Sammenhengen mellom alder og aritmetiske ferdigheter

En kvantitativ undersøkelse av norske enspråklige andreklassinger

Cecilie Pavljuk Alver

Masteroppgave i spesialpedagogikk
Institutt for spesialpedagogikk
Det utdanningsvitenskaplige fakultet

UNIVERSITETET I OSLO

Vår 2019

Sammenhengen mellom alder og aritmetiske ferdigheter

En kvantitativ undersøkelse av norske enspråklige andreklassinger

2019

Sammenhengen mellom alder og aritmetiske ferdigheter. En kvantitativ undersøkelse på norske enspråklige andreklassinger.

Cecilie Pavljuk Alver

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

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Matematikk er et fag som står sentralt i skolen og det er et fag som barn møter fra de går i førsteklasse, og for noen elever vil matematikkferdigheter være ekstra utfordrende å tilegne seg (Nortvedt & Vogt, 2012). For barn som har matematikkvansker eller vansker med innlæring av matematikk vil vansken i høy grad være vedvarende (Jordan, Fuchs & Dyson, 2015; Mazzocco & Myers, 2003), og for de elevene som kommer inn i skolen med svake ferdigheter vil ofte ha en sakte utvikling (Morgan, Farkas & Wu, 2009). Alle barn i Norge som er født samme år vil starte i samme klassetrinn med få unntak, og noen elever vil dermed være nærmere ett år yngre enn de eldste i et klassetrinn (Olsen & Bjørnsson, 2018). Forskjellen i alder innenfor samme aldersgruppe blir ofte kalt for relativ alderseffekt (Dalen & Aune, 2013), og i utdanningssammenheng blir denne effekten sett i sammenheng mellom hvor sterkt barns alder er relatert til prestasjoner på faglige prøver i samme klasserom (Olsen & Bjørnsson, 2018). Relativ alderseffekt har vært et kjent fenomen i kroppsøving og i ulike sportsgrener (Cobley, Abraham & Baker, 2008; Sæther, Peterson & Matin, 2017). Det er også flere studier som viser at de eldste elevene i et klasserom presterer bedre enn sine yngre medelever i akademiske fag, som matematikk (Bedard & Dhuey, 2006; Kawaguchi, 2011; Solli. 2017; Strøm, 2004). Relativ alderseffekt er funnet i studier fra Norge i både sport og akademiske fag hvor de tar for seg undersøkelser som TIMSS og PISA fra 4. trinn og oppover (Dalen & Aune, 2013; Olsen & Bjørnsson, 2018; Solli, 2017). Det er funnet relativ alderseffekt i andre land som Japan (Kawaguchi, 2011), Tyskland (Thoren, Heining & Brunner, 2016) og England (Cobley, McKenna, Baker & Wattie, 2009; Crawford, Dearden & Greaves, 2013). Forskningen viser at elever som går i samme klassetrinn kan ha ulike forutsetninger for å lære basert på alder, og det vil kunne påvirker både elevene og lærerne i hverdagen (Olsen & Bjørnsson, 2018). Relativ alderseffekt vil også kunne gi konsekvenser i hele livsløpet når det gjelder akademiske ferdigheter, atferd og holdninger (Cobley et al., 2009; Dalen & Aune, 2013; Thoren et al., 2016).

Formålet med den foreliggende studien er å undersøke sammenhengen mellom alder og aritmetiske ferdigheter hos norske andreklassinger. Problemstillingen for oppgaver er:

Er det sammenheng mellom alder og aritmetiske ferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?

Denne studien har som hensikt å svare på følgende forskningsspørsmål:

I hvilken grad kan alder forklare unik variasjon i regneflyt hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn? I hvilken grad kan alder forklare unik variasjon i aritmetiske problemløsningsferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbal evnere og kjønn?

Studien er av et ikke-eksperimentelt design og er en tverrsnittstudie med data fra pre-testene i intervensjonsprosjektet *The vocabulary learning challenge* ved Instituttet for Spesialpedagogikk ved Universitetet i Oslo. Utvalget i VLC-prosjektet består av 718 andreklassinger. I denne studien ble aritmetiske ferdigheter målt og nonverbale evner. For å måle dette ble det tatt utgangspunkt i måleinstrumentene *WISC-IV regning* som mål på aritmetiske problemløsningsoppgaver, *Regnefaktaprøven - addisjon og subtraksjon* som mål på regneflyt, og *Raven* som mål på nonverbale evner. På den bakgrunn ble elever som ikke hadde de relevante testene ekskludert fra utvalget i denne studien, samt elever som ikke var norske enspråklige elever og elever som var født i et annet fødselsår. Utvalget i denne studien består av 332 norske enspråklige andreklassinger, fra 12 skoler fordelt på 32 klasser i tre kommuner på Østlandet. Dataen ble analysert ved hjelp av deskriptive analyser, bivariat korrelasjonsanalyse og hierarkisk multippel regresjonsanalyse i IBM SPSS.

De bivariat korrelasjonsanalyse viste ingen sammenheng mellom alder og regneflyt, men viste en svak korrelasjon mellom alder og aritmetiske problemløsningsferdigheter. Alder forklarer 2.7 % av den unike variansen i WISC-IV regning. Det er fremdeles en statistisk signifikant sammenhengen mellom alder og aritmetiske problemløsningsferdigheter etter at det blir kontrollert for nonverbale evner og kjønn. Nonverbale evner, kjønn og alder forklarer til sammen 19.7 % av variasjonen i WISC-IV regning, og hvor alder forklarer 1 % av den unike variasjonen. Samlet støtter resultatene fra aritmetiske problemløsingsoppgaver oppunder funn fra tidligere studier at det finnes en relativ alderseffekt i fag som matematikk. Denne studien er med på vise at en allerede i andreklasse kan se at de som er født tidlig på året presterer bedre enn de som er født sent på året.

Forord

Først og fremst vil jeg takke mine veiledere Riikka Mononen og Terje Ulv Throndsen for uvurderlig veiledning og tilbakemelding på masteroppgaven. Jeg vil også takke de resterende som er med i forskergruppen LeMoWe for all veiledning og spesielt for hjelp til analysene.

Jeg vil også gi en takk til alle i VLC-prosjektet for at jeg har fått ta del i forskningsprosjektet. Det har vært spennende å få lov til å delta i et så stor prosjekt og få lov til å benytte seg av datamaterialet i forbindelsen med denne masteroppgaven.

Avslutningsvis vil jeg takke gjengen på lesesalen. Uten dere hadde prosessen med å skrive masteroppgaven vært betraktelig tyngre, og kaffe- og lunsjpausene med dere har vært en av høydepunktet i løpet dagen.

Mai, 2019

Cecilie Pavljuk Alver

Innholdsfortegnelse

Liste over tabeller:

Liste over figurer:

Liste over vedlegg:

1 Innledning

I Norge går alle barn som er født det samme året i samme klassetrinn med få unntak. Det betyr at forskjellen i alder på et klassetrinn kan være bortimot ett år (Olsen & Bjørnsson, 2018; Solli, 2017). Ny forskning fra Norge viser at når en er født på året kan ha en betydning for prestasjonene på skolen (Andersen & Smestad, 2018). I 2018 kom Olsen og Bjørnsson ut med en bok som tar for seg Trends in International Mathematics and Science Study (TIMSS) og Programme for International Student Assessment (PISA) i Norge de siste tjue årene. Et av temaene de tok for seg er hvordan alderseffekten har endret seg over tid og hvordan den varierer etter alder og klassetrinn. Et av resultatene de fant var at effekten av alder i fjerdeklasse er tydelige i barns prestasjoner i matematikk og naturfag, og det er en faktor som holder seg gjennom skoletiden (Olsen & Bjørnsson, 2018). Denne effekten av alder er også funnet på tvers av landegrenser og knyttet til et bredt spekter av områder som idrett, sjakk, akademiske fag, spesifikke lærevansker, spesialundervisning, med og uten innvandringsbakgrunn (Aune, Pedersen, Ingvaldsen & Dalen, 2017; Cobley et al., 2009; Crawford et al., 2013; Dalen & Aune, 2013; Helsen, Baker, Schorer, Steingröver, Wattie & Starks, 2016; Solli, 2017; Thoren et al., 2016). Forskningen til Olsen og Bjørnsson (2018) viser at elever som går i samme klasse kan ha svært ulike forutsetninger for å lære basert på alder, og de vil ofte konkurrere og vurderes opp mot hverandre (Dalen & Aune, 2013). Relativ alderseffekt viser til at det ikke bare påvirker elevene og lærerne i hverdagen, men kan også gi konsekvenser i hele livsløpet når det gjelder akademiske ferdigheter, atferd og holdninger (Cobley et al., 2009; Dalen & Aune, 2013; Thoren et al., 2016). Bjørnsson trekker frem at dette er noe lærere bør merke seg på bakgrunn av at de yngste elevene kan ha behov for særskilt tilrettelegging, spesielt på barnetrinnet (Andersen & Smestad, 2018).

I 2017 fremmet Camilla Stoltenberg (leder for Folkehelseinstituttet) frem en debatt om at det er et nytt kjønnsgap i skolen. Hun fremhever at karakterene gutter og jenter får fra ungdomsskolen kan prege dem resten av livet, og at det er i denne alderen at de biologiske «aldersforskjellene» er på sitt største. Camilla Stoltenberg sier at jenter som er født tidlig på året biologisk sett kan være flere år eldre enn gutter som er født sent på året, og disse elevene går samme klasse og vurderes opp mot hverandre (Stoltenberg, 2017). Camilla Stoltenberg har sammen med Stoltenbergutvalget fremmet en debatt om fleksibel skolestart på bakgrunn av såkalt modningshypotese som går ut på at flere av guttene som er «umodne» for skolestart med fordel kan starte senere på skolen (NOU, 2019:3). Dette er noe forskningen til Olsen og

Bjørnsson (2018) ikke støtter. De viser til at hvis slik hypotese om at jenter og gutters modningstakt skal være en riktig og viktig faktor for å forstå kjønnsforskjeller i skolen, burde det gjenspeiles i kjønnsforskjeller i gjennomsnittlige prestasjoner og relativ alderseffekt hos de yngste barna. Dette er noe som viser seg å ikke være tilfellet (Andersen & Smedstad, 2018).

Matematikk er også et fag som står sentralt i skolen, og det er et fag som barn møter fra førsteklasse og som vil følge dem til videregående skole (Nortvedt & Vogt, 2012). Et av målene med matematikkfaget er å utvikle kompetansen til den enkelte elev slik at de har redskaper for å løse ulike problemer og oppgaver, samt at faget kan være med å danne grunnlag for senere utdanning og deltakelse i yrkesliv (Nortvedt & Vogt, 2012; Utdanningsdirektoratet, 2013). Utdanningsdirektoratet (2013) trekker frem at matematikkfaget på den bakgrunn «spelar ei sentral rolle i den allmenne danninga ved å påverke identitet, tenkjemåte og sjølvforståing» (s. 2). For noen elever vil denne kompetansen være ekstra utfordrende å tilegne seg. For disse elevene kan det i hverdagslivet føre til bekymringer, og det kan oppstå problemer med å forstå og holde orden på egen privatøkonomi. Vanskene elevene har i matematikk kan også få følger for fremtidige yrkesvalg (Nortvedt & Vogt, 2012).

1.1 Formål

Formålet med denne studien er å undersøke om en kan se forskjeller i alder hos norske enspråklige andreklassinger i matematikk, spesielt aritmetiske ferdigheter. Denne studien vil undersøke om hvordan alderen påvirker norske andreklassingers ferdigheter i aritmetikk, som regneflyt og aritmetiske problemløsningsoppgaver basert på standardiserte kartleggingsverktøy. Bakgrunnen for dette er at matematikk er et fag som står sentralt i skolen og vansker i dette faget kan føre til store bekymringer, samt at forskningen som er gjort på alder (relativ alderseffekt) og matematikkferdigheter i Norge er i hovedsak basert på undersøkelser fra 4.trinn og oppover. Datagrunnlaget fra disse studiene også er hentet fra nasjonale eller internasjonale kartleggingsprøver som nasjonale prøver (Norge) og TIMSS- og PISA-undersøkelsene. Forskningen viser at elever som går i samme klassetrinn kan ha ulike forutsetninger for å lære basert på alder, og at alder påvirker både elevene og lærerne i hverdagen, men kan også gi konsekvenser i hele livsløpet når det gjelder akademiske ferdigheter, atferd og holdninger.

1.2 Problemstilling

Problemstillingen for studien er: *Er det sammenheng mellom alder og aritmetiske ferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?* Studien har som hensikt å svare på følgende forskningsspørsmål:

I. I hvilken grad kan alder forklare unik variasjon i regneflyt hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?

II. I hvilken grad kan alder forklare unik variasjon i aritmetiske problemløsningsferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?

1.3 Oppgavens oppbygning

Det innledende kapittelet presenterer oppgavens innledning med bakgrunn og formål, samt problemstillingen og forskningsspørsmålene for denne studien.

I kapittel 2 presenteres studiens teoretiske og empiriske grunnlag. Først presenteres aritmetisk utvikling med fokus på utvikling av strategier og aritmetiske problemløsningsferdigheter, deretter vansker i aritmetikk. Det redegjøres videre for tidligere studier på relativ alderseffekt i akademiske fag fra ulike aldersgrupper.

I kapittel 3 redegjøres det for oppgavens metodiske tilnærming, utvalg, måleinstrumenter, datainnsamling og etiske vurderinger. Videre presenteres validitet og reliabilitet, samt analysene som skal gjennomføres.

I kapittel 4 presenteres studiens analyser og resultater.

I kapittel 5 presenteres drøftingen av resultater i lys av tidligere teori og empiri, samt i lys av validitetsteori. Avslutningsvis oppsummeres oppgaven og det vil bli presentert hvilke implikasjoner resultatene kan ha for praksis, og forslag på fremtidige studier på relativ alderseffekt.

2 Teoretisk og empirisk grunnlag

I dette kapittelet presenteres studien teoretiske og empiriske grunnlag. Først vil det bli gjort rede for utvikling av aritmetiske ferdigheter; utvikling av aritmetiske strategier og aritmetiske problemløsningsoppgaver. Deretter vil det bli gjort rede for vansker med aritmetikk; vansker med aritmetiske strategier og aritmetiske problemløsningsoppgaver. Videre vil studier på relativ alderseffekt presenteres, som tar for seg alder og akademiske fag, som matematikk, fra barneskole til videregående skole. Avslutningsvis vil studiens teoretiske og empiriske grunnlag oppsummeres.

2.1 Utvikling av aritmetiske ferdigheter

Aritmetikk består ikke bare av en ferdighet, men av flere ferdigheter (Dowker, 1998, 2005). En naturlig følge av dette er at aritmetisk utvikling ikke bare består av en enkel prosess, men av flere som involverer utvikling av ulike komponenter (Dowker, 1998). Grunnleggende aritmetiske ferdigheter inkluderer kunnskap om tall, kunnskap om de fire grunnleggende regneartene (addisjon, subtraksjon, multiplikasjon og divisjon), memorere aritmetisk fakta (f.eks. 5 + 2 = 7), følge aritmetiske prosedyrer og forståelse av aritmetiske prinsipper (Dowker, 1998, 2005; Haskell, 2000). Aritmetisk utvikling involverer også en endring i bruk av strategier, samt hvor nøyaktig og hvilken hastighet som strategier kan bli utført (Geary, 1994). De tidlige matematikkferdighetene barn lærer seg er å telle, identifisere tall, og sammenligne og manipulere mengde (Raghubar & Barnes, 2017). Kunnskap om telling danner et grunnlag for læring og forståelse av addisjon, og kunnskap om addisjon danner videre grunnlag for læring og forståelse av subtraksjon (Siegler & Braithwaite, 2017).

Mestring av grunnleggende aritmetiske ferdigheter er et nøkkelmål i barns tidlige skolegang og utviklingen opptar flere år (Göbel, Watson, Lervåg & Hulme, 2014). Likevel er det flere barn som tilegner seg forståelse av aritmetikk før de har mottatt formell undervisning (Butterworth, 2005; Dowker, 2005; Geary, 1994; Gilmore, Göbel & Inglis, 2018). Noen forskere foreslår at barn har aritmetiske ferdigheter fra de er spedbarn (Moore & Cocas, 2006; Koechlin, Dehaene & Mehler, 1997; Simon, Hespos & Rochat, 1995; Wakeley, Rivera & Langer 2000; Wynn, 1992). For eksempel viser det seg at spedbarn er sensitive til små tall og til øking (addisjon) og redusering (subtraksjon) i mengden ved små sett (Geary, 2000; Wynn, 1992). På bakgrunn av dette kan en anta at barn har en grunnleggende forståelse av at addisjon og subtraksjon påvirker mengde (Wynn, 1992). Selv om det er blandede bevis på om spedbarn er i stand til å forstå og gjennomføre aritmetiske operasjoner så er det robuste bevis på at barnehagebarn (to til fire år) er i stand til å utføre addisjons- og subtraksjonsoppgaver før de har hatt formell undervisning i aritmetikk (Gilmore et al., 2018).

Som nevnt er telling en del av de tidlige matematikkferdighetene og barn tilegner seg disse ferdighetene fra de er små (Fuson, 1988; Gelman & Gallistel, 1978; Huttenlocher, Jordan & Levine, 1994). Barn bruker telling og tallkunnskap til å løse aritmetiske problemer tidlig i utviklingen (Geary, 1994; Siegler & Shrager, 1984), samt at det danner grunnlag for barns aritmetiske problemløsningsferdigheter (Geary, 1994). Barn lærer at prosessen med telling kan gi antall objekter i et sett, det vil si at de har forståelse for kardinal prinsippet (Butterworth, 2005). Barn vil også støtte seg på titallsystemet for å løse aritmetiske problemer, og barns forståelse av den konseptuelle meningen av flersifrede tall er avhengig av titallssystemet (Geary, 1994). Titallsystemet refererer til forståelsen av forskjellen mellom hundrere, tiere og enere, det vil si kunnskap om hvordan tallsymbolene har ulik verdi avhengig av hvilken plassering de har i serien av tall (Aunio & Räsänen, 2015).

Utvikling av kunnskap om de aritmetiske prinsippene vil hjelpe barn å beregne svar på ulike tallkombinasjoner (Baroody & Tiikaninen, 2003; Cowan, 2003; Dowker, 1998). For eksempel kan det kommutative prinsippet (a + b = ?/ b + a = ?) og det assosiative prinsippet, (a + b) + c/a + (b + c), bidra til «snarveier» til beregninger (Russel & Ginsburg, 1984). Utvikling av kunnskap om aritmetiske prinsipper og bruk av telleprosedyrer vil resultere i utvikling av minnerepresentasjon av aritmetisk fakta. Dette fører til bruk av strategier som gjenhenter aritmetisk fakta fra langtidsminne, som videre vil bli brukt for å løse aritmetiske problemløsningsoppgaver (Aunio & Räsänen, 2015; Raghubar & Barnes, 2017; Siegler & Shrager, 1984).

Fra barn er små er forskjellene matematikkferdigheter mellom jenter og gutter som regel heller et unntak enn regelen, men det er funnet signifikante forskjeller i addisjons- og subtraksjonsoppgaver. Imidlertid var effektstørrelsen relativ svak, og når oppgavene ble brutt ned i klassetrinn syntes gutter og jenter å utføre oppgavene like bra (Hutchison, Lyons & Ansari, 2019). Fra PISA-undersøkelsen fra 2015 (15/16-åringer) vises det ingen signifikante forskjeller mellom kjønnene i matematikk, men at det var størst spredning i guttenes prestasjoner enn hos jentene. Det tyder derfor på at det er lite sannsynlig at kjønn påvirker barns matematikkprestasjoner (Nortvedt & Pettersen, 2016).

2.1.1 Aritmetiske strategier

Barn lærer ulike strategitilnærminger gjennom å tenke dem ut selv og gjennom undervisning (Gilmore et al., 2018). For å løse aritmetiske problemer bruker barn et bredt utvalg av strategier og kan bruke disse fleksibelt (Gilmore et al., 2018; Ostad, 1997; Siegler & Braithwaite, 2017; Siegler & Shrager, 1984; Siegler, 1987, 1989), samt at de har også delvis forståelse av aritmetiske prinsipper (Cowan, 2003). Strategiene som blir brukt for å løse aritmetiske problemer blir hovedsakelig delt inn i strategier som involverer bruk av konkreter (f.eks. fingre eller objekter), verbale strategier (mentale) og retrievalstrategier (Gilmore et al., 2018). De blir også delt inn i backup- og retrievalstrategier (Ostad, 2013). Retrievalstrategier («hente-frem-strategier») kjennetegnes ved oppgaver hvor barnet henter frem kunnskaps-enheter fra minnet. Backupstrategier omfavner de øvrige strategiene som ikke retrieval (Siegler & Jenkins, 1989).

Forskere har identifisert flere ulike strategier som brukes til å løse enkle addisjonsproblemer (Baroody, 1987; Carpenter & Moser, 1982; Geary, 2004; Jordan, Hanich & Uberti, 2003; Siegler, 1987), som er oppsummert i tabell 1 (Geary, 1994, s. 62; Gilmore et al., 2018, s. 60). Strategiene skiller seg fra hverandre ved hvor nøyaktige de er, tiden de trenger for å gjennomføres, hvor mye minnekapasitet de krever og i hvor stor bredde de har for å benyttes på de ulike aritmetiske problemene (Siegler & Shrager, 1984, Siegler, 1991). Telleprosedyrene kan utføres ved bruk av hjelpemidler som fingre (fingerstrategier) og ulike objekter eller ved å telle verbalt (verbale strategier) (Siegler & Shrager, 1984).

Tabell 1. Oppsummering av addisjonsstrategier

Strategi	*Beskrivelse*	*Eksempel på løsning av 4 + 8*
Telle alle	Bruke objekter eller fingre til å danne sett for hver addend og deretter telle alle	«1, 2, 3, 4», «1, 2, 3, 4, 5, 6, 7, 8», «1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12»
Telle fra første	Telle fra første addend, fingre kan bli brukt til å holde rede på den andre addend	«5, 6, 7, 8, 9, 10, 11, 12»
Telle fra den største	Telle fra den største addend, fingre kan bli brukt til å holde rede på den andre addenden	«9, 10, 11, 12»

Dekomposisjons/ derivere fakta	Rekonstruere svaret basert på retrieval av deler av svaret	Steg 1: 4 = 2+2 Steg 2: 8+2 = 10 Steg 3: 10+2 = 12
Retrieval	Gjenhenting av kjent fakta	12

De strategiene som gjelder for å løse enkle subtraksjonsoppgaverminner om strategiene som brukes for å løse enkle addisjonsoppgaver, se tabell 1 (Geary, 1994). Strategiene for å løse enkle subtraksjonsoppgaver (f.eks. 3 -1) er (a) *separere fra*; barnet har tre klosser og teller ut disse, deretter blir en kloss tatt bort. De resterende klossene blir telt ut «1, 2» og angir svaret. (b) *legge til*; involverer å starte med antall klosser som utgjør det minste tallet «1», deretter legge til antall klosser inntil det største tallet er nådd «1, 2» og svaret er antall klosser som er lagt til. (c) *telle opp*; involverer å starte å telle fra det minste tallet «1» til det største tallet er nådd «2, 3», og svaret er representert av antall tellinger. (d) *telle ned;* involverer å telle tilbake fra det største tallet «3», antall ganger representert av det minste tallet «1» og svaret er representert ved antall tellinger «2». (e) *retrieval*; gjenhenting av kjent fakta, «2». En annen strategi som er brukt for å løse enkle subtraksjonsoppgaver er henvisning til addisjonsoppgaver (addition reference). For å løse 8 - 2, kan de gjenhente 6 + 2 = 8 for å løse subtraksjonsproblemet (Geary, 1994).

Ved barnehagestart vil et barn kunne respondere på ulike addisjons- eller subtraksjonsoppgaver gjennom gjetting eller ved å gjøre et omtrentlig estimat (Jordan et al., 2003). De vil gradvis lære å regne ut enkle aritmetiske oppgaver gjennom telling av fingre eller andre konkreter (f.eks. 2 + 1 = 3, 2 – 1 = 1) (Jordan et al., 2003; Siegler, 1987). Ved tre årsalderen kan noen barn bruke telling til å addere i hverdagskontekster (Fuson, 1982), og ved denne alderen vil de normalt støtte seg på konkreter for å addere (Geary, 1994). Barn i barnehagealder vil i hovedsak bruke telle alle strategier når de skal løse addisjonsoppgaver (Geary, 1994; Jordan et al., 2003), og de vil kunne bruke direkte retrieval ved enkle tallkombinasjoner som de har kjennskap til fra før, for eksempel 2 + 1 = 3 (Jordan et al., 2003). Bruk av konkreter som ulike objekter eller fingre er et hjelpemiddel for å løse ulike addisjons- og subtraksjonsproblemer (Geary, 1994). Ved spørsmål om «hvor mange er tre kjeks og to kjeks» vil barn typisk telle ut tre objekter, deretter telle to objekter og til slutt telle alle objektene ved å starte fra en (Geary, 1994). Bruk av konkreter hjelper barn å representere tallene som skal telles og hjelper barna å holde øye med tellingen (Carpenter & Moser, 1983).

Bruk av konkreter blir også brukt av barn i fire- til fem årsalderen avhengig av kompleksiteten i problemet og om konkreter er lett tilgjengelig (Fuson, 1982). De fleste fire- til femåringene vil bruke en kombinasjon av fingertelling og verbaltelling til å løse addisjonsproblemer hvis de ikke kan gjenhente svaret direkte. I denne alderen vil barn også kunne løse subtraksjonsproblemer som er formelt presentert. For eksempel «hvis du har tre kjeks og gir en til broren din. Hvor mange har du igjen?» (Siegler & Shrager, 1984). Fra barna er fem- til seks år vil de som regel bruke tellestrategier for å løse subtraksjonsproblemer (Geary, 1994).

Overgangen fra å bruke fingre og konkreter til bruk av verbale tellestrategier er i hovedsak avhengig av barns evne til å mentalt holde kontroll på tall som allerede har blitt telt og de som skal bli telt (Fuson, 1982). For de fleste av barna vil endringen fra bruk av konkreter til verbal telling skje gradvis (Geary, 1994). I første- og andreklasse utvikler barn raskere og mer effektive tellestrategier (Jordan et al., 2003), og vil øke bruken av strategiene som er relativt effektive. De vil også minske bruken av de strategiene som er mindre effektive, som for eksempel gjetting og telle fra en (Siegler & Braithwaite, 2017). Å telle fra første tallet innebærer å angi verdien av det første tallet og deretter telle antall ganger som er lik verdien av det andre tallet (Geary, 1994). Dette innebærer at barnet forstår at å konstatere kardinalverdien på det første tallet er en snarvei til å telle det tallet og at en ikke trenger å starte å telle fra en (Fuson, 1982). Når barnet teller fra det største tallet må barnet identifisere det største tallet og deretter telle fra det tallet (Geary, 1994). I denne strategien kreves det ikke bare at barnet har forståelse at kardinalverdien til addendene kan bli brukt for å gjøre de verbale strategiene mer effektive, men også en forståelse for at hvilken rekkefølge tallene blir addert sammen ikke har en påvirkning på summen, også kalt det kommutative prinsippet (Geary, 1994).

Når barn har utviklet minnerepresentasjon av fakta basert på bruken av de ulike tellerprosedyrene vil de i større grad støtte seg på minnebaserte strategier som direkte retrieval og dekomposisjon (Geary & Hoard, 2005). Ved fem- til åtte årsalderen vil barn kunne telle fra det første tallet og fra det største tallet i addisjonsoppgaver, de vil også bruke dekomposisjonsprosedyren mer (Geary, 1994; Jordan et al., 2003; Siegler & Braithwaite, 2017; Siegler, 1987). Barn vil i denne alderen utføre strategiene hurtigere og nøyaktigere, og danne mer adaptive strategivalg (Siegler & Braithwaite, 2017). Dekomposisjonsstrategien involverer bruk av memorerte addisjons- og subtraksjonsfakta som grunnlag for å løse mer utfordrende problemer (Carpenter & Moser, 1983). Barn har lettere for å huske dobleproblemer (f.eks. 1+1, 2+2+, 3+3) eller bindeproblemer (tie) enn andre kombinasjoner (Ashcraft, 1992). Disse memorerte faktaene kan virke som et grunnlag for å løse andre addisjonsproblemer, for eksempel å løse «6 + 7» ved å gjenhente svaret på «6 + 6» og deretter legge til «1» til den foreløpige summen (Geary, 1994). En annen dekomposisjonsstrategi er strukturert rundt titallsystemet, for eksempel kan «6 + 7» bli løst gjennom å dele «6» til to «3», for deretter å addere en av «3» til «7» for å få «10», og deretter addere «10 + 3 = 13» (Geary, 1994). Ved direkte retrieval kan barn raskt produsere et svar, uten tegn på telling og på bakgrunn av at de husker svaret fra tidligere telleerfaringer og dekomposisjonsstrategier (Geary, 1994).

I løpet av barneskolen (første- til syvendeklasse) vil barn øke evnen til å gjenhente svar på kjente tallkombinasjoner (Jordan et al., 2003) og de vil støtte seg mer på retrievalstrategier (Ostad, 1997). Eldre barn og voksne vil hovedsakelig bruke retrieval (den strategien som blir vurdert som den raskeste) når de kan utføre en oppgave nøyaktig (Siegler & Braithwaite, 2017). Når barn ikke kan gjenhente svaret nøyaktig vil de bruke langsommere strategier, som telle fra en eller telle fra det største tallet (Siegler & Braithwaite, 2017). Ved valg av andre strategier enn retrieval vil barn ofte velge den backupstrategien som vil resulterer i størst nøyaktighet og hastighet for å finne rett svar i forhold til de alternativene som er tilgjengelige (Jordan et al., 2003; Siegler & Braithwaite, 2017; Siegler, 1989). I starten av utviklingen av strategier vil barn ofte bruke backupstrategier på vanskeligere oppgaver (f.eks. 4 + 5, 7 - 3) og direkte retrieval på enklere oppgaver (f.eks. 2 + 3, 3 - 1) (Jordan et al., 2003, Siegler, 1991). Ved slutten av barneskolen vil de fleste barna kunne gjenhente eller rekonstruere summer og differanser automatisk til delvis automatisk (Jordan et al., 2003). Barn er i tidlig alder tilpasningsdyktige ved valg av aritmetiske strategier i den forstand at de fremmer den strategien som gir størst nøyaktighet og som de kan gjennomføre hurtigs (Siegler, 1996). Barn vil dermed variere sin strategibruk fra situasjon til situasjon (Aschcraft, 1992; Siegler & Jenkins, 1989), og ved et bredt utvalg i strategier vil de lettere kunne tilpasse fremgangsmåten til å møte ulike omstendigheter (Siegler & Shrager, 1984, Siegler, 1991).

For jenter og gutter er det lite som tyder på at det er forskjeller i valg av ulike strategiene for å løse enkle aritmetiske problemer (Geary, Bow-Thomas, Liu & Siegler, 1996; Siegler, 1988). Der er derimot funnet at jenter bruker objekter og fingerstrategier mer enn guttene, som bruker retrievalstrategien mer (Carr & Davis, 2001; Jordan, Kaplan, Ramineni & Locuniak, 2008; Shen, Vasilyeva & Laski, 2016).

2.1.2 Aritmetiske problemløsningsoppgaver

Aritmetiske problemløsningsoppgaver er med på å gi et innblikk i barns aritmetiske ferdigheter (Gilmore et al., 2018; Thevenot & Barrouillet, 2015). For å løse aritmetiske problemløsningsoppgaver må barn danne seg et bilde av problemet, hente ut relevant informasjon, velge riktig regneoperasjon og deretter gjennomføre den valgte regne-operasjonen (Geary, 1994; Gilmore et al., 2018). Gjennom grunnskolen vil barns grunnleggende regneferdigheter (f.eks. addisjon og subtraksjon), fleksibilitet knyttet til valg av strategier og matematikkvokabular, sammen med hvilke krav det aritmetiske problemet stiller til minne være med på å avgjøre barnets suksess til å løse aritmetiske problemløsnings-oppgaver (Geary, 1994; Riley, Greeno & Heller, 1983).

Den semantiske strukturen i aritmetiske problemløsningsoppgaver vil også være med å påvirke barns evne til å løse disse problemene (Carpenter & Moser, 1983; De Corte & Verschaffel, 1987; Riley et al., 1983; Jordan et al., 2003), samt påvirke barns valg av strategi (Carpenter & Moser, 1982; De Corte & Verschaffel, 1987; Fuson, 1992). Språkstrukturen i de aritmetiske problemløsningsoppgavene vil kunne påvirke barns evne til å representere meningen med problemet mentalt, som videre vil påvirke problemløsningsferdighetene (Geary, 1994). En av de sentrale trekkene ved barns problemløsningsevner ved aritmetiske problemløsningsoppgaver involverer å utvikle en passende representasjon av problemet og deretter knytte denne representasjonen til den strategien som faktisk vil løse problemet (Geary, 1994).

I aritmetiske problemløsningsoppgaver blir det beskrevet situasjoner hvor det skjer en endring, et bytte eller en beskrivelse av mengde og forholdet mellom disse (Riley et al., 1983; Thevenot & Barrouillet, 2015). Aritmetiske problemløsningsoppgaver fanger et bredt spekter av situasjoner (Gilmore et al., 2018) og disse blir klassifisert etter den semantiske struktur (Geary, 1994). I addisjons- og subtraksjonsproblemløsningsoppgaver skilles det ofte mellom tre situasjoner av problemer: *endringsproblem, sammensetningsproblem* og *sammenligningsproblem*. Endringsproblem refererer til en aktiv eller dynamisk situasjon der noen hendelser endrer verdien av en innledende mengde. Sammensetningsproblemer relaterer seg til statiske situasjoner som involverer to mengder som enten vurderes separat eller i kombinasjon. Sammenligningsproblemer involverer sammenligning av to verdier og differansen mellom dem (Verschaffel, Depaepe & Dooren, 2015; Verschaffel & De Corte, 1997).

Ytterligere forskjeller ble gjort på bakgrunn av den ukjentes natur og retningen av handlingen eller forholdet, noe som resulterer i åtte forskjellige en-stegs addisjons- og subtraksjonsproblemer, som vises i tabell 2 (Gilmore et al., 2018, s. 57). De grunnleggende strategiene som brukes til å regne ut addisjons- og subtraksjonsoppgaver er de samme grunnleggende strategiene som blir brukt til å løse disse aritmetiske problemløsningsoppgavene (Carpenter & Moser, 1982).

Tabell 2. Klassifisering av aritmetiske problemløsningsoppgaver

Klasse	*Ukjent*	*Eksempel*
Endring	Ukjent resultat	Jonas hadde 3 kroner. Han fikk 5 kroner av Thomas. Hvor mange kroner har Jonas nå?
Endring	Ukjent endring	Jonas hadde 3 kroner. Han fikk noen kroner av Thomas. Nå har Jonas 8 kroner. Hvor mange kroner ga Thomas til Jonas?
Endring	Ukjent start	Jonas hadde til å begynne med noen kroner. Han fikk 5 kroner av Thomas. Nå har Jonas 8 kroner. Hvor mange kroner hadde Jonas til å begynne med?
Sammensetning	Ukjent sammensetningsverdi	Jonas hadde 3 kroner. Thomas hadde 5 kroner. Hvor mange kroner har Jonas og Thomas til sammen?
Sammensetning	Ukjent delverdi	Jonas og Thomas hadde 8 kroner til sammen. Jonas hadde 3 kroner. Hvor mange kroner har Thomas?
Sammenligning	Ukjent differanse	Jonas hadde 8 kroner. Thomas har 5 kroner. Hvor mange kroner mer hadde Jonas enn Thomas?
Sammenligning	Ukjent sammenligningsverdi	Jonas hadde 3 kroner. Thomas hadde 5 kroner mer enn Jonas. Hvor mange kroner har Thomas?
Sammenligning	Ukjent referent	Jonas hadde 8 kroner. Han hadde 5 kroner mer enn Thomas. Hvor mange kroner har Thomas?

Mange barn kan løse enkle aritmetiske problemløsningsoppgaver før de starter på skolen (Levine, Jordan & Huttenlocher, 1992). Rundt to- til fire og halvt årsalderen kan barn løse nonverbale aritmetiske problemløsningsoppgaver (Gilmore et al., 2018; Levine et al., 1992).

Eksempel på nonverbale oppgaver er bruk av objekter som blir endret ved enten å tilføye eller ta bort et element (Levine et al., 1992). De fleste barnehagebarn kan løse endringsproblem med ukjent resultat (Geary, 1994; Jordan et al., 2003). Før formell undervisning presterer barn bedre på aritmetiske problemløsningsoppgaver som er innebygget i hverdagsproblemer enn ved ulike tallkombinasjoner som er relativt fri for kontekst (Levine et al., 1992).

Fra fem- til seks årsalderen kan barn utføre addisjons- og subtraksjonsproblemløsnings-oppgaver (Levine et al., 1992). Eksempel på slike oppgaver er «Mikkel hadde 4 baller. Han fikk 1 til. Hvor mange baller har han til sammen?» (Gilmore et al., 2018). Når barn er i fire- til fem årsalderen støtter de seg på verbal telling for å løse enkle aritmetiske problemløsnings-oppgaver (Baroody & Ginsburg, 1986; Siegler & Jenkins, 1989). Ved fem- til seks årsalderen kan barn også løse abstrakte addisjonsoppgaver, og rundt seks- til seks og et halvt årsalderen kan de løse abstrakte subtraksjonsoppgaver (Levine et al., 1992). Eksempel på abstrakte problemløsningsoppgaver er «hvor mye er 4 og 1 til sammen?» (Gilmore et al., 2018). I løpet av første- til tredjeklasse blir barns problemløsningsferdigheter stadig mer sofistikerte (Riley & Greeno, 1988).

I førsteklasse er mange av barna i stand til å utføre enkle problemløsningsoppgaver med ukjent endring (Jordan et al., 2003), men fra de er barnehagebarn viser de få vansker med å løse endringsproblem med ukjent resultat (Riley et al., 1983). I førsteklasse løser barn endringsproblemer med ukjent resultat gjennom telle alle strategien, og senere ved telle fra strategier (Carpenter & Moser, 1982). I løpet av andreklasse kan de fleste barna løse en rekke aritmetiske problemløsningsoppgaver som involverer sammenligning, som for eksempel «Maria har 7 blyanter. Benjamin har 3 blyanter. Hvor mange flere blyanter har Maria enn Benjamin?» (Jordan et al., 2003). Ved slutten av tredjeklasse klarer de fleste barna å mestre mer begrepsmessig sofistikerte aritmetiske problemløsningsoppgaver som involverer sammenligning, «Maria hadde 7 blyanter. Hun har 4 blyanter mer enn Benjamin. Hvor mange blyanter har Benjamin?» (Jordan et al., 2003).

2.2 Vansker med aritmetikk

For noen barn er innlæring av matematikk en kilde til vanskeligheter, spesielt i grunnleggende ferdigheter i aritmetikk (Nöel, 2015). Noen barn har vansker i matematikk selv om de har normal intelligens, utdanningsmuligheter, normalt syn og hørsel eller språk. Vanskene er karakterisert med å være signifikante og vedvarende, og disse vanskene blir ofte referert som dyskalkuli eller spesifikke matematikkvansker (Nelson & Powell, 2017: World Health Organization, 2018). Andre barn kan vise mildere vansker enn de barna som har dyskalkuli, men vil enda ha tilstrekkelig svake vansker ved innlæring matematikk, og da særlig i grunnleggende ferdigheter i aritmetikk (Mazzocco, 2007). Disse barna blir referert til lavt presterende i matematikk, og vanskene kan komme av utilstrekkelig undervisning, manglende eksponering for aritmetiske problemer, dårlig motivasjon eller lav intelligens (Geary, Brown & Samaranayake, 1991; Geary & Hoard, 2005). Forekomsten av dyskalkuli og lavt presterende i matematikk vil variere etter hvilke kriterier og tester som blir brukt. Barn som skårer på eller under 10. persentil i en standardisert matematikktest i to år eller mer blir ofte identifisert som dyskalkuli. Barn som skårer mellom 11. og 25. persentil blir ofte identifisert som lavt presterende i matematikk. Forekomsten av barn som har dyskalkuli er beregnet for å ligge rundt tre til åtte prosent og forekomsten av barn som er lavt presterende i matematikk er beregnet å ligge rundt ti til femten prosent (Geary, 2015, 2017). Samtidig er en cut-off skår kunstig på den bakgrunn av at matematikkferdigheter varierer på et kontinuum. Barn som blir identifisert med dyskalkuli ligger av den grunn på enden av et kontinuum, og diagnosekriteriene kan bli flyttet lavere og høyere. Samtidig vil mange av disse barna ha spesifikke vansker innenfor et område i matematikken (Geary, 2017).

Barna som enten er lavt presterende i matematikk eller har dyskalkuli vil ofte streve eller mislykkes med å tilegne seg aldersadekvate ferdigheter (Dowker, 2005; Price & Ansari, 2013) og vil kunne resultere i betydelige svekkelser i faglige og yrkesmessige funksjon (World Health Organization, 2018). Vanskene kommer ofte til uttrykk gjennom langsom innlæring av tallkonsepter og grunnleggende aritmetikk (Geary, 2017). Barna viser vansker med å lære ferdigheter som tallforståelse, aritmetiske prosedyrer, regneflyt, og de har vansker med innlæring og gjenhenting av aritmetisk fakta (World Health Organization, 2018; Geary, 2004; Nöel, 2015). Barn som har vansker med innlæring av matematikk lærer aritmetisk fakta, men de viser vansker med å huske like mange fakta som barn uten vansker og de glemmer disse faktaene relativ raskt (Geary, 1994, 2017). Flere av barna med matematikkvansker bruker umodne strategier når de skal løse aritmetiske problemløsningsoppgaver (Geary, 2004; Jordan et al, 2003; Nöel, 2015). De teller for eksempel på fingrene for å løse slike oppgaver i flere år enn sine medelever, og de gjør ofte flere feil når de teller. De fleste barna når igjen sine medelever når det kommer til aritmetiske problemløsningsoppgaver, men viser til å ha en mer vedvarende vanske med å huske aritmetisk fakta (Geary, Hoard, Nugent & Bailey, 2012).

Barn som har vansker eller er forsinket med innlæring av tall og grunnleggende aritmetikk er i risikosonen for å falle bak sine medelever i matematikk og henge etter dem gjennom skolegangen (Geary, 2017). Disse barna er også ofte svakere forberedt for høyere nivå i opplæringen i ungdomsskole og videregående skole (Jordan et al., 2003).

Matematikkvansker hos barn er som nevnt i høy grad vedvarende (Jordan et al., 2015; Mazzocco & Myers, 2003). Noen studier viser til at prestasjoner i matematikk er stabile og har økende varians over tid (Aunola, Leskinen, Lerkkanen & Numi, 2004). Samt at utviklingen av matematikkferdigheter er raskere blant de elevene som går inn i grunnskolen med allerede høyere nivå i matematikkferdigheter (Aunola et al., 2004) og at de elevene som opplever vansker med matematikk har tendenser til å fortsette å oppleve vansker med matematikk (Morgan et al., 2009). Andre studier fant at barn ikke nødvendigvis har vedvarende svake vansker i matematikk over tid, men at de går inn og ut av kategorien vansker (Mazzocco & Myers, 2003; Stock, Desoete & Roeyers, 2010). Dette viser til at det kan være utfordrende å predikere hvilke barn som har vedvarende svake matematikkferdigheter (Reikerås & Salomonsen, 2019).

Dowker (2005) fremhever at en ikke kan forstå aritmetiske vansker uten å ta i betraktning frykt for aritmetikk. Hun trekker frem at matematikk er et fag som vekker sterke følelser, og er et fag som barn kan føle en ekstrem motvilje mot. Matematikkangst utvikler seg generelt ikke før det er betydelige forsinkelser i tallforståelsen, og det er sannsynlig at vansker med matematikk vil resultere i frustrasjon, unngåelse, lav mestringstro og potensielt matematikkangst når de skal løse matematikkproblemer. Lav mestringstro og angst vil sammen med en vanske ganske sikkert gjøre innlæringen av matematikk mer utfordrende (Geary, 2017; Rubinsten & Tannock, 2010).

Moll, Kunze, Neuhoff, Bruder og Schulte-Körne (2014) fant i sin studie at det var høyere forekomst av jenter med vansker i aritmetikk enn gutter. Videre er det andre studier som ikke finner at det er høyere forekomst av jenter som har vansker i aritmetikk, og at det ikke er kjønnsforskjeller i forekomsten av dyskalkuli (Barbaresi, Katusic, Colligan, Weaver & Jacobsen, 2005; Devine, Soltèsz, Nobes, Goswami & Szűcs, 2013).

2.2.1 Vansker med aritmetiske strategier

Flertallet av tellefeilene barn gjør involverer enten undertelling eller overtelling med en, og som regel ligger feilen i undertelling med en (Siegler & Shrager, 1984). Tellefeilene som oppstår er ofte på bakgrunn av at barnet mister oversikt på hvilke verdier som har blitt telt og ikke eller så gjør de en prosedyrefeil. Hvis barnet teller fra «4» for å løse «4 + 3» må barnet konstaterer «4» og deretter telle oppover tre ganger, hvis barnet mister oversikt kan føre til at de teller oppover fire eller fem ganger (Geary, 1994). En prosessuell feil kan forekomme av at en bruker teller fra (f.eks. første eller største) strategier (Fuson, 1982). Ved en prosessuell feil teller barnet riktig antall ganger, men starter med feil tall. For å løse «4 + 3» teller barnet oppover tre ganger, men inkluderer «4» som representerer verdien av den første addenden og som første tallet i tellingen av den andre addenden: telle «4, 5, 6» i stedet for «4, 5, 6, 7» (Geary, 1994). Tellefeilene for enkel subtraksjon ser ut til å ligne de type feil som er vanlig sett når barn regner med addisjon (Geary, 1994; Siegler & Shrager, 1984).

Retrievalfeil kan deles inn i fire kategorier*: ren gjetting* (wild guess), *delvis suksess* (near misses), *operasjonsforvirring* (operation confusions) og *tabellfeil* (table errors) (Baroody, 1989; Siegler & Shrager, 1984). Gjetting er vanlig hos barnehagebarn, hvor de for eksempel konstaterer at «4 + 1 = 41» eller at de gjentar en av addendene som svaret (Baroody, 1989). Delvis suksess involverer å gjenhente et svar som er en eller to høyere enn riktig sum og speiler barnets tidligere tellefeil (Siegler & Shrager, 1984). Ved operasjonsforvirring henter barn det riktige svaret til et tilsvarende problem med en annen aritmetisk operasjon, som for eksempel at barnet gjenhenter «12» når de blir spurt om å løse «4 + 3» (Geary, 1994). Disse retrievalfeilene er mest vanlig hos de barna som lærer nye aritmetiske operasjoner (Geary, 1994). Tabellfeil involverer gjenhenting av svaret til et relatert problem, som å gjenhente «12» for problemet «6 +7» når det korrekte svaret til «12» er «5 + 7». Det ser også ut som at tabellfeil er relatert til hvordan aritmetisk fakta blir representert i langtidsminne (Aschcraft, 1992; Geary, 1994).

En god utvikling i aritmetikk kjennetegnes av en gradvis skifte i bruk av backupstrategier (tellestrategier), til økt bruk av retrievalstrategier (Geary, 1993; Ostad, 1999). Barn med vansker i aritmetikk viser å anvende mer umodne strategier for å løse aritmetiske problemer sammenlignet med jevnaldrende i samme klassetrinn (Ostad, 2013). Strategibruken til barn med vansker kjennetegnes av ensidig valg av backupstrategier og de velger de mest primitive backupstrategiene, som telle alt og telle fra første (Ostad, 1997, 2013). Barn med matematikkvansker bruker også relativ lang tid på å finne løsningen på problemet, samt at de ofte begår prosedyre- og minneretrievalfeil (Geary; 1990; Geary et al., 1991). Barna kjennetegnes også ved liten variasjon i valget mellom de ulike strategiprosedyrene og de har en lav endringsgrad i strategivalget fra år til år gjennom grunnskolen. Barn med aritmetiske vansker viser også manglende bruk av retrievalstrategier. Manglende bruk av retrievalstrategier, og hyppig og ensidig bruk av backupstrategier vil kunne representere en kritisk faktor for normal utvikling (Ostad, 2013).

2.2.2 Vansker med aritmetiske problemløsningsoppgaver

Barns mestring av aritmetiske problemløsningsoppgaver påvirkes av alder, hvordan problemet blir representert og hvilke regneoperasjoner som er involvert (Gilmore, et al., 2018). Innenfor de ulike situasjonene i problemløsningsoppgavene er det forskjeller på hvilke problemer barn finner mer vanskelig å løse enn andre (Riley et al., 1983; Verschaffel & De Corte, 1997). De feilene barn normalt gjør i aritmetiske problemløsningsoppgaver er ikke utregningsfeil, men er knyttet til vansker med å danne et bilde av problemet (Gilmore et al., 2018).

Ved aritmetiske problemløsningsoppgaver hvor spørsmålet er knyttet til resultatet (a+b = ?) er det relativt lett for barn å løse, men oppgaver hvor problemet er representert ved at den ukjente verdien er presentert som første tall (?+a = b) eller som andre tall (a+? = b) er vanskeligere (Riley et al., 1983; Thevenot & Barrouillet, 2015). Endringsproblemer med ukjent resultat og sammensetningsproblem med ukjent sammensetningsverdi bruker å være lettere å løse enn oppgaver med endringsproblem med enten ukjent start eller ukjent endring, eller ved sammensetningsproblem med ukjent delverdi (Riley et al., 1983; Verschaffel & De Corte, 1997).

Endringsproblemer er lettere for andre- og tredjeklassinger enn for barnehagebarn og førsteklassinger (Thevenot & Barrouillet, 2015). Ved sammensetningsproblemer er ukjent sammensetningsverdi lettere for barn å løse enn med ukjent delverdi, noe som barn opplever stor vanske med i barnehagealder og førsteklasse (Thevenot & Barrouillet, 2015). De problemene som viser seg å være spesielt vanskelige å løse er sammenligningsproblemer, spesielt ved ukjent referent (Geary, 1994; Thevenot & Barrouillet, 2015; Verschaffel & De Corte, 1997). Sammenligningsproblemer som hjelper barn til å utvikle en mental modell av sammenligningen synes å være enklere å løse enn andre former for sammenligningsproblemer (Geary, 1994).

2.3 Effekten av alder i akademiske fag

Som nevnt innledningsvis går alle barn i Norge som er født det samme året i samme klassetrinn med få unntak (Olsen & Bjørnsson, 2018; Solli, 2017). Denne variasjonen i alder ved skolestart eksisterer fordi skolestart er sterkt knyttet til fødselsdato, og det er som regel en bestemt dato som avgjør om en starter på skolen det samme året eller året etter (Bedard & Dhuey, 2006; Olsen & Bjørnsson, 2018). I Norge er cut-off datoen for skolestart 1. januar, selv om skoleåret starter i august og slutter i juni (Aune, Ingvaldsen, Vestheim, Bjerkeset & Dalen, 2018). Forskjellen i alder innenfor samme aldersgruppe blir ofte kalt for den relative alderseffekten (Dalen & Aune, 2013; Thoren et al., 2016). I utdanningssammenheng blir den relative alderseffekten sett i sammenheng mellom hvor sterkt barns alder er relatert til prestasjoner på faglige prøver i samme klasserom (Olsen & Bjørnsson, 2018). På bakgrunn av at Norges cut-off dato for skolestart er 1. januar gjør at data fra Norge er spesielt egnet for å studere relativ alderseffekt. Dette er fordi elevene vil ha like lang tid i skolen når de deltar i en studie (Olsen & Bjørnsson, 2018).

Relativ alderseffekt har vært et kjent fenomen i kroppsøving og i ulike sportsgrener (Aune et al., 2017; Cobley et al., 2008; Sæther et al., 2017). Grupperingen av ungdom i sport etter kronologisk alder, for både gutter og jenter, pleier å gi fordel til de spillerne som er født tidlig på året i utvelgelsesåret (Sæther et al., 2017). Denne effekten kan også ses i sjakk, som er en sport som involverer kognitive oppgaver og stiller få fysiske krav (Helsen et al., 2016). Det er flere studier viser at de eldste elevene presterer bedre enn sine yngre medelever i akademiske fag, inkludert matematikk (Bedard & Dhuey, 2006; Cobley et al., 2009; Kawaguchi, 2011; Solli, 2017; Strøm, 2004; Thoren et al., 2013). Relativ alderseffekt i akademiske fag og sport er funnet i studier fra blant annet Norge (Dalen & Aune, 2013; Olsen & Bjørnsson, 2018; Solli, 2017), Tyskland (Thoren et al., 2016), Japan (Kawaguchi, 2011) og England (Cobley et al., 2009; Crawford et al., 2013). Relativ alderseffekt er også observert i ulike aldersgrupper, i barneskolealder og i ungdomsskolen (Cobley et al., 2009; Kawaguchi, 2011; Olsen & Bjørnsson, 2018; Thoren et al., 2016; Strøm, 2004), samt i videregående skole (Dalen & Aune, 2013; Solli, 2017).

2.3.1 Studier på relativ alderseffekt i grunnskolen

Studier gjort på relativ alder i grunnskolen viser til at det er forskjeller mellom de eldste og yngste elevene i akademiske fag (Crawford et al., 2013; Olsen & Bjørnsson, 2018). Olsen og Bjørnsson (2018) fant signifikante forskjeller i alder hos norske fjerdeklassinger i fagene matematikk og naturfag fra *Trends in International Mathematics and Science Study* (TIMSS) undersøkelsene (1995, 2003, 2007, 2011 og 2015). Det ble gjennomført analyser for jenter og gutter i begge fagområder og de fant ingen kjønnsforskjeller. Elevenes sosiale bakgrunn ble kontrollert for og den relative alderseffekten var upåvirket av dette (Olsen & Bjørnsson, 2018). Olsen og Bjørnsson (2018) fant også signifikante forskjeller i relativ alder i 8. trinn i naturfag og matematikk fra TIMSS-undersøkelsene, samt i *Programme for International Student Assessment* (PISA) undersøkelsen for 10.trinn. I matematikktesten fra PISA-undersøkelsen fra 2015 representerer forskjellen mellom eldste og yngste elev omtrent 16 poeng. Når de så på kjønn fant de ingen forskjeller i 8.trinn, men i 10.trinn hadde guttene høyere relativ alderseffekt enn jentene. Effekten de fant i 10.trinn var kun beskjedent signifikant. Funnene fra studien viser at resultatene i matematikk faller mellom eldste og yngste fra omtrent 30 poeng i 4. trinn ned til 10-15 poeng i 8. og 10.trinn. Studien viser dermed at relativ alderseffekt avtar i løpet av den obligatoriske skolegangen, men den er enda betydelig for 15/16 åringene som deltar i PISA-undersøkelsen (Olsen & Bjørnsson, 2018). Kawguchi (2011) fant også signifikante forskjeller i relativ alder for 4. trinn i fagene matematikk og naturfag i Japan fra TIMSS-undersøkelsen fra 2003. Japan har et lignende skolesystem som Norge hvor de som er født det samme året går i samme klassetrinn med få unntak. Aune et al. (2018) undersøkte i sin studie norske elevers prestasjoner på nasjonale prøver i regning fra 2012. Studien viser konsistente relativ alderseffekt gjennom 5. trinn, 8. trinn og 9. trinn for både jenter og gutter. Gjennomsnittskårene var systematisk redusert med fødselsmåned hos begge kjønn, men gjennomsnittspoengene var høyere hos guttene enn hos jentene. Studien viser at gutter som er født tidlig på året er overrepresentert av de som har høy skåre, mens jenter som er født sent på året er overrepresentert av de som har lav skåre (Aune et al., 2018).

Det er også funnet signifikante forskjeller hos norske elever i lesing i 4. og 5. trinn fra *Progress in International Reading Literacy Study* (PIRLS) undersøkelsene (2001, 2006, 2011 og 2016) som viser at barn som er født tidlig på året presterer bedre enn elevene som er født sent på året (Gabrielsen & Lundetræ, 2017). Strøm (2004) fant samme tendenser i resultatene i lesing fra PISA-undersøkelsen fra 2000 hos norske tiendeklassinger. De elevene som var født sent på kalenderåret presterte signifikant lavere i lesing sammenlignet med sine eldre medelever. Han fant at effektene var omtrent lik når det ble sjekket for familiebakgrunn, men at det var svake bevis på at elever som er født sent i kalenderåret og som har foreldre med høy utdanning kommer dårligst ut (Strøm, 2004). Thoren et al. (2016) fant alderseffekter i lesing og matematikk til fordel for de relativt eldre tyske andreklassinger. I tredjeklasse var denne effekten mindre, men enda til fordel til de eldre elevene. De fant også at det ikke var noen betydelige forskjeller i relativ alder mellom elever med og uten innvandringsbakgrunn. I 8.trinn fant de at effekten av relativ alder hadde forsvunnet i lesing og hadde reversert seg til fordel for de yngste elevene i matematikk.

Cobley et al. (2009) fant lignende funn i sin studie i Nord-England i alderen 11 til 14 år. Resultatene viste en aldersforskjell i fagene matematikk og naturfag, men ikke i engelsk basert på vurderinger fra den nasjonale læreplanen (Key stage 3). De fant at den relative alderen spesielt påvirket prestasjonene i matematikk (Cobley et al., 2009). I fagene engelsk, matematikk og naturfag presterte jenter signifikant bedre enn guttene, men de eldste elevene skåret generelt høyere uansett aldersgruppe, kjønn og fagområde (Cobley et al., 2009). I en annen studie fra England fant Crawford et al. (2013) forskjeller mellom de eldste og de yngste i akademiske ferdigheter basert på resultater fra den nasjonale læreplanen (Key stage 1, 2, 3, 4). Funnene viser at i syv årsalderen er barn som er født i august måned (de yngste) er 26 prosentpoeng mindre sannsynlig til å nå regjeringens forventede nivå enn sine medelever som er født i september måned (de eldste). Ved slutten av grunnskolen (11 år) er denne forskjellen gått ned til 13 prosentpoeng. Dette tyder på at barn som er yngst i sitt skoleår kommer på nivå med sine medelever over tid, da forskjellen i relativ alder avtar (Crawford et al., 2013). Disse forskjellene er størst ved skolestart og avtar gradvis, men de er enda statistisk- og utdanningssignifikant ved 16 årsalderen (Crawford et al., 2013).

Resultatene fra flere studier tyder på at de eldste i et klassetrinn presterer bedre enn de yngste i fag som matematikk, men at forskjellen er størst ved skolestart og minsker utover klassetrinnene (Bedard & Dhuey, 2006; Crawford et al., 2013; Kawaguchi, 2011; Strøm, 2004; Thoren et al., 2016). På bakgrunn av bredden i alder ved skolestart vil de eldste elevene i klassen sannsynlig være mer modne enn de yngste (Bedard & Dhuey, 2006), samtidig viser relativ alderseffekt til heterogenitet hos elevgruppen i en klasse (Thoren et al., 2016). Elevene som går i samme klassetrinn vil også konkurrere og vurderes mot hverandre (Dalen & Aune, 2013). De yngste elevene er også overrepresentert blant de som ble diagnostiert med en spesifikk lærevanske, og at de har mer spesialundervisning (Cobley et al., 2009; Crawford et al., 2013). Barn som er født i september måned har i forhold til de som er født i august måned 5.4 prosentpoeng høyere sannsynlighet til å få svake spesialpedagogiske behov i elleve årsalderen (Crawford et al., 2013). Det er funnet få bevis på at disse forskjellene vedvarer til voksen alder (Crawford et al.,2013), men elever som starter bak har tendenser til å holde seg bak sine medelever (Geary, 1993). Det å være født sent i skoleåret syns å ha en negativ effekt på barns velvære ved for eksempel at de har lavere selvtillit og at det er større sannsynlighet for at de engasjerer seg i risikofylt atferd i yngre alder (Crawford et al., 2013). Samt at de yngste elevene var mer sannsynlig til å utvikle negative holdninger og atferd mot utdanning (Cobley et al., 2009). Relativ alderseffekt påvirker ikke bare elevene og lærerne i hverdagen, men kan gi konsekvenser i hele livsløpet når det gjelder akademiske ferdigheter, atferd og holdninger (Cobley et al., 2009; Dalen & Aune, 2013; Thoren et al., 2016).

2.3.2 Studier på relativ alderseffekt i videregående skole

Solli (2017) fant i samsvar med tidligere forskning at de eldste elevene presterer signifikant bedre i skolen enn sine yngre medelever basert på gjennomsnittskarakterer fra 10.trinn, men at med tid så tar de yngste igjen de eldste. Denne forskjellen har i tiendeklasse et standardavvik på rundt 20 % mellom eldste og yngste, samt at effekten var relativ lik for jenter og gutter. På langtidsutfallene viser resultatene at barn som er født tidlig på året har større sannsynlighet enn sine yngre medelever for å fortsette direkte til videregående skole og uteksamineres ved 19 årsalderen, samt større sannsynlighet for å starte på universitet og høyskole (Solli, 2017). Dalen og Aune (2013) fant relativ alderseffekt i kroppsøving, norsk skriftlig hovedmål og matematikk. De fant at denne effekten var tydeligst i kroppsøving sammenlignet med de mer kognitive fagene som norsk og matematikk i tiendeklasse og i utdanningsprogrammet studiespesialiserende i VG2 og VG3 på videregående skole. De fant signifikant alderseffekt i kroppsøving i alle undersøkte klassetrinn. Dette er tydeligst ved antall toppkarakterer, hvor de som er født første halvår (jan.-jun.) har høyere antall toppkarakterer enn de som er født siste halvår (jul.-des.). De fant samme tendens i norsk skriftlig hovedmål, men de fant ingen signifikante forskjeller i relativ alderseffekt i matematikk selv om det er en prosentvis nedgang totalt fra de som er født i januar til de som er født i desember (Dalen & Aune, 2013).

2.3.3 Andre årsaker til forskjell i prestasjoner hos eldste og yngste i en alderskohort

Crawford et al. (2013) trekker frem i sin rapport at det er andre potensielle grunner, i tillegg til relativ alderseffekt som gjør at vi ser ulike resultater mellom barn som er født tidlig på året og barn som er født på slutten av året. En av grunnene kan være barns alder når de blir testet (age-at-test effect). Hvis alle barn som er født det samme året gjennomfører en test på samme dato, vil de barna som er født sent på året alltid være yngre enn sine medelever når de gjennomfører tester (Crawford et al., 2013). En annen grunn kan være på bakgrunn av barn starter på skolen i ulik alder (age-of-starting-school effect). De som er født sent på året bruker å være å være opptil ett år yngre enn de som er født tidlig på året. Hvis disse barna ikke er klare til å starte på skolen (for eksempel hvis de ikke klarer å delta i mer formell undervisning som kommer med skolen sammenlignet med barnehagen), kan dette hjelpe å forklare deres lave skårer på tester (Crawford et al., 2013). En annen grunn til ulike resultater mellom de eldste og yngste elevene i et klassetrinn kan komme av mengden undervisning barna har før en test (length-of-schooling effect). Noen land følger en opptakspolicy der barn som er født sent på året starter på skolen ett til to semestre etter de som er født i starten av året. Det betyr videre at de har ett til to semestre mindre med undervisning ved for eksempel nasjonale prøver i England som går på alder (Crawford et al., 2013). Crawford et al. (2013) fant at alder ved testing er en nøkkelfaktor til ulikhetene i skårer i utdannelsesnivå og skårer på kognitive tester mellom barn som er født tidlig på året og sent på året. Dette betyr (i norsk sammenheng) at barn som er født i desember bruker å gjøre det svakere på tester enn barn som er født i januar på bakgrunn av at de er elleve måneder yngre ved testdato. De fant videre at alderen ved skolestart og mengden undervisning forklarte lite av denne forskjellen, men at relativ alder kan være med på å forklare den gjenværende forskjellen.

2.4 Oppsummering av teori og empiri

Aritmetikk består ikke bare av en ferdighet, men av flere (Dowker, 1998, 2005). Disse ferdighetene består blant annet av kunnskap om tall og de fire grunnleggende regneartene, memorere aritmetisk fakta, følge aritmetiske prosedyrer og forståelse av aritmetiske prinsipper (Dowker, 1998, 2005; Haskell, 2000). Utvikling av aritmetiske ferdigheter innebærer en endring i bruk av strategier, og hvor nøyaktig og hurtig strategiene kan bli utført. Tidlig tallkunnskap og telleferdigheter vil være med på å danne grunnlag for barns aritmetiske problemløsningsferdigheter (Geary, 1994).

Strategiene som blir brukt for å løse aritmetiske problemer blir delt inn i strategier som involverer bruk av konkreter, verbale strategier og retrievalstrategier (Gilmore et al., 2018). De blir også delt inn i backupstrategier og retrievalstrategier (Ostad, 2013). Retrievalstrategier kjennetegnes ved at en kan hente frem kunnskapsenheter fra minnet og backupstrategier omfavner de øvrige strategiene (Siegler & Jenkins, 1989). Strategiene skiller seg fra hverandre ved hvor nøyaktige de er, tiden de trenger for å gjennomføres, hvor mye minnekapasitet de krever og i hvor stor bredde de har for å benyttes på de ulike aritmetiske problemene (Siegler & Shrager, 1984, Siegler, 1991). Barn starter med å kunne respondere på ulike addisjons- og subtraksjonsoppgaver først gjennom gjetting eller ved å gjøre et omtrentlig estimat (Jordan et al., 2003), deretter til å løse enkle aritmetiske oppgaver ved bruk av telling (Geary, 1994: Jordan et al., 2003). I løpet første- og andreklasse vil barn utvikle raskere og mer effektive tellestrategier (Jordan et al., 2003), samt retrievalstrategier (Jordan et al., 2003; Ostad, 1997). I aritmetiske problemløsningsoppgaver vil backupstrategier ha en fremtredende rolle i førsteklasse og gjennom barneskolen (Ostad, 1997).

Aritmetiske problemløsningsoppgaver er med på å gi et innblikk i barns aritmetiske ferdigheter (Gilmore et al., 2018; Thevenot & Barrouillet, 2015). Når barn skal løse aritmetiske problemløsningsoppgaver må de danne seg et bilde av problemet, hente ut relevant informasjon, velge riktig regneoperasjon og deretter gjennomføre den valgte regneoperasjonen (Geary, 1994; Gilmore et al., 2018). Aritmetiske problemløsningsoppgaver fanger et bredt spekter av situasjoner (Gilmore et al., 2018) og de blir klassifisert etter deres semantiske struktur (Geary, 1994). I addisjons- og subtraksjonsproblemløsningsoppgaver skilles det ofte mellom tre situasjoner av problemer: *endringsproblem*, *sammensetningsproblem* og *sammenligningsproblem* (Gilmore et al., 2018). Mange barn kan løse enkle aritmetiske problemløsningsoppgaver før de starter på skolen (Levine et al., 1992), og de fleste barnehagebarnene kan også løse endringsproblem med ukjent resultat (Geary, 1994; Jordan et al., 2003). I løpet av første- til tredjeklasse blir barns aritmetiske problemløsningsferdigheter stadig mer sofistikerte (Riley & Greeno, 1988). I løpet av denne perioden vil flere av elevene mestre enkle problemløsningsoppgaver med ukjent endring og oppgaver som involverer sammenligning (Jordan et al., 2003).

For noen elever er innlæring av matematikk en kilde til vanskeligheter, og de viser vansker med innlæringen selv om de har en normal intelligens og utdanningsmuligheter (Nöel, 2015). For mange kommer vanskene til uttrykk gjennom langsom innlæring av tallkonsepter og grunnleggende aritmetikk (Geary, 2017). Barn viser også vansker med å lære ferdigheter som tallforståelse, aritmetiske prosedyrer, regneflyt, og vansker med innlæring og gjenhenting av aritmetisk fakta (World Health Organization, 2018; Geary, 2004; Nöel, 2015). Barna bruker også mer umodne strategier når de skal løse oppgaver som aritmetiske problemløsningsoppgaver (Geary, 2004; Jordan et al, 2003; Nöel, 2015).

Relativ alderseffekt har vært et kjent fenomen i kroppsøving og i ulike sportsgrener (Aune et al., 2017; Cobley et al., 2008; Sæther et al., 2017), samt at relativ alderseffekt kan ses i sjakk (Helsen et al., 2016). Det er også blitt gjort flere studier på relativ alderseffekt og akademiske fag, som matematikk på tvers av landegrenser og aldersgrupper (Bedard & Dhuey, 2006; Cobley et al., 2009; Kawaguchi, 2011; Solli. 2017; Strøm, 2004; Thoren et al., 2013). Generelt viser funnene fra studiene at det finnes relativ alderseffekt i akademiske fag og i matematikk, men at disse effektene avtar i skoleløpet. Funnene viser derimot at det enda er betydelig alderseffekt i ungdomsskolen (Aune et al., 2018; Olsen & Bjørnsson, 2018; Solli, 2017). Crawford et al. (2013) trekker frem andre potensielle grunner, i tillegg til relativ alderseffekt, som gjør at vi ser ulike resultater mellom barn som er født tidlig på året og barn som er født på slutten av året.

3 Metode

I dette kapittelet vil det gjøres rede for forskningsprosessen for denne studien. Denne studien er knyttet til intervensjonsprosjektet *The vocabulary learning challenge* (VLC) ved Instituttet for Spesialpedagogikk, Universitetet i Oslo. Denne studien vil dermed forholde seg til de rammene som dette gir. Først vil det gjøre redes for den metodiske tilnærmingen, utvalg, måleinstrumenter og datainnsamling. Deretter vil det gjøres rede for validitet og reliabilitet. Videre vil analysene for denne studien presenteres. Avslutningsvis vil det etiske perspektivet belyst med fokus på rollen som forsker og etiske vurderinger i en forskningsprosess.

3.1 Design

Forskningsdesign er en helhetlig oversikt over de ulike fasene som er i en forskningsprosess. Forskningsprosessen består av utarbeiding av mål, begrepsrammeverk, forskningsspørsmål, valg av metode, innhenting av data, analyse av data og resultater og konklusjoner (Lund, 2002a). På bakgrunn av problemstillingen *«Er det sammenheng mellom alder og aritmetiske ferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?»* vil studien besvares gjennom en kvantitativ tilnærming. Denne studien har som formål å beskrive sammenhengen mellom ulike variabler, uten å påvirke individene som deltar. Denne studien vil da være av et ikke-eksperimentelt design, også kalt en deskriptiv studie (Kleven, 2002a). I et ikke-eksperimentelt design undersøker en hvordan tingenes tilstand er, uten å forsøke å påvirke tilstanden (Kleven, 2002a). Studien vil beskrive grad av sammenheng mellom de ulike variablene og ikke forklare hvorfor det er en sammenheng (Kleven, 2002a). For å beskrive datamaterialet vil det bli gjort en deskriptiv analyse, og for å beskrive sammenhengen mellom ulike variabler vil det blir gjort bivariat korrelasjonsanalyse og hierarkisk multippel regresjonsanalyse. Denne studien tar for seg datamaterialet fra et målepunkt, og er dermed en tverrsnittstudie. Ved en tverrsnittstudie kan en ikke si noe om utviklingen over tid, men kan si noe om dataen fra et gitt tidspunkt og forholdet mellom variablene på dette tidspunktet (Gall, Gall & Borg, 2007).

Formålet med denne studien blir derfor å se om alder er en mulig prediktor for aritmetiske ferdigheter, og om en eventuell påvirkning kan forklares av andre faktorer. I analysene vil alder i måneder være uavhengig variabel og aritmetiske ferdigheter vil være den avhengige variabelen. Kontrollvariablene i denne studien vil være nonverbale evner og kjønn.

Problemstillingen spør om det er kvalitative forskjeller i det å være født sent på året sammenlignet med det å være født tidlig på året som kan føre til ulike forutsetninger for aritmetiske ferdigheter i andreklasse. Nullhypotesen (H_0) vil i denne studien være at det ikke fremkommer forskjeller i aritmetiske ferdigheter hos de som er født tidlig på året og de som er født sent på året. H_1 vil derfor bli at det er forskjeller i aritmetiske ferdigheter hos de som er født tidlig på året og de som er født sent på året.

3.2 Utvalg

Utvalget består av 332 andreklassinger, 179 jenter og 153 gutter fra 12 skoler fordelt på 32 klasser i tre kommuner på Østlandet. Ved det aktuelle måletidspunktet var elevenes gjennomsnitts alder i måned på 90.81 (SD = 3.41), hvor yngste elev var 84.17 måneder gammel og eldste elev var 97.28 måneder gammel. Rekrutteringen ble gjort ulikt i de tre kommunene, men i korte trekk kontaktet prosjektgruppen i VLC skoleetaten og skoleetatene ga så tilbakemeldinger om hvilke skoler som var iPad-skoler (dvs. at skolen hadde iPad tilgjengelig for alle elevene på trinnet under intervensjonsperioden), som var et krav for å delta i studien. Deretter tok VLC-gruppen kontakt med skoleledelsen ved de enkelte skolene og sendte informasjonsskriv om prosjektet. Ved tilbakemelding om ønsket deltakelse fra skolene og kontaktlærere, henvendte de seg videre til foresatte til alle elevene hvor klassen skulle delta. Det ble sendt ut informasjonsbrev og samtykkeerklæring som foresatte måtte underskrive på hvis elevene skulle delta i prosjektet, og 718 takket ja til å delta i intervensjonsprosjektet. Lærere og ledelsen i de ulike skolene som deltok ble invitert på et felles informasjonsmøte før prosjektstart. På bakgrunn av at min studie tar for seg norske enspråklige andreklassinger ble familietospråklige og minoritetspråklige elever ekskludert (281 elever). Videre ble barn som var født utenfor alderskohorten ekskludert (8 elever) og de elevene som ikke hadde gjennomført alle relevante tester som skal brukes i denne studien (97 elever). Bakgrunnen for at ikke alle elevene hadde gjennomført alle testene var på bakgrunn av manglende ressurser.

3.3 Måleinstrumenter

Denne studien tar utgangspunkt i resultatene fra pre-testene i VLC-prosjektet og informasjon fra spørreskjemaet til foresatte. Testene som ble gjennomført i klasserommet var Raven’s Coloured Progressive Matrice (Raven), British Picture Vocabulary Scale (BPVS), diktat og

Regnefaktaprøven - addisjon og subtraksjon. Testene som ble gjennomført i den individuelle kartleggingen var prosjektspesifikk lesetest, nonordrepitisjon, prosjektspesifikk ordforklaring, Test of Word Reading Efficiency (TOWRE), Wechlser Intelligence Scale for Children (WISC-IV) ordforståelse, WISC-IV regning og prosjektspesifikk lesetest. Denne studien har som formål å se på aritmetiske ferdigheter og kontrollerer for nonverbale evner, og på denne bakgrunn vil det bli tatt utgangspunkt i Regnefaktaprøven - addisjon og subtraksjon, WISC-IV regning og Raven. Samt informasjon fra spørreskjemaet fra foresatte om «barnets kjønn» og «barnets fødselsdato».

3.3.1 Regnefaktaprøven

Regnefaktaprøven (Klausen & Reikerås, 2016) er et normert og standardisert kartleggingsverktøy for 2. til 10. trinn for å måle regneflyt. Regnefaktaprøven består av oppgaver i de fire regneartene: addisjon, subtraksjon, multiplikasjon og divisjon. I VLC-prosjektet ble delprøvene som omhandler addisjon og subtraksjon i tallområdet 0-20 benyttet og er standardisert for 2. til 10. trinn. Hver delprøve består av 45 oppgaver og hvor elevene skal løse flest mulig oppgaver på to minutter i en bestemt rekkefølge, se figur 1 for eksempler. Regnefaktaprøven er laget for å gi en indikasjon på hvor effektive regnestrategier elevene bruker når de skal løse regneoppgaver, og det vil gi en indikasjon på hvordan regneflyten til elevene er. For at elevene skal oppnå optimal regneflyt må de utvikle tallforståelse og effektive regnestrategier. Regnefaktaprøven gir derfor et mål på effektivitet som vil si at den måler nøyaktigheten og regnehastigheten i addisjon og subtraksjon.

7 + 4 =	10 – 2 =
8 + 5 =	12 – 9 =
6 + 7 =	15 – 5 =
4 + 9 =	8 – 5 =
4 + 4 =	7 – 3 =

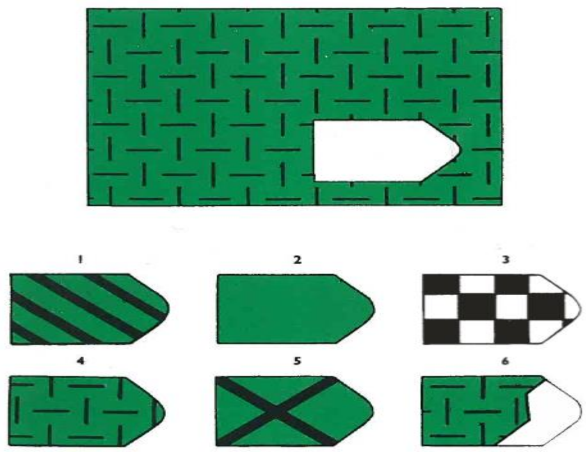
Figur 1. Eksempler på oppgaver i Regnefaktaprøven addisjon og subtraksjon

3.3.2 WISC-IV regning

WISC-IV regning (Wechsler et al., 2003) måler barnets matematiske problemløsningsferdigheter. WISC-IV regning er standardisert for aldersgruppen 6-16 år og består av muntlig presenterte tekstoppgaver som skal løses ved hjelp av hoderegning uten bruk av hjelpemidler. Testen består av 34 oppgaver som øker i vanskelighetsgrad, og hvor hver oppgave har en tidsbegrensning på 30 sekunder. De første fem oppgavene blir presentert med bilder. Oppgavene blir presentert som en fortelling med kjente elementer, uten at det oppgis hvilken strategi som må benyttes for å løse oppgaven. Eksempel på oppgaver fra WISC-IV regning er «Hvis du har 10 epler og spiser 4, hvor mange epler har du igjen?». Testen avbrytes etter fire påfølgende nullpoengssvar (inkludert over tid) og det gis null og ett poeng for hver oppgave.

3.3.3 Ravens Coloured Progressive Matrice

Raven (Raven, 1998) består av en rekke av ikke-verbale oppgaver for å vurdere barns nonverbale evner. I VLC-prosjektet ble Ravens Coloured Progressive Matrice benyttet siden den er beregnet for yngre barn. Testen består av 36 items, hvor de to første oppgavene er øvelsesoppgaver. Barnet blir vist en stimulusbok hvor hver side inneholder et bilde av en rektangel der en bit mangler, og nedenfor er det presentert seks ulike biter som barnet kan velge mellom. Barnet skal deretter peke på den biten de mener passer best i den manglende delen av rektangelet. Testen har ingen tidsbegrensning.



Figur 2. Eksempl på øvingsoppgaver fra Raven

3.4 Datainnsamling

Datainnsamlingen for denne masteroppgaven ble innhentet vår 2018, og ble gjennomført av masterstudenter og vitenskapelige ansatte ved VLC-prosjektet. Alle studentene som deltok fikk en grundig opplæring i testbatteriet i regi av prosjektgruppen, samt at studentene måtte gjennomføre testbatteriet på to barn i åtte årsalderen i forkant av pre-testen. I tillegg til testbatteriet ble det utlevert spørreskjema til elev, foresatte og kontaktlærer. Pre-testen ble gjennomført i perioden januar-februar. Kartleggingen ble utført i tre økter på tre forskjellige dager. Kartleggingen ble først gjennomført på gruppenivå i klasserom, samt kartlegging via app. Deretter ble det gjennomført individuell kartlegging i et eget og avskjermet rom. Dersom elevene ikke var til stede under gruppekartleggingen eller av andre grunner ikke gjennomførte kartleggingen i klasserommet, ble den utført i mindre grupper eller individuelt.

Gruppekartleggingene i pre-testen tok rundt 90 minutter per klasse og app-kartleggingen tok rundt 25 til 30 minutter. Ved gjennomføringen av Raven ble øvelsesoppgavene gjennomgått i plenum, og deretter løste elevene de resterende oppgavene individuelt og i deres eget tempo. Regnefaktaprøven ble gjennomført i to deler hvor en del bestod av 45 addisjonsoppgaver og den andre delen besto av 45 subtraksjonsoppgaver. Den individuelle kartleggingen tok rundt 30 til 60 minutter per elev. Rekkefølgen på de individuelle testene ble systematisk lagt opp slik at det skulle være variasjon i oppgavene og at liknende oppgaver ikke kom etter hverandre. Dette skulle bidra til å holde elevenes motivasjon, mestringsfølelse og interesse oppe. I tillegg ble det utlevert en diplom med klistremerke for hver gjennomførte test og klassene som deltok mottok en liten premie for deltakelse som ble gitt samlet til klassen.

3.5 Validitet og reliabilitet

Cook og Campells validitetssystem blir ofte brukt som en metodologisk referanseramme i kvantitativ forskning (Lund, 2002a). Validitet refererer til om måledata gir et reelt og pålitelig bilde av det fenomenet eller variabelen en har som formål på å måle, eller om data i en betydelig grad er «forurenset» av andre faktorer (Befring, 2015). I en kvantitativ studie er det viktig at en kan trekke valide slutninger. Ved svak validitet vil det påvirke konklusjonen til å bli tilsvarende usikker, og omvendt ved sterk validitet (Lund, 2002a). Validitet handler i den forstand om en kan trekke gyldige slutninger basert på måledata. I denne studien vil det handle om en kan trekke gyldige slutninger om alder påvirker andreklassingenes ferdigheter i aritmetikk. Validitetssystemet omfatter fire kvalitetskrav eller typer validitet; *statistisk validitet*, *indre validitet*, *begrepsvaliditet* og *ytre validitet* (Lund, 2002a; Shadish, Cook & Campbell, 2002). Reliabilitet er et uttrykk på hvor nøyaktig eller konsistent en test måler det den er ment å måle (Kleven, 2002b). I metodekapittelet vil det bli gjort kort rede for de ulike validitetskravene og reliabilitet. Drøftingen av studiens validitet i lys av disse kravene vil bli gjort i kapittel 5.

3.5.1 Statistisk validitet

Statistisk validitet handler om den antatte årsaken og effekten korrelerer, og hvor sterkt de eventuelt korrelerer (Shadish et al., 2002). En undersøkelse vil ha god statistisk validitet hvis det kan trekkes pålitelige slutninger om sammenhengen mellom den avhengige og uavhengige variabelen, eller om siktepunktet er statistisk signifikant og rimelig sterk. Hvis en undersøkelse er statistisk invalid vil sammenhengen mellom variablene eller siktemålet skyldes en målefeil, eller at størrelsen til siktemålet er triviell (Lund, 2002a). Ved statistisk invaliditet vil heller ikke indre validitet være oppfylt, og statistisk validitet kan være oppfylt selv om de tre øvrige validitetstypene ikke er tilfredsstillende (Lund, 2002a). Trusler mot statistisk validitet er brudd på statistiske forutsetninger og lav statistisk styrke (Lund, 2002a), samt upålitelighet i måleinstrumentene (Shadish et al., 2002). Brudd på statistiske forutsetninger kan innebære brudd på normalitet, lik varians og uavhengighet i observasjoner. Dette kan føre til enten en overestimering eller underestimering av størrelse og signifikansen til en effekt (Shadish et al., 2002). Dette kan få konsekvenser for både type I-feil, som vil si at en forkaster en sann nullhypotese og type II-feil, som vil si at en aksepterer en ukorrekt nullhypotese (Lund, 2002a). Lav statistisk styrke vil kunne komme av lite utvalg, strengt signifikansnivå og små effektstørrelser (Gall et al., 2007; Lund, 2002a). Dette vil gi større sannsynlighet for type II-feil (Lund, 2002a). Upålitelighet i måleinstrumentene kan føre til at bivariate forhold reduseres. Reliabilitet burde bli estimert og rapportert for hvert måleinstrument som blir brukt i en studie (Shadish et al., 2002). Reliabiliteten til måleinstrumentene vil bli vurdert senere i dette kapittelet.

3.5.2 Indre validitet

Indre validitet handler om en kan trekke pålitelige slutninger om sammenhengen mellom variablene er kausal, uten at det kan forklares av andre faktorer. Indre validitet vil i ikke-eksperimentelle studier være et problem, og spesielt i tverrsnittstudier som denne studien er (Shadish et al., 2002). Ikke eksperimentelt design er opptatt av å beskrive grad av sammenheng, og vil dermed ha lavere indre validitet på bakgrunn av at sammenhengen alltid vil være forenlig med flere mulige årsakrelasjoner (Kleven, 2002a). Trusler mot indre validitet kan komme av ulike faktorer. En mulig trussel for indre validitet er en tredjevariabler. I denne studien er målet å se på sammenhengen mellom alder og aritmetiske ferdigheter og det vil kontrolleres for nonverbale evner og kjønn. En kan enda ikke være sikker om det er en tredjevariabel som ikke er blitt kontrollert for som kan forklare variasjonen mellom alder og aritmetiske ferdigheter. Resultatene må derfor tolkes i lys av konfunderende faktorer som kan ha hatt en innvirkning i resultatene. En annen trussel er retningsproblem, og det har med å avgjøre hva som er årsak og hva som er virkning (Shadish et al., 2002). En faktor som kan styrke den indre validiteten er valg av design som kontrollerer eksperimentelt eller statistisk for alternative årsaksforhold (Lund, 2002a).

3.5.3 Begrepsvaliditet

Begrepsvaliditet handler om samsvaret mellom hvordan begrepet er teoretisk definert og hvordan begrepet er operasjonalisert gjennom måling (Kleven, 2002b). En studie har god begrepsvaliditet hvis den uavhengige og avhengige variabelen måler de relevante begrepene i forskningsproblemet (Lund, 2002a). En vurdering av begrepsvaliditet er en vurdering på om det målte begrepet oppfører seg slik som det teoretiske begrepet forventer å oppføre seg. Reliabilitet er ikke noe mål i seg selv, men er viktig på bakgrunn av svak reliabilitet svekker begrepsvaliditeten. Under selve datainnsamlingen kan det oppstå feilkilder som kan redusere begrepsvaliditeten siden det reduserer samsvaret mellom begrepet slik det er definert og slik det er operasjonalisert (Kleven, 2002b). Begreper som skal operasjonaliseres i denne studien er aritmetiske ferdigheter i form av regneflyt og aritmetiske problemløsningsferdigheter.

Truslene mot begrepsvaliditet kan deles opp i to hovedgrupper (Kleven, 2002b). Den første er tilfeldige målefeil og den sier ikke at feilene nødvendigvis skyldes tilfeldigheter, men at feilene oppfører seg tilfeldig. De tilfeldige målefeilene kan redusere en eventuell korrelasjon mellom variablene, og lav reliabilitet kan gjøre det vanskelig å empirisk finne sammenhengen som kan være til stede mellom de teoretiske begrepene. Den andre målefeilen er systematiske, som handler om at tendensen av målingene går i samme retning for den samme personen ved gjentatte målinger. De systematiske feilene er feil som fører til at vi får et mer eller mindre skjevt bilde av begrepet som måles. Dette kan enten være fordi indikatorene avdekker enkelte deler av begrepet som måles eller fordi irrelevante forhold som ikke hører til begrepet dukker opp. I praksis vil det kunne være vanskelig å vurdere om en feilkilde er systematisk eller tilfeldig (Kleven, 2002b).

3.5.4 Ytre validitet

Ytre validitet handler om hvor sikkert vi kan generalisere resultatene til og over de relevante individene, situasjonene og tidene som problemstillingen tar for seg (Lund, 2002b). Utvalget vil som regel ikke bestå av alle individene, situasjonene og tidene, men et utvalg. Det er derfor viktig å være tydelig på valgene som er tatt slik at resultatene kan generaliseres på en mest mulig pålitelig måte (Lund, 2002b). Ved god ytre validitet vil en kunne generalisere resultatene (Lund, 2002a) og ved svak ytre validitet vil en ikke kunne generalisere med rimelig sikkerhet. En studie vil på den bakgrunn bli tilsvarende dårlig belyst og helheten bli svekket (Lund, 2002b). Faktorer som kan påvirke den ytre validiteten negativt er interaksjon mellom uavhengig variabler og individer, situasjoner og tider. Hvis effekten varierer mellom disse, vil overgeneraliseringen bli usikker. En annen faktor er hvis individene i utvalget er relativ ensartet og hvis det er en skjevhet i utvalget i forhold til populasjonen (Lund, 2002a). En forutsetning for god ytre validitet er ikke bare basert på individer, situasjoner og tider er valgt ut, men også ut i fra tidligere kunnskap fra undersøkelser og studier. Det er den samlende kunnskapen fra resultatene og informasjonen i og utenfor oppgave som er avgjørende for generalisering (Lund, 2002b).

3.5.5 Reliabilitet

Reliabilitet sier noe om hvor presise målingene og registreringene er, og det refererer til hvilken grad det finnes målefeil i resultatene fra testene (Befring, 2015; Gall et al., 2007). Hvis målingene er konsistente vil den samme personen få tilnærmet samme resultat når en måling blir gjentatt (Kleven, 2002b). Måleinstrumentene som er valgt i denne studien fra testbatteriet til VLC-prosjektet er valgt med bakgrunn i empiri for å kunne besvare problemstillingen. Måleinstrumentene som er brukt i denne oppgaver er alle standardiserte tester, og det vil si at de har blitt gjennomført forskning som viser at de måler det de er ment å måle. En viktig reliabilitetskontroll er intern konsistens, som er en tilnærming for å estimere reliabiliteten på testskårer og involverer undersøkelse av hver enkelt items i en test (Gall et al., 2007). En metode som er mye brukt for å måle dette er Cronbachs alpha. Cronbachs alpha gir uttrykk for den gjennomsnittlige korrelasjonen når en test blir delt og seg imellom korrelert på alle mulige måter (Befring, 2015). Reliabilitetskoeffisientene for måleinstrumentene i denne studien vises i tabell 3.

Tabell 3. Testreliabilitet (Cronbachs alpha)

Måleinstrumenter	Cronbachs alpha
Regnefaktaprøven addisjon	.930
Regnefaktaprøven subtraksjon	.913
WISC-IV regning	.757
Raven	.851

Hver utregning involverer en reliabilitetskoeffisient. Denne koeffisienten varierer mellom verdier på .00 og 1.00. Generelt vil en test som resulterer i en reliabilitetsskår på .80 eller høyere være tilstrekkelig for de fleste forskningsformålene (Gall et al, 2007), men en skår på 0.7 er også akseptabel (Pallant, 2016). Alle måleinstrumentene viser til en reliabilitetskår høyere enn .80 som viser til tilstrekkelig og god reliabilitet, bortsett fra WISC-IV regning som viser en reliabilitetsskår på .757. Denne skåren vil enda gi en akseptabel reliabilitet. Ved høy intern konsistens vil en kunne med rimelig sikkerhet si at måleinstrumentene måler det de er ment å måle.

3.6 Analyse

Her vil de statistiske analysene som vil bli brukt for å analysere datamaterialet presenteres. I denne studien vil det gjennomføres deskriptiv analyse, bivariat korrelasjonsanalyse og hierarkisk multippel regresjonsanalyse. Alle analysene av data blir gjennomført i IBM SPSS statistics 25.

3.6.1 Deskriptiv analyse

Deskriptiv analyse vil bli brukt for å vurdere de ulike variablene (måleinstrumentene) og for å beskrive distribusjonen av skårer vil gjennomsnitt, standardavvik, variasjonsbredde, skjevhet og kurtosis bli analysert. Gjennomsnitt, standardavvik og variasjonsbredde vil bli brukt for å beskrive karakteristikken til skårene og for å se på formen til distribusjonen vil skjevhet og kurtosis bli brukt. Skjevhetverdien gir en indikasjon på symmetrien på distribusjonen og kurtosisverdien gir informasjon om spissheten til distribusjonen (Field, 2018; Pallant, 2016). Hvis distribusjonen er normalfordelt vil skjevhet- og kurtosisvedien være null. En distribusjon som er skjev vil være asymmetrisk og de mest frekvente skårene er samlet i en av endene av skalaen. Hvis det er uteliggere (outliers) i dataen er det tendenser til at distribusjonen vil bli skjevfordelt (Field, 2018). Skjevhetsverdier ± 2 kan føre til at slutningsstatistikken blir upålitelig (Christophersen, 2012). Kurtosis forteller i hvilken grad skårene er klynget sammen i endene av skalaen (kjent som halen) (Field, 2018). *The central limit theorem* vil i store utvalg (f.eks. N = 332) føre til at distribusjonen av data fra utvalget vil være normalfordelt, uansett hvilken form dataen hadde når den ble samlet inn. Ettersom et utvalg blir større kan en være mer sikre på at distribusjonen av dataen fra utvalget vil være normalfordelt (Field, 2018).

3.6.2 Bivariat korrelasjonsanalyse

En korrelasjonsanalyse blir brukt for å beskrive styrken i et forhold mellom to variabler (Gall et al., 2007). Det er ulike typer korrelasjonsanalyser tilgjengelig og disse er avhengig av målenivå og karakteren til datasettet (Pallant, 2016). Pearsons korrelasjons koeffisient (*r*) er en type korrelasjonsanalyse og er laget for variabler på intervallnivå. Den beskriver forholdet mellom to variabler og retningen på forholdet, men kan ikke si noe om årsaksforholdet mellom variablene (Gall et al., 2007). På bakgrunn av dette vil den bli brukt som korrelasjonsanalyse i denne oppgaver. Pearson *r* tar for seg verdier mellom -1 til +1, hvor tegnene foran indikerer om det er en positiv eller negativ korrelasjon. En korrelasjon på null indikerer ingen korrelasjon, og ± 1 indikerer er perfekt korrelasjon (Pallant, 2016). For å tolke korrelasjonene vil retningslinjene til Cohen (1988, s. 79-80) bli brukt: liten korrelasjon (*r*= .10 til .29), medium korrelasjon (*r*=.30 til .49) og stor korrelasjon (*r*= .50 til 1.00). Forutsetninger for bivariate korrelasjoner er linearitet og normalitet. Disse forutsetningene blir sjekket og rapportert før korrelasjonsanalysene bli gjennomført.

3.6.3 Regresjonsanalyse

I en hierarkisk multippel regresjonsanalyse er formålet å identifisere potensielle prediktor variablene (uavhengige variabler) som kan gi en mest mulig presis prediksjon av den avhengige variabelen (utfallsvariabel) (Befring, 2015). Ved hierarkisk vil det si at en introduserer de uavhengige variablene i en spesifisert rekkefølge. Først vil det beregnes en fellesvarians med den første variabelen, videre vil den andre uavhengige variabelen bli tatt inn i analysen og deretter den tredje (Befring, 2015). I analysen vil alder i måneder være uavhengig variabel og aritmetiske ferdigheter (WISC-IV regning, Regnefaktaprøven addisjon og subtraksjon) vil være avhengig variabel. Målnivået i en regresjonsanalyse bør være intervall eller ratio, et unntak er hvis en har en dichotomous uavhengige variabel (med to verdier, f.eks. kjønn) (Pallant, 2016). Kontrollvariablene i denne studien er nonverbale evner og kjønn. På bakgrunn av at det er tre mål på aritmetiske ferdigheter vil det bli gjennomført tre analyser for hver avhengig variabel. Rekkefølgen på introduseringen av de uavhengige variablene er først kontrollvariablene nonverbale evner og kjønn, og deretter alder. Effekten av hver enkelt uavhengig variabel på den avhengige variablene vil her bli suksessivt kontrollert for og for alle de uavhengige variablene som inngår i analysen, og det vil dermed kunne avdekke eventuelle falske relasjoner (Befring, 2015).

For å gjennomføre en multippel regresjonsanalyser er det visse forutsetninger som må oppnås og disse vil bli sjekket i resultatkapittelet og rapportert før analysen blir gjennomført; *Linearitet*, som handler om at utfallsvariabelen burde ha et lineært forhold til hver prediktor variabel. *Selvstendighet*, som menes med at error i modellen ikke er forbundet med hverandre. *Homoscedasticity*, som betyr at residual på hvert nivå av prediktor/prediktorene burde ha den samme variansen. *Normalitet*, som menes med at forskjellen mellom den forventede og observerte verdien ligger på null eller rundt null (Field, 2018).

3.7 Etiske hensyn

Denne studien er en del av et større prosjekt, VLC, og derfor vil mange av de etiske hensynene være knyttet mot de føringene for etiske hensyn som ligger i VLC-prosjektet. Innmelding av prosjektet til Norsk senter for forskningsdata (NSD) og innhenting av samtykke fra informanter ble gjennomført av VLC-prosjektet. Opplysningene som fremkommer i prosjektet blir behandlet konfidensielt og blir anonymisert, samt lagret på Tjenester for Sensitiv Data. Dette står i tråd med den nasjonale forskningsetiske komite for samfunnsvitenskap og humaniora (2016) at opplysninger som kan identifisere enkeltpersoner skal lagres forsvarlig og informasjonen skal behandles konfidensielt og fortrolig.

Barn er en gruppe som er sårbar og konfidensialitet gjelder spesielt når barn deltar i forskning (NESH, 2016). All deltakelse i prosjektet er frivillig og på bakgrunn av at informantene i VLC-prosjektet er under 15 år ble det innhentet samtykke fra foresatte (NESH, 2016). I VLC-prosjektet ble informasjonsskriv, samtykkeskjema og spørreskjema oversatt til de seks vanligste minoritetsspråkene på bakgrunn av at informasjonen skal være tilpasset deltakernes kulturelle bakgrunn og det skal formidles på et språk de forstår (NESH, 2016). NESH (2016) trekker samtidig frem at det er viktig å behandle mindreårige som selvstendig individer og barn som er fylt syv år er i stand til å danne sine egne meninger om en sak, bli informert og bør få mulighet til å kunne si sin egen mening. Derfor er det nødvendig å innhente samtykke fra barna, i tillegg til foresattes formelle samtykke (NESH, 2016). For alle elevene som deltok ble det gitt alderstilpasset informasjon om prosjektet muntlig, og det ble understreket av deltakelsen var frivillig og det er mulig å trekke seg når som helst i prosessen. Det ble innhentet samtykke fra elevene ved hver testsituasjon. I selve testsituasjonen er det viktig at forskere har tilstrekkelig kunnskap om barn slik at metoden og innholdet i forskningen blir tilpasset aldersgruppen som deltar (NESH, 2016). I tråd med dette er at forskningsassistentene som deltok i kartleggingen er masterstudenter ved Instituttet for Spesialpedagogikk og er ansett å ha god kjennskap til barns utvikling, Testbatteriet er tilpasset andreklassinger i den grad det er mulig, både i oppgavetype og omfang, samt at deler av testbatteriet er standardisert for denne aldersgruppen. Rekkefølgen på testene ble også bevisst lagt opp slik at barns motivasjon skulle holdes oppe, og at de skulle oppleve mestringsfølelse og få en god opplevelse.

4 Resultater

I dette kapittelet vil studiens resultater bli beskrevet og analysert. Først presenteres deskriptive analyser av variablene i en tabell og deretter vil det bli gjort en individuell vurdering av Regnefaktaprøven - addisjon og subtraksjon, WISC-IV regning, samt kontrollvariabelen Raven. Videre presenteres resultatet og vurderingen av bivariate korrelasjoner. Deretter presenteres vurderingen av forutsetningene og resultatene på hierarkisk multippel regresjonsanalyse. Avslutningsvis oppsummeres analysene og funnene.

4.1 Deskriptive analyser

Beskrivelsen av distribusjonen på de ulike måleinstrumentene er presentert i tabell 4. Her presenteres måleinstrumentenes utvalgsstørrelse, gjennomsnitt, standardavvik, variasjonsbredde, skjevhet og kurtosis.

Tabell 4. Utvalgsstørrelse (N), gjennomsnitt (M), standardavvik (SD), variasjonsbredde (VB), skjevhet (Skew) og kurtosis (Krt) for målte variabler

Måleinstrument	N	M	SD	VB	Skew	Krt
Regnefaktaprøven addisjon	332	11.62	6.66	0-43	1.32	3.04
Regnefaktaprøven subtraksjon	332	7.66	5.77	0-30	1.09	1.89
WISC-IV regning	332	17.20	3.50	6-31	- .14	.47
Raven	332	26.89	4.97	9-34	-1.19	1.31

4.1.1 Regnefaktaprøven addisjon

Variabelen Regnefaktaprøven addisjon har et gjennomsnitt på 11.62 og standardavvik på 6.66. Skjevhetverdien ligger på 1.32 og kurtosisverdien på 3.04 som viser at fordelingen heller mot venstre side og hvor en stor del av skårene ligger i halen. Q-Q plotet (se vedlegg 2) viser at noen av punktene ligger i en bue under linjen. Dette er med på å indikerer at fordelingen er skjev og avviker fra normalfordelingen (Christophersen, 2012). Samtidig vil «the central limit theorem» føre til at distribusjonen av data fra utvalget på 332 andre klassinger vil være normalfordelt. Fra skjevhets- og kurtosisverdiene kan det tyde på at flere av andreklassingene ikke har oppnådd optimal regneflyt på addisjonsoppgaver (51.5 % av andreklassingene fikk under 10 oppgaver riktig), og at de støtter seg på backupstrategier enn retrievalstrategier som er mer effektive. Dette er også noe som gjenspeiler seg i litteraturen at barn støtter seg mer på backupstrategier i første- og andreklasse (Geary, 1994; Jordan et al., 2003; Siegler & Braithwaite, 2017; Siegler, 1987). Variasjonsbredden har en minimumsverdi på 0 og en maksimumsverdi på 43. Prøven består i av alt 45 oppgaver og viser at det er noen elever som har oppnådd optimal regneflyt i andreklasse (1.5 % fikk over 30 oppgaver riktig). Variabelen møtte kriteriene for normalitet og linearitet. Variabelen vil bli brukt i videre i analysene. Histogram ligger under vedlegg 1.

4.1.2 Regnefaktaprøven subtraksjon

Variabelen Regnefaktaprøven subtraksjon har et gjennomsnitt på 7.66 og et standardavvik på 5.77. Skjevhetsverdien ligger på 1.09 og kurtosisverdien er 1.89, og det vil si at fordelingen lener seg mot venstre side. Q-Q plotet (se vedlegg 4) viser at noen av punktene ligger i en bue under linjen, og er med på å indikerer at fordelingen er skjev. Skjevhetsverdien på 1.09 blir vurdert som akseptabel da den ligger innenfor verdiene ± 2. Dette kan tyde på at det er en ferdighet de fleste andreklassingene enda ikke mestrer (76.2 % av elevene fikk under 10 oppgaver riktig). En faktor som kan påvirke skjevhets- og kurtosisverdien er en feilkilde under selve testingen. Under testingen var det ikke alle elevene som fikk med seg at de skulle bruke regneoperasjonen subtraksjon og ikke addisjon, dette førte til at flere elever fikk skåren 0. Blant utvalget på 332 var det 10.5 % som fikk skåren null. Denne begrensningen vil bli drøftet videre i drøftingskapittelet. Variasjonsbredden har en minimumsverdi på 0 og en maksimumsverdi på 30 av totalt 45 mulige. Dette viser også at det er noen elever som begynner å utvikle optimal regneflyt i subtraksjon. Variabelen vil bli brukt videre i analysene. Histogram ligger under vedlegg 3.

4.1.3 WISC-IV regning

For variabelen WISC-IV regning er gjennomsnittet 17.20 og standardavviket ligger på 3.50. Skjevhetverdien er på - .14 og kurtosisverdien ligger på .47, og viser at distribusjonen av skårer er tilnærmet normalfordelt. Variasjonsbredden har en minimumsskår på 6 og en maksimumsskår på 31, som tyder på at det er en ferdighet som andreklassingene mestrer i ulik grad. 3.3 % av andreklassingene fikk under 10 oppgaver riktig og 15.7 % fikk 20 oppgaver eller mer riktig av totalt 34. På bakgrunn av at distribusjonen av skårer er tilnærmet normal fordelt vil testen bli brukt videre i analysene. Histogram ligger under vedlegg 5 og Q-Q plot under vedlegg 6.

4.1.4 Raven

Kontrollvariabelen Raven har et gjennomsnitt på 26.89 og et standardavvik på 4.97. Skjevheten ligger på – 1.19 og kurtosis på 1.31 og viser at distribusjonen er negativ skjev og verdiene er samlet i blant de høyere skårene og halen er mot de lavere skårene, men skjevhetsverdien ligger innenfor ± 2. Variasjonsbredden er 9-34 av totalt 36 oppgaver som tyder på at det er en test som elevene mestrer i ulik grad, men hvor de fleste elevene ligger blant de høyere skårene. Histogram ligger under vedlegg 7 og Q-Q plot ligger under vedlegg 8.

4.2 Bivariat korrelasjoner

Korrelasjonen mellom alle variablene er presentert i tabell 5. Tabellen viser styrken til korrelasjonene gjennom Pearsons korrelasjonskoeffisient. Innledende analyser ble gjennomført for å sikre at det ikke var noen brudd på forutsetninger angående normalitet og linearitet.

Tabell 5. Oversikt over korrelasjonene mellom målene.

Variabel	1	2	3	4	5
1.Alder i måned	-				
2.Raven	.168**	-			
3. Regnefaktaprøven addisjon	.091	.239**	-		
4.Regnefaktaprøven subtraksjon	.095	.186**	.683**	-	
5.WISC-IV regning	.163**	.425**	.488**	.473**	-

** Korrelasjonen er signifikant på .01 nivå (to-hale test).

Kontrollvariabelen Raven korrelerer svakt eller moderat med variablene som måler aritmetiske ferdigheter og alder i måned. Det er en svak sammenheng mellom alder og Raven, *r* = .168, *p* = .002. Raven har også en svak sammenheng med Regnefaktaprøven addisjon, *r* = .239, *p* = .000, og Regnefaktaprøven subtraksjon, *r* = .186, *p* =. 001. Raven korrelerer moderat med WISC- IV regning, *r* = .425, *p* = . 000. Det vil si at det er en positiv sammenheng mellom nonverbale evner og aritmetiske ferdigheter.

Tabellen viser at det er to av variablene som viser sterk positiv korrelasjon, det er mellom Regnefaktaprøven addisjon og Regnefaktaprøven subtraksjon, *r* = .683, *p* = .000. En vil dermed kunne si med høy sikkerhet at de andreklassingene som skårer høyt på Regnefaktaprøven addisjon vil skåre høyt på Regnefaktaprøven subtraksjon. Lavere skåre på addisjonsregneflyt, vil være forbundet med lave skårer på subtraksjonsregneflyt. Regnefaktaprøven addisjon korrelerer positivt og moderat med WISC- IV regning, *r* = .488, *p* = .000, samt at Regnefaktaprøven subtraksjon også korrelerer positivt og moderat med WISC-IV regning, *r* = .479, *p* = .000. Det vil si at høye skårer på Regnefaktaprøven - addisjon og subtraksjon er assosiert med høye skåre på WISC-IV regning. Svak regneflyt vil være forbundet med svake aritmetiske problemløsningsferdigheter.

Alder i måned viser ingen sammenheng til addisjonsregneflyt, *r* = .091, *p* =.097, det samme gjelder for subtraksjonsregneflyt, *r* = .095, *p* =.084. Alder i måned viser en svak positiv sammenheng til aritmetiske problemløsningsferdigheter, *r* = .163, *p* = .003. Kvadratet av korrelasjons koeffisienten .163 er .027, og alder vil dermed forklare 2.7 % av variansen i WISC-IV regning. Pearsons *r* viser at elever som er født tidlig på året ikke skårer bedre på regneflyt enn sine medelever som er født sent på året innenfor samme alderskohort når de går i andreklasse. Ved aritmetiske problemløsingsferdigheter er det en svak sammenheng at elever som er født tidlig på året skårer bedre enn sine medelever som er født sent på året.

Korrelasjonene viser at alle variablene korrelerer bortsett fra alder og regneflyt, og at korrelasjonene er signifikante på .01 nivå. Dette indikerer at en kan være 99 % sikker på de resultatene som er oppnådd (Pallant, 2016). Ut fra disse analysene vil en bare kunne si noe om sammenhengen mellom variablene og retningen, men en vil ikke kunne si noe om årsaksforholdet.

Siden alder i måned viser ingen eller liten sammenheng med aritmetiske ferdigheter, og nonverbal evner korrelerer med svakt til moderat i aritmetiske ferdigheter, vil det dermed videre bli sett på om alder i måned kan forklare unik variasjon i aritmetiske ferdigheter når det kontrolleres for nonverbal evner og kjønn. For å se hvilke av variablene som kan inkluderes i hierarkisk multippel regresjonsanalyse vil det bli gjort innledende analyser på forutsetninger for å kunne gjennomføre en regresjonsanalyse.

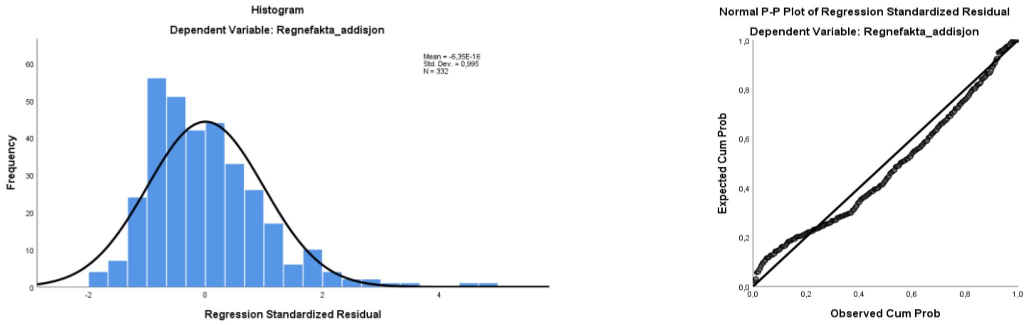
4.3 Kriterier for regresjonsanalyse

Før gjennomføringen av multippel regresjonsanalysen ble det sjekket for forutsetningene som må være oppnådd for å kunne gjennomføre en slik analyse. Disse forutsetningene må være oppnådd hvis en skal kunne generalisere funnene utenfor denne studien (Field, 2018). I denne studien har følgende forutsetninger undersøkt; normalitet, linearitet og uavhengige feil (Durbin-Watson). Det er også gjort vurderinger på om det er noen tilfeller i dataen som kan påvirke regresjonsmodellen (Mahalanobis distance og Cook’s distance).

For å vurdere om fordelingen er normalfordelt vil de fleste verdiene være fordelt rundt null, og forholdet vil være lineært hvis verdiene ligger i en diagonal linje fra venstrebunn til høyretopp. For *Durbin-Watson* testen er akseptable verdier beregnet å ligge mellom 1-3, hvor 2 indikerer at residualene ikke korrelerer (Field, 2018). Akseptable verdier for *Mahalanobis distance* ved tre uavhengige variabler ligger under 16.26 (Tabachnick & Fidell, 2013), og for *Cook’s distance* er det under 1. Hvis det er tilfeller over disse verdiene vil det potensielt skape problemer for regresjonsmodellen (Field, 2018).

4.3.1 Antakelse for regresjon med Regnefaktaprøven addisjon

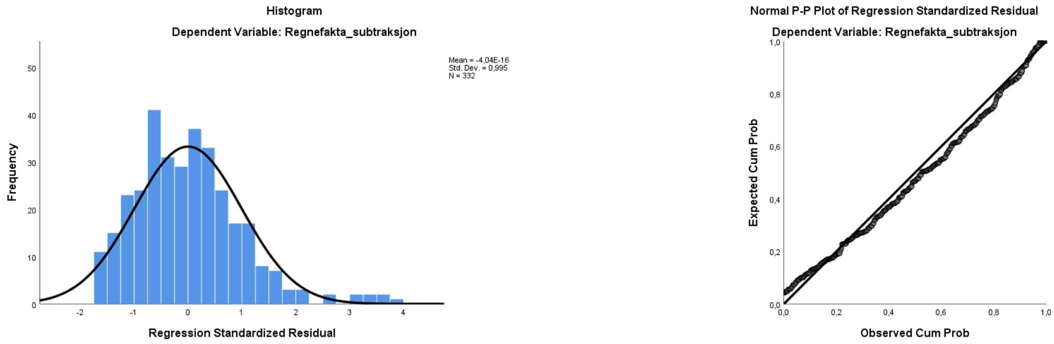
Normalfordeling til Regnefaktaprøven addisjon er svak venstreskjev, hvor det er flere lave verdier og den er lineær (figur 3). Durbin-Watson er på 1.825, og er innenfor grenseverdien. Mahalanobis distance er på 15.95 og Cook’s distance er på .082. Den blir enda vurdert som akseptabel og inkluderes i regresjonsanalysen.



Figur 3. Histogram og P-P plot over regresjon med Regnefaktaprøven addisjon som avhengig variabel

4.3.2 Antakelse for regresjon med Regnefaktaprøven subtraksjon

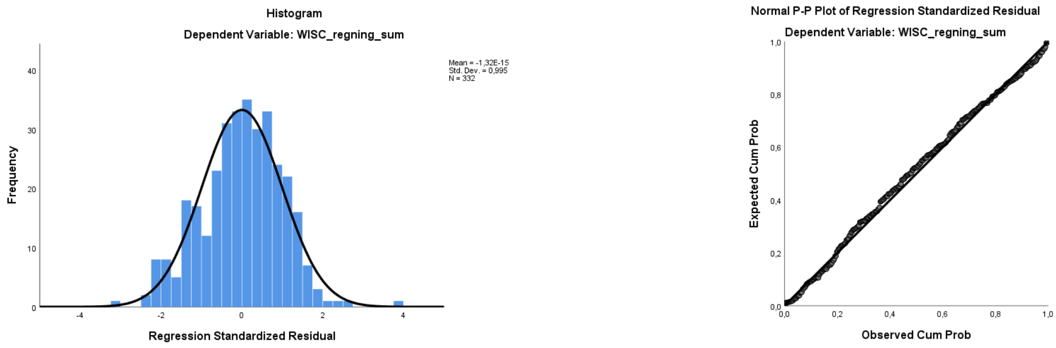
Normalfordelingen til Regnefaktaprøven subtraksjon er svak venstreskjev og den er lineær (figur 4). Durbin-Watson er på 1.678, og er innenfor grenseverdien. Mahalanobis distance er på 15.95 og Cook’s distance er på .050. Variabelen er vurdert som akseptabel og inkluderes i regresjonsanalysen.



Figur 4. Histogram og P-P plot over regresjon med Regnefaktaprøven addisjon som avhengig variabel

4.3.3 Antakelse for regresjon med WISC-IV regning

Variabelen er normalfordelt og er lineær (figur 5). Durbin-Watson er på 1.705, og er innenfor grenseverdien. Mahalanobis distance er på 15.95 og Cook’s distance er på .041. Variabelen er vurdert som akseptabel og inkluderes i regresjonsanalysen.



Figur 5. Histogram og P-P plot over regresjon med WISC-IV regning som avhengig variabel

4.4 Hierarkisk multippel regresjonsanalyse

De avhengige variablene som inkluderes i denne studien er Regnefaktaprøven - addisjon og subtraksjon som mål på regneflyt og WISC-IV regning som mål på aritmetiske problemløsningsferdigheter. Variablene er inkludert på bakgrunn av at de oppfyller forutsetningene for å gjennomføre multippel regresjonsanalyse. De oppnår akseptabel normalitet og linearitet, og de har ingen tilfeller som vil påvirke regresjonsmodellen. I disse analysene undersøkes det om sammenhengen mellom alder og aritmetiske ferdigheter kan være påvirket av andre faktorer. Nonverbale evner (Raven) og kjønn inkluderes i regresjons-analysen som kontrollvariabler. Resultatene fra av de tre ulike regresjonsanalysene er presentert i tabellene 6. Regnefaktaprøven addisjon, 7. Regnefaktaprøven subtraksjon og 8. WISC-IV regning.

Tabell 6. Oversikt over resultatene i regresjonsanalysen for Regnefaktaprøven addisjon

N=332	B	SE(B)	β	t	p	R^2	R^2 Change
Steg 1						.088	
Raven	.328	.071	.245	4.641	.000		
Kjønn	2.331	.703	.175	3.317	.001		
Steg 2						.092	.005

Raven	.313	.072	.233	4.368	.000		
Kjønn	2.410	.705	.181	3.420	.001		
Alder i måned	.134	.105	.069	1.280	.202		

4.4.1 Regnefaktaprøven addisjon

Regresjonsanalysen viser at modell 2 (R^2 = .092) forklarer mer av variasjonen enn modell 1 (R^2 = .088), men denne endringen er ikke statistisk signifikant (*p* = .202). Modell 2 viser at nonverbale evner, kjønn og alder tilsammen forklarer 9,2 % av variasjonen i addisjonsregneflyt, hvor alder forklarer ytterlige 0.5 % av den unike variasjonen. Modellen totalt er statistisk signifikant (*p* = .000). Dersom alder øker med ett standardavvik øker elevenes addisjonsregneflyt med .069 standardavvik. Den variabelen som utgjør det største unike bidraget for å forklare den avhengige variabelen, Regnefaktaprøven addisjon er nonverbale evner (*β* **=** .233, *p* = .000), når de overlappende effektene av de andre variablene statistisk er fjernet. Deretter er det kjønn (*β* = .181, *p* = .001) som forklarer mest og hvor guttene presterer bedre enn jentene. Alder som er den variabelen som forklarer minst og er ikke signifikant (*β* = .069, *p* = .202). Når det ble kontrollert for nonverbale evner og kjønn forklarte alder ingen unik variasjon i addisjonsregneflyt.

Tabell 7. Oversikt over resultatene i regresjonsanalysen for Regnefaktaprøven subtraksjon

N=332	B	SE(B)	β	t	p	R^2	R^2 Change
Steg 1						.044	
Raven	.220	.063	.189	3.511	.001		
Kjønn	1.132	.623	.098	1.818	.070		
Steg 2						.050	.005

Raven	.205	.063	.177	3.204	.001		
Kjønn	1.207	.624	.105	1.934	.054		
Alder i måned	.127	.093	.075	1.365	.173		

4.4.2 Regnefaktaprøven subtraksjon

Regresjonsanalysen viser at modell 2 (R^2 = .050) forklarer mer av variasjonen i Regnefaktaprøven subtraksjon enn modell 1 (R^2 = .044), men denne endringen er ikke statistisk signifikant (p = .173). Modell 2 viser at nonverbale evner, kjønn og alder tilsammen forklarer 5 % av variasjonen i subtraksjonsregneflyt, hvor alder forklarer ytterlige 0.5 % av den unike variasjonen, og modellen totalt er statistisk signifikant (p = .001). Dersom alder øker med ett standardavvik øker elevenes subtraksjonsregneflyt med .173 standardavvik. Nonverbale evner (β = .177, p = .001) er den faktorer som forklarer mest av variasjonen i subtraksjon etterfulgt av kjønn (β = .105, p = .054). Alder (β = .075, p = .173) er den faktoren som forklarer minst av den unike variasjonen i Regnefaktaprøven subtraksjon. Når det ble kontrollert for nonverbale evner og kjønn forklarte alder ingen unik variasjon i subtraksjonsregneflyt.

Tabell 8. Oversikt over resultatene i regresjonsanalysen for WISC-IV regning

N=332	B	SE(B)	β	t	p	R^2	R^2 Change
Steg 1						.187	
Raven	.031	.035	.428	8.598	.000		
Kjønn	.534	.349	.076	1.533	.126		
Steg 2						.197	.010*
Raven	.289	.035	.411	8.138	.000		
Kjønn	.596	.348	.085	1.712	.088		

Alder i måned	.104	.052	.102	2.016	.045

Note. *p =<.05

4.4.3 WISC-IV regning

Regresjonsanalysen viser at modell 2 (R^2 = .197) forklarer mer av variasjonen enn modell 1 (R^2 = .187), og denne endringen er statistisk signifikant (*p* = .045). Modell 2 viser at nonverbale evner, kjønn og alder til sammen forklarer 19.7 % av variasjonen i WISC-IV regning, og alder forklarer ytterlige 1 % av den unike variasjonen. Regresjonsmodellen totalt er statistisk signifikant (*p* = .000). Dersom alder øker med ett standardavvik øker elevenes aritmetiske problemløsningsferdigheter med .102 standardavvik. Nonverbale evner (β = .411, *p* = .000) er den faktorer som forklarer mest av variasjonen i aritmetiske problemløsningsferdigheter etterfulgt av alder (β = .102, *p* = .045). Kjønn (β = .085, *p* = .088) er den faktoren som forklarer minst av den unike variasjonen og er ikke statistisk signifikant. Når det ble kontrollert for nonverbale evner og kjønn bidrar alder til unik variasjon i aritmetiske problemløsningsferdigheter.

4.5 Oppsummering av analyser og resultater

I det foregående kapittelet ble deskriptiv data for Regnefaktaprøven - addisjon og subtraksjon, WISC-IV regning og Raven presentert. Variablene Raven og WISC-IV regning viste til en normalfordeling, mens variablene Regnefaktaprøven - addisjon og subtraksjon viste til en fordeling som var svak venstreskjev. Regnefaktaprøven addisjon hadde en høy kurtosisverdi, men alle variablene ble vurdert som akseptable til å bli brukt i de videre analysene.

De bivariate korrelasjonene viste at nonverbale evner korrelerer signifikant med de andre avhengige variablene Regnefaktaprøven - addisjon og subtraksjon og WISC-IV regning, samt alder. Addisjonsregneflyt og subtraksjonsregneflyt korrelerte sterkt med hverandre, og de korrelerte moderat med aritmetiske problemløsningsferdigheter. Alder viste en svak signifikant sammenheng med WISC-IV regning, men viste ingen sammenheng med Regnefaktaprøven - addisjon og subtraksjon.

Forskningsspørsmålet som gjelder i hvilken grad alder kan forklare unik variasjon i aritmetiske ferdigheter når det kontrolleres for nonverbale evner og kjønn, viste regresjonsanalysen varierende resultater. Regresjonsanalysen hvor Regnefaktaprøven addisjon var avhengig variabel forklarte alder ingen unik variasjon og var ikke statistisk signifikant når det ble kontrollert for nonverbale evner og kjønn. Nonverbale evner og kjønn derimot forklarte unik variasjon og var statistisk signifikant. Regresjonsanalysen av Regnefaktaprøven subtraksjon forklarte alder ingen unik variasjon og var ikke signifikant når det ble kontrollert for nonverbale evner og kjønn. Nonverbale evner viste unik variasjon i subtraksjonsregneflyt og var signifikant, mens kjønn forklarte ingen unik variasjon og var ikke signifikant. Når det ble kontrollert for nonverbale evner og kjønn i WISC-IV regning forklarte alder unik variasjon. Den variabelen som bidro høyest til unik variasjon og som var signifikant var nonverbale evner. Kjønn bidro ikke til unik variasjon i aritmetiske problemløsningsferdigheter og var ikke statistisk signifikant. I Regnefaktaprøven addisjon presterte guttene signifikant bedre enn jentene. Samme tendens kan også ses i Regnefaktaprøven subtraksjon og WISC-IV regning, men her var resultatet ikke signifikant.

5 Drøfting

Problemstillingen for studien er: *Er det sammenheng mellom alder og aritmetiske ferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?* Studien har som hensikt å svare på forskningsspørsmålene:

I. I hvilken grad kan alder forklare unik variasjon i regneflyt hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?

II. I hvilken grad kan alder forklare unik variasjon i aritmetiske problemløsningsferdigheter hos norske enspråklige andreklassinger når det kontrolleres for nonverbale evner og kjønn?

I dette kapittelet drøftes resultatene fra analysene i lys av tidligere gjennomgått teori og empiri om utvikling av aritmetiske ferdigheter og studier på relativ alderseffekt i akademiske fag. Deretter vil studien, samt resultatene drøftes i lys av validitetssystemet til Cook og Campell. Avslutningsvis presenteres hvilke implikasjoner resultatene kan ha for praksis, og fremtidige studier på relativ alderseffekt.

5.1 Resultatene sett i lys av teori og empiri

I denne delen vil resultatene bli sett i lys av teori og empiri som ble presentert i kapittel 2. Mer spesifikt vil det si at resultatene fra Regnefaktaprøven - addisjon og subtraksjon vil bli sett i lys teori om strategiutvikling og resultatene fra WISC-IV regning vil bli sett i lys av teori om aritmetisk problemløsningsoppgaver, samt i lys av tidligere teori og empiri på relativ alderseffekt og akademiske fag.

5.1.1 Alderseffekt i regneflyt

Resultatene fra regresjonsanalysen med Regnefaktaprøven - addisjon og subtraksjon som avhengig variabel viste at alder ikke forklarte unik variasjon i regneflyt, både i addisjonsregneflyt (β = .069, p = .202) og subtraksjonsregneflyt (β = .075, p = .173) når det ble kontrollert for nonverbale evner og kjønn. En mulig grunn til at alder ikke forklarer unik variasjon er at barn blir eksponert for addisjons- og subtraksjonsoppgaver i en tidlig alder. Fra elevene går i barnehagen (to- til fireårsalderen) er de i stand til å utføre addisjons- og subtraksjonsoppgaver, og det er før de har fått formell undervisning i matematikk (Gilmore et al., 2018). Andreklassingene i denne studien var i en periode i utviklingen hvor de stadig utvikler mer effektive strategier som for eksempel telle fra det største tallet og dekomposisjon, samt retrieval (Geary, 1994; Jordan et al., 2003; Siegler & Braithwaite, 2017; Siegler, 1987). Resultatene til Regnefaktaprøven - addisjon og subtraksjon kan tyde på at alle elevene enda ikke har oppnådd optimal regneflyt, og at de støtter seg mer på backupstrategier enn retrievalstrategier. Dette kan være en annen grunn til at alder ikke forklarer unik variasjon i regneflyt.

Gjennomsnittet på antall riktige addisjonsoppgaver var 11.62 av totalt 45 oppgaver, og variasjonsbredden er 0-43. Dette kan være med å støtte opp om at elevene mestrer denne ferdigheten i ulik grad, og at det er flere av elevene som støtter seg på backupstrategiene. Det var 51.5 % av elevene som fikk under 10 oppgaver riktig. Det var 1.5 % av andreklassingene som fikk over 30 oppgaver riktig, og viser at noen elever har oppnådd optimal addisjonsregneflyt. Dette viser til at det er en stor individuell forskjell i prestasjoner, men som ikke kan forklares av alder i måneder. Samtidig viser resultatene fra standardiseringen og normeringen av Regnefaktaprøven som ble gjennomført fra november 2012 til januar 2013 at gjennomsnittet for antall riktige oppgaver på delprøven addisjon 0-20 for 2. trinn er 9.5 (*SD* = 6.2) (Klausen & Reikerås, 2016). Dette viser at andreklassingene i denne studien ligger rundt de forventede gjennomsnittsverdiene for antall riktige oppgaver i 2 trinn. I subtraksjonsoppgavene var gjennomsnittet for antall riktige oppgaver 7.66 og 76.2 % av elevene fikk under 10 oppgaver riktige. Av de 76.2 % var det 10.5 % som fikk null oppgaver riktig og variasjonsbredden er 0-30. Gjennomsnittet som er forventet for 2. trinn i delprøven subtraksjon 0-20 er 7.4 (SD = 4.8) (Klausen & Reikerås, 2016), og gjennomsnittet for andreklassingene i denne studien viser at de ligger på de forventede gjennomsnittsverdiene for antall riktige oppgaver i 2. trinn.

Tidligere studier på relativ alderseffekt i matematikk fant signifikante forskjeller mellom de eldste og de yngste elevene i et klassetrinn (Aune et al., 2018; Crawford et al., 2013; Olsen & Bjørnsson, 2018; Thoren et al., 2016). I denne studien er det derimot ikke funnet relativ alderseffekt i matematikk, mer spesifikt på den delen av aritmetiske ferdigheter som omhandler regneflyt. De tidligere studiene tar som regel utgangspunkt i TIMSS-undersøkelsene og denne undersøkelsen tar for seg fagområdene tall, geometri og statistikk, samt at de tar for seg algebra i 8. og 9. trinn (Universitetet i Oslo, 2017). Det er også en studie som tar for seg den norske nasjonale prøver i regning, som tar for seg fagområdene tall, måling og geometri og statistikk, samt algebra og sannsynlighet i 8. og 9. trinn (Utdanningsdirektoratet, 2017). Siden denne studien tar for seg en liten del av matematikkfeltet vil det kunne være problematisk å si noe om den relative alderseffekten på elevers generelle matematikkferdigheter slik som studiene som er basert på TIMSS-undersøkelsene og nasjonale prøver i regning. Studien viser til at det kan være ulikheter i hvordan alder påvirker elevene i de ulike ferdighetene og fagområdene i matematikk.

I Regnefaktaprøven addisjon forklarer kjønn en unik variasjon (β = .181, p = .001), men viser ingen signifikante forskjeller i Regnefaktaprøven subtraksjon (β = .105, p = .054) og WISC-IV regning (β = .085, p = .088). Resultatene viser at guttene presterer bedre enn jentene i addisjonsregneflyt. Generelt viser forskningen at det ikke er signifikante forskjeller mellom kjønnene (Hutchison et al., 2019; Nortvedt & Pettersen, 2016). Det er imidlertid funnet signifikante forskjeller i addisjons- og subtraksjonsoppgaver, men effektstørrelsen var relativ svak (Hutchison et al., 2019). Noe forskning viser til at det er lite som tyder på at jenter og gutter velger ulike strategier for å løse enkle aritmetiske oppgaver (Geary et al., 1996; Siegler, 1988). Det er derimot noen studier som har funnet at jenter bruker strategier som involverer bruk av konkreter som fingre mer enn gutter. Guttene bruker retrievalstrategien mer og den er regnet som en mer effektiv strategi (Carr & Davis, 2001; Jordan et al., 2008; Shen et al., 2016). Dette kan være en av faktorene som gjør at det er signifikante forskjeller mellom jenter og gutter i Regnefaktaprøven addisjon. I denne studien ble derimot ikke elevenes strategibruk undersøkt, men dette vil være interessant å undersøke i fremtidige studier om jenter og gutter bruker ulike strategier når de skal løse aritmetiske oppgaver.

5.1.2 Alderseffekt i aritmetiske problemløsningsoppgaver

Sammenlignet med resultatene fra regneflyt viser resultatene fra regresjonsanalysen med WISC-IV som avhengig variabel at alder forklarte unik variasjon i aritmetiske problemløsingsferdigheter (β = .102, p = .045) når det ble kontrollert for nonverbale evner og kjønn. Det vil si at dersom alder øker med ett standardavvik øker elevenes aritmetiske problemløsningsferdigheter med .102 standardavvik. Gjennomsnittet på antall riktige oppgaver på WISC-IV regning var 17.20 og variasjonsbredden var 6-31 av totalt 34 oppgaver. I WISC-IV regning var det ingen som fikk null riktige svar, til forskjell fra Regnefaktaprøven - addisjon og subtraksjon, og det var 3.3 % fikk under 10 oppgaver riktig og 15.7 fikk over 20 oppgaver riktig. Dette viser at det er en ferdighet elevene mestrer i ulik grad. Hvordan elevene presterer i aritmetiske problemløsningsoppgaver vil gjennom grunnskolen være knyttet til grunnleggende regneferdigheter som addisjon og subtraksjon, fleksibiliteten de har til valgt av ulike strategier og den semantiske strukturen (Carpenter & Moser, 1983; De Corte & Verschaffel, 1987; Riley et al., 1983; Jordan et al., 2003). Mestringene elevene opplever i aritmetiske problemløsningsoppgaver vil også være påvirket av alder (Gilmore et al., 2018). Dette kan være en av faktorene til at alder har en signifikant påvirkning på aritmetiske problemløsningsferdigheter og ikke på regneflyt. En annen faktor kan være på bakgrunn av at WISC-IV regning består av muntlig presenterte tekstoppgaver som skal løses ved hjelp av hoderegning uten bruk av hjelpemidler bortsett fra de fem første oppgavene som har støtte av bildestimulus (Wechsler et al., 2003). Disse oppgavene krever mer av elevene enn i Regnefaktaprøven - addisjon og subtraksjon. Elevene må i slike oppgaver danne seg et bilde av problemet, hente ut relevant informasjon, velge riktig regneoperasjon og gjennomføre den valgte regneoperasjonen (Geary, 1994; Gilmore et al., 2018).

Det er funnet relativ alderseffekt i aritmetiske ferdighetene som omhandler aritmetiske problemløsningsferdigheter og dette er noe som gjenspeiler seg i tidligere studier fra Norge. Disse studiene viste at de eldste i et klassetrinn presterer bedre enn sine yngre medelever i matematikk (Aune et al., 2018; Olsen & Bjørnsson, 2018). Som nevnt tar disse studiene for seg TIMSS-undersøkelsene og nasjonale prøver i regning, og de tar for seg flere fagområder i matematikk enn det denne studien gjør. Samtidig blir ofte matematiske problemløsningsoppgaver brukt for å kartlegge hvilke ferdigheter elevene har i ulike undersøkelser og tester, og i denne studien er det funnet signifikante forskjeller mellom de eldste i klasserommet og de yngste i aritmetiske problemløsningsoppgaver. Studiene i Norge som omhandler relativ alderseffekt og matematikk tar som regel utgangspunkt i fjerdeklasse og oppover (Aune et al., 2018; Dalen & Aune, 2013; Olsen & Bjørnsson, 2018), men denne studien viser at en kan se relativ alderseffekt allerede i andreklasse. Crawford et al. (2013) fant relativ alderseffekt i syv årsalderen i England i matematikk og Thoren et al. (2016) fant relativ alderseffekt hos tyske andreklassinger i matematikk. Dette viser at en kan se en slik effekt allerede når elevene er i andreklasse, og det på tvers av landegrensene.

Crawford et al. (2013) trekker frem i sin studie andre potensielle grunnen til at en ser ulike resultater mellom elever som er født sent på året og de som er født tidlig, i tillegg til relativ alderseffekt. En av grunnene de trekker frem er alderen til elevene ved testdato. Hvis alle elevene i et klasserom tar en test på samme dato, vil de barna som er født sent på året alltid være yngre enn sine medelever. I denne studien er datasettet basert på elever i ulik alder ved testdato. Dette er på bakgrunn av at elevene som gikk i samme klasse tok testene samtidig eller med få dagers mellomrom. Hvis en hadde testet elevene ved samme alder vil det være mulig at resultatet på Regnefaktaprøven - addisjon og subtraksjon og WISC-IV regning kunne vært annerledes. Hvis elevene hadde hatt samme alder ved gjennomføringen av WISC-IV regning er det mulig at en ikke ville ha sett en signifikant effekt og at det ikke eksiterer relativ alderseffekt i aritmetiske problemløsningsferdigheter. Dette er fordi de yngste elevene i denne studien vil være bortimot ett år yngre enn elevene de sammenlignes med. På bakgrunn av at datamaterialet er basert på alderen elevene hadde ved gjennomføringen av pre-testene kan en ikke kontrollere om effekten vil vært tilstede hvis elevene hadde blitt testet ved samme alder. Dette er en begrensning med denne studien. Crawford et al. (2013) fant i deres studie at alder ved testing var en nøkkelfaktor til ulikhetene i skårer utdannelsesnivå og skårer på kognitive tester mellom elevene som var født tidlig på året og de som var født sent på året. En kan derfor ikke være sikker om det ville oppstått en signifikant forskjell mellom de som er født tidlig på året og de som er født sent på året i aritmetiske problemløsningsferdigheter hvis elevene hadde blitt testet ved samme alder.

En annen grunn Crawford et al. (2013) trekker frem er alder ved skolestart. Bakgrunnen for dette er at de fleste land har en cut-off dato som avgjør når elevene skal starte på skolen. Dette kan være en faktor på bakgrunn av at det ikke er alle barn fra barnehagen som er «modne» for å delta i mer formell undervisning som er et hovedfokus i skolen, og at dette kan være en av grunnene til forskjeller i skårer. De trekker også frem mengde undervisning før testdato som er forklaring for forskjeller mellom de eldste og de yngste. Noen land følger en opptakspolicy som gjør at elever som er født sent på året kan starte på skolen ett til to semestre etter de som er født i starten av året. I Crawford et al. (2013) sin studie betydde dette at noen elever hadde ett til to semestre mindre undervisning ved de nasjonale prøvene i England som går på alder. I Norge et cut-off datoen for skolestart 1. januar, og det er med få unntak. Dette fører dermed til at elever som deltar i norske studier og denne studien vil ha like lang tid på skolen og lik mengde undervisning i forkant av testdato, og at forskjeller i prestasjoner ikke har vært påvirket av at elevene har hatt ulik lengde på skolegangen (Olsen & Bjørnsson, 2018). Dette betyr i denne studien at forskjellene i prestasjon i aritmetiske problemløsningsoppgaver mellom de eldste og de yngste elevene ikke er påvirket av lengden på skolegangen (Olsen & Bjørnsson, 2018). Crawford et al. (2013) fant lignende funn i deres studie. De fant at alderen ved skolestart og mengden undervisning før testdato forklarte lite av forskjellene i skårer mellom de eldste og de yngste. Forskjellen i skårer vil dermed kunne bli forklart av relativ alder.

5.1.3 Oppsummering

I denne studien forklarer alder ingen unik variasjon i regneflyt. Det vil si at alder ikke forklarte unik variasjon i addisjonsregneflyt og subtraksjonsregneflyt når det ble kontrollert for nonverbale evner og kjønn. Grunnen til at alder ikke forklare unik variasjon i regneflyt kan være på bakgrunn av at elevene i en tidlig alder blir eksponert for addisjons- og subtraksjonsoppgaver, og de er i stand til å utføre slike oppgaver i barnehagealder. Elevene i denne studien viser at regneflyt er en ferdighet elevene mestrer i ulik grad, men at de ligger rundt den forventende gjennomsnittsskåren i andreklasse. Dette er også en periode hvor de stadig utvikler strategibruken til å bli hurtigere og mer effektiv. Tidligere studier har funnet relativ alderseffekt i matematikk, men i denne studien er det ikke funnet en slik effekt i den delen av aritmetiske ferdigheter som omhandler regneflyt. I addisjonsregneflyt forklarer kjønn unik variasjon og resultatene viser at guttene presterer bedre enn jentene. Generelt viser forskning at det ikke er forskjeller mellom kjønnene (Hutchison et al., 2019; Nortvedt & Pettersen, 2016), men det er studier som har funnet at jenter bruker strategier som involverer bruk av konkreter og guttene bruker retrievalstrategien mer (Carr & Davis, 2001; Jordan et al., 2008; Shen et al., 2016).

Sammenlignet med resultatet fra regneflyt viser resultatet fra aritmetiske problemløsningsoppgavene at alder forklarte unik variasjon når det ble kontrollert for nonverbale evner og kjønn. Resultatene viser at aritmetiske problemløsningsferdigheter er en ferdighet elevene mestrer i ulik grad, og hvor godt elevene mestrer denne ferdigheten er knyttet til de grunnleggende regneferdighetene, fleksibelt strategivalg, den semantiske strukturen i oppgavene og alder (Carpenter & Moser, 1983; De Corte & Verschaffel, 1987; Gilmore et al., 2018; Riley et al., 1983; Jordan et al., 2003). Funn av relativ alderseffekt i aritmetiske ferdigheter som omhandler problemløsningsferdigheter gjenspeiler seg i tidligere forskning som viser at det finnes relativ alderseffekt i matematikk (Aune et al., 2018; Dalen & Aune, 2013; Olsen & Bjørnsson, 2018). Selv om denne studien tar for seg et lite fagområde i matematikken sammenlignet med andre studier. Tidligere studier viser som denne studien at en kan se relativ alderseffekt allerede i andreklasse (Crawford et al., 2013; Thoren et al., 2016).

Selv om denne studien bare tar for seg et lite fagområde i matematikk på et målepunkt og viser til at aritmetiske ferdigheter i ulik grad viser sammenheng med alder vil denne studien kunne være med å gi et bidrag på forskning som omhandler relativ alderseffekt i matematikk. Studien er viser at en kan se relativ alderseffekt hos norske andreklassinger i aritmetiske problemløsningsoppgaver og slike oppgaver blir ofte brukt for å kartlegge elevers matematikkferdigheter.

5.2 Resultater sett i lys av validitetsteori

Som nevnt i metodekapittelet handler validitet om en kan trekke gyldige slutninger basert på måledata (Lund, 2002a), og det handler om å gjøre vurderinger på i hvilken grad relevant bevis støtter slutningen som å være sann eller usann (Shadish et al., 2002). Videre vil kvaliteten på slutningene for denne studien vurderes.

5.2.1 Statistisk validitet

Statistisk validitet handler om den antatte årsaken og effekten korrelerer, og hvor sterk korrelasjonen eventuelt er (Shadish et al., 2002). Effektmålet i denne studien refererer til korrelasjonskoeffisienten Pearsons *r*, og regresjonskoeffisienten beta (β). Denne studien tar for seg om det er en sammenheng mellom alder og aritmetiske ferdigheter basert på regneflyt og aritmetiske problemløsningsferdigheter, og hvor sterk den eventuelle sammenhengen er. Det handler om å vurdere om denne sammenhengen skyldes tilfeldigheter eller ikke, og om forholdet er statistisk signifikant og rimelig sterkt. Hvis studien er statistisk invalid vil sammenhengen mellom alder og aritmetiske ferdigheter skyldes en målefeil, eller at størrelsen er triviell (Lund, 2002a).

En av truslene er brudd på statistiske forutsetninger. I denne studien viser Regnefaktaprøven - addisjon og subtraksjon til en skjevfordeling som heller mot venstre side. Begge har en skjevhetsverdi som ligger innenfor ± 2, og vil ifølge dette ikke føre til upålitelig slutningsstatistikk. WISC-IV regning viser ingen brudd på statistiske forutsetninger. Regnefaktaprøven addisjon har derimot er kurtosisverdi på 3.04, og Q-Q plotene til Regnefaktaprøven addisjon (vedlegg 2) og subtraksjon (vedlegg 4) viser til svak bue under linjen som kan indikere brudd på normalitet. En konsekvens for denne studien på bakgrunn av dette er at det kan være en over- eller underestimering av størrelsen og signifikansen til effekten av Regnefaktaprøven - addisjon og subtraksjon. Både addisjonsegneflyt og subtraksjonsregneflyt viser i denne studien ingen korrelasjon med alder, som kan vise til en underestimering. Dette kan videre få konsekvenser for både type I-feil og type II-feil. Utvalget i denne studien er derimot over 300 andreklassinger og «the central limit theorem» sier at distribusjonen av utvalget vil være normalfordelt uansett hvilken form dataen hadde når den ble samlet inn. En vil på bakgrunn av «the central limit theorem» og ved skjevhetsverdier under ± 2 kunne si at Regnefaktaprøven - addisjon og subtraksjon viser godkjente verdier for å kunne gjennomføre en regresjonsanalyse.

En annen trussel for god statistisk validitet er lav statistisk styrke, og styrken vil påvirkes av utvalgsstørrelsen, signifikansnivå og effektstørrelse. Statistisk styrke vil øke automatisk med stort utvalg, og ved større utvalg vil en trenge mindre ulikheter, forhold eller effekt for å forkaste nullhypotesen. Dette må samtidig tas med forbehold av at signifikansnivået og effektstørrelsen blir holdt konstant (Gall et al., 2007). I denne studien er utvalget på 332 andreklassinger og på bakgrunn av dette vil en trenge mindre ulikheter, forhold eller effekt for å forkaste nullhypotesen så lenge signifikansnivået og effektstørrelsen blir holdt konstant. Den andre faktoren som er avgjørende for statistisk styrke er nivået på p-verdien hvor en kan forkaste en nullhypotese (Gall et al., 2007). Statistisk styrke kan økes ved å senke nivået av signifikans som trengs for å forkaste nullhypotesen. En statistisk signifikanstest med en p-verdi satt på .10 er sterkere enn en samme test med p-verdi satt på .05 (sterkere vil si at en lettere kan forkaste en falsk nullhypotesen). Ved en p-verdi på .10 øker sannsynligheten for en type I- feil, men det kan muligens føre til fremheving av potensielt viktige ulikheter, forhold eller effekter som ellers ville ha blitt oversett (Gall et al., 2007). I denne studien er korrelasjonskoeffisienten signifikant på .01 nivå (to-halet test), og i Regnefaktaprøven - addisjon og subtraksjon viser det ingen statistisk signifikante forskjeller i korrelasjonen med alder. WISC-IV regning viser derimot det statistisk signifikante forskjeller på dette nivået i korrelasjonen med alder. Regresjonskoeffisienten viser at det ikke er korrelasjon mellom Regnefaktaprøven - addisjon og subtraksjon og alder, selv når det kontrolleres for nonverbale evner og kjønn. WISC-IV regning viser signifikante forskjeller på .05 nivå med alder når det kontrolleres for nonverbale evner og kjønn.

Effektstørrelsen er et estimat på ulikhetene, forholdet eller effekten i populasjonen som blir studert. Ved større observerte ulikheter, forhold eller effekter vil det produseres lavere p-verdier, samt at en vil i større grad oppnå en stor effektstørrelse i utvalget når det er stor effektstørrelse i populasjonen (Gall et al., 2007). Som nevnt er det ingen observerte ulikheter, forhold eller effekter mellom Regnefaktaprøven - addisjon og subtraksjon og alder. Dette kan være på bakgrunn av at det er små effektstørrelser i populasjonen. Hvis en hadde økt p-verdien ville en muligens kunne oppdage potensielle ulikheter, forhold eller effekter, men det ville samtidig ha økt sjansen for en type I-feil. WISC-IV regning viser svak korrelasjon med alder, og korrelasjonen er enda signifikant når det kontrolleres for nonverbale evner og kjønn. Dette viser til et en kan forkaste nullhypotesen på bakgrunn av at det er observert ulikheter, forhold eller effekt.

Upålitelighet i måleinstrumentene er også en trussel mot statistisk validitet. Lav reliabilitet kan føre til feil i estimeringen av størrelsen på observerte effekter og redusere bivariate forhold (Gall et al., 2007; Shadish et al., 2002). På bakgrunn av dette burde reliabilitet bli estimert og rapportert for hvert måleinstrument som blir brukt i en studie (Shadish et al., 2002), og dette er noe som har blitt gjort i denne studien. Regnefaktaprøven addisjon gir en reliabilitetskoeffisient på .930 og Regnefaktaprøven subtraksjon gir en koeffisient på .913. Måleinstrumenter som gir reliabilitetskår .80 eller høyere viser tilstrekkelig reliabilitet og tester som er standardiserte kan oppnå reliabilitetskoeffisient på .90 eller høyere (Gall et al., 2007). Både Regnefaktaprøven - addisjon og subtraksjon er standardiserte tester og får en reliabilitetskår høyere enn .90 og da vil en kunne si med rimelig sikkerhet at måleinstrumentene måler det de er ment å måle. WISC-IV regning er også et standardisert kartleggingsverktøy, men har en reliabilitetskoeffisient på .757. På bakgrunn av at WISC-IV regning er et standardisert verktøy vil den kunne oppnå en reliabilitetskoeffisient på .90 (Gall et al., 2007), samtidig er en skår på .07 eller høyere også beregnet som en akseptabel skår (Pallant, 2016). Testen består av muntlig presentere tekstoppgaver og disse oppgavene skal løses ved hjelp av hoderegning uten bruk av hjelpemidler i løpet av 30 sekunder (Wechsler et al., 2003). Dette stiller store krav til at barna klarer å holde all informasjonen de trenger samtidig som de skal regne ut svaret. Reliabilitetskoeffisienten kan vise til at måleinstrumentet måler andre faktorer enn barns matematiske problemløsningsferdigheter, eller at noen av items i WISC-IV regning måler noe annet enn flertallet av items. Andre faktorer som WISC-IV regning kan måle er arbeidsminne og lytteforståelse. Resultatene samlet tyder på at det er stor grad av indre konsistens i måleinstrumentene som er inkludert i denne studien.

5.2.2 Indre validitet

Indre validitet omhandler hvorvidt en kan trekke pålitelige slutninger om at sammenhengen mellom variablene er kausal og uten at det kan forklares av andre faktorer (Shadish et al., 2002). En trussel på indre validitet er dermed tredjevariabler og retningsproblem. I denne studien er formålet å se om det er en sammenheng mellom alder og aritmetiske ferdigheter. Her vil det være mulig at det er andre faktorer som kan forklare variasjonen i aritmetiske ferdigheter. Det blir kontrollert for kjønn og nonverbale evner, men i regresjonsanalysen forklarer nonverbale evner, kjønn og alder (modell 2) i Regnefaktaprøven addisjon 9.2 % av variasjonen, i Regnefaktaprøven subtraksjon 5 % av variasjonen og i WISC-IV regning forklarer modell 2 19.7 % av variasjonen. I en regresjonsanalyse er også målet å identifisere potensielle uavhengige variabler som kan predikere den avhengige variabelen (Befring, 2015). I regresjonsanalysene i denne studien vil det mest sannsynlig være andre faktorer som forklarer variasjonen i regneflyt og aritmetiske problemløsningsferdigheter som det ikke ble kontrollert for og som kan predikere aritmetiske ferdigheter bedre. De bivariate korrelasjonene viser at addisjonsregneflyt og subtraksjonsregneflyt korrelerer sterk med hverandre, og begge korrelerer moderat med aritmetiske problemløsningsferdigheter. På bakgrunn av dette er det mulig at disse vil forklare unik variasjon, men dette er ikke kontrollert i regresjonsanalysene.

Gjennom de bivariate korrelasjonene og regresjonsanalysene vil en ikke kunne si noe om retningsforholdet. På bakgrunn av at det vil være umulig å avgjøre hva som er årsak og hva som er virkning (Shadish, et al., 2002). En av grunnene er at denne studien er av et ikke-eksperimentelt design og en tverrsnittstudie. I slik forskning er en opptatt av sammenhenger og ikke årsaksforhold. Denne studien vil ha lavere indre validitet på bakgrunn at sammenhengen alltid vil være forenelig med flere mulige årsakrelasjoner (Kleven, 2002a), selv om det det er gjort kontrollert statistisk for alternative årsaksforklaringer (Lund, 2002a).

5.2.3 Begrepsvaliditet

Begrepsvaliditet handler om å uttrykke samsvar mellom det teoretiske begrepet og gjennomført måling (Kleven, 2002b). Begrepet aritmetiske ferdigheter er i denne studien målt gjennom elevenes regneflyt i addisjon og subtraksjon, samt aritmetiske problemløsningsferdigheter. Fra elevene er små består den aritmetisk utvikling av en endring i bruk av strategier, og hvor nøyaktig og hurtig strategiene kan bli utført (Geary, 1994). Hvor nøyaktig og hurtig elevene kan utføre de valgte strategiene vil ses i sammenheng med hvor effektive de er på Regnefaktaprøven - addisjon og subtraksjon, og dermed et mål på hvor god regneflyt elevene har. Reliabilitetsskåren til Regnefaktaprøven addisjon er .930 og for subtraksjon er den .913, samt at de er standardiserte kartleggingsverktøy. Som nevnt er vurderingen av begrepsvaliditet en vurdering på om det målte begrepet oppfører seg slik det teoretiske begrepet forventer å oppføre seg, og selv om reliabilitet ikke er noe mål i seg selv vil svak reliabilitet svekke begrepsvaliditeten (Kleven, 2002b). Hvordan regneflyt er teoretisk definert samsvarer med hvilken reliabilitetskoeffisient de får. Både Regnefaktaprøven - addisjon og subtraksjon får en skår over .90 som er vurdert som god reliabilitet og er standardiserte, og ved en høy intern konsistens som dette vil en kunne si med rimelig sikkerhet at de måler det de er ment å måle.

Under selve datainnsamlingen kan det oppstå feilkilder som kan svekke begrepsvaliditeten (Kleven, 2002b). Ved gjennomføringen av Regnefaktaprøven subtraksjon var det elever som startet å bruke regneoperasjonen addisjon i stedet for subtraksjon på bakgrunn av at testene ble gjennomført etter hverandre og ikke tilstrekkelige instruksjoner ble gitt. Dette førte til at noen elever fikk null som skår. Disse elevene utgjør bare 10.5 % av utvalget og ved en reliabilitetsskår på .913 kan det tyde på at det ikke har hatt for stor innvirkning på begrepsvaliditeten. Dette fører samtidig til en begrensning ved denne studien og i fremtidige studier vil det være viktig å legge større vekt på at elevene har forstått instruksjonene slik at disse misforståelse unngås. Gjennomføringen av Raven i VLC-prosjektet avviker fra hvordan måleinstrumentet er standardisert. Raven ble gjennomført i klasserom og ikke individuelt. Øvingsoppgavene ble gjennomført i plenum, mens elevene gjennomførte resten av testen i sitt eget tempo. Dette kan ha påvirket resultatene elevene fikk på Raven og er en begrensning ved denne studien.

WISC-IV regning har som mål å kartlegge elevers matematiske problemløsningsferdigheter, og i de aritmetiske problemløsningsoppgavene blir det beskrevet situasjoner hvor det skjer endringer, bytter eller en beskrivelse av mengde og forholdet mellom disse (Riley et al., 1983; Thevenot & Barrouillet, 2015). For å kunne løse disse problemene må elevene danne seg et bilde av situasjonen, hente ut relevant informasjon, velge riktig regneoperasjon og gjennomføre den valgte regneoperasjonen (Geary, 1994; Gilmore et al., 2018). De må også velge seg ut en strategi for å løse problemet (Geary, 1994). Oppgavene i WISC-IV regning blir også muntlig presentert uten bruk av hjelpemidler etter de fem første oppgavene og skal løses på 30 sekunder. På bakgrunn av dette er det mulig at begrepet oppfører seg annerledes enn slik det teoretisk er forventet å oppføre seg. WISC-IV regning har også en reliabilitetsskår på .757 som støtter opp at det kan være andre faktorer som blir testet eller at noen av item i måleinstrumentet måler noen annet en flertallet av items, som for eksempel arbeidsminne og lytteforståelse. Samtidig er WISC-IV regning er standardisert kartleggingsverktøy. På bakgrunn av reliabilitetsskåren vil en ikke kunne si mer rimelig sikkerhet at måleinstrumentet måler det den er ment å måle. Samtidig får måleinstrumentet en skår på .07 eller høyere som er beregnet som en akseptabel skår på reliabilitet og testen er standardisert. Dette vil kunne støtte opp om at WISC-IV regning måler det den er ment å måle.

Trusler mot begrepsvaliditet kan være tilfeldige målefeil eller systematiske målefeil. Tilfeldige målefeil menes med at feilene oppfører seg tilfeldige, og kan komme av dagsformen elevene hadde under testsituasjonen, og det kan påvirkes av at noen elever liker matematikk og noen ikke. De tilfeldige målefeilene oppfører seg litt som «flaks og uflaks» (Kleven, 2002b). Systematiske målefeil går mer på om tendensen av målingene går i samme retning ved gjentatte målinger, og dette kan for eksempel komme av at noen elever kan ha matematikkangst. Samtidig er det vanskelig å vurdere om dette er en systematisk målefeil eller en tilfeldig målefeil på bakgrunn av at i praksis vil dette være vanskelig å kunne vurdere. Alle type målinger vil også ha mer eller mindre usikker begrepsvaliditet, og dette er heller ikke et krav som må være perfekt oppfylt for å oppnå god forskning (Kleven, 2002b).

5.2.4 Ytre validitet

Ytre validitet handler om resultatene kan generaliseres til og over de relevante individene, situasjonene og tidene som problemstillingen tar for seg (Lund, 2002b). Utvalget i denne studien består av 332 norske enspråklige andreklassinger. Utvalget består av 179 jenter og 153 gutter fra 12 skoler og er fra 32 klasser. Utvalget er også hentet fra tre kommuner på Østlandet. Ytre validitet vil i denne studien handle om en kan generalisere resultatene fra dette utvalget til populasjonen av norske enspråklige andreklassinger i de tre kommunene, og deretter fra populasjonen i de tre kommunene til målpopulasjonen av norske enspråklige andreklassinger på landsbasis (Lund, 2002b).

En trussel mot ytre validitet er hvis individene i utvalget er relativ ensartet, samt hvis det er en skjevhet i utvalget sammenlignet med populasjonen (Lund, 2002a). Utvalget tar for seg et relativ jevnt utvalg av norske enspråklige jenter og gutter. Dette vil være med å støtte opp at funnene både kan generaliseres til jenter og til gutter. Utvalget kommer også fra tre kommuner på Østlandet fordelt på 12 skoler og 32 klasserom. Det vil også støtte opp om generalisering av resultatene til populasjonen i de tre kommunene og til en generalisering til populasjonen på landsbasis. Elevene vil for eksempel ha ulik bakgrunn, være utsatt for ulike undervisningsmetoder og ha fokus på ulike deler i fag, som videre vil påvirke utvalget til å være en heterogen gruppe.

En forutsetning for å oppnå god ytre validitet er ikke bare basert på individer, situasjoner og tider er valgt ut, men vil også basere seg på tidligere kunnskap fra undersøkelser og studier (Lund, 2002b). I denne studien er det funnet relativ alderseffekt hos norske enspråklige andreklassinger i aritmetiske problemløsningsferdigheter. Dette er i samsvar med tidligere studier på relativ alderseffekt i matematikk i Norge og andre land (Aune et al., 2018; Olsen & Bjørnsson, 2018: Crawford et al., 2013; Thoren et al., 2016). Det er funnet relativ alderseffekt i matematikk hos elever som er syv år (Crawford et al., 2013) og for elever som går i andreklasse (Thoren et al., 2016). I Norge er det funnet aldersforskjell i matematikk elever i 4. og 5. trinn, samt i 8. og 9. trinn (Aune et al., 2018 ; Olsen & Bjørnsson, 2018). Studiene som er gjort på relativ alderseffekt og matematikk tar for seg flere fagområder enn denne studien, men viser at det finnes signifikante forskjeller på flere av fagområdene i matematikk. Det er også funnet alderseffekter i andre fag som naturfag (Olsen & Bjørnsson, 2018) og norsk (Dalen & Aune, 2013), samt i lesing (Gabrielsen & Lundetræ, 2017; Strøm, 2004). Relativ alderseffekt er også et kjent fenomen i kroppsøving og i ulike sportsgrener (Aune et al., 2017; Cobley et al,, 2008; Sæther et al., 2017). Disse effektene kan også ses i andre land som Tyskland (Thoren et al., 2016), Japan (Kawaguchi, 2011) og England (Cobley et al., 2009; Crawford et al., 2013), samt i land som har likt system for skolestart som Norge (Kawaguchi, 2011) og ulikt (Cobley et al., 2009; Crawford et al., 2013; Thoren et al., 2016).

Selv om det ikke er gjort mange studier på aritmetiske ferdigheter hos andreklassinger viser flere studier som tar for seg matematikk i ulike klassetrinn at det finnes relativ alderseffekt i norsk skole og at det er et fenomen som kan ses i andre land, samt i ulike fag og i sport. Dette vil være med på å styrke denne studiens ytre validitet og til en mulig generalisering fra dette utvalget til populasjonen av norske enspråklige andreklassinger i de tre kommunene på

Østlandet, og deretter fra populasjonen i de tre kommunene til målpopulasjonen av norske enspråklige andreklassinger på landsbasis.

5.3 Avslutning

Formålet med denne studien var å undersøke forholdet mellom alder og andreklassingers aritmetiske ferdigheter. Første forskningsspørsmålet omhandlet om det var sammenheng mellom alder og regneflyt når det kontrollertes for nonverbale evner og kjønn. Resultatene fra analysene viste at det ikke er noen sammenheng mellom alder og regneflyt hos norske enspråklige andreklassinger. Det andre forskningsspørsmålet tok for seg om det var en sammenheng mellom alder og aritmetiske problemløsningsferdigheter når det kontrolleres for nonverbale evner og kjønn. Resultatene fra analysen viser at det er en signifikant sammenheng mellom alder og aritmetiske problemløsningsferdigheter, og at alder forklarer unik variasjon i aritmetiske problemløsningsoppgaver. Resultatene viser til at de elevene som er født tidlig på året presterer bedre i aritmetiske problemløsningsoppgaver enn sine medelever som er født sent på året. Nonverbale evner er den faktorer som er målt i denne studien som best predikerer aritmetiske ferdigheter, og kjønn viste bare signifikant sammenheng med Regnefaktaprøven addisjon. Resultatene viste at gutter presterer bedre enn jentene på addisjonsoppgaver.

Denne studien vil ha lavere indre validitet på bakgrunn av at det er en ikke-eksperimentell studie, samt en tverrsnittstudie. Ikke-eksperimentelle studier tar for seg sammenhenger og ikke årsaksforklaringer, og sammenhengen mellom alder og aritmetiske ferdigheter vil alltid være forenelig med flere mulige årsakrelasjoner. Samtidig vil en ikke kunne si noe om hva som er årsak og hva som er virkning. Det er kontrollert statistisk for alternative årsaksforklaringer i denne studien da det ble kontrollert for nonverbale evner og kjønn i regresjonsanalysen.

Aritmetiske ferdigheter er i denne studien målt gjennom regneflyt i addisjon og subtraksjon, og gjennom aritmetiske problemløsningsoppgaver. Regneflyt ble målt gjennom Regnefaktaprøven - addisjon og subtraksjon, og aritmetiske problemløsningsoppgaver gjennom WISC-IV regning. Regnefaktaprøven hadde som formål å måle elevers regneflyt og oppnådde en reliabilitetskoeffisient på .930 på delprøven om addisjon og .913 på delprøven subtraksjon. Dette er vurdert som god reliabilitet og Regnefaktaprøven er et standardisert kartleggingsverktøy, og en vil derfor si med rimelig god sikkerhet at den målte det den er ment å måle. WISC-IV regning hadde som formål å måle barn matematiske problemløsningsferdigheter, og har en reliabilitetsskår på .757. En vil dermed kunne si med rimelig god sikkerhet at den målte det en er ment å måle.

Utvalget i denne studien tar for seg begge kjønn og kommer fra tre kommuner på Østlandet fordelt på 12 skoler og 32 klasserom, og dette støtter opp om en heterogenitet i utvalget. Ved stort utvalg og en heterogen gruppe vil styrke den ytre validiteten og en vil kunne generalisere resultatene på at det som er født tidlig på året presterer bedre i aritmetiske problemløsningsoppgaver enn de som er født sent på året til populasjonen av norske enspråklige andreklassinger i de tre kommunene. Resultatene fra denne studien får også støtte fra tidligere undersøkelser og studier på relativ alderseffekt på matematikk og andre akademiske fag, samt på tvers av landegrenser. Dette vil være med på å styrke den ytre validiteten, og en vil dermed kunne generalisere til populasjonen av norske enspråklige andreklassinger i de tre kommunene, og deretter fra populasjonen i de tre kommunene til målpopulasjonen av norske enspråklige andreklassinger på landsbasis.

5.3.1 Implikasjoner for praksis

Resultatene fra denne studien viser at det er en relativ alderseffekt i aritmetiske ferdigheter, mer spesifikt i aritmetiske problemløsningsferdigheter hvor de eldste presterer signifikant bedre enn sine yngre medelever. Dette kan ha implikasjoner for elever og lærere i skolehverdagen på bakgrunn av at elevene ofte blir vurdert gjennom aritmetiske problemløsingsoppgaver og slike type oppgaver gir et innblikk i elevenes ferdigheter (Gilmore et al., 2018; Thevenot & Barrouillet, 2015). Elever som går i samme klasse vil også ofte konkurrere og bli vurdert opp mot hverandre (Dalen & Aune, 2013). Resultatene fra denne studien tyder på at elevene trenger å få tilpasset undervisningen i matematikk på deres nivå slik at de elevene som er yngst i klasserommet skal få lik mulighet til å utvikle ferdighetene i matematikk og oppleve mestring på lik linje som sine eldre medelever. En av grunnene til dette er at elever som går inn i skolen med svakere ferdigheter har tendenser til å fortsette å prestere svakere enn sine medelever (Aunola et a., 2004; Morgan et al., 2009), og utover skolegangen stilles det stadig høyere krav til matematikkferdigheter. Matematikk er også et fag som kan vekke sterke følelser og vansker i matematikk kan resultere i frustrasjon, unngåelse, lav mestringstro og potensielt matematikkangst (Geary, 2017; Rubinsten &

Tannock, 2010). Tidligere studier viser at de yngste elevene i et klasserom er overrepresentert blant de elevene som ble diagnostisert med en spesifikk lærevanske og de har mer spesialundervisning (Cobley et al., 2009; Crawford et al., 2013). De yngste elevene i klasserommet har større sannsynlighet for å utvikle negative holdninger og atferd mot utdanning (Cobley et al., 2009). Det blir derfor viktig å tilpasse undervisningen til de elevene som er født sent på året på bakgrunn av at når de er født på året burde ikke være en faktor som fører til at noen elever presterer svakere i ulike matematikkoppgaver. Dette er viktig på bakgrunn av at det er en faktor lærere og politikere ikke kan endre eller påvirke. Undervisning og skolehverdagen er derimot faktorer en kan påvirke slik at de elevene som er født sent på året får like muligheter som sine eldre medelever. På bakgrunn av denne studien finner en relativ alderseffekt allerede i andreklasse viser det til at en burde sette inn tiltak tidlig slik at de skal få samme progresjon i matematikkutviklingen som sine eldre medelever. En annen grunn er at tidligere studier finner signifikante forskjeller i 15/16 årsalderen og ved å sette inn tiltak tidlig vil muligens hindre at relativ alderseffekt fører til en vedvarende svake matematikkferdigheter utover skolegangen.

Camilla Stoltenberg har som nevnt fremmet en debatt om fleksibel skolestart på bakgrunn av en modningshypotese som går ut på at flere av guttene som er «umodne» før skolestart med fordel kan starte senere på skolen (NOU, 2019:3). Hun fortalte også at jenter som er født tidlig på året biologisk sett kan være flere år eldre enn gutter som er født sent på året (Stoltenberg, 2017). Resultatene fra denne studien viser at en i aritmetiske ferdigheter ikke kan se at guttene presterer svakere enn jentene i andreklasse, men viser heller til det omvendte. Hvis det er noen elever som kan ha nytte av fleksibel skolestart er det de elevene som er født sent på året uansett om de er jente eller gutt, men dette gjelder for ferdigheter i aritmetikk. Bakgrunnen for at disse elevene kunne hatt nytte av et slikt tilbud er at noen av de yngste elevene kan være bort imot 11 måneder yngre enn de eldste i et klasserom, og disse elevene vil som regel alltid bli vurdert opp mot hverandre.

5.3.2 Videre studier på relativ alderseffekt

Denne studien tar for seg norske enspråklige andreklassinger og er et nytt bidrag i norsk forskning på relativ alderseffekt. Denne studien viser at en kan se en relativ alderseffekt i aritmetiske problemløsningsoppgaver allerede i andreklasse. Denne studien tar for seg alder ved samme testdato, men det ville være interessant å se om det er en slik effekt når elever tar en test ved samme alder. Dette er på bakgrunn av funnene fra studien til Crawford et al. (2013) som fant at alder ved testdato var en nøkkelfaktor til ulikhetene i skårer på ulike tester mellom de elevene som var født tidlig på året og de som var født sent på året. Samtidig vil elever som går i samme klasse i den norske skolen vil alltid kunne ha en aldersforskjell på elleve måneder på bakgrunn av cut-off datoen som er 1. januar.

Tidligere studier tar som regel for seg elever ved et målepunkt og hvor de finner signifikante forskjeller i relativ alderseffekt i barneskolen og ungdomsskolen (Aune et al., 2018; Olsen & Bjørnsson, 2018: Crawford et al., 2013; Thoren et al., 2016). Denne studien viser at en kan finne en relativ alderseffekt i starten av barneskolen i andreklasse, men at slik effekt også kan ses i ungdomsskolen og videregående skole fra tidligere studier(Cobley et al., 2009; Dalen & Aune, 2013; Kawaguchi, 2011; Olsen & Bjørnsson, 2018; Thoren et al., 2016; Solli, 2017; Strøm, 2004). Samtidig viser studiene fra Norge og andre land at denne effekten avtar utover skolegangen (Bedard & Dhuey, 2006; Crawford et al., 2013; Kawaguchi, 2011; Strøm, 2004; Thoren et al., 2016). For å lære mer om hvordan relativ alderseffekten påvirker elevene i skolehverdagen og om det er en vedvarende faktor vil det være hensiktsmessig å gjennomføre longitudinelle studier for å se hvordan elevene som er født sent på året utvikler seg over tid sammenlignet med sine eldre medelever. Samt for å se i hvor stor grad relativ alder påvirker elevene i matematikk og andre akademiske fag. Det vil også være interessant å se hvilke fagområder i matematikk hvor effekten er mest fremtredende. I denne studien er det funnet relativ alderseffekt i aritmetiske problemløsningsferdigheter, men ikke i regneflyt. Ved å studere hvilke fagområder i matematikk og andre akademiske fag hvor en eventuell relativ alderseffekt er mest fremtredende vil en kunne få et mer helhetlig bilde. En vil da kunne få et større innblikk i hvordan det påvirker elevene i skolehverdagen og hvilke områder det er viktig å ha fokus på og støtte opp i undervisningen for de elevene som er født sent på året.

Litteraturliste

Andersen, Ø. & Smestad, T. (2018, 26 oktober). Fødselsmåned påvirker skoleprestasjoner. Hentet 03.04.19, fra https://www.uv.uio.no/cemo/forskning/aktuelle-saker/2018/fodselsmaned-pavirker-skoleprestasjoner.html.

Ashcraft, M. H. (1992). Cognitive arithmetic – a review of data and theory. *Cognition*, 44:1-2, 75-106, doi: 10.1016/0010-0277(92)90051-I.

Aune, T. K., Ingvaldsen, R. P., Vestheim, O. P., Bjerkeset, O. & Dalen, T. (2018). Relative Age Effects and Gender Differences in the National Test of Numeracy – A Population Study of Norwegian Children. *Frontiers in Psychology,* 9:1091, doi: 10.3389/fpsyg.2018.01091

Aune, T. K., Pedersen, A. V., Ingvaldsen, R. P. & Dalen, T. (2017). Relative Age Effect and Gender Differences in Physical Education Attainment in Norwegian Schoolchildren. *Scandinavian Journal of Educational Research*, 61:3, 369-375, doi: 10.1080/00313831.2016.1148073.

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 Research Journal*, 684–704, doi: 10.1080/1350293X.2014.996424

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology, 96*(4), 699-713. doi:10.1037/0022-0663.96.4.699.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L. & Jacobsen, S. J. (2005). Math Learning Disorder: Incidence in a Population-Based Birth Cohort, 1976–82, Rochester, Minn. *Ambulatory Pediatrics*, 5:5, 281-289, doi: 10.1367/A04-209R.1.

Baroody, A. J. & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. I J. Hiebert (red.), *Conceptual and procedural knowledge: The case of mathematics* (s. 75-112). New Jersey: Lawrence Erlbaum Associates, Inc.

Baroody, A. J. & Tiilikainen, S. H. (2003). Two Perspective on Addition Development. I A. J. Baroody & A. Dowker (red.), *The development of arithmetic concepts and skills – constructing adaptive expertise* (s. 75- 125). New Jersey: Lawrence Erlbaum Associates, Inc.

Baroody, A. J. (1987). The Development of Counting Strategies for Single-Digit Addition. *Journal for Research in Mathematics Education*, 18:2, 141-157, doi: 10.2307/749248.

Baroody, A. J. (1989). Kindergartners' mental addition with single-digit combinations. *Journal for Research in Mathematics Education, 20*:2, 159-172, doi: 10.2307/749280.

Bedard, K. & Dhuey, E. (2006). The Persistence of Early Childhood Maturity – International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics*, 121: 4, 1437–1472, doi: 10.1093/qje/121.4.1437.

Befring, E. (2015). *Forskningsmetoder i utdanningsvitenskap.* Oslo: Cappelen Damm Akademiske.

Butterworth, B. (2005). The development of arithemtical abilities. *J Child Psychol Psychiatry,* 46:1, 3-18, doi: 10.1111/j.1469-7610.2004.00374.x.

Carpenter, T. P. & Moser, J. M. (1982). The Development of Addition and Subtraction Problem-Solving Skills. I T. P. Carpenter, J. M. Moser & T. A. Romberg (red). *Addition and Subtraction – A Cognitive Perspective* (s. 9-24). New Jersey: Lawrence Erlbaum Associates, Inc.

Carpenter, T. P. & Moser, J. M. (1983). The Acquisition of Addition and Subtraction Concepts. I R. Lesh & M. Landau (red.). *Acquisition of Mathematics Concepts and Processes* (s. 7-44). New York: Academic Press, Inc.

Carr, M. & Davis, H. (2001). Gender Differences in Arithmetic Strategy Use: A Function of Skill and Preference. *Contemporary Educational Psychology* 26, 330–347, doi: 10.1006/ceps.2000.1059.

Christophersen, K. A. (2012). *IBM SPSS/AMOS – Databehandling og statistisk analyse*. Oslo: Akademika forlag.

Cobley, C., Abraham, C. & Baker, J. (2008). Relative age effects on physical education attainment and school sport representation. *Physical Education and Sport Pedagogy,* 13:3, 267-276, doi: 10.1080/17408980701711983.

Cobley, S., McKenna, J., Baker, J. & Wattie, N. (2009). How Pervasive Are Relative Age Effects In Secondary School Education? *Journal of Educational Psychology*, 101:2, 520-528, DOI: 10.1037/a0013845.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates, Inc.

Cowan, R. (2003). Does It All Add Up? Changes in Children`s Knowledge of Addition Combinations, Strategies, and Principles. I A. J. Baroody & A. Dowker (red.), *The development of arithmetic concepts and skills – constructing adaptive expertise* (s. 35-74). New Jersey: Lawrence Erlbaum Associates, Inc.

Crawford, C., Dearden, L. & Greaves, E. (2013). *When you are born matters: evidence for England*. Hentet 22. 11. 2018, fra http://www.ifs.org.uk/comms/r80.pdf.

Dalen, T, & Aune, T.K. (2013) Relativ alderseffekt ved karaktersetting i skolen. I I. Pareliussen, B. B Moen, A. Reinertsen & T. Solhaug. *FoU i praksis 2012 conference proceedings*, 62-68. Trondheim: Akademika forlag

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *Journal for Research in Mathematics Education, 18*:5, 363-381, doi: 10.2307/749085.

Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Hentet 22.11.18, fra https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf.

Devine, A., Soltész, F., Nobes, A., Goswami, U. & Szűcs, D. (2013). Gender differences in developmental dyscalculia depend on diagnostic criteria. *Learning and Insctruction*, 27, 31-39, doi: 10.1016/j.learninstruc.2013.02.004.

Dowker, A. (1998). Individual differences in normal arithmetical development. I C. Donlan (red.). *The development of mathematical skills* (s. 275-302). East Sussex: Psychology Press Ltd.

Dowker, A. (2005). *Individual Differences in Arithmetic – Implications for psychology, neuroscience and education.* New York: Psychology Press.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics, 5th edition*. London: SAGE Publications, Ltd.

Fuson, K. C. (1982). An Analysis of the Counting-On Solution Procedure in Addition. I T. P. Carpenter, J. M. Moser & T. A. Romberg (red). *Addition and Subtraction – A Cognitive Perspective* (s. 67-81). New Jersey: Lawrence Erlbaum Associates, Inc.

Fuson, K. C. (1988). *Children`s counting and concepts of number.* New York: Springer-Verlag.

Fuson, K. C. (1992). Research on Whole Number Addition and Subtraction. I D. A. Grouws (red.) *Handbook of Research on Mathematics Teaching and Learning (*s. 243-275). New York: Macmillian Publishing Company.

Gabrielsen, E. & Lundetræ, K. (2017). Indikerer de norske PIRLS-resultatene et behov for å justere retningslinjene for skolestartsalder? I E. Gabrielsen (red.). *Klar framgang! Leseferdighet på 4. og 5. trinn i et femtenårsperspektiv* (s. 204-221). Oslo: Universitetsforlaget.

Gall, M. D., Gall, J. P. & Borg, W. R. (2007). *Educational Research: An Introduction, 8th Edition*. New York, NY Pearson Education.

Geary, D. C. & Hoard, M. K. (2005). Learning Disabilities in Arithmetic and Mathematics – Theoretical and Empirical Perspectives. I J. I. D. Campbell (red.). *Handbook of Mathematical* Cognition (s. 253-267). New York: Psychology Press.

Geary, D. C. (1990). A Componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology,* 49:3, 363-383, doi: 10.1016/0022-0965(90)90065-G

Geary, D. C. (1993). Mathematical disabilities - Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*:2, 345-362, doi: 10.1037/0033-2909.114.2.345

Geary, D. C. (1994). *Children`s Mathematical Development – Research and Practical Applications.* Washington, DC: American Psychological Association.

Geary, D. C. (2000). From infancy to adulthood – the development of numerical abilities. *European Child & Adolescent Psychiatry*, 9:2, 11-16, doi: 10.1007/s007870070004.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37:1, 4-15, doi: 10.1177/00222194040370010201.

Geary, D. C. (2015). The Classification and Cognitive Characteristics of Mathematical Disabilities in Children. I R. C. Kadosh & A. Dowker (Red.), *The Oxford Handbook of Numerical Cognition* (s. 767-786). United Kingdom: Oxford University Press.

Geary, D. C. (2017). Dyscalculia at an Early Age. In: Tremblay RE, Boivin M, Peters RDeV, eds. *Encyclopedia on Early Childhood Development* [online]. Hentet 25.02.2019, fra http://www.child-encyclopedia.com/learning-disabilities/according-experts/dyscalculia-early-age

Geary, D. C., Bow-Thomas, C., Liu, F. & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children - Influence of Age, Language, and Schooling. *Child Development,* 67:5, 2022-2044, doi: 10.1111/j.1467-8624.1996.tb01841.x.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition - A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27*:5, 787-797. doi:10.1037/0012-1649.27.5.787

Geary, D. C., Hoard M.K., Nugent L. & Bailey D.H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement - A five year prospective study. *Journal of Educational Psychology,* 104: 1, 206–223, doi: 10.1037/a0025398.

Gelman, R. & Gallistel, C. R. (1978). *The child`s understanding of number*. Cambridge, MA: Harvard University Press.

Gilmore, C., Göbel, S. M. & Inglis, M. (2018). An introduction to mathematical cognition. London; New York: Routledge, Tayler & Francis Group.

Göbel, S. M., Watson, S. E., Lervag, A., & Hulme, C. (2014). Children's arithmetic development - it is number knowledge, not the approximate number sense,that counts. *Psychological Science*, 25:3, 789-798, doi: 10.1177/0956797613516471.

Haskell, S.H. (2000). The determinants of arithmetic skills in young children - some observations. *European Child & Adolescent Psychiatry*, 9:2, 77-86, doi: 10.1007/s007870070011

Helsen, W. F., Baker, J., Schorer, J., Steingröver, C., Wattie, N. & Starks, J. L. (2016). Relative age effects in a cognitive task – A case study of youth chess. *High Ability Studies*, 27:2, 211-221, doi: 10.1080/13598139.2016.1242063

Hutchison, J. E., Lyons, I. M. & Ansari, D. (2019). More Similar Than Different: Gender Differences in Children's Basic Numerical Skills Are the Exception Not the Rule. *Child Development*, 90:1, 66-79, doi: 10.1111/cdev.13044

Huttenlocher, J., Jordan, N. C. & Levine, S. C. (1994). A Mental Model for Early Arithmetic. *Journal of Experimental Psycholog*, 123:3, 284-296, doi: 10.1037/0096-3445.123.3.284.

Jordan, N. C, Hanich, L. B. & Uberti, L. B. (2003). Mathematical Thinking and Learning Difficulties. I A. J. Baroody & A. Dowker (red.), *The development of arithmetic concepts and skills – constructing adaptive expertise* (s. 359-383). New Jersey: Lawrence Erlbaum Associates, Inc.

Jordan, N. C., Fuchs, L. S. & Dyson, N. (2015). Early number competencies and mathematical learning – individual variation, screening, and intervention. I. R. C. Kadosh & A. Dowker (red.) *The Oxford Handbook of Numerical Cognition* (s. 1079-1098). United Kingdom: Oxford university press.

Jordan, N. C., Kaplan, D., Ramineni, C. & Locuniak, M. N. (2008). Development of number combination skill in the early school years: when do fingers help? *Developmental Science*, 11:5, 662-668, doi: 10.1111/j.1467-7687.2008.00715.x.

Kawaguchi, D. (2011). Actual age at school entry, educational outcomes, and earnings. *Journal of The Japanese and International Economies,* 25:2, 64-80, doi: 10.1016/j.jjie.2009.02.002.

Klausen, T. G. & Reikerås, E. (2016). *Regnefaktaprøven*. Stavanger: Lesesenteret

Kleven, T. A. (2002a). Ikke-eksperimentelle design. I T. Lund. (red.), *Innføring i forskningsmetodologi* (s. 265- 286). Bergen: Fagbokforlaget

Kleven, T. A. (2002b). Begrepsoperasjonalisering. I T. Lund. (red.), *Innføring i forskningsmetodologi* (s. 141-183). Bergen: Fagbokforlaget

Koechlin, E., Dehaene, S. & Mehler J. (1997). Numerical Transformations in Five-month-old Human Infants. *Mathematical Cognition*, 3:2, 89-104, doi: 10.1080/135467997387425

Levine, S. C., Jordan, N. C. & Huttenlocher, N. (1992). Development of Calculation Abilities in Young Children. *Journal of Experimental Child Psychology*, 53:1, 72-103, doi: 10.1016/S0022-0965(05)80005-0.

Lund, T. (2002a). Metodologiske prinsipper og referanserammer. I T. Lund. (red.), *Innføring i forskningsmetodologi* (s. 79-123). Bergen: Fagbokforlaget

Lund, T. (2002b). Generaliseringsproblematikk. I T. Lund. (red.), *Innføring i forskningsmetodologi* (s. 125-140). Bergen: Fagbokforlaget

Mazzocco, M. M. M. (2007). Defining and Differentiating Mathematical Learning Disabilities and Difficulties. I D. B. Berch & M. M. M. Mazzocco (red.). *Why is math* *so hard for some children? – The nature and origins of mathematical learning disabilities and disabilities* (s. 29-47). Baltimore: Paul H. Brooks Publishing Co.

Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school age years. *Annals of Dyslexia*, 53:1, 218-253, doi: 10.1007/s11881-003-0011-7

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., Schulte-Körne, G. (2014). Specific Learning Disorder: Prevalence and Gender Differences. *PLoS ONE* 9:7, 1-8 e103537, doi:10.1371/journal.pone.0103537.

Moore, D. S. & Cocas L. A. (2006). Perception Precedes Computation – Can Familiarity Preferences Explain Apparent Calculation by Human Babies? *Developmental Psychology*, 42:4, 666–678, doi: 10.1037/0012-1649.42.4.666.

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities*, 44, 472–488, doi: 10.1177/0022219411414010.

Nelson, G. & Powell, S. R. (2017). A Systematic Review of Longitudinal Studies of Mathematics Difficulty. *Journal of Learning Disabilities*, 51:6 523 –539, doi: 10.1177/0022219417714773.

Noël, M-P. (2015). When Number Processing and Calculation is not your Cup of Tea. I. R. C. Kadosh & A. Dowker (red.) *The Oxford Handbook of Numerical Cognition* (s. 635-646). United Kingdom: Oxford university press.

Nortvedt, G. A. & Pettersen, A. (2016). Matematikk. I M. Kjærnsli & F. Jensen (red.). *Stø kurs – Norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015*. Oslo: Universitetsforlaget.

Nortvedt, G. A. & Vogt, G. O. (2012). Når matematikk blir vanskelig – matematikkvansker i elev- og undervisningsperspektiv. I E. Befring & R. Tangen (red.). *Spesialpedagogikk* (s. 370-384). Oslo: Cappelen Damm Akademiske.

NOU 2019:3 (2019). *Nye sjanser – bedre læring – Kjønnsforskjeller i skoleprestasjoner og utdanningsløpet*. Hentet 05.05.19, fra https://www.regjeringen.no.

Olsen, R.V. & Bjørnsson, J.K. (2018). Fødselsmåned og skoleprestasjoner. I R. V. Olsen & J. K. Bjørnsson.(red.), *Tjue år med TIMSS og PISA i Norge: trender og nye analyser* (s. 76-93). Oslo: Universitetsforlaget.

Ostad, S. A. (1997). Developmental differences in addition strategies - A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology,* 67:3 , 345– 357, doi: 10.1111/j.2044-8279.1997.tb01249.x.

Ostad, S. A. (1999). *Elever med matematikkvansker – Studier av kunnskapsutviklingen i strategisk perspektiv.* Oslo: Unipub forlag

Ostad, S. A. (2013). *Strategier, strategiobservasjon og strategiopplæring med fokus på elever med matematikkvansker*. Trondheim: Læreboka Forlag

Pallant, J. (2016). *SPSS survival anual – a step by step guide to data analysis using IBM SPSS, 6th edition.* London: Open University Press.

Price, G. R. & Ansari, D. (2013) Dyscalculia: Characteristics, Causes, and Treatments. *Numeracy*: 6:1, 1-16, doi: 10.5038/1936-4660.6.1.2.

Raghubar, K. P & Barnes, M. A. (2017). Early numeracy skills in preschool-aged children - a review of neurocognitive findings and implications for assessment and intervention. *The Clinical Neuropsychologist,* 31:2, 329-351, doi: 10.1080/13854046.2016.1259387

Raven, J. C. (1998). The Raven's Progressive Matrices. Oxford Psychologists Press Oxford

Reikerås, E. & Salomonsen, T. (2019) Weak mathematical skills at an early age: persistent or temporary? Children with weak mathematical skills and their development from toddlers to preschoolers, *Early Child Development and Care*, 189:4, 670-682, doi: 10.1080/03004430.2017.1337753.

Riley, M. S. & Greeno, J. G. (1988). Developmental Analysis of Understanding Language about Quantities and of Solving Problems. *Cognition and Instruction,* 5:1, 49-101, doi: 10.1207/s1532690xci0501_2

Riley, M. S., Greeno, J. G. & Heller, J. I. (1983). Development of children`s problem-solving ability in arithmetic. I H. Ginsbrug (red.) *The development of mathematical thinking* (s. 153-196). Orlando, FL: Academic Press, Inc.

Rubinsten, O., & Tannock, R. (2010). Mathematics anxiety in children with developmental dyscalculia. *Behavioral and Brain Functions*, 46:6, 1–13. doi:10.1186/1744-9081-6-46

Russel, R. L. & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction, 1*:2, 217-244, doi: 10.1207/s1532690xci0102_3.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Casual Inference*. Boston: Houghton Mifflin Company.

Shen, C., Vasilyeva, M. & Laski, E. V. (2016). Here, but not there – Cross-national variability of gender effects in arithmetic. *Journal of Experimental Child Psychology*, 146, 50-65, doi: 10.1016/j.jecp.2016.01.016.

Siegler, R. S. & Braithwaite, D. W (2017). Numerical Development. *Annual Review of Psychology*, 68: 187-213, doi 10.1146/annurev-psych-010416-04410.

Siegler, R. S. & Jenkins, E. (1989). *How Children Discover New Strategies*. New Jersey: Lawrence Erlbaum Associates, Inc.

Siegler, R. S. & Shrager, J. (1984). Strategy Choices in Addition and Subtraction – How Do Children Know What to Do? I C. Sophian (red.). *Origins of Cognitive Skills* (s. 229-293). New Jersey: Lawrence Erlbaum Associates, Inc.

Siegler, R. S. (1987). The perils of averaging data over strategies - An example from children's addition. *Journal of Experimental Psychology: General, 116*:3, 250-264, doi: 10.1037/0096-3445.116.3.250.

Siegler, R. S. (1988). Individual Differences in Strategy Choices: Good Students, Not-So-Good Students, and Perfectionists. *Child Development*, 59:4, 833-851, doi: 10.2307/1130252.

Siegler, R. S. (1989). Hazards of mental chronometry: An example from children's subtraction. *Journal of Educational Psychology, 81*:4, 497-506. doi:10.1037/0022-0663.81.4.497

Siegler, R. S. (1991). Strategy choice and strategy discovery. Learning and Instruction, 1:1, 89-102, doi: 10.1016/0959-4752(91)90020-9.

Siegler, R. S. (1996). *Emerging Minds – The Process of Change in Children`s Thinking*. Oxford: Oxford University Press.

Simon, T.J., Hespos, S.J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, 10:2, 253–269, doi: 10.1016/0885-2014(95)90011-X.

Solli, I. F. (2017). Left behind by birth month. *Education Economics*, 25:4, 323-346, doi: 10.1080/09645292.2017.1287881.

Stock, P., Desoete, A., & Roeyers, H. (2010). Detecting children with arithmetic disabilities from kindergarten: Evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities*, 43 (3), 250-268, doi:10.1177/0022219409345011.

Stoltenberg, C. (2017, 10. februar). Vi har skapt et nytt kjønnsgap før vi har rukket å kvitte oss med det gamle. *Morgenbladet*. Hentet 05.05.19, fra https://morgenbladet.no.

Strøm, B. 2004. Student Achievement and Birthday Effects. *Norwegian University for Science and Technology working paper*.

Sæther, S. A., Peterson, T. & Matin, V. (2017). The Relative Age Effect, Height and Weight Characteristics among Lower and Upper Secondary School Athletes in Norway and Sweden. *Sports*, 5:4, 92, doi: 10.3390/sports5040092.

Tabachnick, B. G. & Fidell, L. S. (2013). *Using Multivariate Statistics.* Harlow: Pearson Education, Ltd.

Thevenot, C. & Barrouillet, P. (2015). Arithmetic Word Problem Solving and Mental Representations. I. R. C. Kadosh & A. Dowker (red.) *The Oxford Handbook of Numerical Cognition* (s. 158-179). United Kingdom: Oxford university press.

Thoren, K., Heinig, E. & Brunner, M. (2016). Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time and Classes. *Front. Psychol*. 7:679, . doi: 10.3389/fpsyg.2016.00679.

Universitetet i Oslo (2017, 17. januar). Om TIMSS. Hentet, 03.05.2019, fra https://www.uv.uio.no/ils/forskning/prosjekt-sider/timss-norge/TIMSS/om-timss/.

Utdanningsdirektoratet (2013). *Læreplan i matematikk fellesfag* (MAT1-04). Hentet 30.03.19, fra, https://www.udir.no/kl06/MAT1-04.

Utdanningsdirektoratet (2017, 04. mai). Hva måler nasjonale prøver i regning? Hentet, 03.05.19, fra https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/mestringsbeskrivelser-og-hva-provene-maler/hva-maler-nasjonal-prove-i-regning/.

Verschaffel, L. & De Corte, E. (1997). Word problems – A vehicle for promoting authentic mathematical understanding and problem solving in the primary school. I T. Nunes & P. Bryant (red.). *Learning and Teaching Mathematics – An International Perspective* (s. 69-97). East Sussex: Psychology Press Ltd.

Verschaffel, L., Depaepe, F. & Van Dooren, W. (2015). Individual Differences in World Problem Solving. I. R. C. Kadosh & A. Dowker (red.) *The Oxford Handbook of Numerical Cognition* (s. 953-974). United Kingdom: Oxford university press.

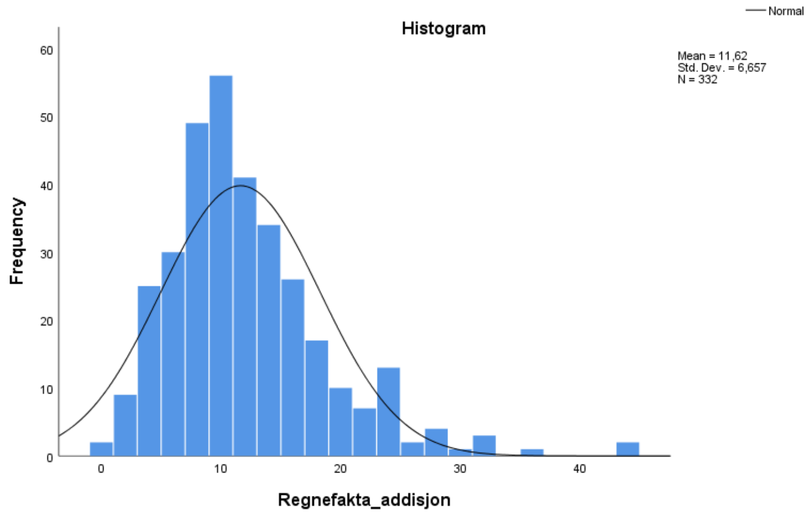
Wakeley, A., Rivera, S., & Langer, J. (2000). Can young infants add and subtract? *Child Development*, 71: 6, 1525–1534, doi: 10.1111/1467-8624.00244.

Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maelender, A. (2003). Wechsler Intelligence Scale for Children (Version 4) [Assessment instrument]. San Antonio: TX: Pearson.

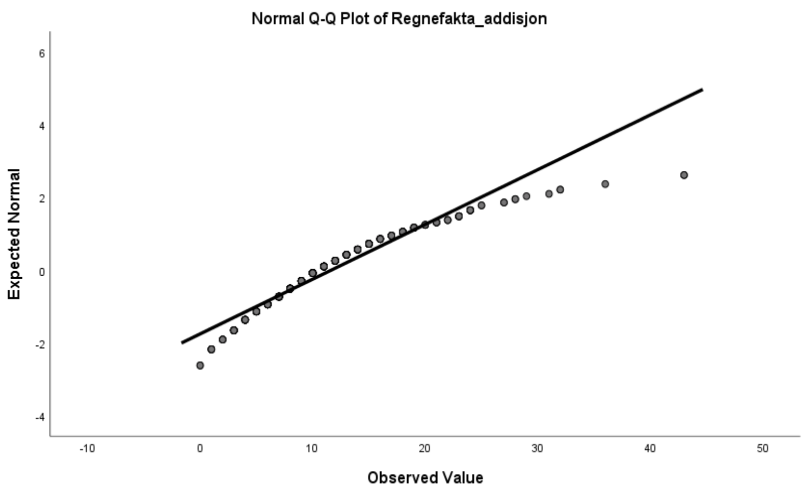
World Health Organization (2018). ICD-11: *International Classification of Diseases 11th Revision*. Hentet 16.01.2019, fra https://icd.who.int/browse11/l-m/en.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358: 6389, 749–750.

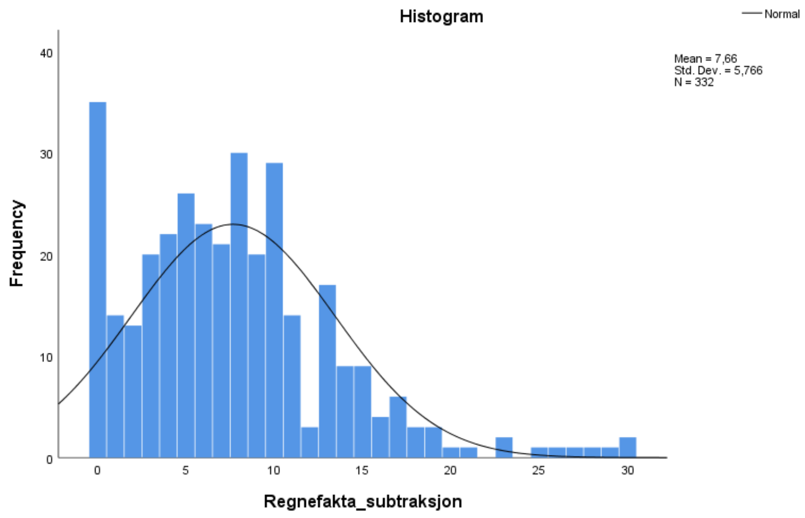
Vedlegg



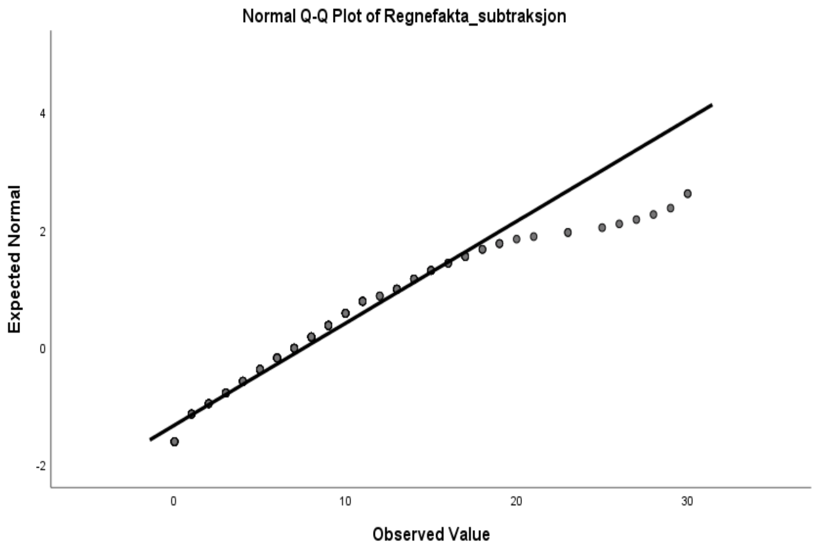
Vedlegg 1 Histogram av Regnefaktaprøven addisjon



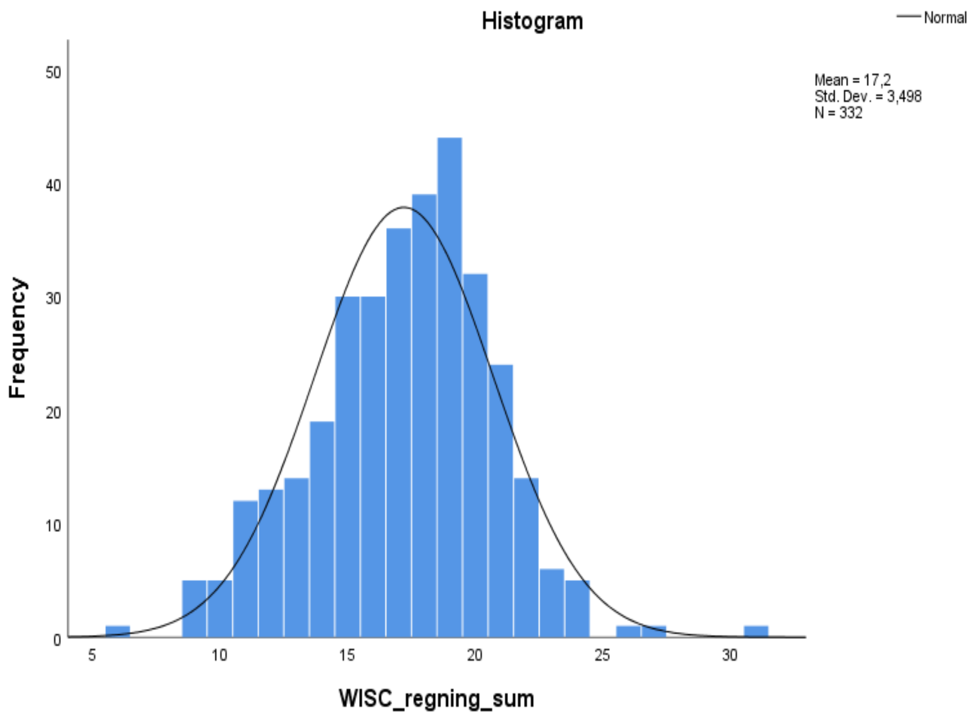
Vedlegg 2 Q-Q plot av Regnefaktaprøven addisjon



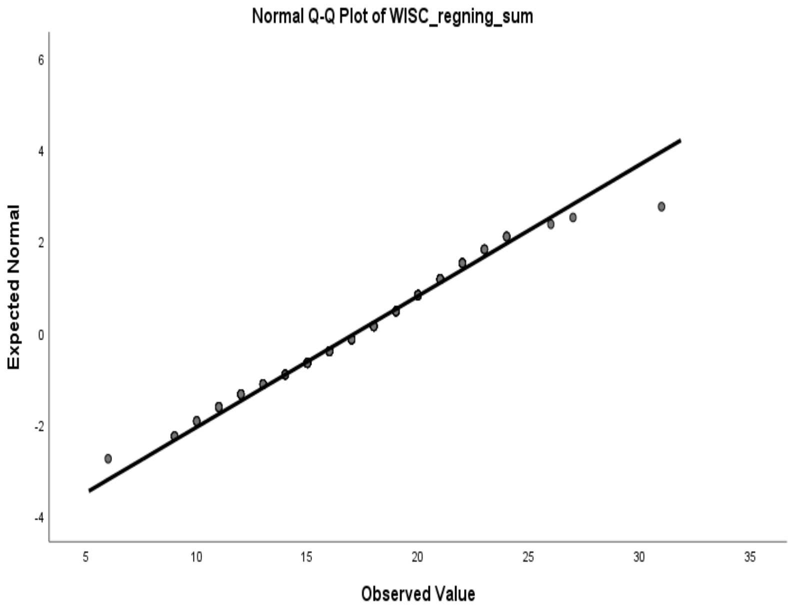
Vedlegg 3 Histogram av Regnefaktaprøven subtraksjon



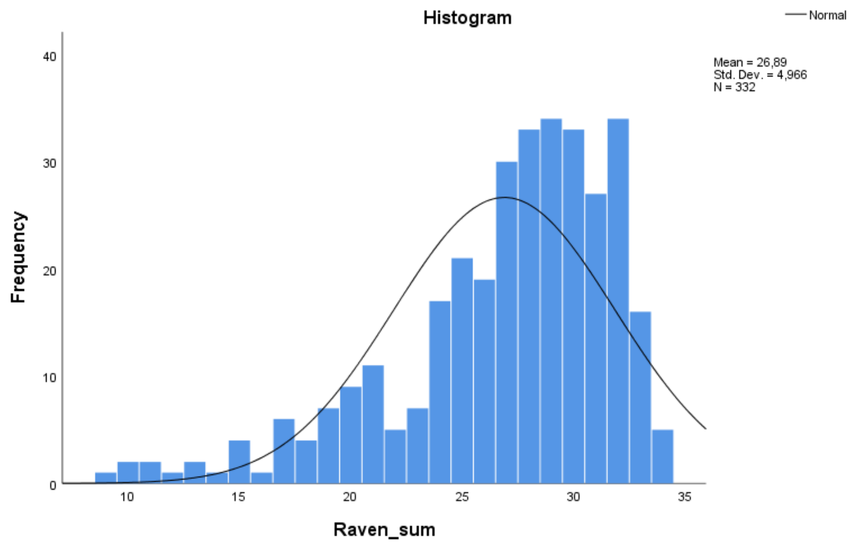
Vedlegg 4 Q-Q plot av Regnefaktaprøven subtraksjon



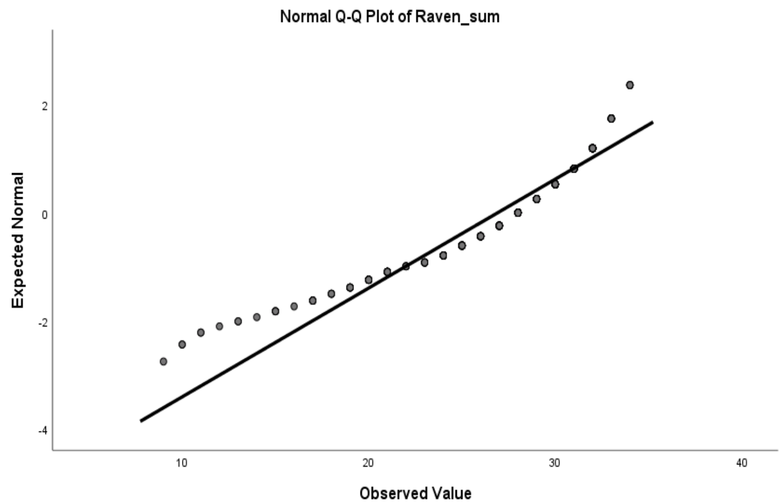
Vedlegg 5 Histogram av WISC-IV regning



Vedlegg 6 Q-Q plot av WISC-IV regning



Vedlegg 7 Histogram av Raven



Vedlegg 8 Q-Q plot Raven