

Arbeidsminne og aritmetiske ferdigheter

En kvantitativ studie om sammenhengen mellom norske barnehagebarns arbeidsminne og deres aritmetiske ferdigheter i 1. klasse

Sina Sabotta



Masteroppgave i spesialpedagogikk
Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultetet
UNIVERSITETET I OSLO

Vår 2020

Arbeidsminne og aritmetiske ferdigheter

En kvantitativ studie om sammenhengen mellom norske barnehagebarns arbeidsminne og deres aritmetiske ferdigheter i 1. klasse

© Sina Sabotta

2020

Arbeidsminne og aritmetiske ferdigheter

En kvantitativ studie om sammenhengen mellom norske barnehagebarns arbeidsminne og deres aritmetiske ferdigheter i 1. klasse

Sina Sabotta

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

Trykk: Reprosentralen, Universitetet i Oslo

IV

Sammendrag

Det å kunne regne blir ansett som en «grunnleggende ferdighet som går på tvers av fag» (Udir, 2018), og norske skoler følger det lovfestede prinsippet om tilpasset opplæring (Lovdata, 2020), for å kunne møte disse kravene, kreves det kunnskap om individuelle variasjoner som påvirker læring for å kunne gi tilpasset opplæring. Man antar blant annet at omtrent 15-20% av barn opp til ungdomsalder har en grad av matematikkvansker (Mononen, 2018), for å gi disse barna faglig støtte ved et tidlig tidspunkt, er det essensielt å ha kunnskap om hvilke indikatorer som tidlig kan vise til en eventuell vanske.

Tidligere forskning viser at matematiske ferdigheter ved skolestart har stor betydning for senere matematiske ferdigheter (Price & Ansari, 2013), i tillegg viser også forskning at arbeidsminne er en sterk prediktor for aritmetiske ferdigheter (Gilmore, Göbel, & Inglis, 2018). Med utgangspunkt i dette forsøker denne studien å utvide kunnskapen omkring sammenhengen mellom ulike delkomponenter ved arbeidsminnet ved 5 årsalder og de aritmetiske ferdighetene addisjon og subtraksjon i første klasse hos norske barn. Dette gjøres med utgangspunkt i forskningsspørsmål som har som formål å belyse en unik grad av variasjon i nettopp denne longitudinelle sammenhengen. Ved å øke kunnskap rundt dette vil man kunne bidra til å forbedre kartleggingsrutiner for å kunne gi barn i faresonen tidlig støtte under opplæringen, i tillegg til å kunne forbedre eller utvikle forskningsbaserte intervensjonsprogrammer.

I den foreliggende studien har det blitt valgt et ikke-eksperimentelt design, med en longitudinell og kvantitative tilnærming. Utvalget som denne studien baserer seg på 254 enspråklig norske barn uten lærevansker, som er en del av en større studie som kalles NumLit: Development of numeracy and literacy in children. Barnas arbeidsminne ble målt ved hjelp av deltestene Listening recall, Maze memory og Backward digit span fra Working Memory Test Battery for Children (WMTB-C), mens aritmetiske ferdigheter ble målt ved hjelp av deltestene addisjon og subtraksjon, hoderegning addisjon og subtraksjon fra Test of Basic Arithmetic and Numeracy Skills (TOBANS) i tillegg til Ordproblemer fra Wechsler Intelligence Scale for Children (WISC), som er en deltest for regning. Kontrollvariabelen Nonverbal intelligens ble målt ved hjelp av Ravens Coloured Progressive Matrices (CPM).

For å belyse sammenhengen mellom arbeidsminnet og aritmetiske ferdigheter ble det gjennomført en bivariat korrelasjonsanalyse og en hierarkisk multippel regresjonsanalyse i det statistiske analyseprogrammet SPSS (IBM, 2019). Resultatene viser statistisk signifikante

forklaringsvariabler samt statistisk signifikante positive korrelasjoner mellom testene for arbeidsminne og aritmetiske ferdigheter.

Forord

Det har vært en lang, spennende og ikke minst lærerik reise frem mot det å nå få lov til å levere denne masteroppgaven. I begynnelsen av 2020 var jeg veldig klar over at dette semesteret kom til å bli veldig spesielt. Det å bruke et helt semester på å skrive en oppgave gir på mange måter en annen studiehverdag enn det man er vant til fra før. I tillegg har det siste halvåret også vært preget av usikkerhet og samfunnsmessige restriksjoner, som har gjort at man har måttet tenke nytt og ikke minst positivt, for å gjøre det beste ut av situasjonen man er i.

Takk til mine veiledere Tonje Amland og Silke M. Göbel, som har bistått med gode tilbakemeldinger og råd, både via videosamtaler og e-mail, gjennom hele skriveprosessen.

Takk til NumLit for å ha gitt meg muligheten til å delta i et forskningsprosjekt og for at jeg kunne bruke deler av datamaterialet til min masteroppgave.

Takk til min familie, for å vise tiltro i meg og for å sørge for små avkoblingsmoment med telefonsamtaler, is på trappa og bading.

Takk til min samboer for faglige diskusjoner, for støtten og oppmuntringene du har gitt meg.

Juni 2020

Sina Sabotta

Innholdsfortegnelse

1	Innledning.....	1
1.1	Tematisk bakgrunn	1
1.2	Avgrensning og problemstilling	2
1.3	Oppgavens oppbygning	3
2	Teoretisk bakgrunn og empiriske funn.....	4
2.1	Aritmetikk.....	4
2.1.1	Begrepsavklaring.....	4
2.1.2	Utvikling av aritmetiske ferdigheter.....	4
2.1.3	Aritmetikk og tekstoppgaver.....	9
2.2	Arbeidsminne.....	10
	Central executive.....	12
	Det verbale arbeidsminne (phonological loop)	13
	Det visuospatiale arbeidsminne (visuospatial sketchpad).....	14
	Episodisk buffer	14
2.2.1	Utviklingen av arbeidsminne	15
2.2.2	Arbeidsminnets betydning for aritmetikk.....	16
2.2.3	Sammenhengen mellom aritmetikk og intelligens	18
3	Metode.....	19
3.1	Kartlegging	20
3.1.1	Kartlegging av arbeidsminne.....	21
3.1.2	Kartlegging av aritmetiske ferdigheter.....	24
3.1.3	Kartlegging av variabelen nonverbal intelligens.....	26
3.2	Validitet	27
3.2.1	Statistisk validitet	29
3.2.2	Indre validitet	30
3.2.3	Begrepsvaliditet.....	30
3.2.4	Ytre validitet.....	31
3.3	Etiske hensyn.....	32
4	Analyse.....	34
4.1	Rensing av data før dataanalysen	35
4.2	Deskriptiv analyse	35

4.2.1	Vurdering av variabelen: Backward Digit Span, WMTB-C	37
4.2.2	Vurdering av variabelen: Maze Memory	38
4.2.3	Vurdering av variabelen: Listening Recall, WMTB-C	39
4.2.4	Vurdering av variabelen: Ordproblemer (Math Word problem).....	40
	40
4.2.5	Vurdering av variabelen: TOBANS addisjon og subtraksjon	41
4.2.6	Vurdering av variabelen: Hoderegning addisjon og subtraksjon	43
4.2.7	Vurdering av variabelen: The Raven's coloured Progressive Matrices	45
4.3	Reliabilitet	46
4.4	Bivariat korrelasjonsanalyse	48
4.5	Hierarkisk Multippel Regresjonsanalyse	50
4.5.1	Forutsetninger for hierarkisk multippel regresjonsanalyse	50
4.5.2	Gjennomføring og resultater av hierarkisk regresjonsanalysen	51
4.6	Oppsummering, analyser	61
5	Resultater i lys av teori og empiri	62
5.1	Vurdering av studiens validitet og reliabilitet	66
5.1.1	Statistisk validitet og reliabilitet.....	66
5.1.2	Indre validitet	68
5.1.3	Begrepsvaliditet.....	70
5.1.4	Kriterievaliditet	72
5.1.5	Spesialpedagogiske implikasjoner	74
5.1.6	Veien videre	75
5.2	Oppsummering	76
	Litteraturliste	78

Antall ord: 20269

1 Innledning

De grunnleggende matematiske ferdighetene er viktige i ulike situasjoner i vårt daglige liv. Ved å tilegne seg slike ferdigheter, legger man et fundament for fremtidig læring på høyere og vanskeligere nivå, noe som bidrar til selvstendighet i hverdagen.

1.1 Tematisk bakgrunn

Det antas at omtrent 15 til 20 % av de som er i ungdomsalder og yngre har en grad av matematikkvansker, noe som medfører en utfordring med å kunne tilegne seg matematiske ferdigheter i samme grad som deres jevnaldere (Mononen, 2018).

Matematikksenteret, nasjonalt senter for matematikk i opplæringen (2018) sier at «regning er en grunnleggende ferdighet som går på tvers av fag». Under Stortingsmeldingen nr. 030 (2003-2004), kultur for læring, defineres «basiskompetansene som lese- og skriveferdigheter, regneferdigheter og tallforståelse ...» (St.meld. nr. 030 (2003-2004), 2018).

Videre sier dette dokumentet at det kreves ulike matematiske ferdigheter som addisjon, subtraksjon, multiplikasjon, divisjon og forholdstall for å kunne mestre både faglige og hverdagslige situasjoner (St.meld. nr. 030 (2003-2004), 2018).

Forskning viser at de matematiske ferdighetene man har så tidlig som ved skolestart, har stor betydning for senere matematiske ferdigheter (Price & Ansari, 2013; Kyttälää, Aunio, Lepola, & Hautamäki, 2014). Det vil derfor være viktig at den grunnleggende matematikkforståelsen etableres allerede i barnehagealder, for å kunne ha gode forutsetninger gjennom skoleløpet, mestring av hverdagen, samt jobbhverdag.

For å kunne gjøre en tidlig innsats gjelder det å finne indikatorer som ligger til grunn for matematiske ferdigheter (Kyttälää, Aunio, Lepola, & Hautamäki, 2014), som tidlig kan vise til en eventuell vanske. Da man ikke kan anta at barn i barnehagealder har utviklet ferdigheter i hverken skriving eller regning, kan det å se på barnas generelle kognitive ferdigheter gi en prediksjon av senere matematiske ferdigheter (Hornung, Schiltz, Brunner, & Martin, 2014). Med prediksjon mener man for eksempel at generelle kognitive ferdigheter kan gi en statistisk beregnet antakelse om et fremtidig forhold, for eksempel en persons senere matematiske ferdigheter. Man antar derfor at, det er en tett kobling mellom den kognitive komponenten

arbeidsminne og aritmetiske ferdigheter (Hitch & McLean, 1999; Baddeley, 2000; Price & Ansari, 2013; Kyttälä, Aunio, Lepola, & Hautamäki, 2014). I denne studien velges det å se på om arbeidsminne kan være en slik indikator, i tillegg til generell intelligens.

Videre er arbeidsminne ansvarlig for å samtidig prosessere og lagre informasjon under kognitive oppgaver (Lefevre, Wells, & Sowinski, 2015; Baddeley, 2000). Antakelsen er derfor at arbeidsminnet spiller en viktig rolle under løsning av matematiske oppgaver.

Det finnes hittil en del forskning som ser på norske barn innenfor dette feltet, men det ser allikevel ut til at det er behov for mer forskning. En metaanalyse viser til at noen individer har en sterkere relasjon mellom arbeidsminne og matematiske ferdigheter enn andre, og at man for å kunne lage gode matematiske intervensjoner for individer med matematikkvansker, behøver å forske mer på blant annet forskjellige typer matematikk og ulike deler av arbeidsminne (Peng, Namkung, Barnes, & Sun, 2016). Den foreliggende studien vil kunne bidra til å gi mer innsikt i arbeidsminnets ulike delkomponenter og deres betydning for tidlige aritmetiske ferdigheter.

Med dette som grunnlag vil mulige funn potensielt kunne ha betydning for å kunne gjøre en tidlig innsats med tanke på å hjelpe barna som viser behov, slik at også disse barna skal få opplevd mestring i daglige og faglige situasjoner.

1.2 Avgrensning og problemstilling

Med utgangspunkt i det overnevnte, lyder min problemstilling derfor som følger:

Kan arbeidsminne ved 5 årsalder predikere aritmetiske ferdigheter i førsteklasse?

Videre har det blitt formulert to forskningsspørsmål, som skal bidra til å besvare problemstillingen:

- I. Hvilken sammenheng er det mellom barnas arbeidsminne og grunnleggende aritmetiske ferdigheter?
- II. Hvilken del av arbeidsminnet korrelerer sterkest med grunnleggende aritmetiske ferdigheter, når det kontrolleres for intelligens?

Med utgangspunkt i teori og empiri, vil denne studien forsøke å besvare problemstillingen samt forskningsspørsmålene.

Den foreliggende studien tar utgangspunkt i et relativt stort utvalg av norske barn som har gjennomført tester innenfor ulike ferdigheter. Denne studien har derfor muligheten til å inkludere flere typer utfallsvariabler.

1.3 Oppgavens oppbygning

I det første og innledende kapittelet vil det gis en beskrivelse av bakgrunnen for det tematiske valget; arbeidsminne og aritmetiske ferdigheter. Videre gis det en redegjørelse for relevansen temaet har for det spesialpedagogiske fagfeltet, med utgangspunkt i problemstilling og forskningsspørsmål.

I det andre kapittelet vil det bli gitt en redegjørelse av den teoretiske og empiriske bakgrunnen for aritmetiske ferdigheter samt arbeidsminne. I delkapittel 2.1 forklares det hva aritmetikk er og hvordan aritmetiske ferdigheter utvikles, mens delkapittel 2.2 vil gi en oversikt over arbeidsminnet, dets ulike delkomponenter, utvikling i tillegg til arbeidsminnets betydning for aritmetikk.

I det tredje kapittelet gis det en beskrivelse av metodene som brukes, i tillegg til at de valgte variablene og kartleggingsverktøy skal beskrives. Det blir også gitt en beskrivelse av validitet, i tillegg til eventuelle trusler mot ulike validitetstyper i kvantitativ forskning. Det vil også bli gitt en beskrivelse av forskningsetiske hensyn.

I det fjerde kapittelet vil den deskriptive analysen og resultatene bli presentert. Det vil bli vist og beskrevet histogrammer av variablenes distribusjoner for så å beskrive reliabilitet i tillegg til å vurdere variablenes reliabilitet. Videre skal resultatene fra den bivariate korrelasjonsanalysen og den hierarkisk multiple regresjonsanalysen legges frem med de tilhørende resultatene, for så å oppsummere analysene.

I det femte kapittelet vil resultatene bli drøftet i forhold til teori og empiri, for så å drøfte resultater i forhold til studiens validitet og reliabilitet. Det vil bli gitt en redegjørelse for implikasjoner for det praktisk-pedagogiske fagfeltet og det vil bli drøftet mulige veier for videre forskning.

2 Teoretisk bakgrunn og empiriske funn

2.1 Aritmetikk

2.1.1 Begrepsavklaring

Aritmetikk defineres som det «å legge til» og «å trekke fra» ensifrede tall (Fuchs, et al., 2006) og kan dermed ansees som bruken av konseptet om tallsystemer og gjennomføringen av oppgaver med addisjon, subtraksjon, multiplikasjon og divisjon (Hindry, 2011). Aritmetiske ferdigheter kan derfor beskrives som evnen til å memorere grunnleggende regler og det å mestre algoritmer som kreves for å løse de fire grunnleggende regnemetodene: addisjon, subtraksjon, multiplikasjon og divisjon (Wright, Ellemor-Collins, & Tabor, 2012).

I den følgende delen skal det presenteres ulike modeller av hvordan ulike forskere antar at aritmetiske ferdigheter utvikles, for så å gi en sammenlikning av disse modellene.

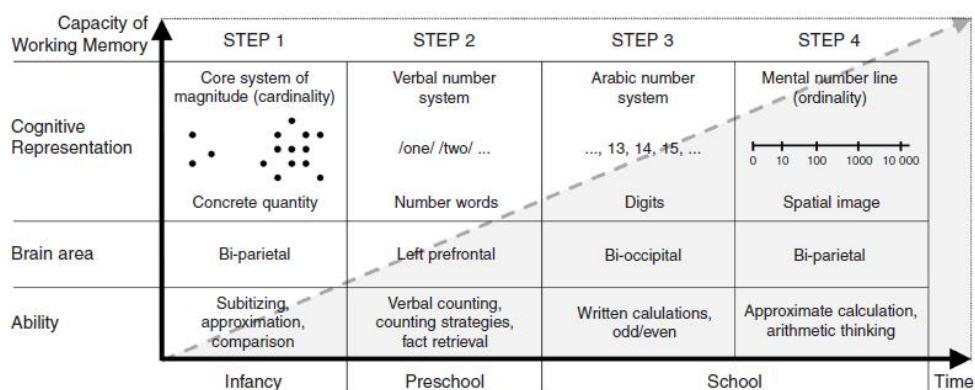
2.1.2 Utvikling av aritmetiske ferdigheter

Fire-steg-utviklingsmodellen er en fremstilling av hvordan numerisk kognisjon utvikles. Modellen ble lagt fram av von Aster og Shalev (2007), og viser en dynamisk utvikling samtidig som den anerkjenner et samspill mellom alder og ferdigheter og mellom numeriske og ikke-numeriske prosessering (Kaufmann, Kucian, & von Aster, 2015). Modellen gir videre en illustrasjon på hvilke kognitive tallrepresentasjoner et barn skal ha utviklet fra barnet er nyfødt, i løpet av barnehagetiden og de første årene på barneskolen (von Aster & Shalev, 2007).

Modellens steg 1 beskriver medfødte kjernesystemrepresentasjoner av grunntall (kardinalatet) med sine tilhørende funksjoner, som for eksempel å nøyaktig benevne representasjoner av små mengder som kan bli fremstilt av ulike elementer (f.eks. rask gjenkjenning av tallet 5 ved å se på en terning med siden hvor 5 øyne synes, subitizing) og å anslå mengder (Approximate number system, ANS). Steg 1 gir altså grunnlaget for den grunnleggende betydningen av tall og antall (von Aster & Shalev, 2007). Von Aster og Shalev (2007) skriver at disse ferdighetene er viktige forutsetninger for at barn skal lære å knytte et antall av objekter,

muntlig talte tallord og senere skriftlige- og arabiske-representasjoner av tall, til hverandre. Steg 2 innebærer å prosessere lingvistiske representasjoner av tall, mens steg 3 innebærer arabiske representasjoner av tall, noe som er forutsetningen for utviklingen av den mentale tallinjen i steg 4. Videre er også ordinalitet en del av steg 4, der ordinalitet utgjør det andre medfødte kjenesystemet av tall. Prinsippet ordinalitet baserer seg på å arrangere elementer i en bestemt rekkefølge, hvor hvert element får en slags numerisk merkelapp (Kadosh & Dowker, 2016). Dette prinsippet benyttes for eksempel ved navngivning av klassetrinn, første, andre, tredje og så videre.

Siden numerisk prosessering er avhengig av etablering av disse tidlige stegene, kan det medføre store konsekvenser for senere matematiske ferdigheter, dersom disse ikke etableres. For eksempel kan en genetisk sårbarhet føre til at steg 1 ikke blir etablert på riktig måte. Med genetisk sårbarhet menes en medfødt genetisk svakhet, som for eksempel utviklingsmessig dyskalkuli, som medfører en svekkelse i utviklingen av grunnleggende matematiske ferdigheter (von Aster & Shalev, 2007). Von Aster og Shalev (2007) foreslår at man da kan lære navnet til et tall på en fonologisk måte, ved for eksempel å si navnet på et tall i tillegg til å benytte memorering og repetering. I andre tilfeller kan språkutviklingen være forsinket, noe som kan medføre at koblingen mellom nonverbale numeriske egenskaper, som å kjenne igjen tre prikker, og deres språklige representasjon, altså verbal uttalte tallord, ikke kan etableres på en aldersmessig måte som i steg 2. Dette kan medføre en forsinkelse i utviklingen av blant annet tellestrategier, aritmetikk og lagringen av tallfakta. Trolig kan dette også gjelde for barn med primærvansker innen oppmerksomhet og arbeidsminne, da disse barna har vansker med lagre og å produsere tellerrekker eller rekker som inneholder elementer av tallord. Videre vil unøyaktigheter og hyppige feil under aritmetisk telling også forsinke eller hindre utvikling av strategier som bidrar til å tilbakekalle aritmetiske fakta (von Aster & Shalev, 2007).



Figur 1: Four-step-developmental model of numerical cognition (von Aster & Shalev, 2007)

Det finnes flere modeller som beskriver tallrepresentasjoner som er involvert i aritmetikk, blant disse er *triple-code modellen*, *abstract-code modellen*, og *encoding complex modellen*.

Triple-code modellen er en kognitiv modell av hvordan man prosesserer tall (Dehaene, 1992). For eksempel er man spesielt avhengig av et godt utviklet kognitivt system, som effektivt kan hente aritmetiske faktakunnskaper fra minnet, under operasjoner som multiplikasjon (Zamarian & Delazer, 2015). I denne modellen deles tallkunnskaper inn i tre hovedelementer som ansees som å være sammenkoblet ved at det er en interaksjon mellom elementene. De tre hovedelementene er; *en visuell-arabisk tallform*, *en auditiv-verbal ordramme*, og *en analog representasjon av numeriske mengder* (Campbell, 2015; Kaufmann, Kucian, & von Aster, 2015; Lindemann & Fischer, 2015). Disse tre elementene bidrar ulikt når oppgaver skal løses (Campbell, 2015).

Det *visuell-arabiske tallformsystemet* bidrar under bearbeiding av input og output av arabiske tall (tall fra 0-9) og ved matematiske operasjoner hvor flersifrede tall er involvert (Göbel, 2015). Forskning viser at dette systemet blir aktivert både når man skal lese eller skrive arabiske tall, og når man skal løse aritmetiske oppgaver som har en skriftlig form (Göbel, 2015; Kaufmann, Kucian, & von Aster, 2015).

Den *auditiv-verbale ordrammen* blir ansett som å være i stand til å koble sammen både muntlig- og skriftlig- numerisk input og output. Den auditiv-verbale ordrammen ser dermed ut til å bidra når man skal telle muntlig, og når tall uttales eller skrives ved å koble sammen verbale- og skriftlige representasjoner av tall. Videre er denne ordrammen også i stand til å tilbakekalle verbale representasjoner av addisjon- og multiplikasjonstabeller som er lagret i langtidsmindet (Campbell, 2015; Göbel, 2015; Kaufmann, Kucian, & von Aster, 2015).

Den *analoge representasjonen av numeriske mengder*, som også blir kalt *omtrentlig tall-sans* (ANS), bearbeider ikke-symbolsk stimuli (Göbel, 2015). Dette gjør den ved å lage en slags mental visuell- og romlig tallinje, som skal hjelpe til når størrelser skal sammenlignes. Denne tallsansens blir ansett som å være aktiv ved estimeringer, når man med nøyaktighet skal gjenkjenne små mengder uten å telle (også kalt *subitizing*), når man skal gjennomføre omtrentlige beregninger og den gir i tillegg støtte til semantiske strategier, som ofte er brukt under utregning av subtraksjonsoppgaver (Campbell, 2015; Göbel, 2015; Kaufmann, Kucian, & von Aster, 2015).

Man kan se at det i hvert av de hovedelementene inngår en verbal faktor, noe som kan tyde på at denne modellen tar utgangspunkt i språkbaserte representasjoner av aritmetikk. Videre

antar denne modellen en additiv og ikke en interaktiv kommunikasjon mellom hovedkategoriene. Dette kan med andre ord tolkes som at hovedkategoriene ikke direkte kommuniserer med hverandre, men benytter seg av språkbaserte representasjoner for å samarbeide.

Modellen nedenfor er en annen modell som kan forklare aritmetiske ferdigheter ved å dele matematiske ferdigheter inn i ulike kategorier, for så å belyse hvordan kategoriene henger sammen og utvikler seg over tid. Denne modellen ble presentert av Aunio og Räsänen i 2015, og den gir en inndeling av matematiske ferdigheter og hvordan de utvikler seg. Ferdighetene kan deles inn i *symbolsk* og *ikke-symbolsk tallforståelse*; *forstå matematiske relasjoner*; *telleferdigheter*; og *grunnleggende aritmetiske ferdigheter*.

I kategorien ikke-symbolsk tallforståelse inngår «*subitising*»; mønster gjenkjenning, og «*approximate number system*» (ANS); sammenlikning av relative størrelser uten å telle, altså det å forstå mengdeforhold. Denne preverbale kunnskapen er relatert til telling, tallmønstre, sammenlikning av mengder, estimering og talltransformasjoner, altså kunnskapen man har før man klarer å verbalisere eller kjenner til de representative tallsymbolene (Aunio & Räsänen, 2015).

Det å forstå matematiske relasjoner dreier seg om konseptet av å ha «et bestemt antall i et sett» (Aunio & Räsänen, 2015). Tidlig matematisk-logiske prinsipper inkluderer kunnskapen om det å trekke slutninger, klassifisere, sammenlikne og å lage 1-1 korrespondanse mellom elementer og tall. Videre er utviklingen av matematisk tekning relatert til det å forstå hva det betyr at et tall er *lik*, *mer* eller *mindre* enn et annet tall.

Grunnleggende aritmetiske prinsipper refererer til forståelsen av additiv sammensetning, *kommutativitet*, *assosiativitet* og *inversjoner*, som gjenspeiler forståelse av forhold mellom variabler i en matematisk oppgave. *Kommutativ forståelse* innebærer at barnet vet at enheter kan bli lagt til i ulik rekkefølge men allikevel gi samme svar. *Assosiativitet* forståelse innebærer at barnet vet at selv om settene er brutt opp og kombinert på nytt, allikevel gir samme svar. Kunnskapen om *operasjonelle symboler* i matematikk, handler om å forstå samt bruke matematiske symboler, som *mindre enn* ($<$), *mer enn* ($>$), *er lik* ($=$), *er ikke lik* (\neq) og å kunne bruke de selvstendig (Aunio & Räsänen, 2015). Forstå *plassverdi og 10-tallsystemet*, dreier seg om å forstå betydningen av ener-, tier- og hundre- plassen i tallsystemet, som er essensiell kunnskap for å vite at ulik plassering gir ulik tallverdi.

Tellferdigheter refererer til kunnskapen om tallord og symboler. Videre viser det til en ferdighet som innebærer å danne koblinger mellom symbol-ord og ord-symbol, altså identifisering og gjenkjenning av tall (Aunio & Räsänen, 2015). Denne kunnskapen er essensiell for å kunne forstå hvordan tallsystemer fungerer og er dermed vesentlig for å utvikle og utvide kunnskapen i de andre matematiske ferdighetene.

Tallordsekvenser er ferdigheten som innebærer å si tallord både fremover og bakover, og å kunne hoppe to'er-, fem'er- og ti'er steg, både fremover og bakover fra et gitt tallord. Å være i stand til å operere med tallordsekvenser er en essensiell forkunnskap for nøyaktige tellekunnskaper. Derfor danner denne ferdigheten også grunnlaget for addisjons- og subtraksjonsoppgaver, dette grunnet i at barn tilegner seg aritmetiske ferdigheter ved å begynne med å bruke tellebaserte strategier (Aunio & Räsänen, 2015). Når man skal telle mengder (Enumeration) trenger man ferdigheter om tallordsekvenser. Disse benyttes også i addisjons- og subtraksjonsoppgaver (aritmetikk). For å mestre dette trenger barn å vite tallordene i riktig rekkefølge, dette for å lage 1-1 korrespondanse (Aunio & Räsänen, 2015).

Grunnleggende ferdigheter i aritmetikk fokuserer på å mestre addisjons- og subtraksjonsoppgaver med tallsymboler. Barn med matematikkvansker har vedvarende vansker med å gjenkalle aritmetiske regler og har derfor vansker med å forstå at store sett av tall er satt sammen av mindre sett (Aunio & Räsänen, 2015). Disse barna bruker derfor ofte strategier for å telle framfor å gjenkalle informasjon.

Ut ifra de ovenfor beskrevne modellene kan man se at utviklingen av matematiske ferdigheter, og dermed også aritmetiske ferdigheter, forutsetter at flere delferdigheter etableres da disse er avhengige av hverandre. To av modellene ser ut til å være dynamiske med et samspill mellom alder og ferdigheter. Fire-steg modellen og Aunio og Räsänen (2015) sin modell gir en beskrivelse av hvordan ferdighetene etableres i forhold til hverandre, fra de mest grunnleggende ferdighetene til mer komplekse. Triple-code modellen strukturerer ikke ferdighetene etter hvilke ferdigheter som ansees til å etableres først, den snarere deler ferdighetene inn i kognitive kategorier som ser ut til å utvikle seg til å bli mere komplekse over tid og dermed legger grunnlag til å løse mer og mer komplekse matematiske problemer.

Videre beskriver de tre modellene ferdighetene subitizing og ANS, prosessering av både verbale- og skriftlige representasjoner av tall, i tillegg til å forstå plassverdisystemet og etablere en mental tallinje. Hverken subitizing eller ANS krever kunnskap om arabiske representasjoner av tall, og blir ansett som en medfødt preverbal ferdighet. Man kan anta at

disse to er ferdigheter som etableres tidlig i utviklingen av aritmetiske ferdigheter. Ut ifra fire-steg modellen og Aunio og Räsänen (2015) kan det se ut som at den neste ferdigheten som etableres er verbale representasjoner, etterfulgt av skriftlige representasjoner. Triple-code modellen nevner også disse ferdighetene, men gir tilsynelatende ikke noe uttrykk for når i utviklingsforløpet disse etableres.

En mental tallinje antas, ifølge fire-steg modellen, å etableres etter å ha begynt formell skolegang. De to andre modellene nevner ikke en omtrentlig alder for når dette etableres, men siden Aunio og Räsänen (2015) beskriver hvordan de ulike ferdighetene utvikler seg i forhold til hverandre, og siden de beskriver at kunnskapen om titallssystemet blir etablert etter verbale- og skriftlige representasjoner, kan det virke sannsynlig at denne delkomponenten utvikles fra omtrent samme tidsrom. Når man har begynt skolegangen, vil utviklingen av matematiske- og aritmetiske ferdigheter sannsynligvis følge det som undervises på skolen. Her gir Aunio og Räsänen (2015) en oversikt over ferdigheter som bygger på de allerede etablerte kunnskapene, subitizing, ANS, bearbeide verbale- og skriftlige representasjoner av tall, i tillegg til å forstå plassverdisystemet og etablere en mental tallinje. Hvorvidt kunnskapsutviklingen for mer komplekse ferdigheter følger Aunio og Räsänens (2015) modell er det vanskelig å si noe om, og siden denne studien i all hovedsak ser på barn opp til første klasse, ansees det ikke som relevant for denne studien å analysere dette nærmere.

2.1.3 Aritmetikk og tekstoppgaver

Matematiske oppgaver med aritmetisk innhold finnes i alle mulige hverdagslige situasjoner. Det er nettopp derfor forskning ofte fokuserer på dette området. For å kunne løse en tekstoppgave må barna gjenskape en representasjon av oppgaven, for så å trekke ut relevant informasjon som for eksempel riktig antall, rett regnemetode, for så å regne ut oppgaven på riktig måte (Gilmore, Göbel, & Inglis, 2018).

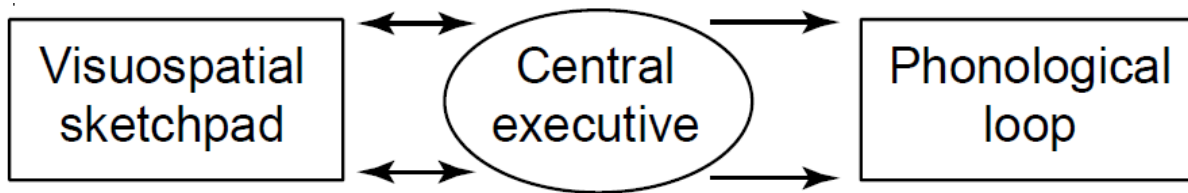
Aritmetikk i ordproblemer, eller tekstoppgaver, handler i stor grad om å mestre hverdagslige situasjoner, som for eksempel når man skal handle. Man vil for eksempel kunne ha bruk for å beregne rabatter, totalsummer samt hvor mye penger man får tilbake.

En viktig faktor for å kunne løse en slik oppgave vil avhenge av den semantiske situasjonen som er beskrevet i ordproblemet (Lindemann & Fischer, 2015). Konteksten og ordvalget i oppgaven vil påvirke grunnlaget for hvilken strategi som velges, i tillegg til å ville kunne påvirke oppgavens struktur og kompleksitet til den mentale representasjonen som trenger å bli konstruert for å løse oppgaven korrekt (Lindemann & Fischer, 2015). I noen oppgaver kan det for eksempel være fordelaktig å forstå konseptet *mer enn* eller *mindre enn*, noe som kan være vanskelig for noen barn. Dersom barnet ikke forstår dette konseptet vil en oppgave som dette være vanskelig å løse: «*Du ser fem kaniner og tre gulrøtter. Hvor mange flere kaniner er det enn det er gulrøtter?*». Dersom man omformulerer denne oppgavens spørsmål til: «*Hvor mange kaniner vil ikke få en gulrot*», vil barn i mye større grad svare korrekt (Hudson, 1983). Barna vil med andre ord også ha bruk for et vokabular som hjelper dem med å forstå aritmetiske tekstoppgaver.

2.2 Arbeidsminne

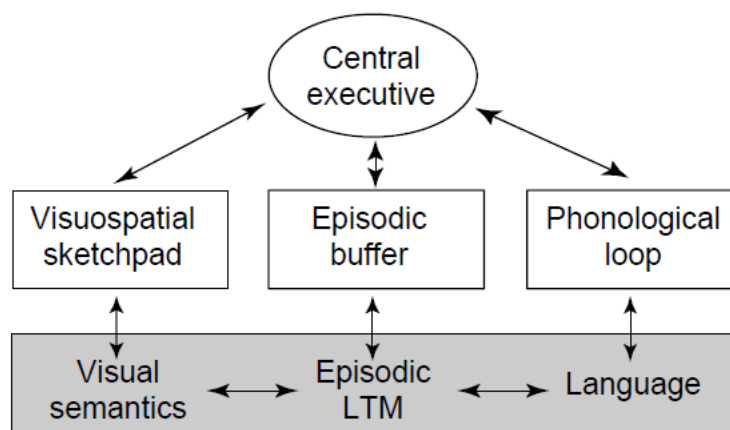
Begrepet arbeidsminne blir brukt på ulike måter og i ulike arenaer innenfor kognitiv forskning. I kognitive psykologi og områder som for eksempel ser på bearbeiding av matematiske oppgaver, viser arbeidsminnet til et domene-spesifikt system med begrenset kapasitet som er ansvarlig for samtidsprosessering og lagring av informasjon under en kognitiv oppgave (Baddeley, 2000; Hornung, Schiltz, Brunner, & Martin, 2014; Kadosh & Dowker, 2016). Videre finnes det også domene-generelle ferdigheter, som omhandler generelle kognitive ferdigheter som for eksempel eksekutive funksjoner (Cragg, Keeble, Richardson, Roome, & Gilmore, 2017).

Baddeley og Hitch presenterte i 1974 en modell av arbeidsminnet (se Figur 2). Denne modellen foreslo at arbeidsminnet består av tre komponenter. Da begrepsbruken ikke er helt konsekvent på hvilke norske ord som skal benyttes for arbeidsminnets delkomponenter, velger jeg å bruke følgende; *visuospatialt arbeidsminne* (også kalt *visuospatial sketchpad* eller *visuospatial skisseblokk*), *Central executive* og *verbale arbeidsminne* (også kalt *fonologisk løkke*, *phonological loop*) (Baddeley, 2000).



Figur 2; Modell av arbeidsminnet etter Baddley og Hitch fra 1974 (Baddeley, 2000)

I senere tid har Baddeley (2000) lagt til komponenten *episodisk buffer* i modellen. Den delen av arbeidsminnet kombinerer informasjon fra de andre delkomponentene. Den muliggjør for eksempel interaksjon mellom det verbale arbeidsminne og langtidshukommelsen, for å skape enn enhetlig episodisk representasjon (Baddeley, 2000). Denne nye modellen (se Figur 3) har et større fokus på hvordan delprosessene samarbeider for å sammenkoble informasjon, i motsetning til modellen fra 1974, som fremstiller delprosessene som adskilte prosesser. Den nye modellen gir dermed, ifølge Baddeley (2000), et bedre grunnlag for å belyse de komplekse aspektene ved arbeidsminnet (Baddeley, 2000).



Figur 3; Ny modell av arbeidsminnet med episodisk buffer etter Baddley og Hitch (Baddeley, 2000)

Bevisst oppmerksomhet antas å være den viktigste faktoren for henting fra bufferen. For eksempel ved å bevisst rette sin oppmerksomhet mot en oppgave hvor det visuospatiale arbeidsminne er involvert, som for eksempel under lagring og manipulering av visuell informasjon (Baddeley, 2000). Den reviderte modellen skiller seg fra den gamle hovedsakelig ved å fokusere på prosessene for å integrere informasjon, snarere enn fokusere på isolering av

delsystemene. Man antar som sagt at den episodiske bufferen er et slags koblingspunkt, hvor informasjon fra delsystemene kombineres. Ved å gjøre dette gir den et bedre grunnlag for å takle mer komplekse aspekter ved utøvende kontroll i arbeidsminnet.

Med utgangspunktet i dette vil man altså kunne anta, at arbeidsminne er delaktig når en matematisk oppgave skal utføres, siden det å løse en matematisk oppgave ofte innebærer å måtte bearbeide og lagre informasjon samtidig.

Man skiller som sagt ofte mellom verbalt arbeidsminne og visuospatialt arbeidsminne (Gilmore, Göbel, & Inglis, 2018). Det verbale arbeidsminnet omhandler lagring og manipulering av verbal informasjon, mens det visuospatiale arbeidsminne omhandler lagring og manipulering av visuell og romlig oppfatninger. Begge disse systemene er styrt av et kjernesystem, den *Central executive*, som er ansvarlig for å tildele oppmerksomhet til de to delene av arbeidsminne. Dersom man har en redusert kapasitet i hver av disse, vil det føre til vansker i matematikk (Gilmore, Göbel, & Inglis, 2018). Svekket kapasitet i det verbale arbeidsminne kan for eksempel føre til vansker med å samtidig holde på aritmetisk informasjon om oppgaven og mellomregningene i hukommelsen, noe som gjøre det vanske å fremkalle aritmetiske fakta.

Mange studier viser at arbeidsminne er en sterk prediktor for aritmetiske ferdigheter. Denne prediksjonen inkluderer da både utregning av verbale- og skriftlige oppgaver (Gilmore, Göbel, & Inglis, 2018). Man kan derfor også se at jo større kapasitet en persons arbeidsminne har, jo bedre viser denne personens matematikkferdigheter seg å være (Gilmore, Göbel, & Inglis, 2018). Baddeleys (2000) modell har blitt undersøkt i flere studier i senere tid (Dehn, 2008) og mange studier støtter denne modellen empirisk.

Central executive

Den Central executive er per i dag en mindre forstått komponent i arbeidsminne (Dehn, 2008). Det ser ut til at området hvor frontallappen ligger, er sterkt involvert i prosesser hvor det antas at den Central executive er delaktig (Baddeley, 2000). Men den ansees til å være kjernen i arbeidsminnet (Dehn, 2008). Mangelen på konkret kunnskap om den Central executive skyldes at konseptet er vanskelig å måle grunnet i dens flersidige funksjonalitet.

Allikevel er det enighet om at individuelle forskjeller i arbeidsminnet hovedsakelig er bestemt av prosesser i den Central executive.

Den Central executive antas å kontrollere den episodiske bufferen, siden den er ansvarlig for å kombinere og koordinere informasjonsflyten mellom den episodiske bufferen, det verbale-, og det visuospatiale arbeidsminne til sammenhengende sekvenser og man antar at slike sekvenser kan gjenskapes bevisst (Baddeley, 2000; Karbach & Kray, 2016). Generelt antar man at den Central executive er involvert når enn et individ skal lagre og bearbeide informasjon samtidig (Dehn, 2008).

Videre kontrollerer dette systemet flyten av informasjon gjennom arbeidsminnet, ved å til enhver tid være involvert når informasjon blir manipulert, som for eksempel ved mental aritmetikk (hoderegning) (Dehn, 2008). Den Central executive ansees også til å være *domene-fri*, da den fungerer som en link mellom delsystemene som er avhengig av verbal eller visuell prosessering. Man antar også at dette systemet ikke har lagringskapasitet i seg selv, men bruker den generelle kapasiteten til arbeidsminnet.

Det verbale arbeidsminne (phonological loop)

Det verbale arbeidsminne ble i utgangspunktet kalt en *verbal løkke*, som henviste til et språkbasert lager, med begrenset kapasitet, som lagrer talebasert informasjon (Dehn, 2008).

I senere tid har denne delen av arbeidsminnet fått ulike navn og har blitt beskrevet som å være ansvarlig for å holde på verbal og akustisk informasjon, ved å benytte seg av midlertidig lagring og et artikulatortisk øvingssystem (Baddeley, 2000; Göbel, 2016).

Man antar at det verbale arbeidsminne inneholder et midlertidig fonologisk lager, noe som gjør den spesielt god til å oppbevare sekvenser av informasjon (Baddeley, 2000). Man kan derfor benytte seg av minnespennsoppgaver, da funksjonsevne til det verbale arbeidsminne gjenspeiles tydelig i slike oppgaver hvor man for eksempel skal gjenta en tallsekvens enten forlengs eller baklengs, direkte etter at den har blitt presentert for den som utfører oppgaven (Baddeley, 2000; Göbel, 2016). Studier viser også til at det verbale arbeidsminne er et robust system som i stor grad er predikerende for kognitive funksjoner og akademisk læring (Dehn, 2008).

Det visuospatiale arbeidsminne (visuospatial sketchpad)

Det antas at denne delen av arbeidsminne holder på visuell (visuo-) og romlig (spatial) informasjon og ligger hovedsakelig i høyre hjernehalvdel (Baddeley, 2000). Det visuelt-romlige arbeidsminnet kan deles inn flere subkomponenter, visuell-, romlig- og trolig også kinestetiske komponenter, som hver for seg refererer til ulike oppfatninger av visuelle og romlige forhold mellom objekter og personer (Baddeley, 2000). Denne delen av arbeidsminnet har altså til oppgave å lagre og å manipulere visuell og romlig informasjon (Gilmore, Göbel, & Inglis, 2018). Slik informasjon kan blant annet være minner for plasseringen til ulike objekter (Dehn, 2008). Det visuospatiale arbeidsminne ser også ut til å spille en viktig rolle under lesing, da man benytter seg av visuell avkoding når man skal lese skrevne bokstaver eller tall samtidig, hvor man beholder en slags visuell ramme som bidrar til at man klarer å holde blikket på et bestemt sted når man leser en tekst. Forskingen er imidlertid ikke helt entydig på om hvor sterk denne delen av arbeidsminne henger sammen med aritmetiske ferdigheter. Dette vil bli nærmere belyst i avsnitt 2.2.2 Arbeidsminnets betydning for aritmetikk.

Episodisk buffer

Den episodiske bufferen er et midlertidig lagringssystem, med begrenset kapasitet (Baddeley, 2000). Denne bufferen er i stand til å koble sammen informasjon fra ulike kilder, som det verbale arbeidsminne og det visuospatiale arbeidsminne. Man antar som sagt at den episodiske bufferen blir styrt av den Central executive ved at den Central executive henter informasjon fra minnet ved å bevisst reflektere rundt den gitte informasjonen, og om nødvendig, manipulere den (Baddeley, 2000). Denne komponenten har blitt lagt til i arbeidsminnemodellen i senere tid, for å forklare innvirkningen langtidshukommelsen har på arbeidsminne (Dehn, 2008).

Den har videre tilgang til episodisk og semantisk langtidshukommelse for å konstruere representasjoner basert på ny informasjon (Dehn, 2008). Noen forskere hevder at den episodiske bufferen bare er en refleksjon av representasjonene som finnes i langtidshukommelsen. Men siden denne komponenten ser ut til å være viktig for læring ved å integrere informasjon fra de ulike delkomponentene i arbeidsminnet og langtidshukommelsen til enhetlige representasjoner, har den episodiske bufferen blitt beholdt som en delkomponent ved arbeidsminnet.

2.2.1 Utviklingen av arbeidsminne

Arbeidsminne blir blant noen forskere betegnet som «[...] the seat of thought and intelligence.» (Camos & Barrouillet, 2018). Dette fordi de anser arbeidsminne som et kognitivt system som konstruerer representasjoner som oppstår i vårt minne, og med den raske manipuleringen eller substitusjonen, lager arbeidsminnet en slags mental strøm som vi oppfatter som en tanke. Man kan derfor benytte prosesseringshastighet for å registrere tiden det tar for å konstruere en representasjon i arbeidsminnet ved å måtte tilbakekalle relevant informasjon fra langtidsmminnet, bruke denne informasjonen og kombinere den med presentert input, for eksempel en aritmetikkoppgave, på en meningsfull måte. Det finnes noe evidens for at prosesseringshastighet øker gjennom barndommen, og at dette henger sammen med den kognitive utviklingen som arbeidsminne er en del av (Camos & Barrouillet, 2018). Men selv om arbeidsminnets kapasitet øker gjennom barndommen, vil den varierer fra hvert individ (Gilmore, Göbel, & Inglis, 2018).

Enkelte forskere mener derfor at prosesseringshastighet henger sammen med en økning i arbeidsminnet, dette grunnet i at eldre barn oppnår en høyere prosesseringshastighet enn yngre barn (Camos & Barrouillet, 2018). Eldre barn har en fordel da de har kortere intervaller mellom for når de mellomlagrer informasjon i arbeidsminnet, før denne informasjonen blir gjenkalt. Denne økningen i prosesseringshastighet henger også sammen med evnen til å tilegne seg strategier for å holde på informasjon. For eksempel antar man at verbale elementer blir opprettholdt gjennom å øve på en artikulatorkisk måte, for eksempel ved å gjentatte ganger si et nytt ord som du skal lære deg. Ved å være i stand til å øve fortere, har eldre barn større mulighet til å gjenkalle flere elementer på begrenset tid, enn det yngre barn har (Camos & Barrouillet, 2018).

Arbeidsminnet ser ut til å være i stand til å lagre og å hente informasjon fra langtidsmminnet, noe som kan tyde på at arbeidsminnet også er involvert i å hente informasjon som fakta om tall (Cragg, Keeble, Richardson, Roome, & Gilmore, 2017). Eksempelvis kan hurtigere gjenkalling av informasjon fra langtidshukommelsen bidra til å blant annet utdype eller kategorisere allerede kjente kunnskap (Camos & Barrouillet, 2018). I tillegg øker bevisstheten, om at strategier er nyttige gjennom barndommen. Jo bedre implementert strategien er, jo flinkere blir man til å bruke den.

Andre studier viser til en aldersrelatert endring i hvor fort informasjon forsvinner fra arbeidsminnet under dual-task oppgaver, mens andre studier ikke finner aldersrelaterte endringer i arbeidsminnet hos barn (Camos & Barrouillet, 2018). Videre ser det ut til at minnet for visuelt presenterte stimuli, for barn i barnehagealder, hovedsakelig avhenger av det visuospatiale arbeidsminne. En annen antakelse er at verbal presenterte oppgaver har direkte tilgang til det verbale minnet, mens visuell presenterte oppgaver får tilgang til minnet indirekte ved å først kode om visuell input til verbal-representasjon.

Det finnes i dag lite forskning på dette feltet som kan gi entydige svar på hvordan arbeidsminnet utvikles, da det som tidligere sagt er vanskelig å kartlegge kognitive prosesser grunnet i at dette er et relativt vagt og uklart konsept (Camos & Barrouillet, 2018).

2.2.2 Arbeidsminnets betydning for aritmetikk

Studier viser at arbeidsminnet er delaktig i å opprettholde prosedyrer som å dele opp mengder og telling (Hubber, Gilmore, & Cragg, 2014). Trolig fordi arbeidsminnet holder på informasjon om oppgaven, lagrer midlertidige svar som for eksempel mellomregninger, og holder kontroll på tellinger.

Flere studier som har observert barn med vansker i matematikk, viser at disse barna også har mindre kapasitet i arbeidsminnet, noe som betyr at de har vansker med å fremkalle informasjon om grunnleggende aritmetiske fakta fra langtidshukommelsen og presterer dårligere på arbeidsminnetester enn jevngamle barn som ligger i normalområdet (Camos & Barrouillet, 2018; Geary, Hoard, Byrd-Craven, & DeSoto, 2004).

Dette kan tyde på at lav kapasitet i arbeidsminnet kan være relatert til at disse barna blant annet behøver mer tid til å telle. Denne saktere telleprosessen resulterer i et svakere arbeidsminne, noe som fører til svakere assosiasjoner i langtidshukommelsen, som gjør det vanskelig for langtidshukommelsen å støtte direkte gjenhenting av informasjon (Camos & Barrouillet, 2018).

Det finnes videre evidens på at det visuospatiale arbeidsminne støtter aritmetiske prosesser (Gilmore, Göbel, & Inglis, 2018). Også annen forskning peker på at aritmetikk støttes av både domene-spesifikke i tillegg til domene-generelle ferdigheter (Cragg, Keeble, Richardson, Roome, & Gilmore, 2017). Det foreslås at det finnes en domene-spesifikk relasjon mellom

arbeidsminne og aritmetiske ferdigheter, hvor det verbale arbeidsminne, med evnen til lagring og manipulering av verbal informasjon, spiller en større rolle enn det visuospatiale arbeidsminne (Cragg, et.al, 2017; Gilmore, Göbel, & Inglis, 2018). Samtidig mener andre forskere at det motsatte er tilfelle ved at det er en sterkere relasjon mellom matematikk og det visuospatiale arbeidsminne enn det verbale arbeidsminne.

Man antar videre at den aritmetiske ferdigheten *subitizing* gjenspeiler en persons visuell-romlige evner, som er en del av arbeidsminnet. Disse evnene gjør at man kan bearbeide små mengder av visuelle elementer samtidig (Piazza, Fumarola, Chinello, & Melcher, 2011).

Videre antar man at den Central executive bidrar under bearbeidingen av komplekse aritmetiske oppgaver (Gilmore, Göbel, & Inglis, 2018). Eksempelvis når man skal holde på mellomregninger i arbeidsminnet mens man foretar en annen utregning, en egenskap man har behov for i regnestykker som krever å låne (ved subtraksjonsoppgaver med tier-overganger) eller bruken av minnetall (ved addisjonsoppgaver med tier-overganger).

Det verbale arbeidsminne ser ut til å være involvert under bearbeiding av mentale aritmetiske oppgaver, som hoderegning, ved å holde og manipulere informasjon (Göbel, 2016)

Siden arbeidsminnet ser ut til å spille en så stor rolle for aritmetikk, har noen forskere foreslått at man kan forbedre aritmetiske ferdigheter ved å trene arbeidsminnet og dermed øke arbeidsminnets kapasitet (Melby-Lervåg & Hulme, 2013). Her er ikke forskningen helt entydig. For selv om noe forskning tyder på at øving kan føre til forbedringer i arbeidsminnet, vet man ikke hvor vidt denne forbedringen i arbeidsminnet også vil forbedre aritmetiske ferdigheter (Melby-Lervåg & Hulme, 2013).

Det ser ut til at arbeidsminnet er viktig for kognitive prosesser, som blant annet bidrar til utviklingen av aritmetiske ferdigheter. Arbeidsminnets kapasitet til å bearbeide og holde på informasjon blir derfor ansett som en sterk prediktor for aritmetiske ferdigheter (Gilmore, Göbel, & Inglis, 2018). Tendensene peker altså på, at større arbeidsminnekapasitet er relatert til bedre aritmetiske ferdigheter (Gilmore, Göbel, & Inglis, 2018).

2.2.3 Sammenhengen mellom aritmetikk og intelligens

Intelligens blir ansett som å være en underliggende faktor for svært mange ferdigheter, deriblant aritmetiske ferdigheter (Kyttälä & Lehto, 2008). Innenfor intelligens finnes det ulike kategorier. Man kan blant annet dele intelligens inn etter to faktorer som blir presentert av Cattell (1971) og Horn (1968), disse er flytende- og krystallisert-intelligens. Flytende-intelligens blir ansett som en evne for abstrakt resonnering, noe som er en forutsetning for å kunne løse aritmetiske oppgaver (Kyttälä & Lehto, 2008). Krystallisert-intelligens ser på tilegnet kunnskap i forhold til alder og utdanning. Videre er det noen studier som hevder at arbeidsminnets kapasitet og intelligens så å si er det samme (Kyttälä & Lehto, 2008). Andre studier har kritisert denne konklusjonen, dette grunnet i at selv om flertallet av studier som har undersøkt dette forholdet har funnet evidens på at det finnes en relasjon, ser ikke styrken på relasjonen ut til å være særlig sterk.

3 Metode

I denne delen vil det gis en redegjørelse av undersøkelsens forskningsmetodiske tilnærming. Det vil først bli gitt en presentasjon av studiens design og utvalg, for så å redegjøre for innsamling av data samt måleinstrument som ble tatt i bruk under datainnsamlingen. Det vil gis en beskrivelse av de statistiske prosedyrene som er blitt brukt under datamaterialets analyse. Videre vil det bli gjort rede for etiske hensyn som ligger til grunn for studien, for så å til slutt redegjøre for kartleggingstestenes reliabilitet. Validitet vil bli sett nærmere på i oppgavens diskusjonsdel.

Denne masterstudiens hovedintensjon er å undersøke hvorvidt et barns arbeidsminne ved femårsalder kan predikere matematikkferdigheter i første klasse. For å kunne gjøre en vurdering skal det brukes data som er samlet inn i forbindelse med NumLit prosjektet. Dataene vil bli fremvist ved å presentere graden av sammenheng som vises mellom utvalgte aritmetiske variabler og variabler for arbeidsminne.

Barna som deltar i studien er ikke påvirket av intervensjon før første måling, og har dermed fulgt en naturlig utvikling i hjemmet.

Siden studiens formål er å belyse et fenomen ved et gitt tidspunkt, for så å analysere dataene innenfor en bestemt variabel ved et senere tidspunkt, ansees den til å være en longitudinell studie. Videre har studien en kvantitativ metodisk tilnærming, da studien «baserer seg på store mengder data» som er innhentet ved bruk av psykometriske tester (Grue, 2015). Svarene ble registrert i et Excel-dokument for videre analyser i SPSS (IBM, 2019) og R-studio.

Datamaterialet vil ikke bli manipulert på noe vis, virkeligheten kommer altså til å bli beskrevet som den er. Studiens design vil dermed være ikke-eksperimentelt (Lund, 2002b).

De statistiske analysene vil bestå av en deskriptiv analyse, reliabilitet, en bivariat korrelasjonsanalyse i tillegg til en hierarkisk multipl regressjonsanalyse.

I den deskriptive analysen gis det beskrivelser av fordelingen til hver enkelt variabel. Den bivariate analysen vil vise i hvilken grad en variabel samvarierer med en annen variabel, og den hierarkisk multipl regressjonsanalyse benyttes for å påvise hvorvidt det er korrelasjon mellom variabler.

3.1 Kartlegging

Utvalget i denne studien er knyttet til NumLit prosjektet som gjennomføres ved instituttet for spesialpedagogikk ved Universitetet i Oslo. NumLit prosjektets utvalg består av 254 enspråklige barn som har norsk som morsmål, og som i utgangspunktet følger normal språkutvikling. Målet med prosjektet er å følge barna fra siste året i barnehagen (omtrent 5 år) opptil 18-årsalderen, for å kartlegge barnas utvikling i forhold til tallforståelse og leseferdigheter, hvordan disse henger sammen med kognitive ferdigheter, i tillegg til deres sosiokulturelle bakgrunn samt kjønn (NumLit, 2020).

Barna denne studien vil ta utgangspunkt i er blitt kartlagt ved omtrent 5 år (siste året i barnehage) og omtrent 6 år (første trinn på barneskolen).

Dataene ble samlet inn av testledere som fikk en grundig opplæring i hvordan kartleggingen skal foregå, og hvordan de ulike kartleggingstestene skal gjennomføres. Barna ble belønnet muntlig av testleder og med klistremerker, for å holde motivasjonen oppe.

Kartleggingsbatteriet som blir benyttet i denne undersøkelsen består av 8 deltester. Nedenfor blir det gitt en oversikt over kartleggingsinstrumentene som blir brukt i denne undersøkelsen.

Tabell 1. Oversikt over variabler og instrumenter som brukes.

Variabler	Deltest	Test
Arbeidsminne	Backward digit recall	WMTB-C
	Maze Memory	WMTB-C
	Listening recall	WMTB-C
Aritmetikk	TOBANS addisjon	TOBANS
	TOBANS subtraksjon	TOBANS
	Hoderegning addisjon	NumLit
	Hoderegning subtraksjon	NumLit
	Ordproblemer	WISC-IV
Nonverbal intelligens	Raven CPM	Raven CPM

3.1.1 Kartlegging av arbeidsminne

I NumLit prosjektet kartlegges arbeidsminne ved hjelp av deltestene *Listening recall*, *Backward digit span*, og *Maze memory* fra testbatteriet *WMTB-C* (Gathercole & Pickering, 2001). Denne studien tar utgangspunkt i testresultatene fra *Backward digit span* og *Maze memory* som ble målt i barnas siste år i barnehage. Resultatene fra *Listening recall* vil bli hentet fra første klasse, da den så ut til å være for vanskelig for barna i barnehagen.

Både *Listening recall* og *Backward Digit span* er tester som ser på den den Central executive delen av arbeidsminne, og baserer seg på Baddeley og Hitch sin modell av arbeidsminne (Pickering, 2006). *Maze memory* måler den visuospatiale delen av arbeidsminne.

Backward Digit span, WMTB-C

Oppgavene er delt inn i blokker, hvor hver blokk inneholder seks oppgaver med samme vanskelighetsgrad. Stoppekriteriet for denne testen er satt til å avslutte dersom barnet mislykkes innenfor et nivå, ved å få tre eller flere feil i den samme blokken.

I disse oppgavene skal barnet huske og gjengi en rekke med tall som leses opp av testlederen. Tallrekken skal gjengis i omvendt rekkefølge. For at barnet lett skal forstå prinsippet i oppgavene, blir det gitt øvingsoppgaver. Underveis i testen blir det gitt poeng for hvert rette svar, dersom barnet får fire rette i en blokk, går man automatisk videre til å administrere neste blokk.

I den første blokken får barnet presentert to tall som skal gjengis i omvendt rekkefølge. Det legges til ett tall for hver nye blokk.

Et eksempel på en oppgave kan se slik ut:

Testleder: «2,3».

Barnet: «3,2».

Listening recall, WMTB-C

Disse oppgavene er også delt inn i blokker, hvor hver blokk består av seks utsagn med lik vanskelighetsgrad. Vanskelighetsgraden økes for hver blokk som blir gjennomført, ved at det legges til et utsagn. Testingen blir avbrutt dersom barnet har svart feil tre ganger i samme blokk.

Før selve kartleggingen får barnet presentert noen øvingsoppgaver, for å bli kjent med prinsippet i oppgavene. Under selve kartleggingen av Listening recall-testen leses det opp et utsagn, barnet må avgjøre om utsagnet er sant eller usant, for så å gjengi det siste ordet i utsagnet.

Et eksempel på en oppgave med en setning er:

Testleder: «Ananas spiller fotball.»

Barnet svarer: «Usann, fotball.»

Et eksempel på en oppgave med to setninger er:

Testleder: «Fisker har langt hår.»

Barnet svarer: «Usann.»

Testleder: «Bøker har sider.»

Barnet svarer: «Sann, hår, sider.»

Maze Memory

I denne oppgaven får barnet presentert todimensjonale labyrinter med gitte utveisruter. En og en utveisrute vises for barnet ved at testlederen presenterer den med fingeren. Utveisruten er en rød linje som går innenfra og ut av labyrinten. Etter at barnet har sett den enkelte utgangsruten i omtrent 3 sekunder, blir barnet bedt om å huske utveisruten, for så å tegne ruten med blyant i et eget, tilsvarende, labyrint. Oppgavene øker i vanskelighetsgrad ved at labyrintene blir større og mer komplekse. Oppgavene er også her delt inn i blokker, med 6 labyrinter per blokk. Veggene i labyrinten har to åpninger, og for hver ny blokk legges det til en ny vegg. Også her blir det gitt øvingsoppgaver. Dersom barnet har fire rett i en blokk, hopper man automatiske videre til neste blokk, men testen stoppes dersom barnet har 3 eller flere feil i samme blokk.

3.1.2 Kartlegging av aritmetiske ferdigheter

Aritmetiske ferdigheter ble målt ved hjelp av *Ordproblemer (Math Word Problem)* (Wechsler, 2003), *TOBANS addisjon og subtraksjon* (Hulme, Brigstocke, & Moll, 2016), i tillegg til *hoderegning addisjon og subtraksjon* (NumLit, 2020). Dataene for aritmetiske ferdigheter som ligger til grunn i denne studien ble målt da barna gikk i første klasse.

Ordproblemer (Math Word Problem)

Aritmetikk ble målt med deltesten Math Word Problem fra WISC-IV, som er et kartleggingsverktøy som måler kognitive evner (Wechsler, 2003). Deltesten består av regneoppgaver som er kontekstbaserte, det innebærer at barnet får presentert regneproblemet muntlig av testlederen i form av en regnefortelling, og barnet må regne ut svaret mentalt. Testen er utarbeidet for å måle både aritmetiske ferdigheter og arbeidsminne, men også konsentrasjon og oppmerksomhet. Videre består testen av 34 enheter og er standardisert for barn fra 6-16 år. Siden noen barn bare var 5 år ved testtidspunktet, la NumLit (2020) til enklere oppgaver i starten av denne testen, for å tilpasse innholdet til alderen.

Vanskelighetsgraden er økende, det ble gitt ett poeng for korrekte svar og testen avsluttes etter fire påfølgende oppgaver med null poeng. I testsituasjonen får barnet øvingsoppgaver. Under selve testingen får barnet 30 sekunder på å besvare oppgaven muntlig etter at testleder har presentert ordproblemet. Grunnet stopperegelen og at barna i denne studien går i 1.klasse, er det kun addisjon og subtraksjon som blir testet.

Et eksempel på dette er:

Testleder: Tore hadde en bil og så fikk han en til av farfar. Hvor mange biler hadde han da?

Barnet: To.

TOBANS addisjon og subtraksjon

Aritmetiske ferdigheter i første klasse ble også målt med *Test of Basic Arithmetic and Numeracy Skills*, TOBANS. Testen er delt inn i to deler, en addisjonsdel og en subtraksjonsdel. Testen måler både regneflyt og nøyaktighet. Siden testen er beregnet for barn mellom syv til elleve år, testes barna i denne studien bare i addisjon og subtraksjon, hvor hver av oppgavearkene består av 60 oppgaver, med tall fra 1 til 9.

Barnet får beskjed om at det nå skal gjennomføres noen enkle oppgaver med tall med tidsbegrensning. Barnet får presentert noen øvingsoppgaver før selve testen starter. Barnet får 60 sekunder for å gjennomføre så mange oppgaver i en deltest som mulig. Barnet får beskjed om å prøve så godt han/hun klarer, og om man skulle møte på en oppgave man ikke får til, kan man hoppe over den og fortsette på neste oppgave. Det blir gitt poeng for hvert riktige svar.

Hoderegning addisjon og subtraksjon

I denne deloppgaven fikk barnet lest opp addisjons- og subtraksjonsoppgaver. Den første delen består av 22 addisjonsoppgaver og den andre delen av 14 subtraksjonsoppgaver.

Det gis 1 poeng for hvert korrekte svar, testingen stoppes etter 4 gale svar på rad.

Et eksempel for addisjon:

Testleder: «Hva blir 1 hvis du legger til 1, altså 1 pluss 1?»

Barnet: «2»

Et eksempel for subtraksjon:

Testleder: «Hva blir 2 hvis du tar bort 1, 2 minus 1?»

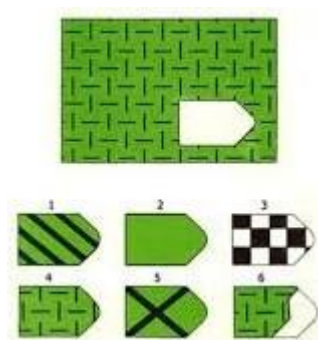
Barnet: «1»

3.1.3 Kartlegging av variabelen nonverbal intelligens

For å måle nonverbal intelligens i barnehage ble det benyttet *The Raven's Coloured Progressive Matrices* (Helland-Riise & Martinussen, 2017).

The Raven's Coloured Progressive Matrices (CPM)

Denne testen er laget for å kartlegge barnas generelle non-verbale kognitive evner (Helland-Riise & Martinussen, 2017). Barna får presentert et bilde hvor det har blitt «klippet» ut en bit. Barnet skal forsøke å skape et meningsfullt bilde ved å velge det riktige bildet av seks presenterte alternativer for å danne et fullstendig bilde. Siden testen tester nonverbal intelligens, er det ingen krav om språkkunnskaper. Testen ble administrert på laptop, den har ingen tidsbegrensning, men stoppet automatisk etter et visst antall feil.



Figur 8. Eksempeloppgave fra: The Raven's Coloured Progressive Matrices

3.2 Validitet

I en kvantitativ studie forsøker man å besvare forskningsspørsmål ved å trekke slutninger. For at slutningene skal være holdbare (valide) bør de basere seg på ulike faktorer som inngår i en forskningsprosess (Lund, 2002a). Dette kan gjøres ved å belyse metodevalgene og om de bidrar til å trekke valide slutninger for å besvare forskningsspørsmålene. I en kvantitativ studie som denne, kan validitet bli definert som i hvilken grad et konsept (variabel) er nøyaktig målt (Heale & Twycross, 2015). Validitet forteller altså noe om hvor gode målene som man benytter for å trekke slutninger i en studie er. En variabel ansees som valid dersom den klarer å måle det den er ment til å måle (de Vaus, 2014; Field, 2018), og refererer dermed til den best mulige tilnærmingen til sannheten om påstander, inkludert påstander om årsaker (Cook & Campbell, 1979). Det finnes imidlertid ulike delkategorier innenfor validitet, som gir ulike måter å undersøke validitet på.

En inndeling finner man hos Field (2018). Denne inndelingen beskriver tre grunnleggende måter å vurdere validitet på; *Criterion validity*, *Content validity* og *Ecological validity*. Dersom man oversetter begrepene, blir de oversatt til henholdsvis *kriterievaliditet*, *innholdsvaliditet* og *økologisk-validitet*. Man finner noen av de samme delkomponentene hos Heale og Twycross (2015). De deler validitet inn i *Criterion validity*, *Content validity* og *Construct validity*. Den sist nevnte validitetstypen kan oversettes til begrepsvaliditet.

Kriterievaliditet hos Field (2018) ser på hvorvidt man kan slå fast at et måleinstrument måler det det faktisk skal måle ved å sammenlikne måleinstrumentet ved hjelp av objektive kriterier. Dette gjøres ved å relatere dataene fra en test til observasjoner fra 'den virkelige verden', altså samfunnet som det er. Man ønsker altså å belyse hvorvidt resultatene i en studie kan overføres til andre individer i samfunnet. Field (2018) deler kriterievaliditet inn i to underkategorier, *predictive validity* (predikerende-validitet) og *concurrent validity* (samtidig-validitet). Den førstnevnte benytter man dersom man skal benytte data til å predikere observasjoner ved et senere tidspunkt (Field, 2018), dette er noe som undersøkes i denne studien ved å se på om en ferdighet i 5 år alder kan predikere en annen ferdighet i førsteklasse. Og *concurrent validity* benyttes dersom man skal beskrive data som er registrert ved samme tidspunkt, som for eksempel ved bruk av et allerede eksisterende måleinstrument i tillegg til et nytt.

Heale og Twycross (2015) beskriver kriterievaliditet ved å gi tre ulike kriterier som skal tilfredstilles for å ha god validitet, disse er; *Convergent validity* (konvergens validitet),

Divergent validity (divergent validitet) og *Predictive validity* (predikerende validitet).

Konvergens validitet skal vise at et måleinstrument er sterkt korrelert til andre instrumenter som måler den samme variabelen, mens divergent validitet skal vise at et måleinstrument er svakt korrelert med måleinstrumenter som måler andre variabler. Predikerende validitet krever at det bør være en høy korrelasjon med fremtidige kriterier, som for eksempel at en skåre på en kognitiv evnetest i ung alder kan predikere matematiske ferdigheter senere i barndommen (Hornung et al., 2014). Slike korrelasjoner kan man eksempelvis undersøke ved å analysere korrelasjonene mellom variablene i en bivariat korrelasjonsanalyse, som er en del av den statistiske analysen i denne studien.

Innholdsvaliditet defineres hos Field (2018) og Heale og Twycross (2015) som å være evidens på at innholdet i en test stemmer overens med det testen er designet for å dekke, altså dersom en test er ment til å måle arbeidsminne, også måler dette. Videre introduseres det også økologisk-validitet (Field, 2018). Denne validitetstypen ser på bevis for at resultatene fra en studie faktisk kan benyttes til å trekke konklusjoner som har overføringsverdi til samfunnet, for denne studien vil det være å finne bevis for at resultatene også kan gjelde for andre i samme aldergruppe som denne studien tar utgangspunkt i.

I tillegg nevner Heale og Twycross (2015) begrepsvaliditet, som referer til hvorvidt man kan trekke en konklusjon om at testresultatene er relaterte til konseptet (variabelen) som blir målt. For å finne bevis for begrepsvaliditet, kan man benytte tre argumenter for validitet; *Homogeneity, Convergence, Theory evidence*. Begrepene kan oversettes til henholdsvis *homogenitet, konvergens* og *teoretiske bevis*. Homogenitet ser på om måleinstrumentet måler det den er ment til å måle, som står i tråd med Field (2018) sin definisjon av innholdsvaliditet. I likhet med konvergens validitet innen kriterievaliditet, oppstår også her konvergens når flere måleinstrument måler et felles konsept, eller lik variabel. Dersom det ikke er like måleinstrument tilgjengelig, vil det ikke være mulig å kommentere dette punktet. Teoretisk bevis refererer til dersom resultatene man får i en studie, stemmer overens med det man finner i forskningsteorier for den gitte variabelen, da er dette et argument for god begrepsvaliditet. I denne studien vil resultatene derfor bli sett i lys av de teoretiske tilnærmingene som gir tidligere kapitler.

I denne studien blir i all hovedsak Cook og Campells (1979) validitetssystem benyttet. Systemet presenterer fire typer av validitet for å belyse gyldighet i studier. De fire typene er; *indre-, ytre-, statistisk- og begrepsvaliditet* (Lund, 2002b). De ulike validitetstypene påvirker

hverandre. For eksempel kan valg av metode føre til at en validitetstype styrkes, samtidig som en annen en kan svekkes. Det vil også bli gitt en vurdering av hvilke faktorer som kan true validiteten. De påfølgende delene i dette kapittelet vil gi en beskrivelse av delkomponentene i tillegg til å ta utgangspunkt i Cook og Campells validitetssystem, da dette systemet ofte benyttes i kausal forskning (Cook & Campbell, 1979).

3.2.1 Statistisk validitet

Statistisk validitet referer til konklusjoner som trekkes ut ifra om det er rimelig å anta at det eksisterer en korrelasjon mellom variabler gitt et spesifisert signifikans-nivå (α -nivå) og variablers målte varians (Cook & Campbell, 1979). For å måle varians tar statistisk validitet utgangspunkt i t-verdier for uavhengige (forklaringsvariabel) og avhengige variabler (responsvariabel), og ser på om sammenhengen mellom disse to variablene er statistisk signifikant (Lund, 2002b). I denne studien er arbeidsminne den uavhengige variabelen og aritmetiske ferdigheter er den avhengige variabelen.

Når man gjennomfører en statistisk analyse, kan det oppstå *trusler* mot den statistiske validiteten. Slike trusler ved signifikanstesting kan oppstå ved brudd på statistiske forutsetninger eller ved lav statistisk styrke, og kalles henholdsvis *Type1-feil* og *Type2-feil*. Type1-feil kan oppstå ved at man forkaster en sann null-hypotese, og Type2-feil oppstår ved å akseptere en gal nullhypotese (Lund, 2002b; Moore, McCabe, & Craig, 2017). Noe som minker sannsynligheten for Type2-feil er å ha et stort nok datautvalg. Et datautvalg som er stort nok gjør det mulig å kunne se variasjoner mellom variabler dersom de finnes (Diez, Barr, & Cetinkaya-Rundel, 2017). Siden statistisk styrke er sammensatt og avhenger av flere faktorer som utvalgsstørrelse, signifikansnivå, varians, effektstørrelse (differanse mellom gjennomsnitt på populasjonsnivå) og bruken av tohalet- eller enhalet test, vil det også være flere faktorer som potensielt kan redusere den statistiske styrken (Lund, 2002b). Målefeil ansees også som en trussel mot den statistiske styrken, men i tillegg kan slike feil også bidra til å redusere begrepsvaliditeten og/ eller den indre- og ytre validiteten. Jo flere målefeil, jo større usikkerhet blir det rundt slutninger og generaliseringer (Lund, 2002b). Målefeil kan oppstå for eksempel dersom man har ustandardiserte tester eller svært varierende testsituasjoner. Dersom den avhengige variabelen får en økning i variasjonen til målefeil, vil dette redusere variabelens reliabilitet.

3.2.2 Indre validitet

Indre validitet referer til en antakelse om gyldighet som vi slutter oss til, for eksempel en antakelse om relasjonen mellom to variabler kan vise til en årsakssammenheng eller at fraværet av en relasjon innebærer fravær av en årsakssammenheng (Cook & Campbell, 1979). Indre validitet dreier seg med andre ord om å belyse hvorvidt korrelasjoner mellom observerte variabler kan forklares (Lund, 2002b). Denne studien ønsker å trekke slutninger om variabelen arbeidsminne kan forklare variabelen aritmetiske ferdigheter. Mulige trusler mot indre validitet kan være faktorer som spiller inn på resultatene som man ikke har kontroll over. Det er derfor nødvendig å diskutere om truslene er reelle eller uaktuelle når man vurderer resultater.

3.2.3 Begrepsvaliditet

Det finnes ulike definisjoner av ordet *begrepsvaliditet*. De Vaus (2014) beskriver begrepsvaliditet (construct validity) som en måte å evaluere en test ut ifra hvor godt den henger sammen med teoretiske forventninger. Cook og Campbell (1979) beskriver begrepsvaliditet til å referere til en antakelse om validitet, som har en generaliserende overførbarhet til en større gruppe individer enn det som er inkludert i en studie (Cook & Campbell, 1979). Hos Lund (2002b) ser det ut til å omhandle begrepsoperasjonalisering i tillegg til å være ment til å belyse hvorvidt man med de operasjonaliserte variabelbegrepene måler det man faktisk skal måle og ikke noe annet (Lund, 2002b). Kleven (2002a) beskriver denne validitetstypen som hvorvidt det er samsvar mellom de teoretiske definisjonene til begrep og om man lykkes med å operasjonalisere disse. Det ser ut til at disse definisjonene refererer til å se på sammenhengen mellom en valgt test og teori som ligger til grunn som henholdsvis taler for eller mot at testen faktisk måler det den er ment til å måle. Selv om man, i dette fagfeltet, bruker noe ulike begrep til omtrent lik eller lik definisjon, mangler det enighet om hva som kan være en trussel mot begrepsvaliditet. Denne uenigheten i seg selv kan ansees som en trussel mot begrepsvaliditeten.

Kleven (2002a) deler begrepet inn i *innholdsvaliditet* og *kriterievaliditet*. Innholdsvaliditet blir ansett som høy dersom måleinstrumentene en studie benytter dekker et utvalg som er representativt for ferdighetene som skal måles. Det benyttes da ofte en skjønnsmessig vurdering for å evaluere dette.

Kriterievaliditet benyttes for å predikere prestasjoner (Kleven T. A., 2002b), og er dermed i tråd med Field (2018) sin underkategori *predictive validity* hvor man benytter datamålinger fra et tidspunkt til å predikere utfall ved et senere tidspunkt.

Trusler mot denne validitetstypen varierer mellom hva slags undersøkelser man gjennomfører, men *systematiske målefeil* blir ansett som en vanlig trussel (Kleven A. T., 2002a). Målefeil kan føre til skjevhet i målingene ved at testene enten måler mindre enn den det skal dekke, eller ved at testen måler mer enn det den er tenkt til.

3.2.4 Ytre validitet

Ytre validitet referer til en antakelse om validitet som tar utgangspunkt i en antatt årsakssammenheng, som er generaliserende på tvers av alternative måleinstrumenter for samme variabel og på tvers av ulike typer individer, omgivelser og tidspunkt (Cook & Campbell, 1979). Man oppnår altså god ytre validitet i en studie dersom den gjør det mulig å trekke generaliserende slutninger (Lund, 2002b). Generaliseringen skal være ikke-statistiske og de skal med andre ord kunne overføres til andre individer i samfunnet, ved ulike tidspunkt og situasjoner, med rimelig sikkerhet. Denne studien tar utgangspunkt i en antakelse om en årsakssammenheng mellom arbeidsminne ved omtrent 5 år og aritmetiske ferdigheter i førsteklasse.

Videre kan man skille mellom *over-generalisering* og *til-generalisering*. Over-generalisering omfatter hvor bredt noe kan generaliseres, mens til-generalisering dreier seg om slutninger som trekkes til en definert populasjonsgruppe, situasjon eller tid (Lund, 2002b). Det kan oppstå flere trusler mot ytre validitet. En er systematiske forhold som gjør det vanskelig å trekke generaliserende slutninger. En annen trussel baserer seg på gjensidig påvirkning eller relasjon mellom individer og uavhengige variabler, situasjonstyper og tider.

Det kan også være en trussel dersom utvalget som deltar i studien ikke er representativ for populasjonen ved at det er for ensartet (Lund, 2002b). Dersom utvalget baserer seg på en veldig spesifikk gruppe vil det være vanskelig å trekke slutninger som kan overføres til andre individgrupper. Den siste trusselen tar utgangspunkt i *ikke-representative individutvalg* (Lund, 2002b). Denne tar utgangspunkt i at man ikke kan generalisere resultater dersom utvalget man tar utgangspunkt i, ikke er representativt for populasjonen som forskningen er ment til å si noe om. For å sikre seg mot denne trusselen kan man trekke et tilfeldig utvalg fra populasjonen.

3.3 Etiske hensyn

Siden denne masterstudien har en tilkobling til en større studie er det allerede blitt søkt og fått godkjenning for gjennomføring ved *Norsk senter for forskningsdata* (NSD). Dette sentret forvalter og er retningslinje givende for behandling personopplysninger i forhold til taushetsplikt, etiske retningslinjer samt datainnsamling (NSD – Norsk senter for forskningsdata, 2018).

De ansvarlige for NumLit studien har derfor på forhånd informert om at deltakelsen er frivillig, og at det er mulighet for å trekke seg underveis, i den forbindelse har barnas foresatte signert på et samtykkeskjema. Datamaterialet blir anonymisert og slettet når studien er avsluttet. Testlederne har fått instruksjoner av forskergruppen på hvordan de skal forholde seg i møte med barna som deltar i prosjektet, da det i forskning som har barn som informantgruppe er særlig viktig å ta hensyn til at de er en sårbar gruppe (NESH, 2016). Barna blir i denne sammenhengen ansett som å være i en sårbar gruppe, da de formelt sett ikke får gitt fritt samtykke. Med fritt samtykke menes det, at samtykket er innhentet uten at vedkommende har opplevd en begrensning i sin egen handlingsfrihet eller ytre press (Dalen, 2011). Selv om foreldre har gitt samtykke om at barnet kan delta i studien, får også barnet informasjon om prosjektet det deltar i, på en alderstilpasset måte, i tillegg til at forskergruppen presiserte for testlederne at deltakelsen er frivillig for barna, og at de kan trekke seg når som helst (NESH, 2016). I den forbindelse er det også viktig å bemerke seg at barn er lojale ovenfor autoritære personer, altså voksne, noe som krever at testlederne skaper en trygg situasjon og relasjon til barnet, slik at barnet opplever at det faktisk kan trekke seg (NESH, 2016).

Videre er det viktig at barna som deltar i et forskningsprosjekt som dette, møter testledere med kunnskap om barns utviklingsforløp og hvilke behov et barn har i ulike faser (NESH, 2016). Det var derfor hensiktsmessig å gjennomføre testingen på et skjermet sted, slik at barnet kunne gjennomføre testene uforstyrret.

Videre må hver enkelt testleder gjøre underveisvurderinger og tilpasninger som for eksempel å bruke et språk som er forståelig for hvert enkelt barn. Det kan også være nødvendig med tilpasninger i form av for eksempel endringer i rekkefølgen på deltestene, eller antall pauser, for å belaste barnet minst mulig i testsituasjonen (de Vaus, 2014).

Videre må man sørge for gode anonymiseringsmetoder. Norge er et relativt lite land, noe som kan gjøre det lett å identifisere enkeltpersoner. Man kan tenke seg at dersom de som skal delta i et forskningsprosjekt tilhører en minoritet, som har noen spesielle særtrekk, vil det øke sannsynligheten for at disse individene vil bli gjenkjent (Dalen, 2011). Det er derfor særlig viktig at man som forsker er oppmerksom på å ikke formidle informasjon som kan skade individene som er del av forskningen (NESH, 2016). Forskningsmaterialet som ligger til grunn i denne studien stiller derfor høye krav til anonymitet og anonymisering. Videre skal de lagrede personopplysningene og det innhentede datamaterialet slettes når studien er gjennomført (NESH, 2016). NumLit prosjektet ved UiO har derfor ikke bare grundige rutiner for sikker oppbevaring og lagring, men også for sletting av personopplysninger.

4 Analyse

I denne delen av studien vil resultatene fra undersøkelsen bli presentert og analysert. Først vil det bli gitt en presentasjon av det endelige antallet elever som vil bli inkludert i analysen.

Derpå vil resultatene fra den deskriptive statistikken bli presentert. Den deskriptive statistikken bruker variablenes data til å gi informasjon om gjennomsnitt, standardavvik, skjevhet og kutrosis, i tillegg vil også reliabilitet bli kommentert. Det vil også bli gitt visuelle fremstillinger av variablenes normalfordelingskurve.

Videre vil det gis en beskrivelse av den bivarierte korrelasjonsanalysen og dens formål, for så å presentere resultater fra studiens hierarkiske multiple regresjonsanalyse.

Korrelasjonen måler retningen (retning gitt av beta-verdi forteller om korrelasjonen er positiv eller negativ) og styrken til en lineær relasjon mellom to kvantitative variabler (Moore, McCabe, & Craig, 2017). Denne korrelasjonen blir vanligvis uttrykket med r . Når man skal gjennomføre en korrelasjonsanalyse med *Pearsons* r som korrelasjons koeffisient, kreves det at dataene er tilnærmet normalfordelte. Normalfordelinger er symmetriske tetthetskurver som er klokkeformet (Løvås, 2013; Moore, McCabe & Craig, 2017), med gjennomsnittet (μ) i midten av fordelingen og dets standardavvik (σ) som sier noe om spredningen rundt gjennomsnittet. Videre viser en positiv r til en positiv forbindelse (positiv retning) mellom variablene, dersom r har en negativ verdi, viser den til en negativ forbindelse (negativ retning) mellom variablene. Korrelasjonen r er alltid et tall mellom -1 (negativ) og 1 (positiv) (Moore, McCabe, & Craig, 2017). Verdier som er nærme 0 indikerer en veldig svak lineær relasjon. Styrken på relasjonene øker når r beveger seg fra 0 og mot enten -1 eller 1. Verdier av r som er nærme -1 eller 1 viser til at datapunktene ligger nærme inntil en rett linje. Dersom man får ekstreme verdier som $r = -1$ eller $r = 1$, ligger datapunktene langs en eksakt linje.

Videre blir det benyttet kontrollvariabler for å styrke undersøkelsens validitet, dette gjøres for å utelukke at en eventuell sammenheng mellom arbeidsminne og aritmetikk ikke skyldes av en annen variabel (Lund, 2002b). Kontrollvariablene i denne undersøkelsen vil være den non-verbal testen for flytende- intelligens, med deltesten Raven- CPM, i tillegg til alder og kjønn. Flytende-intelligens blir inkludert da den har vist seg å være en sterk prediktor for aritmetiske ferdigheter (Kyttälä & Lehto, 2008).

4.1 Rensing av data før dataanalysen

Datasettet består av totalt 254 barn. Noen av disse barna mangler informasjon på enkelte deltester. For å forenkle analysen vil det i denne studien bli fjernet støy i form av manglende datapunkter. Disse manglende datapunktene omtales som non-available (NA) det betyr at noen barn mangler data på ulike deltester. Etter det satte kriteriet har derfor 31 elever blitt fjernet fra datasettet. Da står det igjen 223 elever som blir med i videre analyser.

4.2 Deskriptiv analyse

Den deskriptive analysen beskriver fordelingen til hver enkelt variabel. Dersom en variabel skulle avvike fra normalfordelingen (Gauss-kurve), vil denne analysen også vise dette. Det vil bli fremvist histogrammer med variablenes normalfordeling. Normalfordelingen blir gitt som en rød linje som baserer seg på variablenes fordeling i histogrammene. Histogrammene viser de utvalgte variablene langs x-aksen og langs y-aksen vises det *density*. Density er det engelske ordet for det som på norsk heter densitet, eller tetthet. Histogrammet viser dermed hvor mange prosent som har oppnådd en gitt poengsum på venstre siden. Ved siden av histogrammet er det gitt en skala i ulike gråtoner som går fra 0 – 50. De ulike grå-tonene finner man igjen i histogrammets stopper og representerer antall elever.

Dersom en fordeling ikke er normalfordelt, kan det skyldes *skjevhet* (skewness) eller *kurtosis*. Disse er statistiske mål som benyttes for å beskrive en fordeling dersom den er asymmetrisk eller spiss. Skjevhet beskriver dersom en fordeling er forskjøvet til høyre dersom halen til høyre har større verdier og er mye lengre enn halen til venstre som har lavere verdier og motsatt (Moore, McCabe, & Craig, 2017). Får man en positiv skjevhetsverdi, viser denne til en lengre hale med høye verdier i halen til høyre (høyreskjev fordeling), mens en negativ skjevhetsverdi viser til en lengre hale med lave verdier i halen til venstre (venstreskjev fordeling) (Diez, Barr, & Cetinkaya-Rundel, 2017).

Når absoluttverdien til skjevheten er <0.5 er fordelingen tilnærmet symmetrisk, er absoluttverdien til skjevheten mellom 0.5 og 1, er fordelingen svakt skjev, mens en fordeling med absoluttverdien >1 for skjevhet, viser en svært skjev fordeling (Løvås, 2013). En absoluttverdi er et tall uten fortegn og er alltid positiv, en verdi på -1 vil dermed ha absoluttverdien 1.

Verdien for kurtosis sier noe om toppen i fordelingen. Eksempelvis kan fordelingene være svært tynn, spiss og høy med korte haler, eller flat og bred med lange haler. Får man en positiv kurtosisverdi viser den til en spiss fordeling, kurtoseverdi =0 viser til en normalfordeling, en negativ kurtosisverdi derimot viser til en flatere fordeling (Løvås, 2013). Dersom fordelingen er asymmetrisk, kan dette skyldes at enkelte oppgavesett enten er for vanskelige eller for enkel for en større andel elever enn det som er forventet. Ved oppgaver som er for enkle vil et større antall elever få høyere verdier enn forventet, og man vil oppnå en takeffekt (Field, 2018). I tilfeller hvor oppgavene er for vanskelige, vil man få mange lave resultater, og man får en gulveffekt (Field, 2018).

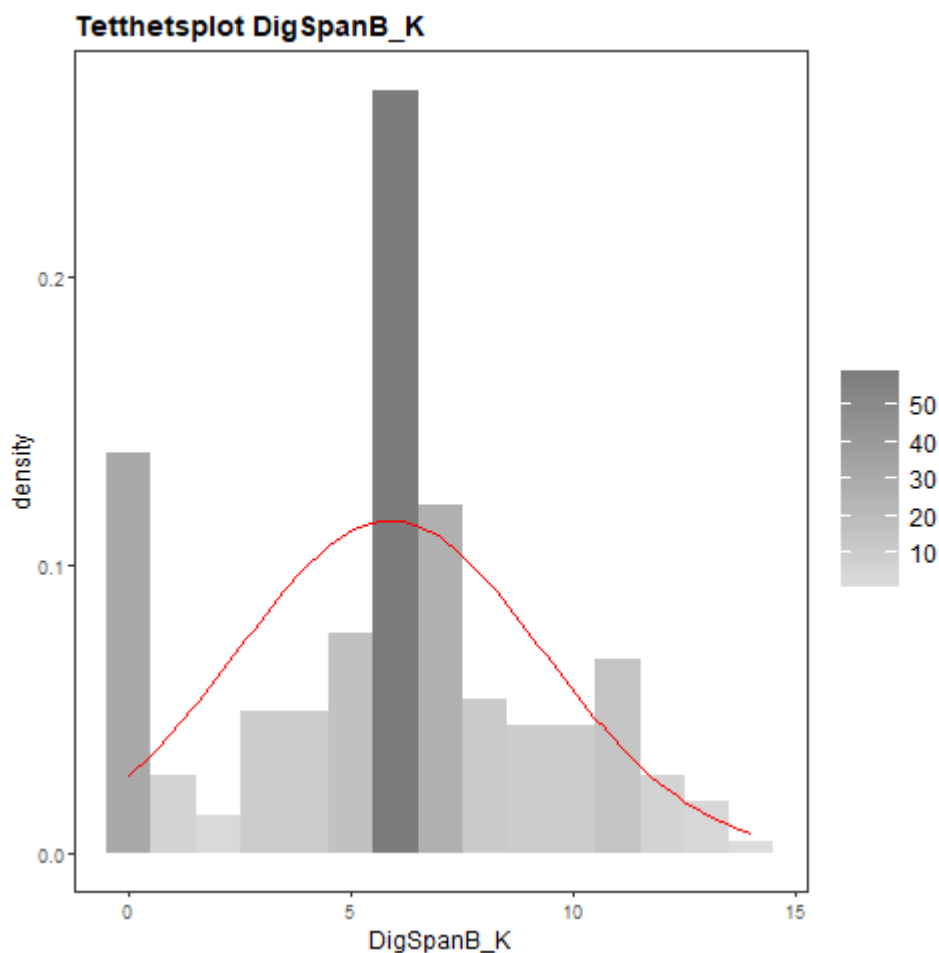
I tabellen nedenfor viser *N* til antall elever er som er med i studiens utvalg. *Range Statistic* viser til differansen mellom *Minimum Statistic* (lavest poengscore) og *Maximum Statistic* (høyest poengscore).

Mean (M) viser til gjennomsnitt. *Std. Deviation* gir variabelnes standardavvik (*s*), som er et mål på spredning rundt gjennomsnittet (Moore, McCabe, & Craig, 2017). Standardavviket =0 bare dersom det ikke er noe spredning, noe som betyr at alle datapunktene har samme verdi. Ellers er $s > 0$, jo mere spredt datapunktene er rundt gjennomsnittet, jo større blir verdien *s* (Moore, McCabe, & Craig, 2017).

Tabell 2. Beskrivende statistikk

	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation	Skewness		Kurtosis	
							Statistic	Std. Error	Statistic	Std. Error
TOBANS addisjon	223	20	3	23	9.91	3.523	.797	.163	.804	.324
TOBANS subtraksjon	223	20	0	20	6.16	3.804	.503	.163	.374	.324
Hoderegning addisjon	223	18	4	22	15.66	4.057	-.717	.163	-.084	.324
Hoderegning subtraksjon	223	14	0	14	9.95	4.117	-1.357	.163	.754	.324
Ordproblemer	223	17	4	21	11.28	3.295	.535	.163	.009	.324
Backward Digit span	223	14	0	14	5.86	3.452	-.090	.163	-.452	.324
Maze memory	223	18	0	18	4.88	3.654	1.001	.163	1.227	.324
Listening recall	223	21	0	21	10.57	4.713	-.878	.163	.443	.324
Raven's CPM	223	28	7	35	17.79	4.410	.315	.163	.465	.324
Valid N (listwise)	223									

4.2.1 Vurdering av variabelen: Backward Digit Span, WMTB-C



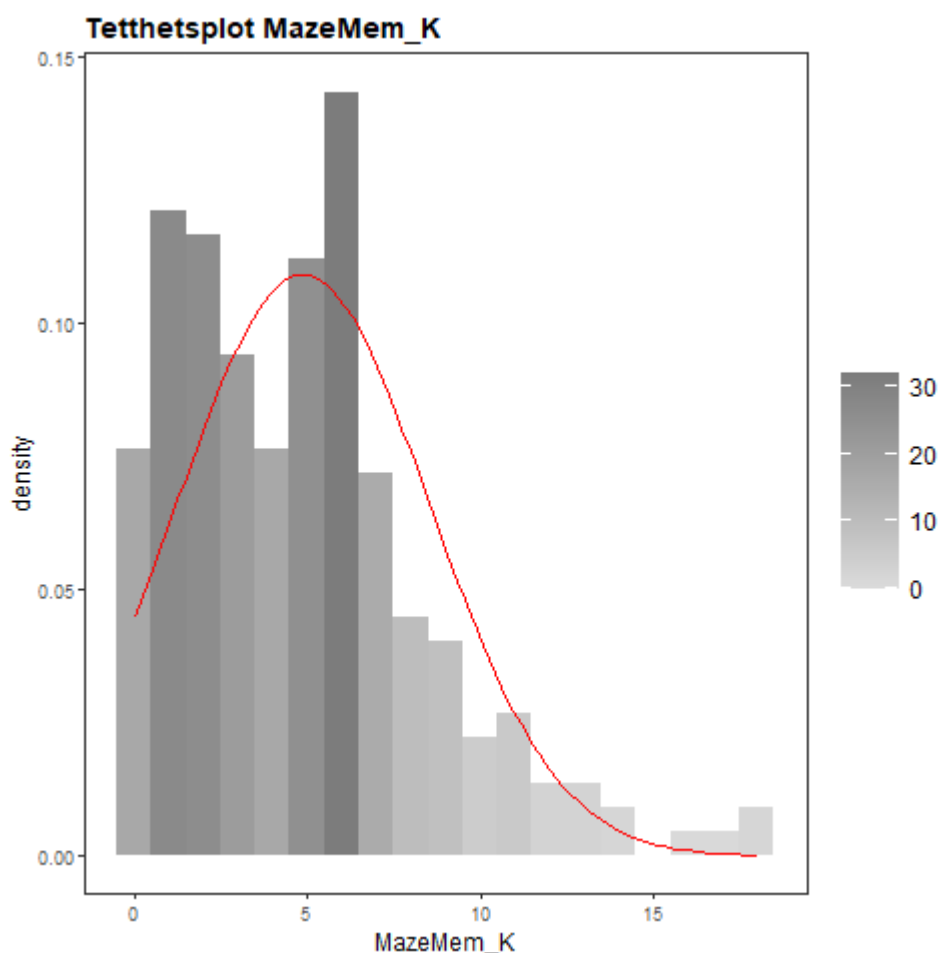
Figur 9. Fordeling av resultater for variabelen Backward Digit Span.

Variabelen Backward Digit Span er en av to kartleggingstester som i denne studien blir brukt for å måle arbeidsminnet i barnehagen. Resultatene viser et gjennomsnitt på 5.86 og et standardavvik, 3.45. Gjennomsnittsverdien gjenspeiler seg i poengscoren flest barn oppnådde, som er omtrent 6 poeng. Med en range på 14, hvor lavest score er på 0 og høyest score på 14, ligger gjennomsnittet også omtrent i normalfordelingens toppunkt.

Verdien på skjevhet er -0.09 , som indikerer en normalfordeling. Verdien for kurtosis er negativ med -0.45 , som indikerer at fordelingen er flat.

På fordelingen kan man se at litt under 15% fikk en score på 0, noe som kan tyde på at testen var for vanskelig for relativt mange av barna da de ble testet i barnehagen.

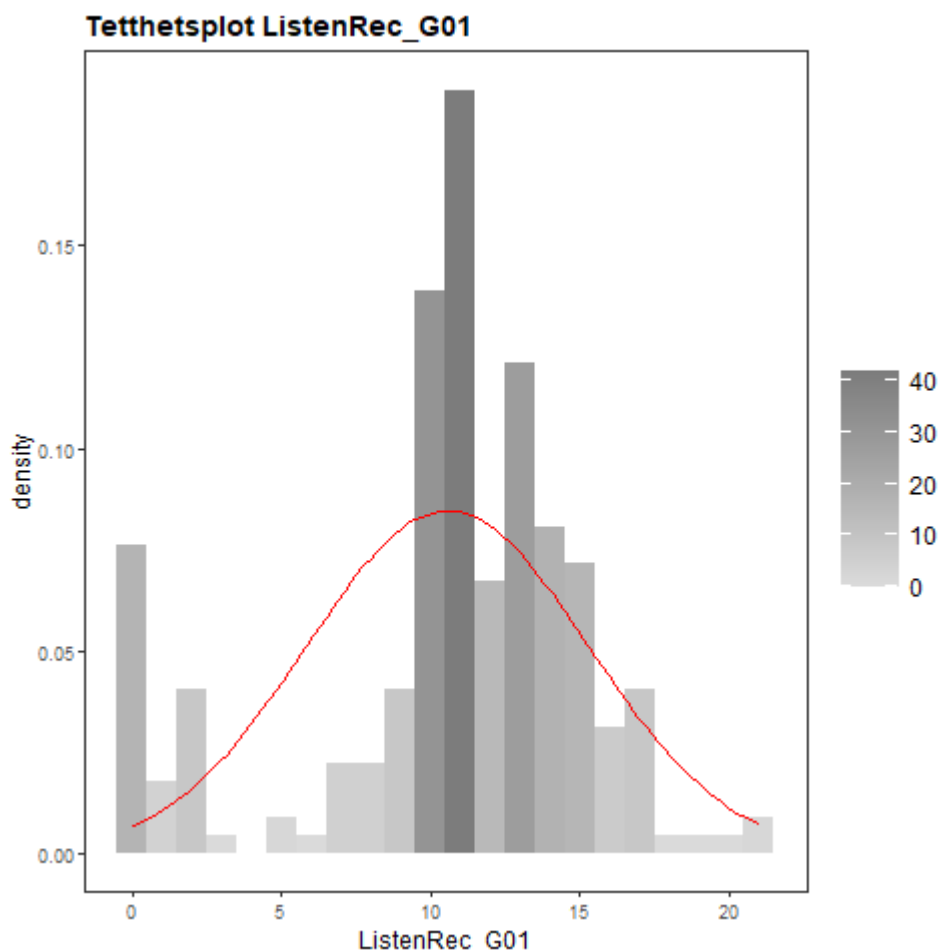
4.2.2 Vurdering av variabelen: Maze Memory



Figur 10. Fordeling av resultater for variabelen Maze memory.

Variabelen Maze memory er det andre målet på arbeidsminne i barnehagen. Variabelen har et gjennomsnitt på 4.88 og et standardavvik på 3.65. Range er på 18, med lavest score på 0 og høyest score på 18. Som man ser på grafen er fordelingen svært høyreskjev, med en skjevhetsverdi på 1.00. Det er med andre ord mange barn som oppnår en lav poengsum sammenliknet med de som scorer høyest. Dette kan peke på at testene var vanskelige for i alle fall de 7.5% som scoret 0 poeng. Kutrosisverdien er på 1.23, som viser til en spiss fordeling.

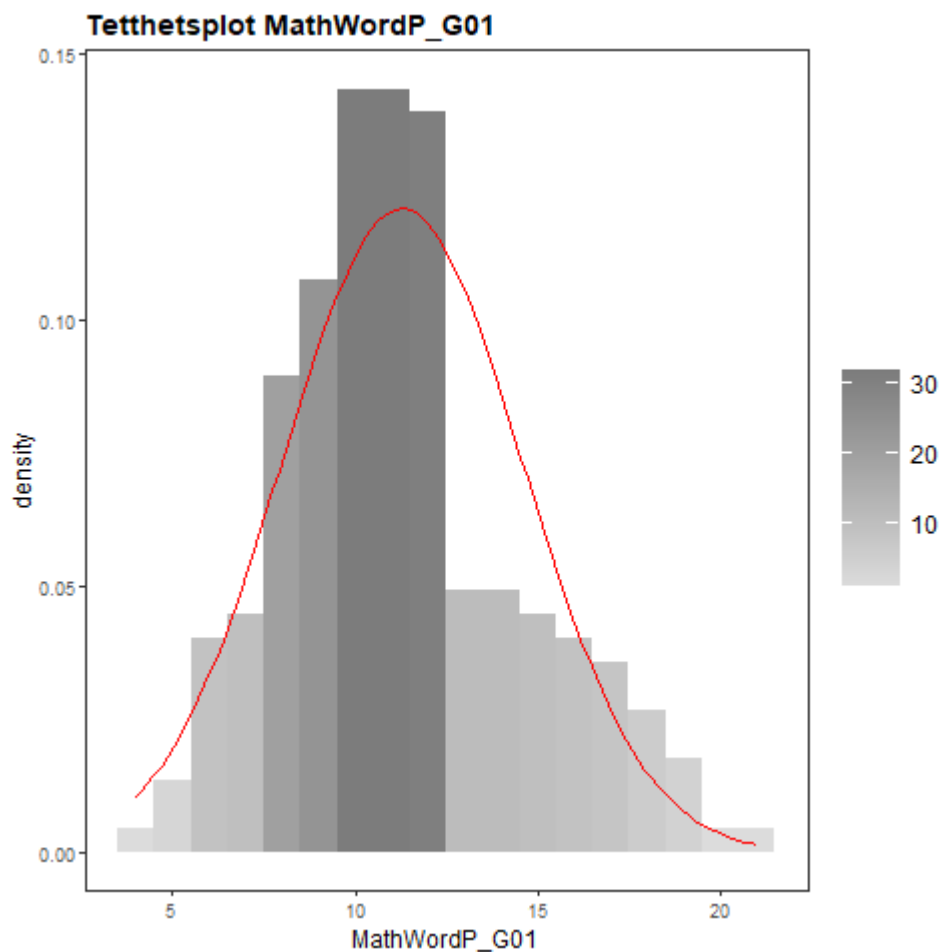
4.2.3 Vurdering av variabelen: Listening Recall, WMTB-C



Figur11. Fordeling av resultater for variabelen Listening Recall.

Listening recall er variabelen som måler arbeidsminne i 1.klasse. Gjennomsnittet ligger på 10.57 med et standardavvik på 4.71. Med en range på 21, med lavest score på 0 og høyest score på 21. Omtrent 7% får en poengscore på 0, som kan tyde på at testen var noe vanskelig for denne aldersgruppen. Skjevhetsverdien er -0.88 , som betyr at fordelingen har en lengre hale med lave verdier på venstre side av fordelingen og er dermed svakt venstreskjev. Kurtosisverdien er $.44$, som viser til en spiss fordeling.

4.2.4 Vurdering av variabelen: Ordproblemer (Math Word problem)

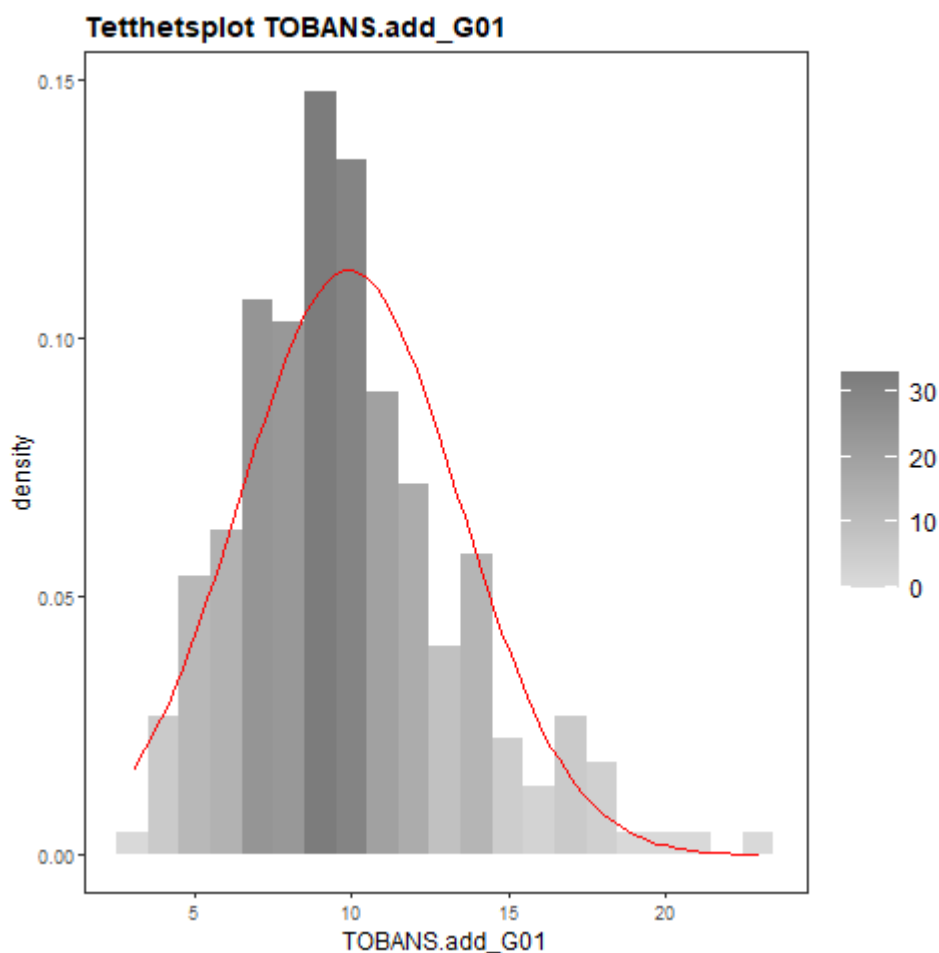


Figur 12. Fordeling av resultater for variabelen Ordproblemer.

Variabelen ordproblemer er et av målene på aritmetiske ferdigheter og har et gjennomsnitt på 11.28 med et standardavvik på 3.30. Med en range på 17, med lavest score på 4 og høyest score på 21, viser til at ingen av barna oppnådde 0 poeng, testens vanskelighetsgrad ser dermed ut til å være egnet for aldersgruppen. Resultatene indikerer svakt høyreskjev fordeling med en skjevhet på .53. En kurtoseverdien på .01, viser til at denne fordelingen er noe spiss, men nærmere normalfordelt.

4.2.5 Vurdering av variabelen: TOBANS addisjon og subtraksjon

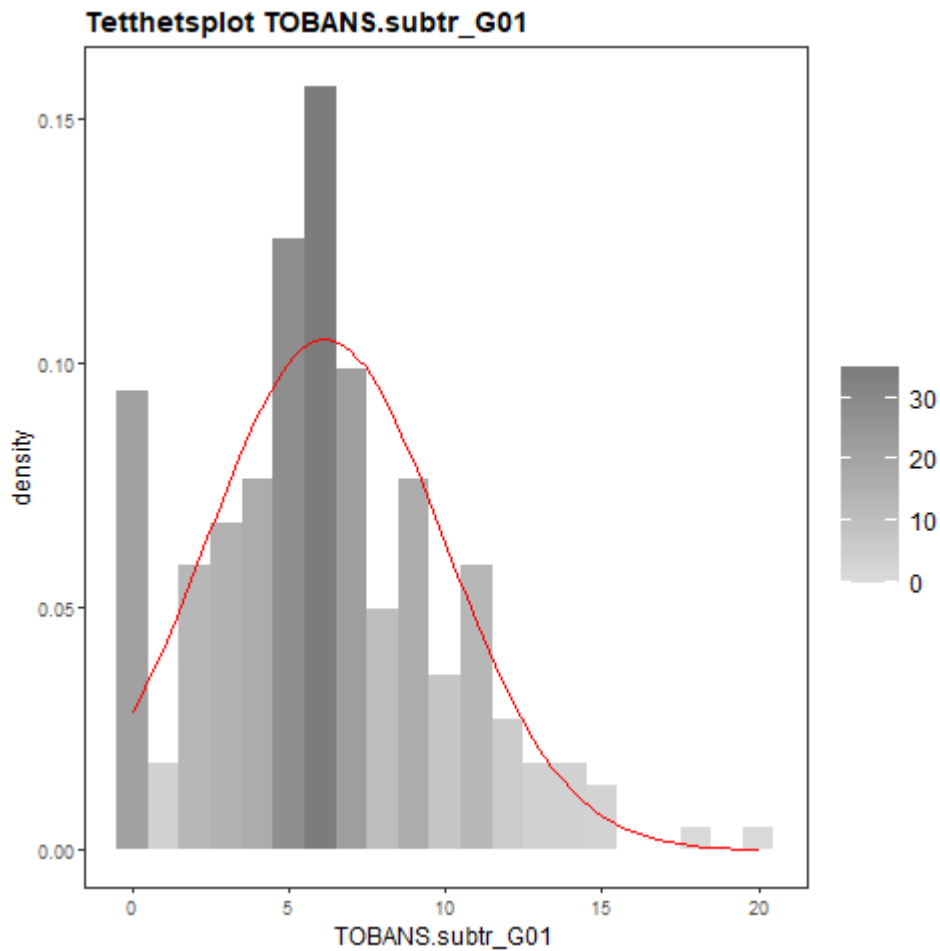
TOBANS addisjon



Figur 13. Fordeling av resultater for variabelen TOBANS addisjon.

TOBANS addisjon er et annet mål for aritmetiske ferdigheter. Gjennomsnittet i denne deltesten ligger på 9.91 med et standardavvik på standardavvik 3.52. Range er på 20, med lavest score på 3 og høyest score på 23. Det ser dermed ut til at testen er egnet for aldersgruppen da ingen fikk en score på 0 poeng. Verdien på skjevhet er på .80, da verdien er positiv, er fordelingen svakt høyreskjev. Kutosisverdien er på .80 og viser dermed til en spiss fordeling.

TOBANS subtraksjon



Figur 14. Fordeling av resultater for variabelen TOBANS subtraksjon.

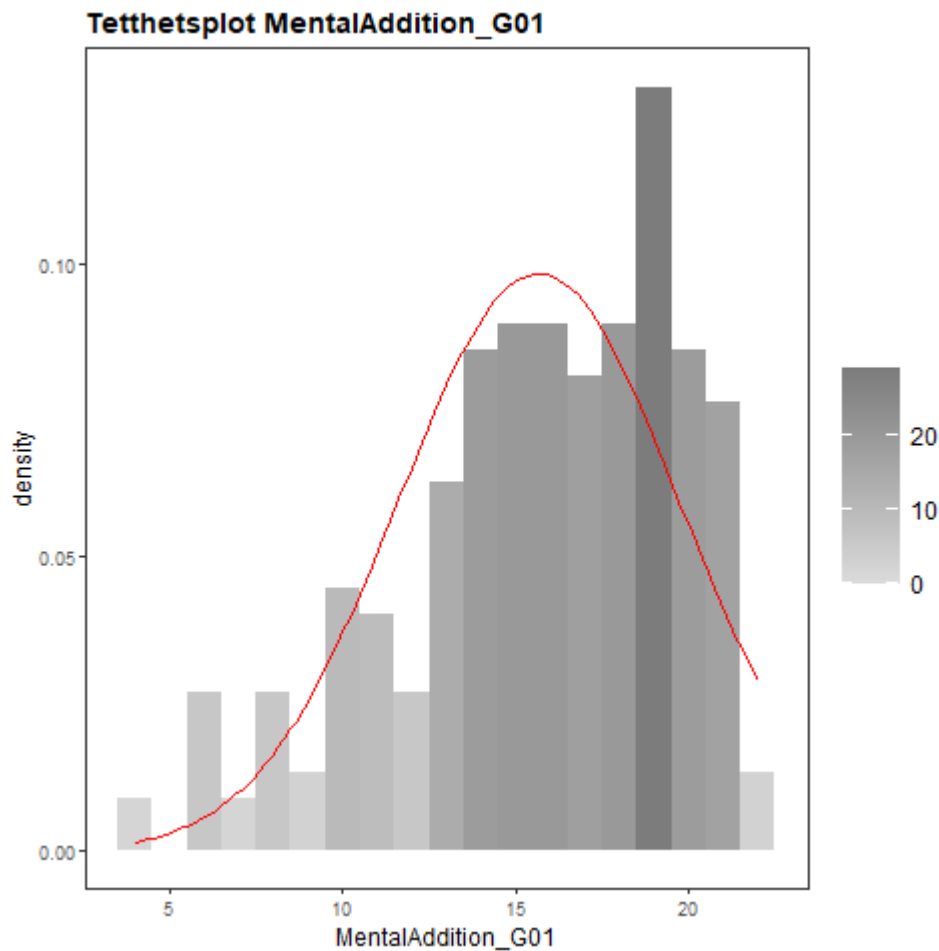
Variabelen TOBANS subtraksjon måler også aritmetiske ferdigheter. Gjennomsnittet ligger her på 6.16 med et standardavvik på 3.80. Range er på 20, med lavest score på 0 og høyest score på 20. Det kan se ut som at denne testen har vært noe for vanskelig, da nesten 10% av barna fikk en score på 0 poeng, noe som kan tyde på en gulveffekt. Dette sammenfaller med forventningene ettersom subtraksjon er ganske nytt for barna ved testtidspunktet, og vil dermed også være vanskeligere på dette stadiet.

Verdien for skjevhet er positiv med .50, noe som gir en svakt høyreskjev fordeling.

Kurtoseverdien er på 3.74 indikerer en spiss fordeling.

4.2.6 Vurdering av variabelen: Hoderegning addisjon og subtraksjon

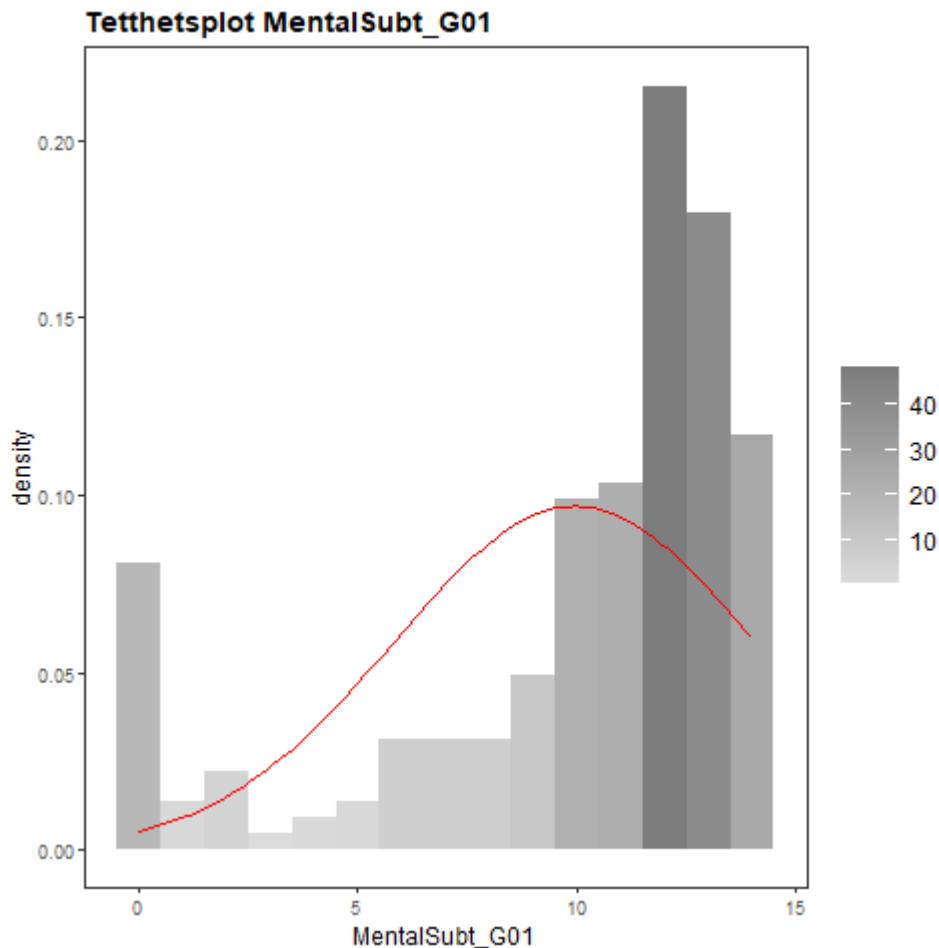
Hoderegning addisjon



Figur 15. Fordeling av resultater for variabelen Hoderegning addisjon.

Denne variabelen er også et mål på aritmetiske ferdigheter. Hoderegning addisjon har et gjennomsnitt på 15.66 og et standardavvik på 4.06. Range er på 18, med lavest score på 4 og høyest score på 22, det ser dermed ikke ut som at deltesten var for vanskelig, tvert imot. Den maksimal mulige poengscoren er 22 poeng. Dette tyder på at testen kanskje var for enkel og man får dermed en takeffekt. Verdien for skjevhet er -0.90 , som gir en svakt venstreskjev fordeling med flere lave verdier i halen til venstre. Kurtoseverdien er på $.50$, som indikerer en spiss fordeling.

Hoderegning subtraksjon



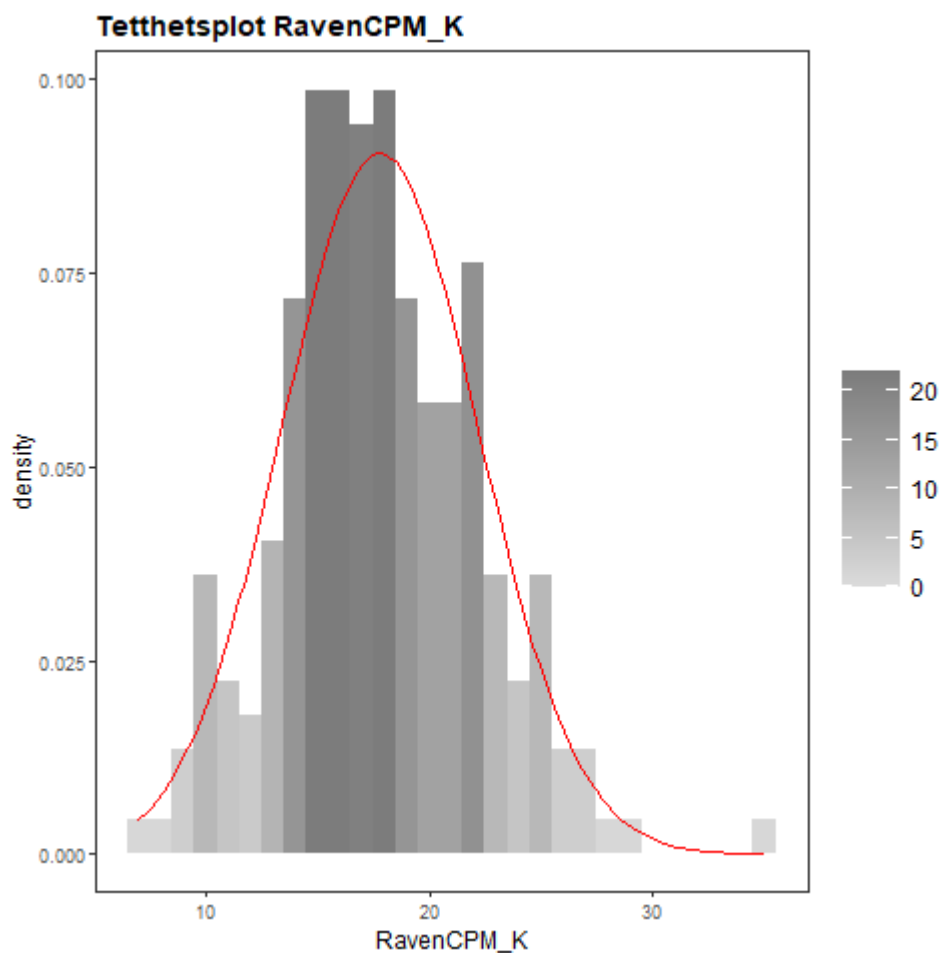
Figur 16. Fordeling av resultater for variabelen Hoderegning subtraksjon.

Hoderegning subtraksjon får et gjennomsnitt på 9.95 og et standardavvik på 4.12. Range er på 14, med lavest score på 0 og høyest score på 14. Oppgaven har totalt 14 oppgaver og mange av barna scorer svært høyt på denne deltesten, samtidig som også er omtrent 7% som scorer 0 poeng. Det kan tyde på at deltesten var for enkel og/eller inneholdt for få oppgaver for de fleste, men samtidig også for vanskelig for noen få, og kan dermed tyde på en takeffekt på den ene siden, men samtidig også på en gulveffekt.

Fordelingen gir en skjevhetsverdi på -1.36, som indikerer en svært venstreskjev fordeling.

Verdiene for kutrosis er på .75, som viser til en spiss fordeling.

4.2.7 Vurdering av variabelen: The Raven's coloured Progressive Matrices



Figur 17. Fordeling av resultater for variabelen Ravens CPM.

Raven's CPM blir benyttet som kontrollvariabel og som måler non-verbal intelligens. Resultatene gir et gjennomsnitt på 17.79 med et standardavvik på 4.410. Range er på 28, med lavest score på 7 og høyest score på 35. Testen ser altså ut til å være godt tilpasset aldersnivået. Skjevhetsverdien er på .32, som gir en symmetrisk fordeling, normalfordeling. Verdien for kutrosis er på .47, som indikerer en spiss fordeling.

4.3 Reliabilitet

Reliabilitet referer til konsistensen til et måleinstrument (Heale & Twycross, 2015; Field, 2018). Et måleinstrument er reliabelt dersom man får det samme resultat når man gjentar prosedyren (de Vaus, 2014; Heale & Twycross, 2015). De Vaus (2014) beskriver tre sider ved reliabilitet. Den første er *kilder til ikke-reliabilitet*, som referer til dårlig eller mangelfullt ordvalg i en oppgave i en test, som fører til at man kan tolke oppgaven på ulike måter. Den andre siden ved reliabilitet er å *teste reliabilitet*. Det finnes ulike måter å teste reliabilitet til en variabel på, blant annet kan man beregne korrelasjonen mellom resultatene, dersom korrelasjonen er høy (ut ifra en satt grense), antar man at oppgavene er reliable (de Vaus, 2014). Dette kan gjøres ved å benytte koeffisienten Chronbach's α (alfa), som skal kommenteres nærmere senere i dette delkapittelet. Den tredje siden er å *øke reliabiliteten*, dette kan gjøres ved å benytte varierte oppgaver for å måle en variabel.

Heale & Twycross (2015) beskriver også tre faktorer som kan bidra til å evaluere reliabilitet. Den første er *Homogeneity* (homogenitet, indre konsistens), denne kan analyseres ved å se på Chronbachs α korrelasjonen mellom items innad i en test. Den andre faktoren er *Stability* (stabilitet) som blir testet ved å bruke *test-retest*, *parallel-* eller *alternativ-form* pålitelighetstesting. Dersom man skal gjennomføre en test-retest, blir testene gjennomført flere ganger, ved ulike tidspunkt, på de samme individene under de samme omstendighetene, for så å sammenlikne et individs testresultater fra de ulike kartleggingstidspunktene. Parallell- eller alternative-former pålitelighet minner om test-retest prinsippet, men her benyttes det et annet måleinstrument for å måle den samme variabelen ved retestingen. Den tredje og siste faktoren er *Equivalence* (ekvivalens), som vurderes gjennom *inter-rate* reliabilitet. Dette er en kvalitativ reliabilitetstest som bestemmes ved å stille det samme spørsmålet til ulike individer, for så å se om svarene sammenfaller.

Reliabilitet forteller altså noe om stabiliteten til målingene. Dersom en test har god reliabilitet, vil man få like testresultater dersom man gjennomfører den samme testen igjen, under de samme forholdene. Dersom målingene ikke er reliable, kan de være en trussel for både statistisk validitet og begrepsvaliditet (Lund, 2002c).

Når man skal gi uttrykk for graden av reliabilitet, benytter man en reliabilitetskoeffisient (de Vaus, 2014). For å vurdere dette benyttes indeksen Chronbach's α (alfa), som er den mest vanlige testen for indre konsistens (Heale & Twycross, 2015), i denne studien. Dette er et

statistisk mål på indre konsistens, som sier noe om variablenes reliabilitet. Med andre ord sier denne verdien noe om hvor godt de enkelte delene i en kartleggingstest korrelerer med hverandre. Koeffisienten for reliabilitet har verdier mellom 0 og 1 (Heale & Twycross, 2015), og jo nærmere 1 denne verdien er, jo bedre er reliabiliteten til variabelen.

Reliabilitetskoeffisienten regnes som god når $\alpha > .70$ (de Vaus, 2014; Heale & Twycross, 2015).

I denne studien ligger Chronbach's $\alpha > .70$ og tar man med størrelsen på utvalget i betraktning, er reliabiliteten på deltestene ansett som å være tilfredsstillende.

Tabell 3. Reliabilitet målt med Chronbach's α

Måleinstrument	Chronbach's α
TOBANS addisjon	.771
TOBANS subtraksjon	.763
Hoderegning addisjon	.764
Hoderegning subtraksjon	.776
Ordproblemer	.767
Backward Digit span	.789
Maze memory	.790
Raven's CPM	.799
Listening recall	.805

4.4 Bivariat korrelasjonsanalyse

Bivariat korrelasjonsanalyse benyttes for å analysere sammenhengen mellom variablene for arbeidsminne, aritmetikk, intelligens i tillegg til alder og kjønn. En slik korrelasjonsanalyse viser korrelasjonen mellom de ulike variablene og hvor sterk den er ved å beregne verdier for r . Positive r -verdier viser til en positiv relasjon mellom to variabler som impliserer en positiv økning i begge variablene, mens en negativ r -verdi viser til en negativ relasjon mellom to variabler impliserer reduksjon i den ene variabelen når den andre variabelen øker. Dersom man skulle få en korrelasjonskoeffisient på 0, altså $r=0$, viser den til ingen relasjon mellom variablene. Innenfor utdanningsforskning regnes en korrelasjon på .10 som svak, .30 som middels og .50 som høy (Field, 2018). For å beregne denne modellen ble det benyttet en tohalet-test for å ikke utelukke henholdsvis en positiv eller negativ korrelasjon mellom to variabler (Lund, 2002b).

Videre viser korrelasjonstabellen hvilke korrelasjoner som er statistisk signifikante, uttrykt med p -verdier. Graden av statistisk styrke blir markert med stjerner, * $p<.05$, ** $p<.01$, *** $p<.001$. Man benytter også p -verdiene for å avgjøre om man kan eller ikke kan forkaste en nullhypotese (Moore, McCabe, & Craig, 2017). Man forkaster nullhypotesen dersom p -verdien er lav. En p -verdi $<.05$ fører til at vi forkaster nullhypotesen i favør til den alternative hypotesen, noe som betyr at det er en signifikant relasjon mellom variablene i den lineære modellen. Ellers kan vi ikke forkaste nullhypotesen.

Tabell 4. Korrelasjoner mellom variabler vist med Pearsons r .

	Kjønn	Alder	Backward Digit span	Maze memory	Raven CPM	TOBANS add.	TOBANS sub.	Hoderegning add.	Hoderegning sub.	Ord- problemer	Listening recall
Kjønn	1	.07	.09	-.08	.00	-.09	-.05	-.08	.00	-.01	.17**
Alder		1	.05	.21**	.13*	.13*	.22**	.17**	.16**	.07	.03
Backward Digit span			1	.20**	.30**	.17**	.22**	.34**	.24**	.37**	.34**
Maze memory				1	.37**	.40**	.36**	.25**	.19**	.22**	.15*
Raven's CPM					1	.24**	.22**	.24**	.17**	.22**	.19**
TOBANS add.						1	.67**	.50**	.33*	.47**	.16**
TOBANS sub.							1	.46**	.52**	.53**	.20**
Hoderegning add.								1	.50**	.56*	.25**
Hoderegning sub.									1	.49**	.24**
Ordproblemer										1	.28**
Listening recall											1

*** $p<.001$
 ** $p<.01$
 * $p<.05$

Ut ifra tabellen kan man se at de fleste variablene har positive verdier som er statistisk signifikante. Dette betyr at korrelasjonen mellom variablene er positiv, altså medfører en positiv økning i en variabel til en positiv økning i den andre variabelen. Variabelen kjønn ser ut til å bare ha en signifikant korrelasjon til Listening recall med $p < .01$. Alder viser en korrelasjon til de fleste variablene, med noen unntak. Backward Digit span, ordproblemer og Listening recall ser ikke ut til å være korrelert med alder. Variablene for arbeidsminne, aritmetikk og intelligens er i de fleste tilfellene statistisk positiv korrelert med $p < .01$. Det er tre par korrelasjoner som skiller seg ut med å være noe mindre korrelerte enn de andre variablene. TOBANS addisjon og hoderegning subtraksjon er noe mindre positivt korrelerte ($r = .33, p < .05$). Videre er også korrelasjonen mellom ordproblemer og hoderegning addisjon noe mindre ($r = .56, p < .05$). Den siste er korrelasjonen mellom variablene Listening recall og Maze memory ($r = .15, p < .05$).

4.5 Hierarkisk Multippel Regresjonsanalyse

En regresjonsanalyse benyttes når man ønsker å finne ut *hvilken* sammenheng det er mellom to variabler (Løvås, 2013). Denne studien ønsker å se på om flere variabler (forklaringsvariabel) kan forklare utfallet til andre variabler (responsvariabler). En *Multippel regresjonsanalyse* brukes for å finne ut hvilke forklaringsvariabler som viser sterkest korrelasjon til responsvariablene.

I denne studien benyttes en *Hierarkisk multippel regresjonsanalyse*, da man kan velge rekkefølgen på variablene som skal inngå i analysen selv. Rekkefølgen velges basert på forklaringsverdien til hver enkelt variabel. Forklaringsvariabler, som man antar kan predikere responsvariablenes utfall, analyseres først i modellen. Denne analysen vil forklare noe varians mellom forklaringsvariablene og responsvariablene.

De neste stegene er å inkludere andre forklaringsvariabler, disse vil gi en forklaringsverdi for de nye forklaringsvariablene, som ikke allerede har blitt forklart av tidligere forklaringsvariabler. Siden variabler korrelerer med hverandre, vil rekkefølgen man velger også kunne påvirke resultatene (Field, 2018). For å gjennomføre analysen benyttes statistikkprogrammet SPSS og R-studio (IBM, 2019).

Beta-verdien er stigningstallet i en regresjonsanalyse. Verdien forteller hvor mye den gjennomsnittlige verdien til en responsvariabel (Y) øker eller minker, når en forklaringsvariabel (X) endres med en enhet (Løvås, 2013). Denne verdien beskriver altså sammenhengen mellom responsvariabelen og forklaringsvariabelen.

4.5.1 Forutsetninger for hierarkisk multippel regresjonsanalyse

Det stilles noen kriterier dersom man ønsker å trekke slutninger basert på en hierarkisk multippel regresjonsanalyse. Størrelsen på utvalget må være stort nok dersom man ønsker å finne statistisk signifikante korrelasjoner. For multippel regresjon beregner man 15 caser for hver prediktorvariabel (uavhengig variabel) (Field, 2018). Ut ifra dette bør denne studien ha minst 90 barn. I denne studien er utvalget $N = 223$.

4.5.2 Gjennomføring og resultater av hierarkisk regresjonsanalysen

Studiens forskningsspørsmål er følgende:

- I. *Hvilken sammenheng er det mellom barns arbeidsminne og grunnleggende aritmetiske ferdigheter?*
- II. *Hvilken del av arbeidsminnet korrelerer sterkest med grunnleggende aritmetiske ferdigheter, når det kontrolleres for intelligens?*

Forskingsspørsmålet tar utgangspunkt i en hypotese om at arbeidsminne i barnehagen kan forklare variasjonen i aritmetiske ferdigheter i førsteklasse. Nullhypotesen vil derfor være at det ikke er noe sammenheng mellom de to variablene. Som tidligere nevnt forkaster man nullhypotesen dersom p-verdien er lav. En p-verdi $<.05$ fører til at vi forkaster nullhypotesen i favør til den alternative hypotesen, noe som betyr at det er en signifikant relasjon mellom variablene i den lineære modellen (Moore, McCabe, & Craig, 2017). Ellers kan vi ikke forkaste nullhypotesen. Man forkaster altså nullhypotesen dersom p-verdien er lav. Ellers kan vi ikke forkaste nullhypotesen.

Det har blitt beregnet en hierarkisk multippel regresjonsanalyse for hver responsvariabel for å kunne besvare dette forskningsspørsmålet. Denne studien velger å sette inn kontrollvariablene; alder og kjønn som første steg i analysen. I det andre steget legges variablene for intelligens i barnehage (The Ravens' CPM) til. I det tredje steget legges arbeidsminne i barnehage (Backward Digit span, Maze memory) til, og i det fjerde og siste steget ble en variabel for arbeidsminne i førsteklasse (Listening recall) lagt til. I analysene vil en og en variabel for aritmetiske ferdigheter i første klasse bli satt inn i den hierarkiske multiple regresjonsmodellen med de ovenfor nevnte stegene.

I tabellen vil de ulike testtidspunktene bli markert med t for å forklare fra hvilke testtidspunkt dataene er hentet fra, t^1 står for data som er innhentet i siste året i barnehagen (ved 5 årsalder), og t^2 står for at dataene er innhentet i første klasse.

TOBANS addisjon og arbeidsminne

Tabell 5. Prediktor for TOBANS addisjon.

Prediktor for TOBANS addisjon (t ²)		▲R ²	β	t	Signifikans F-change
Steg 1					
Alder			-.10	-1.54	.126
Kjønn			.13	2.02	.045
Total R ²	.03	.03			
Steg 2					
Raven's CPM (t ¹)			.22	3.41	.001
Total R ²	.08	.05**			
Steg 3					
Backward Digit span (t ¹)			.09	1.33	.186
Maze memory (t ¹)			.34	5.00	.000
Total R ²	.18	.11***			
Steg 4					
Listening recall (t ²)			.09	1.32	.187
Total R ²	.19	.01			
<hr/>					
N	223				

Note: *** $p < .001$ ** $p < .01$ * $p < .05$

I steg 1 forklarer kontrollvariablene alder og kjønn 2.7% ($F(2,220)=3.02$, $p=.051$) av variansen i TOBANS addisjon, og er marginalt signifikant. Modellen viser ikke signifikante nivåer for kjønn, men variabelen alder er signifikant $<.05$ med $\beta=.13$ og t -verdi= 2.02.

Når man legger til variabelen for intelligens i steg 2, forklarer intelligens 4.9% ($F(1,219)=11.59$, $p=.001$) av variasjonen. Her blir intelligens en svært signifikant $<.01$ prediktor med $\beta=.00$ og $t=3.41$. Alder er ikke en signifikant prediktor lenger.

I steg 3 legges variablene for arbeidsminne i barnehage til i modellen. Modellen viser at arbeidsminne i barnehage alene forklarer 10.8% ($F(2,217)=14.33$, $p<.001$) av variansen i

TOBANS addisjon. Arbeidsminne variabelen Maze memory er sterk signifikant $<.001$, $\beta=.34$ og $t=5.00$. Arbeidsminne variabelen Backward Digit span er ikke signifikant for TOBANS addisjon og intelligens er ikke lenger en signifikant prediktor.

I steg 4 legges variabelen for arbeidsminne i første klasse til, denne forklarer alene 0.7% ($F(1,216)=1.75$, $p=.187$), som dermed ikke viser en ytterligere varians fra steg 3. Denne variabelen er ikke en signifikant prediktor, men Maze memory forblir sterkt signifikant $<.001$, men får en noe lavere $\beta=.33$ og en litt lavere t -verdi $=4.89$. En lavere β -verdi viser til en mindre bratt regresjonslinje, som betyr at korrelasjonen mellom TOBANS addisjon og Maze memory har blitt noe svakere enn det den var i steg 3.

TOBANS subtraksjon og arbeidsminne

Tabell 6. Prediktor for TOBANS subtraksjon.

Prediktor for TOBANS subtraksjon (t ²)		▲ R ²	β	t	Signifikans F-change
Steg 1					
Alder			.22	3.38	.001
Kjønn			-.07	-1.90	.315
Total R ²	.05	.04**			
Steg 2					
Raven's CPM (t ¹)			.19	2.96	.003
Total R ²	.09	.08**			
Steg 3					
Backward Digit span (t ¹)			.14	2.19	.030
Maze memory (t ¹)			.28	4.07	.000
Total R ²	.18	.16***			
Steg 4					
Listening recall (t ²)			.17	1.74	.083
Total R ²	.19	.17			
<hr/>					
N	223				

Note: *** $p < .001$ ** $p < .01$ * $p < .05$

I steg 1 forklarer kontrollvariablene alder og kjønn 5.2% ($F(2,220)=6.01$, $p=.003$) av variansen i TOBANS subtraksjon. Modellen viser ikke noe signifikans for variabelen kjønn, men variabelen alder er signifikant $<.01$ med $\beta=.22$ og t -verdi= 3.38.

I steg 2 legger man til variabelen for intelligens. Intelligens forklarer 3.7% ($F(1,219)=8.79$, $p=.003$) unik variasjon. Intelligens er også en signifikant ($<.01$) prediktor i dette steget, med $\beta=.19$ og $t=2.96$. Variabelen alder forblir en signifikant ($<.01$) prediktor med $\beta=.20$ og $t=3.04$. Til sammen forklarer kontrollvariablene 7.6% av variasjonene for TOBANS subtraksjon.

I steg 3 legges variablene for arbeidsminne målt i barnehage til i modellen. Arbeidsminne alene forklarer 9.0% ($F(2,217)=11.83$, $p < .001$) av variansen i TOBANS subtraksjon og er svært signifikant. Ser man på hver arbeidsminne variabel hver for seg, er Maze memory sterk signifikant $<.001$, $\beta=.28$ og $t=4.07$. Videre er også variabelen Backward Digit span signifikant ($<.05$) for TOBANS subtraksjon, $\beta=.14$ og $t=2.19$.

Alder forblir en signifikant ($<.05$) prediktor med $\beta=.15$ og $t=2.35$, men intelligens er ikke lenger en signifikant prediktor.

Arbeidsminne i barnehage sammen med kontrollvariablene forklarer i denne modellen 17.8% av variansen for TOBANS subtraksjon. Det ser ut som at arbeidsminne i barnehage forklarer det meste av variasjonen.

I steg 4 legges arbeidsminne i første klasse legges til i modellen. Denne variabelen forklarer kun 1.1% ($F(1,216)=3.04$, $p = .083$) av variasjon for TOBANS subtraksjon. Denne variabelen viser med andre ord ingen ytterligere variasjoner. Listening recall og Backward Digit span er ikke en signifikant prediktor i dette steget, men alder forblir en signifikant ($<.05$) prediktor, $\beta=.15$ og $t=2.39$. Maze memory forblir sterkt signifikant $<.001$, men får en noe høyere β -verdi $=.27$ og en noe lavere t -verdi $=3.94$. En høyere β -verdi viser til en brattere regresjonslinje, som betyr at korrelasjonen mellom TOBANS subtraksjon og Maze memory har blitt noe sterkere enn det den var i modell 3.

Hoderegning addisjon

Tabell 7. Prediktor for Hoderegning addisjon.

Prediktor for Hoderegning addisjon (t^2)		ΔR^2	β	t	Signifikans F-change
Steg 1					
Alder			.18	2.67	.008
Kjønn			-.09	-1.33	.184
Total R^2	.04	.03*			
Steg 2					
Raven's CPM (t^1)			.22	3.37	.001
Total R^2	.09	.07**			
Steg 3					
Backward Digit span (t^1)			.29	4.38	.000
Maze memory (t^1)			.13	1.84	.067
Total R^2	.20	.16***			
Steg 4					
Listening recall (t^2)			.15	2.22	.027
Total R^2	.20	.17*			
<hr/>					
N	223				

Note: *** $p < .001$ ** $p < .01$ * $p < .05$

I steg 1 forklarer alder og kjønn 3.7% ($F(2,220)=4.24$, $p = .016$) av variasjonen i hoderegning addisjon. Modellen viser heller ikke her signifikante verdier for kjønn, men variabelen alder er signifikant ($<.01$) med $\beta = .18$ og t -verdi=2.67.

I steg 2 forklarer intelligens 4.7% ($F(1,219)=11.36$, $p = .001$) av variasjonen. Intelligens er en signifikant ($<.01$) prediktor for hoderegning addisjon, med $\beta = .22$ og $t = 3.37$. Alder forblir en signifikant ($<.05$) prediktor med en noe lavere $\beta = .15$ og $t = 2.29$. Til sammen forklarer kontrollvariablene 7.2% av variasjonen for hoderegning addisjon.

I steg 3 legges variablene for arbeidsminne i barnehage til i modellen. forklarer 9.3% ($F(2,217)=12.34$, $p < .001$) av variansen i hoderegning addisjon. Backward Digit span er sterk

signifikant $<.001$, $\beta=.29$ og $t=4.38$, mens variabelen Maze memory ikke er signifikant for hoderegning addisjon. Alder er fortsatt signifikant ($<.05$) og har $\beta=.13$ og $t=1.98$, men intelligens er ikke lenger en signifikant prediktor.

Arbeidsminne i barnehage sammen med kontrollvariablene forklarer denne modellen 8.5% av variansen for hoderegning addisjon.

I det siste steget, steg 4, legges variabelen Listening recall for arbeidsminne i første klasse til. Listening recall forklarer 1.3% ($F(1,216)=4.95$, $p=.027$) av variasjonen. Denne variabelen er her en signifikant ($<.05$) prediktor med en β -verdi $=.15$ og t -verdi $=2.22$.

Backward Digit span forblir sterkt signifikant $<.001$, men får en noe lavere $\beta=.24$ og en litt lavere t -verdi $=3.61$. En lavere β -verdi viser til en mindre bratt regresjonslinje, som betyr at korrelasjonen mellom hoderegning addisjon og Backward Digit span har blitt noe svakere enn det den var i steg 3.

Hoderegning subtraksjon

Tabell 8. Prediktor for Hoderegning subtraksjon.

Prediktor for Hoderegning subtraksjon (t ²)		▲ R ²	β	t	Signifikans F-change
Steg 1					
Alder			.16	2.39	.018
Kjønn			-.01	-.16	.872
Total R ²	.03	.02			
Steg 2					
Raven's CPM (t ¹)			.15	2.29	.023
Total R ²	.05	.04*			
Steg 3					
Backward Digit span (t ¹)			.20	2.93	.004
Maze memory (t ¹)			.10	1.45	.149
Total R ²	.10	.08**			
Steg 4					
Listening recall (t ²)			.18	2.53	.012
Total R ²	.12	.10*			
<hr/>					
N	223				

Note: *** $p < .001$ ** $p < .01$ * $p < .05$

Steg 1 forklarer 2.5% ($F(2,220)=2.84$, $p = .060$) av variasjonen i hoderegning subtraksjon, gitt av variablene alder og kjønn. Modellen forklarer i dette steget ikke en signifikant mengde varians. Ser man på en og en variabel ser man at kjønn ikke er en signifikant prediktor, men variabelen alder er signifikant $< .05$ med $\beta = .16$ og t -verdi=2.39.

Steg 2, med intelligens som ny variabel, forklarer 2.3% ($F(1,219)=5.25$, $p = .023$) av variasjonen. Intelligens er en signifikant $< .05$ prediktor med $\beta = .15$ og $t = 2.29$. Alder forblir signifikant $< .05$ med en β -verdi=.14 og en t -verdi=2.10. Til sammen forklarer kontrollvariablene 3.5% av variasjonene for hoderegning subtraksjon.

I steg 3 forklarer arbeidsminne i barnehage alene 4.9% ($F(2,217)=5.90$, $p= .003$) av variansen i hoderegning subtraksjon. Ser man på hver arbeidsminnevariabel hver for seg ser man at Backward Digit span er sterk signifikant $<.01$, med $\beta=.20$ og $t=2.93$, men at Maze memory ikke er signifikant for hoderegning subtraksjon. Alder og intelligens er ikke lengre signifikante prediktorer. Variabelen arbeidsminne i barnehage sammen med kontrollvariablene forklarer denne modellen 9.7% av variansen for hoderegning subtraksjon.

I steg 4 legges variabelen Listening recall for arbeidsminne i første klasse til. Listening recall forklarer 2.6% ($F(1,216)=6.41$, $p= .012$). Listening recall er signifikant $<.05$ med $\beta=.18$ og $t=2.53$. Digit span blir i dette steget mindre signifikant $<.05$, og får en noe lavere β -verdi $=.15$ og en lavere t -verdi $=2.13$. En lavere β -verdi viser til en mindre bratt regresjonslinje, som betyr at korrelasjonen mellom hoderegning subtraksjon og Backward Digit span har blitt noe svakere enn det den var i modell 3.

Ordproblemer

Tabell 9. Prediktor for Ordproblemer.

Prediktor for Ordproblemer (t ²)		▲R ²	β	t	Signifikans F-change
Steg 1					
Alder			.07	1.04	.298
Kjønn			-.02	-.26	.799
Total R ²	.01	-.00			
Steg 2					
Raven's CPM (t ¹)			.22	3.28	.001
Total R ²	.05	.04**			
Steg 3					
Backward Digit span (t ¹)			.32	4.86	.000
Maze memory (t ¹)			.12	1.72	.087
Total R ²	.16	.14***			
Steg 4					
Listening recall (t ²)			.17	2.57	.011
Total R ²	.19	.17*			
<hr/>					
N	223				

Note: *** $p < .001$ ** $p < .01$ * $p < .05$

Alder og kjønn i steg 1 forklarer 0.5% ($F(2,220)=.56$, $p= .571$). Modellen forklarer ikke en signifikant varians i dettes steget.

I steg 2 forklarer intelligens alene 4.7% ($F(1,219)=10.75$, $p= .001$) av variasjonen. Intelligens er signifikant $<.01$ med $\beta=.22$ og $t=3.28$. Til sammen forklarer kontrollvariablene 3.9% av variasjonene for ordproblemer.

I steg 3 legges variablene for arbeidsminne i barnehage til i modellen. Arbeidsminne forklarer her 11.1% ($F(2,217)=14.41$, $p < .001$) av variansen i ordproblemer. Ser man på hver enkelt arbeidsminnevariabel er Backward Digit span sterk signifikant $<.001$, $\beta=.32$ og $t=4.86$. Variabelen Maze memory er ikke signifikant for ordproblemer. Arbeidsminne i barnehage sammen med kontrollvariablene forklarer denne modellen 16.3% av variansen for ordproblemer. Intelligens er ikke lenger en signifikant prediktor.

I steg 4 legges variabelen Listening recall for arbeidsminne i første klasse til. Listening recall forklarer 2.5% ($F(1,216)=6.58, p=.011$) av variasjonen i ordproblemer. Videre er Listening recall signifikant $<.05$, med $\beta=.17$ og $t=2.57$. Backward Digit span forblir signifikant $<.01$, men får en noe lavere $\beta=.27$ og en litt lavere t-verdi=4.0. En lavere betaverdi viser også her til en mindre bratt regresjonslinje, som betyr at sammenhengen mellom ordproblemer og Maze memory har blitt noe svakere enn det den var i modell 3.

4.6 Oppsummering, analyser

Etter å ha gjennomført en deskriptiv analyse og en hierarkisk multippel regresjonsanalyse med en vurdering av hver enkelt variabel, kan man se at variablene for arbeidsminne gir signifikante p-verdier $<.05$ for alle aritmetikk variabler etter å ha kontrollert for alder, kjønn og intelligens.

Arbeidsminne målt i barnehage ser ut ifra dette ut til å ha en statistisk signifikant korrelasjon til aritmetiske ferdigheter i første klasse i denne studien. Videre var det noen arbeidsminnevariabler som viste seg å være signifikant for noen aritmetiske ferdigheter, men ikke for andre, i tillegg forklarer også arbeidsminnevariabelen, som ble målt i førsteklasse (Listening recall), en signifikant forklaringsverdi.

For TOBANS addisjon og subtraksjon viste Maze memory en sterkt sammenheng, men i tillegg til Maze memory, viste også Backward Digit span signifikante verdier for TOBANS subtraksjon. Både Backward Digit span og Maze memory ser ut til å ha en sammenheng med variasjonen i hoderegning addisjon. Videre ser også Listening recall ut til å ha signifikant forklaringsverdi for hoderegning addisjon. Signifikante prediktorer for hoderegning subtraksjon er Backward Digit span og Listening recall. Når det gjelder ordproblemer, ser det ut til at Backward Digit span er den testen for arbeidsminne som forklarer mest variasjon, men også Listening recall er signifikant.

5 Resultater i lys av teori og empiri

I denne delen skal resultatene drøftes i henhold til teori og empiri som ble gitt tidligere i studien. Studiens validitet vil drøftes med utgangspunkt i validitetssystemet til Cook og Campell (1979). Videre vil også studiens reliabilitet bli diskutert. Det vil bli gjort en vurdering av arbeidsminne som måleinstrument, for så å avslutningsvis vurdere studiens relevans for det spesialpedagogiske fagfeltet i tillegg til eventuelle muligheter for videre forskning.

Sammenheng mellom arbeidsminne ved 5 årsalder og aritmetiske ferdigheter i 1.klasse

For **TOBANS addisjon** er det beta-verdien til Maze memory som gir det høyeste stigningstall ($=.34$). Maze memory får også en høy t-verdi ($= 5.00$) og en signifikans $<.001$. Ut ifra disse verdiene ser man at Maze memory, som måler det visuospatiale arbeidsminne, ser ut til å være en sterk prediktor for TOBANS addisjon. Dette kan ha en sammenheng med at Maze memory er en kartleggingstest hvor barnet må huske en visuell presentert vei uten å måtte forholde seg til verbal input. TOBANS addisjon er en non-verbal aritmetikk test, hvor barnet må løse regnestykker på papir og dermed må bearbeide informasjon som blir gitt visuelt. Det er mulig at barnet blant annet benytter seg av en mental tallinje eller mentale representasjoner av tallstørrelser under utregningen av addisjonsoppgaver, noe som stemmer overens med faglitteraturen som beskriver at det visuospatiale arbeidsminne er viktig for utviklingen av visuelle og rolige representasjoner av tall, i form av for eksempel en mental tallinje (Gilmore, Göbel, & Inglis, 2018). Forskning viser også at det visuospatiale arbeidsminne ser ut til å være viktig for løse skrevne aritmetiske oppgaver ved at det holder styr på plasseringene til tallene i henhold til titallsystemet (Gilmore, Göbel, & Inglis, 2018).

For **TOBANS subtraksjon** er beta-verdien til Maze memory som gir det høyeste stigningstallet ($=.28$). Maze memory får også her en høy t-verdi ($=4.07$) og en signifikans $<.001$. Ut ifra disse verdiene ser man at Maze memory, ser ut til å være en svært sterk prediktor også for TOBANS subtraksjon. Som nevnt i avsnittet over kan det også her ha en sammenheng med at dette er en kartleggingstest hvor barnet må tolke og bearbeide visuelle og skriftlige representasjoner av tall. Det ser ut til at resultatene er i tråd med tidligere forskning, som viser til at det visuospatiale arbeidsminne ser ut til å være viktig ved løsning av skrevne

aritmetiske oppgaver (Gilmore, Göbel, & Inglis, 2018). Det er også her mulig at barnet benytter seg av en mental tallinje eller mentale representasjoner av tallstørrelser under utregningen av subtraksjonsoppgaver for å holde styr på plasseringene i titallsystemet. Videre er også variabelen Backward Digit span signifikant ($<.05$) for TOBANS subtraksjon med $\beta=.14$ og $t=2.19$. Backward Digit span måler som sagt den Central executive og det verbale arbeidsminne. Siden kunnskapen om hvordan man løser en subtraksjonsoppgave er ganske nytt for barna da de ble testet i førsteklasse, antas det at denne beregningen er nokså kompleks for dem ved testtidspunktet. Noe som kan forklare hvorfor den Central executive er en signifikant prediktor for TOBANS subtraksjon. Videre kreves det i en av oppgavene at man må mestre utregning med en tierovergang. Slike evner har vist seg å være involvert i prosesser i den Central executive (Gilmore et al., 2018).

Det er tilsynelatende ingen synlig eller innlysende faktorer ved denne aritmetikktesten som peker på at den også skulle ha involvert det verbale arbeidsminne i prosessen. Det kan skyldes at det verbale arbeidsminnet er sterkere representert i Digit recall oppgaver hvor tallene skal gjentas i presentert rekkefølge, mens når tallene skal gjentas i omvendt rekkefølge, er den Central executive mere involvert (Pickering, 2006).

Sammenlikner man resultatene fra addisjon og Maze memory med resultatene for subtraksjon og Maze memory, ser man noe lavere verdier for β og t , men lik signifikans for subtraksjon. Dette viser til en noe flattere regresjonslinje for TOBANS subtraksjon, sammenliknet med TOBANS addisjon, når man ser på sammenhengen mellom henholdsvis TOBANS subtraksjon - Maze memory og TOBANS addisjon – Maze memory

Både TOBANS addisjon og subtraksjon måler i tillegg til korrekthet, også prosesseringshastighet. Som tidligere beskrevet antar man at prosesseringshastighet henger sammen med den kognitive utviklingen (Camos & Barrouillet, 2018). Grunnet i dette ser man altså også at det er en korrelasjon mellom prosesseringshastighet og arbeidsminne i denne studien.

For **hoderegning addisjon** er det β -verdien til Digit span som gir det høyeste stigningstallet ($=.29$). Backward Digit span får også en høy t -verdi ($= 4.38$) og en signifikans $<.001$. Ut ifra disse verdiene ser man at Backward Digit span, som måler den Central executive og det verbale arbeidsminne, ser ut til være en sterk prediktor for hoderegning addisjon. Dette kan ha en sammenheng med at Backward Digit span er en kartleggingstest

hvor barnet må huske tall som blir presentert muntlig, og må dermed bearbeide muntlig input for å så svare korrekt. Man antar at minnespennsoppgaver, som Backward Digit span, gjenspeiler funksjonsevnen til det verbale arbeidsminne tydelig siden det verbale arbeidsminne inneholder et midlertidig fonologisk lager, som gjør den spesielt godt egnet til å oppbevare sekvenser av informasjon (Baddeley, 2000).

Ut ifra hvordan oppgaven er satt opp øker de i vanskelighetsgrad ved at tallene som presenteres blir større. Dette innebærer at barnet må bruke mellomregninger for eksempel ved hjelp av minnetall ved tieroverganger. Under slike komplekse aritmetiske delprosesser antar man at den Central executive bidrar ved bearbeidingen (Gilmore, Göbel, & Inglis, 2018).

For **hoderegning subtraksjon** er det beta-verdien til Backward Digit span som gir det høyeste stigningstall (= .20). Backward Digit span får også en høy t-verdi (= 2.93) og en signifikans = .004. Ut ifra disse verdiene ser man at Backward Digit span ser ut til å være en sterk prediktor for hoderegning subtraksjon. Dette kan også her ha en sammenheng med at Backward Digit span er en kartleggingstest hvor barnet må huske tall som blir presentert muntlig, og må dermed bearbeide muntlig input for å så svare korrekt. Som sagt inkluderer testen Backward Digit span, arbeidsminnekomponentene verbalt-arbeidsminne og Central executive. I likhet med hoderegning addisjon krever også her de mer kompliserte oppgavene at man mestrer utregninger med tieroverganger. Siden oppgaven også har en verbal faktor, ser man også her at funksjonsevne til det verbale arbeidsminne er tydelig korrelert med hoderegning subtraksjon.

For **Ordproblemer** er det beta-verdien til Backward Digit span som gir det høyeste stigningstall (= .32). Backward Digit span får også en høy t-verdi = 4.86 og en signifikans < .001. Ut ifra disse verdiene ser man at Backward Digit span, som måler verbalt arbeidsminne, ser ut til å være en svært viktig prediktor for Ordproblemer. Dette kan ha en sammenheng med at både Backward Digit span og Ordproblemer er muntlige kartleggingstester hvor barnet må huske og bearbeide muntlig presentert input. Det som er viktig å bemerke seg, er at Ordproblemer-testen også er ment til å kartlegge arbeidsminnet. Barn må altså både ha et godt arbeidsminne, gode aritmetiske ferdigheter i tillegg til språklige kunnskaper for å få bra score. Siden testene blir gjennomført muntlig, ser det ut til at det verbale arbeidsminne spiller en viktig rolle under utregning av muntlig presenterte

ordproblemer, noe som også gjenspeiles i hvilke deler av arbeidsminne testen er ment til å måle. Gjennom bruken av testen Backward Digit span kan man måle både den Central executive og det verbale arbeidsminne (Pickering, 2006). Siden man antar at det verbale arbeidsminne inneholder et midlertidig fonologisk lager og er i stand til å holde på verbal og akustisk informasjon ved å benytte seg av midlertidig lagring, gjør den denne delkomponenten spesielt godt egnet til å oppbevare sekvenser av verbal informasjon (Baddeley, 2000), når oppgaver som ordproblemer skal løses. Ordproblemene er bygget opp slik at man får presentert en setning som inneholder et regnestykke med blant annet både addisjon og subtraksjon. Barnet må velge riktig regnemetode, holde på informasjon om mellomregningen, samtidig som det lytter videre på ordproblemet som blir presentert av testlederen. Siden oppgavene er nokså komplekse, kan man anta at den sentral utøvende bidrar under bearbeidingen av ordproblemene. Som tidligere nevnt antar man at den Central executive skal holde på mellomregninger i arbeidsminnet mens man foretar en annen utregning (Gilmore, Göbel, & Inglis, 2018).

Som tidligere nevnt er det uenigheter om hvorvidt det verbale arbeidsminne, spiller en større rolle enn det visuospatiale arbeidsminne for matematikk (Cragg, et.al, 2017; Gilmore, Göbel, & Inglis, 2018). Ut ifra resultatene i den foreliggende studien kan det se ut som Backward Digit span er den testen som viste seg å være sterkt korrelert til de aritmetikk-testene som er inkludert i denne studien. Backward Digit span tester i størst grad den Central executive, men det verbale arbeidsminnet er også involvert under denne testen siden den er administrert verbalt (Pickering, 2006). Generelt ser det ut til at arbeidsminne ved 5 årsalder kan predikere aritmetiske ferdigheter i førsteklasse. Dette finnes det også evidens for i tidligere forskning, hvor det verbale arbeidsminne ved 6 årsalder har vist seg å kunne predikere matematiske ferdigheter 1 år frem i tid (Göbel, 2016).

5.1 Vurdering av studiens validitet og reliabilitet

I dette kapittelet vil resultatene og slutninger bli drøftet med hensyn på validitets- og reliabilitetskriteriene i henhold til Cook og Campells (1979) og Heale og Twycross's (2015) validitetssystem, i tillegg til tidligere teori og empiri. Siden de ulike validitetstypene i Cook og Campells (1979) system påvirker hverandre, for eksempel ved at valg av metode kan føre til at en validitetstype styrkes, samtidig som en annen kan svekkes (Lund, 2002b), skal validitet benyttes som redskap til å finne den best mulige tilnærmingen til sannheten for antakelser (Cook & Campbell, 1979). Reliabilitet skal, som tidligere nevnt, belyse variabelenes indre konsistens (Lund, 2002b; Heale & Twycross, 2015).

5.1.1 Statistisk validitet og reliabilitet

Statistisk validitet handler som tidligere nevnt om å se på om korrelasjonen mellom den uavhengige og den avhengige variabelen er statistisk signifikant (Lund, 2002b). Man kan benytte seg av statistisk '*kontroll*' for å veie opp for svak indre validitet (Kleven T. A., 2002b). Statistisk kontroll benyttes for å øke den statistiske validiteten gjennom å kontrollere selve analysen for eventuelle alternative hypoteser. I denne studien har det derfor blitt kontrollert for både alder og kjønn for å ekskludere trusselen fra *tredje variabler*.

Trusler mot statistisk validitet og reliabilitet er, som tidligere nevnt, målingsfeil (Lund, 2002b) og lav statistisk styrke (Cook & Campbell, 1979). I denne undersøkelsen er det blitt gjort flere valg som styrker den statistiske validiteten.

Målingsfeil

Målingsfeil kan minke den statistiske styrken i tillegg til å redusere begrepsvaliditeten og/eller den indre- og ytre validiteten. Jo flere målingsfeil, jo større usikkerhet blir det rundt slutninger og generaliseringer (Lund, 2002b). Målingsfeil kan oppstå dersom man har ustandardiserte tester eller testsituasjonen. Dersom den avhengige variabelen får en økning i variasjonen til målingsfeil, vil dette redusere variabelens reliabilitet.

Målingsfeil kan altså oppstå under selve kartleggingen ved at det er eventuelle feil ved selve kartleggingstesten, dersom testlederen gjør feil ved administreringen av testene eller feil som gjøres når det innsamlede datamaterialet skal plottes (Lund, 2002b). Mange av testlederne i

NumLit prosjektet er masterstudenter fra spesialpedagogikk, noe som har gitt testlederne kompetanse og ferdigheter til å kommunisere med barn. Alle testlederne har fått en grundig opplæring i testbatteriet som også inneholder instruksjoner for hvordan hver enkelt test skal gjennomføres. Til tross for dette kan det allikevel oppstå feil som gir en økt fare for målefeil. Noen av testene er blitt digitaliserte og krever dermed mindre administrering og skåring av testlederen. Dette bidrar til å minke målefeil. Videre gir testene i den foreliggende studien gode verdier for reliabilitet, noe som tyder på at eventuelle målefeil ikke ansees som en trussel mot den statistiske styrken (Shadish, Cook, & Campbell, 2002).

Statistisk Homogenitet

Homogenitet er en faktor som benyttes til å evaluere reliabilitet i en studie ved å undersøke signifikansnivået til måleinstrumentene, dette gjøres her ved hjelp av Chronbach's α (Heale & Twycross, 2015). Tidligere i studien ble det vist en tabell med de ulike variablene og deres reliabilitetskoeffisient. Reliabiliteten til hver enkelt variabel ble der målt og vurdert ved hjelp av reliabilitetskoeffisienten Chronbach's α . Samtlige variabler ga verdier $\alpha > .750$, noe som peker på høy indre konsistens og dermed også god reliabilitet. Samtidig bør man betenke at Chronbach's α er sensitiv for antall tester som er inkludert i en studie (Field, 2018). Et høyere antall tester vil kunne gi høyere α -verdier, derfor bør disse verdiene vurderes med omhu. Siden variablene gir tilfredsstillende verdier og det kun er brukt åtte variabler i denne studien, gir det grunnlag for å anta at den indre konsistensen ikke er en trussel for studiens reliabilitet. Siden testene er reliable, er de heller ingen trussel mot begrepsvaliditet (Lund, 2002b).

En annen trussel mot statistisk validitet kan oppstå ved signifikanstesting ved at brudd på statistiske forutsetninger eller lav statistisk styrke oppstår, dette kalles henholdsvis *Type-1 feil* og *Type-2 feil*. Type-1 feil kan oppstå ved at man forkaster en sann nullhypotese, og Type 2-feil oppstår ved å akseptere en gal nullhypotese (Lund, 2002b). Type-2 feil kan oppstå ved lav statistisk styrke, sannsynligheten for å begå en Type-2 feil øker når utvalgsstørrelsen er liten og når signifikansverdiene for α -nivået er satt lavt (Cook & Campbell, 1979). Hvor mye styrke det ligger i en statistisk test kan variere. Utvalgsstørrelsen N viste seg å være tilstrekkelig stor ut ifra de presenterte kriteriene og ansees derfor ikke til å svekke den statistiske validiteten i henhold til å begå Type-2 feil. Videre følger signifikansverdiene et α -nivå ($\alpha > .7$) som blir ansett som akseptabel i forskningssammenheng (de Vaus, 2014).

Multipel regresjonsanalyse og beta-verdier

Videre brukes det en type multipel regresjon i denne studien. Multipel regresjonsanalyse ser på sammenhengen mellom en avhengig variabel og flere uavhengige variabler (Kleven T. A., 2002b). Denne regresjonstypen kan bidra til å bedre den indre validiteten ved ikke-eksperimentelle design. I en hierarkisk multipel regresjonsmodell velger man en gitt rekkefølge for X-variablene. I en slik regresjonsmodell forklarer variablene i det første steget hvor mye den står for av variasjon i Y-variabelen, og variablene i det andre steget forklarer en mengde variasjon som ikke ble forklart i det første steget.

Videre har beta-verdien blitt kommentert i analysen for å belyse hvilken variabel som har den største uavhengige innvirkningen på den avhengige variabelen. Beta-verdiene er *standardiserte regresjonskoeffisienter*, og er dermed ansett som en sammenliknbar størrelse (Kleven, 2002; de Vaus, 2014).

5.1.2 Indre validitet

Indre validitet refererer som sagt til validiteten som påstander kan basere seg på, og hvorvidt det er en kausal sammenheng mellom variablene (Lund, 2002b). Siden denne studien har et ikke-eksperimentelle design, vil det i liten grad være mulig å trekke konkluderende slutninger (Kleven, 2002b). Når validitet skal undersøkes bør man derfor alltid ordlegge seg med vaghet, dette fordi vi aldri kan vite hva som faktisk er sant (Cook & Campbell, 1979). For å styrke den indre validiteten bør man derfor se etter alternative mulige som kan ha påvirket resultatene. I de påfølgende avsnittene vil indre validitet bli vurdert ut ifra ulike forhold som presenteres av Lund (2002b).

Retningsproblemet

En trussel mot indre validitet i ikke-eksperimentelle design er det som blir kalt *retningsproblemet* (Kleven, 2002b; Lund, 2002b). Retningsproblemet tar utgangspunkt i en årsak-virkning-problematikk. Ut ifra beta-verdier kan man for eksempel se om to variabler viser en statistisk sammenheng og ut ifra dette trekke ut informasjon om hvorvidt den ene variabelen kan predikere den andre variabelen, men det vil allikevel ikke være mulig å trekke konkluderende slutninger om årsakssammenhenger. Man kan kun trekke slutninger om en eventuell felles variasjon i variablene, dersom man har kontrollert for tredjevariabler. Det har

derfor blitt kontrollert for statistisk mulige tredjevariabler, for å vurdere og eliminere alternative tolkninger. Som nevnt i avsnittet over, har denne studien benyttet en hierarkisk multippel regresjonsanalyse for å belyse statistiske sammenhenger. Ut ifra hvor sterke eller svake disse sammenhengene er kan man vurdere tendenser i henhold til hypoteser.

Instrumentering

En annen trussel som nevnes av Lund (2002b) er *instrumentering*. Denne trusselen er ment til å belyse hvorvidt sider ved testene kan føre til *golv-* eller *tak-effekt*. Som beskrevet i analysedelen kunne man se tegn på både tak- og gulv-effekt i fordelingene. Testbatteriet som ble benyttet er basert på standardiserte tester, det antas derfor ikke at dette er en trussel mot validitet. Det kan derimot tenkes at testene i noen tilfeller bare var for vanskelige og i andre tilfeller for enkle.

Modning

Modning kan også ansees som en trussel mot indre validitet. Siden dataene som benyttes er innhentet med et års mellomrom, kan man anta at barna har hatt en modningsperiode mellom målingene. Samtidig analyseres det ulike variabler fra de gitte testtidspunktene for å belyse hvorvidt en variabel kan påvirke en annen. Modning ansees derfor ikke som en trussel.

Andre trusler

Andre trusler som *seleksjon*, *fracfall*, *historie* og *atypisk kontrollgruppeatferd* (Lund, 2002b), ansees ikke som relevante blant annet fordi det kun er blitt inkludert barn med data fra både barnehage og førsteklasse, studien har ingen kontrollgruppe, seleksjon og historie ansees ikke som relevante da utvalget er randomisert og er ment til å analysere et representativt utvalg av en populasjon som denne forskningen har tenkt til å si noe om.

5.1.3 Begrepsvaliditet

Begrepsvaliditet omhandler som tidligere nevnt hvorvidt det er samsvar mellom begreps definisjon og hvordan man bruker de operasjonaliserte variabelbegrepene (Kleven A. T., 2002a). Trusler mot denne validitetstypen varierer mellom hva slags undersøkelse man gjennomfører. Det ser ut som at man i dette fagfeltet knytter ulike begrep til omtrent lik eller lik definisjon. Det at det mangler enighet om hvilke begrep man bruker og hvordan disse begrepene defineres kan være en trussel mot begrepsvaliditet. En annen trussel mot begrepsvaliditet er, som tidligere nevnt, systematiske målefeil (Kleven, 2002a; Lund, 2002b). Slike målefeil kan føre til skjevhet i målingene. For å finne bevis som taler for god begrepsvaliditet kan man som tidligere nevnt benytte tre argumenter; homogenitet, konvergens og teoretiske bevis (Heale & Twycross, 2015).

Begrensninger ved kartleggingstestene for aritmetikktestene

TOBANS addisjon og subtraksjon er non-verbale tester. For å kunne løse disse oppgavene må barnet forstå tallsymbolene for tall fra 0-9. Videre er testen tidsbegrenset, noe som gjør at testen i tillegg til å måle nøyaktighet også måler regneflyt. Testene måler altså ikke bare én ferdighet og kan dermed ikke ansees som homogene. Oppgavene øker ikke i vanskelighetsgrad.

Hoderegning addisjon og subtraksjon måler aritmetiske ferdigheter, ved verbale representasjoner av tall. Vanskelighetsgraden er økende ved at tallene i regnestykkene blir større. Men ordlyden for hver oppgave forblir lik. Siden testene ser ut til å kun måle én ferdighet, kan de ansees som homogene.

Siden studien benytter to ulike deltester for å måle henholdsvis addisjon og subtraksjon, kan man se om de ulike regneartene konvergerer. Både addisjons- og subtraksjonsoppgavene er sterkt korrelerte med $p < .01$, og kan dermed være et bevis for konvergens.

Under testen Ordproblemer blir også tiden målt, men denne påvirker ikke resultatene, testen måler dermed ikke regneflyt. I denne testen er vanskelighetsgraden økende. Testen måler en kombinasjon av subtraksjon og addisjon, og når oppgavene blir vanskelige nok, måler testen også enkel divisjon. De aritmetiske oppgavene i denne testen er altså mer krevende enn i TOBANS. I tillegg kreves det at man forstår ordene og de matematiske begrepene i oppgavene. Testen ser dermed ut til å måle flere enn én ferdighet og kan dermed ikke ansees som homogen.

Begrensninger ved kartleggingstestene for arbeidsminne

Studien inkluderer tre arbeidsminnetester, Maze memory, Backward Digit span og Listening recall. Det har blitt valgt tre ulike arbeidsminnetester da de tre testene måler arbeidsminne på ulike måter. Ved å velge ulike tester, som måler ulike deler av arbeidsminnet, kan man skape et mer nyansert bilde av hvilke deler av arbeidsminnet som er inkludert i de ulike aritmetiske beregningene.

Som tidligere beskrevet krever Backward Digit span at man lagrer og samtidig mentalt prosesserer og omstrukturerer informasjonen. Noen antar derfor at delkomponentene den Central executive og det verbale arbeidsminne spiller en viktig rolle når tall skal gjentas i omvendt rekkefølge (Wilde, Strauss, & Tulsy, 2004). Det kan derfor være vanskelig å si noe konkret om hvilke av de delene som faktisk måles ved denne deltesten (Pickering, 2006). Men siden det tilsynelatende er nokså sikkert at deltesten måler arbeidsminnet, kan man anta at testens homogenitet er valid. Et teoretisk bevis som står i tråd med presenterte forskningsteorier er at det verbale arbeidsminne kan måles ved å gjenta tallsekvenser (Baddeley, 2000), og at det verbale arbeidsminne er sterkt relatert til aritmetiske ferdigheter (Cragg, et.al, 2017; Gilmore, Göbel, & Inglis, 2018).

Maze memory er en non-verbal test som brukes for å måle det visuospatiale arbeidsminnet. Denne ble benyttet grunnet i at forskning viser til at minnet for visuelt presenterte stimuli, i barnehagealder, ser ut til å hovedsakelig være avhengig av det visuospatiale arbeidsminne (Camos & Barrouillet, 2018). Siden resultatene i denne studien viser en korrelasjon som står i tråd med tidligere presenterte forskningsteorier om at arbeidsminne kan predikere aritmetiske ferdigheter (Gilmore et al., 2018), kan dette ansees som et teoretisk bevis for validitet. Slik denne studien ser det er det ingen umiddelbare begrensninger eller andre faktorer som måles ved denne testen. Homogenitet for denne deltesten er derfor sikret ved at instrumentet måler kun én ferdighet.

En begrensning ved denne studien er at resultatene som ble benyttet for Listening recall er målt i første klasse, disse resultatene er dermed begrenset i forhold til at de ikke måler det studien egentlig spør etter. Målingene ble allikevel tatt med i analysen for å se om Listening recall eventuelt forklarer en signifikant mengde varians etter å ha beregnet forklaringsverdien til arbeidsminne testene ved barnehagealder. Listening recall skal i all hovedsak måle den Central executive, det kan imidlertid ikke utelukkes at andre delkomponenter av arbeidsminne ikke er delaktige (Pickering, 2006). Men siden det hittil ikke ligger til grunn noen antakelser for at deltesten måler mer enn én ferdighet, antas det at testen er homogen.

Begrensninger ved kartleggingstestene for flytende-intelligens

Ravens CPM ble tidligere beskrevet som en non-verbal test for kognitive ferdigheter. Men denne testen ansees i litteraturen ikke bare til å være et mål på non-verbal intelligens, man anser den også som en visuell-romlig test som kan måle non-verbal resonering (Kyttälä & Lehto, 2008). Man kan med andre ord også interpretare resultatene til å måle det visuospatiale arbeidsminne. Siden det er vanskelig å kartlegge kognitive prosesser fordi konseptet er uklart og vagt (Camos & Barrouillet, 2018) og siden arbeidsminne og flytende-intelligens ansees som kognitive evner, antas det at testen har en tilfredsstillende homogenitet.

Ser man på korrelasjonen mellom det visuospatiale arbeidsminne og Ravens, kan man se om disse konvergerer. I den bivariante analysen kan man se at disse to testen er sterkt korrelerte ($p < .01$). Dersom en eventuell antakelse om at Ravens CPM også måler det visuospatiale arbeidsminne er sann, ser antakelsen om konvergens ut til å være tilfredsstillt.

5.1.4 Kriterievaliditet

Kriterievaliditet har som mål belyse hvorvidt resultatene i en studie kan overføres til andre individer i samfunnet (Field, 2018). Dette står i tråd med Cook og Campell (1979) definisjon av ytre validitet som benyttes for å vise til en antatt validitet, som trekker konklusjoner basert på en antakelse om generaliserbarhet til en kausal relasjon på tvers av ulike grupper individer, omgivelser og tidspunkt (Cook & Campbell, 1979). Man oppnår dermed god ytre validitet i en studie, dersom den gjør det mulig å trekke generaliserende slutninger (Lund, 2002b).

Heale og Twycross (2015) deler kriterievaliditet som sagt inn i tre kriterier som skal tilfredstilles for å ha god validitet. Divergent validitet ansees ikke som relevant for denne studien, da studien kun har inkludert variabler og måleinstrumenter som ansees som relevante for besvarelsen av forskningsspørsmålene.

Konvergens validitet

Konvergens validitet innen kriterievaliditet skal vise at et måleinstrument er sterkt korrelert til andre instrumenter som måler den samme variabelen (Heale & Twycross, 2015). I denne studien har det blitt benyttet flere måleinstrument for å kartlegge aritmetiske ferdigheter og arbeidsminne. I tabellen for bivariat analyse kan man se korrelasjonene, gitt med Pearson's r , mellom de ulike variablene som er inkludert i denne studien. Samtlige variabler i denne studien viste en signifikant korrelasjon til hverandre. Signifikansnivået varierte noe, men for de fleste korrelasjonene var $p < .01$. I tre tilfeller var korrelasjonen noe svakere, med $p < .05$. Den ene korrelasjonen er mellom aritmetikktestene TOBANS addisjon og hoderegning subtraksjon. Det var i midlertid til å anta at korrelasjonen mellom en addisjonsoppgave og en subtraksjonsoppgave skulle være noe mindre, da dette er to ulike regnemåter, i tillegg til at barna har mer opplæring i addisjon enn subtraksjon ved testtidspunktet. Den andre er mellom Maze memory (målt i barnehage) og Listening recall (målt i første klasse). Trolig er denne korrelasjonen noe lavere da testene måler ulike deler av arbeidsminnet og dermed også ulike ferdighetsområder. Den tredje er korrelasjonene mellom hoderegning addisjon og ordproblemer, her er det mulig at måten oppgavene skal løses på spiller inn på resultatene. Begge er muntlige tester hvor man skal bearbeide informasjonen som gis i hukommelsen, men siden ordproblemen ikke består av konkrete regnestykker hvor regnemethoden er oppgitt, er kanskje dette en faktor som gir en noe mindre korrelasjon. Konvergens validitet anses her som god, da samtlige korrelasjoner er innenfor et signifikant nivå ($p < .05$).

Predikerende validitet

Predikerende validitet krever som sagt at det bør være en høy korrelasjon med fremtidige prestasjoner (Heale & Twycross, 2015). Siden denne studien skal belyse hvorvidt en ferdighet i barnehagen kan predikere en annen ferdighet ved et senere tidspunkt, ser dette ut til å være et relevant mål for validitet i denne studien. Også her kan man blant annet i den Bivariate analysen se, at ferdighetene målt i 5 årsalder er sterkt korrelerte, med $p < .01$, til aritmetiske ferdigheter målt i førsteklasse. Videre viser også den Hierarkiske multiple analysen sterke korrelasjoner mellom de ulike delene i arbeidsminnet og aritmetiskeferdigheter etter å ha kontrollert for alder, kjønn og flytende-intelligens. Predikerende validitet ansees også som tilfredsstillt ut i fra signifikante korrelasjoner mellom ferdighetene som ble målt i barnehagen, og ferdighetene i første klasse.

Individhomogenitet og ikke-representative individutvalg

Siden kriterievaliditet og ytre validitet har som mål belyse hvorvidt resultatene i en studie kan begrepsoperasjonaliseres, velges det her også å belyse slutninger som trekkes til en definert populasjonsgruppe for å styrke den ytre validiteten. For å gjøre dette presenterer Lund (2002b) to kategorier; *individhomogenitet* og *ikke-representative individutvalg*. En trussel mot ytre validitet er dersom utvalget som deltar i studien ikke er representativ for populasjonen ved at det er for ensartet (Lund, 2002b). Dersom utvalget baserer seg på en veldig snever og spesifikk gruppe, vil det være vanskelig å trekke slutninger som kan overføres til andre individgrupper. Egenskapene til individutvalget i denne studien har som mål å belyse den normale variasjonen blant enspråklige norske barn uten lærevansker, ved 5 år og i første klasse. Slike homogene egenskaper bidrar til å kunne trekke generaliserbare slutninger om den populasjonen som beskrives i individutvalget, men man kan ut ifra dette ikke trekke slutninger for hele populasjonen. Eksempelvis vil funnen ikke kunne generaliseres til andre aldersklasser eller barn med flerspråklig bakgrunn.

5.1.5 Spesialpedagogiske implikasjoner

Arbeidsminnet er et komplekst system og forskning er ikke entydig hvorvidt det er mulig å forbedre arbeidsminnets kapasitet ved hjelp av spesifikke treningsmetoder (Melby-Lervåg & Hulme, 2013). Ved å tilegne seg kunnskap om hvilke av arbeidsminnets delkomponenter som er involvert i enkle aritmetiske prosesser, gir det mulighet til å identifisere barn som er risikogruppen tidlig.

Denne studien viser en sammenheng mellom barns arbeidsminne i barnehage og deres aritmetiske ferdigheter i første klasse. Det viste seg også at det varierte hvilke delkomponenter ved arbeidsminne som hadde betydning for å løse de ulike aritmetiske oppgavene. Denne studien gir på ingen måte kunnskap om hva som kan være grunnen til en eventuell vanske innen aritmetiske ferdigheter, men den kan bidra med å vise til hvilke indikatorer man kan benytte for å identifisere barn i risikogruppen.

Denne studien viste at arbeidsminnevariablene Backward Digit span og Maze memory, som ble målt i barnehage, predikerer aritmetiske ferdigheter i første klasse. Tidlig identifisering av lærevansker eller risikogrupper er gunstig i forhold til at man ved et tidlig tidspunkt får

hjulpet disse barna. Ved å vite hvilken del av arbeidsminnet som er svekket, kan man eventuelt benytte andre type oppgaver, som involverer andre deler av arbeidsminnet, eller man kan finne strategier som kan kompensere for svekkelsen. Eventuelle intervensjoner bør være forskningsbasert og målrettet for å kunne hjelpe barnet på en god måte. Det at for eksempel arbeidsminnetesten Maze memory er sterk prediktor for TOBANS addisjon, kan gi en indikasjon på hvilke indikatorer man kan rette oppmerksomheten mot for å utvikle en forskningsbasert intervensjon.

5.1.6 Veien videre

Mange av grunntankene som benyttes i spesialpedagogikken ser ut til å basere seg på relativt gamle teorier, som tilsynelatende har blitt lite utfordret i nyere tid. Det kan derfor være av betydning at fremtidig forskning både utfordrer og videreutvikler eksisterende teorier, men også prøve å finne nye tilnærminger. Blant annet peker noen studier på at trening av arbeidsminne ikke har stor overføringseffekt (Melby-Lervåg & Hulme, 2013), kan det allikevel tenkes at trening av strategier og eller trening av det å huske elementer ved matematiske oppgaver har en effekt i seg selv, og dette bør dermed forskes mer på.

En annen interessant tilnærming kan være å implementere variabler for ulike strategier som kan benyttes for å løse aritmetiske oppgaver. Siden forskning viser at prosessene i arbeidsminnet også inkluderer bruk av ulike strategier som telling, tilbakekalling av aritmetiske fakta og/eller semantiske strategier (Aunio & Räsänen, 2015; von Aster & Shalev, 2007; Campbell, 2015), kunne det ha vært interessant om noen av disse strategiene allerede er etablert i barnehagealder, eller eventuelt når i skoleforløpet norske barn benytter seg av de ulike strategiene.

5.2 Oppsummering

For det spesialpedagogiske fagfeltet er det essensielt å ha bred kunnskap om hvilke forhold som kan ha innflytelse på læring. Som beskrevet i denne studien, ser det ut ifra forskning, ut til at arbeidsminne og generelle kognitive evner spiller en viktig rolle for barns skolefaglige utvikling og prestasjoner.

Siden de generelle kognitive ferdighetene og arbeidsminnets kapasitet legger grunnlaget for den faglige utviklingen, også på tvers av fag, er det viktig å kunne utrede barn så tidlig som mulig, slik at også de med store vansker kan ha muligheten til å utvikle kunnskap til å leve et selvstendig liv. For å kunne møte barna der de er i læringsprosessen og ut ifra nivået kunne gi tilpasset opplæring som skal sikre et tilfredsstillende læringsutbytte, er det essensielt å ha kunnskap om forhold som har innflytelse på tidlig læring.

Denne studien viser at arbeidsminnets kapasitet ved omtrent 5 år (siste året i barnehage) ser ut til å kunne predikere aritmetiske ferdigheter i førsteklasse, også når det blir kontrollert for non-verbal intelligens, alder og kjønn, da arbeidsminne viser en unik statistisk signifikant forklaringsverdi for aritmetiske ferdigheter. Videre visste også korrelasjonskoeffisientene i den bivarierte korrelasjonen en tydelig sammenheng mellom de ulike dekomponentene ved arbeidsminne og de ulike aritmetiske deltestene.

Sammenhengen mellom de ulike variablene viser at ulike deler ved arbeidsminne som er involvert i ulike typer aritmetiske oppgaver, noe som ser ut til å kunne begrunnes med hvordan testen for aritmetiske ferdigheter er satt opp, og hvilke deler av arbeidsminne, arbeidsminne-testene måler. Tidligere forskning har som tidligere nevnt funnet evidens for at deler av arbeidsminne kan predikere matematiske ferdigheter 1 år fram i tid (Göbel, 2016), noe som ser ut til å samsvare med funnene i denne studien om at arbeidsminne kan predikere aritmetiske ferdigheter ett år fram i tid.

Denne studien har kun sett på et lite utvalg av eksisterende kartleggingstester som kan forklare variasjonen knyttet til aritmetiske ferdigheter og arbeidsminne. Ved å benytte flere ulike kartleggingstester som måler ulike sider ved en ferdighet, vil man kunne få økt innsikt i hvorfor noen har vansker, og kunnskap som kan bidra til å forebygge eller minke graden av en vanske. Videre bidrar økt kunnskap også til at nye kartleggingsverktøy og intervensjoner blir utarbeidet.

«*Kan arbeidsminne ved 5 år alder predikere aritmetiske ferdigheter i førsteklasse?*» er spørsmålet som ligger til grunn for denne studien. Spørsmålet er ment til å belyse om de ulike delene av arbeidsminne, målt før formell skolegang, kunne predikere aritmetiske ferdigheter i førsteklasse. Tidligere forskning indikerer at arbeidsminne er involvert i ulike aritmetiske prosesser (Hubber et al., 2014; Camos & Barrouillet, 2018; Geary et al., 2004; Gilmore et al., 2018; Cragg et al., 2017; Göbel, 2016). I den foreliggende studien ble det funnet tilsvarende resultater. Resultatene viste også at det visuospatiale arbeidsminne er sterkt korrelert til aritmetiske oppgaver som presenteres skriftlig, mens det testen for det verbale arbeidsminne og den Central executive viser sterkest korrelasjon til muntlige presenterte oppgaver.

Litteraturliste

- Aunio, P., & Räsänen, P. (2015). Core numerical skills for learning mathematics in children aged five to eight years - a working model for educators. *European Early Childhood Education*.
- Baddeley, A. (2000, 11). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, s. 7.
- Camos, V., & Barrouillet, P. (2018). *Working Memory in Development*. Oxon: Routledge.
- Campbell, J. I. (2015). How Abstract is Arithmetic? I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 140-157). Oxford: Oxford University Press .
- Cattell, R. B. (1971). *Abilities: Their growth, structure, and action*. Boston: Houghton Mifflin.
- Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019, 09 23). The matrix reasoning item bank (MaRs-IB): novel, open-access abstract reasoning items for adolescents and adults. *Royal Society open Science*, ss. 1-13.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design & analysis issues for*. Boston: Houghton Mifflin Company.
- Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, ss. 12-26.
- Dalen, M. (2011). *Intervju som forskningsmetode - en kvalitativ tilværmning*. Oslo: Universitetsforlaget.
- de Vaus, D. (2014). *Surveys in Social Research*. London: Routledge.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, ss. 1-42.
- Dehn, M. J. (2008). *Working Memory and Academic Learning*. New Jersey: John Wiley & Sons, Inc.
- Diez, D. M., Barr, C. D., & Cetinkaya-Rundel, M. (2017). *Open Intro Statistics. Third Edition*. OpenIntro.

- Field, A. (2018). *Discovering statistics using IBM SPSS statistics (5. utg.)*. SAGE edge.
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., & Schatschneider, C. (2006). The Cognitive Correlates of Third-Grade Skill in Arithmetic, Algorithmic Computation, and Arithmetic Word Problems. *Journal of Educational Psychology*, ss. 29–43.
- Gathercole, S., & Pickering, S. (2001). *Working Memory Test Battery for Children (WMTB-C)*. Psychological Corporation.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology* 88, ss. 121-151.
- Gilmore, C., Göbel, S. M., & Inglis, M. (2018). *An Introduction to Mathematical Cognition*. Routledge.
- Grue, J. (2015). *Teori i praksis, Analysestrategier i akademisk arbeid*. Bergen: Fagbokforlaget.
- Göbel, S. M. (2015). Number Processing and Arithmetic in Children and Adults with Reading Difficulties. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 696-720). Oxford: Oxford Library of Psychology.
- Göbel, S. M. (2016). Number Processing and Arithmetic in Children and Adults with Reading Difficulties. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Conition* (ss. 696-720). Oxford : University Press.
- Heale, R., & Twycross, A. (2015, 07). Validity and reliability in quantitative studies. *Evid Based Nurs (volume 18, number 3)*, ss. 66-67.
- Helland-Riise, F., & Martinussen, M. (2017). Måleegenskaper ved de norske versjonene av Ravens matriser [Standard Progressive Matrices (SPM)/Coloured Progressive Matrices (CPM)]. *PsykTestBarn*, 2:2.
- Hindry, M. (2011). *Arithmetics*. London: Springer-Verlag.
- Hitch, G. J., & McLean, J. F. (1999). Working Memory Impairments in Children with Specific Arithmetic Learning Difficulties. *Journal of Experimental Child Psychology*, ss. 240–260.
- Horn, J. L. (1968). Organization of abilities and the development og intelligence. *Psychological Review*, ss. 242-259.
- Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014, 04 04). Predicting first-grade mathematics achievement: the contributions of domain-general cognitive abilities,

- nonverbal numbersense, and early numbercompetence. *Frontiers in Psychology*, ss. 1-18.
- Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *The Quarterly Journal of Experimental Psychology*, 67, ss. 936-954.
- Hudson, T. (1983, 02). Correspondences and Numerical Differences between Disjoint Sets. *Society for Research in Child Development*, ss. 84-90.
- Hulme, C., Brigstocke, S., & Moll, K. (2016). *Test of Basic Arithmetic and Numeracy Skills*. Oxford University Press .
- IBM, C. (2019). IBM SPSS Statistics for Windows (Version 26). NY: IBM Corp.
- Kadosh, R. C., & Dowker, A. (2016). *The Oxford Handbook of Numerical Cognition*. Oxford: Oxford University Press.
- Karbach, J., & Kray, J. (2016). Executive Functions. I T. Strobach, & J. Karbach, *Cognitive Training* (ss. 93-106). Sveits: Springer.
- Kaufmann, L., Kucian, K., & von Aster, M. (2015). Development of the Numerical Brain. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 485-501). Oxford: Oxford Library of Psychology.
- Kleven, A. T. (2002a). Begrepsoperasjonalisering. I T. Lund (red.), *Innføring i forskningsmetodologi* (ss. 141-184). Oslo: Unipub.
- Kleven, T. A. (2002b). Ikke-eksperimentelle design. I T. Lund (red.), *Innføring i forskningsmetodologi* (ss. 265-286). Oslo: Unipub.
- Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory The role of visuospatial working memory. *European Journal of Psychology of Education*, ss. 77-94.
- Kyttälä, M., Aunio, P., Lepola, J., & Hautamäki, J. (2014). The role of the working memory and language skills in the prediction of word problem solving in 4- to 7-year-old children. *Educational Psychology*, ss. 674–696.
- Lefevre, J. A., Wells, E., & Sowinski, C. (2015). Individual differences in basic arithmetical processes. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 895-914). Oxford University Press.
- Lindemann, O., & Fischer, M. H. (2015). Cognitive Foundations of Human Number Representation and Mental Arithmetic. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 35-44). Great Britain : Oxford University Press.

- Lovdata. (2020, 05 13). Hentet fra Lov om grunnskolen og den vidaregåande opplæringa (opplæringslova): https://lovdata.no/dokument/NL/lov/1998-07-17-61/KAPITTEL_1#%C2%A71-1
- Lund, T. (2002a). Innføring i forskningsmetode. I T. Lund (red.), *Innføring i forskningsmetodologi* (ss. 9-18). Oslo: Unipub.
- Lund, T. (2002b). Metodologiske prinsipper og referanserammer . I T. Lund (red.), *Innføring i forskningsmetodologi* (ss. 79-124). Oslo: Unipub.
- Lund, T. (2002c). Generaliseringsproblematikk. I T. Lund (red.), *Innføring i forskningsmetodologi* (ss. 125-140). Oslo: Unipub.
- Løvås, G. G. (2013). *Statistikk for universiteter og høyskoler 3. utgave*. Oslo: Universitetsforlaget.
- Melby-Lervåg, M., & Hulme, C. (2013). Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology* 49, ss. 270–291.
- Mononen, R. (2018, 11 27). *Fakta om matematikkvansker* . Hentet fra Universitetet i Oslo, Det utdanningsvitenskapelige fakultet : <https://www.uv.uio.no/tjenester/kunnskap/matematikk-i-spesialundervisningen/fakta%20om%20matematikkvansker>
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2017). *Introduction to the Practice of Statistics 9th edition*. New York: Macmillan education.
- NESH. (2016, 04). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Hentet fra NESH: https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf
- NSD – Norsk senter for forskningsdata. (2018, 11 27). Hentet fra Regjeringen.no: <https://www.regjeringen.no/no/dep/kd/org/etater-og-virksomheter/tilknyttede-virksomheter/nsd--norsk-senter-for-forskningsdata/id440384/>
- NumLit. (2020, 04 7). Hentet fra UiO Institutt for pedagogikk: <https://www.uv.uio.no/iped/forskning/prosjekter/numlit-development-of-numeracy-and-literacy-in-chi/>
- Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A Meta-Analysis of Mathematics and Working Memory: Moderating Effects of Working Memory Domain, Type of Mathematics Skills, and Sample Characteristics. *Journal of Educational Psychology*, ss. 455-473.
- Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, ss. 147–153.

- Pickering, S. J. (2006). *Assessment of Working Memory*. Academic Press.
- Price, G. R., & Ansari, D. (2013). Dyscalculia: Characteristics, Causes, and Treatments. *Numeracy*, 6(1).
- Rossiter, J. R. (2011). *Measurement for the Social*. New York: Springer.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston New York: Houghton Mifflin Company.
- St.meld. nr. 030 (2003-2004)*. (2018, 11 27). Hentet fra Regjeringen.no : <https://www.regjeringen.no/no/dokumenter/stmeld-nr-030-2003-2004-/id404433/sec4>
- Udir. (2018, 11 27). *Matematikksenteret, Nasjonalt senter for matematikk i opplæringen*. Hentet fra Udir: https://www.udir.no/globalassets/filer/laring-trivsel/grunnleggende-ferdigheter/regning_i_alle_fag_130316.pdf
- von Aster , M. G., & Shaley, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology*, ss. 868–873.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children (4)*. San Antonio, TX: Psychological.
- Wilde, N. J., Strauss, E., & Tulskey, D. S. (2004). Memory Span on the Wechsler Scales. *Journal of Clinical and Experimental Neuropsychology*, ss. 539–549.
- Wright, R. J., Ellemor-Collins, D., & Tabor, P. D. (2012). *Developing Number Knowledge*. London: SAGE Publications.
- Zamarian, L., & Delazer, M. (2015). Arithmetic Learning in Adults: Evidence from Brain Imaging. I R. Cohen Kadosh, & A. Dowker, *The Oxford Handbook of Numerical Cognition* (ss. 837-866). Oxford: Oxford Library of Psychology.

