

Kjønnsforskjeller i matematikk

*En kvantitativ undersøkelse av lavtpresterende
førsteklassinger*

Terje Ulv Throndsen



Masteroppgave i spesialpedagogikk
UNIVERSITETET I OSLO

01.06.17

Kjønnsforskjeller i matematikk

En kvantitativ undersøkelse av lavtpresterende førsteklasinger

© Terje Ulv Throndsen

2017

Kjønnsforskjeller i matematikk. En kvantitativ undersøkelse av lavtpresterende
førsteklassinger

Terje Ulv Throndsen

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

Trykk: Reprosentralen, Universitetet i Oslo

IV

Sammendrag

Bakgrunn og formål

I denne undersøkelsen er kjønnsforskjeller i matematikkferdigheter blant lavtpresterende førsteklasinger blitt undersøkt i et utvalg på 120 barn. Formålet med denne oppgaven er å belyse et tema som på mange måter er omstridt, i et felt der det fremdeles er behov for mer forskning for bedre kunne bistå de som strever (Halpern et al., 2007; Price & Ansari, 2013). Kjønnsforskjeller er et kontroversielt tema. Spesielt i matematikkferdigheter ser det ut til å være en standhaftig stereotypi om gutters overlegenhet, til tross for at empirien på i svært liten grad entydig underbygger dette (Furnham, Reeves & Budhani, 2002). Uavhengig av kjønn er spesifikke matematikkvansker sannsynligvis like utbredt som dysleksi, men matematikkvansker er mindre kjent og det foreligger langt mindre forskning på feltet (Geary, 2013b). De negative konsekvensene av lave ferdigheter i matematikk er ofte svært alvorlige for de som er rammet (Duncan et al., 2007), og det er per 2017 fremdeles et betydelig behov for mer detaljert kunnskap om hvordan matematikkvansker opptrer, hvordan matematikkvansker best kan forebygges og hvilke tiltak som har best effekt for å hjelpe de som er rammet. Sentralt i oppgaven er en teori som postulerer at tidlige matematikkferdigheter består av noen biologiske primære og biologiske sekundære kvantitative ferdigheter (Geary, 1995a; 2000). Basert på denne teorien og empiri om sammenhengen mellom IQ og matematikkferdigheter, eksekutive funksjoner og matematikkferdigheter, samt kjønnsforskjeller i matematikkvansker, er disse forskningsspørsmålene utarbeidet:

- 1. Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?*
- 2. Er det kjønnsforskjeller når vi kontrollerer for nonverbal IQ og eksekutive funksjoner.*

Metode

Oppgaven er skrevet i tilknytning til et større forskningsprosjekt, «The effects of mathematical interventions programs for students performing low in mathematics», og all data som er analysert er hentet fra dette prosjektet. Forskningsspørsmålene besvares ved bruk et kvantitativt ikke-eksperimentelt design. Utvalget består av 120 elever som på forhånd er blitt identifisert som lavtpresterende i matematikk. Videre er de blitt testet med ulike matematikktester, i tillegg til blant annet en nonverbal IQ-test og en test av eksekutive funksjoner. Enkelte av resultatene fra ulike deltester i matematikk er undersøkt for kjønnsforskjeller gjennom komparative analyser.

Analyser

Resultatene fra de ulike testene er først analysert for kjønnsforskjeller gjennom en to-halet t-test, deretter er det foretatt multiple regresjonsanalyser hvor IQ og eksekutive funksjoner er kontrollert for. Samvariasjonen i de ulike variablene er analysert ved hjelp av korrelasjonsanalyser. Alle analyser i denne oppgaven er foretatt i det statistiske analyseprogrammet SPSS.

Resultater

Resultatene viser at det er en signifikant forskjell på gutter og jenters prestasjoner på testen som måler den biologisk primære kvantitative evnen *approximate number system (ANS)*. Signifikansen er på .05-nivå, med en effektstørrelse på $d = .47$. Når IQ og eksekutive funksjoner er kontrollert for forklarer kjønn 5.5% av variasjonen i denne variabelen. Dette er et funn som ikke finner støtte i tidligere empiri. I de tre andre variablene som er analysert, Kjennskap til tall, Tekstoppgaver og Aritmetikk fremkommer det ikke kjønnsforskjeller i denne undersøkelsen. Disse resultatene er delvis i tråd med tidligere forskning på kjønnsforskjeller i matematikk.

Forord

Først og fremst; Mange takk til min hovedveileder Monica Melby-Lervåg, og min bi-veileder Anita Lopez-Pedersen, dere har på hvert deres vis gjort et uvurderlig stykke arbeid i å lyse opp veien for meg. Takk til min kone for støtte og oppmuntring, det har vært godt å ha en psykiater i ryggen. Og til mine medstudenter på ISP, dere vet hvem dere er, takk for laget.

Denne oppgaven har tatt tid, og det er sikkert mange ting som kunne vært gjort annerledes, men nå er siste punktum satt. Herfra er det bare resten av livet som gjenstår. En klokere mann enn meg får siste ordet her.

"O, grubl ei over hva der ei kan endres ..."

- H. Ibsen

Innholdsfortegnelse

1. Innledning	1
1.1 Bakgrunn og formål	1
1.2 Problemstilling.....	3
1.3 Begrepsavklaring.....	3
1.4 Avgrensning.....	4
1.5 Oppgavens oppbygning	5
2. Teoretisk og empirisk bakgrunn	6
2.1 Matematikkferdigheter.....	6
2.2 Utvikling av matematikkferdigheter	8
2.2.1 Matematikkvansker.....	8
2.2.2 Biologisk primære kvantitative evner	11
2.2.3 Biologisk sekundære kvantitative evner.....	12
2.2.4 Prediksjon av matematikkferdigheter	13
2.2.6 Kjennskap til tall	15
2.2.7 IQ og matematikkferdigheter	16
2.2.8 Eksekutive funksjoner og matematikkferdigheter	17
2.3 Kjønnsforskjeller i matematikk.....	19
2.3.1 Kjønnsforskjeller blant barn med matematikkvansker	23
2.4 Oppsummering av teori og empiri.....	24
3. Metode	26
3.1 Design	26
3.2 Utvalg	27
3.3 Datainnsamling.....	27
3.4 Testene	28
3.4.1 Matematikktester	28
3.4.2 Raven	30
3.4.3 Tower of London.....	30
3.5 Validitet og reliabilitet	31
3.5.1 Statistisk validitet.....	31
3.5.2 Indre validitet	32
3.5.3 Begrepsvaliditet	33
3.5.4 Ytre validitet	33
3.6 Analyser.....	34
3.7 Etiske hensyn	35
4. Resultater	36
4.1 Deskriptiv analyse	37
4.2 Korrelasjonsanalyse av måleinstrumentene	38
4.3 Variablenes fordeling og reliabilitet.....	39
4.3.1 Vurdering av variabelen ANS	40
4.3.2 Vurdering av variabelen Tekstoppgaver	40
4.3.3 Vurdering av variabelen Kjennskap til tall	41
4.3.4 Vurdering av Aritmetikk	42
4.4 Vurdering av kontrollvariablene	42
4.4.1 IQ	43
4.4.2 Eksekutive funksjoner.....	43
4.5 Deskriptiv analyse av variablene fordelt på kjønn	44
4.5.1 Kjønnsforskjeller i biologisk primære kvantitative evner	45
4.5.2 Kjønnsforskjeller i Tekstoppgaver	45
4.5.3 Kjønnsforskjeller i Kjennskap til tall.....	45
4.5.4 Kjønnsforskjeller i Aritmetikk.....	46

4.5.5 Kjønnsforskjeller i biologisk sekundære kvantitative evner.....	46
4.6 Korrelasjonsanalyse	46
4.7 Regresjonsanalyse	47
4.7.1 Regresjonsanalyse av ANS	48
4.7.2 Regresjonsanalyse av Tekstoppgaver	48
4.7.3 Regresjonsanalyse av Kjennskap til tall	49
4.7.4 Regresjonsanalyse av Aritmetikk	49
4.8 Oppsummering av analyser og funn.....	50
5. Drøfting av resultater	52
5.1 Undersøkelsens validitet og reliabilitet	52
5.1.1 Statistisk validitet.....	52
5.1.2 Indre validitet	54
5.1.3 Begrepsvaliditet	56
5.1.4 Ytre validitet	57
5.2 Resultater i lys av tidligere empiri	58
5.2.1 Kjønnsforskjeller i biologisk primære kvantitative evner	58
5.2.2 Kjønnsforskjeller i Tekstoppgaver	59
5.2.3 Kjønnsforskjeller i Kjennskap til tall	59
5.2.4 Kjønnsforskjeller i Aritmetikk	60
5.2.5 Kjønnsforskjeller i biologisk sekundære kvantitative evner.....	60
5.3 Oppsummering	61
5.3.1 Pedagogiske konsekvenser	61
5.3.2 Fremtidige undersøkelser	61
Litteraturliste	63

Liste over figurer og tabeller

Figur 3.1 Eksempel på oppgave fra Raven.....	30
Figur 3.2 Eksempel på oppgave fra Tower Of London.....	31
Tabell 3.1 Måleinstrumentene for matematikkferdigheter.....	34
Tabell 3.2 Måleinstrumentene for kontrollvariablene.....	35
Tabell 4.1 Måleinstrumentene.....	37
Tabell 4.2 Korrelasjon mellom måleinstrumentene.....	38
Tabell 4.3 Variabler for matematikkferdigheter.....	39
Figur 4.1 Histogram som viser fordeling på «ANS».....	40
Figur 4.2 Histogram som viser fordeling på «Tekstoppgaver».....	41
Figur 4.3 Histogram som viser fordeling på «Kjennskap til tall».....	41
Figur 4.4 Histogram som viser fordeling på «Aritmetikk».....	42
Figur 4.5 Histogram for fordeling på Raven.....	43
Figur 4.6 Histogram for fordeling på Tower Of London.....	44
Tabell 4.4 Kjønnsforskjeller i variablene med effektstørrelse uttrykket i Cohens d.....	44
Tabell 4.5 Korrelasjon mellom variablene uttrykket i Pearson r.....	47
Tabell 4.6 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på ANS. N=120.....	48
Tabell 4.7 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Tekstoppgaver. N=120.....	48
Tabell 4.8 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Kjennskap til tall. N=120.....	49
Tabell 4.9 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Aritmetikk. N=120.....	49
Figur 7. Sidestilte histogram med fordeling på ANS fordelt på kjønn.....	50

1. Innledning

1.1 Bakgrunn og formål

En rekke undersøkelser viser at barn som har lave matematikkrelaterte ferdigheter ved skolestart som regel også vil ligge etter sine medelever i matematikk når de går ut av skolen (Duncan et al., 2007; Geary, 2013a). En longitudinell undersøkelse der barns tallrelaterte kompetanse ble målt ved seks ulike tidspunkt fra førskolen og frem til 3. klasse viste at sammenhengen mellom tidlig tallferdigheter og senere matematikkprestasjoner var sterk og signifikant gjennom hele perioden (Jordan, Kaplan, Ramineni, & Locuniak, 2009). Naturlig nok vil matematiske ferdigheter ved endt skolegang påvirke muligheter en har i for eksempel arbeidsmarkedet og andre områder der en må forholde seg til matematikk også i voksen alder (Geary, 2013a). Videre vet vi at elever som ikke oppnår funksjonelle ferdigheter i matematikk har økt risiko for å bli stående utenfor arbeidslivet og for å ikke kunne delta i demokratiske samfunnsprosesser på en likeverdig måte (Price & Ansari, 2013).

Det er av ovenfornevnte årsaker derfor bekymringsfullt at hele 20,3 prosent av elevene på linjen «Studieforberedende» fikk karakteren 1 på eksamen i matematikk nivå 1P, mens det tilsvarende tallet for matematikkeksamen 1P på yrkesfag samlet sett var 14,5 prosent i 2016 (Utdanningsdirektoratet, 2017). Dette er problematisk på flere nivåer, blant annet trekker kunnskapsminister Torbjørn Røe Isaksen (H) frem risikoen for å i fremtiden ikke få rekruttert nok folk til viktige yrker (Sund, Asbjørnsen, & Zondag, 2015). Kanskje er det først og fremst mest problematisk for elevene selv. I undersøkelser fra England viser resultatene at individer som har dårlig tallforståelse i langt større grad enn befolkningen for øvrig sliter med arbeidsløshet, fysiske og mentale lidelser, samt kriminalitet (Bynner & Parsons, 1997; Parsons & Bynner, 2006). Til tross for at man anslår at det er like mange som sliter med matematikkvansker som dysleksi, finnes det sammelignet med lese- og skrivevansker mindre forskningsbasert kunnskap om matematikkvansker (Geary, 2013b).

Når det gjelder kjønnsforskjeller i matematikkferdigheter er dette et omdiskutert tema (Geary, Saults, Liu, & Hoard, 2000; Halpern et al., 2007). Resultater fra TIMSS (Trends in International Mathematics and Science Study) viser at det på 4.- og 5.trinn ikke er signifikante forskjeller mellom gutter og jenter i Norge (Bergem, 2016), samtidig vet vi at jenter i snitt får bedre eksamenskarakterer i matematikk på videregående, mens det blant guttene er større spredning (Utdanningsdirektoratet, 2017). Vi vet også at gutter er i stort flertall på retninger i høyere utdanning med avansert matematikk, som for eksempel enkelte ingeniørretninger og matematisk-naturvitenskapelige fag (NSD, 2017). Denne typen funn og data inviterer til forskning som studerer kjønnsforskjeller i matematikkprestasjoner blant barn generelt og blant de som er lavtpresterende i matematikk spesielt. Dette vil blant annet kunne gi kunnskap om hvorvidt det er kjønnsforskjeller i matematikkprestasjoner allerede på første trinn eller om forskjellene først og fremst skapes i skolen og ulikhetene kommer senere. Eventuelle kjønnsforskjeller vil kunne være viktige å kjenne til når det skal tilrettelegges for god læring for både jenter og gutter.

Matematikkvansker er et felt som det fremdeles er forsket lite på, tatt i betraktning utbredelsen av denne typen vansker og de potensielt store negative konsekvensene for de som er rammet. Samtidig er det en vedvarende stereotypi om gutters overlegenhet i matematikk. Til sammen gjør dette gjør det interessant å se nærmere på om jeg kan finne kjønnsforskjeller i matematikk blant et utvalg på 120 førsteklasinger. Det er fremdeles et betydelig behov for mer detaljert kunnskap om hvordan matematikkvansker opptrer, hvordan matematikkvansker best kan forebygges og hvilke tiltak som gir mest effekt for å hjelpe de som er rammet. Denne oppgaven vil redegjøre for sentrale teorier og empiriske funn knyttet til utviklingen av matematikkferdigheter i seksårsalder og hva som kan forklare mulige kjønnsforskjeller.

1.2 Problemstilling

Tema:

Kjønnsforskjeller i matematikk.

Forskningsspørsmål:

1. Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?
2. Er det kjønnsforskjeller når vi kontrollerer for nonverbal IQ og eksekutive funksjoner.

1.3 Begrepsavklaring

I denne studien brukes begrepet *lavtpresterende* om de 120 elevene som danner grunnlaget for datamaterialet som analyseres. Et utvalg på 369 elever ble testet med en matematikktest og disse 120 var de 32,52% med lavest skåre. Å definere de 32% svakeste som lavtpresterende er i tråd med andre undersøkelser på området der termen *learning disabled in mathematics* eller *mathematical disabilities (MD)* brukes på de 30% svakeste (Geary, Hoard, & Hamson, 1999) eller de 35% svakeste (Geary, Hamson, & Hoard, 2000). Det legges likevel ikke til grunn i denne studien at alle deltakerene nødvendigvis er *learning disabled in mathematics* eller har *mathematical disabilities*. Det er mer dekkende å si at de har utfordringer i tilegnelsen av matematikkferdigheter, og dermed lavtpresterende i denne sammenheng.

I den engelskspråklige forskningslitteraturen brukes både *sex* og *gender* om kjønn, *sex* omhandler det rent biologiske, mens *gender* referer til en sosial rolle (Halpern et al., 2007). Da denne distinksjonen ikke forekommer på norsk brukes i denne oppgaven utelukkende begrepet kjønnsforskjeller, både der litteraturen på originalspråket omtaler *sex differences* og der det er tale om *gender differences*.

Når det gjelder uttrykk som *biologisk*, *iboende* eller *medfødt*, er dette en henvisning til et potensial som er «klar» for utvikling, forutsatt et støttende miljø. Mennesker er både biologiske og sosiale vesener formet av et komplekst samspill mellom biologi og miljø. At noen har *medfødte* egenskaper betyr ikke uungåelig eller uforanderlig (Halpern et al., 2007), og det er altså ikke i en deterministisk betydning disse begrepene brukes i denne sammenheng.

1.4 Avgrensning

I denne studien er fire sentrale matematikkferdigheter analysert for kjønnsforskjeller. Valget av de fire variablene er basert på tidligere teori og empiri om matematikkferdigheter og mulige kjønnsforskjeller knyttet til disse. IQ og eksekutive funksjoner er også sentrale ferdigheter i matematikkprestasjoner, og det er derfor foretatt hierarkiske multiple regresjonsanalyser for å se om eventuelle kjønnsforskjeller fremdeles finnes etter at variablene *Non-verbal IQ* og *eksekutive funksjoner* er kontrollert for.

Oppgaven er skrevet i tilknytning til et større forskningsprosjekt, «*The effects of mathematical interventions programs for students performing low in mathematics*», som utføres av stipendiat Anita Lopez-Pedersen i samarbeid med forskergruppen *Child Language and Learning*. Gjennom denne undersøkelsen har jeg fått tilgang til resultater fra i alt tolv del-tester, og det kunne vært interessant å se på langt flere sammenhenger. Men med de gitte rammebetingelser som ligger i en masteroppgave har jeg sett det hensiktsmessig å begrense meg til de nevnte analysene. I tillegg vil man ved å gjøre for mange signifikanstester på samme materialet øke sannsynligheten for å gjøre Type II-feil. Eksempelvis vil det om man utfører 25 signifikanstester være 65% sannsynlighet for å oppdage signifikans på slump.

1.5 Oppgavens oppbygning

I kapittel 1 er det blitt redegjort for bakgrunn for valg av tema, oppgavens formål og problemstillingen er presentert. I kapittel 1 finnes også en begrepsavklaring, avgrensning og denne beskrivelsen av oppgavens oppbygning.

I kapittel 2 gjøres det rede for relevant teori og tidligere forskning med empiriske funn knyttet til utvikling av matematikkferdigheter, matematikkvansker, samt hvilken rolle IQ, eksekutive funksjoner og kjønn spiller i dette. Denne redegjørelsen legger grunnlag for å kunne besvare problemstillingen, men er også bakgrunnen for den.

Beskrivelse av studiens design, utvalg og datainnsamlingen vil finnes i kapittel 3. Der vil det også være en gjennomgang av de ulike testene i «*The effects of mathematical interventions programs for students performing low in mathematics*», og en mer utførlig beskrivelsen av de testene som er benyttet for analysene i min studie. Videre vil jeg komme inn på validitet og reliabilitet, og en kort gjennomgang av kommende analyser før kapittelet avsluttes med etiske hensyn.

Kapittel 4 vil presentere resultatene, først de deskriptive analysene og korrelasjonsanalyser, deretter regresjonsanalyser.

Avslutningsvis vil jeg i kapittel 5 drøfte mine funn opp mot teorien og de empiriske funn som ble presentert i kapittel 2, i tillegg til å drøfte analysene og funnen i lys av Cook & Campbells validitetssystem.

2. Teoretisk og empirisk bakgrunn

Først i dette kapitlet vil jeg fokusere på rådende teori og empiri om matematikkferdigheter, hvilke evner og egenskaper som ligger til grunn for matematikkferdigheter og utviklingen av disse. Deretter følger en redegjørelse for definisjoner, kjennetegn og teorier om matematikkvansker. Videre vil jeg presentere en teori om biologisk primære kvantitative evner og biologisk sekundære kvantitative evner, denne teorien er i stor grad bakgrunnen for valg av analyser i denne undersøkelsen. Hvilke evner og ferdigheter som predikerer senere matematikkprestasjoner står sentralt i forskning på matematikkvansker, og en gjennomgang av denne forskningen blir redegjort for. Deretter følger en redegjørelse for betydningen av kjennskap til tall, før en gjennomgang av sammenhengen mellom IQ og matematikkferdigheter, og eksekutive funksjoner og matematikkferdigheter. Siste del av dette kapitlet vies så kjønnsforskjeller og hvordan disse vil kunne spille inn på matematikkferdigheter og matematikkvansker.

2.1 Matematikkferdigheter

Algebra, geometri, måling, nummerering, regning, algoritmisk beregning og ulike former for problemløsning er alle eksempler på områder i matematikken (Fuchs et al., 2006).

Matematikk er dermed et bredt og sammensatt felt. På 1. trinn er det først og fremst tall og telling, aritmetikk og enkle tekstoppgaver som det er forventet at elevene skal mestre (Utdanningsdirektoratet, 2013). Disse ferdighetene kalles ofte for lavere ordens ferdigheter, i kontrast til algebra, brøk og lignende som går inn under høyere ordens ferdigheter. Det finnes ikke spesifiserte kompetansemål for 1. trinn, men etter 2. trinn sier er kompetansemålene at eleven blant annet skal (a) kunne telle til 100, (b) bruke tallinja til beregninger og vise tallstørrelser, (c) utvikle, bruke og samtale variert om regnestrategier for addisjon og subtraksjon av tosifrede tall, samt å vurdere hvor rimelige svarene er, og (d) doble og halvere (Utdanningsdirektoratet, 2013).

Flere undersøkelser har sett på hvilke evner som ligger til grunn for matematikkferdigheter. Blant annet har Aunio & Räsänen (2015) i en gjennomgang av longitudinelle undersøkelser analysert resultatene fra disse for å se hvilke ferdigheter som kan betegnes som

kjerneferdigheter, og som dermed er av størst betydning for prestasjoner i matematikk hos barn i alderen 5 til 8 år. I sin oppsummering av resultatene fra de ulike undersøkelsene mener Aunio & Räsänen at de viktigste ferdighetene i utøvelse av matematikk er disse fire; (1) symbolsk og ikke-symbolsk tallforståelse, (2) forståelse av matematiske relasjoner, som plassverdisystemet og titallsystemet, (3) telleferdigheter og (4) grunnleggende regneferdigheter (Aunio & Räsänen, 2015). Også De Smedt, Noël, Gilmore & Ansari (2013) viser i sin gjennomgang av hjerne- og adferdsforskning at evnen til å prosessere numerisk mengde legger grunnlag for matematikkferdigheter. Tilsvarende de funn som Aunio & Räsänen (2015) presenterer, finner også De Smedt et al. (2013) at både symbolsk og ikke-symbolsk mengdeforståelse har betydning for barns matematikkferdigheter. Et av de mest robuste funn i gjennomgang av denne forskningen viser at de barna som best er i stand til å både fastslå hvilket av to tallsymboler som er størst, og i hvilken av to grupperinger av objekter det er flest, også gjør det bedre i matematikk (De Smedt et al., 2013). Selv om forskningen ennå ikke har avgjort hvorvidt ikke-symbolsk mengdeforståelse, eller symbolsk mengdeforståelse er avgjørende for matematikkforståelse, er det på det rene at de begge er av stor betydning for barns matematikkferdigheter (Aunio & Räsänen, 2015; De Smedt et al., 2013).

Ulike studier og meta-studier vektlegger ulike egenskaper som essensielle for barns matematikkferdigheter, men det er ikke konsensus om nøyaktig hvilke som er de mest avgjørende (Fuchs et al., 2006; Geary, 2013; Geary, Hamson, et al., 2000; Gobel, Watson, Lervag, & Hulme, 2014). Likevel ser det ut til å være en bred enighet om at en for å ha gode ferdigheter i matematikk ikke er avhengig av kun én egenskap, men i langt større grad en kombinasjon flere (Aunio & Räsänen, 2015; Geary, Hamson, et al., 2000).

2.2 Utvikling av matematikkferdigheter

Kunnskap om utvikling av matematikkferdigheter er nødvendig for å lettere kunne avdekke elever som strever med matematikk. Det er gjort flere undersøkelser på hvordan matematikkferdigheter utvikler seg fra spedbarnsalder som viser at barn allerede i sitt første leveår er i stand til å oppfatte antall, og å skille mellom grupperinger med ulikt antall objekter der forskjellen er stor, for eksempel seks mot tolv (Feigenson & Carey, 2005; Xu & Spelke, 2000). Den første symbolske kvantitative kunnskap barn tilegner seg er tallordene (Geary & vanMarle, 2016). Allerede fra 2-årsalderen ser vi at barn viser forståelse for at disse tallordene beskriver et bestemt antall objekter, selv om det på et så tidlig stadium dreier det seg om svært basal diskriminering av mengde (Aunio & Räsänen, 2015). Fra 3-årsalderen kan langt de fleste barn si tallordene, men ikke nødvendigvis i riktig rekkefølge, dette lærer de derimot i 4-årsalderen. Omtrent på samme tid kan de også telle objekter, til tross for at tallordene de sier og pekingen på objektene ikke alltid stemmer overens (Aunio & Räsänen, 2015). Ganske små barn vil som regel kunne si hvilken som er størst eller minst når man sammenligner to objekter, men også etterhvert som barnet blir eldre, rangere etter størrelse når det er et høyere antall objekter. Fra 5-årsalderen vil barn kunne forstå sammenhengen mellom tallord og antall. De kan telle objekter mens de peker («en-til-en-korrespondanse»), og vite at det er det samme antallet objekter uavhengig av i hvilken rekkefølge objektene telles. Ved skolestart vil barn med normalutvikling kunne gjenkjenne for eksempel tallet 5 og telle korrekt videre derfra. I utviklingsfasen rundt skolestart ser man en vesentlig økning i barns evne til operere med tallordrekkefølgen også til problemløsning (Aunio & Räsänen, 2015).

2.2.1 Matematikkvansker

Matematikkferdighet er en kontinuerlig variabel som følger en normalfordeling i befolkningen, dette gjør at det er litt tilfeldig hvor man setter kuttpunkt for hva som er vansker eller ikke (Geary, Hoard, Byrd-Craven, Nugent & Numtee, 2007). I tillegg til å skille de som har vansker fra de med normalutvikling er det også vanlig å skille mellom de som er lavtpresterende i matematikk, og de med en mer alvorlig lærevanske, gjerne da kalt dyskalkuli eller spesifikke matematikkvansker (Geary, 2013b). Uansett hvor man setter kuttpunkt, eller hva man velger å kalle de som har vansker med å utvikle gode ferdigheter, er

vi avhengige av å studere hele fordelingen for å få kunnskap om hvorfor noen strever, mens andre får veldig gode ferdigheter i matematikk.

Alvorlige lærevansker knyttet til tall og regneferdigheter er sannsynligvis like utbredt som lærevansker knyttet til leseferdigheter (Butterworth, 2005). Et estimat antyder at det for begge vanskene dreier seg om et sted mellom 3% og 6% (Price & Ansari, 2013). Hos Ostad (2010) anslås det at rundt 10% av elevene i grunnskolen har lærevansker i matematikk, mens Geary (2013b) opererer med et estimat på 7%. Uavhengig av hva som er det nøyaktige omfanget, er det snakk om et signifikant antall barn som har lave ferdigheter i matematikk (Swanson & Jerman, 2006) og som dermed har behov for hjelp og støtte. I diagnosemanualen ICD-10 finner vi under F81.2 denne definisjonen av matematikkvansker: *«Spesifikk forstyrrelse i regneferdigheten (sic.) som ikke bare kan forklares ved generell psykisk utviklingshemming eller utilstrekkelig undervisning. Regnevanskene omfatter manglende evne til å beherske basale regnearter som addisjon, subtraksjon, multiplikasjon og divisjon, snarere enn mer abstrakte matematiske ferdigheter som trengs i algebra, trigonometri, geometri eller komplekse beregninger»* (Helsevesenet, 2017). At det finnes konkrete diagnosekriterier er derimot ikke det samme som at det finnes en enkel oppskrift for å avdekke matematikkvansker hos barn. Det finnes eksempelvis ingen fast test som benyttes i diagnostisering og ingen fast kutt punkt som sier hva som er en spesifikk lærevanske i matematikk og hva som bare er lavtpresterende (Geary, 2013b).

Noen klassiske kjennetegn ved barn med matematikkvansker er de vedvarende problemene med å utvikle helt enkle ferdigheter knyttet til forståelse av antall og hva tallsymbolene representerer. Som vi skal se senere er dette også sentrale prediktorer for matematikkferdigheter i populasjonen, på tvers av hele normalfordelingen. Videre knytter vanskene seg også til problemer med å utvikle evnen til å telle objekter og å se sammenhengen med antall objekter og tallord (Reeve & Gray, 2015). Barn med matematikkvansker kjennetegnes i tillegg ved at de har vedvarende problemer med å utvikle effektive strategier for enkle regneoppgaver (Aunio & Räsänen, 2015). Et annet kjennetegn på barn med matematikkvansker er at de også har en ensidig og uhensiktsmessig bruk av det som kalles backupstrategier (Fuchs & Fuchs, 2002; Geary, 2011a).

Når det gjelder årsaksforklaringer på matematikkvansker er det blant annet gjort undersøkelser hvis resultater peker på at vanskene i matematikk kan relateres til svakheter i det semantiske langtidsmiynet (Geary et al., 1999; Swanson & Jerman, 2006) og i arbeidsmiynet, noe som igjen gjør det vanskelig å lagre aritmetisk fakta i langtidsmiynet (Karagiannakis & Cooreman, 2015). En slik svakhet i miynet vil føre til at erfaringer fra tidligere utførte regneoperasjoner ikke har noen overføringsverdi, eller som Ostad (2010) beskriver det, at eleven ikke kan hente frem regneoppgaver med svaret som én meningsbærende enhet. Felles for både barn med dyskalkuli og de som karakteriseres lavtpresterende er at de har vansker knyttet til forståelse for numeriske størrelser og at de har vansker med å gjenkalle grunnleggende aritmetiske fakta fra langtidsmiynet. Videre kjennetegnes begge grupper ved at de bruker lengre tid på å lære matematiske prosedyrer, og at disse vanskene og forsinkelsene ikke kan tilskrives lav intelligens (Geary, 2011b).

Ulike metoder er blitt brukt for å definere dyskalkuli, flere studier har lagt et IQ-diskrepanskriterium til grunn, altså at prestasjonene i matematikk er tilstrekkelig langt under det en kan forvente basert på den målte IQ hos den enkelte (Barahmand, 2008; Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005). Det finnes også andre typer diskrepanskriterier, en variant legger vekt på en uoverensstemmelse mellom hvordan eleven presterer i ulike fag. Om eleven har gjennomsnittlige prestasjoner i skriftspråkfag, men sliter i matematikk, kan dette styrke mistanken om en spesifikk lærevanske i matematikk. En tredje variant vil være elevens alder, altså at elevens prestasjoner i matematikk er under det en forventer av elever på samme klassetrinn (Ostad, 2010). I følge Geary et al. (2007) kjennetegnes barn med alvorlige lærevansker i matematikk ved at prestasjonene i matematikk ligger under et gitt «kutt punkt». Dette kutt punktet varierer fra relativt mild, på den trettiende persentilen, til mer restriktiv, på den tiende eller femte persentilen. Mye av forskning på matematikkvansker har hatt ganske rause kutt punkt-kriterier, og vil derfor ha sammenblandet barn med relativt udramatiske vansker, med barn som har mer alvorlige vansker (Geary, 2007). Denne typen bruk av kutt punkt kalles også prokuradefinisjoner, og er også blitt brukt i diagnostisering av barn med matematikkvansker (Mazzocco & Thompson, 2005). Ved å bruke avgrensninger basert på slike prokuradefinisjoner, for eksempel ved å bruke en matematikktest for å avdekke de med matematikkvansker, må man merke seg at det er fare

for å ekskludere falske negativt (de som har matematikkvansker, men som av ulike årsaker ikke fanges opp av en gitt test) og inkludere falske positive (de som ikke har vansker, men som av andre årsaker skårer lavt på testen) (Mazzocco & Thompson, 2005).

Noe tilsvarende diskrepanskriterium som vektlegger alder for å diagnostisere dyskalkuli har vært å bruke et forsinkelses-kriterium, her defineres matematikkvansken som dyskalkuli om eleven presterer på et nivå tilsvarende elever som er to år yngre (Gross-Tsur, Manor, & Shalev, 1996; Ramaa & Gowramma, 2002). Videre finnes det en variant hvor det å se etter manglende positiv effekt av intervensjoner er brukt for å diagnostisere dyskalkuli. I denne varianten legges det til grunn at det kan defineres som dyskalkuli dersom matematikkvansken vedvarer til tross for at det blir iverksatt tiltak som har vist effekt i andre tilfeller (Desoete, Roeyers, & De Clercq, 2004).

2.2.2 Biologisk primære kvantitative evner

Sentralt i denne masteroppgaven er en teori som beskriver to ulike systemer som ligger til grunn for matematikkferdigheter. Denne teorien postulerer at det finnes et tverrkulturelt, medfødt sett med evner som er av stor betydning for utviklingen av kvantitative ferdigheter. Det er flere studier som viser at vi mennesker har noen iboende kvantitative ferdigheter, eller evner, som er medfødte og tverrkulturelle (Gallistel & Gelman, 1992; Geary, 1995a; 2000; Temple & Posner, 1998). Geary (1995a) kaller disse ferdighetene for *biologisk primære kvantitative evner*, og de beskrives som kognitive kompetanser som dukker opp i forbindelse med barns naturlige aktiviteter, eksempelvis lek (Lin & Geary, 1998). Felles for disse biologisk kvantitative evnene som oppstår i løpet av barnas første leveår og gjennom førskolealder, er at de ser ut til å springe ut fra et kognitivt system som er biologisk, eller arvelig, og at de er universelle både i uttrykk og utvikling .

Disse primære kvantitative ferdighetene består av en delvis medfødt forståelse av (a) *numerosity*, (b) ordinalitet, (c) telling, og (d) enkel aritmetikk (Geary, 1995a; 2000). *Numerosity*, er et begrep som i denne sammenhengen beskriver evnen til å presist fastslå antallet i et lite sett med gjenstander eller hendelser uten å telle, kan best oversettes med *intuitiv mengdeoppfatning*. Den samme evnen omtales også som *subitizing* (Hannula,

Räsänen, & Lehtinen, 2007). Denne intuitive mengdeoppfatningen, som normalt vil gjøre seg gjeldende fra spedbarnsalderen og vare livet ut, vil som regel begrense seg til å fastslå det nøyaktig antallet i sett med 4 eller færre ting (Geary, 1995a; 2000; Hannula et al., 2007).

Den medfødte grunnleggende evnen til å forstå mer enn og mindre er *ordinalitet*. I tidlig barndom, før barnet har lært tallordene, antar man at denne evnen begrenser til mengder under 5 (Geary, 1995a; 2000). Duncan et al. (2007) fant i resultatene fra sine meta-analyser av 6 longitudinelle datasett at ordinalitet også var en sterk prediktor for senere matematikklæring. I følge Geary (1995a; 2000) viser undersøkelser at det tidlig i utviklingen ser ut til å være tellesystem allerede før barnet gjennom språkutviklingen har tilegnet seg tallord. Dette preverbale tellesystemet benyttes på sett med 4 eller færre elementer (Geary, 2000). Med økte språkferdigheter og læring av tallord ser det ut til å være en medfødt forståelse av at tallordene kan benyttes til telling, måling og enkel regning (Geary, 2000). Den fjerde av de biologisk primære kvantitative evnene er enkel aritmetikk, og denne evnen kan spores tidlig i utviklingen der man ser sensitivitet overfor økning og minking i mengde med små sett. I tidlig barndom ser det ut til at denne evnen er begrenset til addisjon eller subtraksjon innenfor sett av to, for så å gradvis øke til å inkludere sett med flere elementer (Geary, 2000). Førskolebarn ser ut til å kunne legge sammen mengder opp til tre objekter ved å benytte seg av en form for preverbal telling (Geary, 1995a).

2.2.3 Biologisk sekundære kvantitative evner

Voksnes funksjonelle tallforståelse kan måles gjennom blant annet deres evner til å løse enkle tekstopp-gaver som krever hel-tall aritmetikk, brøker, enkel algebra og måling (Geary, 2013a). Denne typen problemløsning avhenger av biologisk kvantitative sekundærferdigheter. Disse biologiske sekundærferdighetene er kulturelt betinget, til tross for at de bygger på de biologisk primære kvantitative evner, som jo ikke er kulturelt betinget. Det er viktig å skille mellom primære og sekundære kvantitative evner blant annet fordi den normative utviklingen av sekundærevner varierer mellom kultur og generasjoner (Geary, 1995a). I det at de biologisk sekundære kvantitative er kulturelt betinget ligger det at disse evnene utvikles i det man kan kalle en *unaturlige* kontekst og ved hjelp av *unaturlige* aktiviteter, det vanligste eksempelet på dette vil være en undervisningssituasjon i skolen. De fleste av

matematikkferdighetene som forventes at barn og voksne skal inneha er ikke medfødte, men bygger på nettopp de medfødte matematiske primæregenskapene som er beskrevet tidligere. Likevel vil sekundærferdighetene, i motsetning til primærferdighetene, kreve organisert trening med høy grad av fokus og mye repetisjon (Geary, 1995a; 2000). Et kjennetegn ved biologiske sekundære kvantitative evner vil derfor være at dette er egenskaper som ikke oppstår i miljøer uten målrettede og organiserte aktiviteter som er utformet med tanke på å tilegne seg nettopp disse ferdighetene (Geary, 1995a).

Under biologiske sekundære kvantitative evner regnes (a) tall og telling, (b) regning/aritmetikk og (c) tekstopp-gaver (Geary, 1995a; 2000) og jeg vil her kort redegjøre for hvordan de enkelte evnene defineres. Kunnskap om tall og telling innebærer å forstå sammenhengen mellom de arabiske tallsymbolene, den verbale representasjonen av tallene og verdien tallordet representerer (Geary et al., 1999). Det arabiske tallsystemet kan sies å være det vanskeligste kvantitative konsept som barn i grunnskolen må lære, og læringen er helt avhengig av den pedagogiske praksis (Geary, 2000). Telling forutsetter forståelse av en-til-en korrespondansen, altså at kun ett tallord tildeles hvert telte objekt, og forståelse av at det siste tallet man teller representerer den totale mengden objekter i settet (Geary et al., 1999). I de fleste industrialiserte land er det også forventet at barn i grunnskolen tilegner seg grunnleggende aritmetisk kalkulasjon og etterhvert lærer strategier for å løse mer komplekse aritmetiske problemer, eksempelvis $396+744$. For å tilegne seg evnen til å utføre aritmetisk kalkulasjon er man avhengig av å memorere grunnleggende tallfakta og tilhørende prosedyrer for kalkulasjon, noe som avhenger av pedagogiske instruksjoner og øving (Geary, 2000). Også i arbeid med tekstopp-gaver vil øving og pedagogiske instruksjoner være nødvendig for å oppnå forståelse. Noe av det som gjør tekstopp-gaver utfordrende er å identifisere hva slags type regneoperasjon som skal benyttes for å løse oppgaven, og å overføre den verbale representasjonen til en matematisk representasjon (Geary, 2000).

2.2.4 Prediksjon av matematikkferdigheter

I ordet prediksjon ligger det å anta eller beregne fremtidige forhold (SNL, 2009), og det vil derfor være longitudinelle studier som følger barn over tid som er best egnet til å si noe om sammenhengen mellom utviklingene av ulike evner og ferdigheter på et tidlig tidspunkt og senere matematikkferdigheter. Tverrsnittstudier som ser på et utvalg barn på ett tidspunkt er

således lite egnet til å predikere fremtidige matematikkferdigheter, og kan altså bare si noen om en sammenheng på et gitt tidspunkt. Som nevnt er matematikkferdigheter en kontinuerlig variabel som følger en normalfordeling i befolkningen, og det vil også være slik at mangelfull eller sen utvikling av de evner og egenskaper som predikerer gode ferdigheter i matematikk vil kunne predikere matematikkvansker.

Geary (2013a) undersøker tre sentrale og konkurrerende teorier for hvordan barn tilegner seg matematikkferdigheter, og om det foregår ved hjelp av (a) et medfødt system som anvendes for å representere omtrentlige størrelser, (b) et system for oppmerksomhetskontroll som eksplisitt behandler kvantitative symboler, eller (c) de logisk problemløsende egenskapene som legger til rette for å tilegne seg kunnskap om relasjonen mellom tall (Geary, 2013a). Studiene det vises til tyder på at hver og en av disse kompetanseområdene er involvert i tidlig kvantitativ læring, men i ulik grad og til ulike steder i læringsprosessen (Geary, 2013a). Av de tre teoriene Geary (2013a) undersøker har én lenge har vært dominerende. Denne teorien postulerer at utviklingen av regneferdigheter avhenger av det medfødte Approximate Number System - ANS. ANS beskrives som et primitivt mentalt system for nonverbal representasjon av antall som gradvis utvikles fra spedbarnsalder og gjennom voksenlivet (Mazzocco, Feigenson & Halberda, 2011a). Noe av årsaken til at ANS lenge har vært ansett som avgjørende når det kommer til utvikling av matematikkferdigheter, er at man gjerne finner svak ANS hos elever med lærevansker i matematikk (Mazzocco, Feigenson, & Halberda, 2011b; Price & Ansari, 2013). Også i longitudinelle studier hvor ANS har blitt målt før skolestart viser resultatene at ANS er en tydelig prediktor for senere matematikkferdigheter (Libertus, Feigenson, & Halberda, 2011; Mazzocco et al., 2011b). I Libertus, Feigenson & Halberda (2011) sin longitudinelle undersøkelse fant de korrelasjon mellom barns ANS og matematikkferdigheter, selv når de kontrollerte for alder og verbalferdigheter, og Libertus med kolleger mener således at disse funnene bidrar til bevis for hypotesen at sammenhengen mellom ANS og senere matematikkferdigheter (Libertus et al., 2011).

Forskningen på matematikkutvikling spriker altså når det gjelder nøyaktig hvilke faktorer som best predikerer matematikkferdigheter, men det ser ut til å være bred enighet om at den sterkeste prediktor for senere prestasjoner i matematikk er tidlige kvantitative evner og tidlige

matematikkferdigheter (Aunio & Räsänen, 2015; Geary, Bailey, & Hoard, 2009). Duncan et al. (2007) fant i sine analyser av seks longitudinelle datasett at elevenes matematikkprestasjoner ved skolestart var det beste prediktor for senere prestasjoner i matematikk (Duncan et al., 2007). En sterk sammenheng mellom tidlige og senere prestasjoner i matematikk ser vi også i resultater fra nasjonale prøver; hvordan elever gjør det på 5. trinn har stor betydning for resultater på nasjonale prøver på 8. trinn, og at resultater fra nasjonale prøver på 8. trinn igjen har stor betydning for avgangresultater fra 10. trinn. Videre vet vi at avgangresultater i matematikk fra 10. trinn er av stor betydning for resultater i matematikk på videregående skole (Kunnskapsdepartementet, 2014).

2.2.6 Kjennskap til tall

Til tross for at flere undersøkelser vektlegger ANS som en viktig prediktor for senere matematikkferdigheter (Libertus et al., 2011; Mazzocco et al., 2011a) har andre undersøkelser funnet resultater som peker på at det er andre evner som er bedre egnet som prediktor (Göbel et al., 2014). Gjennom en longitudinell undersøkelse hvor 173 (testtidspunkt 1) og 165 (testtidspunkt 2) 6/7-åringer ble testet med 11 måneders mellomrom fant Göbel, Watson, Lervåg og Hulme (2014) grunnlag for å hevde at det ikke er ANS, men kjennskap til tall som er mest avgjørende utviklingen av regneferdigheter. I denne studien fant de ingen resultater som gav grunnlag for å hevde at manglende ANS hindrer utvikling av aritmetiske ferdigheter hos 6- og 7-åringer. Derimot viste resultatene at evnen til se sammenhengen mellom flersifrede arabiske tall og dets korrekte «verbale merkelapp» var en sterkere prediktor for den aritmetiske utvikling i denne alderen (Göbel et al., 2014)

Tilsvarende funn viser Desoete (2014) i en treårig undersøkelse hvor 471 barn ble fulgt fra barnehagealder og opp til 2. klasse. Denne undersøkelsen viser at barns kunnskap om tall er av avgjørende betydning for senere matematikk. I denne undersøkelsen av Stock, Desoete & Roeyers (2010) fant man at ved å kartlegge blant annet barnas telleferdigheter og forståelse av tallrepresentasjon klarte forskerne å korrekt diagnostisere 87,50 prosent av de barna som viste seg å ha matematikkvansker (Desoete, 2014). Også i en longitudinell undersøkelse av Aunola og kolleger (2004) som fulgte 194 finske barn gjennom tre år fra barnehage til

andreklasser viser resultatene at telleferdigheter, og dermed kjennskap til tall, var den sterkeste prediktoren for senere prestasjoner i aritmetikk (Aunola et al., 2004).

2.2.7 IQ og matematikkferdigheter

Intelligens, regnes sammen med arbeidsminne og prosesseringshastighet, som generelle kognitive evner og en kombinasjon av nivået på disse evnene vil derfor påvirke læring på mange akademiske områder, inkludert matematikk (Geary, Hoard, Nugent, & Bailey, 2012). Av de nevnte kognitive evnene anses intelligens ofte som en nøkkelfaktor når det kommer til matematikkprestasjoner (Kyttälä & Lehto, 2008), og intelligensmålinger er konsistente prediktorer for akademiske prestasjoner (Walberg, 1984). I enkelte undersøkelser viser resultatene at intelligens er en tydelig prediktor også for senere matematikkferdigheter, blant annet en femårig longitudinell studie med 70,000 elever som ble publisert i 2006. Der fant Deary, Strand, Smith og Fernandes (2007) at intelligens målt når elevene var 11 år forklarte nesten 60% av variasjonen på nasjonale matematikkprøver når eleven var 16 år. Tilsvarende viser Kyttälä & Lehto (2008) til en studie utført av Kuusinen & Leskinen i 1986, hvor resultatene viste at intelligens tester som ble utført mellom 1. og 4. trinn forklarte 32% av matematikkprestasjonene på 9. trinn (Kyttälä & Lehto, 2008).

I en longitudinell undersøkelse som målte utviklingen av den mentale tall-linje representasjonen hos barn med lærevansker i matematikk ble også sammenhengen med IQ undersøkt (Geary, Hoard, Nugent, & Byrd-Craven, 2008). Undersøkelsen fulgte 62 lavtpresterende barn, de skåret under 25. persentilen på matematikktesten som ble brukt for utvelgelse, gjennom 1. og 2. trinn, og her viste resultatene at det var en sammenheng mellom de matematiske ferdighetene knyttet til tall-linjen, og blant annet IQ. Sammenhengen med IQ var spesielt viktig på 1. trinn, der høyere IQ-skåre var relatert til blant annet mer presis plassering av tall på tall-linjen og hyppigere bruk av en lineær strategi (Geary et al., 2008). Denne sammenhengen mellom IQ og matematikkprestasjoner knyttet til tall-linjen i første klasse henger sammen med evnen til logisk og systematisk tenking, som igjen er aspekter ved intelligens som legger til rette for forståelse av den logiske strukturen i tall-linjen (Geary et al., 2008).

Man finner altså ofte en sammenheng mellom matematikk og intelligens i longitudinelle studier (Deary et al., 2007; Geary et al., 2008; Kyttälä & Lehto, 2008), men det er heller ikke slik at man alltid finner en klar sammenheng mellom IQ og matematikk. I en undersøkelse publisert i 2007 ble 170 barn testet i starten og slutten av første klasse. Her ble sammenhengen mellom kognitive evner og matematikkprestasjoner analysert, og disse resultatene viste at intelligensen som ble målt i denne testen ikke hadde noen direkte påvirkning på elevenes målte matematikkferdigheter (Passolunghi, Vercelloni, & Schadee, 2007). Barn med matematikkvansker skårer riktignok som regel noe under gjennomsnittet (Geary et al., 2012), men dette forklarer imidlertid ikke nødvendigvis hvorfor de strever i matematikk, for selv om de skårer under gjennomsnittet, er de likevel innenfor normalområdet (Geary, 2013a). En ting er at IQ og matematikk korrelerer på et tidspunkt, og at nonverbal IQ ser ut til å kunne forklare opp til 60% av variasjonen i matematikkferdigheter på senere tidspunkt, men selv da gjenstår det 40% som ikke kan forklares av IQ. Det påpekes derfor fra flere hold at den nøyaktige sammenhengen mellom matematikk og IQ ikke er helt avklart (Blackwell, Trzesniewski, & Dweck, 2007; Geary, 2011b).

2.2.8 Eksekutive funksjoner og matematikkferdigheter

Tidligere ble eksekutive funksjoner behandlet som et enhetlig system, men eksekutive funksjoner anses nå for å bestå av flere ulike funksjoner som er distinkte, men har en nær forbindelse til hverandre (Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Eksekutive funksjoner er vanligvis beskrevet som et sett med mentale evner som inkluderer (a) arbeidsminne, som gjør en i stand til å mentalt holde på informasjon, (b) en kontrollferdighet som gjør en i stand til å ivareta relevant informasjon samtidig som man ignorerer eller stenger ute irrelevant og distraherende informasjon og (c) en kognitiv fleksibilitet som gjør at man kan skifte strategi avhengig av hva den enkelte oppgaven krever (Fuhs, Hornburg, & McNeil, 2016; Miyake et al., 2000). I denne sammenhengen er arbeidsminne-funksjonen ofte kalt *Updating*, og evnen er nødvendig også for å manipulere og oppdatere informasjon som er midlertidig lagret i arbeidsminne (Van der Ven et al., 2012). Studier som har sett på sammenhenger mellom eksekutive funksjoner og matematikkvansker finner nesten utelukkende sterke sammenhenger mellom updating/arbeidsminne og

matematikkferdigheter (De Smedt et al., 2009; Passolunghi, Mammarella, & Altoe, 2008; Van der Ven et al., 2012).

Om den egenskapen som er knyttet til å undertrykke forstyrrende informasjon eller adferd til fordel for mer målrettede handlinger brukes begrepet *inhibition* (Bull, Espy, Wiebe, Sheffield, & Nelson, 2011). Inhibition er i følge enkelte studier uløselig knyttet til en annen av de eksekutive funksjoner, nemlig *shifting* (Van der Ven et al., 2012). Shifting, eller kognitiv fleksibilitet er nødvendig for skifte frem og tilbake mellom ulike strategier eller tenkesett for å så presist og effektivt som mulig løse en gitt oppgave (Miyake et al., 2000). Dette antas å ha betydning for matematikkferdigheter fordi man i matematikken kan benytte seg av flere ulike strategier for å løse en oppgave, og vil da være eksekutive funksjoner som benyttes når en tar avgjørelsen om hvilken strategi som er mest hensiktsmessig i den enkelte oppgave (Murray, Hillaire, Johnson, & Rappolt-Schlichtmann, 2015). U hensiktsmessige strategier er således et tegn på matematikkvansker (Fuchs & Fuchs, 2002; Geary, 2011a; Ostad, 2010).

Selv om eksekutive funksjoner er en sekkebetegnelse som inneholder flere ulike evner eller egenskaper, er det vanskelig å skille disse fra hverandre når man ser på de eksekutive funksjonenes rolle i å løse matematikkproblemer (Fuhs et al., 2016). Årsaken til at de ulike underkomponentene i eksekutive funksjonene er vanskelig å måle enkeltvis er at de i stor grad overlapper, samtidig som det er fare for målefeil fordi testresultatene også vil påvirkes av barnas språkferdigheter, motoriske ferdigheter, eller andre ferdigheter som kreves for å utføre den aktuelle testen (Fuhs et al., 2016; Miyake et al., 2000).

Flere har sett på sammenhenger mellom eksekutive funksjoner og matematikkferdigheter (Bull, Espy, & Wiebe, 2008; Bull & Scerif, 2001; Clark, Sheffield, Wiebe, & Espy, 2013; Van der Ven et al., 2012), og resultater fra stadig flere studier viser en sterk og vedvarende sammenheng mellom barns eksekutive fungeringsevne og hvordan de presterer i matematikk (Bull et al., 2008; Lee, Bull, & Ho, 2013; Lee, Ng, Bull, Pe, & Ho, 2011). Blant andre hos Karagiannakis og Cooreman (2015) sies det at en mulig årsak til barns matematikkvansker kan forklares i en kognitiv svikt som rammer de eksekutive funksjonene. Til tross for at sammenhengen mellom eksekutive funksjoner og matematikkferdigheter har blitt grundig studert, er det fremdeles uklart hvordan dette forholdet utvikler seg over tid fordi det mangler

gode longitudinelle undersøkelser (Van der Ven et al., 2012). Passolunghi med kolleger (2008) viste at sammenhengen mellom kognitive ferdigheter og matematikk endret seg over tid. I deres longitudinell undersøkelse viste resultatene at der den fonologisk løkken, IQ og den eksekutive funksjonen *updating* i første klasse bidro til matematikkprestasjoner, var det i andre klasse kun *updating* som viste seg å være en prediktor (Passolunghi et al., 2008). At dette mønsteret endrer seg over tid viser at utviklingen av kognitive evner, som eksekutive funksjoner, er en dynamisk og kompleks prosess (Van der Ven et al., 2012). Selv om det fremdeles ikke finnes en universell definisjon av eksekutive funksjoner, er det etterhvert en bred enighet om at de egenskapene som paraplybegrepet *eksekutive funksjoner* rommer er avgjørende når det kommer til å utvise fleksibilitet og evnen til å tilpasse seg i nye situasjoner. Da nettopp læring er en slik situasjon antas det at eksekutive funksjoner spiller en nøkkelrolle tilegnelsen av matematikkferdigheter (Van der Ven et al., 2012).

2.3 Kjønnforskjeller i matematikk

En britisk undersøkelse publisert i 2002 viser at foreldre, ved å bruke IQ-skår, vurderer sine guttebarns intelligens på området «Matematikk og logikk» som signifikant høyere enn sine pikebarns intelligens på samme område. Dette er altså snakk om hva foreldrene tror, og ikke resultater av IQ-tester. Differansen er størst i fedrenes vurdering av barna, med henholdsvis 110 for sine guttebarn, og 98 for sine pikebarn. Fra mødrenes side vurderes jentene til 104 og guttene til 110 (Furnham, Reeves & Budhani, 2002). Også i andre undersøkelser ser det ut til å være en utbredt oppfatning blant voksne at gutter «naturlig» er flinkere enn jenter i matematikk (Nosek et al., 2009). Heldigvis, kan man si, finnes det flere metastudier som viser at det ikke er signifikant forskjell på gutter og jenters matematikkprestasjoner på barneskolen (for eksempel Lindberg, Hyde, Petersen, Linn og Hinshaw (2010)) I den grad det er forskjeller fant både Hyde, Fennema & Lamon (1990) og Steele (2003) at forskjellene ikke bare var minimale, men at jentene på enkelte matematikkområder gjorde det bedre enn guttene.

Også i Norge viser tall fra SSB at jenter i enkelte sammenhenger gjør det bedre enn gutter i matematikk; i 2016 var standpunkt karakteren i matematikk fra grunnskolen i snitt 3,7 for jenter mens den for gutter var 3,4 (SSB, 2016). Når det kommer til resultater fra TIMSS

(Trends in International Mathematics and Science Study) ser man at det på 4./5. trinn ikke er signifikante forskjeller mellom gutter og jenter i Norge og Sverige, samtidig er det små men signifikante forskjeller i de andre referanselandene; i Finland skårer jentene i snitt 9 poeng høyere enn guttene, mens i England, Danmark og USA er det guttene som i snitt skårer signifikant høyere med henholdsvis 6 poeng i England og Danmark, og 7 poeng i USA (Bergem, 2016). I denne sammenheng må det nevnes at karakterer nok anses som mindre «rene» mål på ferdigheter enn det resultater fra tester gjør. Karakterer settes ofte ut fra mer skjønnsmessige vurderinger, mens tester utført i forskningsøyemed har strengere vurderingskriterier. Jenter skårer for eksempel dårligere enn gutter på standardiserte tester i matematikk der stoffet ikke er tett knyttet opp mot det de har gjennomgått i matematikkundervisningen, til tross for at de får bedre karakterer (Halpern et al., 2007).

Det er altså en stereotypi som sier at jenter er underlegne gutter i matematikk, til tross for at det i liten grad finnes resultater fra vitenskapelige undersøkelser som kan bekrefte denne stereotypien (Furnham et al., 2002; Nosek et al., 2009). Derimot foreligger det en betydelig mengde forskning som taler for at det finnes kjønnsforskjeller når det kommer til spesifikke evner som har betydning for matematikkprestasjoner (Furnham, Reeves & Budhani, 2002; Geary, 1995b, 1996). Mye av forskningen på kjønnsforskjeller i matematikk har tatt for seg ulikheter i de tre kognitive områdene (a) verbale evner, (b) visuospatiale evner, og (c) kvantitative evner (Halpern et al., 2007). Sammenhengen mellom verbale evner og matematikkprestasjoner gjør seg gjeldene på flere områder; både for å forstå tidvis komplekse instruksjoner direkte knyttet til problemløsning og fordi det er et generelt krav om språkferdigheter i utdanning (Halpern et al., 2007). Til tross for at jenter ofte presterer bedre enn gutter på områder som krever verbale evner, som for eksempel lesing (Halpern et al., 2007), finner man i enkelte undersøkelser knyttet til matematikkferdigheter kjønnsforskjeller i gutters favør når det kommer til å løse tekstoppgaver allerede i førsteklasse. En tendens som også fortsetter utover gjennom skolegangen (Harnisch, Steinkamp, Tsai, & Walberg, 1986; Lummis & Stevenson, 1990). Også Geary (1996) påpeker at gutter konsistent gjør det bedre enn jenter på tekstoppgaver.

En forklaringsmodell som ofte trekkes frem for hvorfor gutter presterer bedre i matematikk enn jenter på høyt nivå, er kjønnsforskjeller i spatiale evner (Geary, 1995b; Halpern et al., 2007; Zhou, 2015). Geary (1995b) fremsetter en tese om at det er kjønnsforskjeller i spatiale evner som gjør at menn har et vedvarende fortrinn når det kommer til mental manipulasjon av tredimensjonal representasjon av informasjon, navigasjon og til å forutse banen til objekter i bevegelse. Dette er alle ferdigheter som avhenger av visuospatiale evner. Denne tesen har i mange tilfeller blitt bekreftet empirisk. I en review av Halpern et al. (2007) trekkes det frem at en stor mengde forskning gjennom de siste 25 år har avdekket kjønnsforskjeller nettopp i enkelte mål på visuospatiale evner, som igjen kan forklare at gutter prestere bedre på høye nivå i matematikk. At eldre gutter har høyere skårer på spatiale evner viser også resultatene fra en undersøkelse av Geary, Sauls, Lie & Hoard (2000) som så på sammenhengen mellom kjønn og matematikkferdigheter. I denne undersøkelsen, der 113 gutter og 123 jenter på 19 år ble testet på ulike områder, blant andre regning, IQ og spatiale evner, viste resultatene at det ikke var kjønnsforskjeller i IQ, men at guttene hadde signifikant høyere skåre på både regning og spatiale evner (Geary, Sauls, et al., 2000).

Når det gjelder kjønnsforskjeller knyttet til kvantitative evner følger ikke disse et forutsigbart utviklingsmønster og de varierer ut fra hvilken type kvantitative ferdighet som måles (Halpern et al., 2007). I de første skoleårene mens matematikk ennå handler om grunnleggende regning og hastighet ser jenter ut til å gjøre det noe bedre enn gutter, men denne forskjellen utjevnes, for så å bli i guttenes favør etterhvert som matematikken blir mer komplisert og avhengig av spatiale ferdigheter (Halpern et al., 2007). Empiri som bekrefter jenters fortrinn i aritmetikk i tidlig skolealder finnes blant annet i en studie fra 2011 der 1556 elever mellom 8 og 11 år ble testet for ulike kognitive ferdigheter (Wei, Liu, & Barnard-Brak, 2015). I denne undersøkelsen hadde jentene høyere skårer enn guttene i blant annet aritmetikk og oppgaver knyttet til kjennskap til tall, mens guttene skåret bedre på oppgaver knyttet til spatiale evner (Wei et al., 2015).

I en metastudie som tok for seg 100 undersøkelser om kjønnsforskjeller i matematikk utført mellom 1963 og 1988 så man på blant annet sammenhengen mellom kjønn og prestasjoner i problemløsningsoppgaver. I denne sammenhengen beskrives problemløsningsoppgaver som

en type oppgaver der eleven må benytte seg av kjent kunnskap i nye situasjoner. En oppsummering av resultatene viste at det enten ikke var kjønnsforskjeller, eller så gjorde jentene det marginalt bedre i denne type oppgaver i barne- og ungdomsskolen. Likevel oppstod det en forskjell som var noe tydeligere i favør av gutter på den samme typen oppgaver i videregående skole og i høyere utdanning (Hyde et al., 1990). Den samme metastudien så også på kjønnsforskjeller i regning, som i denne sammenhengen definert som oppgaver med bruk av bare algoritmiske prosedyrer for å finne et enkelt numerisk svar. Tilsvarende resultatene fra kjønnsforskjeller i problemløsningsoppgaver, gjorde jenter det også her marginalt bedre enn gutter i regning på barneskolen, men utover i utdanningsløpet ble det ikke lenger målt kjønnsforskjeller (Hyde et al., 1990).

Også i en tverrsnittstudie av Carr og Davis (2001), hvor 42 gutter og 42 jenter i første klasse løste regneoppgaver med noen gitte instruksjoner så man kjønnsforskjeller. I det ene settet kunne de løse oppgavene med den strategien de selv ønsket, mens det andre settet var lagt opp som et spill med regler for hvilke strategi man måtte benytte seg av på de enkelte oppgavene. Resultatene viste at der elevene selv kunne velge strategi valgte jentene stort sett å benytte seg av konkrete, mens guttene stort sett benyttet seg av retrievalstrategi. I settet der de ulike strategiene var bestemt og kontrollert for hver oppgave viste det seg at guttene mestret å løse oppgaver også med konkrete, mens jentene ikke var i stand til å benytte retrievalstrategier i like stor grad som guttene (Carr & Davis, 2001).

Hvorvidt det er kjønnsforskjeller i matematikk vil altså avhenge av blant andre hvilke ferdigheter i matematikk man ser på, på hvilket tidspunkt i livet de er analysert og i hvilken kontekst analysene er blitt utført i (Halpern et al., 2007). En metastudie av Lindberg med kolleger (2010) fra 2010 analyserte funn fra 242 undersøkelser utført mellom 1990 og 2007, med totalt 1.286.350 deltakere. Resultatet her viste at den samlede vektete effektstørrelse totalt var på $d = +0.05$, hvilket vil si at det er en 98.01% overlapp mellom de to gruppene og ved å trekke ut en tilfeldig gutt er det bare 51% sjanse for at han er bedre enn gjennomsnittet for jenter. Forskjellene er altså så marginale at det nesten ville være en overdrivelse å si at forskjellen er «til fordel» for gutter. Kort oppsummert har tidligere forskning vist at

kjønnsforskjellene i matematikkprestasjoner er svært små og noen ganger i favør av gutter, andre ganger i favør av jenter (Lindberg et al., 2010).

2.3.1 Kjønnsforskjeller blant barn med matematikkvansker

Som med matematikkferdigheter og matematikkvansker generelt, vil det også når det gjelder kjønnsforskjeller være en glidende overgang og mye av det som er nevnt tidligere om kjønnsforskjeller i matematikkferdigheter vil være relevant for kjønnsforskjeller blant barn med matematikkvansker. I en metastudie fra 2013 hvor en ser på kjønnsforskjeller i dyskalkuli viser resultatene at kjønnsforskjeller i dyskalkuli ikke overraskende avhenger av diagnosekriterier (Devine, Soltesz, Nobes, Goswami, & Szucs, 2013). I den studien vises det til at resultater fra tidligere undersøkelser om matematikkvansker ikke er konsistente når det gjelder kjønnsfordelingen av de med vansker. Eksempelvis finner Dirks, Spyer, van Lieshout og de Sonneville (2008) at utbredelsen av dyskalkuli er litt høyere blant jenter enn gutter på fjerde- og femte-trinn. Her var det brukt en kuttpunkt på tiende- og tjuefemte-persentilen på standardiserte nederlandske tester. Til sammenligning viste resultatene fra en undersøkelse av Barbaresi et al. (2005) at forekomsten av dyskalkuli var høyere for gutter enn jenter uavhengig av både barnas alder og hvordan dyskalkuli var definert. I en gjennomgang av ulike undersøkelser om kjønnsforskjeller i blant barn med matematikkvansker konkluderer Devine et al. (2013) med at det ikke finnes kjønnsforskjeller i utbredelse av matematikkvansker der man bruker tydelige diagnosekriteriene, for eksempel ved å sette et absolutt kuttpunkt eller ved å legge spesifikke diskrepanskriterier til grunn. Selv om gutter er overrepresentert blant mange lærevansker, tyder ikke data per 2013 på at gutter hverken er over- eller underrepresentert blant barn med lærevansker i matematikk (Devine et al., 2013). Derimot er det ofte en bredere variasjon i gutters skårer på mål for kvantitative evner, noe som nødvendigvis betyr at flere gutter enn jenter er representert blant både de sterkeste og de svakeste (Halpern et al., 2007).

2.4 Oppsummering av teori og empiri

Temaet for denne oppgaven er kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger, og vil bli belyst ved å svare på forskningsspørsmålene;

- 1. Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?*
- 2. Er det kjønnsforskjeller når vi kontrollerer for nonverbal IQ og eksekutive funksjoner.*

I dette kapittelet har vi sett at matematikkprestasjoner og ferdigheter er komplekst sammensatt og at det krever flere ulike evner og egenskaper for å ha gode ferdigheter. Empirien er ikke entydig, og resultater fra ulike undersøkelser kan peke i ulike retninger, men det man kan slå fast er at tidlige kvantitative evner vil predikere tidlige matematikkferdigheter som igjen i stor grad predikerer senere matematikkprestasjoner. Nøyaktig hvilke ferdigheter som er mest sentrale er det fremdeles ikke mulig å slå fast. En sentral teori beskriver en todeling av kvantitative ferdigheter, hvorav de første går inn under biologisk primære kvantitative evner, som er iboende, tverrkulturelle og uavhengige av formell opplæring. Dernest følger de biologisk sekundære kvantitative evnene, som bygger på de primære, men til forskjell fra disse krever formell læring for å utvikles. Denne teorien er lagt til grunn for utformingen av forskningsspørsmålene, og dermed også i analysene av materialet i denne undersøkelsen.

Videre er det vist til ulike undersøkelser som tar for seg forholdet mellom IQ og matematikkferdigheter, og eksekutive funksjoner og matematikkferdigheter. Heller ikke her er empirien entydig, men at non-verbal IQ har en viss sammenheng med matematikkprestasjoner er tydelig både i gruppesammenligningstudier og i longitudinelle studier. Det samme gjelder i stor grad for betydningen eksekutive funksjoner har for matematikkprestasjoner. Den nøyaktige sammenhengen er uklar, men det er bred enighet om at evner som er knyttet til dette området er av betydning for matematikkferdigheter. Det synes å være en oppfatning blant mange at gutter har et slags naturlig fortrinn når det kommer til matematikkprestasjoner. Til tross for at det er flere menn som tar høyere

utdanning med avansert matematikk spriker forskningen på om dette kan forklares av iboende, eller biologiske, kjønnsforskjeller til menns fordel. Det er gjort en rekke undersøkelser på kjønnsforskjeller i matematikkferdigheter, både blant normalpresterende, og blant barn med lærevansker i matematikk, og som jeg har vist til i dette kapitlet er ikke resultatene entydige.

Basert på empirien som er presentert i kapittel 2 er det nærliggende å ha en nullhypotese som utgangspunkt for analysene. En nullhypotese, H_0 , vil si at det ikke er noen sammenheng mellom variablene. I dette tilfellet at det ikke er noen sammenheng mellom variabelen *kjønn* og de fire variablene som representerer matematikkferdigheter, men som resultatene vil vise er det på området *primære kvantitative evner* signifikante kjønnsforskjeller. Resultatene som presenteres i kapittel 4, og betydningen av disse vil bli drøftet i kapittel 5.

3. Metode

I dette kapittelet vil det gjøres rede for den metodiske tilnærming, datainnsamlingen og analyseprosessen. Denne undersøkelsen er tilknyttet prosjektet *The effect of mathematical intervention programs for children performing low in mathematics*, heretter kalt Prosjektet, og vil dermed måtte forholde seg til de rammebetingelsene dette gir. Disse betingelsene vil bli redegjort for her. Videre vil forhold knyttet til validitet og reliabilitet kort forklares før jeg presenterer analysene som skal utføres. Avslutningsvis vil jeg gå inn på min rolle som forsker og løfte vurderinger knyttet til etikk.

3.1 Design

Forskningsdesign kan defineres som det helhetlige opplegget for et forskningsarbeid. Fra formål og problemstillinger, gjennom datainnsamlingsmetoder og dataanalyse frem til konklusjon og diskusjon utgjør til sammen forskningsdesignet. I dette tilfellet er det for problemstillingen « *Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?* » valgt en kvantitativ metodisk tilnærming hvor designet er en ren observasjonsstudie hvor det ikke foretas noen form for manipulasjon, og i så måte er designet ikke-eksperimentelt. At designet er ikke-eksperimentelt vil si at vi ikke forsøker å gjøre en påvirkning som endrer tingens tilstand for så å undersøke denne endringen, men at man simpelthen vil undersøke «virkeligheten» slik den fremstår (Kleven, 2002b). Alle data er samlet inn på ett tidspunkt og materialet egner seg dermed for en tverrsnittanalyse (Midtbø, 2007).

I denne undersøkelsen er formålet å studere kjønn som en mulig påvirkningsfaktor for matematikkferdigheter, og om forskjeller som oppdages i så fall kan forklares av andre faktorer. I analysene vil kjønn være uavhengig variabel mens avhengig variabel vil være matematikkferdigheter, kontrollert for IQ og eksekutive funksjoner. Problemstillingen spør om det er kvalitative forskjeller i egenskap av å være gutt eller jente som fører til ulike forutsetninger for matematikkprestasjoner. H_0 er da at det ikke vil fremkomme forskjeller i matematikkprestasjoner hos gutter og jenter, eller i så fall at denne forskjellen kan forklares

av noe annet enn testobjektens kjønn. H_1 vil i så måte være at det finnes kjønnsforskjeller, og at disse ikke kan forklares av andre variabler.

3.2 Utvalg

Utvalget består av 120 førsteklasinger, 68 jenter og 52 gutter, fra ni skoler i to kommuner på Østlandet. Den ene kommunen ønsket selv å delta etter å ha blitt kjent med Prosjektet da pilotering av screeneren ble gjennomført der. I den andre kommunen gikk forespørselen til kommunalsjefen og enhetslederne. Disse var interessert og det ble så avholdt et informasjonsmøte der Prosjektet ble nærmere presentert for kommunalsjef og enhetsledere. Etter å ha brakt invitasjonen videre til rektorene på de respektive skolene ble resultatet at alle skolene i begge kommunene ønsket å delta i studien. Videre gikk det fra Prosjektet, gjennom skolene, ut et informasjonsskriv til de foresatte av cirka 410 elever. Av disse takket 29 nei til deltakelse, 369 takket ja, og en liten gruppe på 12 responderte ikke på henvendelsen. Alle de 369 som hadde takket ja til deltakelse gjennomførte høsten 2016 en screener, og av disse ble de 120 (32,52%) som presterte lavest plukket ut til videre deltakelse. Det ble så foretatt en randomisering av 120 utvalgte deltakerne hvorav noen ble plassert i intervensjonsgruppen, og de andre i en kontrollgruppe. Eksklusjonskriterier fra studien var kun uttredelse for nevrologiske dysfunksjoner.

3.3 Datainnsamling

Datainnsamlingen ble foretatt av totalt åtte forskningsassistenter i tillegg til doktorgradsstipendiaten som leder Prosjektet. De i alt 12 deltestene ble utført i to omganger med henholdsvis fem og syv tester i hver sesjon som tok mellom 45 minutter og en time. Testingen ble todelt både av praktiske årsaker og for at elevene i størst mulig grad skulle være skjerpet i testsituasjonen, en lengre økt ville nødvendigvis føre til mer utålmodighet og sviktende konsentrasjon mot slutten av testingen. Alle tester ble utført en-til-en med elevene på grupperom der ingen andre var til stede. Testleder jobbet bevisst med å skape tillit og trygghet i testsituasjonen slik at barna kunne prestere sitt beste, og at det dermed ble en positiv opplevelse for barnet å delta. Videre skulle også testleder holde elevens fokus på oppgaven ved være oppmuntrende og anerkjennende uten å være ledende. Som motivasjon

fikk barna i starten av hver sesjon utdelt et diplom som de kunne skrive navnet sitt på. Hver diplom hadde en kolonne med tall nedover som stemte overens med antall deltester i den enkelte sesjon, og for hver utførte deltest fikk barnet et klistremerke som kunne settes ved siden antallet oppgaver som var utført. Etter hver sesjon fikk barna diplommet med seg hjem som en takk for deltakelsen.

3.4 Testene

Testbatteriet bestod av i alt 12 deltester. Av disse 12 har jeg benyttet meg av 5 tester som måleinstrumenter i mine analyser; (1)ThinkMath, (2)WISC-IV-Regning, (3)TOBANS ikke-symbolisk sammenligning og TOBANS tall-sammenligning (4)Tower Of London og (5) Ravens Matriser. De siste 6 testene er ikke lagt til grunn for analysene i denne oppgaven og vil derfor heller ikke bli beskrevet nærmere, men for ordens skyld nevnes de kort her; (7) Listening Recall - Test av arbeidsminne, (8) Baklengs tallhukommelse - test av arbeidsminne, (9) Test av ordforståelse, (10) TOWRE - Test Of Word Reading Efficiency, (11) Woodcock-Johnson - test av lytteforståelse, og sist (12) Speed of Processing - test av prosesseringshastighet.

3.4.1 Matematikktester

Prosjektets anliggende er å måle effekten av en matematikkintervensjon på 1. trinn og det foreligger derfor en rekke tester som måler ulike former for matematikkferdigheter. Den mest omfattende testen er utviklet av forskergruppen selv og den er derfor ikke publisert eller standardisert. Prosjektets egen test heter ThinkMath og den administreres ved at testleder gir detaljerte instruksjoner fra et medfølgende instruksjonshefte. Ferdigheter som måles her er relasjonelle ferdigheter i matematikk, telleferdigheter, problemløsningsferdigheter og grunnleggende aritmetiske ferdigheter i totalt 100 enkeltoppgaver.

I de første oppgavene er det de relasjonelle ferdigheten som måles, for å se i hvilken grad eleven har kunnskap om tallenes forhold til hverandre, altså hva som er «én mer enn», «én mindre enn», «den andre blomsten» på en rekke av flere, «den fjerde bilen» på en rekke av flere. I tillegg kommer også kjennskapen til relasjonen mellom tallsymbolene i oppgaver hvor eleven skal sette et merke på det minste av tre tall som står ved siden av hverandre.

I de første oppgavene som måler telleferdigheter står to tall ved siden av hverandre, eleven skal så merke av tallet som kommer etter;

3 4 _

2	1	5	6
---	---	---	---

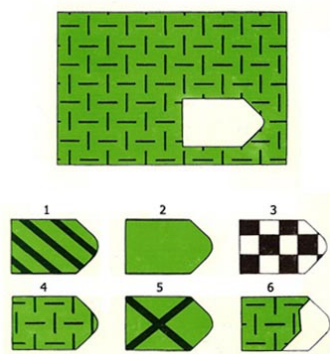
I de neste oppgavene, som også måler telleferdigheter, skal eleven telle hvor mange små kvadrater som befinner seg inne i ulike bokser og skrive tallsymbolet som tilsvarer antall kvadrater. Videre er det åtte problemløsningsoppgaver av den typen man finner i tekstoppgaver, men på grunn av elevens alder og leseferdigheter leses oppgaven opp av testleder. Disse oppgavene måler det som i kapittel 2 omtales som problemløsningsoppgaver. Til sist i ThinkMath er det det 10 regneoppgaver med addisjon og 10 regneoppgaver med subtraksjon, disse oppgavene er lagt til grunn for variabelen *Aritmetikk*.

TOBANS står for *Test of Basic Arithmetic and Numeracy Skills* og det finnes totalt 8 TOBANS tester som måler ulike ferdigheter. Alle TOBANS-testene bruker standardiserte skårer. «TOBANS Dot comparison» måler elevens approximate number system, mens i «TOBANS - Digit comparison» (symbolsk) er det kjennskap til tallsymboler som måles. I «TOBANS Dot comparison» (ikke-symbolsk sammenligning), skal eleven merke av på en boks med sirkler. To og to bokser er sidestilt og det er ulikt antall sirkler i de to boksene, eleven skal uten å telle merke av den boksen der det er flest sirkler. Først gjør eleven tre øvingsoppgaver før oppgavene med tidtaking igangsettes. Også i «TOBANS - sammenligne tall» er det 3 øvingsoppgaver først, deretter igangsettes oppgavene med tidsbegrensning.

I testen WISC-IV Regning skal eleven løse en rekke regneoppgaver ved hjelp av hoderegning. Oppgavene presenteres muntlig, og skal besvares innenfor en gitt tidsgrense på 30 sekunder. Testen skal avsluttes når eleven har avgitt galt svar, eller ikke svart innenfor tidsbegrensning 4 ganger på rad. Selv om oppgavene presenteres muntlig har de mange likhetstrekk med tekstoppgaver, som at det ikke oppgis hvilken strategi som må benyttes for å løse oppgaven og at den fremstilles som en «historie» med kjente elementer. Eksempelvis; «Per har tre blyanter, han får to til av moren sin. Hvor mange blyanter har Per?». Det legges i denne studien til grunn at disse oppgavene måler det som i kapittel 2 omtales som «problemløsningsoppgaver».

3.4.2 Raven

Ravens matriser er en serie ikke-verbale evnetester ment for bedømmelse av generelle evner. I denne sammenhengen er det Ravens CPM - Coloured Progressive Matrices som er benyttet fordi den er beregnet på yngre barn. I stimulusheftet er det totalt 36 «bilder» hvorav de 2 første er øvingsoppgaver. Hver side inneholder et rektangel hvor en bit mangler og nedenfor er det 6 ulike biter, merket 1-6. Barnet skal så ved å peke eller si tallet på biten de mener passer best i «hullet» i bildet. Ravens matriser måler evnen til å på grunnlag av kjent informasjon kunne foreta en logisk utledning av nye relasjoner og mønstre, og testen er ikke tidsbegrenset.

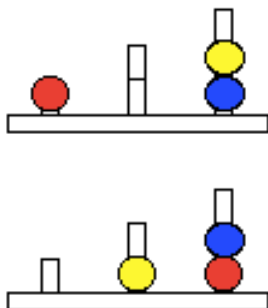


Figur 3.1 Eksempel på oppgave fra Raven.

3.4.3 Tower of London

TOL tester eksekutive funksjoner og den utføres ved at eleven løser oppgaver på en PC. På dataskjermen vises en todelt ramme, hvor det på den øverste delen av rammen er tre ulikt farget kuler plassert på tre pinner med ulik høyde. På den nederste delen er det de samme kulene og pinnene, men kulene er plassert annerledes enn de øverste. Testobjektet skal så flytte «sine» kuler på den nederste delen av skjermen slik at de får lik formasjon som kulene øverst. For hver ny oppgave vil kulene være ulikt plassert og det vil være ulikt antall trekk som kreves for å løse oppgaven, og også ulikt antall trekk eleven får utføre før programmet går videre til neste oppgave. Det er også begrenset hvor mye tid eleven har til rådighet på hver enkelt oppgave. Skåringen gjøres av programmet og skjer automatisk (Culbertson & Zillmer, 1999). Vanskelighetsgraden og antall trekk som behøves for å løse oppgaven vil øke utover i testen. Oppgavene i TOL anses som planleggingsoppgaver fordi man antar at eleven

vil utføre en mer effektiv problemløsning om de ulike trekkene som behøves for gjennomføringen av oppgaven er planlagt før man begynner selve flytting. (Riccio, Wolfe, Romine, Davis, & Sullivan, 2004).



Figur 3.2 Eksempel på oppgave fra Tower Of London

3.5 Validitet og reliabilitet

Cook og Campbells validitetssystem er en metodologisk referanseramme innen kvantitativ forskning (Lund, 2002). For å belyse et forskningsproblem er det vesentlig at slutningene som trekkes er valide. Validitet sier noe om gyldigheten av slutningene i en studie, og det skiller mellom fire ulike typer validitet; statistisk validitet, indre validitet, begrepsvaliditet og ytre validitet korrekte. Likevel er ikke denne validiteten absolutt, men det er snakk om grad av validitet (Shadish, Cook, & Campbell, 2002). Trusler mot validitet vil være spesifikke grunner til at de slutninger vi trekker er helt eller delvis gale. Ved å være bevisst disse truslene kan man i sin forskning forsøke å minimere sjansene for at de skal svekke slutningens validitet (Shadish et al., 2002). I metodekapittelet følger en kort gjennomgang validitetskravene, hvor også reliabilitet inngår, mens drøftingen av undersøkelsens validitet i lys av disse kravene vil komme i kapittel 5.

3.5.1 Statistisk validitet

Statistisk validitet omhandler sammenheng mellom variablene, hvorvidt de korrelerer og i så fall hvor sterk denne korrelasjonen er (Shadish et al., 2002). I undersøkelser med svak statistisk validitet øker sannsynligheten for å begå både Type I-feil, som vil si å forkaste en sann nullhypotese, og å begå en Type II-feil, som vil si å akseptere en gal nullhypotese. Typiske trusler for den statistiske validiteten vil være brudd på statistiske forutsetninger og

lav statistisk styrke (Lund, 2002). Brudd på den statistiske forutsetninger øker sjansen for å begå både Type I- og Type II-feil, og kan skyldes at utvalget bryter med forutsetningen om normalitet og lik varians. Som det nevnes i 1.4 *Avgrensning* kan brudd på statistiske forutsetninger også oppstå om man «fisker» etter resultater ved å foreta signifikanstesting av mange korrelasjoner. Med statistisk styrke menes en tests evne til å oppdage sammenhenger som finnes i populasjonen, og statistisk styrke omtales ofte som sannsynligheten for at en test vil avvise en gal nullhypotese (Shadish et al., 2002). Om en undersøkelse har lav statistisk styrke vil da sjansen for å begå en Type II-feil øke. Sentrale trusler i denne sammenhengen er nært knyttet til studiens reliabilitet, og handler om i hvilken grad data er fri for tilfeldige målingsfeil, eller mer generelt hvor pålitelig målingene er. Med tilfeldige målingsfeil menes ikke det nødvendigvis at feilene er forårsaket av tilfeldigheter, men at de opptrer tilfeldig. Svak reliabilitet kan også skyldes upresise datainnsamlingsmetoder. Når det gjelder forholdet mellom reliabilitet og validitet er ikke dette et motsetningsforhold hvor man må velge å prioritere det ene eller andre. Å vektlegge god reliabilitet i en undersøkelse vil ikke gå utover validiteten (Kleven, 2002a).

3.5.2 Indre validitet

For å oppnå høy grad av indre validitet forutsetter det at vi kan vise til en kausal sammenheng mellom uavhengig og avhengig variabel, og at denne sammenhengen ikke kan forklares av andre forhold (Shadish et al., 2002). I gruppesammenligningstudier som denne, hvor formålet er å sammenligne to undergruppers resultater på avhengige variabler, matematikkferdigheter, og se om forskjeller i resultatene kan forklares av uavhengig variabel, i dette tilfellet kjønn, vil det være svært vanskelig å trekke helt sikre konklusjoner. Årsaken er at i en ikke-eksperimentell undersøkelse kan en statisk sammenheng alltid kunne forklares med flere mulige kausalrelasjoner (Kleven, 2002b). Sentrale trusler i denne undersøkelsen vil være tredjevariabler og retning. Med retningsproblemet menes at hva som er årsak og hva som er virkning vil være umulig å fastslå. Kjønn kan i teorien være en årsak til forskjell i variasjonen mellom gutter og jenter, altså at det skyldes iboende egenskaper som er ulike mellom gutter og jenter. Men kjønn kan også være en konsekvens i den forstand at gutter i teorien kan ha blitt fulgt opp bedre i matematikkundervisningen fordi læreren har en tanke om at gutter har mer anlegg for matematikk, og at det derfor er viktigere å prioritere gutter i undervisningen.

På denne måten vil gutter prestere bedre, ikke som følge av indre egenskaper, men som følge av ytre oppfølging. Tredjevariabel som trussel vil være alle andre mulige årsaker til variasjon mellom kjønnene enn det at de er ulike kjønn. Eksempelvis vil det kunne oppstå variasjoner som skyldes en ubevisst holdning hos testleder som gjør at det ene kjønn konsekvent tillegges godvilje i svarene slik at skårene ikke blir kjønnsnøytrale.

3.5.3 Begrepsvaliditet

En sentral trussel i pedagogisk forskning vil være begrepsvaliditeten. Vi studerer teoretiske begreper som i seg selv ikke er direkte observerbare, og begrepsvaliditeten vil derfor henge sammen med i hvilken grad man har lyktes med begrepsoperasjonaliseringen (Kleven, 2002a). I følge Kleven (2002a) er selve kjernen i målingsproblemet i fag som pedagogikk nettopp det at man må benytte seg av synlige indikatorer for å måle abstrakte begreper. Begreper som skal operasjonaliseres i denne undersøkelsen er blant annet *kjønn*, og *matematikkferdigheter*, mer spesifisert *biologisk primære kvantitative evner*, *biologisk sekundære kvantitative evner*, *tekstoppgaver*, *kjennskap til tall og aritmetikk*. Det er viktig å merke seg at begrepsvaliditet ikke bare gjelder operasjonaliseringer av måleinstrumenter, men det gjelder også andre operasjonaliseringer, som for eksempel utvalg og kontekst. Dette henger tett sammen med ytre validitet. Dersom for eksempel ikke utvalget er operasjonalisert på en valid måte og har svak begrepsvaliditet, kan man heller ikke generalisere resultatene til målpopulasjonen. Også reliabiliteten spiller inn her; en test med dårlig reliabilitet kan aldri ha god begrepsvaliditet, fordi den som følge av dårlig operasjonalisering inneholder «støy» og dermed måles ikke rene begreper (Kleven, 2002a).

3.5.4 Ytre validitet

For at en undersøkelse skal ha god ytre validitet forutsetter det at slutninger som trekkes har en overføringsverdi. Overføringen kan være fra smal til bred, altså at variasjon man finner i en liten gruppe også vil gjelde for en større gruppe. Eller overføringen kan gjelde fra bred til smal, eksempelvis at variasjon man finner i en gruppe også vil gjelde for en mindre gruppe eller et enkelt individ. Det kan også være tale om overføring til en tilsvarende gruppe, altså at det man finner i et utvalg også vil gjelde for et lignende utvalg (Shadish et al., 2002) Når det gjelder ytre validitet, er det vanskelig å oppnå i kun én undersøkelse, og man kan gjøre flere

undersøkelser for å styrke den ytre validiteten i resultatene (Lund, 2002). For generalisering er det viktig at gruppen med individer som er testet er representative for gruppen det skal generaliseres til, i dette tilfelle lavtpresterende førsteklasinger.

3.6 Analyser

Av de 12 deltestene i Prosjektet er det som nevnt 5 stykker jeg har benytte meg av i mine analyser. Testen TOBANS symbolsk/ikke-symbolsk er delt i to med separate items og resultater; symbolsk - *digit comparison* og ikke-symbolsk - *dot comparison/ANS*. For å få variabler som best måler de ulike matematikkområdene som er redegjort for tidligere i oppgaven, er enkelte items fra måleinstrumentene satt sammen til nye variabler. Tabell 3.1 viser hvilke items fra de ulike måleinstrumentene som er benyttet i ulike variablene. Som det fremkommer av tabellen er målene kategorisert i henhold til teorien om biologisk primære og biologisk sekundære kvantitative evner som det ble redegjort for i kapittel 2.

Tabell 3.1 Måleinstrumentene for matematikkferdigheter

Område	Variabel	Måleinstrument
Biologisk primære kvantitative evner	ANS	TOBANS Dot comparison
	Tekstoppgaver	ThinkMath oppgave 73-80 WISC-IV Regning
Biologisk sekundære kvantitative evner	Kjennskap til tall	ThinkMath oppgave 15-72 TOBANS Digit comparison
	Aritmetikk	ThinkMath oppgave 81-100
	Regneflyt	TOBANS addisjon

Kontrollvariablene *Nonverbal IQ* og *Eksekutive funksjoner* måles av henholdsvis deltestene Raven og Tower of London og det er dermed ikke behov for å gjøre endringer for å få spissede variabler.

Tabell 3.2 Måleinstrumentene for kontrollvariablene

Område	Variabel	Måleinstrument
Eksekutive funksjoner	Eksekutive funksjoner	Tower Of London
Nonverbal IQ	Nonverbal IQ	Raven

3.7 Etiske hensyn

Denne oppgaven er en del av et større prosjekt, og denne oppgavens etiske hensyn vil i mange henseender være knyttet opp mot de føringer for etiske hensyn som ligger i dette Prosjektet. De formelle kravene, som innsamlingen av samtykke fra informanter og meldt studien inn til Norsk Senter for Forskningsdata (NSD), er utført av Prosjektet. I dette tilfellet, hvor informantene er barn under 15 år, er det foreldrene som har gitt sitt samtykke til barnas deltakelse (NESH, 2016). Likevel er en også avhengig av at barna selv er villige til å delta i undersøkelsen, både fordi det er de som skal utføre de nødvendige oppgavene i testene og fordi det ville være uetisk å presse barna til å delta i noen de ikke ønsker.

I en forskningsundersøkelse der en behandler sensitive opplysninger om barn er det spesielt viktig at etiske hensyn ivaretas, den informasjon man som forsker får om deltakeren må bli behandlet konfidensielt, slik at bruken og formidlingen av informasjonen ikke skader deltakeren (NESH, 2016). Om enkeltresultater kobles opp mot den enkelte deltakeren kan dette ha negative konsekvenser for individet. Det er derfor viktig at forskeren er redelig og ikke benytter materialet til annet enn det informantene har gitt sin tillatelse til. Som forsker må en ta hensyn til barnets egen vilje ved deltakelse. Barn er en sårbar informantgruppe, både fordi det ikke er de selv har tatt avgjørelsen om deltakelse, men også fordi barn heller ikke nødvendigvis har den samme kunnskapen om konsekvensene av å gi forskerne informasjon. Kravet om konfidensialitet gjelder derfor særskilt ved barns deltakelse i forskning (NESH, 2016). Hos Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora understrekes også et annet hensyn når barn er informanter; *«Barn har et spesielt krav på beskyttelse, og ved mistanke om for eksempel mishandling eller omsorgssvikt har forskeren også en opplysningsplikt og må melde fra til barnevernet. Dette gjelder for alle og uten hensyn til taushetsplikt.»*(NESH, 2016).

4. Resultater

I dette kapittelet vil temaet «Kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger» belyses ved å besvare forskningsspørsmålene:

- 1. Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?*
- 2. Er det kjønnsforskjeller når non-verbal IQ og eksekutive funksjoner er kontrollert for?*

Gjennomsnitt, standardavvik og verdier i de ulike måleinstrumentene vises her i tabell 4.1 med de rene målene sammen med skjevhet, kurtosis og reliabilitetsanalyser. Deretter følger en kort vurdering av disse målene før korrelasjonen mellom måleinstrumentene presenteres i tabell 4.2. Som vist i tabell 3.1 i forrige kapittel er ulike items fra måleinstrumentene satt sammen til variabler som måler biologisk primære og biologisk sekundære kvantitative evner, hvorav den siste er tredelt. Disse fire variablenes fordeling og reliabilitet vises i tabell 4.3 og følges så av en vurdering av hver enkelt variabel. Deretter følger en vurdering av fordeling og reliabilitet i kontrollvariablene IQ og eksekutive funksjoner. Etterfulgt av de deskriptive analysene presenteres en korrelasjonsanalyse av de fire variablene og kjønn, for de samme variablene analyseres i en hierarkisk multippel regresjonsanalyse hvor IQ og eksekutive funksjoner kontrolleres for.

Kapittelet avsluttes med en oppsummering av funn.

4.1 Deskriptiv analyse

I tabell 4.1 presenteres måleinstrumentene. Tabellen viser en oversikt over de ulike testenes gjennomsnitt, standardavvik, skjevhet, kurtosis og reliabilitet målt i Cronbach's alpha verdier. Som nevnt i metodekapittelet slås disse sammen til mål for henholdsvis biologisk primære, og biologisk sekundære kvantitative ferdigheter, hvor den siste kategorien igjen er delt i tre underkategorier; Aritmetikk, Tekstoppgaver og Kjennskap til tall.

Tabell 4.1 Måleinstrumentene

Tester	N	M	SD	Skew	Krt	C alpha
ThinkMath	120	43.19	13.212	.142	.289	.925
W-Regning	120	9.43	2.786	.315	.661	.743
Digit comp.	120	13.95	3.728	.000	.072	.866
Dot comp.	120	9.00	2.663	-.497	-.051	.701
Raven	120	17.71	4.495	.001	-.124	.742
TOL	120	7.99	3.73	-.530	-.814	.821

Som det fremkommer av tabell 4.1 er verdiene for alle måleinstrumentene i større eller mindre grad tilnærmet normalfordelte og har god eller tilfredsstillende reliabilitet. I en normalfordeling vil skjevheten ligge nær 0, og jo nærmere +1 eller -1, jo skjevere vil fordelingen være. For kurtosis gjelder samme prinsipp om nærhet til null, men her sier vi at jo nærme +1, jo spissere er fordelingen, og om den er nærmere -1, jo flatere er fordelingen. Da items fra de ulike testene vil slås sammen til variabler som mer presist måler egenskapene som omtales i forskningsspørsmålet, vist i tabell 3.1, vil ikke måleinstrumentenes skjevhet og kurtose beskrives nærmere. Den testen som tabellen viser har størst avvik i skjevhet og kurtose, Tower Of London blir ikke analysert for kjønnsforskjeller og er sammen med Raven en kontrollvariabel.

4.2 Korrelasjonsanalyse av måleinstrumentene

I tabell 4.2 vises korrelasjonen mellom de ulike måleinstrumentene.

Tabell 4.2 Korrelasjon mellom måleinstrumentene

	ThinkMath	W-Regning	Digit comp.	Dot comp.	Raven	TOL
ThinkMath	1	.699**	.500**	.235**	.477**	.211*
W-Regning	.699**	1	.429**	.265**	.512**	.275**
Digit comp.	.500**	.429**	1	.484**	.304**	.183*
Dot comp.	.235**	.265**	.484**	1	.295**	.234**
Raven	.447**	.512**	.304**	.295**	1	.387**
TOL	.211*	.275*	.183*	.234**	.387**	1

** . Korrelasjon er signifikant på .01-nivå (2-halet)

* . Korrelasjon er signifikant på .05-nivå (2-halet)

Tabell 4.2 viser at det er høy grad av korrelasjon mellom de ulike måleinstrumentene og alle korrelasjonene mellom matematikkmåleinstrumentene er signifikante på .01 nivå. I de videre analysene vil alle items fra WISC-Regning være slått sammen med 8 items fra ThinkMath til variabelen *Tekstoppgaver*. Korrelasjon mellom disse to variablene er på .699 og korrelasjonen er signifikant på .01- nivå. Dette er et solid empirisk grunnlag for å slå sammen disse to variablene uten at det vil true begrepsvaliditeten. Digit comparison vil i de videre analysene være slått sammen med item 15-72 i ThinkMath, og som vi ser av tabellen har de en korrelasjon på .500, som regnes som moderat, og signifikant på .01-nivå. Det kan dermed også forsvares empirisk å slå disse itemsene sammen til variabelen *Kjennskap til tall*. Den siste av de tre variablene under biologisk sekundære kvantitative evner er *Aritmetikk*. Denne variabelen består av kun items fra ThinkMath. Totalt sett er det høy eller moderat korrelasjon som er signifikant på .01-nivå mellom de tre måleinstrumentene ThinkMath, WISC-Regning

og Digit comparison, og sammen med teori fra kapittel 2 foreligger det dermed et empirisk grunnlag for å slå sammen de nevnte items til tre ulike variabler som hører inn under biologisk sekundære kvantitative evner. Den matematikkvariabelen som har lavest korrelasjon med de andre, Dot comparison måler ANS og vil dermed være undersøkelsens mål på biologisk primære kvantitative evner.

4.3 Variablenes fordeling og reliabilitet

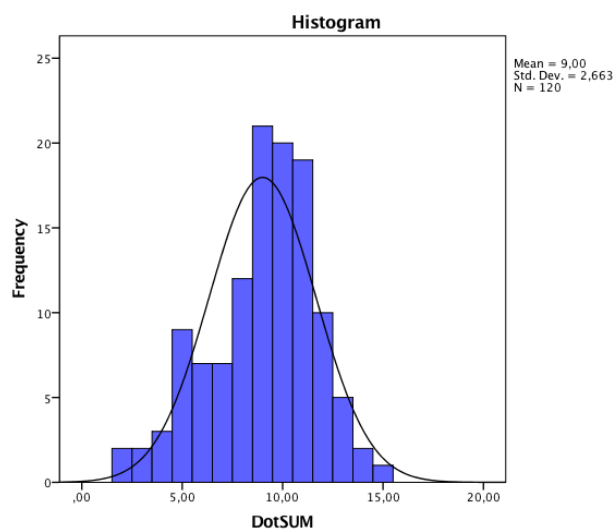
I tabell 4.3 presenteres gjennomsnitt, standardavvik, skjevhet og kurtosis, samt Cronbach alpha for variablene som er undersøkt for kjønnsforskjeller.

Tabell 4.3 Variabler for matematikkferdigheter

	Område	N	M	SD	Skew	Krt	C alpha
Primære kvantitative evner	ANS	120	9.000	2.663	-.497	-.051	.701
	Tekstoppgaver	120	11.258	3.912	.392	.238	.799
Sekundære kvantitative evner	Kjennskap til tall	120	40.800	10.054	.114	.272	.903
	Aritmetikk	120	4.942	3.952	.423	-.166	.879

4.3.1 Vurdering av variabelen ANS

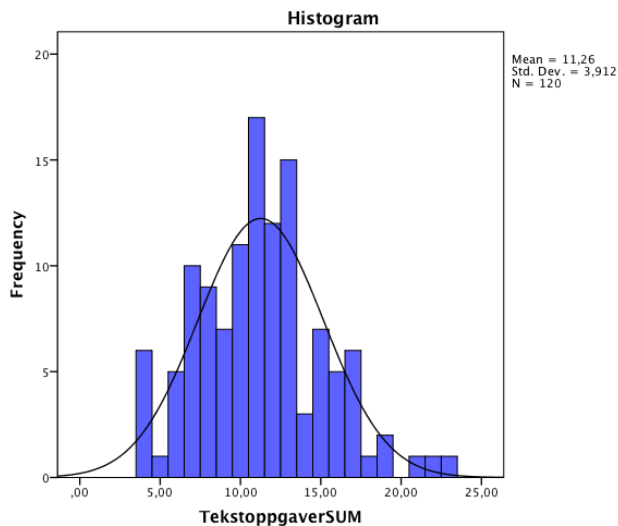
Approximate number system, som defineres under det som hos Geary (1995a) kalles *biologisk primære kvantitative evner* måles her av testen «TOBANS Dot comparison».. Denne målingen har en Cronbach's alpha på .701, hvilket vil si at testen har en akseptabelt høy reliabilitet. Som det fremkommer av histogrammet under, viker resultatet noe fra å være normalfordelt, med en skjevheten i denne variabelen på $-.497$ som vil si at fordelingen er tydelig venstreskjev og altså bærer preg av en takeffekt. Med kurtosis på $-.051$ er fordelingen ikke spesielt flat.



Figur 4.1 Histogram som viser fordeling på «ANS»

4.3.2 Vurdering av variabelen Tekstoppgaver

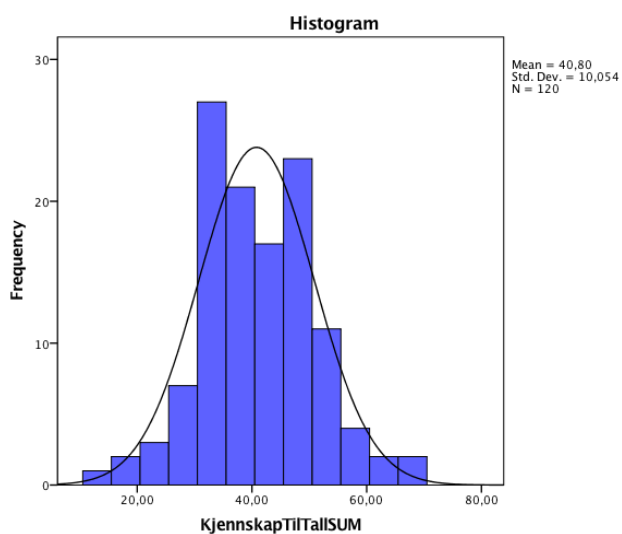
Tekstoppgaver tilhører biologisk sekundære kvantitative evner, og består av alle items fra *WISC-IV Regning* sammen med de 8 oppgavene i *ThinkMath* som måler problemløsningsferdigheter av typen man finner i tekstoppgaver. Disse oppgavene leses opp for elevene i testsituasjonen og eleven skal i *ThinkMath* tegne svaret som riktig antall objekter eller skrive ned svaret i form av et tall. I *WISC-IV Regning* skal eleven avgi svaret muntlig. Som det fremkommer har denne variabelen Cronbach's alpha på .799, og reliabiliteten er derfor å anse som god. Fordelingen er, som det vises i figur 2, relativt normalfordelt. Skjevhet er $.392$, så den drar seg litt mot å være høyreskjev. Kurtosis er på $.238$ og kurven er dermed ikke spesielt spiss.



Figur 4.2 Histogram som viser fordeling på «Tekstoppgaver»

4.3.3 Vurdering av variabelen Kjennskap til tall

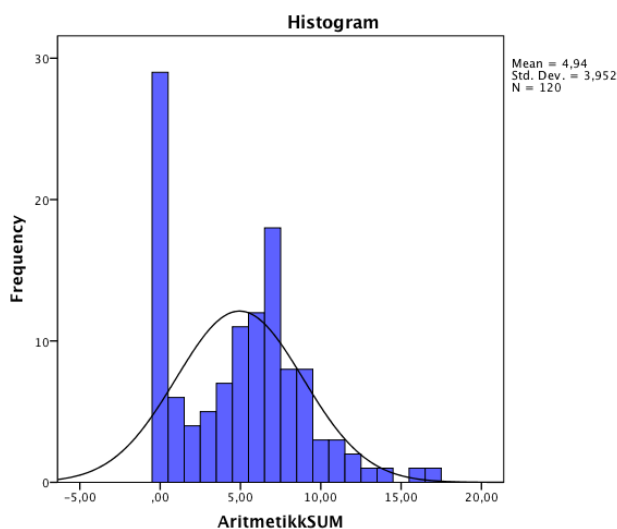
Variabelen for *Kjennskap til tall* tilhører biologisk sekundære kvantitative evner, og består av oppgavene 15-72 i *ThinkMath*, som alle måler nettopp kjennskap til tall og telleferdigheter, samt *TOBANS Digit comparison* som også måler kjennskap til tall. Denne variabelen er tilnærmet normalfordelt og har en Cronbach's alpha på .903 som betyr en svært god reliabilitet. Skjevhet er .114, mens variabelens kurtosis er .272 og altså ikke veldig høyreskjev, og heller ikke spesielt spiss.



Figur 4.3 Histogram som viser fordeling på «Kjennskap til tall»

4.3.4 Vurdering av Aritmetikk

I denne variabelen, som også tilhører biologisk sekundære kvantitative evner, består av items 81-100 i ThinkMath og er 10 addisjon- og 10 subtraksjonoppgaver. Cronbach's alpha for denne variabelen er .879 og vi kan derfor si at testen har god reliabilitet, men som vi ser av histogrammet for denne variabelen er det en tydelig gulveffekt med en skjevheten på .423. Dette skyldes sannsynligvis at mange av elevene på førstetrinn ikke ennå har lært grunnleggende aritmetikk, spesielt subtraksjon, og dermed er det svært mange med veldig lave skårer på disse oppgavene.



Figur 4.4 Histogram som viser fordeling på «Aritmetikk»

4.4 Vurdering av kontrollvariablene

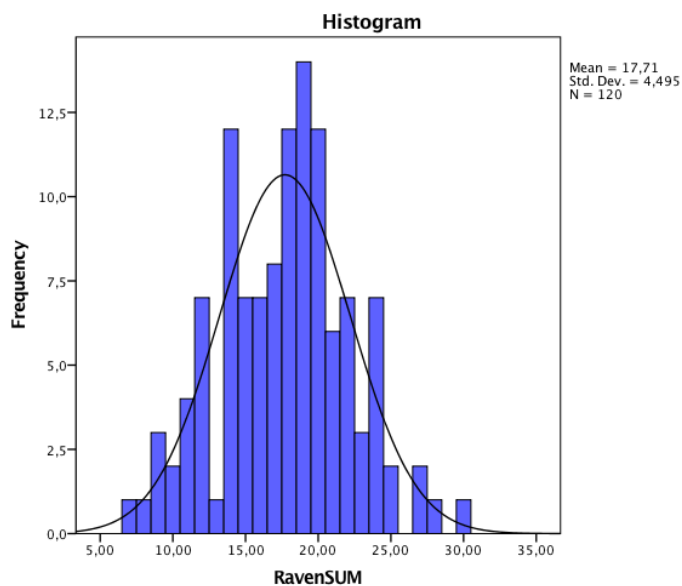
Nonverbal IQ og eksekutive funksjoner er målt ved hjelp av henholdsvis Raven og Tower Of London. Disse variablene benyttes i analysene for å se om kjønnsforskjeller finnes etter at IQ og eksekutive funksjoner er kontrollert for. De vanligste intelligenstestene er designet på en slik måte at det ikke vil være kjønnsforskjeller i IQ-skårene (Neisser et al., 1996) og kjønnsforskjeller i IQ og eksekutive funksjoner er ikke denne studiens anliggende. Kjønnsforskjeller på disse områdene er dermed ikke analysert. Her følger histogrammer og kommentarer av de to variablenes fordeling og reliabilitetskoeffisient.

4.4.1 IQ

Variabelen for IQ slik den fremstår målt med testen Raven er relativt normalfordelt.

Resultatene viser en skjevhet på .001 og variabelen er dermed hverken høyre- eller venstreskjev. Med kurtosis på $-.124$ kan man si at kurven er en anelse flat.

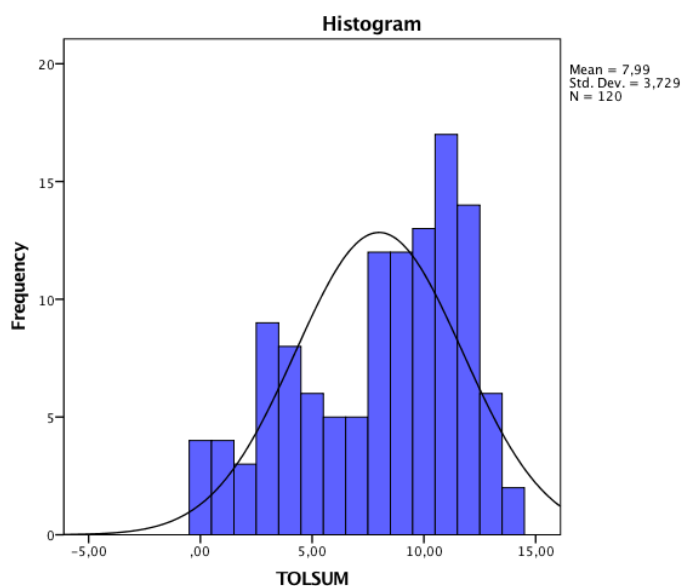
Reliabilitetskoeffisient i Cronbach's alpha er på $.742$. og denne testen har dermed en tilfredsstillende reliabilitet.



Figur 4.5 Histogram for fordeling på Raven

4.4.2 Eksekutive funksjoner

Eksekutive funksjoner er målt ved hjelp av testen Tower Of London, og ut fra figur 4.6 ser det ut til at verdiene er preget av en ikke ubetydelig takeffekt på testen. For TOL er skjevheten $.530$. Kurtosis er $.814$ og fordelingen er altså ganske venstreskjev og svært spiss sammenlignet med en normalfordeling. Reliabiliteten for denne målingen er god med Cronbach's alpha $.821$.



Figur 4.6 Histogram for fordeling på Tower Of London

4.5 Deskriptiv analyse av variablene fordelt på kjønn

Delt inn etter kjønn ser resultatene slik ut:

Tabell 4.4 Kjønnforskjeller i variablene med effektstørrelse uttrykket i Cohens d

		Gutter			Jenter			Cohens d
		N	M	SD	M	SD	N	
Primære kvantitative evner	ANS	68	8.471	2.668	9.692	2.517	52	-.467*
	Tekstoppgaver	68	10.868	4.239	11.769	3.411	52	-.231
Sekundære kvantitative evner	Kjennskap til tall.	68	41.015	11.347	40.519	8.159	52	.0491
	Aritmetikk	68	4.868	4.435	5.038	3.254	52	-.0428
	Sekundære kvantitative e.	68	56.750	18.241	57.327	12.295	52	-.0362

* Korrelasjonen er signifikant på .05 nivå med to-halet test

4.5.1 Kjønnforskjeller i biologisk primære kvantitative evner

I analysen av biologisk primære kvantitative evner, her representert ved ANS som måles av testen TOBANS Dot comparison, viser sumskårene at jenter samlet sett har en gjennomsnittsskåre på 9.69 mens guttenes gjennomsnitt er på 8.47. Med et standardavvik for jentene på 2.52, og 2.67 for guttene viser en t-test for uavhengige utvalg at det med to-halet test er en signifikans på .012 og altså signifikant på .05 nivå, mens effektstørrelsen er Cohen's d , på -.467. At effektstørrelsen fremstår som negativ betyr i dette tilfellet at den er i jentenes favør fordi tallene for gutter, som er lavere, er lagt inn først i utregningen. En effektstørrelse, Cohen's d , på 0.47 betyr at 68 % av jentene har prestert bedre enn meanskåren for guttene, 81 % av de to gruppene vil overlapse, og det er 63% sjans for at en tilfeldig trukket jente vil ha en høyere skåre enn en tilfeldig trukket gutt. Det er vanlig å regne effektstørrelser på .5 som moderate, og vi kan dermed si at i *Primære kvantitative evner* er det signifikante kjønnforskjeller på .05-nivå og effektstørrelsen er moderat.

4.5.2 Kjønnforskjeller i Tekstoppgaver

I *Tekstoppgaver* viser en t-test for uavhengige utvalg at det ikke er signifikante forskjeller på gutter og jenters prestasjoner. Signifikansen er på .212, og det vil si at det er 21% sjans for at vi tar feil om vi sier at ulikhetene i gjennomsnitt og standardavvik for gutter og jenter skyldes kjønnforskjeller. Effektstørrelsen for variasjonen her er $d = -0.23$, at den er negativ vil si at den er i jenters favør. Hvilket betyr at 59% av jentene skårer bedre enn meanskåren til guttene og at det er 91% overlapp mellom jenter og gutter. Om man plukker en vilkårlig jente er det bare 56% sjans for at hun har høyere skår enn en tilfeldig utvalgt gutt.

4.5.3 Kjønnforskjeller i Kjennskap til tall

En t-test for uavhengig utvalg viser at det i variabelen *Kjennskap til tall* gir oss en signifikans på .790, og det er dermed overveiende sannsynlig at forskjellen på gutter og jenters mean og standardavvik ikke skyldes kjønn, men tilfeldigheter. Effektstørrelsen uttrykket i Cohens $d = .049$ vil i tillegg si at kun 52% av guttene vil ha skåret høyere enn meanskåren til jenter og hele 98% av de to gruppene vil overlapse.

4.5.4 Kjønnforskjeller i Aritmetikk

Heller ikke i Aritmetikk er det resultater som peker i retning av at det er kjønnforskjeller. Resultatene i en t-test for uavhengige utvalg viser at signifikansen i forskjellen mellom mean og standardavvik for gutter og jenter er .816, og dermed er det overveiende sannsynlig at denne forskjellen skyldes tilfeldigheter. Om man skulle hevde at forskjellen i variansen mellom gutter og jenter skyldes kjønn vil det være hele 82% sjanse for at man tar feil og altså begår en Type I-feil. For Aritmetikk er effektstørrelsen $d = -.0428$, og også her vil den negative effektstørrelsen komme av at den er i favør av jentene. Denne effektstørrelsen vil si at også her har kun 51% av jentene skåret høyere enn guttenes meanskår, og det er kun 51% sjanse for at en tilfeldig utvalgt jente har høyere skår enn en tilfeldig utvalgt gutt.

4.5.5 Kjønnforskjeller i biologisk sekundære kvantitative evner

Når man legger sammen sumskårene for de tre variablene som hører til under Biologisk sekundære kvantitative evner viser en t-test signifikans på .845, altså at det ikke er signifikant sammenheng mellom variasjonen i meanskårer og standardavvik for gutter og jenter. Denne variasjonen har en effektstørrelse på -0.362 , hvilket vil si at den lille variasjonen som vises er i favør av jenter. En effektstørrelse på $-.0362$ betyr at kun 51% av jenten hadde en høyere skår enn guttens mean, og ved å trekke en tilfeldig jente er det bare 51% sjanse for at hun har høyere skår enn en tilfeldig utvalgt gutt. Hele 99% av de to gruppene vil overlappe.

4.6 Korrelasjonsanalyse

Ulike ferdigheter i matematikk henger tett sammen. Som tabell 4.5 viser, korrelerer alle matematikkvariablene signifikant med hverandre, noe som heller ikke var uventet ut fra undersøkelsens utvalgsstørrelse. I en slik sammenheng anses 120 elever som et relativt stort utvalg, og det er ikke uvanlig at variabler i store utvalg korrelerer. De fleste variablene her korrelerer på .01 nivå og det er kun *Primære kvantitative evner* og *Aritmetikk* som korrelerer på .05 nivå. Ser vi derimot på hvordan *Kjønn* korrelerer med de ulike matematikkferdighetene er det kun signifikant korrelasjon med *Primære kvantitative evner*. Signifikansen på korrelasjonen mellom *Kjønn* og *Primære kvantitative ferdigheter* er .012, og altså signifikant på .05 nivå.

Tabell 4.5 Korrelasjon mellom variablene uttrykket i Pearson *r*

	ANS	Tekstoppgaver	Kjennskap til tall	Aritmetikk	Kjønn
ANS	1	.298**	.297**	.215*	.228*
Tekstoppgaver	.298**	1	.633**	.611**	.115
Kjennskap til tall	.297**	.633**	1	.649**	-.025
Aritmetikk	.215*	.611**	.649**	1	.022
Kjønn	.228*	.115	-.025	.022	1

** . Korrelasjon er signifikant på .01 nivå (2-halet)

* . Korrelasjon er signifikant på .05 nivå (2-halet)

4.7 Regresjonsanalyse

Sammenhengen mellom variabelen *Kjønn* og de fire variablene for ulike matematikkferdigheter (1)*Primære kvantitative evner*, (2)*Tekstoppgaver*, (3)*Kjennskap til tall* og (4)*Aritmetikk* er analysert ved hjelp av regresjonsanalyser. Regresjon er en statistisk metode som benyttes når en skal studere sammenhengen mellom ulike variabler (Pedhazur, 1997). I denne studien vil kjønn være den uavhengige variabelen, den avhenger altså ikke av de andre variablene, mens de fire ulike matematikkvariablene vil være avhengige variabler fordi de kan antas være relatert til uavhengig variabel.

4.7.1 Regresjonsanalyse av ANS

Tabell 4.6 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på ANS. N=120

Steg	Variabel	Multipel korrelasjon	Proporsjonsforklart varians	Signifikans
1	Eksekutive funksjoner & IQ	.322	.104	.002
2	Kjønn	.398	.055	.007

ANS er den eneste ferdigheten hvor det i følge t-testen er signifikante kjønnsforskjeller på gjennomsnittskårene. På tabell 4.6 ser vi at sammen med IQ og eksekutive funksjoner forklarer kjønn 15,8% av variasjonene i *Primære kvantitative evner*. Når vi så kontrollerer for IQ og eksekutive funksjoner forklarer kjønn alene bare 5.5% . Selv om graden av påvirkning av variasjonen er liten, er den likevel signifikant.

4.7.2 Regresjonsanalyse av Tekstoppgaver

Tabell 4.7 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Tekstoppgaver. N=120

Steg	Variabel	Multipel korrelasjon	Proporsjonsforklart varians	Signifikans
1	Eksekutive funksjoner & IQ	.489	.239	.000
2	Kjønn	.503	.014	.148

Kjønn forklarer i følge denne analysen cirka 1.3% av variasjonene i *Tekstoppgaver* etter at Eksekutive funksjoner og IQ er kontrollert for. Totalt forklarer kjønn sammen med IQ og Eksekutive funksjoner cirka 25% av variasjonen i *Tekstoppgaver*.

4.7.3 Regresjonsanalyse av Kjennskap til tall

Tabell 4.8 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Kjennskap til tall. N=120

Steg	Variabel	Multipel korrelasjon	Proporsjonsforklart varians	Signifikans
1	Eksekutive funksjoner & IQ	.429	.184	.000
2	Kjønn	.430	.001	.790

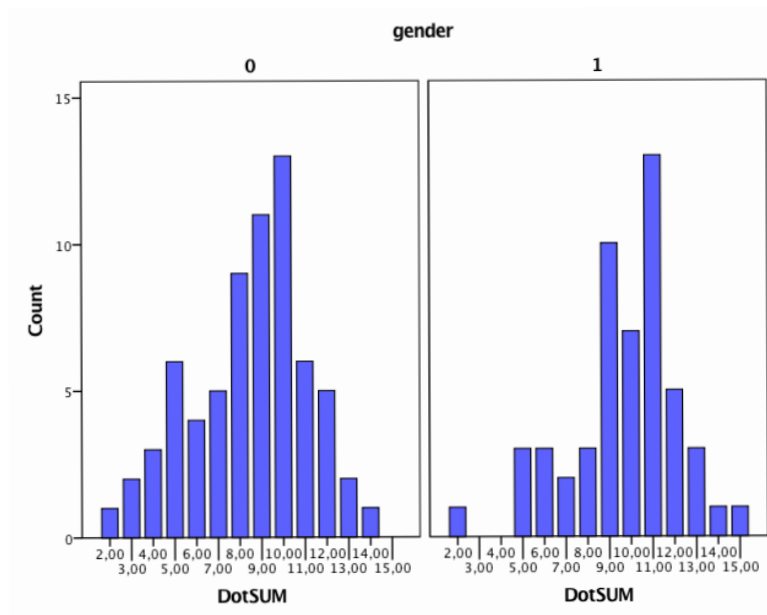
I variabelen Kjennskap til tall viser regresjonsanalysen at kjønn kun forklarer 0.1% av variansen etter at IQ og eksekutive funksjoner er kontrollert for. IQ og eksekutive funksjoner på sin side forklarer cirka 18% av variasjonen.

4.7.4 Regresjonsanalyse av Aritmetikk

Tabell 4.9 Regresjonsanalyse som viser forklaringseffekten av «Eksekutive funksjoner & IQ» og kjønn på Aritmetikk. N=120

Steg	Variabel	Multipel korrelasjon	Proporsjonsforklart varians	Signifikans
1	Eksekutive funksjoner & IQ	.354	.126	.000
2	Kjønn	.355	.000	.816

Heller ikke sammenhengen mellom kjønn og Aritmetikk var signifikant med t-test for uavhengige utvalg. Som det fremkommer av tabell 4.9, forklarer kjønn 0% av variansen i Aritmetikk etter at IQ og eksekutive funksjoner er kontrollert for. I følge denne analysen forklarer IQ og eksekutive funksjoner cirka 12% av variasjonen i Aritmetikk.



Figur 7. Sidelitte histogram med fordeling på ANS fordelt på kjønn.
*0 = gutter, 1= jenter.

Figur 7 viser histogrammer som viser fordelingen på ANS, og det kan se ut som om guttene ligger nærmere en normalfordeling enn det jentene gjør. Totalt sett er denne variabelen ganske venstreskjev, og etter histogrammene å dømme er det jentenes skårer som i størst grad bidrar til denne takeffekten.

4.8 Oppsummering av analyser og funn

I kapittel 4 har først de deskriptive analysene blitt presentert i form av en tabell som viser en oversikt over de ulike testenes gjennomsnitt, standardavvik, skjevhet, kurtosis og reliabilitet målt i Cronbach's alpha-verdier. I denne tabellen kommer det frem at variablene som avviker mest fra normalfordeling er ANS og Aritmetikk, som preges av henholdsvis en takeffekt og en gulveffekt.

Gjennom en t-test hvor gutter og jenters mean er sammenlignet viser resultatene at signifikante kjønnsforskjeller kun finnes i ANS, og at forskjellen er i jenters favør. Effektstørrelsen er på $d = .47$ og er dermed å regne for moderat. I de tre andre variablene som er analysert for kjønnsforskjeller ved hjelp av en t-test finnes det ingen signifikante forskjeller mellom kjønnene.

Materialet er videre analysert ved hjelp av regresjonsanalyser hvor resultatene viser at at kjønn i liten grad forklarer variasjonen i de ulike matematikkferdigheter når IQ og eksekutive funksjoner er kontrollert for. Størst påvirkning i variasjon har kjønn på ANS, hvor det forklarer 5.5% etter at IQ og eksekutive funksjoner er kontrollert for.

5. Drøfting av resultater

Denne undersøkelsen har til hensikt å besvare følgende forskningsspørsmål:

- 1. Er det kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger på oppgaver som har til hensikt å måle henholdsvis biologisk primære og biologisk sekundære kvantitative evner?*
- 2. Er det kjønnsforskjeller når non-verbal IQ og eksekutive funksjoner er kontrollert for?*

I dette kapittelet vil problemstillingene og resultatene først drøftes i sammenheng med Cook & Campbells (1979) validitetssystem hvor også reliabilitet vil inngå. Resultatene av undersøkelsen vil så bli diskutert i lys av tidligere teori og empiri før oppgaven avsluttes med tanker om pedagogiske konsekvenser og mulige videre undersøkelser

5.1 Undersøkelsens validitet og reliabilitet

I tråd med Cook & Campbells validitetssystem (Cook et al., 1979) blir denne drøftingen presentert i lys av de fire validitetstypene. Reliabilitet vil i bli behandlet i hovedsak under statistisk validitet.

5.1.1 Statistisk validitet

Statistisk validitet handler om to beslektede statistiske slutninger vedrørende korrelasjonen i kausale slutninger; (1) hvorvidt den antatte årsak-virkning korrelerer og (2) hvor sterkt de korrelerer (Shadish et al., 2002). I dette tilfelle dreier den statistiske validiteten seg om hvorvidt det er forskjeller mellom to grupper, og hvor store forskjellene i så fall er. I denne undersøkelsen er kjønnsforskjeller analysert ved hjelp av t-test for uavhengige utvalg som på variabelen for biologisk primære kvantitative evner, ANS, viser signifikans på .05-nivå med en effektstørrelsen på Cohen's $d = -.467$. En effektstørrelse på .5 anses som middels (Gall, Gall, & Borg, 2003). Her er altså forskjellen signifikant og rimelig stor. Ut fra en nullhypotese signifikantesting er et vanlig krav at resultatet er statistisk signifikant hvis $p < .05$, hvilket er tilfelle i denne undersøkelsen (Shadish et al., 2002). Resultatet her kan tolkes på to ulike måter. Man kan konkludere med at det er kjønnsforskjeller i primære

kvantitative evner, med fare for å begå en Type I-feil, eller man kan konkludere med at forskjellene skyldes tilfeldigheter, men da med risiko for å begå en Type II-feil. Kjønnforskjellene er også kontrollert for IQ og eksekutive funksjoner gjennom en regresjonsanalyse. Disse analysene viser at kjønn forklarer 5.5% av variasjonene i primære kvantitative evner etter at IQ og eksekutive funksjoner er kontrollert for. I følge analyseresultatene kan vi altså si at det er 95% sannsynlighet for at det er kjønnforskjeller i primære kvantitative evner som ikke skyldes tilfeldigheter, men i og med at kjønn forklarer bare 5.5% av denne variasjonen vil usikkerheten rundt denne konklusjonen være større enn om kjønn hadde forklart mer (Shadish et al., 2002). I resultatene for ANS er det også en betydelig takeffekt, hvilket svekker den statistiske styrken fordi det gjør det vanskeligere å oppdage sammenhenger.

For de tre variablene Tekstoppgaver, Kjennskap til tall og Aritmetikk viser signifikanstesting at det ikke er noen signifikant sammenheng med kjønn, det samme gjelder når de tre er slått sammen til samlevariabelen Biologisk sekundære kvantitative evner. En to-halet t-test viser at signifikansen er på henholdsvis .21, .79, .82 og .85 for de fire variablene. Også disse resultatene kan tolkes på to måter; som at det ikke er noen sammenheng mellom de fire variablene og kjønn, eller som at det er en sammenheng som er blitt oversett fordi undersøkelsen har lav statistisk styrke. Kjennskap til tall, Aritmetikk og samlevariabelen Biologisk sekundære kvantitative evner har alle en signifikans på rundt .80, hvilket altså vil si at det er 80% sannsynlig at vi har rett når vi konkluderer med at det ikke er noen forskjell på gutter og jenters prestasjoner på disse områdene.

En potensiell trussel mot den statistiske validiteten gjør seg likevel gjeldene da variabelen for Aritmetikk er preget av en gulveffekt, hvilket gjør at det kan være variasjoner som ikke fremkommer av testen. Årsaken er at redusert variasjonsbredde svekker forholdet mellom den variabelen og andre variabler (Shadish et al., 2002). Når det gjelder effektstørrelsen i variasjonene for disse tre variablene er den statistisk ubetydelig og det bidrar til å styrke sjansen for at det er riktig å akseptere en H_0 for disse tre variablene. Variabelen Tekstoppgaver har betydelig høyere signifikans enn de tre andre, men den er fremdeles langt unna å være signifikant på .05-nivå. Også for Tekstoppgaver er effektstørrelsen liten, og

regresjonsanalysen viser at kjønn bare forklarer 1,3% av variasjonen etter at IQ og eksekutive funksjoner er kontrollert for. Det er dermed naturlig å akseptere en H0 også for denne variabelen.

Den statistiske styrken i en undersøkelse vil svekkes av dårlig reliabilitet (Shadish et al., 2002). Studiens reliabilitet handler om i hvilken grad data er fri for tilfeldige målingsfeil, eller mer generelt hvor pålitelig målingene er. Med tilfeldige målingsfeil menes ikke det nødvendigvis at feilene er forårsaket av tilfeldigheter, men at de opptrer tilfeldig (Shadish et al., 2002). Slike feil vil jevne seg ut over tid, men da dette er en gruppesammenligningstudie hvor målingene er utført på et tidspunkt kan dette være en reell trussel mot studiens reliabilitet. En slik trussel vil også kunne påvirke studiens begrepsvaliditet da måleresultatene ikke i samme grad vil måle det de utgir seg for, fordi resultatene skyldes i stor grad tilfeldigheter. Svakheter i målingene og målefeil som svekker forholdet mellom to variabler, kan være forårsaket av utrygge elever i testsituasjonen, eller upresise instruksjoner fra testleder, eller ved at testleder konsekvent noterer feil skåring av resultater på skåringsskjema. Utrygge elever kan skyldes at det oppleves skremmende eller ubehagelig å være i en situasjon der de blir bedt om å prestere gjennom å utføre oppgaver sammen med en voksenperson de ikke kjenner. Det kan føre til at de ikke presterer sitt beste og resultatene dermed blir påvirket av ubehag i testsituasjonen, hvilket vil gå utover undersøkelsens reliabilitet og dermed også undersøkelsens statistiske validitet (Shadish et al., 2002).

5.1.2 Indre validitet

I et eksperiment kan man stille seg spørsmål om det var stimulansen i eksperimentet som ga endringen i det spesifikke tilfellet (Shadish et al., 2002). Denne undersøkelsen er ikke-eksperimentell, hvilket bør utelukke at det har vært noe i undersøkelsen som har påvirket resultatene. Men det at undersøkelsen er ikke-eksperimentell gjør det også svært vanskelig å trekke konklusjoner om årsaksforholdet fordi det i en statistisk sammenheng alltid vil kunne være forårsaket av flere mulige kausalrelasjoner (Shadish et al., 2002). De små kjønnsforskjellene som viser seg i Primære kvantitative evner, ANS, kan i teorien skyldes svært mange andre årsaker enn barnas kjønn. For å styrke tilliten til at det er kjønn som er den mest sannsynlige årsaken til variasjonene i ANS kan man vurdere andre mulige

årsaksforklaringer. Jo flere tolkninger som kan elimineres som usannsynlig, jo sterkere er tilliten til den mest sannsynlige tolkningen av resultatet (Kleven, 2002b).

I dette tilfellet er det umulig å finne støtte i tidligere forskning som sier at det er kjønnsforskjeller til fordel for jenter i primære kvantitative evner, og dermed er et mer sannsynlig at forskjellene i denne undersøkelsen skyldes tilfeldigheter eller andre forhold. Det utelukkes ikke at jentene i dette utvalget er har bedre ANS-ferdigheter, men det er ikke mulig å slå fast at dette er som følge av biologiske kjønn. I denne studien vil den indre validiteten kunne sies å være god dersom man kan vise at det er en statistisk signifikant sammenheng mellom variablene, at årsaken, kjønn, eksisterer før virkningen, primære kvantitative evner, inntreffer og at ingen andre forklaringer på denne sammenhengene er sannsynlige (Shadish et al., 2002). I dette tilfelle er det umulig å avgjøre tvetydigheten rundt retning, hva som er årsak og hva som er virkning. Modning kan også true den indre validiteten. I teorikapittelet påpekes det at gutter først senere i utdanningsløpet gjør det bedre enn jenter (Hyde et al., 1990), jenters høyere skårer på ANS kan dermed skyldes at gutter modnes saktere. Sammenhengen mellom kjønn og ANS kan være spuriøs, men på grunn av undersøkelsens ikke-eksperimentelle design er ikke dette mulig å avgjøre. Videre trues den indre validiteten i denne undersøkelsen av tredjevariabler. Mulige tredjevariabler her kan være at foreldres eller læreres syn på kjønnsforskjeller i matematikk har bidratt til at de har viet jenter mer oppfølging og oppmuntring for å utligne en antatt kjønnsforskjell til fordel for gutter.

For variablene som tilhører Biologisk sekundære kvantitative evner er det ingen signifikante kjønnsforskjeller, og dermed heller ikke noe krav til forklaring som gjør det plausibelt at det ikke er andre bedre forklaringer enn kjønn på variasjonen. En trussel mot indre validitet sier at eksponering for testen gjør at elever kan dra nytte av dette på en post-test. I dette tilfellet er testene foretatt kun en gang, men elevene ble eksponert for mange av de sammen oppgavene som i ThinkMath da de tok screenet noen måneder før testingen ble utført. Likevel anses ikke dette som en aktuell trussel i dette tilfellet. Det ville forutsatt at bare et av kjønnene dro nytte av dette, og det er dermed lite sannsynlig at det har bidratt til at kjønnsforskjeller som er der ikke oppdages.

5.1.3 Begrepsvaliditet

Begrepsvaliditet sier noe om graden av samsvar mellom den teoretiske definisjonen av begrepet og hvordan det defineres gjennom operasjonaliseringen, altså i hvilken grad det er samsvar mellom hva man sier skal måles, og hva som faktisk måles (Kleven, 2002). I denne studien vil begrepsvaliditeten dreie seg om hvorvidt det testene fanger opp faktisk er ferdigheter i matematikk, IQ og eksekutive funksjoner. Primære kvantitative evner er operasjonalisert til resultatene av en ikke-symbolsk-mengdeoppfatningstest. Kjennskap til tall er operasjonalisert til resultatene av oppgave 15-72 i ThinkMath og TOBANS Digit comparison. Ferdigheter i tekstopp-gaver er operasjonalisert til resultater av oppleste tekstopp-gaver i ThinkMath og WISC-IV Regning. Aritmetikk er operasjonalisert som resultatene av totalt 20 addisjon og subtraksjonopp-gaver fra ThinkMath. Eksekutive funksjoner er operasjonalisert til resultatene av testen Tower of London, mens nonverbal IQ er operasjonalisert til resultatene av testen Raven. Graden av begrepsvaliditet i denne undersøkelsen vil da være avhengig av i hvilken grad de ulike testene måler det de utgir seg for å måle.

Hver eneste operasjonalisering av begreper innebærer en underrepresentasjon av det man ønsker å måle, samtidig som man måler noe som ikke er en bevisst del av operasjonaliseringen, dette kompliserer slutninger som trekkes (Shadish et al., 2002). Testenes validitet vil i så måte legge grunnlag for validiteten i slutningen som trekkes. I denne undersøkelsen er mye av grunnlaget for analysene hentet fra testen ThinkMath, dette er ikke en standardisert test og den vil i derfor kunne true undersøkelsens begrepsvaliditet (Lund, 2002). Et annet eksempel på en trussel mot begrepsvaliditeten angår variabelen Tekstopp-gaver. I denne testen leses oppgaven opp for eleven og måler dermed ikke bare evnen knyttet til å løse tekstopp-gaver. Eleven må også forså oppgaven som blir lest, og ulike elementer kan gjøre at eleven ikke oppfatter oppgaven riktig. Det kan være en uoppdaget språkvanske, eller at eleven av ulike årsaker er uoppmerksom når oppgaven leses. Dette vil i så fall føre til en lavere skåre på denne testen, enn det reelle nivået eleven er på. Vider vil ufullstendig beskrivelse eller forklaring av begreper også true begrepsvaliditeten man vil kunne trekke feilaktige slutninger. Upresise datainnsamlingsmetoder vil også kunne føre til svak reliabilitet. Begrepsforvirring vil også være en trussel mot begrepsvaliditet (Shadish et

al., 2002) da studien inneholder flere begreper som definerer egenskaper i nært slektskap. Om man ikke har klart å redegjøre for de ulike begrepene kan dette føre til feilvurderinger av resultatene. Det er likevel nærliggende å anta at testene i denne undersøkelsen ikke i nevneverdig grad truer begrepsvaliditeten, da de i hovedsak er standardiserte og hyppig anvendte tester.

5.1.4 Ytre validitet

Den ytre validiteten sier noe om hvem resultatene er gyldige for og må vurderes blant annet i sammenheng med konteksten datainnsamlingen har foregått i. Cook (1979) stiller spørsmålet: Til hvilken populasjon, situasjon og variabler kan denne effekten generaliseres. Den ytre validiteten gjør seg altså gjeldene i slutninger som trekkes om hvorvidt årsak-virkning forholdet står seg over variasjoner i personer, situasjoner, behandlingsvariabler og målevariabler (Shadish et al., 2002). Trusler mot ytre validitet vil være muligheten for at funnene gjort i dette utvalget ikke vil vært oppdaget i et annet utvalg. Det er umulig å forsikre seg om at de sammen kjønnsforskjellene i prestasjoner på Primære kvantitative evner, ANS, ikke vil finnes i et annet utvalg, eller at man i et annet utvalg vil finne kjønnsforskjeller på områder som av ulike årsaker ikke ble avdekket i denne undersøkelsen. Det kan også forholde seg slik at kjønnsforskjeller i matematikk ville vært avdekket i dette utvalget om man hadde gjort analyser av andre variabler enn Kjennskap til tall, Tekstoppgaver og Aritmetikk. Selv om heller ikke disse funnene uten videre ville la seg generalisere til populasjonen; også i et slikt tilfelle vil de sammen forbeholde om generalisering gjøre seg gjeldende.

Videre er settingen en trussel mot den ytre validiteten (Shadish et al., 2002). Man kan ikke utelukke at resultatene ville sette annerledes ut om undersøkelsen hadde vært gjort i en annen setting. Man må ta høyde for at omgivelsene testingen forgår i påvirker resultatene. Det er likevel vanskelig å se for seg hvordan denne trusselen skal ha en negativ påvirkning på denne studiens ytre validitet, da også hovedvekten av tidligere forskning på kjønnsforskjeller i matematikk konkluderer med at forskjellene er marginale. Om funnene i denne undersøkelsen hadde pekt på store kjønnsforskjeller ville kravene til å redegjøre for den ytre validiteten vært mer avgjørende. Her er det tale om nullfunn som er i tråd med mye av det som finnes i

tidligere empiri, og dermed vil ikke resultatene fra denne undersøkelsen være av praktisk betydning med behov for endring av praksis.

5.2 Resultater i lys av tidligere empiri

I gjennomgangen av de statistiske analysene i forrige kapittel viste resultatene at kjønn i svært liten grad er i stand til å forklare matematikkprestasjoner. Det ikke er noen signifikant sammenheng mellom variablene Kjønn og (1)Kjennskap til tall, (2)Tekstoppgaver, eller (3)Aritmetikk. Det ser riktignok ut til å være en signifikant sammenheng mellom variablene Kjønn og Primære kvantitative evner. Funnene vil her bli drøftet variabel for variabel.

5.2.1 Kjønnforskjeller i biologisk primære kvantitative evner

Biologiske primære kvantitative evner ble målt gjennom testen Dot comparison som måler elevenes prestasjoner på ANS. I følge teorien er de primære kvantitative evnene en type egenskaper som er medfødte, tverrkulturelle og uavhengig av formell opplæring (Gallistel & Gelman, 1992; Geary, 1995a). At disse evnene er medfødte vil i så fall si at eventuelle kjønnforskjeller i liten grad vil være et resultat av ytre påvirkning, men som nevnt innledningsvis i denne oppgaven, er ikke det at egenskaper er medfødte det samme som at de er uungåelig eller uforanderlige (Halpern et al., 2007). Til tross for at tidligere studier har gjort funn som underbygger at det er kjønnforskjeller i enkelte spesifikke evner som har betydning for matematikkprestasjoner (Furnham et al., 2002; Geary, 1996), foreligger det ikke kjent empiri som kan bidra til å bekrefte kjønnforskjeller til fordel for jenter i ANS. Det ville derfor være nærliggende å anta at slike kjønnforskjeller heller ikke ville bli funnet i denne undersøkelsen, men i følge analysene er det en signifikant forskjell i resultatene til fordel for jenter. En t-test for uavhengige utvalg viser at det med to-halet test er en signifikans på .012 og dermed signifikant på .05 nivå, hvilket betyr at det er 95% sjans for at vi tar feil om vi sier at forskjellen skyldes tilfeldigheter. Likevel er det, som redegjort for i 5.2 Validitet og reliabilitet, flere andre faktorer som kan forklare den signifikante kjønnforskjellen på denne variabelen.

5.2.2 Kjønnforskjeller i Tekstoppgaver

Resultatene fra analysene av kjønnforskjeller i Tekstoppgaver viser at det ikke er signifikante forskjeller i variasjonen til gutter og jenter på denne testen. Signifikansen er på $.212$, og det vil si at det er 21% sjans for at vi tar feil om vi sier at ulikhetene i gjennomsnitt og standardavvik for gutter og jenter skyldes kjønnforskjeller. Effektstørrelsen på $d = .23$ er til fordel for jenter. Ferdigheter i tekstoppgaver er i denne undersøkelsen målt med WISC-IV Regning og åtte oppgaver fra ThinkMath. Alle oppgavene er problemløsningsoppgaver av den typen man finner i tekstoppgaver, men oppgavene ble gitt muntlig, og i WISC-IV Regning ble også svarene avgitt muntlig. Resultater fra enkelte undersøkelser har vist kjønnforskjeller til fordel for gutter i løsning tekstoppgaver allerede fra førsteklasse (Harnisch et al., 1986; Lummis & Stevenson, 1990). Geary (1996) peker på at gutter er konsistent bedre enn jenter i tekstoppgaver, men tilsvarende funn er altså ikke gjort her, og man kan spekulere i hvorfor. Det kan skyldes at det ikke er kjønnforskjeller i dette utvalget, men det kan i teorien også skyldes at testen av ulike årsaker ikke har fanget opp denne variasjonen.

5.2.3 Kjønnforskjeller i Kjennskap til tall

Heller ikke i analysene av Kjennskap til tall viser resultatene at det er signifikante kjønnforskjeller. Variabelen Kjennskap til tall er målt med oppgave 15-72 i ThinkMath og TOBANS Digit comparison og ved en to-halet signifikanstest viser resultatene $p = .79$. Det er dermed ikke grunnlag for å si at det er kjønnforskjeller i denne variabelen. Empirien som er redegjort for i denne oppgaven legger til grunn at kjennskap til tall er en av de mest essensielle ferdighetene for å prestere godt i matematikk (Aunola et al., 2004; Desoete, 2014; Göbel et al., 2014). Det finnes også noen eksempler på funn som peker i retning av at jenter i enkelte tilfeller gjør det bedre enn gutter på oppgaver som måler kjennskap til tall; i den tidligere nevnte undersøkelsen av Wei et al. (2015) viste resultatene at jentene hadde høyere skårer enn gutten på denne typen oppgaver. Undersøkelsen til Wei et al. (2015) er utført på barn mellom 8 og 11 år, altså noe høyere alder enn de barna som er undersøkt her, hvilket kan være en årsak til at vi ikke ser de samme tendensene. Videre er studien av Wei et al. (2015) utført i Asia, og det kan ikke utelukkes at kulturelle forskjeller er utslagsgivende for kjønnforskjeller. Kjennskap til tall hører inn under biologisk sekundære kvantitative evner,

og disse er ikke tverrkulturelle, men avhengig av opplæring og undervisning (Geary, 2000). Det finnes også en rekke eksempler på undersøkelser som ikke finner kjønnsforskjeller i kjennskap til tall, og min studie føyer seg i så måte inn i rekken av disse.

5.2.4 Kjønnsforskjeller i Aritmetikk

I variabelen aritmetikk finner vi ikke signifikante kjønnsforskjeller og gjennom en hierarkisk regresjonsanalyse viser resultatene at kjønn forklarer 0% av variasjonen i denne variabelen etter at IQ og eksekutive funksjoner er kontrollert for. Det finnes noe tidligere empiri som peker på at jenter ser ut til å gjøre det noe bedre enn gutter de første skoleårene mens matematikk ennå handler om grunnleggende regning, men at denne forskjellen utjevnes (Halpern et al., 2007). Også i en studie fra 2011 der 1556 elever mellom 8 og 11 år ble testet, hadde jentene høyere skårer enn guttene på blant annet aritmetikk (Wei et al., 2015). Selv om det finnes empiri på kjønnsforskjeller i aritmetikk er ikke den typen funn dominerende når det kommer til kjønnsforskjeller i matematikkvansker. Resultatene i mine analyser er ganske talende og neppe til å ta feil av, en to-halet signifikanstest gir resultatet $p = .816$, og det vil derfor være vanskelig å hevde at det kan være kjønnsforskjeller i utvalget her. Med noen unntak støttes dette i tidligere forskning, funn fra 242 undersøkelser med totalt 1.286.350 deltakere viste at det er 98.01% overlapp mellom de to gutter og jenter, og altså i svært liten grad grunnlag for å si at det er systematiske kjønnsforskjeller i matematikk.

5.2.5 Kjønnsforskjeller i biologisk sekundære kvantitative evner

De tre variablene Tekstoppgaver, Kjennskap til tall og Aritmetikk måler alle ferdigheter som hører inn under biologisk sekundære kvantitative evner og er derfor også blitt slått sammen for å analyseres for kjønnsforskjeller. Denne sammenslåingen kan begrunnes også empirisk, da korrelasjon mellom de tre underområdene er signifikant på .01 nivå, og korrelasjonene er på over .6, noe som kan regnes som moderat i denne sammenheng. Analysert for kjønnsforskjeller viser en t-test signifikans på .845, altså at det ikke er signifikant sammenheng mellom variasjonen i meanskårer og standardavvik for gutter og jenter. Det er med andre ord 85% sjans for at vi tar feil om vi sier at variasjonen i gutter og jenters resultater skyldes noe annet enn tilfeldigheter. Biologiske sekundære kvantitative evner er i

følge teorien avhengig av systematisk trening og undervisning (Geary, 2000) Man kan dermed spekulere i om ikke tidligere funn som peker på kjønnsforskjeller i disse ferdighetene er et resultat av ulik opplæring for gutter og jenter, og ikke et resultat av medfødte og iboende egenskaper som er ulike for gutter og jenter.

5.3 Oppsummering

Basert på de undersøkelsene som er blitt presentert i kapittel 2 kan man slå fast at det sprikende funn, og dermed ikke mulig å trekke en entydig konklusjon om kjønnsforskjeller i matematikkferdigheter. I så måte er funnene i denne undersøkelsen i tråd med tidligere undersøkelser; det er vanskelig å argumentere for at det er kvalitative forskjeller i barns kjønn som bidrar til forskjeller i matematikkprestasjoner, men man kan ikke benekte at enkelte små forskjeller fremkommer fra tid til annen.

5.3.1 Pedagogiske konsekvenser

Det fremkommer ikke resultater i denne undersøkelsen som gir grunnlag for konkluderende slutninger om kjønnsforskjeller i matematikk blant lavtpresterende førsteklasinger. Likevel er det, som nevnt innledningsvis i denne oppgaven, fremdeles mye som står ubesvart når det gjelder matematikkvansker og fremdeles er det mange elever i skolen som ikke får den hjelpen de trenger. Til tross for at resultatene som er presentert her ikke medfører et konkret behov for endring i pedagogisk praksis, vil jeg benytte anledningen til å påpeke viktigheten av at tilstrekkelig støtte gis til de som strever med matematikk, uavhengig av kjønn.

5.3.2 Fremtidige undersøkelser

Undersøkelser om kjønnsforskjeller i intellektuelle evner generelt, og matematikk spesielt, har en lang og tidvis kontroversiell historie (Benbow & Stanley, 1980; Feingold, 1992). Selv om mange har funnet små forskjeller på enkelte evneområder som spiller inn på matematikkprestasjoner (Carr & Davis, 2001; Geary, 1995b; Halpern et al., 2007). Det ser i mange tilfeller ut til at kjønnsforskjeller i matematikk er noe som utvikler seg over tid, og det ville være av stor interesse å forsøke å avdekke mer nøyaktig når denne forskjellen begynner å manifestere seg, og på hvilke områder. Til tross for i all hovedsak nullfunn i denne omgang skulle det likevel vært interessant med en undersøkelse som ser på

kjønnsforskjeller i større utvalg og med enten andre spesifikke variabler som arbeidsminne og visuospatiale evner, eller mer generelle matematikkprestasjoner enn det som er gjort i denne undersøkelsen. Det kunne for eksempel være en videreutvikling av denne studien, der en kan benytte seg av samme type tester i en longitudinell undersøkelse der et utvalg testes på årlig basis gjennom hele skolegangen.

Litteraturliste

- Aunio, P., & Räsänen, P. (2015). *Core numerical skills for learning mathematics in children aged five to eight years - a working model for educators*. European Journal of Early Childhood Education Research.
- 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.
- Barahmand, U. (2008). *Arithmetic disabilities: training in attention and memory enhances arithmetic ability*. Research Journal of Biological Sciences, 3(11), 1305-1312.
- Barbarese, 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*, Ambulatory Pediatrics, 5(5), 281-289.
- Benbow, C. P., & Stanley, J. C. (1980). *Sex differences in mathematical ability: Fact or artifact?* Science, 210(4475), 1262-1264.
- Bergem, O. K. (2016). 2 Hovedresultater i matematikk Vi kan lykkes i realfag. Hentet 18.mars 2017 fra <https://www.idunn.no/vi-kan-lykkes-i-realfag/2-hovedresultater-i-matematikk>
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). *Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention*. Child Development, 78(1), 246-263.
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). *Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years*. Developmental Neuropsychology, 33(3), 205-228
- Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). *Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement*. Developmental Science, 14(4) 679-692.
- Bull, R., & Scerif, G. (2001). *Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory*. Developmental Neuropsychology, 19(3), 273-293.

- Butterworth, B. (2005). Developmental dyscalculia. I J. I. D. Campbell (Red.), *Hand- book of mathematical cognition* (Vol. 93, pp. 455-467). New York: PsychologyPress.
- Bynner, J., & Parsons, S. (1997). *Does Numeracy Matter? Evidence from the National Child Development Study on the Impact of Poor Numeracy on Adult Life*. Hente 13. februar 2017 fra <http://oggiconsulting.com/wp-content/uploads/Bynner-Parsons-Does-numeracy-matter.pdf>
- Carr, M., & Davis, H. (2001). *Gender Differences in Arithmetic Strategy Use: A Function of Skill and Preference*. *Contemporary Educational Psychology*, 26(3), 330-347.
- Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). *Longitudinal Associations between Executive Control and Developing Mathematical Competence in Preschool Boys and Girls*. *Child Development*, 84(2), 662-677.
- Cook, T. D., Campbell, D. T., Fankhauser, G., Reichardt, C. S., McCain, L. J., & McCleary, R. (1979). *Quasi-experimentation : design & analysis issues for field settings*. Boston: Houghton Mifflin Co.
- Culbertson, C. W., & Zillmer, E. (1999). *Tower of London Drexel University (TOL DX): Examiner's Manual: Multi-Health Systems Incorporated (MHS)*
- De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). *Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade*. *Journal of Experimental Child Psychology*, 103(2), 186-201.
- De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). *How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior*. *Trends in Neuroscience and Education*, 2(2), 48-55.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). *Intelligence and educational achievement*. *Intelligence*, 35(1), 13-21
- Desoete, A. (2014). Predictive indicators for mathematical learning disabilities/dyscalculia in kindergarten children. I S. Chinn (Red.), *The Routledge international handbook of dyscalculia and mathematical learning difficulties* (s. 90-100). Oxon: Routledge.
- Desoete, A., Roeyers, H., & De Clercq, A. (2004). *Children with mathematics learning disabilities in Belgium*. *Journal of Learning Disabilities*, 37(1), 50-61.

- Devine, A., Soltesz, F., Nobes, A., Goswami, U., & Szucs, D. (2013). *Gender differences in developmental dyscalculia depend on diagnostic criteria*. *Learning and Instruction*, 27, 31-39.
- Dirks, E., Spyer, G., van Lieshout, E. C., & de Sonneville, L. (2008). *Prevalence of combined reading and arithmetic disabilities*. *Journal of Learning Disabilities*, 41(5), 460-473.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C. (2007). *School readiness and later achievement*. *Developmental Psychology*, 43(6), 1428-1446.
- Feigenson, L., & Carey, S. (2005). *On the limits of infants' quantification of small object arrays*. *Cognition*, 97(3), 295-313
- Feingold, A. (1992). *Sex differences in variability in intellectual abilities: A new look at an old controversy*. *Review of Educational Research*, 62(1), 61-84.
- Fuchs, L. S., & Fuchs, D. (2002). *Mathematical problem-solving profiles of students with mathematics disabilities with and without comorbid reading disabilities*. *Journal of Learning Disabilities*, 35(6), 563-573
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., Fletcher, J. M. (2006). *The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems*. *Journal of Educational Psychology*, 98(1), 29-43.
- Fuhs, M. W., Hornburg, C. B., & McNeil, N. M. (2016). *Specific early number skills mediate the association between executive functioning skills and mathematics achievement*. *Developmental Psychology*, 52(8), 1217-1235.
- Furnham, A., Reeves, E., & Budhani, S. (2002). *Parents Think Their Sons Are Brighter Than Their Daughters: Sex Differences in Parental Self-Estimations and Estimations of Their Children's Multiple Intelligences*. *The Journal of Genetic Psychology*, 163(1), 24-39.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research : an introduction* (7th ed.). Boston: Allyn and Bacon.
- Gallistel, C. R., & Gelman, R. (1992). *Preverbal and verbal counting and computation*. *Cognition*, 44(1), 43-74.

- Geary, D. C. (1995a). *Reflections of evolution and culture in children's cognition. Implications for mathematical development and instruction*. *American Psychologist*, 50(1), 24-37.
- Geary, D. C. (1995b). *Sexual Selection and Sex-Differences in Spatial Cognition*. *Learning and Individual Differences*, 7(4), 289-301.
- Geary, D. C. (1996). *Sexual selection and sex differences in mathematical abilities*. *Behavioral and Brain Sciences*, 19(2), 229-&.
- Geary, D. C. (2000). *From infancy to adulthood: the development of numerical abilities*. *European Child and Adolescent Psychiatry*, 9 Suppl 2, II11-16.
- Geary, D. C. (2011a). *Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study*. *Developmental Psychology*, 47(6), 1539-1552.
- Geary, D. C. (2011b). *Consequences, Characteristics, and Causes of Mathematical Learning Disabilities and Persistent Low Achievement in Mathematics* *Journal of Developmental and Behavioral Pediatrics* (Vol. 32, s. 250-263).
- Geary, D. C. (2013a). *Early Foundations for Mathematics Learning and Their Relations to Learning Disabilities*. *Current Directions in Psychological Science*, 22(1), 23-27.
- Geary, D. C. (2013b). *Learning Disabilities in Mathematics – Recent Advances*. In H. L. H. Swanson, K.R. & Graham, S. (Ed.), *Handbook of Learning Disabilities – 2nd edition* (s. 239-255). New York: The Guilford Press.
- Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). *Predicting Mathematical Achievement and Mathematical Learning Disability With a Simple Screening Tool The Number Sets Test*. *Journal of Psychoeducational Assessment*, 27(3), 265-279.
- Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). *Numerical and arithmetical cognition: a longitudinal study of process and concept deficits in children with learning disability*. *Journal of Experimental Child Psychology* 77(3), 236-263.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). *Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability*. *Child Development*, 78(4), 1343-1359.
- Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). *Numerical and Arithmetical Cognition: Patterns of Functions and Deficits in Children at Risk for a Mathematical Disability*. *Journal of Experimental Child Psychology*, 74(3), 213-239.

- 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.
- Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). *Development of number line representations in children with mathematical learning disability*. *Developmental Neuropsychology*, 33(3), 277-299.
- Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). *Sex differences in spatial cognition, computational fluency, and arithmetical reasoning*. *Journal of Experimental Child Psychology*, 77(4), 337-353.
- Geary, D. C., & vanMarle, K. (2016). *Young children's core symbolic and nonsymbolic quantitative knowledge in the prediction of later mathematics achievement*. *Developmental Psychology*, 52(12), 2130-2144.
- Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). *Developmental dyscalculia: Prevalence and demographic features*. *Developmental Medicine & Child Neurology*, 38(1), 25-33.
- 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.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). *The Science of Sex Differences in Science and Mathematics*. *Psychological Science in the Public Interest*, 8(1), 1-51.
- Hannula, M. M., Räsänen, P., & Lehtinen, E. (2007). *Development of Counting Skills: Role of Spontaneous Focusing on Numerosity and Subitizing-Based Enumeration*. *Mathematical Thinking and Learning*, 9(1), 51-57.
- Harnisch, D. L., Steinkamp, M. W., Tsai, S.-L., & Walberg, H. J. (1986). *Cross-national differences in mathematics attitude and achievement among seventeen-year-olds*. *International Journal of Educational Development*, 6(4), 233-244.
- Helsevesenet. (2017). *ICD-10 : Den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer*. Oslo: Kompetansesenter for I. T. i helsevesenet.

- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). *Gender differences in mathematics performance: a meta-analysis*. *Psychological Bulletin*, 107(2), 139-155.
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). *Early math matters: kindergarten number competence and later mathematics outcomes*. *Developmental Psychology*, 45(3), 850-867.
- Karagiannakis, G., & Cooreman, A. (2015). Focused MLD intervention based on the classification of MLD subtypes. I S. Chinn (Red.), *The Routledge international handbook of dyscalculia and mathematical learning difficulties* (s. 265-276). Oxon: Routledge.
- Kleven, T. A. (2002). Begreperoperasjonalisering. I T. Lund (Red.), *Innføring i forskningsmetodologi*. (s. 141-183). Oslo: UNIPUB.
- Kleven, T. A. (2002b). Ikke-eksperimentelle design. I T. Lund (Red.), *Innføring i forskningsmetodologi*. (s. 265-286). Oslo: UNIPUB.
- Kunnskapsdepartementet. (2014). *NOU 2014:7 Elevenes læring i fremtidens skole - Et kunnskapsgrunnlag*. Oslo: Kunnskapsdepartementet.
- Kyttälä, M., & Lehto, J. E. (2008). *Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence*. *European Journal of Psychology of Education*, 23(1), 77.
- Lee, K., Bull, R., & Ho, R. M. H. (2013). *Developmental Changes in Executive Functioning*. *Child Development*, 84(6), 1933-1953.
- Lee, K., Ng, S. F., Bull, R., Pe, M. L., & Ho, R. H. M. (2011). *Are patterns important? An investigation of the relationships between proficiencies in patterns, computation, executive functioning, and algebraic word problems*. *Journal of Educational Psychology*, 103(2), 269-281.
- Libertus, M. E., Feigenson, L., & Halberda, J. (2011). *Preschool acuity of the approximate number system correlates with school math ability*. *Developmental Science*, 14(6), 1292-1300.
- Lin, J., & Geary, D. C. (1998). *Numerical Cognition: Age-Related Differences in the Speed of Executing Biologically Primary and Biologically Secondary Processes*. *Experimental Aging Research*, 24(2), 101-137.

- Lindberg, S. M., Hyde, J. S., Petersen, J. L., Linn, M. C., & Hinshaw, S. P. (2010). *New Trends in Gender and Mathematics Performance: A Meta-Analysis*. *Psychological Bulletin*, 136(6), 1123-1135.
- Lummis, M., & Stevenson, H. W. (1990). *Gender differences in beliefs and achievement: A cross-cultural study*. *Developmental Psychology*, 26(2), 254.
- Lund, T. (2002). Metodologiske prinsipper og referanserammer. I T. Lund (Redd.), *Innføring i forskningsmetodologi* (s. 79-124). Oslo: Fagbokforlaget.
- Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011b). *Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia)*. *Child Development*, 82(4), 1224-1237.
- Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). *Preschoolers' Precision of the Approximate Number System Predicts Later School Mathematics Performance*. *PLOS ONE*, 6(9), e23749.
- Mazzocco, M. M. M., & Thompson, R. E. (2005). *Kindergarten Predictors of Math Learning Disability*. *Learning Disabilities Research & Practice*, 20(3), 142-155.
- Midtbø, T. (2007). *Regresjonsanalyse for samfunnsvitere : med eksempler i SPSS*. Oslo: Universitetsforlaget
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). *The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis*. *Cognitive Psychology*, 41(1), 49-100.
- Murray, E., Hillaire, G., Johnson, M., & Rappolt-Schlichtmann, G. (2015). Representing, acting, and engaging: UDL mathematics. I S. Chinn (Red.), *The Routledge international handbook of dyscalculia and mathematical learning difficulties* (s.393-405). Oxon: Routledge.
- Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., Urbina, S. (1996). *Intelligence: Knowns and unknowns*. *American Psychologist*, 51(2), 77-101.
- NESH. (2016). *Forskningsetiske retningslinjer for samfunnskunnskap, jus og humaniora*. Hentet 9. februar 2017 fra <https://www.etikkom.no>

- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y. Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F. Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W. Somogyi, M., Akrami, N., Ekehammar, B., Vianello, M., Banaji, M. R. Greenwald, A. G. (2009). *National differences in gender–science stereotypes predict national sex differences in science and math achievement*. *Proceedings of the National Academy of Sciences*, 106(26), 10593-10597.
- NSD. (2017). *Database for statistikk om høgre utdanning*. Hentet 15. februar 2017 fra <http://dbh.nsd.uib.no>
- Ostad, S. A. (2010). *Matematikkvansker : en forskningsbasert tilnærming*. Oslo: Unipub.
- Parsons, S., & Bynner, J. (2006). *Does numeracy matter more?* (National research and development centre for adult literacy and numeracy). Hentet 12. februar 2017 fra <http://eprints.ioe.ac.uk/4758/1/parsons2006does.pdf>
- Passolunghi, M. C., Mammarella, I. C., & Altoe, G. (2008). *Cognitive abilities as precursors of the early acquisition of mathematical skills during first through second grades*. *Developmental Neuropsychology*, 33(3) 229-250.
- Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). *The precursors of mathematics learning: Working memory, phonological ability and numerical competence*. *Cognitive Development*, 22(2), 165-184.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research : explanation and prediction (3rd ed.)*. Forth Worth: Harcourt Brace College Publishers.
- Price, G., & Ansari, D. (2013). *Dyscalculia: Characteristics, Causes, and Treatments*. *Numeracy*, 6(1)
- Ramaa, S., & Gowramma, I. (2002). *A systematic procedure for identifying and classifying children with dyscalculia among primary school children in India*. *Dyslexia*, 8(2), 67-85.
- Reeve, R. A., & Gray, S. (2015). Number difficulties in young children: deficits in core number? I S. Chinn (Red.), *The Routledge International Handbook of Dyscalculia and Mathematical Learning Difficulties* (s. 44-59). Oxon: Routledge.

- Riccio, C. A., Wolfe, M. E., Romine, C., Davis, B., & Sullivan, J. R. (2004). *The Tower of London and neuropsychological assessment of ADHD in adults*. *Archives of Clinical Neuropsychology*, 19(5), 661-671.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company
- SNL. (2009). *Prediksjon*. Store norske leksikon. Hentet 13. april 2017 fra <https://snl.no/prediksjon>
- SSB. (2016). *Karakterer ved avsluttet grunnskole, 2016*. Hentet 5. februar 2017 fra <https://www.ssb.no/utdanning/statistikker/kargrs/aar/2016-10-06>
- Steele, J. (2003). *Children's Gender Stereotypes About Math: The Role of Stereotype Stratification I*. *Journal of Applied Social Psychology*, 33(12), 2587-2606.
- 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.
- Sund, I. B., Asbjørnsen, E. H., & Zondag, M. H. W. (2015). *Kunnskapsministeren bekymret: Flere stryker i matte*. NRK.no. Hente 15.mars 2017 fra https://www.nrk.no/norge/kunnskapsministeren-bekymret_-flere-stryker-i-matte-1.12520084
- Swanson, H. L., & Jerman, O. (2006). *Math Disabilities: A Selective Meta-Analysis of the Literature*. *Review of Educational Research*, 76(2), 249-274.
- Temple, E., & Posner, M. I. (1998). *Brain mechanisms of quantity are similar in 5-year-old children and adults*. *Proceedings of the National Academy of Sciences*, 95(13), 7836-7841.
- Utdanningsdirektoratet. (2013). *Læreplan i matematikk fellesfag*. Oslo: Kunnskapsdepartementet.
- Utdanningsdirektoratet. (2017). *Eksamen fellesfag*. Hentet 14. mars 2017 fra <https://skoleporten.udir.no/rapportvisning/videregaende-skole/laeringsresultater/eksamen-fellesfag/nasjonalt>
- Van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2012). *The development of executive functions and early mathematics: a dynamic relationship*. *British Journal of Educational Psychology*, 82 (Pt 1)

- Walberg, H. J. (1984). *Improving the productivity of America's schools*. Educational leadership, 41(8)
- Wei, T., Liu, X., & Barnard-Brak, L. (2015). *Gender Differences in Mathematics and Reading Trajectories among Children from Kindergarten to Eighth Grade*. Research in Education, 93(1), 100-119.
- Xu, F., & Spelke, E. S. (2000). *Large number discrimination in 6-month-old infants*. Cognition, 74(1), B1-b11.
- Zhou, X. C., D. (2015). When and why numerosity processing is associated with developmental dyscalculia. I S. Chinn (Red.), *International handbook of dyscalculia and mathematical learning difficulties* (s. 78-89). London: Routledge.