doi.org/10.3758/s13423-014-0699-x>
- Bartholomew, D. J., Allerhand, M. & Deary, I. J. (2013). Measuring mental capacity: Thomson's Bonds model and Spearman's g-model compared. *Intelligence*, 41(4), 222-233. <https://doi.org/10.1016/j.intell.2013.03.007>
- Bartholomew, D. J., Deary, I. J. & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review*, 116(3), 567–579. <https://doi.org/10.1037/a0016262>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T. & CATALISE consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PloS one*, 11(7). <https://doi.org/10.1371/journal.pone.0158753>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T. & the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *The Journal of Child Psychology and Psychiatry*, 58(10), 1068-1080. <https://doi.org/10.1111/jcpp.12721>
- Bлага, O. M., Shaddy, D. J., Anderson, C. J., Kannass, K. N., Little, T. D. & Colombo, J. (2009). Structure and continuity of intellectual development in early childhood. *Intelligence*, 37(1), 106–113. <https://doi.org/10.1016/j.intell.2008.09.003>
- Bornstein, M. H., Hahn, C.-S., Putnick, D. L. & Suwalsky, J. T. D. (2014). Stability of core language skill from early childhood to adolescence: A latent variable approach. *Child Development*, 85(4), 1346-1356. <https://doi.org/10.1111/cdev.12192>
- Bors, D. A. & Vigneau, F. (2003). The effect of practice on Raven's Advanced Progressive Matrices. *Learning and Individual Differences*, 13(4), 291-312. [https://doi.org/10.1016/S1041-6080\(03\)00015-3](https://doi.org/10.1016/S1041-6080(03)00015-3)
- Brybaert, M., Stevens, M., Mandera, P. & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, 7, Artikel 1116. <https://doi.org/10.3389/fpsyg.2016.01116>

- Calder, S. D., Brennan-Jones, C. G., Robinson, M., Whitehouse, A. & Hill, E. (2023). How we measure language skills of children at scale: A call to move beyond domain-specific tests as a proxy for language. *International Journal of Speech-Language Pathology*, 25(3), 440-448. <https://doi.org/10.1080/17549507.2023.2171488>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40(153-193). <https://doi.org/10.1037/h0059973>
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1–22. <https://doi.org/10.1037/h0046743>
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Houghton Mifflin Company.
- Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Elsevier Science Publishers.
- Christensen, H., Batterham, P. J. & Mackinnon, A. J. (2013). The getting of wisdom: Fluid intelligence does not drive knowledge acquisition. *Journal of Cognition and Development*, 14(2), 321-331. <https://doi.org/10.1080/15248372.2012.664590>
- Creswell, J. W. & Guetterman, T. C. (2021). *Educational research: Planning, conduction, and evaluating quantitative and qualitative research* (6. utg.). Pearson Education Limited.
- Dollinger, S. J., Matyja, A. M. & Huber, J. L. (2008). Which factors best account for academic success: Those which college students can control or those they cannot? *Journal of Research in Personality*, 42(4), 872-885. <https://doi.org/10.1016/j.jrp.2007.11.007>
- Duff, D., Tomblin, J. B. & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew effect. *Journal of Speech, Language, and Hearing Research*, 58(3), 853–864. https://doi.org/10.1044/2015_JSLHR-L-13-0310
- Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997). *The british picture vocabulary scale* (2. utg.). Nelson Publishing Company.
- Eckhoff, N. (1997). *Elever med generelle lærevansker: Opplæring i den videregående skolen*. Ad Notam Gyldendal.
- Egeland, J. (2022). Theories of independent intelligences as a Lakatosian research program. *Philosophia*, 50, 2441–2456. <https://doi.org/10.1007/s11406-022-00565-1>

- Ferrer, E. & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, 40(6), 935–952. <https://doi.org/10.1037/0012-1649.40.6.935>
- Ferrer, E., McArdle, J. J., Shaywitz, B. A., Holahan, J. M., Marchione, K. & Shaywitz, S. E. (2007). Longitudinal models of developmental dynamics between reading and cognition from childhood to adolescence. *Developmental Psychology*, 43(6), 1460–1473. <https://doi.org/10.1037/0012-1649.43.6.1460>
- Ferrer, E., O’Hare, E. D. & Bunge, S. A. (2009). Fluid reasoning and the developing brain. *Frontiers in Neuroscience*, 3(1), 46–51. <https://doi.org/10.3389%2Fneuro.01.003.2009>
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, 21(1), 93-101. <https://doi.org/10.1177/0956797609354084>
- Gallinat, E. & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: a meta-analysis. *Journal of Speech, Language, and Hearing Research*, 57(4), 1363-1382. https://doi.org/10.1044/2014_JSLHR-L-12-0363
- Gathercole, S. E., Willis, C. S., Emslie, H. & Baddeley, A. D. (1992). Phonological memory and vocabulary development during early school years: A longitudinal study. *Developmental Psychology*, 28(5), 887–898. <https://doi.org/10.1037/0012-1649.28.5.887>
- Gertner, B. L., Rice, M. L. & Hadley, P. A. (1994). Influence of communicative competence on peer preferences in a preschool classroom. *Journal of Speech & Hearing Research*, 37(4), 913–923. <https://doi.org/10.1044/jshr.3704.913>
- Gignac, G. E. (2014). Dynamic mutualism versus g factor theory: An empirical test. *Intelligence*, 42, 89–97. <http://dx.doi.org/10.1016/j.intell.2013.11.004>
- Gignac, G. E. (2016). Residual group-level factor associations: Possibly negative implications for the mutualism theory of general intelligence. *Intelligence*, 55, 69-78. <https://doi.org/10.1016/j.intell.2016.01.007>
- Gilje, N. (2019). *Hermeneutikk som metode: Ein historisk introduksjon*. Samlaget.
- Griffiths, S., Kievit, R. A. & Norbury, C. (2022). Mutualistic coupling of vocabulary and non-verbal reasoning in children with and without language disorder. *Developmental Science*, 25(3). <https://doi.org/10.1111/desc.13208>

- Grigorenko, E. L. & Sternberg, R. J. (Red.). (2002). *The general factor of intelligence: How general is it?* Lawrence Erlbaum Associates.
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8(3), 179–203. [https://doi.org/10.1016/0160-2896\(84\)90008-4](https://doi.org/10.1016/0160-2896(84)90008-4)
- Hagen, Å. M. M., Melby-Lervåg, M. & Lervåg, A. (2014). Barn som strever med språk og leseferdigheter: Hvilke tiltak virker? *Norsk tidsskrift for logopedi*, 4, 6-13.
- Hagtvet, B. E., Lyster, S.-A. H., Melby-Lervåg, M., Næss, K.-A. B., Hjetland, H. N., Engevik, L. I., . . . Kruse, J. (2011). Ordforråd i førskolealder og senere leseferdigheter: En metaanalytisk tilnærming. *Spesialpedagogikk*, (1), 34-49.
- Haier, R. J., Colom, R., Schroeder, D. H., Condon, C. A., Tang, C., Eaves, E. & Head, K. (2009). Gray matter and intelligence factors: Is there a neuro-g? *Intelligence*, 37(2), 136-144. <https://doi.org/10.1016/j.intell.2008.10.011>
- Hamaker, E. L., Kuiper, R. M. & Grasman, R. P. P. P. (2015). A critique of the cross lagged panel model. *Psychological Methods*, 20(1), 102-116. <https://doi.org/10.1037/a0038889>
- Harris, J. C. (2013). New terminology for mental retardation in DSM-5 and ICD-11. *Current Opinion in Psychiatry*, 26(3), 260-262. <https://doi.org/10.1097/YCO.0b013e32835fd6fb>
- Haugen, R. (2016). Lærevansker - historikk og begrepsdefinisjoner. I R. Haugen (Red.), *Barn og unges læringsmiljø 4 - med vekt på lærevansker* (s. 15-39). Cappelen Damm Akademisk.
- Helland, T. (2019). *Språk og dysleksi* (2. utg.). Fagbokforlaget.
- Helland-Riise, F. & Martinussen, M. (2017). Måleegenskaper ved de norske versjonene av Ravens matriser [Standard Progressive Matrices (SPM)/Coloured Progressive Matrices (CPM)]. *PsykTestBarn*, 7(2). <https://doi.org/10.21337/0055>
- Hjetland, H. N. & Johansen, T. B. (2009). *Kjønnsforskjeller innen språkutvikling? En studie av kjønnsforskjeller innenfor barns tale- og skriftspråksferdigheter i fire- og femårsalderen* [Masteroppgave]. <https://www.duo.uio.no>
- Hjetland, H. N., Lervåg, A., Lyster, S.-A. H., Hagtvet, B. E., Hulme, C. & Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *Journal of Educational Psychology*, 111(5), 751-763. <https://doi.org/10.1037/edu0000321>
- Hoff, E. (2014). *Language development* (5. utg.). Wadsworth Cengage Learning.

- Hofman, A. D., Kievit, R. A., Stevenson, C. E., Molenaar, D., Visser, I. & van der Maas, H. L. J. (2018). The dynamics of the development of mathematics skills: A comparison of theories of developing intelligence. <https://doi.org/10.31219/osf.io/xa2ft>
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75(3), 242–259. <https://doi.org/10.1037/h0025662>
- Horn, J. L. & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology*, 57(5), 253–270. <https://doi.org/10.1037/h0023816>
- Hulme, C., Snowling, M. J., West, G., Lervåg, A. & Melby-Lervåg, M. (2020). Children’s language skills can be improved: Lessons from psychological science for educational policy. *Current Directions in Psychological Science*, 29(4). <https://doi.org/10.1177/0963721420923684>
- Jaeggi, S. M., Buschkuhl, M., Jonides, J. & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, USA*, 105(19), 6829-6833. <https://doi.org/10.1073/pnas.0801268105>
- Jakobsen, T. G. (2021). *Vitenskapsfilosofi og kritisk realisme – et ikke-antroposentrisk alternativ*. Fagbokforlaget.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.
- Jensen, A. R. (2002). Psychometric g: Definition and substantiation. I R. J. Sternberg & E. L. Grigorenko (Red.), *The general factor of intelligence: How general is it?* (s. 39-53). Lawrence Erlbaum Associates.
- Kane, M. J., Hambrick, D. Z. & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131(1), 66–71. <https://doi.org/10.1037/0033-2909.131.1.66>
- Karbach, J. & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive control and working memory training in younger and older adults. *Psychological Science*, 25(11), 2027–2037. <https://doi.org/10.1177/0956797614548725>
- Kearney, M. W. (2017). Cross-lagged panel analysis. I M. Allen (Red.), *The SAGE Encyclopedia of Communication Research Methods*.
- Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M., Moutoussis, M., . . . Dolan, R. J. (2018). Developmental cognitive neuroscience using

- latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience*, 33, 99-117. <https://doi.org/10.1016/j.dcn.2017.11.007>
- Kievit, R. A., Hofman, A. D. & Nation, K. (2019). Mutualistic coupling between vocabulary and reasoning in young children: A replication and extension of the study by Kievit et al. (2017). *Psychological Science*, 30(8), 1245–1252. <https://doi.org/10.1177/0956797619841265>
- Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T. & Dolan, R. J. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science*, 28(10), 1419–1431. <https://doi.org/10.1177/0956797617710785>
- Kohs, S. C. (1920). The block-design tests. *Journal of Experimental Psychology*, 3(5), 357–376. <https://doi.org/10.1037/h0074466>
- Kovacs, K. & Conway, A. R. A. (2019). A unified cognitive/differential approach to human intelligence: Implications for IQ testing. *Journal of Applied Research in Memory and Cognition*, 8(3), 255-272. <https://doi.org/10.1016/j.jarmac.2019.05.003>
- Kristoffersen, K. E., Rygvold, A.-L., Klem, M., Valand, S. B., Asbjørnsen, A. & Næss, K.-A. (2021). Terminologi for vansker med språk hos barn og unge – en konsensusstudie. *Norsk tidsskrift for logopedi*, (3), 6–23.
- Krämer, S., Möller, J. & Zimmermann, F. (2021). Inclusive education of students with general learning difficulties: A meta-analysis. *Review of Educational Research*, 91(3), 432-478. <https://doi.org/10.3102/0034654321998072>
- Kuschner, E. S. (2013). Nonverbal intelligence. I F. R. Volkmar (Red.), *Encyclopedia of Autism Spectrum Disorders*. Springer.
- Kvist, A. V. & Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence*, 36(5), 422-436. <https://doi.org/10.1016/j.intell.2007.08.004>
- 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*, 74(9), 37–45.
- Marchman, V. A. & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11(3). <https://doi.org/10.1111/j.1467-7687.2008.00671.x>
- McArdle, J. J., Ferrer-Caja, E., Hamagami, F. & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual

- abilities over the life span. *Developmental Psychology*, 38(1), 115–142.
<https://doi.org/10.1037//0012-1649.38.1.115>
- McCallum, R. S. (2017). Context for nonverbal assessment of intelligence and related abilities. I R. S. McCallum (Red.), *Handbook of nonverbal assessment* (2. utg., s. 3–19). Springer.
- McGregor, K. K., Oleson, J., Bahnsen, A. & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders*, 48(3), 307–319. <https://doi.org/10.1111/1460-6984.12008>
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1–10. <https://doi.org/10.1016/j.intell.2008.08.004>
- Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291. <https://doi.org/10.1037/a0028228>
- Melby-Lervåg, M., Lervåg, A., Lyster, S.-A. H., Klem, M. & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children’s vocabulary development. *Psychological Science*, 23(10), 1092–1098. <https://doi.org/10.1177/0956797612443833>
- Melby-Lervåg, M., Redick, T. S. & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of «far transfer»: Evidence from a meta-analytic review. *Perspectives on Psychological Science*, 11(4), 512–534. <https://doi.org/10.1177/1745691616635612>
- Miller, L. T. & Lee, C. J. (1993). Construct validation of the Peabody Picture Vocabulary Test—Revised: A structural equation model of the acquisition order of words. *Psychological Assessment*, 5(4), 438–441. <https://doi.org/10.1037/1040-3590.5.4.438>
- Nagy, W., Herman, P. & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253. <https://doi.org/10.2307/747758>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2. utg.). Cambridge University Press.
- NESH. (2021). *Forskningsetiske retningslinjer for samfunnsvitenskap og humaniora* (5. utg.). De nasjonale forskningsetiske komiteene.
- Norbury, C., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of

- language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 47(11), 1247-1257. <https://doi.org/10.1111/jcpp.12573>
- Nutley, S. B., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K. & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science*, 14(3), 591–601. <https://doi.org/10.1111/j.1467-7687.2010.01022.x>
- Nyeng, F. (2012). *Nøkkeltbegreper i forskningsmetode og vitenskapsteori*. Fagbokforlaget.
- Næss, K.-A. B., Engevik, L. I., Hokstad, S. & Mjøberg, A. G. (2017). Vokabularstimuleringstiltak for barn med Down syndrom: Hva sier forskningen? *Norsk tidsskrift for logopedi*, (1).
- Pearson Clinical Assessment's Scientific Council. (2023). *Standardized clinical assessment for practitioners: A primer*. Pearson. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/featured-topics/assessment-primer-whitepaper.pdf>
- Peng, P. & Kievit, R. A. (2020). The development of academic achievement and cognitive abilities: A bidirectional perspective. *Child Development Perspectives*, 14(1), 15-20. <https://doi.org/10.1111/cdep.12352>
- Peng, P., Wang, T., Wang, C. & Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin*, 145(2), 189–236. <https://doi.org/10.1037/bul0000182>
- Peters, L. & Ansari, D. (2019). Are specific learning disorders truly specific, and are they disorders? *Trends in Neuroscience and Education*, 17. <https://doi.org/10.1016/j.tine.2019.100115>
- Primi, R. (2002). Complexity of geometric inductive reasoning tasks: Contribution to the understanding of fluid intelligence. *Intelligence*, 30(1), 41–70. <https://doi.org/10.1016/S0160-2896%2801%2900067-8>
- Quinn, J. M., Wagner, R. K., Petscher, Y. & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, 86(1), 159-175. <https://doi.org/10.1111/cdev.12292>
- Rindermann, H., Flores-Mendoza, C. & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences*, 20(5). <https://doi.org/10.1016/j.lindif.2010.07.002>

- Ringdal, K. (2018). *Enhet og mangfold: Samfunnsvitenskapelig forskning og kvantitativ metode* (4. utg.). Fagbokforlaget.
- Rogde, K., Hagen, Å. M. M., Melby-Lervåg, M. & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, 15(4).
<https://doi.org/10.1002/cl2.1059>
- Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88(2), 245–258. <https://doi.org/10.1037/0033-2909.88.2.245>
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27-47. <https://doi.org/10.1177/2515245917745629>
- Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in Medicine*, 26(1), 20-36. <https://doi.org/10.1002/sim.2739>
- Rygvoid, A.-L. (2017). Språk- og talevansker. I A.-L. Rygvold & T. Ogden (Red.), *Innføring i spesialpedagogikk* (5. utg., s. 161-195). Gyldendal Akademisk.
- Salthouse, T. A., Pink, J. E. & Tucker-Drob, E. M. (2008). Contextual analysis of fluid intelligence. *Intelligence*, 36(5), 464-486. <https://doi.org/10.1016/j.intell.2007.10.003>
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillian.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951. <https://doi.org/10.1111/lang.12077>
- Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching*, 52(2), 261–274.
<https://doi.org/10.1017/S0261444819000053>
- Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R. & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, 109(8), 1120–1140. <https://doi.org/10.1037/edu0000193>
- Schuth, E., Köhne, J. & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction*, 49, 157-165.
<https://doi.org/10.1016/j.learninstruc.2017.01.005>

- Schweizer, K. & Koch, W. (2002). A revision of Cattell's Investment Theory: Cognitive properties influencing learning. *Learning and Individual Differences*, 13(1), 57-82. [https://doi.org/10.1016/S1041-6080\(02\)00062-6](https://doi.org/10.1016/S1041-6080(02)00062-6)
- Selig, J. P. & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. I B. Laursen, T. D. Little & N. A. Card (Red.), *Handbook of developmental research methods* (s. 265-278). Guilford Press.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shahabi, S. R., Abad, F. J. & Colom, R. (2018). G, mutualism, and development: Cross-sectional evidence from Iranian schoolchildren. *Personality and Individual Differences*, 135, 222-228. <https://doi.org/10.1016/j.paid.2018.07.021>
- Sigstad, H. M. H. (2017). Utviklingshemming: Hva er utviklingshemming? Definisjoner og beskrivelser. I A.-L. Rygvold & T. Ogden (Red.), *Innføring i spesialpedagogikk* (5. utg., s. 132-160). Gyldendal Akademisk.
- Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M. & Hulme, C. (2020). Dyslexia and developmental language disorder: Comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry*, 61(6), 672-680. <https://doi.org/10.1111/jcpp.13140>
- Snowling, M. J. & Hulme, C. (2021). Annual research review: Reading disorders revisited – the critical importance of oral language. *Journal of Child Psychology and Psychiatry*, 62(5), 635–653. <https://doi.org/10.1111/jcpp.13324>
- Snowling, M. J., Hulme, C. & Nation, K. (2020). Defining and understanding dyslexia: Past, present and future. *Oxford Review of Education*, 46(4), 501-513. <https://doi.org/10.1080/03054985.2020.1765756>
- Spearman, C. (1904). «General Intelligence», objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292. <https://doi.org/10.2307/1412107>
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. The Macmillan Company.
- Stephens, R. G., Dunn, J. C. & Hayes, B. K. (2018). Are there two processes in reasoning? The dimensionality of inductive and deductive inferences. *Psychological Review*, 125(2), 218-244. <https://doi.org/10.1037/rev0000088>
- Sternberg, R. J., Kaufman, J. C. & Grigorenko, E. L. (2008). *Applied intelligence*. Cambridge University Press.

- Strong, G. K., Torgerson, C. J., Torgerson, D. & Hulme, C. (2011). A systematic meta-analytic review of evidence for the effectiveness of the 'Fast ForWord' language intervention program. *Journal of Child Psychology and Psychiatry*, 52(3), 224–235. <https://doi.org/10.1111/j.1469-7610.2010.02329.x>
- Saar, V., Komulainen, E. & Levänen, S. (2023). The significance of nonverbal performance in children with developmental language disorder. *Child Neuropsychology*, 29(2), 213-234. <https://doi.org/10.1080/09297049.2022.2077324>
- te Nijenhuis, J., van Vianen, A. E. M. & van der Flier, H. (2007). Score gains on g-loaded tests: No g. *Intelligence*, 35(3), 283-300. <https://doi.org/10.1016/j.intell.2006.07.006>
- Thagaard, T. (2018). *Systematikk og innlevelse: En innføring i kvalitative metoder* (5. utg.). Fagbokforlaget.
- Thomson, G. (1951). *The factorial analysis of human ability* (5. utg.). University of London Press.
- Toffalini, E., Giofrè, D., Pastore, M., Carretti, B., Fraccadori, F. & Szűcs, D. (2021). Dyslexia treatment studies: A systematic review and suggestions on testing treatment efficacy with small effects and small samples. *Behavior Research Methods*, 53, 1954–1972. <https://doi.org/10.3758/s13428-021-01549-x>
- van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M. & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, 113(4), 842–861. <https://doi.org/10.1037/0033-295X.113.4.842>
- van Viersen, S., de Bree, E., Verdam, M., Krikhaar, E., Maassen, B., van der Leij, A. & de Jong, P. F. (2017). Delayed early vocabulary development in children at family risk of dyslexia. *Journal of Speech, Language, and Hearing Research*, 60(4), 937-949. https://doi.org/10.1044/2016_JSLHR-L-16-0031
- Verdens helseorganisasjon. (2019). *ICD-11: International Classification of Diseases 11th Revision*. <https://icd.who.int/>
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22(2), 217-234. <https://doi.org/10.1017/S0142716401002041>
- Wechsler, D. (2009). *Wechsler intelligence scale for children. Norsk versjon* (4. utg.). Pearson Assessment.

