

Inge Vang Ellefsen

Sammenligning av tallforståelseskompetanse hos elever på 1. trinn på data og papir

En kvantitativ studie med kvalitative
observasjoner

Masteroppgave i matematikdidaktikk

Veileder: Trygve Solstad

Mai 2021

Inge Vang Ellefsen

Sammenligning av tallforståelseskompetanse hos elever på 1. trinn på data og papir

En kvantitativ studie med kvalitative observasjoner

Masteroppgave i matematikdidaktikk
Veileder: Trygve Solstad
Mai 2021

Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for lærerutdanning



Kunnskap for en bedre verden

Sammendrag

Denne studien sammenligner tallforståelseskompetanse hos elever på 1. trinn med følgende forskningsspørsmål:

Hvilke forskjeller i tallforståelseskompetanse finner man dersom elever på 1. trinn løser identiske matematikkoppgaver på data og på papir?

Dette forskningsspørsmålet vil gi noen svar på om det vil spille noen rolle om hvilken metode som brukes for å måle tallforståelse, som er svært relevant i dag da kartleggingsprøver i regning på 1. og 3. trinn vil gå fra å gjøres på papir til digitalt om kort tid.

For å søke noen svar på forskningsspørsmålet har til sammen 45 elever på 1. trinn løst to identiske oppgavesett hvor det ene blir gjennomført på data med chromebook og det andre blir gjennomført på papir. For å analysere resultatene fra disse oppgavesettene er det blitt gjort en variansanalyse (repeated measures ANOVA) gjennom statistikkprogrammet SPSS (Statistical Package for the Social Sciences). Grunnlaget for måling av tallforståelse i denne studien er basert på det teoretiske rammeverket FONS (foundational number sense), et rammeverk som via åtte ulike komponenter til sammen skal kunne måle tallforståelsen hos barn.

Resultatene fra studien viser at det ikke er vesentlige forskjeller i målt tallforståelse hos elevene; de scorer ganske likt uavhengig om oppgavene ble løst på papir eller på chromebook.

Studien gir likevel noen indikasjoner og observasjoner på at oppgavesettet på data i visse sammenhenger kan være å foretrekke som testverktøy, men som har mer med fokus på oppgaveløsning og miljøet rundt testsituasjonen å gjøre enn selve tallforståelsen. Dette kan man imidlertid ikke sette to streker under svaret på uten videre fordi det handler om tilgang, og erfaring med, chromebooks hos elevene. Det kan også se ut som det generelt var større motivasjon hos elevene til å gjennomføre oppgavesettet på data kontra oppgavesettet på papir.

Resultatene viser også at det er forskjell på første og siste gjennomførelse uansett hvilken rekkefølge elevene løste oppgavesettene i. Scoren ble høyere andre gang oppgavene ble løst.

Jeg diskuterer også validiteten til oppgavesettet på papir og motivasjonen til å gjennomføre oppgavesettene gjennom kvalitative observasjoner underveis i datainnsamlingen. Disse diskusjonene resulterer i at det er større trusler til validiteten opp mot oppgavesettet på papir kontra på data, og at det generelt var større motivasjon til å gjennomføre oppgavesettet på data kontra på papir.

Abstract

This study compares number sense competence in 1st grade students with the following research questions:

What differences in number sense competence are found if students in 1st grade solve identical mathematics problems on computer and on paper?

This research question will provide some answers as to whether it will matter what method is used to measure number sense, which is very relevant today as map tests in mathematics in the 1st and 3rd grade will go from being solved on paper to digital in a short time.

To seek some answers to the research question, a total of 45 students in the 1st grade have solved two identical problem sets where one is solved on a computer by using a chromebook and the other is solved on paper. To analyze the results from these sets of tasks, an analysis of variance (repeated measures ANOVA) has been performed through the statistical program SPSS (Statistical Package for the Social Sciences). The basis for measuring number sense in this study is based on the theoretical framework FONS (fundamental sense of numbers), a framework that via eight different components together will be able to measure numeracy in children.

The result from the study shows that there are no significant differences in measured students' number sense; they score pretty much regardless of whether the tasks were solved on paper or on chromebook.

The study nevertheless provides some indications and observations that the set of tasks on computer in certain contexts may be predictable as a test tool, but which has more to do with focusing on problem solving and the environment around the test situation than the number sense. One must still be careful about saying so because it is about access, and experience with, chromebooks to the students. It may also appear that there was generally greater motivation among the students to complete the assignment set on computer versus the assignment set on paper.

The results also shows that there is a difference between the first and last completion regardless of the order in which the students solved the problem sets. The score was higher the second time the problems were solved.

I also discuss the validity of the task set on paper and the motivation to complete the task sets through qualitative observations during the data collection. These discussions result in greater threats to the validity of the task set on paper versus on computer, and that there was generally greater motivation to complete the task set on computer versus on paper.

Forord

Denne masteroppgaven er et verdig punktum i min 6-årige lærerutdannelse. Det siste året har vært meget spesielt, og det har til tider vært utfordrende å skrive en masteroppgave samtidig som det har herjet en pandemi rett utenfor døren. Uansett, når oppgaven nå er ferdig, er det mange som fortjener en stor takk for mye og god hjelp på veien.

Aller først vil jeg takke alle foreldre, elever og lærere som sa seg villig til å være en del av forskningsprosjektet mitt. Uten dere ville ikke oppgaven min blitt noe av.

En ekstra stor takk til den dyktige veilederen min Trygve Solstad som har tilrettelagt for gjennomføring av datainnsamling, bidratt med meget gode og konstruktive tilbakemeldinger gjennom hele skriveprosessen og generelt vært en trygghet i en litt fremmed forskerhverdag.

Jeg vil også takke mine medstudenter på lesesal G425 for gode og reflekterende samtaler om oppgaven det siste halvåret samt lange kaffepauser og utallige runder med «runs of eleven». Spesielt Håkon Kristian Tangen Ljøen, min medkompanjong fra dag 1 på lærerstudiet, fortjener en stor takk for at studietilværelsen har blitt så bra som den har vært.

I tillegg vil jeg takke familie, venner, kollegaer og samboer som har sett igjennom oppgaven og kommet med tips og triks. Spesielt vil jeg takke samboer Tiril for måten hun har akseptert at hodet mitt har bestått av masteroppgaven og tallforståelse på papir vs. data det siste året som sikkert ikke alltid virker like spennende.

Til høsten starter en helt ny hverdag som lærer, noe jeg gleder meg veldig til. Det er også vanskelig å ikke være stolt av meg selv som nå har skrevet ferdig en masteroppgave, noe som var vanskelig å se for seg for et par år siden.

Inge Vang Ellefsen

Trondheim, mai 2021.

Innhold

Figurer	xi
Tabeller	xi
Forkortelser/symboler	xi
1 Innledning	1
1.1 Tema for forskningsområde	1
1.2 Grunnleggende ferdigheter	2
1.3 Vurdering av regning som grunnleggende ferdighet	2
1.4 Forskningsspørsmål	3
2 Teori – Tallforståelse og vurdering	4
2.1 Læringsteori	4
2.2 Tallforståelse	5
2.2.1 Foundational Number Sense (FONS)	5
2.2.1.1 Åtte komponenter av FONS	6
2.2.2 Andre definisjoner av tallforståelse opp mot FONS	7
2.3 Vurdering	8
2.3.1 Summativ og formativ vurdering	8
2.3.2 Digital vurdering i matematikk	9
3 Metode	10
3.1 Valg av metode	10
3.1.1 Kvantitativ forskning	10
3.1.2 Observatør	10
3.2 Datainnsamling	11
3.2.1 Utvalg av deltakere	11
3.2.2 Valg av oppgaver	12
3.3 Gjennomføring av datainnsamling	13
3.3.1 Pilotundersøkelse	15
3.3.1.1 Resultater pilotundersøkelse	15
3.3.2 Endringer fra pilotundersøkelsen til datainnsamlingen	16
3.3.3 Gjennomførelse	16
3.4 Retting av oppgaver	17
3.5 Metode for analyse: Variansanalyse gjennom SPSS	19
3.6 Forskningens reliabilitet og validitet	20
3.7 Etske hensyn	22
3.7.1 Forskning på barn	22
3.7.2 Samtykke	22

4	Resultater.....	23
4.1	Forskjeller i oppgavesett som helhet.....	23
4.1.1	Forskjeller papir og data.....	23
4.1.2	Forskjeller første og siste gjennomførelse.....	26
4.2	Forskjeller i FONS-kategorier.....	27
4.2.1	Forskjell papir og data.....	28
4.2.2	Forskjell første og siste gjennomførelse.....	29
4.3	Forskjell i enkeltoppgaver.....	29
4.3.1	Forskjell papir og data.....	30
4.3.2	Forskjell første og siste gjennomførelse.....	31
4.4	Forskjeller i gjennomføring.....	31
4.4.1	Lærerens og elevenes rolle under gjennomførelsen av begge oppgavesettene 31	
4.4.2	Oppgaveforklaringens betydning.....	33
5	Diskusjon.....	34
5.1	Oppsummering av funn opp mot forskningsområde.....	34
5.2	Mine funn opp mot tidligere forskning.....	35
5.3	Betydning av funn.....	35
5.3.1	Betydning av funn papir vs. data.....	35
5.3.2	Betydning av funn første vs. siste gjennomførelse.....	37
5.3.3	Betydning av funn fra gjennomførelsen.....	38
5.4	Studiens begrensninger.....	38
5.5	Videre forskning.....	40
	Referanser.....	41
	Vedlegg.....	45

Figurer

Figur 1 Eksempel på oppgave i oppgavesettet på data	13
Figur 2 Eksempel på oppgave med hjelpestreker som ga riktig svar	18
Figur 3 Eksempel på oppgave med hjelpestreker som ga feil svar	19
Figur 4 Resultater papir og data med standardfeil	24
Figur 5 Resultater oppgavesett med rekkefølgen på gjennomføring	25
Figur 6 Resultater oppgavesett papir og data med linjemodell	25
Figur 7 Linjemodell som viser forskjell på første og siste gjennomførelse.....	26
Figur 8 Forskjell første og siste gjennomførelse for alle respondentene	27
Figur 9 Forskjell papir og data for hver FONS-kategori med standardfeil.....	28
Figur 10 Forskjell først og sist for hver FONS-kategorier med standardfeil	29
Figur 11 Forskjell papir og data for hver enkelt oppgave.....	30
Figur 12 Forskjell første og siste gjennomførelse for hver enkelt oppgave.....	31

Tabeller

Tabell 1 Oversikt over hva som ble gjort og når	14
Tabell 2 De fire gruppene man får ved å bruke motvektprinsippet blant respondentene	15
Tabell 3 Resultater pilotundersøkelse	15
Tabell 4 Oversikt over observasjoner gjort underveis i datainnsamlingen.....	21
Tabell 5 P-verdiene fra repeated measures ANOVA-test.....	24

Forkortelser/symboler

EKVA	Enhet for Kvantitative Utdanningsanalyser
FONS	Foundational Number Sense
NESH	Nasjonale Forskningsetiske Komité for Samfunnsvitenskap og Humaniora
NTNU	Norges teknisk-naturvitenskapelige universitet
NSD	Norsk Senter for Forskningsdata
SPSS	Statistical Package for the Social Sciences

1 Innledning

1.1 Tema for forskningsområde

På barneskolen er man interessert i å vurdere kompetanse i regning hos elever slik at man kan tilpasse undervisningen til deres nivå. Et godt eksempel på dette er gjennom kartleggingsprøver i regning (Utdanningsdirektoratet, 2020c) som blir gjennomført på barneskolen. På 1-3. trinn skal disse kartleggingsprøvene måle elevenes begrepsforståelse og regneferdigheter og målet med disse prøvene er at man kan identifisere elever som befinner seg under en definert bekymringsgrense slik at de kan få ekstra tilrettelegging i faget. På 1. og 3. trinn er kartleggingsprøven frivillig for elevene, mens den er obligatorisk på 2. trinn slik at i løpet av de tre første årene på skolen skal alle elevene ha gjennomført minimum en kartleggingsprøve i regning. Konkret måler prøvene i regning på 1-3. trinn kompetanse i fem temaer: tallrekka, tallinja, regneferdigheter, telleferdigheter og tallbegrep.

Dersom man slår sammen temaene som kartleggingsprøvene i regning måler, kan man si at man hovedsakelig måler tallforståelse hos elevene. Det kan man si på grunn av definisjoner på matematisk kompetanse og tallforståelse fra ulike forskere, hvor begrepene går hånd i hånd. Kilpatrick, Swafford og Findell (Kilpatrick et al., 2001) definerer matematisk kompetanse ut fra fem komponenter, hvor alle disse komponentene er avhengig av hverandre og sammenflettet:

- Konseptuell forståelse: forståelse av matematiske begreper, operasjoner og relasjoner.
- Prosessuell flyt: kan utføre prosedyrer fleksibelt, nøyaktig, effektivt og passende.
- Strategisk kompetanse: evne til å formulere, representere og løse matematiske problemer.
- Adaptiv resonnering: evne til å tenke logisk, reflekterende, forklarende og begrunnet.
- Produktiv disposisjon: evnen til å se matematikk som fornuftig, nyttig og verdifullt kombinert med tro på egne evner i faget.

Temaene i kartleggingsprøvene på 1-3. trinn, støtter seg først og fremst til det å ha en konseptuell forståelse av faget som da konsentrerer seg om forståelse av begreper, operasjoner og relasjoner, og som det vil være vesentlig at man behersker slik at de andre komponentene kan utvikles og senere også måles. En kartleggingsprøve på de nevnte trinnene måler altså elevenes konseptuelle forståelse.

Disse fem komponentene fra definisjonen av matematisk kompetanse kommer også til syne i blant annet McIntosh, Reys og Reys (McIntosh et al., 1992) sin definisjon av begrepet tallforståelse, som viser sammenhengen mellom matematisk kompetanse og tallforståelse. Deres definisjon på tallforståelsesbegrepet er bygd opp av tre elementer og resulterer i at man ser nytten av tall og at matematikk har en viss regelmessighet:

1. Generell forståelse av tall og operasjoner.
2. Denne generelle forståelsen blir brukt hensiktsmessig i matematisk resonnering i arbeid med tall og regneoperasjoner.

3. Refleksjon til å bruke denne kunnskapen for å kommunisere, prosessere og tolke informasjon.

Her ser man at for å beherske tallforståelse må man først bygge opp en generell forståelse av tall og operasjoner. I dette ligger de samme aspektene som man finner i det å ha en konseptuell forståelse av faget. En kartleggingsprøve på 1-3. trinn som måler den konseptuelle forståelsen vil med andre ord også samtidig måle grunnelementet i tallforståelsesbegrepet basert på denne definisjonen.

1.2 Grunnleggende ferdigheter

I stortingsmelding 030, kultur for læring (2003-2004) kom begrepet grunnleggende ferdigheter for første gang til syne (St.meld. nr. 030 2003-2004). Disse ferdighetene ble introdusert som «sentrale ferdigheter som er grunnleggende redskaper for læring og utvikling» (St.meld. nr. 030 2003-2004) og oppsto som en mer meningsfull og samlende beskrivelse av den helhetlige kompetansen i skolen enn det de altomspennende definisjonene av basiskompetansene gjorde. Vi snakker om de fem grunnleggende ferdighetene lesing, skriving, regning og muntlige- og digitale ferdigheter. I den nye fagfornyelsen blir det sagt at «skolen skal legge til rette for og støtte elevenes utvikling av de fem grunnleggende ferdighetene gjennom hele opplæringsløpet» (Utdanningsdirektoratet, 2020a). Det blir argumentert for at disse grunnleggende ferdighetene er «nødvendige redskaper for læring og faglig forståelse og viktige for utviklingen av elevenes identitet og sosiale relasjoner, og for å kunne delta i utdanning, arbeid og samfunnsliv» (Utdanningsdirektoratet, 2020a).

På mange måter er undervisning på 1. og 2. trinn lagt opp slik at man må se disse grunnleggende ferdighetene både i sammenheng med hverandre og på tvers av fag (Utdanningsdirektoratet, 2020a). Sammenhengen mellom de grunnleggende ferdighetene er veldig tydelig i matematikkfaget (Utdanningsdirektoratet, 2020b). I matematikk må man for eksempel kunne lese for å sortere informasjon, analysere og vurdere form og innhold i sammensatte tekster, man må beherske å kunne skrive for å skrive ned en løsning på oppgaven ved hjelp av riktige tegn og representasjoner og man må uttrykke seg muntlig på en meningsfull måte dersom man skal forklare en løsningsstrategi eller problem på en oppgave. Dette viser at norskfaget, hvor man øver seg på blant annet de grunnleggende ferdighetene lesing og skriving, spiller en direkte rolle i utviklingen av regneferdigheter og at fagene spiller ulike roller i utviklingen av de forskjellige ferdighetene.

1.3 Vurdering av regning som grunnleggende ferdighet

Som tidligere nevnt blir regning som grunnleggende ferdighet vurdert gjennom for eksempel kartleggingsprøver fra Utdanningsdirektoratet. Formålet med disse kartleggingsprøvene er å «finne elever som trenger ekstra oppfølging» (Utdanningsdirektoratet, 2020c) ut fra en definert bekymringsgrense. Disse kartleggingsprøvene i regning gjøres skriftlig som vil si at elevene må skrive svarene sine på papir, men de får hjelp til å lese og forstå oppgavene. Det vil si at når elever på barneskolen skal vise tallforståelse og at de behersker den grunnleggende ferdigheten «å kunne regne», må de gjennomføre skriftlige kartleggingsprøver. Ut fra egen erfaring er det stor variasjon i skolestarteres grunnleggende ferdigheter og det finnes mange elever på 1. trinn som så vidt har lært seg å holde i en blyant. Denne erfaringen får støtte fra forskningslitteratur, for eksempel skriver Hekneby (Hekneby, 2005) at de fleste seks- sjuåringer befinner seg på de tre første stadiene av utforskning av skriftsystemet, som vil

si enten skriverabling, bokstavutforsking og helordsskriving. Med andre ord vil det si at de ikke har knekt skriftspråkskoden ennå. Med dette i minnet stiller jeg spørsmål til hvor mye krav til skriving det stilles på de skriftlige kartleggingsprøvene i regning. Er det slik at elever kan prestere dårligere som en direkte konsekvens av måten man gjennomfører slike kartleggingsprøver på og i verste fall bli vurdert som svakere i matematikk enn de egentlig er?

Denne måten å måle regning som grunnleggende ferdighet på vil endre seg i framtiden. Per dags dato er forskere ved enhet for kvantitative utdanningsanalyser (EKVA) i gang med å utvikle kartleggingsprøver i regning for 1. og 3. trinn på oppdrag fra Utdanningsdirektoratet (Heie, 2021). Selv om kartleggingsprøvene blir digitale i framtiden har de likevel samme hovedfunksjon, de skal fremdeles finne elever under en definert bekymringsgrense slik at man kan tilrettelegge for tilpasset undervisning. Et argument for at kartleggingsprøvene blir digitale i framtiden er at man ser at de yngste elevene viser større utholdenhet når de arbeider med digitale oppgaver og klarer derfor å arbeide lengre med oppgavene med større konsentrasjon (Heie, 2021). Et annet argument er at man sparer mye tid på å ha oppgavene digitalt, slik at lærere kan bruke tiden til å fortolke elevenes resultater og utvikling enn å bruke samme tid på å rette prøvene.

På bakgrunn av spørsmålet om elevene får vist fram sin matematiske kompetanse basert på skriftlige kartleggingsprøver og at kartleggingsprøver i regning blir digitale i fremtiden har jeg gjennomført et forskningsprosjekt der målet var å undersøke om det er forskjell i tallforståelseskompetansen hos et utvalg 1.klasseelever dersom de gjennomfører et oppgavesett på papir og et identisk oppgavesett på data. Med andre ord vil jeg undersøke om 1.klasseelevene viser en annen kompetanse i regning dersom man går helt bort fra måling basert på den grunnleggende ferdigheten «å kunne skrive» og om det har en funksjon at kartleggingsprøvene vil bli digitale i framtiden. Dersom det viser seg at det er stor forskjell i resultatene kan man få kunnskaper og erfaringer om hvilke oppgavesett som passer best til å vurdere kompetanse i regning som grunnleggende ferdighet.

1.4 Forskningsspørsmål

Ut fra tema og problemområde for denne forskningen vil forskningsspørsmålet for dette masterprosjektet være:

Hvilke forskjeller i tallforståelseskompetanse finner man dersom elever på 1. trinn løser identiske matematikkoppgaver på data og på papir?

For å svare på dette forskningsspørsmålet har til sammen 45 elever på 1. trinn løst to oppgavesett med matematikkoppgaver hvor det ene er et oppgavesett på data som blir løst på chromebook, mens det andre er et oppgavesett på papir som blir løst med blyant og viskelær. Alle oppgavene på de to oppgavesettene er helt identiske, som betyr at det er nøyaktig de samme oppgavene i begge oppgavesettene og de blir instruert helt likt. Det er svarscoren og observasjoner fra gjennomførelsen på disse oppgavesettene som vil være grunnlaget for å besvare forskningsspørsmålet og for å vurdere elevenes tallforståelse.

2 Teori – Tallforståelse og vurdering

I dette forskningsprosjektet skal jeg se på to ulike vurderingsformer for tallforståelse til 1.klasseelever ved hjelp av forskjellige matematikkoppgaver. I dette kapittelet er det derfor hensiktsmessig å forklare begrepene tallforståelse og vurdering nøyere. I tillegg henger vurdering sammen med læringsteori da jeg i dette forskningsprosjektet skal vurdere resultatene fra individuell løsning av oppgaver, som er forbundet med individuelle kognitive ferdigheter og dermed kognitiv læringsteori.

2.1 Læringsteori

Dette forskningsprosjektet skal altså vurdere en gruppe elevers kognitive ferdigheter i matematikk der og da på to ulike plattformer. Kognitive ferdigheter kan vinkles mot den kognitive utviklingsteorien, en læringsteori utviklet for å prøve å forklare utviklingen av hvordan mennesker tenker, lansert av Jean Piaget (Piaget, 2001). I denne læringsteorien ser man på læring som organiseringer av indre tankeprosesser gjennom kognitive skjemaer som fungerer som byggesteiner i barns tenking (Lyngsnes & Rismark, 2007; Piaget, 2001). Fra man blir født og gjennom hele livet vil man skape og utvikle kognitive skjemaer som er basert på erfaring og viten som et individ er i besittelse av på et gitt tidspunkt gjennom hele livet (Renolen, 2015). Når ny læring og kunnskap skal tilegnes, vil denne kunnskapen enten kunne plasseres inn i allerede eksisterende kognitive skjemaer, eller så er den nye kunnskapen så fremmed og krever at man må opprette helt nye kognitive skjemaer. Dersom ny kunnskap kan plasseres i eksisterende og kjente skjemaer kalles dette assimilasjon. Om kunnskapen er så ny at det må opprettes et nytt kognitivt skjema kalles dette akkomodasjon (Lyngsnes & Rismark, 2007; Piaget, 2001).

For å beskrive hva vi kan forvente på forhånd av elevene som er med i forskningsprosjektet kan vi se på hvor langt de har kommet i utviklingen av tankeprosesser. Piaget deler inn den kognitive utviklingen i fire forskjellige stadier, og vi kan se at elevene befinner seg i det pre-operasjonelle stadiet, som vil si tankesett for barn mellom to og sju år (Piaget, 2001). Navnet på stadiet henviser til operasjoner, som Piaget definerte som «høyere mentale prosesser» (Renolen, 2015). I den pre-operasjonelle fasen kommer intelligens fram gjennom bruken av symboler, språkbruken modnes, hukommelse og fantasien til barn er i utvikling og tankeprosessene skjer på en ulogisk måte som ikke kan gjøres reversibel, det vil si at man ikke kan resonnerer til motsatte operasjoner enn det som faktisk har skjedd (Huitt & Hummel, 2003). Fra fire til seks år har barnet også lært å bruke intuisjon, som betyr at det kan løse kognitivt krevende problemer uten at de klarer å forklare hvordan de løste problemet (Renolen, 2015).

Det man kan forvente i det pre-operasjonelle stadiet kan knyttes direkte opp mot matematikk, noe som Bobby Ojose har gått nærmere inn på (Ojose, 2008). I denne fasen kan man forvente at det er mangel på logiske resonnementer hos barna. I matematikk vil dette bety at man ikke tenker rasjonelt, for eksempel kan man få til operasjonen hvor man legger sammen 4 og 5 og får til sammen 9, men får ikke til den motsatte operasjonen, hvor man trekker 4 fra 9 for å få 5. Det er også forventet at barn i dette stadiet har oppfatninger som er basert på kun ett aspekt, for eksempel at man har

samme mengde vann i to like beholdere, men når det samme vannet i den ene beholderen går til en større beholder slik at vannivået blir lavere, vil et barn i dette stadiet tro at det er mindre vann i den nye beholderen.

Med den kognitive utviklingsteorien i mente handler dette forskningsprosjektet om å vurdere de nåværende kognitive ferdighetene i matematikk til en gruppe elever. Disse elevene er altså på et tidlig nivå i utviklingen av tankesett, og man må dermed ta i betraktning at disse kognitive ferdighetene vil utvikle seg i forskjellige retninger senere ved assimilasjon og akkomodasjon for disse elevene. Det vil med andre ord si at det kan hende at noen elever presterer høyt på disse oppgavesettene, men senere vil slite med matematikk, eller at noen elever presterer lavt på disse oppgavesettene, men senere vil mestre matematikk av ulike grunner.

2.2 Tallforståelse

2.2.1 Foundational Number Sense (FONS)

Tallforståelse er et begrep som mange forskere har definert på forskjellige måter. I dette forskningsarbeidet vil tallforståelsesbegrepet være direkte knyttet til Andrews & Sayers (Andrews & Sayers, 2015) sin definisjon da oppgavesettet som er brukt i dette forskningsprosjektet er basert på deres tolkning av tallforståelsesbegrepet. Definisjonen til Andrews & Sayers (Andrews & Sayers, 2015) handler om at man ser på tallforståelse som en sammensetning av 8 ulike komponenter som til sammen utgjør fellesbetegnelsen FONS (foundational number sense). Disse ulike komponentene legger vekt på forskjellige deler av matematikken, men alle disse til sammen skal vurdere barns tallforståelse, først og fremst barn på 1. trinn. Ikke bare kan FONS anses som en definisjon av tallforståelsesbegrepet, men det er også blitt utviklet til å være et rammeverk for å vurdere tallforståelse hos barn som kan benyttes i ulike kulturer. Det kan man si ettersom forfatterne selv bekrefter at dette er tilfellet da de selv har testet ut dette rammeverket på 1. klasseelever i tre forskjellige land, Sverige, England og Ungarn, og elevene fra disse landene var mottakelig for rammeverket.

FONS skal med andre ord beskrive tallforståelsen hos barn (fem-seks år) hovedsakelig i 1. klasse hvor man får muligheten til å vurdere tidlig i hvilken grad elever får til å operere fleksibelt med tall og mengder. Skaperne av dette rammeverket, Andrews & Sayers (Andrews & Sayers, 2015), mener at ved å bruke rammeverket til å måle tallforståelse på barn i tidlig alder har man også et godt instrument til å forutse hvilke matematiske prestasjoner elevene vil oppnå i fremtiden. Dermed kan man slå fast at de elevene som presterer høyt på oppgavesettene i dette forskningsprosjektet sannsynligvis også vil komme til å prestere høyt i matematikk i fremtiden. Det kan man si basert på tidligere forskning, for eksempel gjennom Watts m.fl sin forskning som viser at barns matematiske ferdigheter allerede i barnehagen kan forutse hvilke matematiske prestasjoner man kan oppnå i tenårene (Watts et al., 2014). Dette krever imidlertid også at barna må møte formelle matematiske instruksjoner tidlig i skolen, for eksempel lære seg å lese og skrive i matematikken, for å bevare og utvikle de matematiske ferdighetene og holdningene (Davasligil, 2004; Robinson et al., 1997).

2.2.1.1 Åtte komponenter av FONS

FONS består som tidligere nevnt av åtte forskjellige komponenter hvor alle komponentene er en sammensetning av flere teoretiske aspekter (Andrews & Sayers, 2015) og alle knytter seg til tallforståelse hos barn i tidlig skolealder.

2.2.1.1.1 Tallidentifikasjon

Barn kan gjenkjenne tallsymboler og assosierer tallsymbolene med riktig vokabular og mening. For eksempel innebærer det at de kan identifisere et tallsymbol ut fra en mengde av flere tallsymboler og gi korrekt navn til tallsymbolet. Dersom man ikke behersker tallidentifikasjon, vil det gi matematiske problemer senere. Spesielt gjelder det subitisering, som betyr at man kan se på et sett med objekter og med en gang gjenkjenne hvor mange objekter det er snakk om. Dersom man behersker tallidentifikasjon tidlig, vil man være mer i stand til å håndtere regning med flersifrede tall senere (Clarke & Shinn, 2004; Desoete et al., 2012; Gersten et al., 2005; Koontz, 1996; Lembke & Foegen, 2009; Malofeeva et al., 2004, i Andrews & Sayers, 2015).

2.2.1.1.2 Systematisk telling

Dette innebærer at man kan telle fra 0 til 20 og bakover igjen, men også at man kan telle videre fra et tilfeldig tall innenfor denne mengden. På den måten har man en forståelse av at hvert tall inntar en fast posisjon i rekkefølgen av tall, som vil være nødvendig for å tilegne seg aritmetisk kompetanse (Gersten et al., 2005; Griffin, 2004; Jordan & Levine, 2009; Lipton & Spelke, 2005; Passolunghi et al., 2007, i Andrews & Sayers, 2015).

2.2.1.1.3 Forhold mellom tall og mengde

Forhold mellom tall og mengde handler om at man forstår en-til-en korrespondanse mellom tallets navn og antall enheter tallet representerer og man forstår også at det siste tallet man kommer til når man teller representerer den hele mengden av det man har talt. Barn som ikke har et slikt forhold mellom tall og mengde har større sjanse for å få matematiske vansker senere (Jordan & Levine, 2009; Kroesbergen et al., 2009; Malofeeva et al., 2004; Mazzocco et al., 2011; Van luit & Schopman, 2000, i Andrews & Sayers, 2015).

2.2.1.1.4 Mengdediskriminering

Her har man bevissthet om størrelse og sammenligning mellom forskjellige størrelser. Dette innebærer at man forstår hva mindre enn og større enn betyr, for eksempel at mengden 8 er større enn mengden 6 men mindre enn mengden 10. Dermed er ikke telling lengre en oppramsing av en memorert liste av tallord, men man forstår prinsippene bak hvert tall man teller (Clarke & Shinn, 2004; Gersten et al., 2005; Griffin, 2004; Ivrendi, 2011; Jordan et al., 2006; Lembke & Foegen, 2009; Lipton & Spelke, 2005; Yang & Li, 2008, i Andrews & Sayers, 2015).

2.2.1.1.5 Forståelse av forskjellige representasjoner for et tall

Dette handler om at man forstår at et tall kan bli representert på forskjellige måter og at disse kan fungere som forskjellige referansepunkter. For eksempel kan en forståelse av tallinja være relevant når man senere skal jobbe med emner innenfor aritmetikk, en forståelse for at en mengde som en representasjon av et tall kan gi senere en bedre forståelse av numeriske strukturer og en forståelse av at man bruke fingrene til enkel aritmetikk og telling tidlig kan gi større grunnlag til økt matematisk kompetanse senere.

(Ivrendi, 2011; Jordan et al., 2007; van Nes & van Eerde, 2010; Yang & Li, 2008, i Andrews & Sayers, 2015).

2.2.1.1.6 Estimering

God FONS vil også si at barn kan estimere, gjennom størrelser eller objekter. Dette involverer at man går fra forskjellige representasjoner for et tall, for eksempel plassere et tall på riktig plass på en tom tallinje. Estimering viser seg å være en viktig faktor i regningskompetanse senere, særlig når man blir utfordret med nye situasjoner (Berch, 2005; Booth & Siegler, 2006, 2008; Gersten et al., 2005; Ivrendi, 2011; Jordan et al., 2006, 2007; Malofeeva et al., 2004, i Andrews & Sayers, 2015).

2.2.1.1.7 Enkel aritmetisk kompetanse

Barn kan utføre enkle aritmetiske operasjoner. Dersom man for eksempel kan løse enkle addisjons- og subtraksjonsstykker har det blitt bevist at det kan være en indikator for senere matematisk suksess. Når barn løser matematikkoppgaver skriftlig vil de først utvikle kompetansen til å løse ikke-verbale matematikkoppgaver, for eksempel et regnestykke med to forskjellige tallsymboler før de utvikler kompetansen til å løse problemløsningsoppgaver som presenteres på vanlig språk. Dette på grunn av barns eksisterende erfaringer med å kombinere fysiske objekter (Geary et al., 2009; Ivrendi, 2011; Jordan & Levine, 2009; Krajewski & Schneider, 2009; Levine et al., 1992; Malofeeva et al., 2004; Yang & Li, 2008, i Andrews & Sayers, 2015).

2.2.1.1.8 Bevissthet om tallmønstre

Bevissthet om tallmønstre betyr at man er i stand til å identifisere et manglende tall i for eksempel en tallrekke. Dette forsterker ferdigheter i telling og vil senere være gunstig i arbeid med aritmetiske operasjoner. Det viser seg at dersom man ikke behersker dette, for eksempel ikke ser at det mangler et tall i en tallrekke, er dette en av de største indikasjonene på senere matematiske vansker (Berch, 2005; Chard et al., 2005; Clarke & Shinn, 2004; Gersten et al., 2005; Lembke & Foegen, 2009, Van Luit & Schopman, 2000, i Andrews & Sayers, 2015).

2.2.2 Andre definisjoner av tallforståelse opp mot FONS

Som tidligere nevnt er begrepet tallforståelse noe som blir definert på forskjellige måter med vekt på ulike aspekter. Berch (Berch, 2005) har skrevet en artikkel og laget en oversikt over 30 forskjellige definisjoner av tallforståelse fra ulike forskere hvor man ser hvor utrolig stort omfang dette begrepet har. Denne artikkelen er skrevet i 2005, som vil si at det er en god del år før Andrews & Sayers (Andrews & Sayers, 2015) la ut sin beskrivelse av FONS, og som forklarer hvorfor ikke FONS er nevnt i artikkelen. Dersom man slår sammen alle disse definisjonene vil man se at dersom man har god tallforståelse vil dette utgjøre at man har en bevissthet, intuisjon, anerkjennelse, kunnskap, dyktighet, evne, lyst, følelse, forventning, prosess, konseptuell struktur og mental tallinje til tall og matematikk. Det er med andre ord vanskelig å konkludere med hva tallforståelse egentlig er basert på alle disse forskjellige definisjonene, og Gersten (Gersten et al., 2005) peker på at ulike forskere ikke definerer tallforståelse på nøyaktig samme måte. Dette møter kritikk hos blant annet Andrews & Sayers (Andrews & Sayers, 2015), som stiller spørsmål til hvorfor forskere definerer begrepet på så vage, altomfattende og forskjellige måter når alle de samme forskerne viser enighet i hvor viktig det er med god tallforståelse og hvilke konsekvenser manglende tallforståelse kan gi i matematikk senere. Denne kritikken er utgangspunktet for FONS hvor forfatterne har

slått sammen flere viktige aspekter ut fra ulike teoretiske retninger til å bli en samlet definisjon på begrepet i motsetning til de allerede eksisterende definisjonene.

2.3 Vurdering

I dette forskningsarbeidet blir vurdering et sentralt tema ettersom grunnlaget for forskningen er å vurdere tallforståelseskompetanse til 1. klasseelever, og det er derfor hensiktsmessig å forklare dette begrepet nøyere.

2.3.1 Summativ og formativ vurdering

Når man snakker om vurdering i skolen skiller man gjerne mellom to typer vurderingsformer: summativ og formativ vurdering. Man ble først kjent med disse to vurderingsbegrepene på 1960-tallet hvor begrepene ble utviklet og tatt i bruk av Michael Scriven (Scriven, 1967). Formativ vurdering handler om at man tar tak i kunnskap hos elever der og da slik at man kan diskutere dette og planlegge videre steg, mens summativ vurdering handler om å vurdere kunnskapen til elevene systematisk som et sluttprodukt (DES/WO, 1988). Tradisjonelt kommer summative vurderinger til syne gjennom sluttvurderinger som karakterer etter prøver, eksamener, standpunktprøver osv., mens formative vurderinger kan være underveisvurderinger gjennom samtaler og tester underveis i et semester med hensikt om å avdekke lærestoff elevene sliter med og for å forme elevenes videre læringsarbeid. Med dette ser vi et skille mellom begrepene hvor formativ vurdering har selve læringsprosessen som formål mens summativ vurdering handler om å bedømme de endelige prestasjonene i læringsprosessen (Helle & Burder, 2021).

Dette skillet mellom begrepene gjør at det kan se ut som de to begrepene er motsetninger av hverandre og ikke kan stilles likt. Akkurat dette aspektet er diskutert blant forskere, og for eksempel Bennett har sett nøye på rammene rundt formativ vurdering og viser til at de to vurderingsformene ikke er motsetninger av hverandre, men tvert imot henger nøye sammen (Bennett, 2011). Han mener i korte trekk at enhver form for formativ vurdering hvor man skal gi tilbakemeldinger for videre læring krever at man på forhånd har gjennomført en eller annen form for summativ vurdering slik at man er i stand til å gi disse formative tilbakemeldingene og at definisjonen av formativ vurdering derfor er mer omfattende og sammensatt enn det andre forskere viser til, hvor begrepet knyttes til kun en læringsprosess eller en test. Da det virker som om man ikke har en entydig definisjon av begrepet, stiller han spørsmålsteget ved om man kan dokumentere effekten av nyskapende ideer innenfor emnet fordi man er uenig om hva man faktisk ser etter (Bennett, 2011). Bennett definerer derfor formativ vurdering til å være en gjennomtenkt integrering av en læringsprosess og utformet metodikk eller instrumentering for å nå et mål (Bennett, 2011). Det vil med andre ord si at formative vurderinger skjer i hele læringsfasen hvor man reflekterer over hvilke metoder som skal til for at så mange som mulig når læringsmålene og hvordan man skal vurdere om læringsmålene er nådd eller ikke. Denne definisjonen viser at det er en sammenheng mellom formativ og summativ vurdering og ikke et skille mellom dem.

For å koble opp summativ og formativ vurdering opp mot relevante aspekter i denne studien, kan man se på formålet med kartleggingsprøvene som gjøres i regning. Hensikten med kartleggingsprøvene i regning er som tidligere nevnt å identifisere elever som trenger ekstra oppfølging, det vil si at de skal være et verktøy i å tilrettelegge for tilpasset opplæring. Man kan derfor si at kartleggingsprøvene er et formativt

vurderingsverktøy som skal brukes til å forme videre undervisningsopplegg for at elevene skal utvikle seg i faget så godt som mulig. Dersom man tilrettelegger på en hensiktsmessig måte basert på resultatene på en kartleggingsprøve vil det også gi resultater på sikt som kan påvirke summative vurderinger senere i utdanningsløpet. Kartleggingsprøvene i regning er med andre ord et godt eksempel på en formativ vurdering som samtidig viser samspillet mellom formative og summative vurderinger, noe Bennett trekker fram i forrige avsnitt (Bennett, 2011).

2.3.2 Digital vurdering i matematikk

Det er mange fordeler ved å bruke teknologien til digitale vurderinger i matematikk. Stacey & Wiliam (Stacey & Wiliam, 2012) har skrevet en artikkel som tar for seg fordelene ved å bruke digitale vurderinger i matematikk, og som vil være relevant i min forskning da jeg bruker et oppgavesett på data til å måle tallforståelseskompetansen hos en gruppe elever. Ved å bruke digitale vurderingsprogrammer til å vurdere kunnskap hos elever oppstår det nye muligheter til å velge ut hvilke oppgaver som skal brukes, hvordan oppgavene presenteres, hvordan elevene løser oppgavene, hvordan matematiske bevis blir identifisert av elever og hvordan bevis kan samles (Almond et al., 2003). I tillegg trekker Stacey og William (Stacey & Wiliam, 2012) frem gode argumenter på teknologiske potensialer og spesifikke muligheter man får når man bruker databaserte oppgavesett til å måle kunnskap hos elever. Blant annet trekker de frem muligheten man har til å være fleksibel ved at forskjellige elever kan gjennomføre de samme oppgavene eller testene på ulikt tidspunkt. I tillegg nevner de muligheten man har til å rotere objekter tredimensjonalt. Ikke minst trekkes det også fram at svarscorene på testene kommer mye raskere når elevene gjør dataoppgaver på en server kontra at de skal gjøres på papir, uten mulighet til menneskelige feil som å telle feil osv. Til slutt nevner de at man har muligheten til å legge til rette for oppgaver mer billedlig enn språklig i dataoppgaver, for eksempel med «drag-and-drop»-oppgaver og bruken av «hotspots» som oppgaver i oppgavesettet på data som jeg bruker i dette forskningsarbeidet inneholder. Typiske «drag-and-drop»-oppgaver er at man må flytte riktig tall eller svar fysisk over i en boks eller et annet objekt for å svare på oppgaven, noe som ikke er like lett å legge til rette for i oppgaver som løses med skriving. Andrews og Sayers mener at alle disse fordelene med digital vurdering i matematikk gjør at man får et bedre bilde på de rent matematiske kunnskapene hos elevene.

Selv om det finnes mange fordeler med å bruke teknologi til å gjennomføre vurderinger i matematikk kan det også by på noen utfordringer, som baserer seg på ulik kompetanse, erfaringer eller kjennskap man har til teknologien. Ut fra egne observasjoner i lærerrollen finnes det familier og elever på skolen som ikke har tilgang til digitale ressurser hjemme av forskjellige årsaker, eller så finnes det lærere som heller ikke behersker nye teknologiske produkter. Det betyr at for disse elevene eller lærerne så kan det være en ulempe å få gjennomført slike digitale vurderinger da man rett og slett mangler erfaringer for denne vurderingsplattformen og det krever at man har brukt tiden og satt seg inn i det på forhånd.

3 Metode

I denne studien er formålet å se om prøveformen, det vil si enten matematikkoppgaver på papir eller på data, har noe å si for resultatet og om disse to prøveformene måler den samme eller ulik tallforståelseskompetanse. Derfor vil jeg benytte meg av en metode som får målt et helhetlig bilde av riktige svar på oppgavene, både fra oppgavene på data og oppgavene på papir, slik at jeg kan måle og sammenligne kompetansen elevene viser. I tillegg vil den helhetlige vurderingen av disse to prøveformene være basert på kvalitative observasjoner under gjennomføringen av datainnsamlingen og hele forskningsprosessen som spesielt knytter seg til validiteten i studien og elevenes motivasjon til gjennomføring.

3.1 Valg av metode

3.1.1 Kvantitativ forskning

For å kunne svare på problemstillingen vil jeg undersøke svarene i matematikkoppgaver på data og på papir til et tilstrekkelig antall respondenter. Dette er med andre ord kvantitativ forskning hvor målet er å få vite noe som angår en hel befolkning eller kunnskap om mer avgrensede grupper (Christoffersen & Johannessen, 2012). I kvantitativ forskning blir de som er med i forskningen kalt enheter. Dersom enhetene er personer som har svart på et spørreskjema, blir de kalt for respondenter (Christoffersen & Johannessen, 2012). Da mitt forskningsprosjekt baserer seg på elever på 1. trinn som gjennomfører en spørreundersøkelse med matematikkoppgaver vil jeg videre bruke begrepet respondenter om disse elevene videre. I kvantitativ forskning får man med andre ord inn data fra mange personer slik at hensikten blir å konkludere generelt ut fra dataen som er samlet inn kontra å gå i dybden på kompetansen eller erfaringen til noen få. Dette betyr også at det er aspekter i denne forskningen som ikke kan svares på, som baserer seg på kvalitativ data. Aspektene dreier seg hovedsakelig om hvorfor resultatene blir som de blir knyttet opp mot respondentenes kognitive ferdigheter og affektive sider i temaet som blir testet.

3.1.2 Observatør

I tillegg til kvantitativ forskning, fant jeg det hensiktsmessig å være en deltakende observatør under gjennomføringen av datainnsamlingen. Deltakende observatør betyr at man er fysisk til stede under datainnsamlingen, men uten å gjennomføre oppgavene selv (Tjora, 2013). Det innebærer også at deltakerne i studien måtte gjøres klar over at jeg kom til å være til stede under datainnsamlingen og være forberedt på at jeg kom til å ta notater underveis, og at jeg hele tiden ville delta synlig gjennom aktivitetene de gjennomførte (Tjora, 2013). Valget på min rolle som deltakende observatør gjorde at jeg, i tillegg til et helhetlig tallmateriale av et sluttprodukt, hadde muligheten til å gjøre kvalitative observasjoner og refleksjoner underveis som særlig kunne knyttes til validiteten til studien og resultatene. Nettopp dette trekkes fram som en av fordelene ved observasjon som metode, hvor man får denne muligheten til å observere aspekter som kan påvirke resultatene i forskningen, i dette tilfellet også validiteten, som kan rettes mot hvor bastant man føler kan konkludere forskningens funn (Cohen et al., 2018). De kvalitative observasjonene ble dokumentert gjennom feltnotater underveis i

datainnsamlingen som ble nedskrevet på et eget dokument i etterkant av hver gjennomførelse. Fokuset for de kvalitative feltnotatene var å legge merke til aspekter som kunne knyttes til motivasjon hos respondentene og validiteten til studien som brukes for å gi en helhetlig vurdering av de to prøveformene.

3.2 Datainnsamling

I dette forskningsarbeidet ble det brukt primærdata for å sammenligne tallforståelseskompetansen hos respondentene. Det vil si at dataen som ble samlet inn var spesifikt rettet for å svare på forskningsspørsmålet i dette forskningsarbeidet (Hox & Boeijs, 2005). Målet for datainnsamlingen var at jeg skulle få inn materiale som kunne bli brukt som grunnlag til å vurdere og sammenligne tallforståelseskompetansen som respondentene viste i to oppgavesett, hvor et oppgavesett var på data og et oppgavesett var på papir. Dette resulterte i at respondentene måtte løse to oppgavesett med flere matematikkoppgaver. Oppgaveløsingen ble gjort i grupper på 4-5 respondenter per gruppe. Det var flere grunner til at det var få elever på hver gruppe. Først og fremst handlet det om at det skulle bli målt individuell tallforståelseskompetanse og det var vesentlig å luke ut muligheter for samhandling og samarbeid respondentene imellom gjennom oppgavesettene som kan skje dersom man har for store grupper. Mest optimalt var at man tok ut respondentene en og en for å være mer sikker på at det var individuell kompetanse som ble målt. Da hver gjennomføring tok i snitt 15-20 minutter ville dette imidlertid tatt veldig lang tid. Kombinasjonen av oversiktlige grupper på 4-5 og min rolle som deltakende observatør gjorde likevel at jeg var sikker på at det ble målt individuell kompetanse hos respondentene da jeg kunne notere ned hvilke respondenter jeg for eksempel la merke til kopierte andres svar under gjennomførelsen og luke ut disse når svarene senere skulle analyseres.

3.2.1 Utvalg av deltakere

Deltakerne i dette forskningsprosjektet var elever på 1. trinn fra en skole i Trondheim. Dette trinnet ble valgt ut etter dialog og godkjenning mellom skolens ledelse og NTNUs ansvarlig for forskningsprosjektet. For at respondentene skulle være deltakende i dette forskningsarbeidet måtte de ha godkjenning fra foreldrene i forkant ettersom respondentene var under 15 år (NESH, 2016). Respondentene fikk i tillegg også når som helst muligheten til å trekke seg fra forskningen dersom de ikke hadde lyst til å være med. Det var med andre ord ikke gjort noen faglige vurderinger som grunnlag for hvem som skulle delta på dette forskningsarbeidet. På dette trinnet var det rundt 70 elever og det ble sendt ut forespørsel om barnet deres kunne få være med på forskningsprosjektet til samtlige foresatte på det utvalgte trinnet, totalt rundt 70. Av de rundt 70 som ble forespurt var det 45 som fikk godkjenning fra foreldrene, alle innenfor samme trinn fordelt på 4 forskjellige klasser. Det vil si at i dette forskningsprosjektet var det samlet inn 45 svar på matematikkoppgaver på papir og 45 svar på matematikkoppgaver på data for å svare på forskningsspørsmålet. Jeg hadde ingen kjennskap til noen av elevene fra før slik at det ikke var dannet noen form for inntrykk om respondentene før gjennomførelsen. For å svare på matematikkoppgavene på data ble det brukt chromebooks, noe som respondentene hadde brukt en god del før i den vanlige undervisningen, slik at de var nokså vant med verktøyet fra før.

Under gjennomførelsen av datainnsamling kom respondentene inn i grupper på 4-5 respondenter per gruppe. Denne inndeling av grupper var tilfeldig, men alle gruppene tilhørte samme klasse slik at det var enklere å ta de ut. For å senere kunne sammenligne

resultatene fikk hver av respondentene et kallenavn under gjennomførelsen, som var det samme kallenavnet i oppgavesettet på data og oppgavesettet på papir. For å ha kontroll på hvilken respondent som var tilknyttet kallenavnet de har fått ble det opprettet et dokument som viste navnet på respondenten og hvilket kallenavn som tilhørte denne respondenten. På grunn av personvern var dette dokumentet kryptert og det var kun jeg som hadde tilgang til dette.

3.2.2 Valg av oppgaver

For å kunne måle tallforståelseskompetanse hos respondentene måtte det bli brukt oppgaver som målte akkurat dette. I samarbeid med NTNU lagde jeg et oppgavesett med totalt 26 oppgaver som målte forskjellige tallforståelseskomponenter som til sammen ga et helhetlig bilde av tallforståelseskompetansen til respondentene. Disse tallforståelseskomponentene var basert på rammeverket FONS (se kapittel 2). Vedlegg 2 viser oppgavesettet på papir hvor alle oppgavene er samlet i tillegg til oppgaveinstruksen for hver oppgave og hvilken kategori av FONS oppgaven tilhører.

FONS legger altså opp til åtte ulike komponenter som har til hensikt å beskrive tallforståelse hos elever i 1. klasse. I denne studien hadde jeg og veileder blitt enige om å prioritere kun fem av disse komponentene av forskjellige årsaker. Blant annet fordi noen av oppgavene i oppgavesettet kunne sies å måle flere komponenter samtidig og på grunn av tidsaspektet, da det ville tatt lang tid å gjennomføre oppgavesett med alle åtte komponentene. De fem komponentene av FONS som oppgavene var basert på (Andrews & Sayers, 2015) og som vi brukte til å måle tallforståelseskompetanse i dette forskningsprosjektet var:

1. Tallidentifikasjon (TI): Det var totalt seks tallidentifikasjonsoppgaver i oppgavesettet mitt.
2. Systematisk telling (ST): Det var totalt fem oppgaver av systematisk telling i oppgavesettet mitt.
3. Tall og mengde (T&M): Det var totalt fire oppgaver om tall og mengde i oppgavesettet mitt.
4. Estimering (E): Det var to oppgaver som handler om estimering i oppgavesettet mitt.
5. Aritmetisk kompetanse (AK): Det var ni oppgaver som måler aritmetisk kompetanse i oppgavesettet mitt.

Vi ser også at jeg har skrevet forkortelser av de ulike komponentene i parentes. Disse forkortelsene vil komme til syne igjen under resultatdelen av dette forskningsprosjektet når jeg skal knytte resultatene opp mot hver kategori av FONS for å måle tallforståelsen til respondentene.

I utgangspunktet var disse oppgavene kun et oppgavesett på data som er utviklet av ILU/NTNU. Den første oppgaven min var dermed å oversette disse oppgavene slik at jeg kunne skrive de ut og gi de til respondentene i papirformat så de kunne løses med blyant og viskelær. Målet med oversettelsen av oppgavesettet på data til et oppgavesett på papir var at oppgavene skulle være så identiske som overhodet mulig slik at jeg kunne sammenligne svarscoren etterpå. Det vil si at oppgavene var nøyaktig de samme, hvordan man forklarte de ulike oppgavene likeså og rammene rundt oppgaveløsingen skulle også være identisk slik at det ikke er ytre påvirkninger rundt selve oppgavene som kunne påvirke resultatene og dermed validiteten til studien.

Opggavene var varierte, men som oftest var de flervalgsoppgaver. De oppgavene som ikke var av dette slaget, var estimeringsoppgaver hvor respondentene skulle plassere riktig tall på en tom tallinje. Uansett hvordan oppgavene var, ville respondentene i hver oppgave bli presentert med et matematisk problem hvor det var flere alternativer som kunne være riktig svar. Det matematiske problemet ble presentert av en innlest stemme i oppgavesettet på data og av meg når de gjennomførte oppgavene på papir hvor jeg sa nøyaktig det samme som den innleste stemmen. I flervalgsoppgavene var det kun et riktig svar. Som oftest kunne respondentene velge mellom 4 ulike svar, men dette varierte. Noen oppgaver hadde bare 3 ulike svar, mens andre hadde hele 12.

Da respondentene gjennomførte oppgavesettet på data, brukte de altså chromebooks til gjennomføring. Med disse kunne man trykke på skjermen slik at man kunne løse oppgavene ved å bruke fingrene til å trykke på svar eller flytte tall til hvor svarene skulle være som var ganske enkelt å beherske for respondentene. Figuren under viser eksempel på en av oppgavene i oppgavesettet på data som ble løst på chromebook. Her ser vi i tillegg at det er en figur av ei jente som man kan trykke på dersom man vil høre oppgaveinstruksen en gang til og en pil som man trykker på for å komme til neste oppgave.



Figur 1 Eksempel på oppgave i oppgavesettet på data

3.3 Gjennomføring av datainnsamling

Datainnsamlingen i feltet tok totalt seks dager fordelt på fire uker. Da inkluderer jeg kun de dagene jeg var ute i felt og ikke tiden det tok å skaffe samtykke, danne grupper til gjennomføring og lage oppgaver og generelt forarbeid. Den første av de seks dagene var en pilotundersøkelse, mens de fem andre var brukt til datainnsamling.

UKE/DAG	HVA BLE GJORT?
43	Pilotundersøkelse.
Tirsdag	- Tre respondenter gjennomførte en pilotundersøkelse med oppgaver på papir.
46	Gjennomføring av datainnsamling.
Mandag	- Respondentene fra A + B klasse gjennomførte oppgavesettet på data på chromebooks.
Tirsdag	- Respondentene fra C + D klasse gjennomførte oppgavesettet på papir.
47	Gjennomføring av datainnsamling.

Mandag	- Respondentene fra A + B klasse gjennomførte oppgavesettet på papir.
Tirsdag	- Respondentene fra C + D klasse gjennomførte oppgavesettet på data på chromebooks.
49	Gjennomføring av datainnsamling for de som manglet på en av de andre dagene.
Onsdag	- Gjennomføring av oppgavesett for de respondentene som ikke var til stede under en av gjennomførelsene. Totalt var det tre respondenter fra A-lassen som ikke gjennomførte oppgavesettet på data i uke 46. Det vil også si at disse respondentene egentlig skulle gjennomføre det første oppgavesettet på chromebook, men ettersom de ikke var til stede, har disse respondentene hatt sin første gjennomførelse på papir istedenfor. Det vil bli tatt hensyn til dette under analysen av datamaterialet slik at disse respondentene kategoriseres som respondenter som gjennomførte oppgavesettet på papir før data kontra data før papir som de egentlig skulle.

Tabell 1 Oversikt over hva som ble gjort og når

Dersom man studerer denne tabellen med oversikten over hva som ble gjort, ser vi hovedsakelig at uke 46 og 47 ble brukt på datainnsamling. Vi ser også at man kan dele inn respondentene i to ulike grupper, hvor den ene gruppen er respondentene som gjennomførte først oppgavesettet på data før papir (senere referert som DP), og den andre gruppen er respondentene som gjennomførte oppgavesettet på papir før data (senere referert som PD). Det var hensiktsmessig med en uke mellom hver gjennomføring da jeg prøvde å utelukke at respondentene la merke til at de gjorde de samme oppgavene. Under gjennomførelsen i uke 47 kunne jeg observere at det kunne virke som om ingen av respondentene la merke til at de hadde gjort disse oppgavene fra før ettersom det var ingen som sa noe om dette eller viste andre tegn til at de gjenkjente oppgavene. Jeg stilte også et oppfølgingsspørsmål etter datainnsamling denne uken der jeg spurte om de synes de gjorde flere matematikkoppgaver denne uka kontra forrige uke. 29 av respondentene mente de gjorde flere oppgaver på papir, fem mente at de gjorde like mange oppgaver i begge oppgavesettene mens to mente at de gjorde flere oppgaver på data og resten hadde ikke noen formening. Dette gir indikasjoner på at de ikke la merke til at det var de samme oppgavene.

Selv om ingen respondenter sa noe om at de la merke til at det var samme oppgaver kan man ikke utelukke at noen faktisk la merke til det. Det betyr at jeg må ta hensyn til at i dette forskningsarbeidet har de deltagende respondentene fått målt den samme tallforståelseskompetansen basert på samme oppgaver to ganger, et aspekt kalt gjentatte mål, og det dukker opp i hovedsak to ulike konsekvenser som kan ha påvirkning til hvordan respondentene svarer på oppgavene (Field, 2017). Den ene konsekvensen er en praktisk konsekvens hvor respondentene presterer forskjellig fra den ene gangen til den andre ettersom de er kjent med oppgavene fra før. Den andre konsekvensen handler om kjedsomhet da det rett og slett kan hende at respondentene kan prestere annerledes den andre gangen på grunn at de er lei av oppgavene ettersom de har gjort det en gang før.

Det er svært vanskelig å eliminere disse to konsekvensene helt, men man kan ta ulike valg underveis i forskningsarbeidet for å motvirke at disse konsekvensene kan påvirke resultatene. I dette forskningsarbeidet har jeg valgt det som benevnes som

motvektsprinsippet kalt (Field, 2018) og baserer seg på man har to grupper som gjennomfører det samme, bare i motsatt rekkefølge. Det betyr at jeg har delt inn alle respondentene i to grupper hvor ca. halvparten av respondentene gjorde først oppgavesettet på data mens den andre halvdel gjorde oppgavesettet på papir, også byttet de neste uke. Man har med andre ord to grupper og to ulike måter å gjennomføre på, og får dermed dannet fire ulike grupper basert på hvem som gjennomførte hva først og sist. Det vil være hensiktsmessig å skille mellom disse fire gruppene når jeg senere skal sammenligne respondentenes første og siste gjennomførelse.

	OPPGAVER UKE 46	OPPGAVER UKE 47	FORKLARING
Klasse A + B	Oppgavesett på data (Gruppe 1)	Oppgavesett på papir (Gruppe 2)	Klasse A + B gjennomfører oppgavesettet på data først (uke 46), deretter oppgavesettet på papir (uke 47).
Klasse C + D	Oppgavesett på papir (Gruppe 3)	Oppgavesett på data (Gruppe 4)	Klasse C + D gjennomfører oppgavesettet på papir først (uke 46), deretter oppgavesettet på data (uke 47).

Tabell 2 De fire gruppene man får ved å bruke motvektsprinsippet blant respondentene

3.3.1 Pilotundersøkelse

I pilotundersøkelsen var målet å få et inntrykk av hvordan oppgavesettet på papir fungerte, og hvor lang tid man kunne regne med at en respondent brukte på å løse oppgavesettet slik at planleggingen av datainnsamlingen kunne blitt gjort basert på disse erfaringene. Det ble gjort et utvalg oppgaver fra oppgavesettet på data som jeg oversatte til papirversjon. Pilotundersøkelsen bestod av 24 matematikkoppgaver. I disse matematikkoppgavene var 6 kategorier av FONS representert, og hver kategori fikk 4 oppgaver hver. Det vil si at det var en ekstra kategori av FONS representert i pilotundersøkelsen kontra oppgavesettet som ble brukt i datainnsamlingen senere. Under gjennomføring av pilotundersøkelsen ble det gjort et utvalg av tre respondenter i en av klassene i samarbeid med klassens kontaktlærer. Det ble tatt ut en høyt presterende-, middels presterende- og lavt presterende elev basert på kontaktlærerens vurderinger og disse gjennomførte oppgavesettet en og en. Ut fra disse gjennomførelsene kunne man konstatere at respondentene brukte ganske forskjellig tid på oppgavesettet og totalscoren på oppgavene også var veldig varierende.

3.3.1.1 Resultater pilotundersøkelse

HVEM	TID BRUKT	SVARSCORE
Lavt presterende elev	15 minutter 18 sekunder	7/24
Middels presterende elev	11 minutter 12 sekunder	13/24
Høyt presterende elev	8 minutter 59 sekunder	23/24

Tabell 3 Resultater pilotundersøkelse

Basert på disse resultatene så jeg at jeg kunne forvente at elevene kom til å bruke veldig variert tid på oppgavesettet og svarscoren kom også til å være veldig ulik. Da den sterke respondenten klarte alle oppgavene ganske kjapt og fikk nesten alt riktig måtte de andre to respondentene bruke mye lengre tid på hver oppgave og det var heller ikke alle oppgavene de fikk til eller skjønnte.

3.3.2 Endringer fra pilotundersøkelsen til datainnsamlingen

Basert på resultatene og erfaringene fra pilotundersøkelsen kunne jeg og veileder vurdere og endre på oppgavesettet til datainnsamlingen. Det ble gjort flere endringer fra pilotundersøkelsen til datainnsamlingen.

Det første vi endret var selve oppgavene i oppgavesettet. Basert på tiden respondentene brukte i pilotundersøkelsen vurderte vi at vi fint kunne legge til flere oppgaver, slik at det nye oppgavesettet ble med 26 oppgaver istedenfor 24. Vi bestemte oss også for å gå bort fra FONS-kategorien mengdediskriminering, slik at det ble fem FONS-kategorier representert istedenfor seks i det nye oppgavesettet. Det gjorde vi fordi oppgavene som knyttet seg til mengdediskriminering handlet om at respondentene skulle forstå begrepene mindre enn og større enn, og vi følte at vi hadde oppgaver fra før i andre FONS-kategorier som kunne knyttes til akkurat dette. For eksempel oppgave 7 og 8 (se vedlegg 2), hvor oppgaven er å finne ut hvilket tall som kommer etter 5 og før 8 som i prinsippet handler om at man også forstår mindre enn og større enn. I tillegg ønsket vi å ikke vektlegge alle FONS-komponentene likt, men heller ha flest oppgaver om aritmetisk kompetanse. Det gjorde vi for å balansere oppgavesettet slik at aritmetisk kompetanse skulle spille en større rolle enn det gjorde i oppgavesettet i pilotundersøkelsen for å se om respondentene kunne anvende de andre FONS-kategoriene til å løse aritmetiske oppgaver, for eksempel bruke systematisk telling til å løse addisjonsstykker.

Det ble også gjort endringer som handlet om hvor identisk oppgavene og gjennomførelsen var i de to oppgavesettene. En av disse endringene dreide seg om at jeg ikke ga identiske instruksjoner i pilotundersøkelsen til oppgavene på papir slik de er presentert i oppgavesettet på data. For at oppgavesettene skulle være helt identiske måtte jeg fokusere videre på å kopiere nøyaktig hvordan oppgavene ble instruert i oppgavesettet på data og skrive dem ned ordrett. Disse instruksjonene på hver oppgave i oppgavesettet på data ble grunnlaget for hvordan oppgavene i oppgavesettet på papir ble presentert som jeg ga til hele gruppen samlet i datainnsamlingen.

En annen endring under selve datainnsamlingen kontra pilotundersøkelsen var at vi gikk fra å gjennomføre en og en til grupper på 4-5 respondenter da vi så at det kom til å ta veldig lang tid dersom man skulle ta ut alle elevene en og en. Selv om det var 4-5 respondenter per gruppe var det lett å holde oversikten over alle sammen og notere ned aspekter man la merke til underveis som kan knyttes til resultatene og validiteten til studien.

3.3.3 Gjennomførelse

Gjennomførelsen av datainnsamling ble gjort på samme måte i begge oppgavesettene. Etter oppstarten i klasserommet med kontaktlæreren kom jeg inn og hentet de respondentene som skulle være med. De kom til ferdig oppsatt chromebook eller oppgaveark med blyant og viskelær med god avstand på et eget lukket rom, slik at det var bare å begynne å løse oppgavene. På forhånd hadde jeg skrevet ned kallenavnet til respondentene på hver sin chromebook eller oppgaveark. Før oppgaveløsingen fikk de noen instruksjoner de måtte forholde seg til, som ble gjennomgått i fellesskap:

- Det ble presisert at dette ikke var noen farlig test og at de var der for å hjelpe til å forske på matematikkoppgaver og at det ikke var noen konkurranse om å svare riktig eller raskt, men at alle skal gjøre sitt beste og svare det de tror er riktig.

Det var for at respondentene ikke skulle føle på presset til å prestere som kanskje ville gå ut over prestasjonsevnen der og da.

- Det ble også presisert at det ikke var lov til å prate sammen eller samarbeide underveis.
- De fikk instruksjoner om at dersom oppgavene var for vanskelige eller de ikke skjønnte oppgavene, var det greit å hoppe over oppgaven eller bare gjette på det svaret de tror er riktig.
- De fikk beskjed om at dersom de sto fast kunne de spørre meg om hjelp, men jeg kom ikke til å hjelpe de med riktig svar på oppgaven. Jeg kunne bare hjelpe til med å forstå oppgaven.
- Det ble også gjennomført en eksempeloppgave hvor respondentene fikk erfare hvordan oppgavene ble gjennomført. Denne eksempeloppgaven ble gjennomført slik at man ikke fikk vite det riktige svaret på oppgaven, men bare hvordan oppgavene var satt opp.

Den eneste vesentlige forskjellen i gjennomførelsen av oppgavesettet på data kontra oppgavene på papir var at de hadde muligheten til å gå videre til neste oppgave selv når de gjennomførte oppgavesettet på data. Det gjorde også at man kunne starte med oppgavene på forskjellig tidspunkt og respondentene ble også ferdig på ulikt tidspunkt. Da gikk de bare tilbake til klasserommet igjen for å ikke forstyrre de som enda satt med oppgavene. I gjennomførelsen av oppgavene på papir måtte jeg gi instruksjoner på hver oppgave i fellesskap. Det vil si at vi måtte vente til alle sammen var ferdig med en oppgave før vi gikk videre på neste. For å sikre meg om at alle var ferdig med oppgaven så laget vi en regel om at de skulle rekke opp hånda når de var ferdige med oppgaven, slik at jeg kunne se når vi kunne gå videre til neste oppgave.

3.4 Retting av oppgaver

Etter at respondentene løste oppgavene, ble besvarelsene rettet. Resultatene både på oppgavene på papir og på data ble skrevet i hvert sitt Excel-regneark.

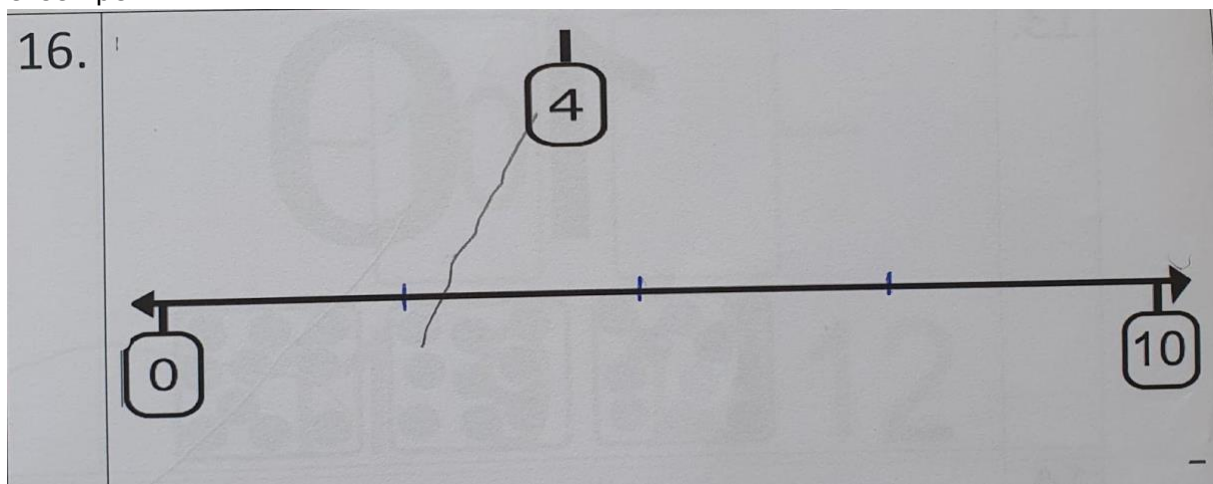
Begge oppgavesettene resultater var rettet på samme måte i Excel-regnearket. I den første kolonnen på regnearket sto respondentene sine kallenavn mens i den siste kolonnen sto totalscoren. I hver kolonne mellom kallenavnet og totalscoren sto enten tallene 1 eller 0. 1 betydde riktig svar på oppgaven mens 0 betydde feil svar, eller ingen svar, på oppgaven. Denne rettingen ble gjort manuelt av meg i oppgavene på papir, mens resultatene på oppgavesettet på data ble gjort automatisk og tilsendt meg av ansvarlig for programvaren på NTNU. På oppgavesettet på data fikk jeg, i tillegg til svarscore, også muligheten til å se på hvor lang tid respondentene brukte per oppgave og hvilket svar de svarte.

I rettingen av de fleste oppgavene på papir var det enkelt å se om respondentene svarte riktig eller galt på oppgavene, men særlig to oppgaver ga meg utfordringer underveis. Disse oppgavene var estimeringsoppgavene, som også var de eneste oppgavene som heller ikke ble rettet automatisk i oppgavesettet på data. I estimeringsoppgavene skulle respondentene plassere to tall, 4 og 8, på en åpen tallinje fra 0 til 10. I papiroppgaven måtte de sette strek fra 4- eller 8-tallet, som sto ovenfor tallinja, til den åpne tallinja mens i oppgaven på data kunne de fysisk flytte de samme tallene til den åpne tallinja. Da tallinjen var åpen betyr det at det ikke var lagt inn andre tall enn 0 og 10 på tallinja og

heller ingen indikasjoner på hvor tallene kunne stå, for eksempel med flere tomme ruter langs tallinja.

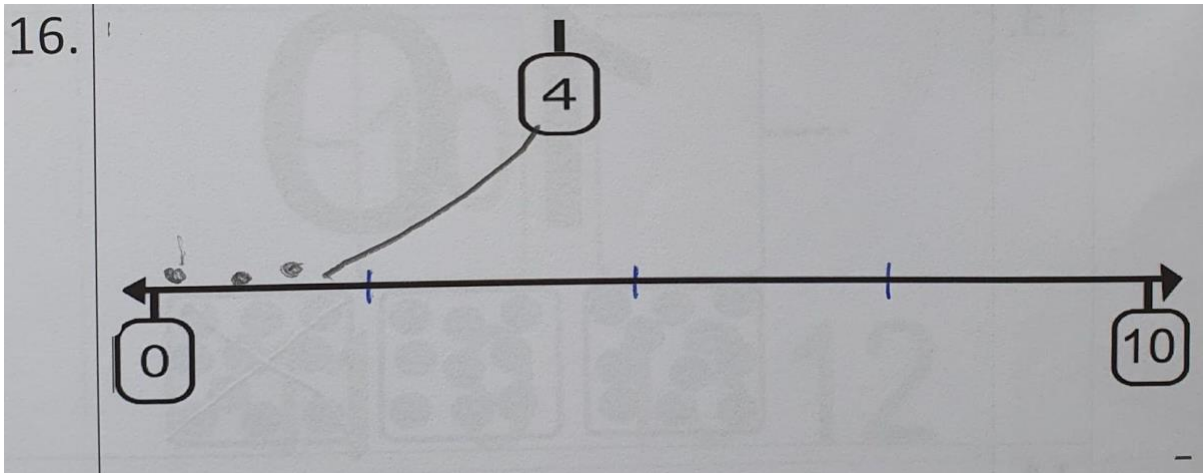
Jeg måtte dermed ta et valg om hva jeg ville godkjenne som riktig svar og ikke og forsøke å følge dette valget i begge oppgavesettene slik at rettingen skjedde på samme grunnlag. Hovedregelen i oppgaven der de skulle plassere 4-tallet ble at jeg ville godkjenne alle svar fra 3 til 5 mens med 8-tallet ville jeg godkjenne alle svar fra 7-9. Denne regelen ble enkelt praktisert under rettingen av oppgavesettet på data. Det kunne jeg gjøre fordi Excel-regnearket som jeg fikk tilsendt inneholdt opplysninger om hva respondentene svarte og blant annet ble det regnet ut hvor respondentene plasserte tallene basert på antall pixler totalt i tallinjen og antall pixler på hvor tallet ble plassert. Dermed fikk vi opplyst nøyaktige tallverdier for hvor respondentene plasserte de to tallene og det var bare å godkjenne alle disse svarene man så lå innenfor 3 og 5 eller 7 og 9.

I rettingen av disse oppgavene på papir måtte jeg ty til litt mer skjønn ettersom man ikke har de samme teknologiske mulighetene til å kalkulere svarene basert på pixler. Jeg måtte dermed bruke andre metoder, og valgte derfor å dele opp tallinjen i fire deler for å nyansere svarene. I oppgaven der respondentene skulle plassere 4-tallet godkjente jeg alle svar som var plassert litt over $\frac{1}{4}$ av tallinja, det vil si litt over 2.5 fram til litt under $\frac{1}{2}$ av linja ettersom det er et viktig aspekt i oppgaven at 4 ikke er halvparten av 10. I oppgaven der respondentene plasserte 8-tallet godkjente jeg alle svarene som var mellom litt før $\frac{3}{4}$ av tallinjen, som vil si rundt 7, til litt før 10, som vil si rundt 9. Det var noen besvarelser som var vanskelig å skille der jeg måtte inn å tegne hjelpestreker, for eksempel:



Figur 2 Eksempel på oppgave med hjelpestreker som ga riktig svar

Denne figuren viser et eksempel hvor hjelpestrekene var avgjørende for å vise om respondenten fikk riktig svar. Her ser vi at respondentens plassering av 4-tallet er over $\frac{1}{4}$ av tallinja og langt innenfor halvparten, som er innenfor det jeg godkjente som riktig svar. I etterkant ser jeg at jeg kunne ha gjort dette på en bedre måte for å få helt nøyaktige inndelinger av tallinja, for eksempel brukt en linjal med nøyaktig inndeling fra 0 til 10 som jeg kunne ha lagt over alle besvarelsene og dermed umiddelbart sett om streken lå innenfor de samme grensene som oppgavesettet på data.



Figur 3 Eksempel på oppgave med hjelpestreker som ga feil svar

Denne figuren viser et eksempel hvor hjelpestrekene var avgjørende for å vise at respondenten ikke fikk godkjent svaret. Her ser vi at selv om respondenten har telt riktig og viser forståelse for at 4-tallet kommer som nummer 4 tallrekka, er plasseringen av 4-tallet under $\frac{1}{4}$ av tallinja, som ikke er innenfor det jeg godkjente som riktig svar.

3.5 Metode for analyse: Variansanalyse gjennom SPSS

Da jeg har gjennomført en kvantitativ undersøkelse av et problemområde, vil jeg bruke statistikk i analysene av dataene. I denne oppgaven er den kvantitative rådataen blitt behandlet gjennom Statistical Package for the Social Sciences (SPSS) hvor man har muligheten til å lage forskjellige tabeller og fremstillinger av datamaterialet og hvor man også har muligheten til å analysere dataen på forskjellig måter.

Metoden for å måle tallforståelseskompetansen hos respondentene vil være gjennom en variansanalyse (ANOVA). I tillegg er det blitt gjort t-tester for å sammenligne to forskjellige gjennomsnitt for å finne ut om de er like (Skog, 2005). De gjennomsnittene som i første omgang er relevant å sammenligne med en t-test i denne studien er respondentenes score i de to ulike oppgavesettene. Ved en ANOVA har man muligheten til å sammenligne data fra to eller flere grupper samtidig (Skog, 2005). En ANOVA vil være hensiktsmessig da jeg skal se på data fra to grupper hvor den ene gruppen er respondentenes resultater på oppgavesettet på data mens den andre gruppen er de samme respondentenes resultater på oppgavesettet på papir. I tillegg vil en ANOVA i dette tilfellet også se om rekkefølgen har noe å si, det vil si om det er forskjell i resultatene om respondentene løste oppgavesettet på papir før oppgavesettet på data eller motsatt. Det at det er de samme respondentene som inngår i begge gruppene gjør også at jeg må ta en «repeated measures ANOVA» i SPSS, som betyr at analysen av datamaterialet tar hensyn til at det er de samme respondentene i hver gruppe. Til hjelp med å finne riktig måte å analysere dataen på har jeg fått god hjelp av Andy Field sin bok om alle mulige statistiske muligheter man har gjennom bruken av SPSS (Field, 2017).

For å ta en repeated measures ANOVA-test er dataen plottet inn i SPSS. Gjennom analyseverktøyet i SPSS får man muligheten til å se ulike verdier knyttet til en repeated measures ANOVA-test, blant annet en p-verdi gitt mellom 0-1. P-verdien brukes til å tolke hvor stor sannsynlighet det er for at det man observerer er tilfeldig i utvalget eller

ikke, og for å akseptere eller forkaste en nullhypotese. Nullhypotesen vil alltid ha utgangspunkt i at det ikke er forskjell mellom gruppene. For å akseptere eller forkaste en nullhypotese trenger man å angi p-verdien opp mot en signifikansverdi for å bestemme hvor stor usikkerhet man aksepterer, eller med andre ord hvor stor sannsynlighet det er for at man tar feil konklusjon. Vanligvis angir man signifikansverdien til å være 0,05. Det betyr at dersom p-verdien er mindre enn 0,05 betyr det at man kan forkaste nullhypotesen noe som resulterer i at det mest sannsynlig er en forskjell mellom gruppene. Dersom p-verdien er større enn 0,05 betyr det det motsatte, at man kan akseptere at nullhypotesen stemmer. Av og til kan man forkaste eller akseptere nullhypotesen på feil grunnlag, og det finner to måter å ta feil konklusjon på. De kalles type 1 og type 2 feil (Field, 2017). Type 1 feil er hvis nullhypotesen er sann, men vi forkaster den. Det betyr at dataene våre viser at det er en forskjell mellom gruppene, mens det i virkeligheten ikke er en forskjell. Type 2 feil er hvis nullhypotesen ikke er sann, men vi forkaster den ikke. Det betyr at dataene våre ikke viser at det er en forskjell mellom gruppene, mens det i virkeligheten er en forskjell.

I tillegg har jeg brukt SPSS til å lage grafiske framstillinger som brukes til å studere resultatene nøyere. Disse grafiske framstillingene er representert som søylediagram, hvor jeg for eksempel får muligheten til å sammenligne hver oppgave for å se om det var noen forskjell i svarscoren mellom oppgavesettet på data og på papir i de ulike kategoriene av FONS eller per oppgave. I mange av søylediagrammene er det i tillegg med standardfeil, som forklarer oss feilmarginen for gjennomsnittene som er funnet.

3.6 Forskningens reliabilitet og validitet

Målet med dette forskningsprosjektet er å se om det er forskjell om man måler tallforståelse til 1.klasseelever med to ulike prøveformer, hvor den ene prøveformen er skriftlig på papir og den andre er på data. Svarscoren og resultatene fra dette forskningsarbeidet kan i første omgang ikke generaliseres til å gjelde alle elever på 1. trinn, til det trenger man langt større dataomfang i ulike miljøer. Selv om resultatene ikke kan generaliseres til alle elevene på 1. trinn gir svarene fra studien likevel en indikasjon på om prøveformen har noe å si. Det kan jeg si ettersom jeg har fordelen med at det er de samme elevene som har gjort begge oppgavesettene for å gi et svar på forskningsspørsmålet. Elevene i denne studien har dermed fått testet ut to prøveformer, og dersom det viser seg at det er en statistisk signifikant forskjell mellom prøveformene har jeg et godt grunnlag til å si at det samme vil gjelde for andre elever på 1. trinn.

Med andre ord vil det være interessant å knytte dette emnet opp mot videre forskning mot blant annet reliabilitet som ifølge Cohen (Cohen et al., 2018) hovedsakelig handler om i hvilken grad man kan gjennomføre samme studie til samme type respondenter, som i dette tilfelle vil si andre 1.klasseelever, og forvente omtrent de samme svarene. Dette kan man som tidligere nevnt ikke alltid forvente, men man kan si at denne studien er en reliabilitetsstudie likevel ettersom respondentene gjennomfører først matematikkoppgaver på en måte og senere skal de samme respondentene gjennomføre de samme matematikkoppgavene, bare på en annen måte. Respondentene gjennomfører dermed en test etterfulgt av en retest, som knytter seg til reliabilitet som ekvivalens (Cohen et al., 2018). Dersom det viser seg at svarene på denne type test-retest-metoden viser lav grad av reliabilitet, som vil si at det er stor forskjell på svarene i en av prøveformene, kan det derfor være verdt å forske på hvorfor resultatene blir forskjellige, men dette er ikke noe som denne studien tar høyde for.

I forskning må man også underveis reflektere over validitet, som kan oversettes til gyldigheten knyttet til forskningen. Tjora (Tjora, 2013) knytter validitetsbegrepet i forskning opp mot spørsmålet om de svarene man finner i sin forskning faktisk er svar på det spørsmålet man forsøker å stille. Cohen (Cohen et al., 2018) går så langt og sier at dersom svarene ikke gir svar på det man faktisk stiller spørsmål om, er forskningen rett og slett ubrukelig. Da jeg i dette forskningsarbeidet stiller spørsmål som handler om tallforståelseskompetanse hos hver enkelt respondent er det viktig at jeg er sikker på at det er nettopp denne respondenten sine svar som kommer fram. Som et ledd i å vurdere dette legger jeg ved utdrag fra observasjonene jeg gjorde meg underveis i datainnsamlingen fra uke 46 og uke 47.

UKE/DAG	OBSERVASJONER
Uke 46	
Mandag	Gjennomføring datainnsamling for klasse A + B, oppgavesett på data. Respondentene fulgte instruksjoner til punkt og prikke. Ble ferdige på forskjellige tidspunkt, men det gikk veldig fint. Stille og rolig jobbing, god konsentrasjon og ingen forstyrrelser til tross for utsatt rom. Min rolle ble veldig passiv mens de løste oppgavene.
Tirsdag	Gjennomføring datainnsamling for klasse C + D, oppgaver på papir. Større krav til min egen rolle. Respondentene virker lettere påvirket av omgivelser rundt og det er vanskeligere å konsentrere seg om oppgavene og lytte til meg. Observerer at noen respondenter ser på andre sine svar eller deler svar høyt, særlig på grunn av at man må vente på at alle skal bli ferdig med hver oppgave. Min rolle går fra å være passiv til å være veldig delaktig.
Uke 47	
Mandag	Gjennomføring datainnsamling for klasse A + B, oppgaver på papir. Mer delekultur og juksekultur sammenlignet med forrige uke. Lavere konsentrasjon. Må forklare oppgavene mye mer nøye selv om de gjorde de samme oppgavene for en uke siden.
Tirsdag	Gjennomføring datainnsamling for klasse C + D, oppgavesett på data. Respondentene ble positivt overrasket at de skulle gjøre matematikk på chromebook. Jobber i sin egen boble uten å forstyrre de andre i rommet. Observerer at noen respondenter sier at "dette er gøy".

Tabell 4 Oversikt over observasjoner gjort underveis i datainnsamlingen

Ut fra disse observasjonene ser man at man kan stille spørsmål til om man faktisk samler inn enkelte respondenter sine svar for å vurdere tallforståelseskompetansen deres i noen tilfeller, særlig under gjennomførelsen på papir. Min rolle som deltakende observatør gjorde at jeg kunne se at noen respondenter ikke stolte på seg selv og prøvde å se på andre sine svar under hele gjennomførelsen. Av de 45 svarene som er samlet inn vil jeg luke bort fire av dem når svarscoren skal analyseres rett og slett på grunn av at jeg ikke er sikker på om det faktisk er deres svar og kompetanse man finner i oppgavene eller om de er kopiert fra andre. I tillegg vil jeg luke bort ytterligere en besvarelse på grunn av at det var en respondent som ikke ville fullføre oppgavene, verken på papir eller data.

3.7 Etiske hensyn

Med tanke på etiske hensyn følger dette forskningsarbeidet de retningslinjer utarbeidet av NESH, som er den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH, 2016). Disse retningslinjene er fulgt etter deres standard i dette forskningsarbeidet og i samarbeid med norsk senter for forskningsdata (NSD) har informasjonen om forskningsarbeidet blitt sendt inn og godkjent ut fra deres krav (Referansenummer: 659374). For at det skulle bli godkjent som forskningsarbeid var det særlig følgende to punkter som måtte være nøyte utarbeidet i søknaden om godkjenning:

3.7.1 Forskning på barn

Dette er et forskningsarbeid som baserer seg på forskning på barns tallforståelseskompetanse. Når barn og unge deltar i forskning, har de særlige krav på beskyttelse (NESH, 2016). Som forsker trenger man kunnskap om barn til å tilrettelegge metode og innhold til aldersgruppen som deltar. Som student på 6. året og gjennom jobb i skolen i snart fem år er dette noe jeg har lært meg og erfart og i tillegg til et samarbeid med NTNU har vi kommet fram til oppgaver og metode vi mener passer til alderen. Likevel, når denne forskningen baserer seg på tallforståelseskompetanse hos flere respondenter vil det si at ikke alle kan få til alle oppgavene, noe som også er hensikten med studien. Det vil si at det er tatt i betraktning at barn nede i fem og seks år er i utvikling og har forskjellige ferdigheter på dette stadiet og det er ønskelig at mange ikke skal få til alle oppgavene. NESH trekker også fram at det er viktig at respondentene i undersøkelsen fikk informasjon om at dette var frivillig. Dette prøvde jeg å opprettholde ved å gi foreldrene muligheten til å trekke seg i etterkant dersom de ønsket det, men også ved å gi informasjon til respondentene om at dersom de ikke ønsket å være med gikk det helt greit. Det virket ikke som om noen av respondentene tenkte at de ikke hadde lyst til å være med, men NESH trekker fram at barn ofte er mer villige til å adlyde autoriteter enn det voksne er og at de føler at de ikke kan protestere. Dette ble også tilrettelagt og jeg opplevde en respondent som gikk fra oppgavene fordi han ikke hadde lyst til å gjennomføre dem. Da fikk han lov til å forlate rommet uten noen problemer.

3.7.2 Samtykke

I dette forskningsarbeidet er jeg pliktig i å gi informasjon om forskningen og forskningens formål ettersom jeg har informasjonsplikt ovenfor deltakerne om hvilke personopplysninger som vil bli brukt i forskningen. Et detaljert informasjonsskriv til foreldre var derfor nødvendig for å gjennomføre datainnsamlingen (vedlegg 1). Dette informasjonsskrivet inneholdt blant annet hva forskningen handlet om, hva dette innebærer for respondentene, hvilke personopplysninger som kommer til å være med og et samtykkeskjema hvor foreldrene kunne krysse av for hva de ville at barna deres skulle være med på med underskrift. Målet med informasjonsskrivet var at all informasjonen om forskningen kommer fram i selve skrivet uten at foreldrene selv skal måtte oppsøke den. I denne forskningen er respondentenes navn med, og jeg var pliktig i å informere om dette. Navnet er for øvrig på et eget dokument som er kryptert, og det er kun jeg som har tilgang til dette dokumentet. Som tidligere nevnt er det kun de elevene som har fått samtykke av foreldrene til å være med som er respondenter i forskningsarbeidet, og respondentene fikk også muligheten til å si nei selv til å være med. I samarbeid med NSD ble derfor dette informasjonsskrivet utarbeidet nøyte slik at samtykket i dokumentet er fritt, informert og uttrykkelig som NESH trekker fram at samtykket må være (NESH, 2016).

4 Resultater

I denne delen av oppgaven vil jeg presentere resultatene fra forskningsprosjektet. Først vil jeg se på forskjellene i begge oppgavesettene som helhet, som betyr at her vil svaret på forskningsspørsmålet bli presentert. Jeg finner ut at i det helhetlige bildet er det ikke noen forskjell på svarscoren på papir og data, men arbeidet med å se på resultatene gjør at jeg også finner ut at det er forskjell på resultatene for første og siste gjennomførelse uansett hvilket oppgavesett som ble gjort først. Disse to funnene vil først bli sett på helhetlig for hele oppgavesettet. Deretter vil det helhetlige bildet bli delt inn i mindre deler, hvor jeg først skal se på begge funnene opp mot resultatene fra hver FONS-kategori og til slutt se på funnene opp mot resultatene på hver oppgave. Jeg finner også ut at det er vesentlige forskjeller på hvordan formatet påvirker gjennomføringen av oppgavene, for eksempel knyttet til validiteten av studien. Dette vil bli presentert mot slutten av kapittelet.

4.1 Forskjeller i oppgavesett som helhet

Kapittel 4.1 vil svare på forskningsspørsmålet og resultatene i sin helhet.

4.1.1 Forskjeller papir og data

I denne delen av kapittelet vil jeg studere resultatene for å få et overordnet bilde av forskningsspørsmålet:

Hvilke forskjeller i tallforståelseskompetanse finner man dersom elever på 1. trinn løser identiske matematikkoppgaver på data og på papir?

For å gi et kvantitativt svar på denne problemstillingen har jeg plottet rådataene inn i SPSS og utført en repeated measures ANOVA-test av disse. I SPSS legger jeg inn tre variabler som vil svare på problemstillingen, hvor den ene variabelen viser hvor mange riktige svar hver respondent har på papir, den andre viser hvor mange riktige svar hver respondent har på data og den siste er for å vise hvilken rekkefølge respondentene gjorde oppgavesettene i (DP eller PD). Riktige svar på data og på papir er derfor uavhengige variabler mens rekkefølgen oppgavesettene ble gjort i er en avhengig variabel. På den måten får jeg sammenlignet resultatene fra oppgavesettet på papir og på data samtidig opp mot hvilken rekkefølge oppgavesettene ble gjort i. Når dette er gjort kan jeg lese av p-verdien, for å akseptere eller forkaste om det er en forskjell i resultatene på de to oppgavesettene. Jeg har satt signifikansverdien til 0,05 som betyr at dersom p-verdien er mindre enn 0,05, vil det bety at man kan forkaste nullhypotesen om at det ikke er forskjell i resultatene mellom oppgavesettene. Dersom p-verdien er større enn 0,05 vil det bety det motsatte, i dette tilfellet at man kan konkludere eller akseptere at det ikke er en forskjell i resultatene på de to oppgavesettene. I tillegg har vi faktoren som er om rekkefølgen på oppgavesettene har noe å si for resultatene. Er p-verdien mindre enn 0,05 her kan vi forkaste nullhypotesen om at rekkefølgen på oppgavesettene ikke har noe å si på resultatene mens er p-verdien høyere enn 0,05 kan man akseptere at rekkefølgen på oppgavesettene ikke har noe å si på resultatene. Vi fikk disse p-verdiene:

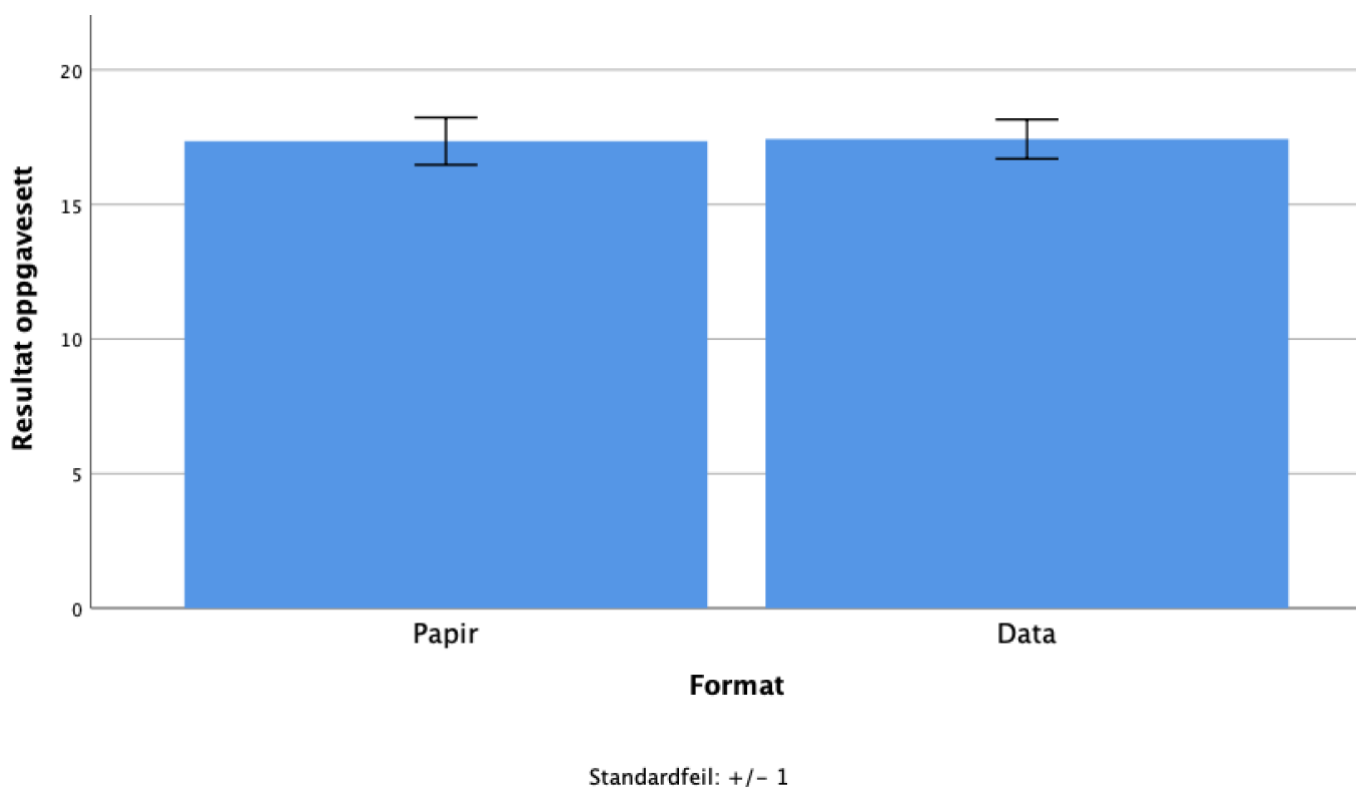
Kilde	P-verdi
Format	0,754

Rekkefølge	0,392
Format * Rekkefølge (tolkes som <i>tidspunktet</i> (første og siste gjennomførelse))	0,005

Tabell 5 P-verdiene fra repeated measures ANOVA-test

P-verdiene fra repeated measures ANOVA-testen gjør at vi kan tolke resultatene med at p-verdien for formatet ($p=0,754$) betyr at vi kan akseptere nullhypotesen om at det ikke er forskjell i resultatene mellom de to oppgavesettene. P-verdien for rekkefølgen ($p=0,392$) gjør også at vi kan akseptere nullhypotesen om at det ikke er forskjell på hvilken rekkefølge oppgavesettene ble gjort i og det har ikke noe å si om respondentene er i den ene gruppen eller den andre. P-verdien for formatet opp mot rekkefølgen ($p=0,005$), som i praksis betyr forskjellen på respondentenes første og siste gjennomførelse, viser derimot at vi kan forkaste nullhypotesen om at resultatene ikke er forskjellige under første og siste gjennomførelsen. Denne p-verdien blir støttet opp mot en t-test mellom respondentenes første og siste gjennomførelse som også viste statistisk signifikant effekt med $p=0,005$ og vil være et betydningsfullt funn ved siden av forskningsspørsmålet i denne studien.

Vi kan også se på resultatene grafisk:



Figur 4 Resultater papir og data med standardfeil

Figur 4 viser gjennomsnittscoren for alle respondentene på papir og data. Ut fra denne fremstillingen ser vi at gjennomsnittscoren er nesten helt lik på begge oppgavesettene og det er omtrent ingen forskjell i resultatene. Totalt kunne man få 26 poeng da det var 26 oppgaver i oppgavesettene.

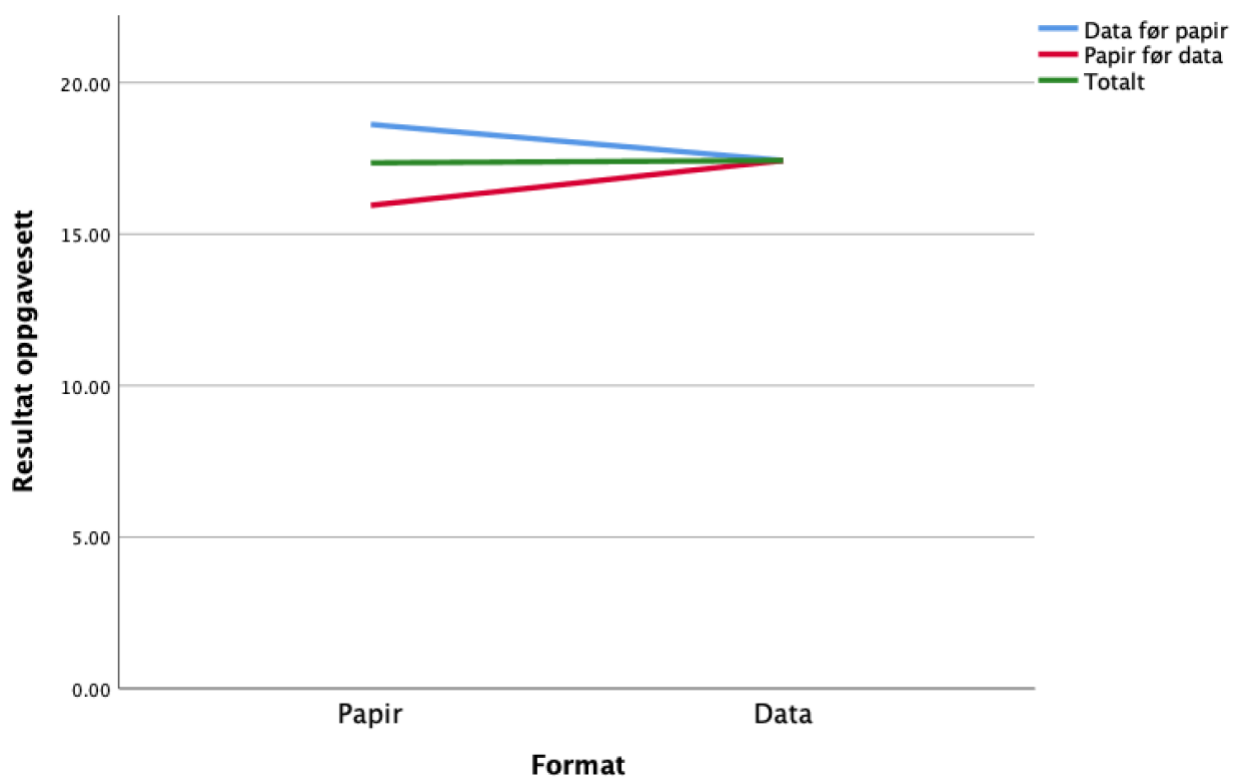
I tillegg kan vi se på resultatene på de to oppgavesettene ut fra de to forskjellige gruppene av respondenter, hvor den ene gruppen er respondentene som gjorde

oppgavene på data før papir og motsatt (DP og PD). Verdiene er hentet fra den beskrivende statistikken som kom til syne under repeated measures ANOVA-testen:

Data	DP	17.43
	PD	17.42
	Total	17.43
Papir	DP	18.62
	PD	15.95
	Total	17.35

Figur 5 Resultater oppgavesett med rekkefølgen på gjennomføring

Verdiene fra figur 5 kan man også fremstille med en linjemodell:



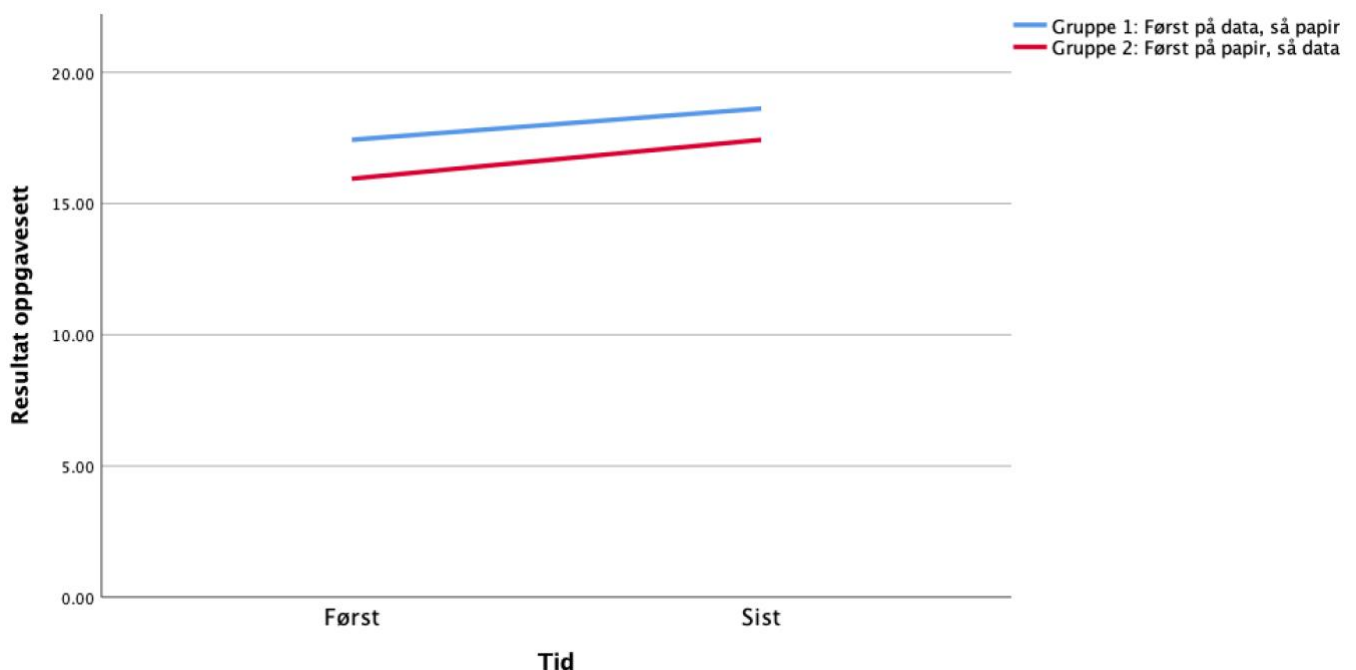
Figur 6 Resultater oppgavesett papir og data med linjemodell

Denne fremstillingen viser gjennomsnittlig poeng totalt for respondentene i begge oppgavesettene, med et skille om de gjennomførte først på data og så papir eller motsatt. Den blå linjen representerer de respondentene som gjennomførte først på data og så på papir. Her ser vi at gjennomsnittscoren er høyere i oppgavesettet på papir enn oppgavesettet på data. Dette resulterer i at vi ser at linjen peker nedover, som vil si at respondentene samlet sett fikk høyest gjennomsnittspoengsum på oppgavesettet på papir, som var det oppgavesettet de gjennomførte sist. Den røde linjen representerer respondentene som gjennomførte først på papir, og så på data. Her ser vi at gjennomsnittscoren er høyere i oppgavesettet på data enn på oppgavesettet på papir. Dette resulterer i at linjen har motsatt retning, hvor den peker oppover. Det vil si at respondentene samlet sett fikk høyest gjennomsnittspoengsum på oppgavesettet på data, som også var det de gjennomførte sist. I tillegg er det en grønn linje som viser

gjennomsnittscoren til hver av oppgavesettene for alle respondentene til sammen. Vi ser her at denne linjen er tilnærmet flat, som vil si at denne linjen viser at det ikke er noen stor forskjell på totale riktige svar mellom de to oppgavesettene, som vi også ser i figur 4. Denne fremstillingen bekrefter p-verdien for formatet ($p=0,754$) i tabell 5 hvor vi ser, ut fra den grønne linja, at det samlet sett ikke er forskjell i gjennomsnittscoren til respondentene i de to oppgavesettene. I tillegg viser denne fremstillingen at det er forskjell i første og siste gjennomførelse hvor vi ser at gjennomsnittscoren til respondentene er høyere under siste gjennomførelse uansett om de gjennomførte først på papir eller på data.

4.1.2 Forskjeller første og siste gjennomførelse

Funnene over ga oss et annet funn i studien, at det er forskjell på resultatene under første og siste gjennomførelse. Hvis vi ser på resultatene opp mot hvilket av oppgavesettene respondentene gjennomførte først og sist kan det fremstilles slik:

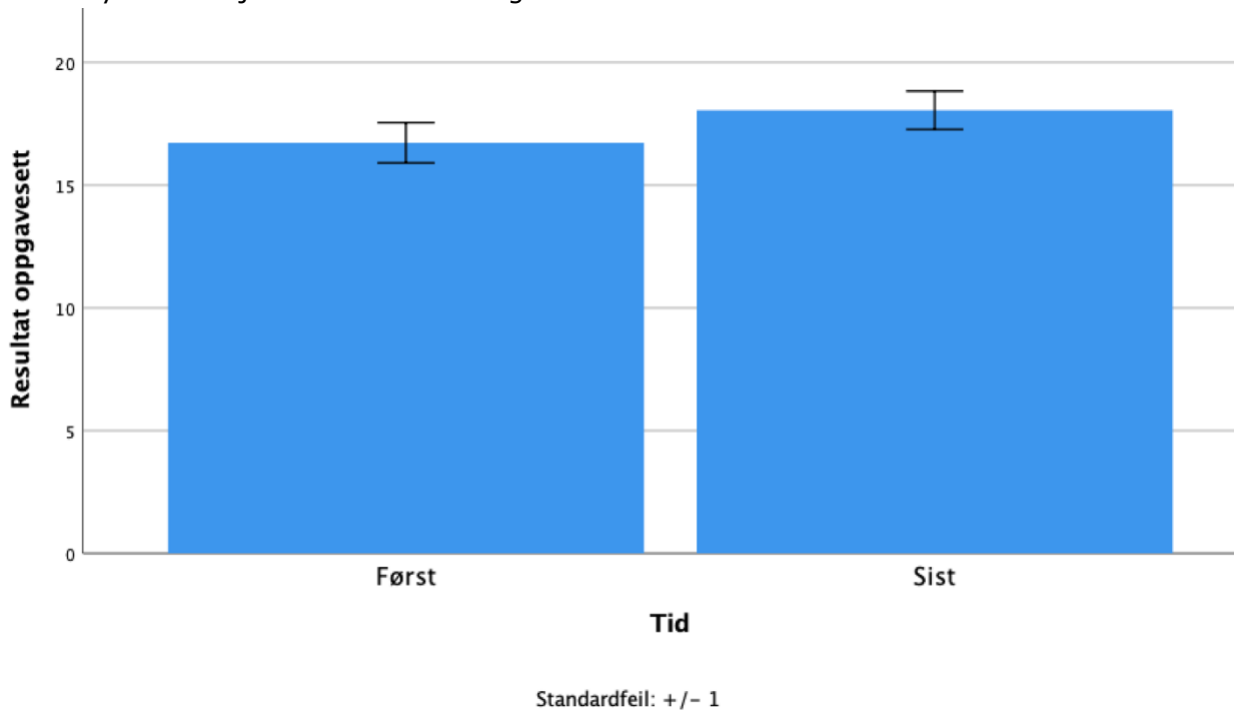


Figur 7 Linjediagram som viser forskjell på første og siste gjennomførelse

Her ser vi to grupper, hvor gruppe 1 (blå linje) er respondentene som gjennomførte først oppgavesettet på data og så på papir, mens gruppe 2 (rød linje) er respondentene som gjennomførte først oppgavesettet på papir og så på data. Denne fremstillingen viser, ved at begge linjene peker oppover og at gjennomsnittscoren i oppgavesettet er høyere siste gjennomførelse, at i det store og det hele så presterer begge gruppene bedre under siste gjennomførelse uansett hvilken rekkefølge de fikk de ulike formatene på. Vi ser også at den blå linjen har høyere score først og sist sammenlignet med den røde. En mulig tolkning til dette kan være basert på tilfeldigheter, hvor det er forskjell på gruppens gjennomsnittscore, men som ikke er statistisk signifikant. En annen tolkning kan være at rekkefølgen på formatet har betydning, som figur 5 kanskje indikerer, hvor det kan se ut som det var fordelaktig å gjennomføre oppgavesettet på papir etter man hadde gjennomført et oppgavesett på data fra før.

Ved å bruke motvektsprinsippet (se metodedelen, kapittel 3.3, s. 14-15), kan vi slå sammen de 4 gruppene som oppstår til å bli to grupper hvor en gruppe er

respondentenes første gjennomførelse (gruppe 1 + gruppe 3) og en annen gruppe er respondentenes siste gjennomførelse (gruppe 2 + gruppe 4). Ved å slå sammen disse gruppene skiller man heller ikke på hvilken type oppgavesett respondentene gjennomførte først da man har inkludert begge prøveformene. Når disse gruppene er slått sammen kan man fremstille forskjellen mellom første og siste gjennomførelse grafisk og knytte det fysisk opp mot p-verdien ($p=0,005$) fra ANOVA-testen og t-testen som betyr at forskjellen er statistisk signifikant.

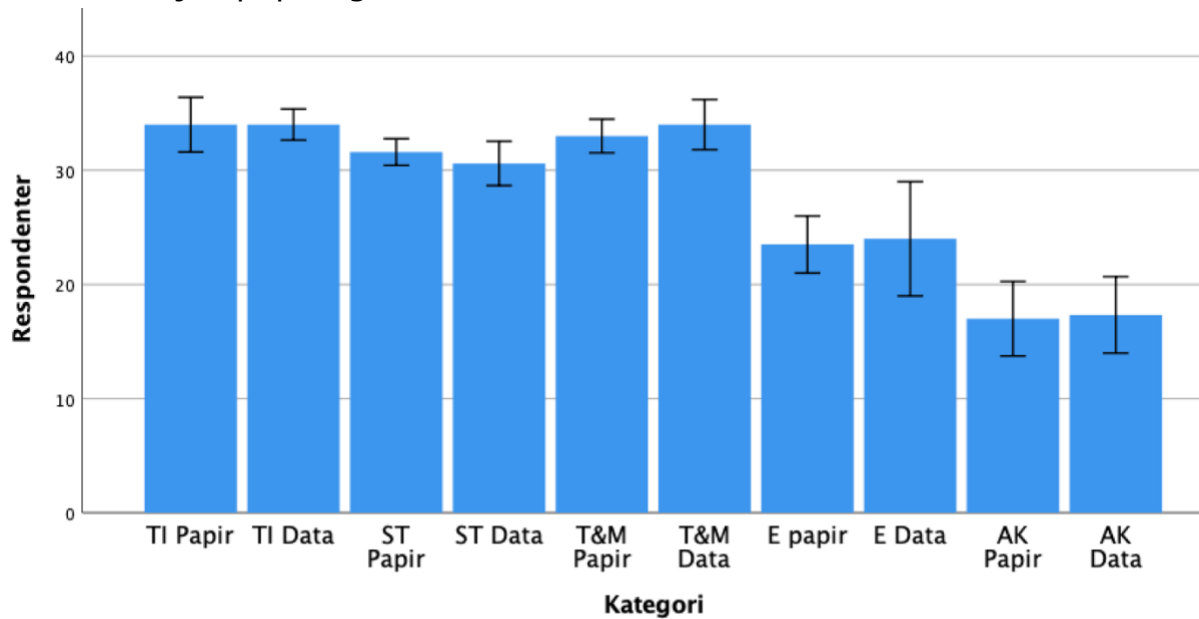


Figur 8 Forskjell første og siste gjennomførelse for alle respondentene

4.2 Forskjeller i FONS-kategorier

I kapittel 4.2 vil jeg dele opp det helhetlige bildet for å se om det kan være forskjeller i de ulike FONS-kategoriene, selv om det helhetlige bildet viser at det ikke er noen forskjell. Det kan være interessant å gjøre fordi det kan hende at noen FONS-kategorier er lettere på papir eller på data, som kanskje vil skjule seg i det helhetlige gjennomsnittet. For å gjøre det har jeg sett på resultatene på de oppgavene som tilhører samme kategori.

4.2.1 Forskjell papir og data

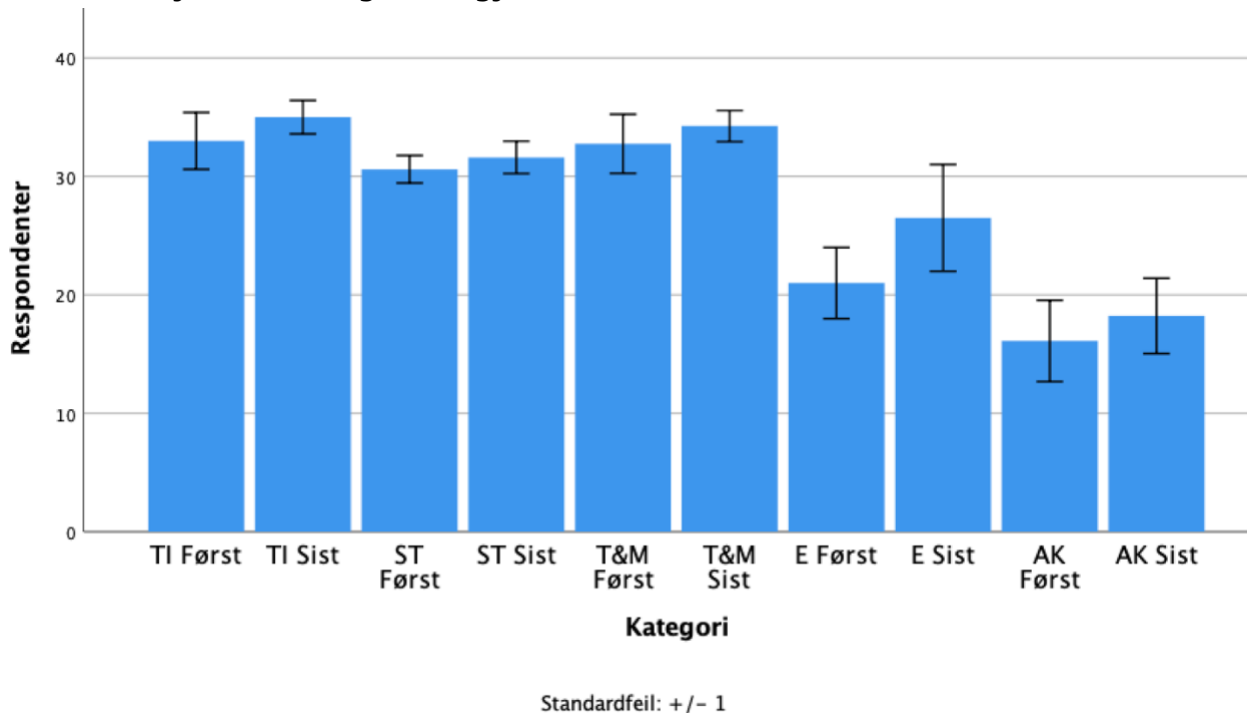


Standardfeil: +/- 1

Figur 9 Forskjell papir og data for hver FONS-kategori med standardfeil

Figuren over viser gjennomsnittet av hvor mange av de 40 respondenter som klarte oppgavene innenfor hver FONS-kategori, hvor fokuset er på papiroppgavene vs. dataoppgavene. Vi ser her at for hver kategori er også forskjellen små, som støtter seg opp mot det helhetlige funnet. Gjennomsnittsverdiene forteller oss at i to kategorier, tallidentifikasjon og estimering, er gjennomsnittet omtrent likt, i systematisk telling er gjennomsnittet litt bedre i favør oppgavene på papir mens i tall og mengde og aritmetisk kompetanse er gjennomsnittet litt bedre i favør data.

4.2.2 Forskjell første og siste gjennomførelse



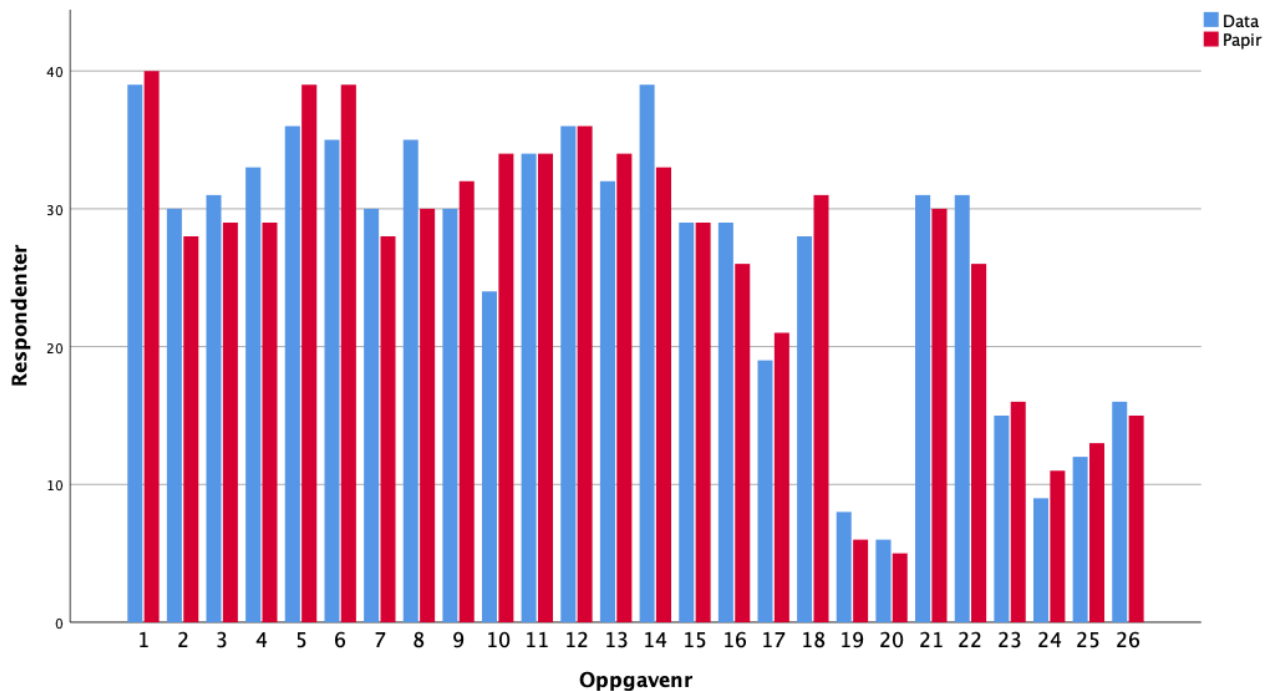
Figur 10 Forskjell først og sist for hver FONS-kategorier med standardfeil

Figuren over viser gjennomsnittet av hvor mange av de 40 respondentene som klarte oppgavene opp mot hver FONS-kategori, hvor fokuset er på første gjennomførelse vs. siste gjennomførelse. Det helhetlige bildet viser at respondentene presterer bedre på siste gjennomførelse, og vi ser at dette også er tilfellet når vi deler det helhetlige bildet inn i hver FONS-kategori. Spesielt ser vi at søylen er høyere under siste gjennomførelse i estimeringsoppgavene, som betyr at flere respondenter klarte estimeringsoppgavene siste gjennomførelse.

4.3 Forskjell i enkeltoppgaver

Vi skal til slutt dele det helhetlige bildet inn i hver enkelt oppgave. Det gjør vi for å se om det likevel er noen enkeltoppgaver som ikke stemmer overens med det vi allerede har funnet ut, hvor resultatene er at det ikke er noen forskjell på papir og data og at respondentene presterer bedre under siste gjennomførelse.

4.3.1 Forskjell papir og data



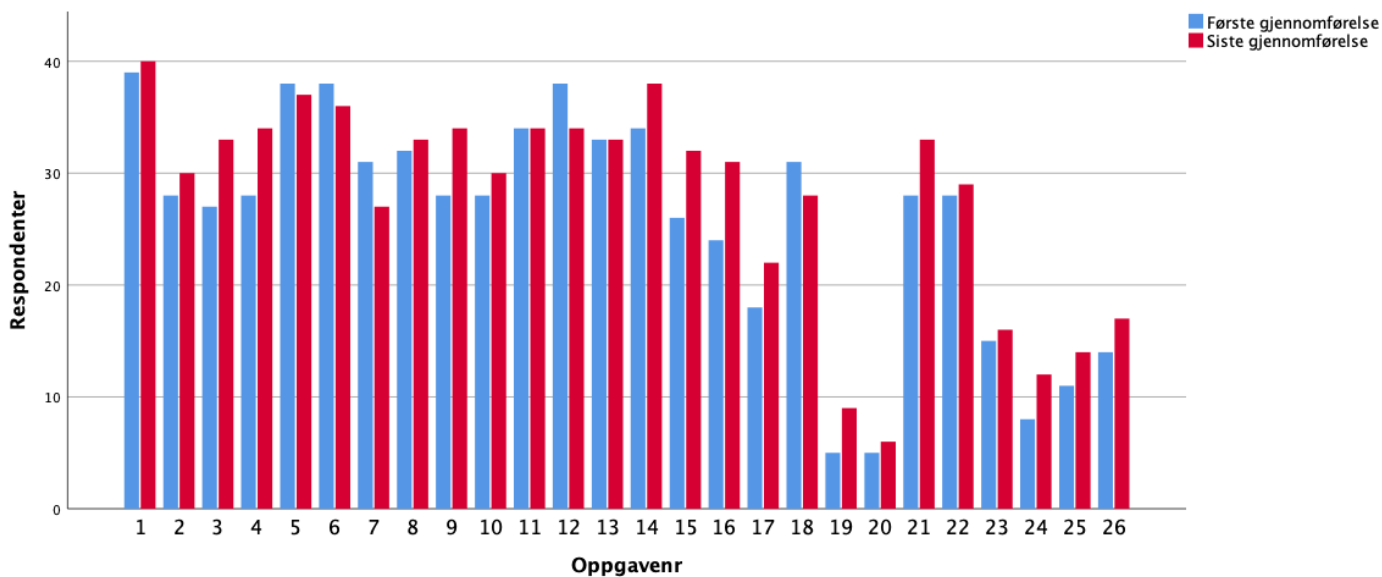
Figur 11 Forskjell papir og data for hver enkelt oppgave

Figuren over viser hvor mange av de 40 respondentene som klarte hver oppgave på papir og data. Denne fremstillingen viser at det er noen små forskjeller for hver oppgave som kan bety at noen oppgaver er enklere på papir eller på data, men som ikke vises i det helhetlige bildet. I flesteparten av oppgavene er disse forskjellene ganske små, men særlig to oppgaver skiller seg ut hvor det kan se ut som disse oppgavene favoriserer en av prøveformene. Vi kan se på hvilke oppgaver det gjelder, og om det finnes noen naturlig forklaring på hvorfor det er store forskjeller i papir og data eller om det bare er en tilfeldighet.

Opgave 10: 34 respondenter har riktig svar på papir, mens det er 24 som har riktig svar på data. Det vil si at det er 10 stykker som ikke fikk til oppgaven på data, men som fikk den til på papir. I denne oppgaven skal respondentene vise at de forstår hvor den andre er i en rekke på fem. Grunnen til at denne oppgaven favoriserer papir kan være av ulike årsaker. Det kan være tilfeldig at det er dette avviket, men det kan også komme av at under oppgavesettet på data var dette den første oppgaven som ikke inneholdt noen form for tall eller siffer, som vil si at det var en ny oppgavetype som respondentene ikke var kjent med fra før. Eller så var denne oppgaven enklere å få til på papir da respondentene hadde muligheten til å fysisk telle seg bortover med blyanten de hadde tilgjengelig.

Opgave 14: 33 respondenter har riktig svar på papir, mens det er 39 som har riktig svar på data. I denne oppgaven skal respondentene telle antall baller i en sirkel og vise med riktig siffer hvor mange baller det er. Grunnen til at denne oppgaven kan favorisere data kan være at hotspots, som Stacey og Wiliam (Stacey & Wiliam, 2012) nevner som et teknologisk potensial, var avgjørende til at respondentene så at de markerte riktig tall, og det ble veldig klart, uten at de måtte tegne kryss selv, hvilket tall som var deres svar. Igjen, det kan også bare være tilfeldig.

4.3.2 Forskjell første og siste gjennomførelse



Figur 12 Forskjell første og siste gjennomførelse for hver enkelt oppgave

Figur 12 over viser alle 26 oppgavene og hvor mange av respondentene som klarte de forskjellige oppgavene, hvor den blå søylen viser første gjennomførelse mens den røde søylen viser siste gjennomførelse. Her ser vi at for alle 26 oppgavene er det kun seks oppgaver som har større andel riktige svar under den første gjennomførelsen, det er en oppgave som har like stor andel riktige svar i begge gjennomførelsene mens det er hele 19 oppgaver som har større andel riktige svar under den siste gjennomførelsen. Dette beviser at respondentene presterer bedre under den siste gjennomførelsen, men man ser likevel dette ikke gjelder for alle oppgavene. De seks oppgavene som har en større andel riktige svar under den første gjennomførelsen er oppgave 5, 6, 7, 12, 18 og 20. En kvalitativ studie kunne gitt mer nøyaktige svar på hvorfor disse oppgavene har en større andel riktige svar under den første gjennomførelsen, men jeg kan likevel gjøre noen antakelser. Ingen av oppgavene gir umiddelbare tanker om hvorfor respondentene presterte bedre under første gjennomførelse på akkurat disse, mens det ikke var tilfellet på de fleste andre. Det kan være tilfeldig, det kan skyldes tilfeldigheter som gjetting eller slurvefeil eller at respondentene var ufokuserte akkurat i gjennomføringsøyeblikket. Det kan også skyldes at respondentene akkurat hadde vært igjennom lignende oppgaver fra før.

4.4 Forskjeller i gjennomføring

Til slutt i resultatkapittelet vil jeg presentere andre funn som kom til syne under datainnsamlingen og forskningsarbeidet som vil være relevant for videre forskning på emnet. Spesielt opp mot organisering av slike prøver og validiteten til videre studier.

4.4.1 Lærerens og elevenes rolle under gjennomførelsen av begge oppgavesettene

Dette forskningsarbeidet viser at for elevenes del spiller det ikke en veldig stor rolle om de løser oppgavesett på papir eller på data for å måle tallforståelseskompetansen deres. Men etter å ha gjennomført datainnsamlingen til dette forskningsarbeidet selv, hvor jeg har gjennomført et oppgavesett på papir og på data har jeg erfart aspekter med tanke på

lærerens- og elevenes rolle i slike situasjoner hvor man skal samle inn data for å vurdere kompetansen til elever.

I gjennomførelsen av papiroppgavene merket jeg at det var mange faktorer som gjorde at jeg fort mistet kontroll og oversikt over respondentene, faktorer som man ikke måtte ta hensyn til under gjennomførelsen på data. Som vi ser på observasjonene/refleksjonene som jeg la til i metodedelen (tabell 4), ser vi at under gjennomførelsen på papir måtte jeg luke bort fire besvarelser på grunn av at jeg la merke til at disse respondentene kopierte ofte andre respondenter sine svar under gjennomførelsen. Dette kom som en konsekvens av at det oppstod dødtid under gjennomførelsen da man hele tiden måtte vente på at sistemann skulle bli ferdig med hver oppgave. Lærerens, eller min, rolle ble dermed veldig påvirket av arbeidsbetingelser som jeg opplevde som overveldende og vanskelig å håndtere da jeg måtte gi instruksjoner for hver oppgave samtidig som jeg måtte følge med om respondentene kopierte andres svar, om respondentene fulgte med og om det var noen som trengte hjelp med å forstå oppgaven. Da vi brukte et litt uegnet rom med mange forstyrrelser av andre elever på skolen utenfor, måtte jeg også jobbe for at respondentene skulle ha fokus på å få gjennomført oppgavesettet. Jeg observerte fort at blikkene til respondentene vandret rundt og de trengte ofte å få samme beskjed flere ganger på grunn av at fokuset deres forsvant. De ble også fort slitne og noen av dem sa ganske fort at de var sultne eller tørste. Med andre ord ble rollen min under gjennomførelsen på papir oppfattet som ganske kaotisk og omfattende da det var mange flere faktorer å ta i betraktning enn å bare gjennomføre et oppgavesett og respondentene ble også fort slitne og lei.

Disse faktorene som kunne skape uro og validitetsutfordringer var ikke til stede når de samme respondentene løste samme oppgaver, bare på data. Vi brukte også samme uegnet rom, men forstyrrelsene rundt ble bare ignorert. Respondentene fikk instruksjoner, gikk til hver sin ferdigoppsatte chromebook, tok på seg headset og satte i gang med oppgavesettet uten noen forstyrrelser eller kopiering av andre sine svar. Respondentene jobbet med oppgavesettet i hvert sitt tempo, ettersom man selv kunne velge når man gikk videre med oppgavene. De ble også ferdig i ulikt tempo uten at det hadde noe å si for de andres gjennomførelse. Min rolle under denne gjennomførelsen var med andre ord preget av mye bedre arbeidsbetingelser og jeg hadde muligheten til å følge opp respondentene på en helt annen måte. Da respondentene sa at de var tørste eller sultne under gjennomførelsen på papir, var det ingen som sa dette når de gjorde oppgavene på data. Istedenfor var det noen som kom med utsagn som «det her er gøy!» under gjennomføringen av oppgavesettet på data eller reagerte med smil da de så oppstilte chromebooks. De aller fleste virket også ivrig på å komme i gang med oppgavene. Dette tolker jeg som at de hadde en god opplevelse under oppgavesettet på data. Gjennomførelsen på data gikk dermed mye fortere enn gjennomførelsen på papir, og de fleste respondentene trodde faktisk at de gjorde mange flere oppgaver på papir enn på data som en konsekvens av dette. Dette kan jeg vise til gjennom feltnotatene jeg gjorde underveis hvor jeg spurte og noterte ned om respondentene følte de gjorde flere oppgaver på data eller på papir som tidligere nevnt (se metodedelen, 3.3 s. 14), hvor 29 av respondentene mente de gjorde flere oppgaver på papir.

Tabell 4 i metodedelen hvor refleksjonene gjort underveis i datainnsamlingen er med på å underbygge disse påstandene.

4.4.2 Oppgaveforklaringens betydning

Et eksempel på et funn som knytter seg til oppgaveforklaringens betydning handler om hvordan man forklarer en oppgave som knytter seg mer til et validitets-perspektiv for forskningsarbeidet. Vi ser på oversikten over antall riktige svar i papir- og dataoppgavene at de to oppgavene med færrest riktige svar var oppgave 19 og 20 (se figur 11). I disse oppgavene fikk respondentene opp en mengde i en rød boks og en mengde i en blå boks og 12 alternativer for riktig svar under. I oppgave 19 var alternativene i ordnet rekkefølge med tallene 1-12 mens i oppgave 20 var tallene 1-12 i uordnet rekkefølge. Oppgaveinstruksen for både oppgavene på papir og data var lik: «Hvor mange flere baller er det i den røde (eller blå) boksen?» Det vil si at respondentene skulle sammenligne disse mengdene (4 og 2, 3 og 5) og krysse av for riktig differanse mellom mengdene. I utgangspunktet er jo dette også ganske lave tall og jeg forventet dermed at disse oppgavene skulle være ganske greie å få til, men slik ble det altså ikke. De aller fleste respondentene svarte 4 på oppgave 19 og 5 på oppgave 20. Det kan derfor virke som om det ble for mye abstrakt og utydelig informasjon i oppgaveinstruksen for respondentene slik at de forsto oppgave 19 som «hvor mange baller er det i den røde boksen» og oppgave 20 som «hvor mange baller er det i den blå boksen» og gikk helt bort fra at det er to mengder som skal sammenlignes.

Hvis vi sammenligner disse svarene med to andre oppgaver, oppgave 23 og 24, ser vi at disse to oppgavene er ganske lik oppgave 19 og 20. Her blir også respondentene presentert med to ulike mengder hvor en mengde er i en rød boks og en mengde er i en blå boks og 12 forskjellige alternativer som svar. Her skal også respondentene sammenligne mengdene og finne differansen mellom dem og krysse av for riktig svar, enten i en ordnet eller uordnet rekkefølge. Vi ser også at tallene i disse to oppgavene er mye høyere enn i oppgave 19 og 20 og ettersom det var få riktige svar med de lavere tallene så skulle man forvente at disse oppgavene, med høyere tall, burde gi enda færre riktige svar. Ut fra framstillingen av resultatene ser vi at oppgave 23 og 24 har fått flere riktige svar, slik at denne antakelsen ikke stemmer. Hovedgrunnen til dette må være den vesentlige forskjellen mellom oppgavene, som ikke handler noe om størrelsen på tallene, men som dreier seg om hvordan man forklarer oppgaven. I denne oppgaven fikk respondentene disse instruksene i begge oppgavesettene: «Det skal være like mange baller i begge boksene. Hvor mange flere baller skal den røde boksen ha?» Denne formuleringen gir et mye bedre bilde av at dette er to mengder som skal sammenlignes og at oppgaven dreier seg om å finne differansen mellom tallene da oppgaveinstruksen tar i bruk bokser med baller og stiller et spesifikt spørsmål om hvor mange flere baller den ene boksen ha for at det skal være like mange som den andre.

Dette viser viktigheten av at dersom man skal måle tallforståelse hos elever på 1. trinn er en presis og informativ oppgaveinstruks veldig viktig for at elevene skal forstå oppgavens rammer og betydning. Upresise oppgaveinstruksjoner har vi nå sett kan påvirke validiteten til et forskningsarbeid hvor man skal måle tallforståelseskompetanse ved at elevene kan misforstå oppgavene, slik at få riktige svar på en oppgave ikke betyr at elevene ikke får til det matematiske i oppgavene, men rett og slett misforstår hensikten med oppgaven.

5 Diskusjon

5.1 Oppsummering av funn opp mot forskningsområde

I denne masteroppgaven har jeg forsøkt å finne svar på følgende forskningsspørsmål:

Hvilke forskjeller i tallforståelseskompetanse finner man dersom elever på 1. trinn løser identiske matematikkoppgaver på data og på papir?

Resultatene som er presentert i forrige kapittel kan oppsummeres med at det helhetlige bildet viser at det ikke er forskjell i respondentenes resultater og hva de får vist fram av kompetanse innenfor akkurat disse oppgavene dersom de løser identiske matematikkoppgaver på data og på papir. Det kan man si ettersom p-verdien fra repeated measures ANOVA-testen viser at man kan akseptere nullhypotesen om at det ikke er en forskjell mellom de to prøveformene ($p=0,754$). Når jeg studerte det helhetlige bildet nøyerer opp mot hver FONS-kategori fremstilt med et søylediagram med standardfeil fant jeg ut at det heller ikke her var en betydelig forskjell mellom resultatene på papir og på data, som også blir et argument opp mot hovedfunnet av forskningsspørsmålet. Til slutt sammenlignet jeg riktige svar på papir og data opp mot hver av de 26 oppgavene fremstilt med et søylediagram, hvor jeg fant ut at selv om det helhetlige bildet viser at det ikke er forskjell mellom papir og data, finnes det forskjeller for hver enkelt oppgave. Spesielt gjaldt det to oppgaver.

Under arbeidet med analysen opp mot forskningsspørsmålet fant jeg i tillegg ut at respondentene generelt presterer bedre på den siste gjennomførelsen, uansett hvilket oppgavesett de gjennomførte først. I det helhetlige bildet så vi at det stemmer for begge gruppene, uansett om gruppene først gjorde oppgavesettet på data og så på papir eller motsatt. Jeg har også sett at dette mønsteret gjelder også hver FONS-kategori, hvor hver kategori har større andel riktige svar under siste gjennomførelse. Spesielt estimeringskategorien viste en større andel riktige svar under den siste gjennomførelsen. Deler man opp FONS-kategoriene videre til hver enkelt oppgave har vi sett at av de 26 oppgavene var det hele 19 oppgaver som flere respondenter fikk til under den siste gjennomførelsen sammenlignet med den første.

Det er også blitt gjort andre funn under forskningsarbeidet, hvor et av funnene er relevant for forskningsspørsmålet, men som går mer på organiseringen av slike prøver enn selve resultatene. Dette funnet handlet om de forskjellige rollene til både meg og respondentene under selve datainnsamlingen som opplevdes veldig forskjellig når vi gjennomførte de to oppgavesettene. Blant annet opplevde jeg mye bedre arbeidsbetingelser opp mot uro og validitetsaspekter under gjennomførelsen av oppgavesett på data kontra når vi gjennomførte samme oppgavesett, bare på papir. Det siste funnet handler om viktigheten av hvordan man forklarer en matematikkoppgave til barn for å unngå at barn misforstår oppgaven, og som dermed knytter seg til validitetsaspekter rundt en slik forskning.

5.2 Mine funn opp mot tidligere forskning

Det er gjort flere forskningsarbeid hvor man har sammenlignet om oppgaver på papir og data har en ekvivalens. Noyes og Garland (Noyes & Garland, 2008) skrev i 2008 en artikkel som tok for seg tidligere forskning på området, hvor de har laget en oversikt over forskjellige studier, med få og flere deltakere, mellom 1981 og 2006 som alle handlet om å sammenligne oppgaver gjort på papir eller på data på en eller annen måte. Disse studiene sammenligner resultater fra oppgaver på papir og data innenfor helt andre områder enn matematikk for eksempel lesehastighet og tester innenfor personlighet og psykologi. Man kan se, ut fra oppsummeringen av hovedfunnene fra de ulike studiene som står i tabellene, at noen tester viser bedre resultater i favør enten papir eller data, men veldig mange av studiene viser at det ikke er funnet noen signifikant forskjell mellom de to formene for gjennomføring. Dette støtter seg opp mot resultatene fra min forskning hvor det i det store og hele ikke var noen forskjell mellom formene og viser at dette ikke gjelder kun for matematikk.

En av forskningsartiklene som Noyes og Garland (Noyes & Garland, 2008) refererer til er Hargreaves (Hargreaves et al., 2004) som har gjennomført en studie hvor hun stiller spørsmålet til om det har noen betydning for barns opptreden i matematikk om oppgaver gjennomføres på papir eller på data, og som derfor kan i stor grad sammenlignes med min studie. I denne studien finner man veldig mange elementer som er likt min studie, hvor det for eksempel er fokus på at måling i nasjonale prøver gjøres på papir og om det har noe å si på resultatene om disse nasjonale prøvene heller gjøres på data. I denne studien er det også brukt identiske oppgavesett på papir og data som bakgrunn på vurdering av resultatene med forskjellige typer oppgaver, blant annet oppgaver om brøk, prosent og statistikk. Her er det i motsetning til min studie brukt to forskjellige oppgavesett hvor den ene gruppen gjennomførte først et oppgavesett på papir og senere identisk oppgavesett på data, mens den andre gruppen gjennomførte først et annet oppgavesett på papir og senere identisk oppgavesett på data. Resultatene fra denne studien viser et generelt bilde hvor det er litt bedre score i oppgavesettene på data, men som samtidig ikke var statistisk signifikant i alle tilfellene. De fant også ut at det var noen oppgaver på data som ga begrensninger på måten et spørsmål kunne bli svart på, som man kan lære av til videre tester. Totalinntrykket til forfatterne var at de fant flere positive sider ved å gjennomføre oppgavesettet på data. Dette er sammenlignbart med mine funn, hvor vi har sett at det ikke er noen forskjell på data og papir i det store og hele og jeg har også trukket fram positive sider ved å gjennomføre vurderinger i matematikk på data. Disse sammenlignbare funnene er positive aspekter knyttet til at for eksempel kartleggingsprøver gjøres på data i fremtiden (Heie, 2021), som viser at prøvene, uansett om det er papir eller data, måler det samme, og at det vil være trygt å gå fra papir til data. I tillegg det faktum at respondentene ga uttrykk for en bedre opplevelse under oppgavesettet på data i denne studien.

5.3 Betydning av funn

5.3.1 Betydning av funn papir vs. data

I innledningen av masteroppgaven diskuterte jeg om det hadde noe å si for tallforståelseskompetansen til 1. klasseelever at de gjennomførte kartleggingsprøver i regning på papir når noen så vidt har lært seg å holde i en blyant. Med denne inngangen til forskningsprosjektet hadde jeg indirekte en hypotese om at resultatene på oppgavesettene var bedre på data enn på papir på grunn av dette aspektet, men vi har

nå sett at dette ikke stemmer. Det betyr at om akkurat disse elevene får kartlagt ferdigheter i regning, så vil de mest sannsynlig vise fram samme kompetanse uansett om de gjennomfører en kartleggingsprøve på data eller på papir. Men det trenger ikke gjelde alle, for det kan jo hende at akkurat disse elevene i dette forskningsprosjektet har mye erfaring enten på chromebook eller med skriving, som kanskje ikke alle andre har. Funnet om at det ikke er forskjell i det helhetlige bildet mellom papir og data kan man knytte opp mot den økte trenden der matematikkoppgaver og tester gjøres på data og at kartleggingsprøver vil bli gjort på data i framtiden (Heie, 2021). Med dette forskningsprosjektet i bakhånd kan man derfor si at det har ikke mye å si for elevenes del at kartleggingsprøvene gjøres på data i framtiden ettersom de i det store og hele presterer like bra uansett. For lærere sin del vil det derimot ha gunstige konsekvenser blant annet på grunn av at man sparer mye tid på å ha kartleggingsprøvene på data som Heie og Stacey og Wiliam også poengterer (Heie, 2021; Stacey & Wiliam, 2012). Dermed kan man også si at selv om det ikke har direkte konsekvenser der og da for elevene om kartleggingsprøvene gjøres på den ene eller andre måten vil det gi gevinst på sikt ettersom lærere sparer mye tid og kan bruke denne tiden på å reflektere over elevenes utvikling istedenfor å rette prøver.

Selv om svaret på forskningsspørsmålet i dette tilfellet er at i det store og hele var det ikke forskjell i resultatene eller kompetansen som respondentene viser fram akkurat der og da mener jeg likevel at, etter å ha prøvd ut begge formene for oppgavesett, det å bruke oppgavesett på data var fordelaktig av særlig to grunner. Den ene grunnen handler om aspekter knyttet til selve gjennomførelsen, som ble nevnt i avsnittet over og i resultatkapittelet (se kapittel 4.4.1, s. 31-32). I gjennomførelsen av oppgavesettet på papir måtte jeg som tidligere nevnt luke bort fire besvarelser på grunn av at disse besvarelsene var trusler mot validiteten til studien. Disse truslene kom bare til syne under gjennomførelsen av oppgavesettet på papir ettersom man måtte vente på at alle sammen skulle være ferdig med hver oppgave før man gikk videre, som man slapp under gjennomførelsen på data. For respondentenes del var det også knyttet større entusiasme og begeistring rundt gjennomførelsen av oppgavesettet på data, slik at både for deres og min del ble det knyttet flere positive erfaringer og holdninger rundt gjennomførelsen på data kontra på papir.

Den andre grunnen støtter seg til de teknologiske potensialene til Stacey og Wiliam (Stacey & Wiliam, 2012), som ble veldig synlige i arbeidet med å samle sammen data og analysere dataen etterpå da jeg fikk erfaringer både fra gjennomførelse av oppgavesett på papir og på data. Jeg har tidligere trukket fram at jeg kunne være fleksibel med når elever kunne gjennomføre de samme oppgavene i oppgavesettet på data, som førte til at elevene var stort sett på forskjellige oppgaver hele tiden og ikke fikk til å se på andre sine svar, som var en mulighet under gjennomførelsen på papir. Ved rettingen av oppgavesettene brukte jeg også mye lengre tid på å rette oppgavesettene på papir. Da det i tillegg var en mulighet for menneskelig feil under denne rettingen, måtte jeg også dobbeltsjekke alle svarene en gang til. Alt i alt tok dette flere timer. Rettingen av oppgavesettene på data ble gjort automatisk, og jeg fikk dem tilsendt av NTNU etter noen minutter uten å bruke noe tid på selve rettingen. Istedenfor å bruke tid på å rette oppgavesett på papir får man verdifull tid til å bruke til for eksempel å skape gode relasjoner til elevene, analysere svarene grundigere for å skape seg et bilde av misoppfatninger, bruke tid på å skape et godt skole-hjem-samarbeid eller helt andre ting som ikke bestandig er like lett å gjøre i en travel lærerhverdag.

5.3.2 Betydning av funn første vs. siste gjennomførelse

Funnet som handler om at resultatene var bedre under siste gjennomførelse uansett om respondentene gjorde oppgavesettet på papir eller data først kan bety at det har skjedd en eller annen form for læring. Det kan være flere årsaker til denne forbedringen som skjedde til tross for at det ikke ble gitt noen form for tilbakemeldinger etter den første gjennomførelsen. Det kan hende at elevene har jobbet med tallforståelse på en eller annen måte mellom gjennomføringene, som påvirket resultatene på den siste gjennomførelsen. Det kan også hende at elevene pratet litt seg imellom om oppgavene og hva de svarte, som kan gjøre at resultatene ble bedre. Med andre ord kan det være faktorer utenfor min kontroll som hadde betydning for at resultatene ble bedre under siste gjennomføring. Det ville derfor vært fordelaktig i en senere gjennomføring og hatt en kontrollgruppe for å kontrollere denne effekten, som kan si oss mer om hvorfor resultatene ble bedre siste gjennomførelse.

Man kan knytte denne erfaringen gjort på læring opp mot hvordan man kan bruke oppgavesett som formative vurdering til å gi økt læring. Forskningslitteraturen har delte meninger når det gjelder læringseffekten av formative vurderinger, det vil si om det er bevist om formative vurderinger gir bedre resultater på tester eller eksamener. Studier av for eksempel Kluger & DeNisi og Kingston & Nash (Kingston & Nash, 2011; Kluger & DeNisi, 1996) fant ingen tydelig effekt mellom formative vurderinger og økt læring, mens resultatene fra en lik studie av Black og Wiliam (Black & Wiliam, 1998) viser det motsatte, at formative vurderinger gir økt læring. Selv om mitt forskningsprosjekt ikke er direkte koblet mot formativ vurdering kan man bruke erfaringene og resultatene opp mot den formative vurderingen gjennom kartleggingsprøver som viser at dersom man gjennomfører et oppgavesett kan man få bedre resultater og økt læring den andre gangen man gjennomfører et tilsvarende oppgavesett. Man kan med andre ord bruke kartleggingsprøver til å gjennomføre oppgavene, prate om oppgavene i etterkant sammen med elevene, hvor sannsynligheten da blir større for at det har skjedd en læring til neste gjennomførelse. Til slutt kan denne utviklingen sees mot den endelige summative vurderingen når elevene en dag går ut av skolen, hvor kanskje kartleggingsprøvene har bidratt til å tilrettelegge undervisningen og gjort at den summative vurderingen ble bedre.

Funnet fra dette forskningsarbeidet som peker mot at elever presterer bedre andre gang de gjennomfører et oppgavesett relaterer seg også til at lærere kobler læring og aktiviteter i klasserommet direkte opp mot hvordan testene for vurdering er gjennomført (Popham, 2001). For eksempel kan lærere i Norge legge opp undervisningen direkte opp mot kartleggingsprøver og nasjonale prøver for å prøve å få så bra resultat som mulig. Dette skjer fordi lærere og skoler gjerne blir målt etter hvor bra elevene gjør det på slike nasjonale standardiserte prøver, og presset på lærere og administrasjonen er stort til å få bedre resultater (Posner, 2004). I California har for eksempel dette ført til at skoler har blitt stengt (Posner, 2004) ettersom de tidligere har hatt en politikk hvor skoler får mer økonomiske midler ut fra hvor stor forbedring på disse testene skolene har hatt. Resultatene fra dette forskningsprosjektet indikerer til at kjennskap til de konkrete oppgavene kan heve prestasjonen selv uten undervisning og tilbakemelding, og kan på noen måter rettes til undervisning direkte opp mot for eksempel kartleggings- eller nasjonale prøver. Dette skjedde til tross for at designet for forskningsprosjektet, ved for

eksempel motvektsprinsippet, var laget for å kontrollere at tidspunkt og rekkefølge ikke hadde noen innvirkning på resultatene, men så viste det seg at det hadde litt innvirkning likevel.

Det er delte meninger om hvor bra det er å ha en slik direkte kobling i undervisningen opp mot kartleggingsprøver. Noen forskningsartikler argumenterer for at det er akkurat slik det skal være, hvor man trekker fram at poenget med nasjonale standardiserte tester er for å sjekke hvor man ligger an med tanke på målene for læreplanen, og hvis man øver seg på disse i klasserommene så er jo det akkurat hva man ønsker at lærerne skal gjøre (Posner, 2004). I Norge mener derimot blant annet Utdanningsforbundet at denne type press og fokus på å prestere bra på nasjonale prøver er meningsløs og viser fram eksempler på at resultatene på nasjonale prøver har blitt dårligere når det har kommet prøvesett til skolene (Rambøl, 2015). Det trekkes frem at grunnen til dette er at presset på elevene og lærerne blir for stort slik at man rett og slett blir for stresset av det. I tillegg er intensjonen til kartleggingsprøver om å kartlegge om noen elever ligger under et kritisk nivå, og hvis man da øver til disse prøvene vil ikke kartleggingsprøvene være det verktøyet som det er utviklet til å være (Rambøl, 2015).

5.3.3 Betydning av funn fra gjennomførelsen

De andre funnene i forskningsprosjektet er som tidligere nevnt viktigheten av hvordan matematikkoppgaver blir forklart til barn i 1. klasse for at de ikke skal bli misforstått og erfaringene på hvordan gjennomføringen av de to oppgavesettene var. Disse funnene har betydning av ulike årsaker. Funnet som handler om viktigheten av hvordan matematikkoppgaver blir forklart til barn i 1. klasse har betydning opp mot validiteten til slike typer studier hvor man skal vurdere regnekompetanse til barn. Det er med andre ord veldig viktig at for at man kan vurdere regnekompetanse til barn på 1. trinn må oppgaveformuleringene være så presise og gjennomtenkte som mulig slik at barn ikke svarer feil på oppgavene på grunn av upresis formulering kontra manglende regnekompetanse. Dersom feilene oppstår på grunn av upresis formulering kan dette bety at barn faktisk forstår det matematiske i oppgaven, men får feil svar på grunn av upresise formuleringer. Dermed kan resultatene bli av feil grunnlag. Funnet som går på selve gjennomføringen av de to oppgavesettene med respondentene har betydning opp mot at for eksempel kartleggingsprøver på barneskolen vil bli gjort på data i fremtiden (Heie, 2021). Min erfaring med dette forskningsprosjektet er at det er flere fordeler med slike prøver på data som for eksempel mindre uro, mer oversikt over alle deltakerne og generelt positive holdninger til gjennomførelsen blant elevene. I tillegg vil slike kartleggingsprøver på data sikre elementer som har betydning for validiteten til prøven og at lærerne vil spare en god del tid med å rette oppgavene.

5.4 Studiens begrensninger

Jeg skal være forsiktig med å generalisere funnene i dette forskningsprosjektet da denne studien har noen begrensninger. Et viktig aspekt knyttet til dette er at selv om jeg har argumentert for metodikken med at dette er en kvantitativ studie er det likevel ikke i nok til å være et argument som man ukritisk kan rette mot andre skoler i ulike miljøer. Det krever mer forskning for å kunne argumentere at dette gjelder generelt for alle 1. klasseelever, blant annet for elever på andre skoler som ikke har like stor erfaring og tilgang med chromebooks.

Faktorer som kan ha noe å si for resultatene av studien er for eksempel skolens tilgang til digitale ressurser for å gjennomføre oppgavesettet på data og hvor godt vant respondentene er med i bruken av disse ressursene. Respondentene som er med i dette forskningsprosjektet har brukt chromebooks mye i undervisningen fra før, men det er ikke sikkert alle har like stor tilgang og erfaring med disse. En annen faktor handler om språklig forståelse. I og med at hver av oppgavene blir forklart muntlig av meg eller en innlest stemme må man jo faktisk forstå hva som blir sagt, som ikke nødvendigvis er en selvfølge i for eksempel miljøer med mange flerspråklige elever eller dersom den innleste stemmen er en annen dialekt enn det man selv prater. En siste faktor handler om hvordan lærerne har lagt opp undervisningen på forhånd av gjennomførelsen og hvor langt man har kommet i matematikkundervisningen. I mitt tilfelle har lærerne på trinnet hatt veldig fokus på at elevene først skal forstå matematiske begreper før de begynner med tallinnlæring. Det kunne man se når respondentene jobbet med tallidentifikasjonsoppgavene hvor mange kunne fortelle at de ikke hadde jobbet med tall over 10 før. Respondentene løste med andre ord matematikkoppgaver om noe de ikke hadde lært noe om på skolen fra før. Dermed kan man forvente at respondentene i dette forskningsprosjektet har lavere score på tallidentifikasjonsoppgavene enn elever på andre skoler hvor de har startet tidligere med tallinnlæring.

I tillegg så vi at under gjennomførelsen av oppgavesettet på papir måtte jeg luke ut fire besvarelser som jeg anså var en trussel til validiteten av studien. Jeg kan ikke med 100% sikkerhet si at jeg fikk luket bort alle sammen ettersom det var ganske krevende å holde oversikten over alle respondentene, da det kan hende at noen fikk til å kopiere andres svar uten at jeg fikk det med meg. Dette er en begrensing til studien, og det ville vært hensiktsmessig og hatt en ekstra person med i gjennomførelsen som ville hatt som hovedoppgave å notere ned aspekter underveis i gjennomførelsene slik at det hadde vært enklere å luke bort besvarelser.

I denne forskningen er tallforståelsen til respondentene målt gjennom kun fem av åtte komponenter av det teoretiske rammeverket FONS. Da FONS er en helhetlig vurdering av tallforståelse kan man med andre ord si at jeg mangler noen komponenter og har noen hull i målingen av tallforståelseskompetansen til respondentene. De tre komponentene som ikke er med er mengdediskriminering, forståelse av forskjellige representasjoner for et tall og bevissthet om tallmønstre. Likevel ser man at noen av oppgavene i oppgavesettene dekker disse tre komponentene, som viser at selv om FONS er en helhetlig vurdering av tallforståelseskompetanse basert på ulike komponenter så ser man at komponentene går hånd i hånd og av og til overlapper hverandre. Selv om oppgavesettene mangler tre FONS-komponenter kan jeg ikke se at det har vesentlige konsekvenser for den helhetlige vurderingen av tallforståelse til respondentene. Et eksempel på dette er systematisk telling-oppgavene hvor respondentene skal telle en mer eller en mindre enn et gitt tall. Dette betyr at man skal telle, men samtidig viser det også at man har en bevissthet til størrelser og mengder og kan forskjellen mellom en mindre eller en mer enn som mengdediskriminering handler om. Et annet eksempel er forhold mellom tall og mengde-oppgavene hvor respondentene skal telle antall baller for å vise en korrespondanse mellom tallets navn og antall enheter tallet representerer. Det betyr også at man har en viss forståelse av forskjellige representasjoner av et tall, hvor for eksempelet tallet 7 tilsvarer sju baller og at dette betyr det samme.

5.5 Videre forskning

Som en konsekvens av mine funn dukker det opp aspekter det ville vært interessant å se på videre. Videre forskning kan med dette være knyttet til denne studiens begrensinger for å se om mine funn også kan relateres til andre skoler i ulike miljøer. Dersom man ser at dette stemmer også hos andre, vil man bli mer sikker på at man kan konkludere mer generelt. Hovedsakelig om forskningsspørsmålet, om det er forskjell på tallforståelseskompetanse om man gjennomfører oppgavesett på data og papir, men også om den siste gjennomførelsen gir bedre resultat enn den første gjennomførelsen.

Dette krever ikke nødvendigvis at man bruker identisk metode som det er brukt i dette forskningsprosjektet, men det kommer an på hvilket spørsmål man stiller seg til hva man skal forske på. Skal man teste tallforståelsen til spesifikt skolestartere kan det være hensiktsmessig å bruke samme type metode, oppgavesett og like rammer rundt gjennomførelsen slik at man kan sammenligne resultatene opp mot mine funn. Dersom man ønsker et mer generelt svar, for eksempel forske på om man får samme resultat i oppgavesett på papir eller på data uansett hvor gamle elevene er, så trenger man ikke følge oppskriften som er brukt i dette forskningsarbeidet.

Det vil derfor være veldig interessant å forske videre på eldre elever for å se om de også får det samme resultatet på papir og på data, men det kan også være interessant å mikse litt mellom metodikken og teorien for å utvide konklusjonshorisonten. Man kan for eksempel bruke samme metodikk, men helt andre matematikkoppgaver. Kanskje viser det seg at i noen typer matematikkoppgaver presterer elever mye bedre på den ene formen av oppgavesettene. For eksempel finnes det digitale matematikkoppgaver fra ulike situasjonskontekster som løses med lek og spill, og det ville vært interessant å sammenligne disse type oppgavene opp mot samme type oppgaver, bare på papir. Denne type oppgaver på data hvor man bruker spill til innlæring av matematikk har gitt positiv signifikante resultater basert på tidligere forskning, for eksempel fra en større forskning av Kebritchi (Kebritchi et al., 2010). Eller så kan man bruke identiske oppgaver, men helt annen metodikk, for eksempel bruke større grupper per gjennomførelse, at den som gjennomfører har en annen rolle enn deltakende observatør eller at man bruker sekundærdata som datamateriale. På den måten får man enda større grunnlag til å konkludere om mine funn gjelder generelt for elever i samme alder, og ikke på grunnlag av den spesifikke metodikken og teorien som er grunnlaget for denne studien.

Denne studien gjør også at jeg lurer på hvordan resultatene hadde vært om man testet alle komponentene av FONS i senere skolealder, for eksempel starten av 2. klasse for å se om resultatene er knyttet opp mot at respondentene i dette forskningsprosjektet er helt i startfasen av skolegangen sin. Men heller ikke for langt ut i skolegangen da FONS hovedsakelig baserer seg på tallforståelse hos barn i de aller tidligste skoleårene. Med andre ord er det utallige muligheter for hvordan man kan teste tallforståelse til barn på papir og på data på flere måter enn det jeg har gjort i denne studien.

Referanser

- Almond, R. G., Steinberg, L., & Mislevy, R. J. (2003). *A four-process architecture for assessment delivery, with connections to assessment design*. University of California Los Angeles Center for Research on Evaluations, Standards and Student Testing (CRESST).
- Andrews, P., & Sayers, J. (2015). Identifying Opportunities for Grade One Children to Acquire Foundational Number Sense: Developing a Framework for Cross Cultural Classroom Analyses. *Early Childhood Education Journal*, 43(4), 257–267. <https://doi.org/10.1007/s10643-014-0653-6>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Berch, D. B. (2005). Making Sense of Number Sense: Implications for Children With Mathematical Disabilities. *Journal of Learning Disabilities*, 38(4), 333–339. <https://doi.org/10.1177/00222194050380040901>
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Christoffersen, L., & Johannessen, A. (2012). *Forskningsmetode for lærerutdanningene*. Abstrakt Forlag.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8. utg.). Routledge.
- Davasligil, Ü. (2004). Early Prediction of High Mathematical Ability. *Gifted and Talented International*, 19(2), 76–85. <https://doi.org/10.1080/15332276.2004.11673040>
- DES/WO. (1988). *National Curriculum Task Group on Assessment and Testing—A report*. DES.
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th edition). SAGE Publications.
- Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early Identification and Interventions for Students With Mathematics Difficulties. *Journal of Learning Disabilities*, 38(4), 293–304. <https://doi.org/10.1177/00222194050380040301>
- Hargreaves, M., Shorrocks-Taylor, D., Swinnerton, B., Tait, K., & Threlfall, J. (2004). Computer or paper? That is the question: does the medium in which assessment questions are presented affect children's performance in mathematics? *Educational Research*, 46(1), 29–42. <https://doi.org/10.1080/0013188042000178809>
- Heie, M. (2021). Slik skal forskerne kartlegge de minste barnas regneferdigheter. *UiO Institutt for lærerutdanning og skoleforskning*. <https://www.uv.uio.no/ils/om/aktuelt/aktuelle-saker/2021/digitaleproverfordeminste.html>
- Hekneby, G. (2005). *Elevens håndskrift skriftforming i grunnskolen*. Universitetsforl.
- Helle, L., & Burder, T. (2021). *Formativ vurdering*. https://snl.no/formativ_vurdering
- Hox, J. J., & Boeije, H. R. (2005). Data Collection, Primary vs. Secondary. *Encyclopedia of Social Measurement*, 1.
- Huitt, W., & Hummel, J. (2003). Piaget's Theory of Cognitive Development. I *Educational Psychology Interactive*. Valdosta State University. <http://chiron.valdosta.edu/whuitt/col/cogsys/piaget.html>
- Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education*, 55(2), 427–443. <https://doi.org/10.1016/j.compedu.2010.02.007>

- Kilpatrick, J., Swafford, J., Findell, B., National Research Council (U.S.), & Mathematics Learning Study Committee. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press. <https://openlibrary.org/books/OL17062503M>
- Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28–37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Lyngsnes, K., & Rismark, M. (2007). *Didaktisk arbeid* (2. utg.). Gyldendal Akademisk.
- Mcintosh, A., Reys, B. J., & Reys, R. E. (1992). A Proposed Framework for Examining Basic Number Sense. *For the Learning of Mathematics*, 12(3), 2–44.
- NESH. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. (4. utg.). De nasjonale forskningsetiske komiteene.
- Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, 51(9), 1352–1375. <https://doi.org/10.1080/00140130802170387>
- Ojose, B. (2008). Applying Piaget's Theory of Cognitive Development to Mathematics Instruction. *The Mathematics Educator*, 18(1), 26–30.
- Piaget, J. (2001). *The psychology of intelligence*. Routledge.
- Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16–20.
- Posner, D. (2004). What's Wrong with Teaching to the Test? *Phi Delta Kappan*, 85(10), 749–751. <https://doi.org/10.1177/003172170408501009>
- Rambøl, A. H. (2015, september 29). Øver på kartleggingsprøver. *Klassekampen*. <https://arkiv.klassekampen.no/article/20150929/PLUSS/150929792>
- Renolen, Å. (2015). *Forståelse av mennesker innføring i psykologi for helsefag*. Fagbokforl.
- Robinson, N. M., Abbott, R. D., Berninger, V. W., Busse, J., & Mukhopadhyay, S. (1997). Developmental Changes in Mathematically Precocious Young Children: Longitudinal and Gender Effects. *Gifted Child Quarterly*, 41(4), 145–158. <https://doi.org/10.1177/001698629704100404>
- Scriven, M. S. (1967). The methodology of evaluation. In R.W. Tyler, R.M Gagne, & M.S. Scriven (Eds.). *Perspectives of curriculum evaluation*, 39–83.
- Skog, O.-J. (2005). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming*. Gyldendal norsk forlag.
- Stacey, K., & Wiliam, D. (2012). Technology and Assessment in Mathematics. I M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Red.), *Third International Handbook of Mathematics Education* (s. 721–751). Springer New York. https://doi.org/10.1007/978-1-4614-4684-2_23
- St.meld. nr. 030 2003-2004. *Kultur for læring*. Regjeringen. <https://www.regjeringen.no/no/dokumenter/stmeld-nr-030-2003-2004-/id404433/?ch=4>
- Tjora, A. (2013). *Kvalitative Forskningsmetoder i Praksis* (2. utg.). Gyldendal Akademisk.
- Utdanningsdirektoratet. (2020a). 2.3 *Grunnleggende ferdigheter*. <https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/grunnleggende-ferdigheter/?lang=nob>
- Utdanningsdirektoratet. (2020b). *Grunnleggende ferdigheter—Læreplan i matematikk 1.–10. Trinn (MAT01-05)*. <https://www.udir.no/lk20/mat01-05/om-faget/grunnleggende-ferdigheter>

Utdanningsdirektoratet. (2020c). *Kartleggingsprøver*. <https://www.udir.no/eksamen-og-prover/prover/kartlegging-gs/>

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's Past Is Prologue: Relations Between Early Mathematics Knowledge and High School Achievement. *Educational Researcher*, *43*(7), 352–360. <https://doi.org/10.3102/0013189X14553660>

Vedlegg

Vedlegg 1: Informasjonsskriv foresatte (4 sider).

Vil du delta i forskningsprosjektet mitt for å finne hvilke forskjeller det finnes når elever løser identiske matematikkoppgaver digitalt og på papir?

Til foresatte for elever på 1. trinn ved Flatåsen Skole

Jeg er masterstudent ved NTNU, Institutt for Lærerutdanning, og skal gjennomføre et kort forskningsprosjekt på skolen til ditt barn. I dette skrivet gir jeg deg informasjon om målene for prosjektet og hva deltakelse vil innebære for ditt barn.

Formål

Formålet med prosjektet er å samle inn datamaterialet til en masteroppgave i matematikdidaktikk. Den nåværende problemstillingen er: «Hvilke forskjeller finner man når elevene løser identiske matematikkoppgaver digitalt og på papir?». Resultatene av studien vil bli brukt i en masteroppgave ved NTNU.

Hvem er ansvarlig for forskningsprosjektet?

NTNU, Institutt for Lærerutdanning er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Ønsket mitt er å bruke så mange som mulige elever på 1. trinn for å få størst mulig dataomfang. Det vil si at dersom dere samtykker til å være en del av forskningsprosjektet vil barnet ditt sine svar på oppgavene som blir gitt bidra til å løse problemstillingen min. Gjennomføringen av prosjektet er avklart med ledelsen på skolen.

Hva innebærer det for deg å delta?

Hvis du velger å delta i prosjektet, innebærer det at elevene først vil løse matematikkoppgaver på chromebook. En ukes tid senere vil elevene løse identiske oppgaver, bare på papir. Elevene vil få utdelt tilfeldige kallenavn slik at jeg får muligheten til å sammenligne de digitale oppgavene og oppgavene på papir. Det vil ikke bli tatt lydopptak eller bilder under gjennomføring og elevene vil få utdelt kallenavn.

Totalt vil det kanskje ta rundt 40-60 minutter per elev som blir tatt ut. Dette vil gjøres i løpet av skoledagen, slik at elevene blir tatt ut av klasserommet i dette tidspunktet. Jeg vil være deltakende observatør under gjennomføringen. Det vil si at jeg kommer til å notere ned aspekter jeg legger merke til underveis i gjennomføringen som kan bli brukt i masteroppgaven. Disse notatene vil ikke inneholde informasjon slik at man kan identifisere enkelte barn og jeg vil referere til kallenavnene. Om dere ønsker kan dere få se oppgavene på forhånd ved å ta kontakt med masterstudent eller veileder.

Det er frivillig å delta

Det er frivillig å delta i prosjektet, både for dere og elevene. Elevene har uansett siste ord, som vil si at dersom dere samtykker men elevene ikke har lyst, vil ikke de gjennomføre prosjektet. Hvis dere velger å delta, kan dere når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg vil da bli anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg og det vil heller ikke påvirke elevens forhold til skolen eller lærere dersom dere velger å ikke delta eller senere trekke deg.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Vi vil bare bruke opplysningene om ditt barn til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Datamaterialet vil kun bli sett av masterstudent Inge Vang Ellefsen og veilederne ved NTNU Gunnhild Saksvik-Raanes og Trygve Solstad. Datamaterialet vil kunne diskuteres med medstudenter, men medstudentene vil ikke få tilgang til personopplysningene. Opplysningene i prosjektet vil bli holdt konfidensielt og i samsvar med personvernregelverket. Deltakerne i prosjektet vil bli anonymisert, og vil ikke kunne gjenkjennes i en eventuell publikasjon. Det vil ikke bli brukt hverken video eller bilder i datainnsamlingen. Det eneste som inneholder personopplysninger er elevens navn på dette samtykkeskjema og et dokument som kobler sammen elevenes navn til hvilket kallenavn de har fått. Dette dokumentet vil krypteres og kun bli brukt av meg og lagret på en minnepenn på NTNU.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal etter planen avsluttes 5. September 2021. Alle personvernopplysninger vil som sagt bli slettet ved prosjektslutt.

Dine rettigheter

Så lenge ditt barn kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om ditt barn basert på ditt samtykke.

På oppdrag fra NTNU, institutt for lærerutdanning har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- NTNU, Institutt for Lærerutdanning ved: Masterstudent Inge Vang Ellefsen (ingeve@stud.ntnu.no) eller veiledere Gunnhild Saksvik-Raanes (gunnhild.b.saksvik@ntnu.no) og Trygve Solstad (trygve.solstad@ntnu.no)
- NSD – Norsk senter for forskningsdata AS, på epost (personverntjenester@nsd.no) eller telefon: 55 58 21 17.
- NTNU sitt personvernombud, Thomas Helgesen, epost (thomas.helgesen@ntnu.no)

Med vennlig hilsen

Gunnhild Saksvik-Raanes

& Trygve Solstad

Inge Vang Ellefsen

Prosjektansvarlig

Student

(Forsker/veileder)

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «*forskjeller som finnes når elevene løser identiske matematikkoppgaver digitalt og på papir*» og har fått anledning til å stille spørsmål. Jeg samtykker til:

_____ (barnets navn) kan delta i prosjektet.

Jeg samtykker til at barnet mitt kan:

- delta i deltakende observasjon
- svare på oppgaver i matematikk

(Signert av foresatt/prosjektdeltaker, dato)

Vedlegg 2: Oppgavesett på papir, lærerversjon m/instrukser for hver oppgave (9 sider).

Matematikkoppgaver

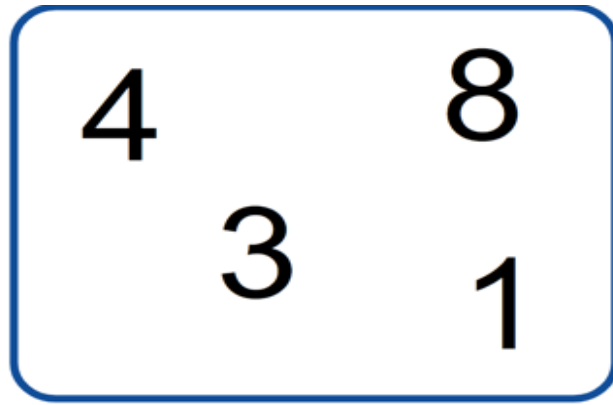


NAVN: _____

1.

TEGN KRYSS PÅ
TALLET TRE

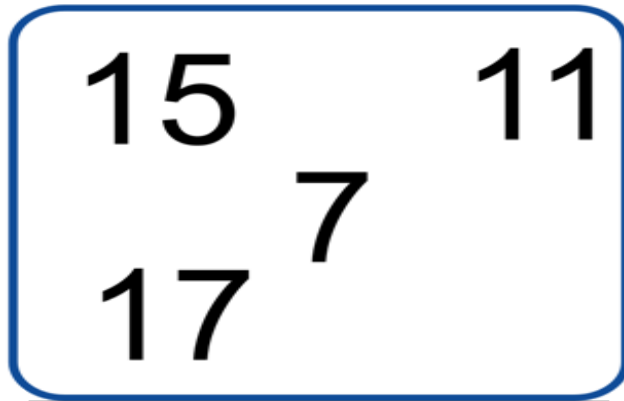
(TI)



2.

TEGN KRYSS
PÅ TALLET
SYTTEN

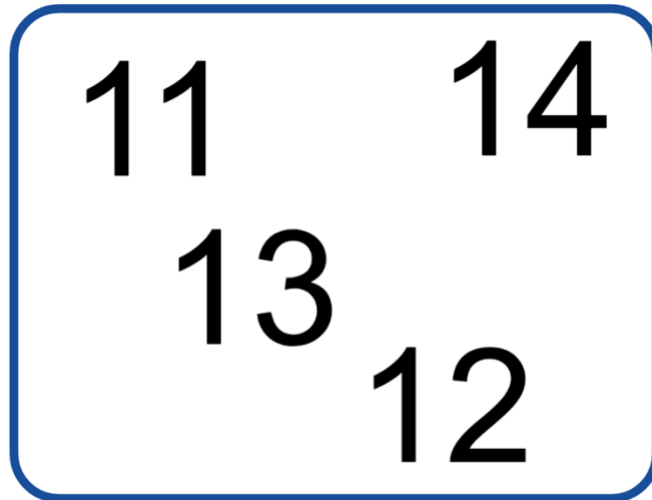
(TI)



3.

TEGN KRYSS
PÅ TALLET
ELLEVE

(TI)



4.

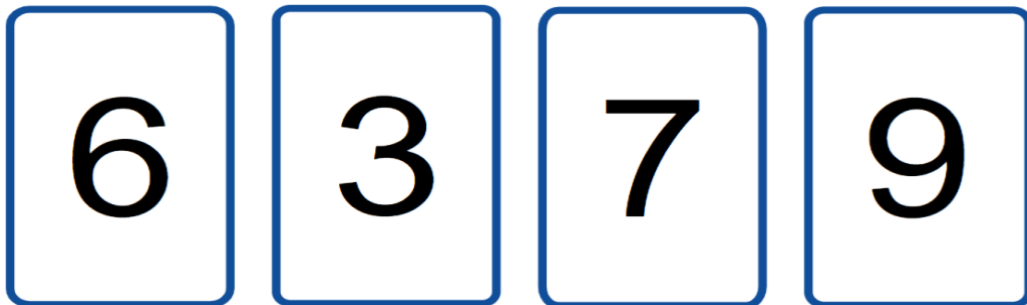
TEGN KRYSS
PÅ TALLET
TRETEN

(TI)



5.

TEGN KRYSS
PÅ TALLET
SEKS
(TI)



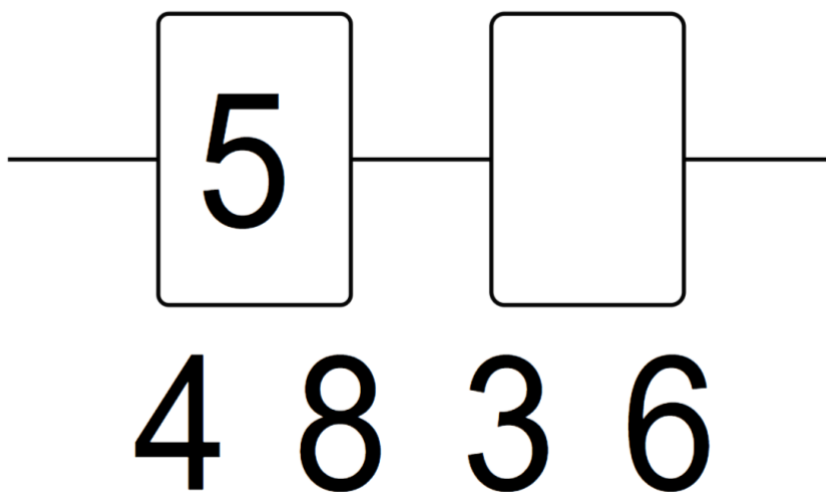
6.

TEGN KRYSS
PÅ TALLET
ÅTTE
(TI)



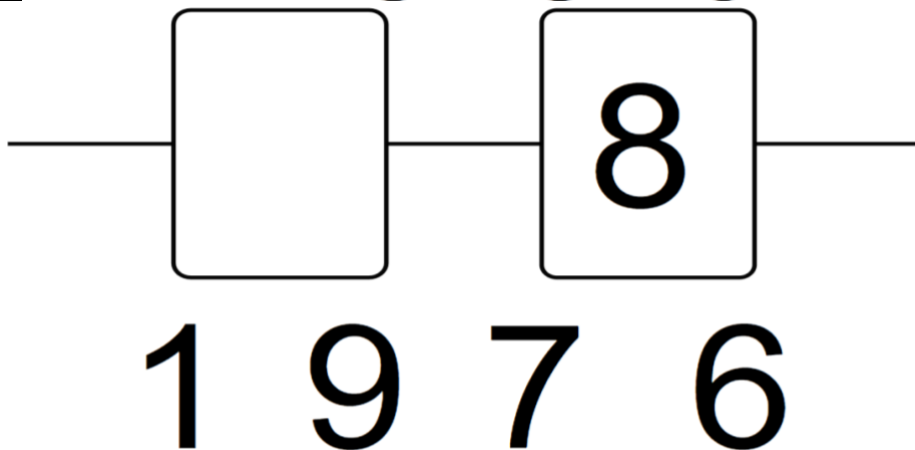
7.

TEGN
STREK FRA
TALLET
SOM
KOMMER
ETTER FEM
INN I
BOKSEN
(ST)



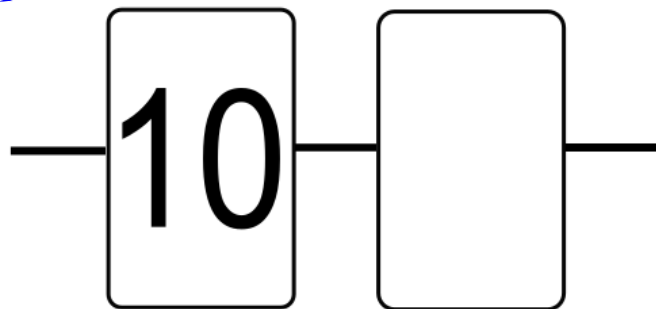
8.

TEGN
STREK FRA
TALLET
SOM
KOMMER
FØR ÅTTE
INN I
BOKSEN
(ST)



9.

TEGN STREK
FRA TALLET
SOM
KOMMER
ETTER TI
INN I
BOKSEN
(ST)



11 9 13 12

10.

TEGN KRYSS
PÅ DEN
ANDRE
STJERNEN I
REKKEN (ST)



11.

TEGN
KRYSS PÅ
DEN
FEMTE
STJERNEN
I REKKEN
(ST)



12.

TEGN KRYSS
DER DU SER
LIKE MANGE
BALLER SOM
TALLET
(T&M)

7



13.

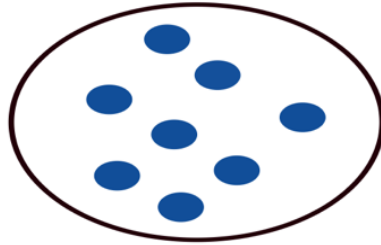
TEGN
KRYSS DER
DU SER
LIKE MANGE
BALLER
SOM TALLET
(T&M)

10



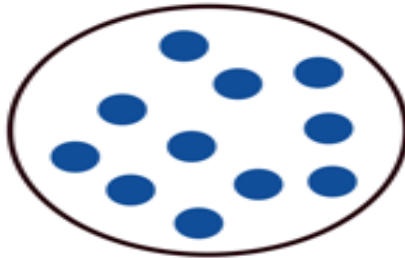
14.

TEGN
KRYSS PÅ
TALLET SOM
VISER HVOR
MANGE
BALLER DET
ER I
RUNDINGEN
(T&M)



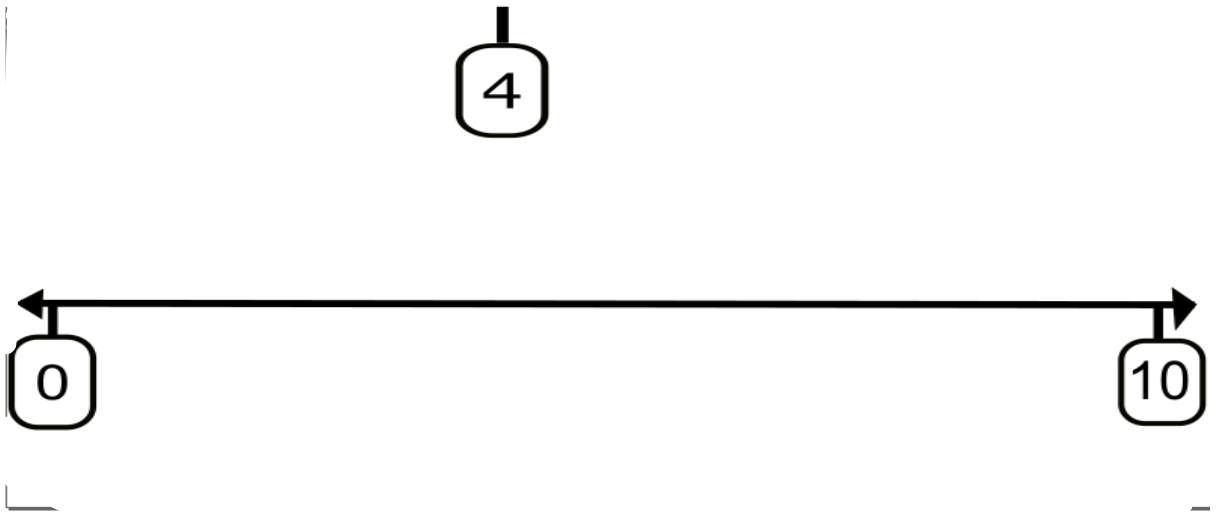
15.

TEGN
KRYSS PÅ
TALLET SOM
VISER HVOR
MANGE
BALLER DET
ER I
RUNDINGEN
(T&M)



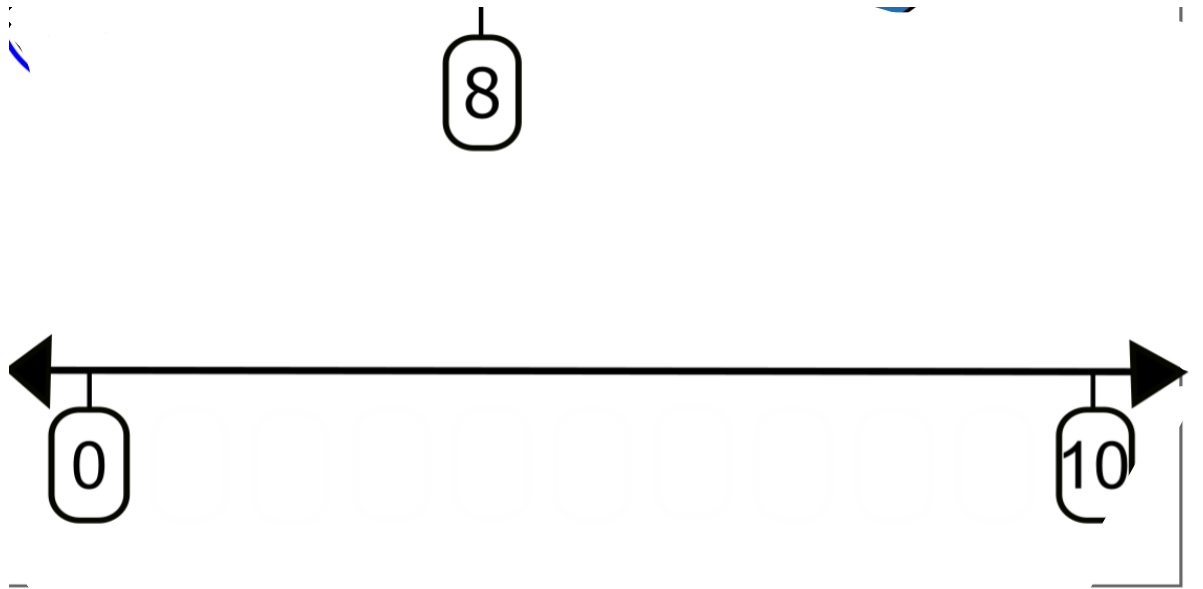
16.

TEGN
STREK FRA
TALLET TIL
RIKTIG
PLASS PÅ
LINJA (E)



17.

TEGN
STREK FRA
TALLET TIL
RIKTIG
PLASS PÅ
LINJA (E)



18.

DU HAR TO
BALLER OG
FÅR FEM
BALLER TIL,
HVOR
MANGE
BALLER HAR
DU DA?
(AK)



19.

HVOR
MANGE
FLERE
BALLER ER
DET I DEN
RØDE
BOKSEN?
(AK)



1 2 3 4 5 6 7 8 9 10 11 12

20.

HVOR
MANGE
FLERE
BALLER ER
DET I DEN
BLÅ
BOKSEN?
(AK)



10 7 8 12 11 6 1 3 9 5 4 2

21.

HVA ER
SEKS OG TO
TIL
SAMMEN?
TEGN
KRYSS PÅ
RIKTIG
SVAR (AK)

$$6+2$$

1 2 3 4 5 6 7 8 9 10 11 12

22.

HVA ER FEM
OG FIRE TIL
SAMMEN?
TEGN KRYSS
PÅ RIKTIG
SVAR (AK)

$$5+4$$

10 7 8 12 11 6 1 3 9 5 4 2

23.

DET SKAL VÆRE
LIKE MANGE
BALLER I BEGGE
BOKSENE. HVOR
MANGE FLERE
BALLER SKAL DEN
RØDE BOKSEN
HA? TEGN KRYSS
(AK)

6

10

10 7 8 12 11 6 1 3 9 5 4 2

24.

DET SKAL VÆRE
LIKE MANGE
BALLER I BEGGE
BOKSENE. HVOR
MANGE FLERE
BALLER SKAL DEN
RØDE BOKSEN
HA? TEGN KRYSS
(AK)

11

15

1 2 3 4 5 6 7 8 9 10 11 12

25.

$$16+2$$

HVA ER
SEKSTEN OG
TO
TILSAMMEN?
TEGN KRYSS
PÅ RIKTIG
SVAR (AK)

14 15 16 17 18 19 20 21 22 23 24 25

26.

$$11+4$$

HVA ER
ELLEVE OG
FIRE
TILSAMMEN?
TEGN KRYSS
PÅ RIKTIG
SVAR (AK)

11 13 7 14 9 16 5 12 6 8 15 10

