UiT Norges arktiske universitet

Fakultetet for humaniora, samfunnsvitenskap og lærerutdanning

# Eksemplifisere for å briljere

## En kvantitativ studie av effekten til eksempler, i problemgenereringssituasjoner

Johansen, Fredrik & Kvikstad, Ole

Masteroppgave i matematikkdidaktikk, LER-3903, mai 2022

# Innholdsfortegnelse

# Tabelliste:



# Figurliste:

# Forord

Denne mastergradsavhandlingen er avslutningen på vårt femårige studium på lærerutdanningen for 5.-10. trinn ved Universitetet i Tromsø – Norges arktiske universitet. Arbeidet med denne mastergraden har gitt oss et større perspektiv på bruken av eksempel i undervisning, og gjennomføring av problemgenerering som en del av utforskende undervisning i matematikk.

Vi ønsker med dette å takke våre veiledere Oskar Jensen Wang og Per Øystein Haavold, for deres innspill på tematikk, gjennomføring og selve avhandlingen. Vi vil også takke elevene og lærerne som har latt oss gjennomføre studien vår, og som har bidratt til interessante funn! Vi vil takke medstudenter for innspill og oppløftende kommentarer under jobbingen. Jeg, Fredrik, ønsker særlig å takke min kone, Marthe Solberg Johansen, for all støtten og tilretteleggingen du har bidratt med det siste året. Jeg, Ole, ønsker særlig å takke min flotte samboer, Hanne Ruistuen, for all støtten du har gitt dette året. Dere er virkelig enestående og har vært gode medspillere under skrivingen av denne avhandlingen! Vi verdsetter dere enormt mye.

Til slutt ønsker vi å takke hverandre for et godt samarbeid, og et fint vennskap! Jobbingen med denne avhandlingen har medført mange gode minner, diskusjoner og mye latter.

Tromsø, mai 2022

Fredrik Solberg Johansen & Ole Martin Kvikstad

## Sammendrag

Utforskende undervisning blir stadig mer aktuelt i skolen, og en finner elementer fra nettopp dette i LK-20. Problemgenerering kan i mange tilfeller være en stor del av den utforskende undervisningen, og i flere tilfeller selve grunnstenen. Studien er gjennomført på to byskoler, med totalt 25 deltakere. For å måle elevene innenfor problemgenerering har vi brukt en kreativitetsskala, da kreativitet i matematikk ofte kan sees i sammenheng med høyt prestasjonsnivå. For å se nærmere på tematikken har vi forsøkt å svare på følgende problemstilling og forskningsspørsmål. 1) *I hvilken grad skårer elever på kategoriene fleksibilitet, flyt og originalitet, når de får presentert et eksempel i en oppgave?* 2) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de får presentert et eksempel?* 3) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de ikke får presentert et eksempel?* For å svare på overnevnte spørsmål har vi brukt kvantitativ metode, nærmere spesifisert eksperiment med kontroll og eksperimentgruppe med post-test. Fra testene kunne man se klare forskjeller på kontroll og eksperimentgruppene, i favør eksperimentgruppen. Både måloppnåelse og bruk av eksempel virker til å ha en positiv effekt.

# 1 Innledning

Mastergradsavhandlingen vår er basert på en eksperimentell, komparativ case studie, hvor vi sammenlikner bruken av eksempel og ikke, innenfor problemgenerering i undersøkende undervisning, når det kommer til elevers kreativitet. I avhandlingen redegjør vi for sammenhengen mellom utforskende undervisning, problemgenerering og kreativitet, hvordan eksperimentet har blitt utformet, analysert og gjennomført, samt drøfter resultatene våre opp mot tidligere forskning og teori.

## 1.1 Bakgrunn for valg av tema

I kjerneelementene i LK20 skrives det om utforsking og problemløsning, resonnering og argumentasjon, modellering og anvendinger og representasjon og kommunikasjon (Utdanningsdirektoratet, 2020). Alle disse begrepene er sentrale i utforskende undervisning. En kan jobbe med, og trene disse ferdighetene hver for seg, eller sammen, i nettopp utforskende undervisning. Denne «nye» måten å arbeide på har som mål å gjøre elevene om til forskere, hvor de selv må ta stilling til en rekke situasjoner, som tema, *problemgenerering*, problemløsning, argumentasjon og det å skape en felles overførbar læring (Utdanningsdirektoratet, 2020). Det er skrevet flere artikler og bøker innenfor flere av temaene, men et noe neglisjert område er problemgenerering.

Problemløsning er ikke noe fremmedord i matematikk i norsk skole, og er av vår erfaring knyttet til en morsom aktivitet én gjør ved siden av den ordinære undervisning. Når det er snakk om utforskende undervisning, kan det virke naturlig å ha med elementer av problemløsning, men også problemgenerering. For vår del er det et nokså nytt begrep, som verken har vært nevnt, eller gjort bruk av, i undervisningen vi har observert fra praksisperioder gjennom utdanningen. Grunnet den utforskende undervisningens art, vil det være rart å aldri berøre denne ferdigheten i matematikkfaget.

Van Harpen og Sriraman (2013, s. 201) skriver at viktigheten til problemgenerering, er understreket i læreplaner i flere land. Mens det, etter vårt skjønn, ikke er like sentralt i opplæringen i Norge. Silver og Cai (2005, s. 129) henviser til Einstein & Infeld (1938) og Hadamard (1945), og skriver at problemgenerering som en del av undersøkende matematikk, normalt sett er like viktig som problemløsning. Dette knytter klart viktigheten av problemgenerering sammen med viktigheten av problemløsning. Elgrably og Leikin (2021, s. 3) skriver at kreativitet i skolematematikk ofte er knyttet til problemløsning og -generering. Videre skriver de at problemgenerering kan bli brukt for å utvikle matematisk kreativitet. Silver (1997, s. 75) skriver at problemgenerering er knyttet til nøkkelaspektet av de moderne og klassiske forestillingene om kreativitet. Her kan en se at problemløsning og -generering er nært knyttet til hverandre, og knyttet opp til kreativitet i matematikk.

LK20 sin bruk av utforsking og problemløsning, samt koblingen til kreativitet i matematikk, ønsker vi å gjøre som Elgrably og Leikin (2021), Van Harpen og Sriraman (2013) og Silver og Cai (2005), nettopp å undersøke problemgenerering, særlig i henhold til kreativitet i matematikk. I våre øyne virker det særs interessant å undersøke sammenhengen mellom problemgenerering og matematisk kreativitet.

## 1.2 Formål, problemstilling og struktur

Formålet med studien er å undersøke om elevers kreativitet, innenfor problemgenerering, påvirkes ved presentering av et eksempel. Her videreføres forskning gjort av Stoyanova (1997), Ellerton (1986) og Van Harpen og Sriraman (2013). De bruker «problem posing», noe som vi oversetter til «problemgenerering», og bruker heretter «problemgenerering» i oppgaven for enkelhets skyld. Elevenes kreativitet i studien måles ved bruk av samme rammeverk som Van Harpen og Sriraman (2013), hvorpå oppgavene som blir generert blir kategorisert, og elevene blir tildelt en skåre innenfor flyt, fleksibilitet og originalitet ut fra oppgavene de lager, samt antall trivielle og uløselige oppgaver.

Formålet og avgrensningen til studien førte til følgende problemstilling og forskningsspørsmål:

*I hvilken grad påvirkes elevers kreativitet i problemgenerering ved bruk av eksempel?*

1) *I hvilken grad skårer elever på kategoriene fleksibilitet, flyt og originalitet, når de får presentert et eksempel i en oppgave?*
2) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de får presentert et eksempel?*

3) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de ikke får presentert et eksempel?*

Problemstillingen ønsker å sette lys på hvorvidt et eksempel i undersøkende undervisning i matematikk, har noe påvirkning på elevers kreativitet eller ikke. Det første forskningsspørsmålet har som hensikt å undersøke hvorvidt elever skårer bedre eller ikke, når det kommer til tre hovedelementer i kreativitet (se kap. 2.4). De to siste forskningsspørsmålene har som hensikt å undersøke hvorvidt elever med høy måloppnåelse i matematikk skårer bedre, eller dårligere, når det kommer til de tre hovedelementene i kreativitet, innenfor problemgenerering.

Kapittel 2 redegjør for relevant teori og rammeverket for studien. I kapittel 3 redegjør vi for, og begrunner, våre metodiske valg. I kapittel 4 presenterer vi resultatene, og analyserer funnene som fremkommer av studien. I kapittel 5 drøfter vi sammenhengen mellom resultatene og relevant teori, gruppene opp mot hverandre, og resultatene våre i lys av undersøkende undervisning. Siste kapittel tar for seg en avslutning, og her fremmer vi forslag til eventuell videre forskning på temaet.

# 2 Teori

I dette kapittelet gjør vi rede for teorien som brukes i studien, og er relevant for avhandlingen. Dette er et utvalg av teori og tidligere forskning, basert på det som kan gi relevante synspunkt til vår oppgave. Teorien som blir presentert representerer ikke nødvendigvis hele spekteret av forskning på området. Ettersom studien tar for seg et rammeverk som tidligere er gjennomført, presenteres rammeverket til Van Harpen og Sriraman (2013) i kapittel 2.8. Grunnet studiens natur presenteres teori angående problemgenerering (kapittel 2.2), problemløsning, kreativitet (kapittel 2.4) og forholdet dem mellom (kapittel 2.5).

## 2.1 Utforskende undervisning

I læreplanen i dag (LK-20) kan en finne spor etter utforskende undervisning. Ord som utforske, lage, forklare og vurdere går igjen i kompetansemålene etter 7. trinn (Utdanningsdirektoratet, 2020). Her kan det forstås at det er ønskelig at elevene skal utforske i større grad i matematikk, samt forklare og vurdere hva de har lært, og hva de gjør når de arbeider i matematikkfaget. Artigue og Blomhøj (2013, s. 1) definerer «inquiry based learning» (utforskende undervisning) til å la elevene arbeide på samme måte som en matematiker gjør, gjennom å stille nye spørsmål. Noe av begrunnelsen til å jobbe utforskende ligger til elevenes nedgang i interesse for kjernefag innen realfag med årene. Artigue og Blomhøj (2013, s. 1) skriver at tanken bak bruken av utforskende undervisning er at en tidlig skal gi elevene innblikk i en forskers hverdag. På denne måten ønsker en å fange elevenes interesse for faget, og dermed motivere til videre karriere innenfor dette. Dette kan en stille som en direkte motsetning til hvordan en tidligere har jobbet med realfagene, gjennom deduktive metoder.

Utforskende undervisning har røtter fra John Dewey, som er kjent for sitt utsagn «Learning by doing». Det betyr ikke nødvendigvis at en lærer bare av å gjøre, men en lærer av å reflektere og gjennom handlinger. En kan trekke linjer fra dette til dagens modeller for utforskende undervisning, eksempelvis Blomhøj (2016, s. 92-93) sin tredelte modell for undersøkende undervisning. Blomhøj (2016, s. 92-93) sin tredelte modell for undersøkende undervisning. De tre delene, eller fasene i Blomhøjs modell er iscenesettelse, elevenes selvstendig utforskende arbeid, og felles refleksjon og faglig læring. Den første fasen innebærer all praktisk informasjon og organisering av klasserommet. Den andre fasen omhandler det faktiske arbeidet elevene gjør. Elevene skal selv formulere problemer, finne løsningen på problemet, og til sist evaluere og oppsummere det de har lært. Det er gjennom den siste fasen, «felles refleksjon og faglig læring» elevene får mulighet til å konsolidere løsningene og strategiene elevene kom opp med.

Utforskende undervisning kan ofte innebære gruppearbeid. For å ha velfungerende grupper er det, ifølge Imsen (2020, s. 439), viktig at en har en gjensidig avhengighet i gruppen. Med andre ord betyr det at gruppen innad, med eventuell veiledning fra lærer, må tildele viktige oppgaver eller roller til alle deltakere i gruppen. På denne måten kan hver enkelt elev føle at de er positivt avhengige av de andre deltakerne, og at de står samlet om samme mål- nemlig å løse oppgaven de er tildelt. For at et gruppearbeid skal være godt fungerende er det viktig med sosiale ferdigheter og kommunikasjonskompetanse. Imsen (2020, s. 439) nevner igjen at det her er spesielt viktig med positivt gjensidig avhengighet. For å få til et godt samspill i gruppen er det viktig at alle gruppedeltakere individuelt tar ansvar, både for egne oppgaver/roller de er tildelt, og for det kollektive samspillet.

I forkant av et gruppearbeid er det viktig å vurdere hvilke gruppestørrelser, som er hensiktsmessig å bruke til oppgaven elevene får. En større gruppe vil naturligvis ha høyere sannsynlighet for å kollektivt inneha et bredere kunnskaps- og ferdighetsspekter. På den andre siden vil en større gruppe kreve en høyere sosial kompetanse fra elevene, slik at gruppemedlemmene kan regulere seg selv (Johnson et al., 1996, s. 65). Videre skriver Johnson et al. (1996) at elever bør plasseres i heterogene grupper basert på måloppnåelse. På denne måten kan elevene forklare for hverandre, og de kan kollektivt få en mer kompleks og dypere samtale og forståelse.

## 2.2 Problemgenerering

Problemgenerering kan sees på som et produkt av to deler, omformulering av et problem eller generering av et nytt problem. Hvor en kan lage problemet før, under eller etter en har fått løsningen. (Silver, 1994, s. 19). Her kan en tolke det til at svaret i seg selv ikke er det en ser på, men måten problemet er formulert på. En kan anse problemgenerering som et kritisk aspekt ved en lærers jobb, både ved å stille problemer til elever, samt å gjøre elevene bedre til å generere problemer selv. Lærere kan også bruke problemgenerering for å få en bedre forståelse av elevenes matematikkforståelse. Cai et al. (2015, s. 6) skriver at problemgenerering har mulighet til å heve læringen i matematikk, og grunnet dette har det en sentral plass i forskning og skolens praksis. Videre skriver Cai et al. (2015, s. 10) at selv om elever kan generere matematiske problemer, er de ikke alltid av høy kvalitet, samt at det kan virke som at elevene øker vanskelighetsgraden på problemene, hvis de har mer erfaring med å løse slike problemer selv. Hvorvidt elever presenterer vanskelige oppgaver, samt deres forståelse av matematikk, har en klar tråd til problemgenerering. Cai et al. (2015, s. 16) henviser til Silver (1997) hvor de drar frem at det er et forhold mellom elevers mulighet til å generere problemer, og deres utvikling av kreativ «fluency», «flexibility» og «novelty». Her oversetter vi det til henholdsvis flyt, fleksibilitet og originalitet.

### 2.2.1 Problemgenerering i utforskende undervisning

Generering av problemer har en klar tråd til kreativ tenking, samtidig som det er ulikt og kanskje viktigere enn problemløsning (Jay & Perkins, 1997, s. 257). Her kan en forstå det slik at løsningen i seg selv nødvendigvis ikke er det viktigste, ettersom fokuset er rettet mot det å generere et problem. Studier om matematikk handler ofte om hvordan det forventes at en skal streve med å nå frem ett korrekt svar, selv om det er et felt som passer med kreativ og original tenkning. Det er enighet om at kreativitet omfatter det å generere noe nytt, dog er det ikke en konsensus angående andre viktige konseptuelle forhold (Matsko & Thomas, 2015, s. 127). En kan se en tråd mellom kreativitet og problemgenerering, i den forstand av å skape noe nytt. Kreativ tenkning manifesterer seg i annerledes tenkning innenfor psykologi, noe som samsvarer med at evnen til å stille sentrale forskningsspørsmål ofte kan ansees som en indikator på eksepsjonelt talent i matematikk (Van Harpen & Sriraman, 2013, s. 204). Elevers evne til å generere sentrale spørsmål i matematikk, kan kobles til elevers divergerende tenkning og kreativitet. Her samsvarer dette med at en kan se indikatorer på eksepsjonelt talent i matematikk. Elevene som genererer kreativt sett gode problemer, kan vise matematisk talent. Silver (1994, s. 20) skriver at problemløsning i lang tid har blitt sett på som et karakteristisk trekk ved eksepsjonelt talent eller kreativ aktivitet, samt at det er en klar link mellom kreativitet og problemgenerering. Noe som kan indikere at en elevs kreativitet er med på å påvirke elevens problemgenereringsevne.

Ellerton (1986) fant at mer dyktige elever genererte mer beregningsvanskelige oppgaver, sammenliknet med sine jevnaldrende. I likhet med Ellerton, fant Krutetskii (1976) at elever med høyere evne klarte å se et problem og presentere det lettere, sammenliknet med jevnaldrende som trengte hint eller ikke klarte å generere et problem. Krutetskii (1976) og Ellertons (1986) forskning viser til at elever med høyere evne og dyktighet i matematikk i større grad klarte å presentere problemer. Med dette sagt viste Van Harpen og Sriraman (2013) i sin forskning at selv elever som hadde avansert matematikk, slet med å presentere gode og/eller nye problemer. Noe som tyder på en motsats av Ellerton og Krutetskii. Problemgenerering trenger ikke å skje i sammenheng med å løse et problem, men kan være det å skape et helt nytt problem ut fra en spesifikk situasjon eller erfaringer. De som kan anses som eksperter bruker betydelig mer tid, sammenliknet med de som er nybegynnere, når det kommer til problemgenerering (Silver, 1994, s. 20). Her viser det til at det kan være slik at elever med høyere evne og dyktighet, bruker mer tid på å generere problemer, noe som kan være grunnen til at de lager flere/bedre problemer. Cai et al. (2015, s. 15) viser til Van Harpen og Sriraman (2013) som fant at det er en sterk link mellom matematisk kunnskap og suksess i problemgenerering. Dette samsvarer med det både Ellerton og Krutetskii skriver om dyktige elever og presentering av problemer. Dog kan dette virke litt selvmotsigende, ettersom tidligere nevnt så kan elever med avansert matematikk slite med å presentere gode og/eller nye problemer.

Stoyanova (1997, s. 6) klassifiserer problemgenereringsoppgaver ut fra frihetsgrader. Her forekommer det tre grader, ut fra hvor mye informasjon deltakeren får presentert i oppgaven; Den åpne oppgaven, den semi-strukturerte oppgaven og den strukturerte oppgaven. En problemgenereringsoppgave kan klassifiseres ut fra en rekke parametere, slik som: informasjon gitt, informasjon som kan antas, mulige løsninger, overflødig informasjon, manglende informasjon og oppgaver med klare eller uklare mål. Stoyanova (1997) sine frihetsgrader blir brukt i Van Harpen og Sriraman (2013, s. 205) sin studie av elevers kreativitet i problemgenereringssituasjoner. Videre forklarer de at en problemgenereringssituasjon er definert som fri når en blir bedt om å generere problemer fra en naturlig gitt eller konstruert situasjon. En oppgave er semi-strukturert når en blir bedt om å generere problemer ut fra en åpen situasjon, og blir invitert til å utforske strukturen i situasjonen og fullføre den ved å bruke ferdigheter, kunnskaper, relasjoner og begreper fra matematiske erfaringer. En strukturert oppgave er når oppgaven er basert på et spesifikt problem. Vår studie tar for seg de tre aktuelle frihetsgradene, ettersom det allerede er dokumentert forskning som støtter opp under at dette kan brukes i problemgenereringssituasjoner, og innenfor kreativitet i matematikk.

## 2.3 Problemløsning

Problemløsning i matematikk har lenge vært et viktig aspekt i å lære bort, og lære, matematikk. Temaet har blitt en del av læreplaner verden over, både gjennom å lære om problemløsning og lære gjennom problemløsning (Liljedahl et al., 2016, s. 1). Her forstår en at det også har blitt en del av norsk opplæring, som nevnt tidligere. Dette er et tema som er skrevet om i lang tid, tidlig av Pólya (1957) og i nyere tid av Liljedahl et al. (2016), Novick og Bassok (2005) og Robertson (2016) for å nevne noen.

Et problem kan defineres på ulike måter, og kan også løses på ulike måter. Novick og Bassok (2005) skriver at en blir konfrontert med problemer på en daglig basis. Videre henviser de til Duncker og Lees (1945) som skriver at et problem oppstår når en person har et mål, men ikke vet hvordan målet skal nås. Når en ikke kan gå fra den gitte situasjonen til den ønskede situasjonen ved enkel handling, så må en ty til tenkning. Her kan en forstå det slik at når en ikke enkelt kan bevege seg fra situasjon A til situasjon B uten tenkning, så står en fremfor et problem. Dette støttes av Robertson (2016, s. 2), som skriver at en står overfor et problem når det er en forskjell fra hvor en er nå, og hvor en ønsker å være. Videre vises det til at et kjennetegn ved et problem, er at det er et mål en ønsker å nå, gjennom handling, men måten å komme seg dit er ikke umiddelbart åpenbar. Robertson (2016, s. 3) refererer til Reese (1994) og skriver at hvis det ikke er noe som hindrer fremgangen mot et mål, så er det ikke et problem å oppnå målet. Videre skrives det at en ikke står overfor et særlig problem når en kan bruke kunnskap for å komme seg til målet. Med dette sagt kan det poengteres at noe som kan være et problem for en person, trenger ikke nødvendigvis være et problem for andre. Eksempelvis elever med høy kontra lav måloppnåelse i matematikk. Her kan eleven med lav måloppnåelse tenke at en oppgave er et problem, og eleven med høy måloppnåelse ser på dette som enkel repetisjon. Dette kan være grunnet ulik kompetanse i faget, eller erfaring med liknende problemer fra tidligere.

En av de tidligste analysene av problemløsning innenfor matematikk, kom fra Pólya (1957), som delte problemløsning opp i fire faser ut fra kognitive prosesser:

1. Forstå problemet
2. Se hvordan ulike gjenstander er tilknyttet, hvordan det ukjente er linket til dataen, lage en plan
3. Gjennomføre planen

4. Se tilbake på det løste problemet, se gjennom og diskutere det

(Robertson, 2016, s. 4)

Her kan en se en tydelig inndeling av problemløsning og hvordan dette kan gjennomføres. Videre viser Robertson (2016, s. 4) til at andre har lagt til flere faser i denne inndelingen. Eksempelvis Hayes (1989) som la til «å finne problemet» i starten og «sammenfatte gevinster» til slutt (Robertson, 2016, s. 4). Robertson (2016, s. 8) skriver at det å studere problemløsning er relativt enkelt. En viser problemet til deltakerne, og venter og ser hva som skjer. Det er typisk at forskerne har manipulert instruksjonene, eksempelvis med et hint.

Elevers oppnåelse i matematikk er ofte vurdert ved å få dem til å løse problemer. Forskning viser til at elevers suksess i problemløsning er knyttet deres evner til problemgenerering (Cai et al., 2015, s. 15). Her kan en se at suksess innenfor problemløsning, og suksess innenfor problemgenerering kan knyttes til hverandre. Videre hevder L. D. English (1998) at evnen til å løse problemer blir forbedret gjennom problemgenerering, samt en bredere forståelse av matematikk. Her kan en tenke seg til at problemgenerering, i gitt grad, bidrar til økning av elevers evner til å løse problemer (Bonotto & Santo, 2015, s. 105). Det kan også tenkes at evnen til å generere problemer er med på å øke elevenes evne til å løse problemer. Prosessen problemløsning begynner med et tydelig mål, hvor en må lene seg på relevant tidligere erfaring, et såkalt repertoar (Liljedahl et al., 2016, s. 12). De skriver videre at tidligere kunnskap er et nøkkelelement i problemløsningsprosessen. Ved å ha tidligere kunnskaper som er relevante til problemløsning, eller gode evner til problemgenerering, skriver flere at dette kan være med på å øke ens evne til å løse problemer. Liljedahl et al. (2016, s. 12) drar videre frem at tidligere kunnskap innenfor problemløsning er med på å påvirke forståelsen av problemet, samt hvilken strategi som skal brukes. De understreker at tidligere kunnskap og erfaring er alt en har når en først skal angripe problemet. Som vist over henvises det til Pólya (1957), noe Liljedahl et al. (2016, s. 12-13) også drar frem. Her understreker de poenget med repertoar av tidligere erfaringer.

## 2.4 Kreativitet

En kan se på forskning på kreativitet generelt som noe splittende og polart. Eksempelvis i psykologi er det noen som ser på det som effekten av konvergent tenkning, og andre ser på det som divergent tenkning (Van Harpen & Sriraman, 2013, s. 202). Viktigheten av kreativitet er vanskelig å overestimere, og kreativitet er viktig for å tilpasse seg til en stadig utviklende verden med en kontinuerlig utvikling av teknologi og forskning (Leikin, 2013, s. 386). Forskning på kreativitet i matematikk drar ofte frem Silver (1997), og kjerneelementene innenfor kreativitet; flyt, fleksibilitet og «novelty»/originalitet.

Leikin (2013, s. 386) skriver at kreativitet i matematikk er en spesiell form for kreativitet, som har en åpenbar viktighet. Samt at skolematematikk burde tilby hver student med mulighetene til å smake på matematisk kreativitet, og at eleven skal få muligheten til å innse eget potensial innenfor kreativitet i matematikk. Leikin (2013, s. 387) viser til Piirto (1999) og skriver at matematisk kreativitet kan ansees som en spesifikk form for kreativitet, som tar for seg den logisk deduktive naturen i matematikken. Videre skrives det at det er vanskelig å gi en presis og bredt akseptert definisjon av matematisk problem, samt at det kanskje kan være umulig å oppnå. Her ser en at kreativitet er ansett som bredt innenfor matematikk, og samtidig er vanskelig å gi en spesifikk definisjon. Leikin (2013, s. 387) skriver også at matematisk kreativitet er et karakteristika av avansert matematisk tenkning, som er reflektert i evnen til å formulere matematiske objektiver og finne et forhold mellom dem. Leikin og Elgrably (2020, s. 3) viser til Ervynck og Tall (1991) som koblet matematisk kreativitet med avansert matematisk tenkning, og anså det som evnen til å formulere matematiske objektiver, og finne et forhold mellom dem. Dette er i tråd med hva Leikin (2013, s. 387) skriver om avansert tenkning, matematiske objektiver og forholdet dem imellom. Videre skriver Elgrably og Leikin (2021) at Ervynck og Tall (1991) sin definisjon av kreativitet inkluderer evnen til å utvikle strukturer som passer inn i den logisk deduktive disiplinen, og Ervynck og Tall (1991) kobler kreative løsninger i matematikk med innsikt og intuisjon.

Van Harpen og Sriraman (2013, s. 204) refererer til Usiskin (2000) sitt hierarki med åtte lag av matematisk talent, hvor det vises til at elever som er begavede og/eller kreative i matematikk har potensiale til å bevege seg opp profesjonelt, med passende affeksjons- og instruksjonsstilaser, når de beveger seg fra grunnskolen og inn mot universitetet. Videre viser de til at begavede/kreative elever har vært av spesiell interesse for forskere innenfor matematikkundervisning. Dette er noe som tyder på at matematisk kreativitet er et felt som er ønskelig å forske på. Silver (1997, s. 75) hevder at utforskende undervisning, som inneholder problemløsning og problemgenerering, kan hjelpe elever til å utvikle mer kreative tilnærminger til matematikk. Her kan det forstås at jobbing med problemløsning og -generering har noe å si for elevers kreativitet innenfor matematikk. Van Harpen og Sriraman (2013, s. 205) viser til at det hevdes at gjennom bruken av aktiviteter og oppgaver som omhandler problemløsning og -generering, kan lærere øke elevers kapasitet når det kommer til hoveddimensjonene i kreativitet; Flyt, fleksibilitet og originalitet. Dette er også i tråd med hva Leikin (2013, s. 388) skriver om utviklingen av kreativitet gjennom problemløsning. Leikin og Elgrably (2020, s. 3) skriver også at problemløsning og -generering er verktøy som utvikler kreativitet i matematikk. De tre overnevnte henviser til alle Silver (1997). Silver (1997, s. 76) skriver at problemgenerering lenge har vært sett på som en karakteristika for kreativitet. Videre skriver Silver (1997, s. 76 & 78) at koblingen til kreativitet ligger i samspillet mellom problemløsning og -generering. Dette samspillet innebærer formulering, forsøk på å løse, reformulere og til slutt løse problemet. Silver (1997, s. 76 & 78) viser også til at kjerneelementene i kreativitet er flyt, fleksibilitet og «novelty». Novelty forstås her som genererte problemer som er ulik de andre genererte problemene, i oppgaven blir dette referert til som originalitet. De tre nevnte kjerneelementene i kreativitet, er fokuset i denne oppgaven, og det som blir vektlagt i henhold til kreativitet.

## 2.5 Forholdet mellom problemløsning, problemgenerering og kreativitet

Problemgenerering kan sees på som en form for kreativ aktivitet, som kan inneholde oppgaver som involverer semi-strukturerte situasjoner og menneskelige interaksjoner (Bonotto & Santo, 2015, s. 103). Problemløsning er et bredt konsept, vanligvis relatert til generering av et nytt problem som respons til et krav om å lage et sett med problemer, eller et enkelt problem (Elgrably & Leikin, 2021, s. 4). Her kan en tenke seg til at det kan være en kobling mellom problemløsning og -generering, som kan knyttes opp mot kreativitet.

Silver (1994, s. 23) skriver at problemgenerering kan ha en verdi når det kommer til å hjelpe elever til å bli flinkere på problemløsning. Dette kan være en grunn til hvorfor problemgenerering er en del av læreplanen flere plasser. Videre skriver Silver (1994, s. 23) at problemgenerering har blitt inkorporert i Japansk undervisning, som et eksperiment for å se om det kan bidra til å øke elevenes kompetanse i problemløsning. Det dras også frem at elever, i USA, som har muligheten til å skrive sine egne problemer for å så løse dem, gjorde det bedre på tester, enn elevene som bare har løst tekstbokproblemer (Silver, 1994, s. 23). Silver (1997), viser til Silver (1994), og skriver at problemgenerering, sammen med problemløsning, er sentralt i matematikk og naturen som er matematisk tenkning. Videre knyttes problemløsning sammen med kreativitet, gjennom samspillet mellom problemløsning og -generering.

Problemgenerering og problemløsningsoppgaver har blitt brukt av flere for å identifisere kreativitet hos individer (Silver, 1997, s. 76). Her kan en dra frem Balka (1974) og Van Harpen og Sriraman (2013). Balka (1974) fikk deltakerne til å generere problemer som kunne bli besvart på bakgrunn av informasjon fra historier om virkelige situasjoner. Denne studien analyserte deltakernes var ut fra flyt, fleksibilitet og originalitet. Van Harpen og Sriraman (2013) sammenliknet elevers kreativitet, i form av flyt, fleksibilitet og kreativitet, på tvers av to land og tre områder. Studien tok for seg problemgenerering i hver gruppe, og ga skåre i flyt fra kjerneelementene i kreativitet, og sammenliknet gruppene opp mot hverandre. For å nevne noen. Silver (1997, s. 76) drar frem at problemgenerering og -løsning er veletablert for å vurdere kreativitet hos deltakere.

Kreativitet i matematikk er normalt sett assosiert med problemgenerering og -løsning (Elgrably & Leikin, 2021, s. 3). De drar også frem Silver (1997), og fremhever at kreativ problemløsning er forbundet med mental fleksibilitet. Leikin (2013, s. 388) henviser også til Silver (1997) som foreslår at utviklingen av kreativitet gjennom problemløsning følger flyt, fleksibilitet og novelty/originalitet. Videre viser Leikin (2013, s. 388) til Ervynck og Tall (1991)(1991), Krutetskii (1976), Pôlya (1973) og Silver (1997) som alle mener at å løse matematiske problemer, på mange måter, er nært relatert til personlig matematisk kreativitet.

## 2.6 Hint/transfer/Worked-example

Renkl (1999, s. 478) skriver om funn fra tidligere forskning, som viser at «worked-example» er effektivt for å lære seg nye ferdigheter i matematikk. Videre argumenteres det for at «worked-example frigjør arbeidsminne til å løse oppgaven, i stedet for å bruke det til å skape nye løsninger på egenhånd, gjennom problemløsning. Et annet argument for å bruke «worked-examples» er at målet med oppgaveløsningen blir å lære seg ferdigheten som trengs for å løse denne type oppgaver, hvor eksempelvis problemløsningsoppgaver har som mål å løse oppgaven, og ikke nødvendigvis lære seg metoden. Renkl (1999, s. 478) nevner at denne måten å tilegne seg matematiske ferdigheter, ikke nødvendigvis skaper læring i seg selv, men at det avhenger av hvor godt det enkelte individet forklarer den gitte måten å løse oppgaven på, til seg selv. Med andre ord vil en elev som slavisk følger en oppskrift uten å prøve å forstå den sannsynligvis ikke lære like godt, sammenlignet med en som nøye forklarer seg selv hvordan metoden fungerer. Skemp (1976, s. 2) definerer dette som instrumentell og relasjonell forståelse. En elev som utfører oppgaven gjennom å bruke «worked-example» som en oppskrift for å løse oppgaven med suksess vil ikke nødvendigvis forstå ferdigheten, til tross for at eleven får riktig svar, altså instrumentell forståelse. Motparten til dette vil forsøke å lære seg mekanismene i en formel eller et eksempel, og lære seg konseptet eller eksemplet på et dypere nivå.

### 2.6.1 Positiv/negativ overføring av løsningsstrategi

Ved analogiske oppgaver kan en forsøke å bruke tidligere lærte løsningsstrategier til å løse en ny oppgave. Ved å overføre løsningsstrategi må oppgavene være like nok i premissene slik at en overført strategi er tilstrekkelig, og riktig. Bassok (2003, s. 343) skriver at å overføre en strategi fra en oppgave til en annen, og vellykket løse problemet, kalles positiv overføring (positiv transfer). Når en bruker en strategi fra en tidligere oppgave for å løse en ny oppgave, hvor løsningen ikke er tilstrekkelig eller feil, kalles det negativ overføring (Bassok, 2003, s. 343).

### 2.6.2 Analogisk transfer

En måte å løse tilsynelatende nye problemer, kan være å bruke løsningsstrategier fra tidligere løste oppgaver. Et eksempel på dette er oppgavene en kan finne i flere læreverk i slutten av hvert kapittel, som oppsummerer det en tidligere har lært. Her kan en ofte bruke tidligere lærte strategier for å løse disse nye oppgavene. En skiller mellom strukturelle likheter (structural similarities), og overflate-likheter (surface similarities) (Bassok, 2003, s. 343). Likheter i informasjon gitt i oppgaver som ikke er relevant for løsningen av oppgaver, kaller Bassok (2003, s. 344) for overflate-likheter. Dette kan for eksempel være navn, sted, objekt, tall og så videre. Under er et eksempel på to oppgaver som har overflatelikheter, samtidig som de har strukturelle ulikheter.

> **Oppgave 1:**
> **Ole og Fredrik har 20 frimerker, og skal selge de til 5 personer. Hvor mye får hver person?**
> **Oppgave 2:**
> **Ole og Fredrik har 20 frimerker og skal selge de for 5 kroner per stykk. Hvor mye tjener Ole og Fredrik?**

Her er personene like, objektene er like, og tallene er like, og er derfor overfladisk lik (Bassok, 2003, s. 344). Fortsatt vil ikke en elev kunne bruke samme løsningsstrategi fra oppgave 1, til å løse oppgave 2, ettersom oppgavene har strukturelle ulikheter. På den andre siden kan to oppgaver ha overfladiske ulikheter, men fortsatt ha strukturelle likheter. Dermed kan to oppgaver være tilsynelatende ulik, men fremdeles kunne løses på samme måte. Under kommer eksempel på dette.

**Oppgave 1:**

**For å feste en planke til en vegg trenger man 12 spiker. For å bygge en vegg trenger Ole 17 planker. Hvor mange spikere trenger Ole for å bygge hele veggen?**

**Oppgave 2:**

**Ole har en gjennomsnittsfart på 18 kilometer i timen, og sykler i 4 timer. Hvor langt sykler Ole?**

Disse eksemplene er strukturelt like, ettersom løsningen på oppgavene er å multiplisere det ene tallet med det andre. Overfladisk ser oppgavene ulike ut, siden det er forskjellige tall, enheter, objekter og personer.

Bassok (2003, s. 344) skriver at det er mer sannsynlig at en legger merke til likheter i overflaten av en oppgave, enn likheter i strukturen til en oppgave. Overfladiske likheter i oppgaver med strukturelle ulikheter kan på den måten føre til negative overføringer av løsningsstrategi. Samtidig kan overfladiske ulikheter i oppgaver med strukturelle likheter hindre positive overføringer av løsningsstrategi (Bassok, 2003, s. 345).

## 2.7 Tidligere forskning

Dette kapittelet tar for seg tidligere forskning på problemgenerering, og hvilke effekter dette kan ha på elever. Den tidligste forskningen som forekommer under, er fra 1986, og den nyeste fra 2022. Ettersom det er en god del forskning som omhandler dette temaet, har vi valgt ut relevant forskning som kan tilby et mer nyansert bilde av temaet.

Stoyanova (1997, s. iii) skriver at selv om problemgenerering har blitt anbefalt som en fornuftig aktivitet i matematikk, så er forskning på situasjoner med problemgenerering hvor elever er involvert en mangelvare. Dog er dette noen år siden. Van Harpen og Sriraman (2013, s. 201) skriver at viktigheten av problemgenerering er understreket i utdanningsdokumenter i mange land, noe som kan tyde på at fokuset er endret.

I en studie gjennomført av Ellerton (1986) ble to elevgrupper valgt, ut fra deres resultater på fem matematiske spørsmål. For å gjennomføre oppgavene trengte elevene en rekke resonneringsferdigheter. Her ble gruppene fordelt inn i to grupper, skilt av mer eller mindre dyktige i matematikk, henholdsvis fem av fem rette og null av fem rette. Gruppene ble fordelt fra et utvalg på 154 elever, hvorpå det ble åtte elever i hver dyktighetsgruppe og alderen på elevene var 11-13 år. Elevene ble bedt om å lage matematiske problemer som de trodde ble vanskelig for en venn å løse (Ellerton, 1986, s. 261 & 263). Ellerton (1986) fant at på generell basis så var problemene til de dyktigere elevene mer beregningsmessig vanskelige, og trengte gjerne flere operasjoner for å kunne svare på. De dyktigere elevene genererte også mer komplekse tallsystemer, samt så viste de mer flyt i bruk av algoritmer og matematisk språk. Det var lite som viste til at mindre dyktige elever planla problemene, i motsetning til mer dyktige elever hvor svarene på problemet var mer konsistent med problemet, som viser noe planlegging (1986, s. 266). Videre skriver Ellerton (1986) at genererte problemer reflekterer en elevs miljø, hvorpå elever som bruker tekstbøker mye i større grad blir å generere et problem som er likt et tekstbok-problem. Mer dyktige elever genererte få «word problems» (noe vi anser som tekstbok-problemer), noe som kan vise til en større grad av kunnskaper i abstrakt form for presentasjon, eller at de ikke anser «word problems» som noe vanskelig (1986, s. 268). Ellerton (1986, s. 269) skriver til slutt at resultatene i studien viser hvordan mer dyktige elever bruker erfaringene sine mer flytende i en ny situasjon.

Zhang et al. (2022, s. 3) gjennomførte en studie i 16 klasser på 3 skoler i Kina, med totalt 669 elever på 6. trinn. Studien tar for seg formatet fra tidligere forskning, ved navn «PPTask-Num» test. I forkant ble det også gjennomført en pilotstudie. Studien til Zhang et al. (2022, s. 4) ble gjennomført i to runder, hvorpå første runde inneholdt problemgenerering og andre runde problemløsning. Begge varte i 45 minutter, og førstnevnte ble gjennomført en måned før den andre. Svarene i problemgenereringsrunden ble kategorisert i matematiske spørsmål, ikke-matematiske spørsmål og påstander, i henhold til Silver og Cai (1996). Videre ble det også delt inn etter om problemet var mulig å løse eller ikke. Resultatene på første oppgave viste at 25,6% av de genererte problemene var matematisk løselige og andre oppgave hadde 94,2% løselige oppgaver. De viser til at elevene klarte i rimelig grad å generere problemer som var løselige, og få problemer som var ikke-matematiske. Studien viser til at det virker som elevene i studien klarte å generere problemer som ikke bare var løselige, men også noe komplekse. Funnene til Zhang et al. (2022, s. 13) korresponderer med funnene til Leung og Silver (1997) sin studie. Resultatene, til Zhang et al. (2022), viser at deltakerne generelt var mer suksessfull med problemgenerering, hvis oppgaven inneholdt numerisk informasjon. Det Zhang et al. (2022, s. 13) videre fremhever er at funnene samsvarer med tidligere forskning gjort av L. D. English (1998). Denne studien kom frem til at de genererte flere problemer under oppgaver med kontekst. Zhang et al. (2022, s. 13) poengterer at dette kan være grunnet at elevene har mulighet til å bruke elementene i konteksten til å generere flere problemer.

Studien gjennomført av Stoyanova (1997, s. 49-51) tar for seg endringer som forekommer når problemgenerering blir introdusert i et klasserom. Den ble gjennomfør som et eksperiment i to stadier, en pilotstudie og en hovedstudie. Pilotstudien ble gjennomført på åtte økter på en time, i en gruppe på 40 elever. Formålet med pilotstudien var å samle data for å selektere deltakere til hovedstudien, som inneholdt både problemløsning og problemgenerering. I tillegg til de to gruppene, ble en gruppe på 112 elever valgt til å delta for å undersøke strategier for problemgenerering. Alderen i gruppene i hovedstudien strakk seg fra 7 til 13 år, hvor hoveddelen var 12 år gammel, og 12-13 år gammel i gruppen for problemgenereringsstrategier. Den ene testen, «the problem-posing test», var basert på premissene at det skulle inneholde situasjoner som gikk under kategoriene fri, semi-strukturert og strukturert. Innunder hver kategori, skulle deltakerne generere problemer (Stoyanova, 1997, s. 62-63). Videre skriver Stoyanova (1997) at studien utforsket deltakernes karakteristika når det kom til problemgenerering innenfor fri, semi-strukturert og strukturerte oppgaver. Stoyanova (1997, s. 193-194) konkluderer med at deltakerne brukte fire hovedkategorier når det kom til problemgenerering, reformulering, rekonstruering, etterligning og oppfinne (invention). Det kommer frem at studentene så ut til å ha en naturlig kapasitet til å generere problemer på bakgrunn av gitte kalkulasjoner. Deltakerne anerkjente selv at problemene de genererte under kategorien fri var relativt like problemene de hadde sett i tekstbøker eller løst på skolen. Avslutningsvis skriver Stoyanova (1997, s. 285-286) at problemgenereringsaktiviteter i matematiske klasserom har blitt anbefalt å være en del av opplæringsdokumenter i flere land. Her vises det til hvordan problemgenerering kan være en del av undervisningen, samt i hvilke aktiviteter dette kan passe inn.

I likhet med overnevnte studier, har Van Harpen og Sriraman (2013) studert bruken av problemgenerering. Deres vinkling er direkte myntet mot kreativitet i matematikk, og hvorvidt problemgenerering er med på å påvirke dette. Studien tar for seg tre klasser, en fra USA og to fra Kina, hvorpå alle er på «high school». Til sammen analyserte de prøvene til 129 studenter, 30 fra USA og 99 fra to ulike distrikter i Kina. I forkant av problemgenereringstesten ble det gjennomført en kunnskapstest i matematikk, dette for å måle elevenes kunnskap i faget. For gjennomføringen av problemgenereringstesten tok de for seg rammeverket til Stoyanova og Ellerton (1996), som inkluderte fri, semi-strukturert og strukturerte oppgaver (Van Harpen & Sriraman, 2013, s. 206-207). I samsvar med forståelsen av kreativitet i matematikk, ble svarene analysert i form av flyt, fleksibilitet og originalitet. Deltakerne ble spesifikt bedt om å generere så mange og så vanskelige oppgaver som de klarte, dog genererte ikke deltakerne mange vanskelige problemer (Van Harpen & Sriraman, 2013, s. 215-216). Studien finner indikasjoner på at det er forskjeller på problemgenereringsevnene i de tre klassene, og at den ene gruppen i Kina (Jiaozhou disktriktet) genererte flere ikke-løselige oppgaver. Dette står i kontrast til det Cai og Hwang (2002) fant gjennom sin studie av yngre elevers problemgenerering i matematikk. Van Harpen og Sriraman (2013, s. 218) drar frem L. English og Sriraman (2010) sin kritikk av at problemgenerering ser ut til å ha nådd en blindgate. Videre skriver Van Harpen og Sriraman (2013, s. 218) at problemgenerering har fått lite oppmerksomhet som et aspekt av kreativitet i matematikk, og understreker at det er nødvendig med mer forskning på dette innenfor matematisk læring.

Cai og Hwang (2002, s. 404-407) gjennomførte en studie på 253 elever fra Kina og USA, med henholdsvis 155 og 98 deltakere. De ønsket å undersøke elevenes evner til generativ tenkning i problemgenerering. Her skulle elevene gjennomføre tre par  oppgaver innenfor problemgenerering, hvor de skulle generere et enkelt, et moderat og et vanskelig problem. Hvert svar ble kategorisert ut fra om det var en forlengelse av oppgaven, eller ikke. Studien viste at amerikanske elever genererte flere forlengelsesproblemer, sammenliknet med de kinesiske, samt at de kinesiske elevene genererte fler ikke-forlengelsesproblemer enn de amerikanske elevene i første oppgave. I den andre oppgaven var det mer sannsynlig at de kinesiske elevene genererte problemer med den gitte informasjonen, og her genererte amerikanske elever flere forlengelsesproblemer enn de kinesiske elevene (Cai & Hwang, 2002, s. 412-414). Videre skriver de at ut fra studien så var det lite beviser på at det var mer sannsynlig at de amerikanske elevene genererte bilderelaterte problemer, samt at prosentandelen kinesiske og amerikanske elever som genererte distinkte problemer var nær identiske. De fant også at de kinesiske elevene genererte problemer i sekvens etter hverandre, som kan vise til et mønster, i større grad enn de amerikanske. Studien viser at de amerikanske elevene genererte problemer i sekvens i en mindre styrt retning (Cai & Hwang, 2002, s. 418-420).

## 2.8 Rammeverk

Van Harpen og Sriraman (2013) gjennomførte en studie, hvor de undersøkte elevers kreativitet i problemløsning, på tvers av land og kulturer. De sammenliknet tre klaser, to fra Kina og en fra USA, opp mot hverandre. Mer spesifikt analyserte de kreativiteten til «High School» elever når det kom til problemgenerering i geometriske scenarioer (Van Harpen & Sriraman, 2013, s. 201)

Studien deres tar for seg deltakere fra tre ulike områder, hvorpå de valgte ut elever som bedriver avansert matematikk. Til å starte med var det 68 deltakere fra Jiaozhou, og 73 fra Shanghai, to byer i Kina. Fra USA deltok 77 stykker. For å kunne delta i studien måtte deltakerne gjennom 4 tester, hvorpå noen ikke gjennomførte én eller to av dem. Grunnet dette falt utvalget til 55 stykker fra Jiaozhou, 44 fra Shanghai og 30 fra USA (Van Harpen & Sriraman, 2013, s. 206-207). Her ser vi at studien originalt skulle hatt flere deltakere, men grunnet fravær under testene falt antallet noe. Fra USA var 13 av deltakerne «Pre-calculus students» (et fag som innebærer matte på et nivå slik at elevene skal kunne delta i calculus etterpå), og 17 hadde en avanserte calculus fag, alle fra Shanghai var i 11. klasse og alle fra Jiaozhou gikk i 12. klasse. Fra USA deltok 13 gutter og 17 jenter, 19 jenter og 25 gutter fra Shanghai og 18 jenter og 27 gutter fra Jiaozhou (Van Harpen & Sriraman, 2013, s. 207). Studien deres viser altså at noen i USA har mer avansert calculus enn andre, samt at kjønnsfordelingen strekker seg i favør av gutter.

Måleinstrumentene som ble brukt var både en matematisk kunnskapstest og en matematisk problemgenereringstest. Her ble førstnevnte brukt for å måle deltakernes kunnskap i generell matematikk, de brukte en test kalt «National Assessment of Educational Progress (NAEP), ettersom denne testen passet studien (Van Harpen & Sriraman, 2013, s. 207). Videre henviser Van Harpen og Sriraman (2013, s. 207) til Stoyanova og Ellerton (1996) sitt rammeverk for matematisk problemgenerering («problem posing»), videre henviser Van Harpen og Sriraman (2013) til Stoyanova (1997) og Cai (2000) som de brukte for å utvikle problemgenereringstesten. Testen bestod av tre oppgaver, i ulike frihetsgrader, og var som følger:

*Task 1 Free problem-posing situation:There are 10 girls and 10 boys standing in a line.Make up as many problems as you can that use the information in some way.*

*Task 2 Semi-structured problem-posing situation:In the picture below (Fig.1), there is atriangle and its inscribed circle. Make up as many problems as you can that are insome way related to this picture. The problems could also be real-life problems.Again, do not limit yourself to the problems you have seen or heard of—try tothink of as many possible and challenging mathematical problems as you can.*

*Task 3 Structured problem-posing situation: Last night there was a party at your cousin'shouse and the doorbell rang 10 times. The first time the doorbell rang only oneguest arrived. Each time the doorbell rang, three more guests arrived than hadarrived on the previous ring. 1. How many guests will enter on the 10th ring? Explain how you found your answer. 2. Ask as many questions as you can that are in some way related to this problem*

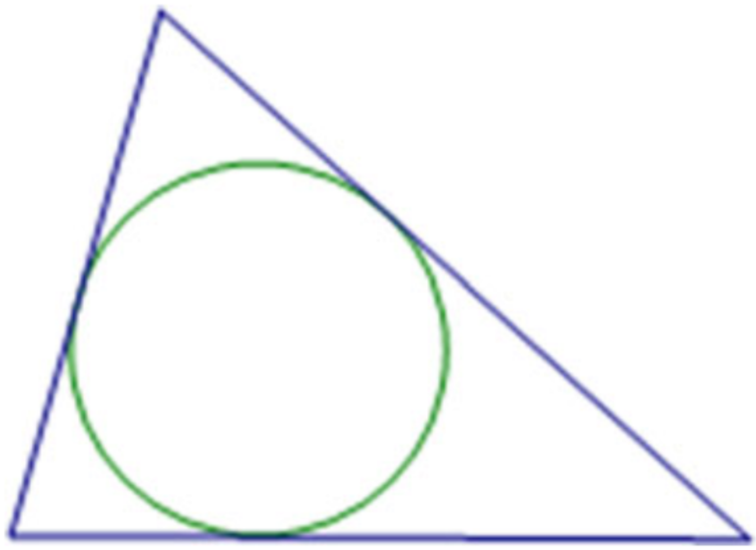

**Fig. 1** Semi-structured problem posing situation

***Figur 1**: Figur til problemgenereringstest, Van Harpen & Sriraman.*

*(Van Harpen & Sriraman, 2013, s. 207-208)*

For å motivere deltakerne til å gjøre sitt beste, lagde de et scenario, der de ba deltakerne om å forestille seg en konkurranse som omhandlet problemgenerering med alle «High Schools» i byen. De inkluderte i tillegg at den skolen som lagde best og/eller flest problemer, skulle få en belønning. Samt tok de frem en fiktiv karakter «Jenny», som angivelig hadde laget 5 gode problemer til hver oppgave, og hadde skrytt av at ingen kunne gjøre det like bra som henne. Til slutt ba de deltakerne om å ikke avgrense seg til problemer de har hørt eller sett, men prøve å tenke på så mange og så vanskelige problemer som de kunne (Van Harpen & Sriraman, 2013, s. 208).

Testene ble foretatt på 50 min hver, og hadde to ukers mellomrom mellom hver test. Kunnskapstesten ble også brukt for å evaluere forskjeller mellom de tre gruppene. I etterkant av testene ble dataene analysert ut fra deres flyt og fleksibilitet. Alle svarene ble skrevet inn i et Word-dokument, hvorpå hver gruppes svar ble separert slik at de kunne skille mellom gruppene. Problemene ble kategorisert ut fra hvilken type problem det var, og om det var noen som ikke tilhørte i en spesifikk gruppe, kom de under «andre». Studien foretok en analyse av originaliteten til hvert enkelt genererte problem, relativt til andre grupper også, hvorpå et problem var originalt om det ble generert av 10% eller færre deltakere. Inn under flyt forekom det analyse av genererte problemer for hvorvidt de var uløselige («nonviable») eller trivielle (Van Harpen & Sriraman, 2013, s. 209-212).

# 3 Metode

Her gjør vi rede for og begrunner metodene vi har gjort brukt av i oppgaven. Ettersom studien tar utgangspunkt i Van Harpen og Sriraman (2013) sin studie, er metoden vår ganske så lik deres. Her forekommer det noe ulikt, ettersom vi undersøker kreativitet med og uten eksempel, mens de undersøkte kreativitet på tvers av grupper uten noen tiltak.

## 3.1 Valg av og bruk av kvantitativ metode

Kvantitativ metode bygger på et post-positivistisk syn på virkeligheten (Postholm et al., 2018, s. 91). Det vil si at en definerer virkeligheten som forutsigbar, kontrollerbar og objektiv, med mer, som videre fører til forskning som produserer absolutte lover (Cohen et al., 2018, s. 16). I virkeligheten er dette vanskelig å tenke på som en sannhet når en driver forskning på mennesker, ettersom en umulig kan kontrollere alle variabler ved et menneske (Postholm et al., 2018, s. 91). Dermed vil også kvantitativ metode kreve fortolkning av målingene en gjør, som i seg selv kan innebære feilkilder. For å se på hvordan elevers arbeid med problemgeneringsoppgaver utarter seg ved bruk av eksempel kontra ikke, fant vi det hensiktsmessig å bruke kvantitativ metode, slik at vi kunne tallfeste effekten, samtidig som vi kunne minimere feilkilder knyttet til forsker-informant samhandling (Christoffersen & Johannessen, 2012, s. 17).

Studiens metode baserer seg på studien til Van Harpen og Sriraman (2013), og hvordan de gjennomførte det. Her forekommer det noen forskjeller, grunnet elevenes alder, praktiske årsaker og ulikheter i hva som spesifikt skal forskes på. Deres studie tar for seg tre grupper elever, to grupper i Kina og en gruppe i USA, og sammenlikner deres resultater. Vår studie tar for seg to grupper elever, en gruppe med eksempeloppgave til oppgavene, og en uten, innad i samme kommune, og sammenlikner resultatene mot hverandre. Her er det viktig å poengtere at studien til Van Harpen og Sriraman (2013) bruker Stoyanova (1997) for å utvikle problemgenereringstesten. Studien er å anse som kvantitativ med et eksperimentelt casestudie-design, ettersom det forekommer en testgruppe og en kontrollgruppe som gjennomfører en test uten stimuli i forkant eller etterkant.

I lærebøker, og i klasserommet er det vanlig med bruk av eksempler, som en hjelp for elevene når de skal arbeide selvstendig eller i grupper. Dette kan være eksempler som er originale eller tradisjonelle, eksempler med andre tall, men samme form for utregning eller eksempler med en indirekte instruksjon om hva som skal gjøres. Utforskende undervisning er spesielt, der elevene i noen tilfeller må lage problemene de skal løse selv. Vi ønsker derfor å gjennomføre en case-studie, hvor vi ser om et eksempel til en oppgave har en effekt på elevers kreativitet. En gruppe blir presentert tre oppgaver i ulike frihetsgrader og et eksempel til hver oppgave, den andre gruppen får samme tre oppgaver, uten eksempel. Dette blir en gruppe med tiltak, i form av eksempel, sammenliknet med en kontrollgruppe. Videre ønsker vi å se på eventuelle forskjeller som kan fremkomme av gruppenes tester. I tillegg ønsker vi å se om et eksempel kan ha noen effekt på elevers kreativitet når det kommer til generering av problemer i matematikk.

Enhver forskning i «socialstudies» kan sies å være en casestudie (Cohen et al., 2018). En casestudie kan ta for seg blant annet et fenomen, en skoleklasse, en gruppe mennesker, eller et kollegium ved en arbeidsplass. Formålet til en casestudie kan være å forsøke å danne begreper eller teorier, som en senere kan generalisere til en større populasjon (Wæhle et al., 2020). For eksempel kan dette være å undersøke et fenomen i én skoleklasse, for så å gjøre bruk av informasjonen innhentet fra casestudien når en driver skoleutvikling på regionalt eller nasjonalt nivå. Vi valgte en casestudie for å se nærmere på hvordan en elev presterer når den arbeider med problemgenereringsoppgaver, og hva som vil gagne elevene best. I dette tilfellet- eksempeloppgaver, eller ikke.

Vår studie tar for seg, i likhet med Van Harpen og Sriraman (2013), elevers kreativitet når det kommer til problemgenereringsoppgaver. Ulikt studien deres foretar vi en problemgenereringstest hvorpå en gruppe får stimuli i form av eksempel på et problem til hver enkelt oppgave, og en kontrollgruppe uten stimuli. Oppgavene tar utgangspunkt i ulike frihetsgrader, noe både Stoyanova (1997) og Van Harpen og Sriraman (2013) også brukte. Frihetsgradene er følgelig fri, semi-strukturert og strukturert. En dypere forklaring av oppgavene forkommer i kapittel 3.5.

## 3.2 Forskningsdesign/-strategi

Prosjektet har et design som kan sammenliknes med et eksperimentelt design, dette innebærer at utvalget ble trukket grunnet tilgjengelighet og egnethet i form av alder. Noe som kalles *convenience sample* (Creswell, 2014, s. 158). Studien tar for seg elevers genererte problemer innenfor temaet problemløsning og problemgenerering, hvorpå to gruppers svar blir sammenliknet med hverandre. Den ene gruppen får lagt ved et eksempel på en løsning, og den andre gruppen får ikke dette. Studiens forskningsstrategi er nært sammenlignbar med det Postholm et al. (2018) skriver om kontrollerte eksperimentelle casestudie, hvor en gruppe med tiltak sammenlignes før og etter tiltaket er tatt i bruk (Postholm et al., 2018, s. 70). Herunder er det et «kontrollert eksperiment», ettersom det er vi som har kontrollen over tiltaket. Nærmere definert er det et post-test-kontroll- og eksperimentgruppe-design, ettersom vi bare gjennomfører en test, hvor vi gir en eksperimentgruppen et stimuli, mens kontrollgruppen ikke får det (Cohen, et. al., 2018, s. 403). Ifølge Cohen et. al. (2018, s. 392-393) kjennetegnes kontrollerte eksperiment ved flere faktorer. For at en skal kunne kalle det et kontrollert eksperiment er utvalget en viktig faktor. Et tilfeldig utvalg er viktig. Fordelingen av deltakerne på eksperiment- og kontrollgruppen er også sentral. Utvalget skal fordeles tilfeldig på de to gruppene, men det er fortsatt viktig at det ikke er store forskjeller mellom gruppene. For vår studie vil eksempelvis elevenes måloppnåelse og kjennskap til temaet være viktig. Videre er kjennskap til, og kontroll av variablene viktig. For å kunne si noe om årsak-virkning er det elementært at en beholder de samme variablene for begge gruppene, bortsett fra variabelen en ønsker å undersøke. En må i tillegg forsikre seg om at det ikke foregår krysskontaminering mellom kontroll og eksperimentgruppene. Krysskontaminering vil for vår del eksempelvis være samtale eller sniktitting mellom elever som sitter vedsiden av hverandre, men er i forskjellige grupper (kontroll-eksperiment).

Studien sammenlikner to grupper opp mot hverandre, ettersom studiens natur gjør det vanskelig å sammenligne før og etter tiltak. Hensikten er også sammenlignbar med det Kokinov et al. (1997) sin studie av effekten til kontekststimulus når det kommer til problemgenerering. Vårt forskningsdesign tar for seg eksempel i Kokinov et al. (1997) sin hint-kategori, samt sammenligner en gruppe opp mot en annen. Det foretas endringer i begge gruppene, hvorpå den ene gruppen får stimuli i form av eksempel, og den andre gruppen ikke får noe annet enn oppgavene. Fordelingen blir slik at det er både med og uten eksempel i begge klassene, ettersom den ene klassen har mer erfaring med undersøkende undervisning.

Dog er dette ikke omfattende erfaring. Gruppene blir henholdsvis referert til som eksperimentgruppen og kontrollgruppen videre i oppgaven.

## 3.3 Utvalg

Utvalget tar utgangspunkt i deltakere som er tilgjengelige gjennom bekjentskap via studiet og praksisperioder i studiet. Studiens originale gruppestørrelse var 35 elever, fordelt på 2 byskoler, med 19 elever i en klasse og 16 elever i den andre klassen. Deltakerne i studien ble til sammen 25 elever, ettersom noen falt fra grunnet fravær den aktuelle dagen, og to elevers oppgaver ble tatt ut av studien. Her ble dette bestemt av oss, ettersom elevene fikk hjelp utenfra, noe som gjør besvarelsen ugyldig. Gruppene ble fordelt til 11 og 14 elever, henholdsvis med eksempel og uten eksempel, dette er grunnet at fordelingen tok sted før gjennomføringen, og ingen endringer ble foretatt på test-dagen. Den opprinnelige størrelsen står i tråd med Borg og Gall (1979, s. 194-195) som skriver at et utvalg på 15 deltakere i hver gruppe er stort nok. Skolene gruppene er fordelt på befinner seg i samme kommune, på samme klassetrinn og er derfor under samme læreplan, samt skal nå samme opplæringsmål i løpet av 7. års trinn. Begge skolene er også betegnet som byskole. Fra byskole B deltok 11 av 16 elever, 4 elever var borte den dagen, og 1 elevs besvarelse ble forkastet, ettersom eleven fikk hjelp fra støttelærer som var i klassen. Det var derfor vanskelig å kontrollere hva som ble sagt til eleven, og hvilken type hjelp eleven fikk. Fra byskole A deltok 14 av 19 elever, 4 var borte og 1 elevs besvarelse måtte forkastes siden denne ikke kunne anses som gyldig ettersom eleven fikk hjelp til å gjennomføre testen. Elevene fra byskole B har noe erfaring med problemløsning-/problemgenereringsoppgaver, men denne erfaringen er ikke dokumentert. I tillegg var det vanskelig å fastslå hva «noe erfaring» betyr. På byskole A ble det sagt av matematikklærer i klassen at elevene ikke hadde noe erfaring med verken problemløsning eller problemgenereringsoppgaver.

Gruppene ble fordelt ut fra elevenes måloppnåelse i matematikk, slik at det skulle være noenlunde like mange i hver gruppe. For å inndele i høy, middels og lav måloppnåelse fikk vi ved byskole A se på kartleggingsprøvene til elevene. Ut fra skåren til elevene tildelte vi elevene en måloppnåelse. Dette ble gjort i henhold til kartleggerens veiledning. Intervallet 80-120 i skåre var middels, og alt over og under dette intervallet var henholdsvis høy og lav måloppnåelse. Ved byskole B fikk vi elevenes måloppnåelse i matematikk presentert av kontaktlæreren, og ut fra dette ble gruppene fordelt. Videre ble hver enkelt elev tildelt et tall innenfor sin måloppnåelse, og så randomisert fordelt inn i gruppen med eksempel eller uten eksempel. Ut fra dette ble gruppefordelingen originalt 17 med eksempel og 18 elever uten eksempel. Grunnet frafall endte gruppestørrelsene på 11 og 14 elever, henholdsvis med og uten eksempel. Herunder ble det 4, 4 og 3 elever med eksempel og henholdsvis høy, middels og lav måloppnåelse i matematikk, og 3, 8 og 3 elever uten eksempel med henholdsvis høy, middels og lav måloppnåelse. Originalt skulle dette vært 6, 7 og 4 elever med eksempel, og 4, 8 og 5 elever uten eksempel, igjen henholdsvis med høy, middels og lav måloppnåelse.

## 3.4 Beskrivelse av oppgavene.

Oppgavene i studien er hentet fra Van Harpen og Sriraman (2013). De gjennomførte en studie som så på kreativitet i problemgenerering, hvor de fant at elever på «high school» selv slet med å generere oppgaver. Årsaken til å velge oppgaver som tidligere er brukt, er at de er utformet av forskere, og utprøvd tidligere, i tillegg er de også blitt klassifisert. Dette betyr også at andre forskere har designet og analysert oppgavene, og deres validitet innenfor feltet, som derfor kan være med på å styrke bruken av oppgavene. Vi anså det som en styrke å bruke oppgaver som har litteratur og forskning som argumenterer for bruken av dem, samt kobler de til kreativitet. Siden vår oppgave omhandler kreativitet og matematikk, kan det være, for vår del, lurt å bruke tidligere styrket forskning innenfor feltet. Vi anså dog den ene oppgaven, oppgave 2, fra Van Harpen og Sriraman (2013) som for avansert for elever på 7. års-trinn i den norske skolen. Grunnet dette omformulerte vi oppgave 2, slik at den skulle passe bedre til deltakerne. Oppgaven ble re-designet slik at kjernen skulle bevares, og samtidig holde samme struktur som Van Harpen og Sriraman (2013) sin oppgave. Ut fra dette klassifiseres den i samme kategori. Her er det verdt å nevne at oppgaven vedlikeholder sitt tema angående geometri, og sin grad av tilgjengelig informasjon.

Den første oppgaven i vår studie faller inn under «fri problemgenereringssituasjon» ettersom elevene blir bedt om å generere problemer ut fra en konstruert situasjon, som også teoretisk sett kunne vært mulig. Oppgave nummer to er en «semi-strukturert problemgenereringssituasjon», ettersom at deltakerne blir presentert en åpen situasjon hvor de får muligheten til å utforske strukturen i oppgaven, samt bruke sine ferdigheter, kunnskaper og begreper fra tidligere matematiske erfaringer. Dog er denne noe endret, som nevnt over, så en kan problematisere dette ved å argumentere for at den ikke er semi-strukturert siden den ikke er anerkjent av litteratur som en semi-strukturert oppgave. Etter vårt skjønn faller den inn under denne kategorien, og er derfor med i studien. Den siste oppgaven er en «strukturert problemgenereringssituasjon» ettersom elevene blir presentert en oppgave som er basert på et spesifikt problem. Alle tre oppgavene er i tråd med frihetsgradene som er skrevet om av Van Harpen og Sriraman (2013, s. 205).

Oppgavene i studien er oversatt til Norsk etter beste evne, slik at elever på 7. årstrinn skal kunne forstå innholdet og ikke slite med å gjennomføre problemgenereringen. Ved oversettelsen er det også gjort et forsøk på å vedlikeholde kjernen i oppgavene i størst mulig grad, slik at den ikke forsvinner, og at oppgavene er noenlunde likt tidligere forskning. De tre oppgavene som ble gitt til elevene er som følger:

**Oppgave 1:** *Det står 10 gutter og 10 jenter i kø. Lag så mange oppgaver du kan, der du bruker denne informasjonen på en eller annen måte.*

**Oppgave 2 (følger med en illustrasjon):** *Alle vinklene er like store, og én side i trekanten er 5 cm. Lag så mange oppgaver/spørsmål du kan, som kan knyttes til figuren* (figur 3.2.2.1) *eller teksten, på en eller annen måte.*

**Oppgave 3:** *I går hadde Kåre en bursdagsfest, og det ringte på døra ti ganger. Første gang dørklokka ringte kom det én gjest. Hver gang dørklokka ringte etter det, kom det tre gjester mer enn det gjorde forrige gang dørklokka ringte.*

*1- Hvor mange gjester kom det da dørklokka ringte for tiende gang? Forklar hvordan du fant svaret ditt?*

*2- Lag så mange oppgaver/spørsmål som du klarer, som kan knyttes til situasjonen over, på en eller annen måte.*

***Figur 2****: 3.4 Figur til semi-strukturert oppgave problemgenerering*

For å prøve å få elevene til å lage så vanskelige oppgaver som mulig, oppfordret vi til å lage oppgaver som de trodde vennene, læreren eller vi ikke kom til å klare. Samtidig oppfordret vi til at de skulle lage så mange oppgaver som de klarte på den disponible tiden. En litt annerledes tilnærming enn Van Harpen og Sriraman (2013, s. 208) gjennomførte, hvor de dro frem et fiktivt scenario, samt at den gruppen som genererte flest og/eller best oppgaver fikk belønning. Verdt å nevne er at de ulike kategoriene ikke alltid er godt nok til å plassere et problem klart innenfor en kategori, slik at noen elementer av en oppgave kan tilhøre forskjellige kategorier. Dog kan noen oppgaver gå like mye i to, eller flere, kategorier og havner derfor i kategorien «sammensatte oppgaver».

Da vi diskuterte hvilken type hint vi skulle velge kom vi fram til at vi skulle bruke eksempeloppgaver. Valget falt på eksempel til løsning av oppgaven siden det er noe de aller fleste elever i norsk skole er kjent med, ettersom alle lærebøker bruker «worked example» for å vise løsningsmetoder. Videre måtte vi velge hvilke typer eksempel vi skulle gi elevene. Vi endte med to alternativer- et åpenbart eksempel, eller et vi kunne anta var originalt. Eksempelvis for oppgave 1 ville et åpenbart eksempel vært «hvor mange stod det i køa totalt?». Etter diskusjon rundt dette valgte vi å gjøre bruk av det originale eksemplet. Dette gjorde vi for å vise elevene i eksperimentgruppen hvordan en kunne løse oppgaven. Gjennom å gi et åpenbart eksempel ville vi kanskje peilet eksperimentgruppen til å produsere flere trivielle oppgaver. Det er jo nødvendigvis ønskelig at eksemplet skal være til hjelp for elevene, slik at de kan produsere best mulig oppgaver.

## 3.5 Datainnsamling

I forkant av studien fordelte vi elever i to grupper, hvorpå eksperimentgruppen fikk et eksempel til hver oppgave, og kontrollgruppen bare fikk oppgavene. Gruppene ble fordelt ut fra måloppnåelse i matematikkfaget, ettersom vi ønsket å ha noenlunde like mange av hver måloppnåelse i hver gruppe. Studien tok for seg randomisering når det kom til utvalg, slik at det skulle være så valid som mulig.
Datainnsamlingen i de to forskjellige klassene ble foretatt så nærme som mulig hverandre, for å unngå at den ene eller andre klassen fikk mer tid til å jobbe med faget. Det var viktig for oss at testene kunne gjennomføres i løpet av en time, slik at vi skulle oppta minst mulig av skoledagen til elevene.

Innhentingen av data foregikk ved at hver enkelt elev ble tildelt et hefte med til sammen 7 ark, hvorpå forsiden hadde overskrift og deltakernavn. Her har vi i forkant ført inn navnene til elevene, for å unngå å bruke tid på dette i selve gjennomføringen. Elevene fikk også helt like hefter, i samme farge med samme forside (sett bort fra navnet på heftet), for å unngå oppmerksomhet rundt hvorvidt de har like eller ulike tester. Det ble opplyst at elevene fikk 45 minutter til disposisjon, og det ble informert om gjenstående tid etter 30 og 40 minutter. Elevene fikk tildelt 3 sider med oppgaver, samt rutenett under teksten (se vedlagte oppgaveark). Til hvert enkelt oppgaveark fikk elevene også tildelt et ark som var dekt av ruter. Her var intensjonen at elevene skulle føre inn oppgavene de genererte, men dog ble det aldri spesifisert at de måtte føre det inn der.

Like før vi startet selve testen forsøkte vi å motivere elevene til å gjøre det best mulig. Vi valgte å gjøre en egen vri på Van Harpen Van Harpen og Sriraman (2013, s. 208) sin motivasjonsbit, ettersom vi mente det var mer passende for våre elever. For å motivere ga vi de utfordringen «prøv å lag noen oppgaver som er så vanskelige at verken Ole, Fredrik, eller mattelærer i klassen klarer å løse den».

Elevene fikk spesifisert at de ikke fikk kommunisere med sidekamerater, eller andre, og at mobiltelefon og kalkulator ikke kunne brukes. Videre ble det også poengtert at de kunne bruke blyant/pen og linjal om de ønsket det. Elevene ble ikke informert om at de ikke kunne bruke noen andre form for hjelpemidler, ettersom vi ikke anså det som et problem, samt hvis en elev ønsket å bruke noe annet for å gjennomføre testen var de velkomne til dette. De fikk også opplyst at de kunne stille spørsmål om de var usikre på noe. For å kunne sikre at vi hadde en lik tilnærming til de enkelte elevene, bestemte vi oss for at Ole skulle besvare alle henvendelser, på begge datainnsamlingsdagene. Elevenes spørsmål ble nesten eksklusivt besvart med «det kan jeg ikke svare på.», «prøv å les oppgaven på nytt» og ordrette gjentakelser av oppgavene. Dette gjorde vi for å unngå å hjelpe elevene på noe vis, for å sikre at vi beholdt variablene for hver enkelt elev så lik som mulig. Ved kommentarer på at elevene var ferdige med å gjennomføre testen, før tiden var over, ble det spurt om de var sikre. Ved et definitivt ja tok vi inn heftet, og ba eleven om å hente seg en bok å lese i, dette for å ikke forstyrre andre. Hvis eleven antydet at han/hun ikke var helt sikker på om han/hun var ferdig, prøvde vi å motivere eleven med å si «kan du muligens lage noen oppgaver til?».

Etter gjennomføringen noterte vi ned elevenes deltakernummer på heftet, for å kunne føre dette inn med hensyn til elevenes anonymitet. Elevenes deltakernummer og navn ble ikke lagret på samme plass, slik at det ikke skal være mulig å spore opp en spesifikk elev ut fra deltakernummer. Elevenes genererte oppgaver ble ført inn i egne Word-dokumenter, hvor deltakernummer også medfulgte, for å kunne holde oversikt over antall oppgaver hver enkelt elev hadde generert. Samt for å kunne bruke til videre analyse.

I etterkant av studiet skal elevsvarene destrueres, i henhold til søknaden sendt til NSD, og dokumentene skal slettes. Elevenes kjønn, og andre karakteristika, er ikke innhentet, ettersom dette ikke er relevant for studien. Dette for å bevare elevenes anonymitet.

## 3.6 Analysemetode

For å analysere elevsvarene skrev vi hvert enkelt svar inn i et Word-dokument, samt fordelte poeng skjematisk i et Excel-dokument for å kunne holde en viss oversikt. Hvert enkelt svar ble satt hos tilhørende elevnummer, både i Word og i Excel. I tillegg har også elevenes måloppnåelse blitt notert ned, og antall problemer er analysert ut fra dette.

Svarene ble skrevet inn i et Word-dokument hvor hver elevs genererte problemer, ble lagt under elevens delegerte tall. Antall svar i hver enkel kategori ble ført inn i et Excel-dokument. Vi lagde et Excel-ark for hver av de to skolene. For å få en god oversikt over oppgavene ble alle oppgavene fordelt inn i 6 ulike Word-dokument. Fordeling så slik ut:
Oppgave 1, med eksempel.
Oppgave 1, uten eksempel.
Oppgave 2, med eksempel.
Og så videre.
For ordens skyld; Elevenes svar, med korresponderende elevtall er ikke lagret på samme sted, slik at elevene forblir anonyme.

I likhet med metoden for gjennomføringen av testingen, samsvarer analysemetoden med Van Harpen og Sriraman (2013, s. 209). De bruker uttrykket «non-viable», som direkte oversatt blir «ikke levedyktige», det har vi oversatt til «uløselig»/«ikke løselig». Elevenes genererte problemer ble først sjekket om de var løselige eller ikke. Her er det vår egen mening om hva som er ansett som løselige og ikke løselige oppgaver. Dette har vi først sjekket hver for oss, for så å møtes for å se om vi hadde samsvarende meninger, og videre diskutere oppgavene. Problemene vi anså som uløselige ble notert ned, og tatt ut fra annen analyse, men blir diskutert i seg selv senere i oppgaven. Eksempelvis ble problemet «hvor høy er jeg?» og «10 gutter og 10 jenter er i en kø til vitensenteret, men det er stenger om en halv time. Hvor mange aktiviteter kan de gjøre?» markert som uløselig. Begge nevnte oppgavene kan ikke løses uten tilleggsinformasjon, og den førstnevnte er i tillegg ikke relevant til testen, og derfor ansett som uløselig. Elevenes flytskåre bestemmes ut ifra antall løselige oppgaver generert dividert med totalt antall oppgaver. Oppgavene som ble ansett som løselige, ble videre analysert og gitt en skåre for deres flyt, fleksibilitet og originalitet.

I tillegg til antall løselige oppgaver ble de også analysert for trivialitet. Her tilegnet vi en prosentandel trivielle oppgaver, i likhet med Van Harpen og Sriraman (2013, s. 212). I likhet med løselig/ikke løselig ble trivielle oppgaver bestemt av oss. Oppgaver vi anså som svært enkle for elevene i aldersgruppen, ble determinert som triviell. Dette gjorde vi først hver for oss, for så å diskutere dette etterpå. Eksempelvis ble oppgaver som «det står 10 gutter og 10 jenter i kø, hvor mange er de til sammen?» merket som triviell. Generelt sett ble oppgaver som bare inneholdte enkle subtraksjon og addisjonsutregninger merket som triviell, ettersom vi anser at en elev i 7. klasse, stort sett, enkelt kan gjøre dette i hodet, eller ved bruk av en enkel algoritme. I tillegg ble enkle divisjon og multiplikasjonsoppgaver merket som triviell. Eksempel på dette er 20 dividert med 2, eller 10 multiplisert med 2.

Elevenes svar ble kategorisert ut fra ulike matematiske kategorier, hvor svar vi mente tilhørte i flere kategorier ble satt i kategorien «Sammensatte oppgaver», og oppgaver som ikke passet i noen kategorier ble plassert i «andre». Sistnevnte er i tråd med Van Harpen og Sriraman (2013, s. 210). Kategoriene problemene ble plassert i er «Fire regnearter», «Geometriske beregninger», «Likninger», «Prosent/brøk», «Algebra», «Sammensatte oppgaver» og «Andre». Ettersom elevenes generering av oppgaver/problemer kan plasseres i ulike kategorier, kan det forekomme oppgaver som har elementer som kan tilhøre i andre kategorier også. Altså kan en elevs oppgave omhandle geometri, samtidig som den omhandler de fire regneartene, og derfor ha elementer av to ulike kategorier. Eksempelvis kan en oppgave som etterspør summen av sidelengdene til en figur, omhandle geometri i den forstand at den innebærer bruken og forståelsen av figurens utforming, samt fire regnearter i den forstand av addisjon av sidelengdene. Dette har vi ellers ikke kategorisert som sammensatte oppgaver, ettersom oppgaver innenfor geometri, som oftest, benytter seg av de fire regneartene for å løses. Under kommer eksempler til hver av de ulike kategoriene:

Sammensatte oppgaver er forbeholdt oppgaver som har flere ledd i utregningen, eksempelvis:

> **Ole og Fredrik skulle kjøpe julegaver til familien sin. De bestemte seg for å kjøpe gaver sammen. De skulle kjøpe kaffekopper til Oles 8 familiemedlemmer, og Fredriks 12 familiemedlemmer. En kopp kostet 69,- kroner. Hvis de kjøpte 20 kaffekopper fikk de 17% rabatt, hvor mye måtte hver av de betale?**

Her må løseren av oppgaven først regne ut hvor mye det blir for 20 kopper, for så å trekke fra rabatten. Til slutt må de regne ut hvor mye hver av de skal betale. Dermed benytter oppgaven seg av kategorien fire regnearter og prosent/brøk, og blir kategorisert som en sammensatt oppgave.

Likninger:

> **Elevene stod i kø for å komme inn på sirkus. Totalt hadde de 2500 kroner, hvorav 500 var fra klassekassa. Resterende penger var lommepenger til elevene, og alle hadde like mye. Hvor mye lommepenger hadde hver elev?**

Prosent/brøk:

> **Én femtedel av elevene fikk gå forbi køen. Hvor mange stod igjen?**

Geometriske beregninger:

> **Hvor stort er arealet til trekanten over?**

Algebra:

> **Hvor mange gjester kom det da det ringte på den tusende gangen?**

Oppgaven over må ikke nødvendigvis løses ved bruk av algebra, men det virker lite hensiktsmessig å løse den manuelt med opptelling. Derfor valgte vi å plassere disse oppgavene her, ettersom den åpenbart best løses ved hjelp av et algebraisk uttrykk.

Andre:

**Hvordan kan du finne ut hvor mange som kom totalt?**

Selv om løsningen til denne oppgaven kan være algebra/likninger/fire regnearter, så kategoriserer vi denne som andre, siden selve løsningen ikke kan klart plasseres innenfor én kategori.

Fire regnearter:

**Alle som stod i køen hadde 73 kroner, hvor mye hadde de totalt?**

I likhet med Van Harpen og Sriraman (2013, s. 210) ble elevenes totale antall løselige problemer definert som elevens flytskåre. Antall kategorier eleven klarte å generere løselige problemer til, ble elevens fleksibilitetsskåre.

Originaliteten for hvert enkelt problem ble determinert ut fra hvor sjeldent svaret var. Våre deltakere befinner seg innenfor samme opplæringsplan, og innenfor samme kommune, samt at kontroll- og eksperimentgruppen er blandet på to byskoler. Her blir problemenes originalitet determinert ut fra gruppen de befinner seg i, ikke byskolen de går på. Dette er ulikt Van Harpen og Sriraman (2013, s. 210) som determinerer originalitet ut fra klassen elevene befinner seg i, og samtidig likt ettersom det blir innad i gruppene som sammenliknes med hverandre. De analyserte ikke originaliteten på tvers av gruppene ettersom det er relativt innad i gruppen, noe vi gjør ettersom gruppene er blandet fra to byskoler. Ved å anse problemer som er generert av 10% eller færre elever, blir problemet definert som originalt. I vårt tilfelle vil det si problemer som er generert av én elev, ikke flere, er originale. Dette er også i tråd med hva Van Harpen og Sriraman (2013, s. 211) definerte som originale, dog er gruppestørrelsene svært ulike. Grunnet dette så kan en argumentere for at det burde vært en annen prosentandel, ettersom besvarelsen til én elev blir en relativt høy prosentandel. Igjen i likhet med dem, ble problemer som ble ansett som trivielle, ikke ansett som originale selv om under 10% av elevene genererte dette problemet. Eksempelvis blir det genererte problemet «hva er 10 + 10?», ikke determinert som originalt, ettersom matematikken er på et relativt lavt nivå i forhold til elevenes alder (Van Harpen & Sriraman, 2013, s. 211). Når vi bedømte hvorvidt et problem var originalt, var det vår egen dømmekraft som ble satt til grunn for om det var originalt eller ikke. Hvis to elevsvar var svært strukturelt like, så ble de bedømt uoriginale. Her valgte vi å kun legge vekt på strukturen i oppgaven, ettersom det er strukturen som bestemmer løsningsstrategi (Bassok, 2003, s. 343). Eksempelvis ble oppgavene under ikke definert som original, ettersom de hadde store strukturelle likheter:

> **10 personer sykler til bursdagen, 12 ble kjørt og resten gikk. Hvor mange gikk?**
>
> **10 gjester tok av seg skoene og resten gjorde ikke det. Hvor mange hadde skoene på?**

Til tross for at den første oppgaven har flere momenter, i at den ber løseren om å trekke to tall fra totalt antall gjester, merket vi disse som uoriginal. Oppgavene over dreier seg om totalt antall gjester, minus et gitt antall gjester. Verdt å merke seg er at selv om disse problemene i utgangspunktet ser trivielle ut, er de ikke det, siden løseren først må løse det relativt sett komplekse problemet «hvor mange gjester kom totalt?».

Elevsvarene våre ble først kategorisert individuelt av hver enkelt av forfatterne, så sammenliknet. Etter dette ble svarene kategorisert og diskutert sammen, og i samsvar med hverandre Van Harpen og Sriraman (2013, s. 211).

### 3.6.1 Validitet

Kvantitativ forskning bør forsøke å betrakte en rekke punkter for å styrke validiteten. Herunder kan en nevne kontrollerbarhet, mulighet for å replisere forskningen, randomisering av utvalg og objektivitet med flere (Cohen et al., 2018, s. 246-247).
En skiller mellom indre og ytre validitet. For vår del vil det si at viktigheten av et nøye beskrevet metodekapittel er stor. Ved å ta utgangspunkt i Van Harpen og Sriraman (2013) sin forskning, kan vi bruke en tidligere utprøvd metode, designet, beskrevet og kvalitetssikret av forskere. Fortsatt er det viktig at tilpasningene vi gjør blir beskrevet nøye, slik at dette kan etterprøves så nøyaktig som mulig.

#### 3.6.1.1 Indre validitet

Cohen et al. (2018) beskriver en rekke trusler mot en studies indre validitet. Under skal vi prøve å betrakte disse.

Forskerens objektivitet, eller mangel på objektivitet, kan være en trussel for en studies validitet (Cohen et al., 2018, s. 247). For å forsøke å oppnå større objektivitet har vi separat fra hverandre analysert alle oppgavene, og diskutert disse sammen etterpå. På denne måten minsker vi sannsynligheten for at en enslig persons mening blir lagt til grunne for kategoriseringen. I tillegg til dette har vi etter cirka en måned på nytt gjennomgått rådata, for å se om vi ville kategorisere de enkelte svarene fra rådataen på en annen måte. I tillegg til dette hadde vi ikke navnet til elevene på testresultatene, for å unngå eventuell forskjellsbehandling.
Ettersom vi ikke har sett noen forskning som tar for seg problemgenerering med- og uten eksempeloppgaver, hadde vi ingen formening om hvordan utfallet ville bli, ei heller noen forhåpninger.

Utvalget vi har er som beskrevet i 3.4 utvalg basert på *convenience sample*. De elevene vi valgte ut ble ikke valgt av noen annen årsak enn at det var de informantene vi hadde tilgjengelig. Videre hadde vi ikke kjennskap til noen av elevene i utvalget på forhånd, og visste dermed ikke noe om hvordan måloppnåelse elevene hadde. Ei heller om elevene hadde kjennskap til problemløsning eller problemgenerering fra tidligere. Da vi randomiserte utvalget delte vi de to ulike klassene i to, slik at vi fikk noen elever fra hver skole i de to ulike gruppene (eksperiment og kontrollgruppene). Vi fordelte de videre ut fra måloppnåelse, slik at det ikke skulle bli en skjevfordeling av høyt presterende/lavt presterende elever i noen av gruppene. Ved å fordele på denne måten forsøkte vi å eliminere feilkilder knyttet til kompetanseskjevheter i gruppene. På den andre siden var det flere elever som var syke på datainnsamlingsdagene. Dermed ble ikke gruppene helt som tiltenkt, og det ble derfor forskjeller i gruppestørrelsene og kompetansenivået. Dette var utenfor vår kontroll, men fortsatt kan vi ikke si noe om hvordan det ville påvirket datasettet om disse elevene var til stede.

Videre handler indre validitet om hvorvidt en kan si noe om årsak-virkning sammenhengen i studien som er gjort. Under dette er punkter som bias i selektering av utvalg, modning mellom observasjoner og «instrument reactivity». Cohen et al. (2018, s. 252) skriver at «instrument reactivity» er hvordan et individ reagerer i samhandling med målingssystemet, være seg test, video, intervju og så videre. En elev som vet at handlingene de gjør, eller samtalen de har blir filmet, tatt opp eller på annen måte lagret, kan handle annerledes enn hva eleven normalt gjør. Da vi innledningsvis, i både informasjonsskriv og på datainnsamlingsdagen, formidlet at dette var utenfor normal undervisning og vurdering, kan vi ha bidratt til at enkelte elever handlet på en annen måte enn de ellers ville gjort i en vanlig undervisningssituasjon. Dette kan være en trussel mot forskningens indre validitet, men fortsatt umulig å unngå av etiske hensyn som er nødvendig å ta. Mange punkter fra Cohen et al. (2018, s. 252-253) dreier seg om endringer som spontant skjer mellom observasjoner, slik at alle disse er eliminert ved at vi ikke har før og etter-test, ettersom vi kun utførte én test med kontroll og eksperimentgruppen.

#### 3.6.1.2 Ekstern validitet

Ekstern validitet dreier seg om hvorvidt resultatene er overførbare til et annet utvalg, eller en populasjon (Cohen et al., 2018, s. 254). I tillegg til dette handler det også om generaliseringer til lignende situasjoner, uten å veie opp kontekstuelle forskjeller- eksempelvis vil det i vårt tilfelle være en test-situasjon sammenlignet med en klasseromssituasjon, hvor læreren i klassen gir eksempeloppgave til oppgaven elevene skal gjøre. Situasjonen er naturligvis forskjellig fra en vanlig matematikktime med elevenes egen matematikklærer, slik at dette kan påvirke resultater.

Cohen et al. (2018, s. 254) beskriver en rekke momenter som kan være trusler for en studies eksterne validitet. Blant annet vil det å generalisere ut fra unntak være en trussel mot validiteten. For å hindre dette har vi synliggjort tabellen for hele klassen, slik at eventuelle unntak kommer frem. Videre er det viktig å forholde seg til samme metode for begge gruppene. Alle variabler blir her viktige. Det vi sier innledningsvis til elevene, hvordan de sitter, hvilken tid på dagen de utfører testen osv. Dette er i stor grad ting vi hadde mulighet til å kontrollere, slik at både eksperimentgruppen og kontrollgruppen fikk samme betingelser og miljø for gjennomføring. I tillegg var eksperiment og kontrollgruppen fordelt ut på de to skolene, slik at miljø, tidligere undervisning, og så videre skal påvirke resultatene i minst mulig grad.

På generelt grunnlag vil viktigheten av et nøye metodekapittel være stor. Det er viktig at de individuelle variablene som gjelder for vårt utvalg skal komme frem så godt som mulig.

Ettersom vi ønsket at gruppene skulle være helt vilkårlig trukket, måtte dette forberedes på forhånd. Naturligvis var det umulig å vite hvem som var borte på gjennomføringsdagen. For å bruke minst mulig av elevenes tid på gjennomføringen ble testene utført med de originale gruppene, uten hensyn til skjevheter som måtte forekomme grunnet elevfravær. Dette er noe som åpenbart kan argumenteres at er en svakhet i oppgaven.

### 3.6.2 Reliabilitet

Måling vil alltid kunne innebære en viss feilmargin. Det er spesielt vanskelig å fullstendig eliminere feilmarginer når en foretar målinger på abstrakte fenomener ved mennesker, ettersom mennesket konstant er i utvikling (Postholm, 2018, s. 224). En skiller mellom reliabilitet som stabilitet, reliabilitet som ekvivalens og reliabilitet som intern konsistens (Cohen, 2018, s. 268-269). Det er en rekke trusler til en studies reliabilitet, og alle aspekter med mennesket spiller en rolle. Elevenes motivasjon, hvilken tid på dagen de gjennomfører testen, personlighet, miljø, og så videre kan ha betydning for resultatet (Cohen, 2018, s. 279-280). Dette er noe som er gjennomgående for alle studier som undersøker atferd eller læring, og en må dermed være kritisk når en tolker resultater og konklusjoner.

#### 3.6.2.1 Reliabilitet som stabilitet

Cohen et al. (2018, s. 268) skriver om reliabilitet som stabilitet. For å se på en studies reliabilitet kan en se på resultatene den gir over tid. Det vil si å utføre testen flere ganger i en test - retest situasjon. En annen måte å vurdere en studies stabilitet på er å utføre samme test på en annen gruppe mennesker, som er like (Cohen et al., 2018, s. 268). Her må alle faktorer som kan spille inn tas høyde for. For vår del vil det være alder, mengde undervisning i det aktuelle temaet og måloppnåelse som er de viktigste faktorene. En kan ikke gå ut ifra at det ikke er flere faktorer som vil spille inn på resultatene, men naturligvis vil det være vanskeligere å finne et lignende utvalg jo flere faktorer en betrakter. Begge måtene å teste studiens reliabilitet på var vanskelig for oss å gjennomføre, ettersom vi hadde begrenset tilgang til informantene. Videre hadde vi ikke anledning til å rekruttere like mange informanter som samtidig var like våre opprinnelige informanter, for å sjekke stabilitet over likt utvalg. Dette blir derfor følgelig en svakhet i vår studies reliabilitet.

#### 3.6.2.2 Reliabilitet som ekvivalens

Ved vår studie vil reliabilitet som inter-rater reliabilitet være aktuell. Inter-rater reliabilitet dreier seg om enigheten mellom forskere (Cohen et al., 2018, s. 269). Vi er to forskere, som skal bedømme det samme datasettet. Alle svarene gitt på testene skal kategoriseres, for vår del innenfor enten fire regnearter, likninger, geometri, prosent/brøk, algebra, andre og sammensatte oppgaver. I tillegg skal de vurderes til triviell/ikke triviell, samt løselig/uløselig. Derfor vil menneskelige faktorer være av betydning for resultatene. Cohen et al. (2018, s. 279) nevner flere aspekter som kan være en trussel for en kvantitativ studies reliabilitet. En kan gjøre feil når en kategoriserer, ulike forskere kan bedømme ulikt, og en kan være lite konsekvent i bedømmelsen. Eksempelvis kan en være veldig streng i starten, for så å løsne opp etter hvert, når en tar for seg mange tester på et og samme tidsrom. For å håndtere disse truslene til reliabiliteten, bedømte vi oppgavene elevene leverte hver for oss, for så å sammenlikne resultatene, og diskutere svarene. Videre bedømte vi også prøvene etter om lag en måned senere, for å igjen se om vi fant uoverensstemmelser, eller svar vi ønsket å endre kategorisering på.

Ved sammenlikning av våre resultater fant vi 4 uoverensstemmelser. Det vil si at vi var enige ved 163 av 167 mulige tilfeller. Ved å regne ut inter-rater enigheten som en prosent får vi $\frac{163}{167} = 0{,}976$, altså en enighetskåre på 97,6%. Da vi diskuterte disse uenighetene kom vi frem til at vi i utgangspunktet var nokså enige. Disse uoverensstemmelsene skyldtes at vi vurderte alle prøvene på en gang, og at det oppstod en viss tretthet i aktiviteten, som førte til feilvurdering av svarene.

#### 3.6.2.3 Reliabilitet som intern konsistens

I vårt prosjekt henter vi data ved å kun gjennomføre én test. For å teste reliabiliteten til dataen vi henter kan vi dermed ikke sjekke test / re-test opp mot hverandre. Derfor måler vi testens reliabilitet som intern konsistens med en Cronbachs alpha verdi (Cohen, 2018, s. 774). Gjennom å sjekke Cronbachs alpha-verdien finner vi en verdi på 0,801, som tilsier at det er rimelig å anta at den indre konsistensen er god. Ved å sjekke om det er enkelte verdier som trekker verdien opp eller ned, finner vi laveste Cronbachs alpha på 0,748, og den høyeste verdien på 0,811. Dette tilsier at det er få momenter med datasettet som påvirker verdien betydelig opp eller ned.

## 3.7 Forskningsetikk

I forkant av studien sendte vi en søknad til Norsk senter for forskningsdata (NSD), dette for å få gjennomføre studien med den tiltenkte hensikten. Ettersom studiens deltakere er barn, ble det innhentet samtykke fra foresatte for å kunne gjennomføre studien. Herunder ble det understreket at elevene, og foresatte, når som helst hadde muligheten til å trekke eleven ut av studien om ønskelig. Dette også uten noen form for begrunnelse eller avklaring. Grunnet vår studie som foregår tett på mennesker, har vi tatt etiske betraktninger med utgangspunkt i retningslinjene til *Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH).* Herunder er del B «hensyn til personer» (punkt 15 til 30) særs sentral. Her går vi kort innpå noen av punktene som er relevante for studien, men ikke alle ettersom det er begrenset med plass. Punkt 15 omhandler samtykke til å delta i forskning, hvor samtykket er frivillig, informert og utvetydig, samt dokumenterbart. Punkt nummer 16 innebærer redusert samtykkekompetanse, og er relevant ettersom studien foretas på barn. Nummer 17 omhandler beskyttelse av barn, og understreker at samtykke må som hovedregel innhentes fra både foresatte og barna selv (NESH, 2021, s. 17). Her har vi, som sagt, innhentet samtykke fra foresatt, samt informert elevene om at deres deltakelse var helt frivillig og at de kunne trekke seg om de ønsket det. Etter testen var gjennomført fikk elevene beskjed om hva som hadde blitt gjort, og hvordan dokumentene kom til å bli oppbevart. Denne innebar også tydeliggjøring av at de kom til å være anonyme i oppgaven, og at elevenes navn og deltakernummer ikke på noen måte skal kunne bli koblet sammen. Elevenes besvarelser ble anonymisert, og identiteten til elevene skal ikke være mulig å innhente gjennom å lese studien vår. Dette er i tråd med NESH (2021, s. 20-21) punkter 19 og 20. Dataen som er innhentet, både resultater, deltakernavn og deltakernummer, er lagret konfidensielt ved bruk av to-faktorsautorisering. Herunder er deltakernavn lagret en plass, og deltakernummer en annen plass, slik at det ikke kan kobles til deltakerne. I tråd med punkt 24 hos NESH (2021, s. 23) skal elevenes deltakelse, besvarelser og deltakernummer destrueres i etterkant av studien og innlevert oppgave, som det også ble opplyst om i søknaden til NSD.

Det at noen elever fikk eksempel til oppgavene som skulle løses, samt at studien innehar oppgaver som er brukt på deltakere av høyere alder (og har derfor kanskje mer avansert og kanskje mer inngående kompetanse i matematikk), kan være en potensiell innvending mot studien. Her kan det argumenteres for at det er gitt «for vanskelige» oppgaver til elevene, og derfor ikke er forsvarlig i forskningsøyemed. Studien viser ikke tydelig om elevene har slitt med å gjennomføre oppgavene, eller om de har hatt noen innvendinger mot gjennomførelsen av dette. Det er etter vårt syn oppgaver som kan brukes i ulike aldre, ettersom det ikke skal vise noe direkte kompetanse i studien i form av korrekt/ikke korrekt besvarelse, samt en oppgave som er blitt tilpasset elevgruppen, og derfor ikke uforsvarlig å presentere dette til gruppen.

# 4 Resultater

Under forekommer resultatene til studien som er gjennomført. Vi vil fremme de fleste aspekter ved datasettet. Herunder flyt, fleksibilitet og originalitetskåre, gjennomsnittlig antall trivielle og uløselige oppgaver, samt hvordan svarene fordeler seg på oppgave 1, 2 og 3. Elevene er hovedsakelig delt inn i eksperimentgruppe og kontrollgruppe, men under disse deler vi de videre inn i lav, middels og høy-måloppnåelse. Vi vil bruke median hvor det virker hensiktsmessig for å utdype enkelte tall, hvor for eksempel én elev har svært mye høyere skåre enn resten av gruppen. Til slutt sammenligner vi de to gruppene mot hverandre.

## 4.1 Beregning av elevenes skåre

Elevenes genererte oppgaver ble analysert for flyt, fleksibilitet, originalitet, antall uløselige problemer og antall trivielle problemer. Her forekomer beregningen på følgende måte.

Elevenes flytskåre ble determinert ut fra antall løselige problemer som ble generert. Her er elevenes trivielle svar også presentert med, mens uløselige oppgaver er tatt ut. Fleksibilitetsskåre ble determinert ut fra antall kategorier elevenes svar forekom i, denne skåren kan skille seg noe fra flytskåren, ettersom elevene kan presentere flere problemer som faller inn under samme kategori. Antall originale svar ble determinert ut fra hvor unike de var i henhold til gruppen sin. Hvis et problem ble generert av 10% eller færre, ble det ansett som originalt. Antall uløselige problemer, ble determinert ut fra hvor mange problemer som ikke kunne løses innad i gruppen. Dette tallet kan en se i lys av totalt antall genererte problemer, minus antall løselige problemer. Antall trivielle problemer ble fastslått ut fra hvor mange problemer som ble ansett som for enkle for elevenes nivå. Her forekommer det en subjektiv vurdering tatt av forfatterne, som først ble vurdert individuelt, så diskutert sammen for å determinere om et problem var trivielt eller ikke. Dette gjelder for samtlige beregninger. Her ble det foretatt gjennomsnitt og median innenfor hver gruppe med henhold til måloppnåelse, og hele gruppens gjennomsnitt og median. Dette for å fremstille skåren på en mest hensiktsmessig måte. Alle beregningene er gjort likt med Van Harpen og Sriraman (2013, s. 210-215) sine analyser, ettersom vår studie baserer seg i stor grad på deres.

## 4.2 Gruppe med eksempel

Tabell 4.2.1 viser elevgruppen og antall svar innenfor hver kategori. Det forekommer også antall svar som ikke er løselige, antall svar som er trivielle og fleksibiliteten til hver enkelt elev.

***Tabell 4.2 1**: Sum svar - elever med eksempel. H = Høy, M = Middels, L = Lav*

| Elevnr. | 15 | 16 | 17 | 29 | 5 | 6 | 7 | 34 | 21 | 24 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Måloppnåelse | H | H | H | H | M | M | M | M | L | L | L |
| Fire regnearter | 1 | 13 | 3 | 3 | 3 | 0 | 4 | 4 | 0 | 1 | 3 |
| Sammensatte oppgaver | 2 | 0 | 9 | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| Geometriske beregninger | 3 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Likninger | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Prosent/brøk | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 |
| Algebra | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Andre | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| Sum | 9 | 15 | 15 | 5 | 4 | 0 | 4 | 10 | 0 | 2 | 3 |
| Uløselig | 1 | 0 | 2 | 0 | 1 | 2 | 0 | 6 | 0 | 0 | 1 |
| Triviell | 1 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |
| Fleksibilitet | 6 | 2 | 5 | 3 | 2 | 0 | 1 | 4 | 0 | 2 | 2 |
| Original | 4 | 8 | 11 | 1 | 2 | 0 | 0 | 2 | 0 | 0 | 1 |

Tabellen viser antall løselige, ikke-løselige og trivielle svar for elevgruppen med eksempel. Her vises det også til antall svar innenfor hver kategori. Antall svar innenfor hver kategori er noe sprikende, med et klart høyere antall på «fire regnearter». Elevene med høy måloppnåelse i matematikk stod for 44 av 67 av alle svar, klart flere enn elevene med lav og middels, som stod for de gjenværende 23 svarene. Elevene med høy måloppnåelse har flest oppgaver som er trivielle, men prosentvis har de fortsatt lavere antall trivielle svar, med 11% trivielle, mot 13 % trivielle svar for middels og lav måloppnåelse. Elevene med høy måloppnåelse produserte 3 uløselige svar, elevene med middels og lav måloppnåelse produserte 10.

Samlet for gruppen kan en se at de produserte 67 løselige svar, hvorav 8 var trivielle som tilsvarer 11%. 13 av de totalt 80 svarene var uløselige, som tilsvarer 16%.

***Tabell 4.2 2**: Sum gjennomsnitt flytskåre per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Flyt | Median |
|---|---|---|---|
| Høy | 4 | 11 | 12 |
| Middels | 4 | 4,5 | 4 |
| Lav | 3 | 1,667 | 2 |
| Samlet | 11 | 6,090 | 4 |

I tabell 4.2.2 ser vi gjennomsnittlig flytskåre fordelt på måloppnåelse, og gjennomsnittet for gruppen samlet. En kan se at det er en synkende trend fra høy måloppnåelse til middels og lav måloppnåelse. Snittet for hele gruppen får en noe høyere gjennomsnittlig flytskåre enn middels måloppnåelse, grunnet en relativt sett, svært mye høyere skåre for elevene med høy måloppnåelse enn lav og middels.

***Tabell 4.2 3**: Sum gjennomsnitt uløselige oppgaver per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Uløselig | Median |
|---|---|---|---|
| Høy | 4 | 0,75 | 0,5 |
| Middels | 4 | 2,25 | 1,5 |
| Lav | 3 | 0,333 | 0 |
| Samlet | 11 | 1,182 | 1 |

Elevene med høy måloppnåelse hadde i gjennomsnitt færre uløselige oppgaver enn elevene med middels måloppnåelse. Elevene med lav måloppnåelse hadde færrest uløselige, med bare 1 av totalt 6 oppgaver som var uløselig. Sammenliknet med gjennomsnitt for elevene fordelt på måloppnåelse ser vi at snittet for alle elevene er høyere enn for elevene med høy og lav måloppnåelse, og betydelig lavere sammenliknet med middels, som hadde desidert flest uløselige oppgaver.

***Tabell 4.2 4**: Sum gjennomsnitt trivielle oppgaver per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Triviell | Median |
|---|---|---|---|
| Høy | 4 | 1,25 | 1,5 |
| Middels | 4 | 0,5 | 0 |
| Lav | 3 | 0,333 | 0 |
| Samlet | 11 | 0,727 | 0 |

Tabell 4.2 4 viser gjennomsnittet av antall trivielle oppgaver fordelt på måloppnåelse. En kan se at antall trivielle oppgaver går ned med synkende måloppnåelse. Elever med høy måloppnåelse ligger over gjennomsnittet, mens middels og lav måloppnåelse ligger under snittet. Tabellen viser at elevene med høy måloppnåelse produserte flest antall trivielle svar i snitt. Dog forteller den ikke hvor mange oppgaver som er produsert totalt. Elevene med høy måloppnåelse produserte 44 oppgaver, hvorav 5 var trivielle, tilsvarende 11%. Elevene med lav og middels produserte til sammen 23 oppgaver, hvorav 3 var trivielle, tilsvarende 13%.

***Tabell 4.2 5**: Sum gjennomsnitt fleksibilitet per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Fleksibilitet | Median |
|---|---|---|---|
| Høy | 4 | 4 | 4 |
| Middels | 4 | 1,75 | 1,5 |
| Lav | 3 | 1,333 | 2 |
| Samlet | 11 | 2,454 | 2 |

Tabell 4.2.5 viser elevenes fleksibilitet. Det kommer frem av tabellen at fleksibiliteten går ned med måloppnåelse. Elevene med høy måloppnåelse scorer over dobbelt så høyt på fleksibilitet sammenliknet med elever med middels, og 3 ganger så høyt som elever med lav måloppnåelse. I gruppen med høy måloppnåelse er det generert klart flere enn snittet for hele gruppen. Gruppene med middels og lav måloppnåelse ligger godt under gjennomsnittet.

*Tabell 4.2 6: Sum gjennomsnitt originale svar per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Originalitet | Median |
|---|---|---|---|
| Høy | 4 | 6 | 6 |
| Middels | 4 | 1 | 1 |
| Lav | 3 | 0,333 | 0 |
| Samlet | 11 | 2,636 | 1 |

Originaliteten til elevene viser en tendens til å synke med synkende måloppnåelse. Gjennomsnittet for gruppen faller mellom elevene med høy og middels måloppnåelse. Både elevene med middels og lav måloppnåelse ligger godt under snittet for gruppen, grunnet elevene med høy måloppnåelses relativt sett betydelig høyere skåre.

*Tabell 4.2 7: Antall svar gitt fordelt på oppgave 1, 2 og 3*

| | | | | |
|---|---|---|---|---|
| Elevnr. | Måloppnåelse | Oppgave 1 | Oppgave 2 | Oppgave 3 |
| 15 | Høy | 3 | 3 | 3 |
| 16 | Høy | 8 | 3 | 4 |
| 17 | Høy | 5 | 5 | 5 |
| 29 | Høy | 4 | 1 | 0 |
| 5 | Middels | 3 | 1 | 0 |
| 6 | Middels | 0 | 0 | 0 |
| 7 | Middels | 2 | 0 | 2 |
| 34 | Middels | 5 | 0 | 5 |
| 21 | Lav | 0 | 0 | 0 |
| 24 | Lav | 1 | 0 | 1 |
| 1 | Lav | 3 | 0 | 0 |
| Sum | | 34 | 13 | 20 |
| Sum høy | | 20 | 12 | 12 |
| Sum middels | | 10 | 1 | 7 |
| Sum lav | | 4 | 0 | 1 |

Fra tabell 4.2.7 kommer det tydelig frem at elevene med høy måloppnåelse står for flesteparten av svarene som er gitt. Sammenlagt på de 3 oppgavene ga elevene med høy måloppnåelse 44 oppgaver, elevene med middels leverte 18 oppgaver, og elevene med lav måloppnåelse leverte 5 oppgaver. Fra 4.1.1 ser en at det gir henholdsvis gir gjennomsnitt per elev 11, 4,5 og 1,67 oppgaver per elev. I tabellen over kan en se at det er klart flest svar på den åpne oppgaven, oppgave 1. Elevene i eksperimentgruppen leverer færrest på oppgave 2, som er den semistrukturerte oppgaven.
Verdt å merke her er at de uløselige oppgavene ikke kommer frem i denne statistikken, dette er noe som tas frem og drøftes i diskusjonen.

## 4.3 Gruppe uten eksempel

Tabell 4.3.1 viser elevgruppen og antall svar innenfor hver kategori. Det forekommer også antall svar som ikke er løselige, antall svar som er trivielle og fleksibiliteten til hver enkelt elev.

*Tabell 4.3. 1: Sum svar for elevene uten eksempel. H = Høy, M = Middels, L = Lav*

| Elevnr. | 19 | 27 | 30 | 8 | 9 | 11 | 12 | 14 | 31 | 33 | 35 | 22 | 23 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Måloppnåelse | H | H | H | M | M | M | M | M | M | M | M | L | L | L |
| Fire regnearter | 3 | 7 | 0 | 4 | 2 | 4 | 0 | 1 | 3 | 5 | 0 | 0 | 2 | 6 |
| Sammensatte oppgaver | 0 | 3 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Geometriske beregninger | 1 | 2 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 1 |
| Likninger | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Prosent/brøk | 1 | 2 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| Algebra | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Andre | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| Sum | 6 | 16 | 3 | 5 | 11 | 5 | 0 | 1 | 3 | 6 | 9 | 0 | 2 | 8 |
| Uløselig | 0 | 1 | 1 | 0 | 0 | 4 | 2 | 0 | 0 | 0 | 1 | 2 | 0 | 1 |
| Triviell | 2 | 3 | 1 | 2 | 3 | 3 | 0 | 0 | 3 | 3 | 7 | 0 | 2 | 6 |
| Fleksibilitet | 4 | 6 | 2 | 2 | 5 | 2 | 0 | 1 | 1 | 2 | 3 | 0 | 1 | 3 |
| Original | 2 | 7 | 0 | 1 | 5 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

Tabell 4.3.1 viser antall løselige, ikke-løselige og trivielle svar for elevgruppen uten eksempel. Her vises det også til antall svar innenfor hver kategori. Antall svar innenfor hver kategori er noe sprikende, med et klart høyere antall på «fire regnearter». Elevene med høy måloppnåelse i matematikk hadde 25 svar, som igjen er klart flere enn elevene med middels og lav måloppnåelse, som hadde henholdsvis 5 og 3 svar. En ser at elevene med lav måloppnåelse har klart flere svar som er ikke-løselige og trivielle, henholdsvis 3 og 8. Samlet sett viser tabellen at elevene uten eksempel hadde 11 svar som ikke var løselige og 28 svar som var trivielle, samt at fleksibiliteten er høyere blant elever med høy måloppnåelse.

*__Tabell 4.3. 2__: Sum gjennomsnitt flytskåre per elev fordelt på måloppnåelse*

| | Antall | Gj. Flytskåre | Median |
|---|---|---|---|
| Høy | 3 | 8,333 | 6 |
| Middels | 8 | 5 | 5 |
| Lav | 3 | 3,333 | 2 |
| Samlet | 14 | 5,357 | 5,5 |

I tabell 4.3.2 ser vi gjennomsnittlig flytskåre per elev. Tabellen viser at elever med høy måloppnåelse har en høyere flytskåre, sammenliknet med middels og lav måloppnåelse.
Vi kan se at elevene med høy og lav måloppnåelse er noenlunde utenfor gjennomsnittet, henholdsvis over og under, mens elevene med middels måloppnåelse er rett under snittet for utvalget. Ut ifra tabellen ser vi at flytskåren synker sammen med måloppnåelse.

*__Tabell 4.3. 3__: Sum gjennomsnitt uløselige oppgaver per elev fordelt på måloppnåelse*

| | Antall | Gj. Uløselig | Median |
|---|---|---|---|
| Høy | 3 | 0,667 | 1 |
| Middels | 8 | 0,875 | 0 |
| Lav | 3 | 1 | 1 |
| Samlet | 14 | 0,857 | 0,5 |

Fra tabell 4.3.3 kan vi se en synkende trend. Elevene med høy måloppnåelse leverte færrest, og under gjennomsnittet. På den andre siden leverte elevene med middels og lav måloppnåelse over gjennomsnitt, dog var bare elevene med middels måloppnåelse marginalt over gjennomsnittet. Verdt å merke seg er medianen for middels måloppnåelse, som er 0. Fra tabell 4.1.1 kan en se at én elev står for 4 av 7 uløselige oppgaver, slik at snittet dras betydelig opp. Dette følger videre for medianen for gruppen samlet, som er en del lavere enn gjennomsnittet.

***Tabell 4.3. 4**: Sum gjennomsnitt trivielle oppgaver per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Triviell | Median |
|---|---|---|---|
| Høy | 3 | 2 | 2 |
| Middels | 8 | 2,625 | 3 |
| Lav | 3 | 2,667 | 2 |
| Samlet | 14 | 2,5 | 2,5 |

For denne kategorien kan en se en klar forskjell fra høy til middels og lav måloppnåelse. Elevene med høy måloppnåelse ligger under snittet for gruppen, mens elevene med lav og middels måloppnåelse har tilnærmet likt gjennomsnitt, rett over snittet for hele gruppen.

***Tabell 4.3. 5**: Sum gjennomsnitt fleksibilitetsskåre per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Fleksibilitet | Median |
|---|---|---|---|
| Høy | 3 | 4 | 4 |
| Middels | 8 | 2 | 2 |
| Lav | 3 | 1,333 | 1 |
| Samlet | 14 | 2,071 | 2 |

Over kan en se fra tabell 4.3.5 at elevene med høy måloppnåelse har en betydelig høyere fleksibilitetskåre enn resten av elevene. Sammenlignet med middels og lav måloppnåelse har de henholdsvis dobbelt, og tre ganger så høy fleksibilitetskåre. I forhold til gjennomsnittet til hele gruppen skårer elevene med høy måloppnåelse dobbelt så høyt. Elevene med lav og middels skårer henholdsvis tilnærmet på gjennomsnittet, og godt under gjennomsnittet, med samsvarende median. En ser en tydelig trend på synkende fleksibilitet ved synkende måloppnåelse i faget.

***Tabell 4.3. 6****: Sum gjennomsnitt originale svar per elev fordelt på måloppnåelse*

| | Antall elever | Gj. Originale svar | Median |
|---|---|---|---|
| Høy | 3 | 3 | 2 |
| Middels | 8 | 1 | 0,5 |
| Lav | 3 | 0 | 0 |
| Samlet | 14 | 1,214 | 0 |

Antall originale svar er tydelig høyere hos elevene med høy måloppnåelse. Fra høy til middels og lav er det henholdsvis 3, 1 og 0 originale svar i snitt per elev. Snittet for gruppen ligger på 1,214, og havner derfor over både middels og lav måloppnåelse. Fra tabell 4.2.1 kan en se at elevene med lav måloppnåelse hadde totalt 0 originale svar. Det er en tydelig trend på at høyere måloppnåelse gir bedre originalitet.

Tabell 4.3.7 under viser antall svar hver enkelt elev har gitt på hver oppgave, samt elevenes måloppnåelse i matematikk.

***Tabell 4.3.*** *7: Antall svar per elev fordelt på oppgave 1, 2 og 3*

| Elevnr. | Måloppnåelse | Oppgave 1 | Oppgave 2 | Oppgave 3 |
|---|---|---|---|---|
| 19 | Høy | 3 | 1 | 2 |
| 27 | Høy | 6 | 4 | 6 |
| 30 | Høy | 0 | 1 | 2 |
| 8 | Middels | 4 | 1 | 0 |
| 9 | Middels | 4 | 3 | 4 |
| 11 | Middels | 3 | 1 | 1 |
| 12 | Middels | 0 | 0 | 0 |
| 14 | Middels | 1 | 0 | 0 |
| 31 | Middels | 3 | 0 | 0 |
| 33 | Middels | 3 | 0 | 3 |
| 35 | Middels | 5 | 4 | 0 |
| 22 | Lav | 0 | 0 | 0 |
| 23 | Lav | 1 | 1 | 0 |
| 3 | Lav | 4 | 4 | 0 |
| Sum | | 37 | 20 | 18 |
| Sum høy | | 9 | 6 | 10 |
| Sum middels | | 23 | 9 | 8 |
| Sum lav | | 5 | 5 | 0 |

For gruppen uten eksempeloppgaver er det elevene med middels måloppnåelse som leverer totalt flest oppgaver med 40 oppgaver, samtidig er de den klart største gruppen med 8 elever. Elevene med høy måloppnåelse lever 25, mens elevene med lav måloppnåelse leverer 10. Elevene med høy måloppnåelse leverer noenlunde likt på alle de tre oppgavene, mens elevene med middels leverer klart flest oppgaver på oppgave 1. Totalt for hele gruppen ser vi en dalende trend, hvor det leveres flest oppgaver på oppgave 1, færre på 2, og minst på oppgave 3. Fra tabell 4.2.1 kan en se at selv om elevene med middels måloppnåelse leverer klart flest oppgaver, har de fortsatt et lavere gjennomsnitt leverte oppgaver per elev enn elevene med høy måloppnåelse.  Elevene med høy måloppnåelse leverer 8,33 oppgaver per elev, middels leverer 5 per elev, og lav leverer 3,333 per elev.

## 4.4 Gruppene sammenliknet med hverandre

I dette delkapitlet skal vi se på resultatene fra gruppen med eksempel og gruppen uten eksempel, sammenliknet med hverandre. Først vil vi se på gruppene fordelt på måloppnåelse mot hverandre, så gruppene generelt mot hverandre. Det foretas også en analyse av resultatene til elevgruppene, hvor tallene sammenliknes opp mot hverandre ut fra antall deltakere og andel svar. Gruppen med eksempler blir presentert først.

### 4.4.1 Elevenes flytskåre

Ved å se på tabell 4.2.2 og tabell 4.3.2 ser vi at det hos begge gruppene synker flytskåren med synkende måloppnåelse. Elevene med høy måloppnåelse med eksempler har klart høyere gjennomsnittsskåre sammenlignet med elevene uten eksempler, med relativt sett samsvarende median. Elevene med middels og lav måloppnåelse, uten eksempler har derimot høyere flytskåre, med henholdsvis 5 mot 4,5 og 3,333 mot 1,667. Fortsatt har gruppen med eksempler høyere gjennomsnitts flytskåre totalt sett, grunnet elevene med høy måloppnåelse sin høye skåre.

Vår studie viser at elever som fikk presentert et hint i form av eksempel, fikk en høyere flytskåre enn elevgruppen uten eksempel. Som vist i tabellene (Finn frem tabellene til flytskåre for hver gruppe) fikk elevene med eksempel en flytskåre på 6,091, og elevene uten eksempel fikk en skåre på 5,357. Dette er et gjennomsnitt per elev innenfor hver enkelt gruppe.

Ved å se på elevenes flytskåre ut fra måloppnåelse, ser vi at elever med høy måloppnåelse presterer bedre. Den gjennomsnittlige flytskåren for elever med høy måloppnåelse, i gruppen som ikke fikk eksempel, er 8,333, med en median på 6. Dette kan også tyde på at det er noen elever som har generert tydelig flere problemer enn resterende. Elevene med middels og lav måloppnåelse hadde henholdsvis 5 og 3,333 i gjennomsnittlig flytskåre, og henholdsvis 5 og 2 som median. I gruppen med eksempel viser det tydelig at de med høy måloppnåelse i matematikk skårer bedre på flyt. Høy, middels og lav måloppnåelse har henholdsvis 11, 4,5 og 1,667 som gjennomsnittlig flytskåre. Medianen er 12 for høy måloppnåelse, 4 for middels og 2 for lav måloppnåelse.

Ut fra tallene over ser vi tydelig at elever med høy måloppnåelse skårer bedre på flyt enn elever med middels og lav måloppnåelse. Samtidig viser tallene at elever med middels og lav måloppnåelse, skårer bedre på flyt i gruppen uten eksempel. Den største forskjellen der er på elever med lav måloppnåelse, hvor gruppen uten eksempel skårer tilnærmet dobbelt så høyt. Elevene med høy måloppnåelse skårer tydelig mye bedre i gruppen med eksempel, sammenliknet med gruppen uten eksempel. Her har gruppen uten eksempel et snitt på 8,333 og gruppen med eksempel et snitt på 11.

Ser en elevgruppene ut fra måloppnåelse mot hverandre får vi følgende. Elever med høy måloppnåelse uten eksempler genererte totalt 33 av 122 problemer, 27,05%, selv om de utgjorde ca. 21% av deltakerne. Middels måloppnåelse genererte totalt 68 av 122 problemer, 55,74%, med en gruppeandel på ca. 57%. Den siste gruppen genererte totalt 21 av 122 problemer, 17,21%, med en gruppestørrelse på ca. 21%. Her ser vi en tendens til at elevene med høy måloppnåelse genererte en god del av problemene, selv om gruppen bestod av 21% av deltakerne. Den største gruppen, middels måloppnåelse, sto for 55,74% av problemene, noe en kan tenke er naturlig ettersom denne var størst med 57% av deltakerne. Dog er det vanskelig å si at det er derfor de genererte flest oppgaver, selv om dette kanskje er nærliggende. Gruppestørrelsen kan være en avgjørende faktor for hvorvidt antallet problemer generert blir større eller mindre, men vi kan ikke si sikkert at om de to andre gruppene var like store at de hadde generert om lag like mange problemer.

Elevgruppen med eksempler, med hensyn på måloppnåelse, har en flytskåre som er fordelt som følger. Høy måloppnåelse genererte totalt 52 av 88 problemer, 59,09%, med en gruppestørrelse på ca. 36%. Gruppen med middels måloppnåelse genererte totalt 29 av 88 problemer, 32,96%, med en gruppestørrelse på ca. 36% og lav måloppnåelse genererte 7 av 88, 7,96%, og var en gruppestørrelse på ca. 27%. Her ser en tydelig forskjell på hvor mange problemer som er generert i gruppene, hvorpå gruppen med høy måloppnåelse har en klart større andel. Dette kan sammenliknes med middels måloppnåelse som er en like stor gruppe, men sto for 32,96% av problemene. En kan argumentere for at dette er grunnet måloppnåelsen i matematikk, ettersom gruppene er like store, og begge har samme utgangspunkt for å prestere, men selve grunnen er vanskelig å stadfeste.

Gruppene mot hverandre viser en forskjell, der gruppen med eksempel skårer bedre på flyt, kontra gruppen uten eksempel. Det som kanskje er mest interessant er at elevene med høy måloppnåelse presterer betydelig bedre, samt har en dobbelt så høy median, noe som tyder på at elevene seg imellom skårer noe likt i denne gruppen.

### 4.4.2 Elevenes uløselige oppgaver

Ved sammenlikning av de to ulike gruppene, se tabell 4.2.3 og tabell 4.3.3, ser en at eksperimentgruppen leverer flere uløselige oppgaver i gjennomsnitt per elev ved både høy og middels måloppnåelse. Videre har eksperimentgruppen ved middels måloppnåelse tydelig flere uløselige oppgaver enn kontrollgruppen. Snittet for gruppene sett bort i fra måloppnåelse i matematikk er også lavere for elevene i kontrollgruppen.

Her forekommer resultatene og drøftingen av antall uløselige problemer per gruppe, og for måloppnåelse innad i gruppene. Det forekommer også en analyse av antall uløselige oppgaver generert, og dette er sett i lys av totalt antall oppgaver generert, dette inkluderer også trivielle oppgaver.

Elevgruppen med eksempler genererte totalt 13 uløselige problemer, av en total på 88. Dette tilsvarer en andel på 14,77%. Gruppen genererte i snitt 1,182 uløselige problemer, med en median på 1. Dette viser at elevgruppen i snitt ikke genererte særlig mange uløselige oppgaver. Kontrollgruppen genererte 12 uløselige problemer, av en total på 122, noe som tilsvarer en andel 9,84%.  Denne gruppen genererte i snitt 0,857 uløselige oppgaver med en median på 0,5. Her ser vi at elevgruppen med eksempler genererte flere uløselige problemer i snitt, og har en høyere andel av uløselige oppgaver.

Ser en på gruppen i henhold til måloppnåelse, i gruppen uten eksempler, genererte elever med høy måloppnåelse 0,667 uløselige problemer, med en median på 1. Elevgruppen med lav måloppnåelse genererte i snitt 1 uløselige problemer per elev, med en median på 1, og middels måloppnåelse genererte i snitt 0,875 uløselige oppgaver, med en median på 0. Tallene viser at antall uløselige oppgaver per elev er noe høyere desto lavere måloppnåelsen i faget er, for elevene uten eksempler. Elevgruppen med eksempler og høy måloppnåelse, genererte i snitt 0,75 uløselige problemer med en median på 0,5. Middels genererte 2,25 uløselige problemer og lav måloppnåelse genererte 0,333 uløselige problemer. Medianen er henholdsvis 1,5 og 0. Her viser tallene at elever med middels måloppnåelse genererte tydelig flere uløselige problemer, enn lav og høy måloppnåelse.

Sett mot gruppen uten eksempler, genererte elevgruppen uten eksempler færre antall uløselige problemer innenfor middels og høy måloppnåelse, mens lav måloppnåelse genererte tre ganger så mange uløselige problemer.

Elevene med høy måloppnåelse uten eksempler genererte totalt 2 uløselige oppgaver, av en total på 12 uløselige problemer, noe som tilsvarer en andel 16,67%. Sett i sammenheng med totalt antall oppgaver generert av høy måloppnåelse, tilsvarer dette 2 av 33, som er en andel på 6,06%. Gruppen med middels måloppnåelse uten eksempler sto for 7 av 12 uløselige problemer, en andel på 58,33%. Ut av totalt antall problemer denne gruppen genererte, var det 7 uløselige av 68 totalt genererte, en andel på 10,29%. Den siste gruppen genererte totalt 3 uløselige oppgaver, 25% av totalt antall uløselige problemer for hele gruppen uten eksempler. Til sammen genererte lav måloppnåelse 21 problemer, hvorav 3 var uløselige, en andel på 14,29%. Her ser vi tydelig at elevgruppen med høy måloppnåelse har en lavere andel av uløselige problemer (6,06%) sett i sammenheng med totale antall problemer generert av gruppen. Deretter følger middels måloppnåelse med 10,29% og til slutt lav måloppnåelse med 14,286%. Ser en bare på antall uløselige problemer i gruppen, ut fra det totale antallet uløselige problemer, er det klart at middels måloppnåelse har generert flest med 58,33%.

I gruppen uten eksempler genererte elevene med høy måloppnåelse totalt 3 av 13 uløselige oppgaver, en andel på 23,08%. Sett i sammenheng med totalt antall problemer denne gruppen genererte, er det 3 av totalt 52 problemer, 5,77%. Gruppen med middels måloppnåelse sto for 9 av 13 uløselige oppgaver, tilsvarende 69,23%. Tar én hensyn til totalt antall genererte problemer, var 9 av 29 problemer uløselige, en andel på 31,03%. Lav måloppnåelse genererte 1 av 13 uløselige problemer, 7,69%. Totalt genererte de 7 problemer, med en andel uløselige problemer på 14,29%. Her er det ikke like stort sprang fra høy til middels måloppnåelse når det kommer til andel uløselige problemer, ut fra alle genererte i gruppen, som for samme måloppnåelse med eksempler. Dog viser tallene at elever med lav måloppnåelse genererte tydelig større andel uløselige problemer i forhold til totalen i gruppen. Ser en antall uløselige problemer i gruppen til sammen, viser det at middels måloppnåelse har en tydelig stor andel uløselige problemer, sammenliknet med lav og høy måloppnåelse.

### 4.4.3 Elevenes trivielle oppgaver

I tabell 4.2.4 og 4.3.4 ser en to ulike trender. I eksperimentgruppen produseres det flere trivielle oppgaver med økt måloppnåelse, kontra kontrollgruppen som produserer færre trivielle oppgaver med økende måloppnåelse. Likevel har eksperimentgruppen færre trivielle per elev enn det kontrollgruppen har i alle ulike nivå av måloppnåelse. Følgelig blir snittet for de to gruppene svært forskjellig. Hvor eksperimentgruppen har 0,727 trivielle svar i gjennomsnitt per elev, mot kontrollgruppens 2 trivielle svar per elev, nesten 3 ganger så mye.

Resultatene til studien tar for seg ulike kategorier, og kjerneelementer til kreativitet, på bakgrunn av studien til Van Harpen og Sriraman (2013). Her har vi kategorisert noenlunde likt dem, med noen tilpasninger ettersom nivåene kan argumenteres å være ulike. Innenfor trivielle problemer, har vi definert dem som genererte problemer som vi anser som for enkle for nivået elevene skal befinne seg når de går på 7. trinn. Dette kan problematiseres ved å dra frem hvorvidt en oppgave er for enkel for noen, og vanskelig for andre. Oppgavene som er ansett som trivielle er basert på en subjektiv vurdering, og ut fra hva kompetansemålene etter 7. årstrinn omhandler. Med dette sagt er oppgavene først vurdert individuelt av hver av oss, for så å diskutere dem sammen og komme til en enighet.

Elevgruppen uten eksempler genererte til sammen 35 trivielle problemer, av totalt 122 problemer. Dette tilsvarer en andel på 28,69% av totale oppgaver generert. Her forekom det et gjennomsnitt på 2,5 trivielle oppgaver per elev, med en median på 2,5. Gruppen med eksempler presenterte totalt 88 problemer, hvorav 8 ble definert som trivielle. Dette er en andel på 9,09%, som er betydelig lavere andel enn kontrollgruppen. Eksperimentgruppen hadde et gjennomsnitt på 0,727 trivielle problemer generert per elev, og en median på 0. Her ser en tydelig at kontrollgruppen presenterer vesentlig flere trivielle problemer, både i andel og i snitt per elev.

Ser en på antall trivielle problemer fordelt på måloppnåelse får vi følgende resultater. Gruppen med høy måloppnåelse i kontrollgruppen genererte 2 trivielle problemer i snitt per elev, med en median på 2. Elevene med middels og lav måloppnåelse genererte henholdsvis 2,625 og 2,667 trivielle problemer i snitt, og hadde median på henholdsvis 3 og 2. Tallene her viser at elevene med høy måloppnåelse genererte et tydelig færre antall trivielle problemer enn de to andre, mens middels og lav genererte noenlunde samme antall. Dog ser en at medianen til gruppen med middels måloppnåelse er noe høyere, som kan tyde på at gruppen i sin helhet har flere elever som genererer flere trivielle problemer. Eksperimentgruppen, med hensyn på måloppnåelse, genererte 1,25, 0,5 og 0,333 trivielle problemer, henholdsvis høy, middels og lav måloppnåelse. Her ser en at antall trivielle oppgaver øker noe i takt med måloppnåelsen i matematikk. Hvorvidt dette skyldes at elever med høy måloppnåelse prøvde å generere vanskelige problemer uten suksess eller ikke, er ikke mulig å si. Medianen til høy, middels og lav måloppnåelse er henholdsvis 1,5, 0 og 0. Dette viser til at noen av elevene i middels og lav gruppen presenterte tydelig flere trivielle problemer, sammenliknet med resten av gruppen.

Elevene med høy måloppnåelse i eksperimentgruppen sto for 5 av 8 trivielle problemer, noe som tilsvarer 62,5% av alle trivielle problemer. Middels måloppnåelse sto for 25% (2 problemer) trivielle, og lav måloppnåelse sto for 12,5% (1 problem). Her er det også verdt å dra frem at høy måloppnåelse sto for totalt 52 av 88 genererte problemer, noe som tilsvarer 59,09%. Ser en prosenten til høy måloppnåelse i lys av problemer de genererte, tilsvarer dette 5 av 52, som er 9,62%. Elevene med middels måloppnåelse genererte totalt 29 av 88, 32,96%, og antall trivielle problemer tilsvarer 6,89%. Elevene med lav måloppnåelse genererte totalt 7 av 88 problemer, tilsvarende 7,96%. Andel trivielle oppgaver tilsvarer 14,29%. Her ser vi tydelig at elevene med lav måloppnåelse genererte en større andel trivielle problemer av antall genererte problemer. Middels måloppnåelse genererte nesten halvparten så mange som lav måloppnåelse, og høy måloppnåelse genererte noe mindre enn middels måloppnåelse. Selv om gruppen med høy måloppnåelse sto for hoveddelen, 62,5%, av alle trivielle problemer i gruppen, genererte de også hoveddelen av totalt antall problemer 59%. Det er derfor rimelig å anta at den høye andelen av trivielle problemer ikke nødvendigvis dreier seg om at elever med høy måloppnåelse produserer flere trivielle problemer enn middels og lav måloppnåelse, men heller at de produserer flere problemer, og elevgruppen dermed produserer flere trivielle problemer, som et biprodukt.

I kontrollgruppen, sett med henhold til måloppnåelse, produserte elevene med høy måloppnåelse 6 av 35 trivielle problemer. Dette tilsvarer 17,14% av alle trivielle problemer. Middels måloppnåelse sto for 21 av 35, som tilsvarer 60%, og lav måloppnåelse sto for 8 trivielle problemer, 22,86%. Gruppen med høy måloppnåelse produserte til sammen 33 av 122 oppgaver, 27,05%, av totalen og hadde en andel trivielle oppgaver på 18,182%. Middels måloppnåelse genererte 68 av 122 problemer, 55,74%, av totalen. Andel trivielle av problemene generert av gruppen tilsvarer 30,88%. Lav måloppnåelse sto for 21 av 122 problemer, 17,21%, hvorav 8 var trivielle. Dette tilsvarer en andel på 38,09%. Som tallene viser, ser vi at elevene med høy måloppnåelse genererte en klart mindre andel trivielle problemer, ut fra totalt antall problemer den gruppen genererte. Elevene med lav måloppnåelse genererte tydelig en større andel på 38%, mens middels måloppnåelse genererte ca. 31% trivielle problemer av det totale antallet til gruppen.

### 4.4.4 Elevenes fleksibilitetsskåre

Ser en på tabell 4.2.5 og 4.3.5, kan det tyde på at fleksibilitetsskårene ser relativt like ut. De fordeler seg begge utover tilnærmet like mange kategorier, og begge gruppene viser en synkende fleksibilitetskåre med synkende måloppnåelse. Likevel er gjennomsnittlig fleksibilitetskåre for hele eksperimentgruppen litt større enn kontrollgruppen.

Elevgruppene har skåret noe ulikt når det kommer til fleksibilitet. Herunder har elevgruppen med eksempler en gjennomsnittskåre på 2,455, med en median på 2. Dette viser til at gruppens sprang fra laveste til høyeste fleksibilitetsskåre ikke er stor. Elevgruppen uten eksempel har en gjennomsnittlig fleksibilitetsskåre på 2,071, samt en median på 2. Her ser vi at gruppen uten eksempler har et gjennomsnitt som er nærmere medianen i gruppen. Gruppene opp mot hverandre viser ikke en stor forskjell, men gruppen med eksempler skårer høyere på fleksibilitet, selv med 3 færre deltakere. Dog kan dette være en grunn både for og mot hvorfor skåren er ulik. Gruppen med eksempler kunne skåret noe høyere, eller lavere hvis det hadde vært like mange deltakere som den andre gruppen.

Fordelt på måloppnåelse har elevgruppen med eksempler fått en annen fleksibilitetsskåre. Her har elever med høy måloppnåelse i matematikk fått en skåre på 4 og en median på 4, elever med middels måloppnåelse har en skåre på 1,75 og median på 1,5 og elever med lav måloppnåelse har fått en skåre på 1,333 og median på 2. Her igjen viser elever med høy måloppnåelse en høyere skåre, enn de andre elevene. Dette kan trolig være grunnet kompetansen de har i matematikk kontra de andre. Dog er dette også en skåre som er over dobbelt så høy som hver av de andre måloppnåelsene. Den gjennomsnittlige skåren og medianen er noenlunde lik innenfor hver gruppe.

Elevgruppen uten eksempler skårer relativt sett likt gruppen med eksempler, selv når det er fordelt på måloppnåelse. Her ser vi at elevene med høy måloppnåelse får en gjennomsnittlig fleksibilitetsskåre på 4 og median på 4, noe som er likt elevgruppen med eksempler. Elevene med middels måloppnåelse fikk en skåre på 2 og median på 2, og gruppen med lav måloppnåelse skårer 1,333 og en median på 1. Igjen viser gruppen med høy måloppnåelse en større skåre enn gruppene med middels og lav. Dette kan også være grunnet den høyere måloppnåelsen i faget.

Med tallene nevnt over, ser vi at det ikke er en stor forskjell mellom gruppene, men at gruppen med eksempler skårer noe høyere enn gruppen uten eksempler. Forskjellen i skåre fra elever med høy måloppnåelse til de med middels og lav, er noe høyere for elevene i gruppen med eksempler kontra gruppen uten. Selv med én mindre deltaker i gruppen uten eksempler, viser de en lik skåre som gruppen med eksempler for elever med høy måloppnåelse.

### 4.4.5 Elevenes originalitet

Ut fra tabell 4.2.6 og 4.3.6, viser begge gruppene samme trend – synkende originalitetskåre med synkende måloppnåelse. Fortsatt har eksperimentgruppen en klart høyere originalitetsskåre, nesten dobbelt så mye som kontrollgruppen. Fra tabellene kan en se at det skyldes at elevene i eksperimentgruppen leverte relativt sett svært mange oppgaver i forhold til kontrollgruppen. Elevene med lav måloppnåelse i både eksperiment- og kontrollgruppen leverte totalt henholdsvis 1 og 0 originale oppgaver.

Elevenes originalitet er definert ut fra hvor unike svarene er. I vårt tilfelle er det, som nevnt tidligere, tatt utgangspunkt i at et svar er originalt hvis det er generert av 10% eller mindre av deltakerne i gruppen. Dog er dette en faktor som er diskutabel ettersom det er tatt en subjektiv vurdering fra forfatternes side. Originalitet kan muligens være med på å vise hvorvidt deltakerne velger å tenke annerledes enn meddeltakere, og hvorvidt de er kreative når det kommer til det å tenke ulikt andre. Her kan dette tenkes som elevenes evne til å «tenke utenfor boksen».

Elevgruppen uten eksempler genererte gjennomsnittlig 1,214 originale svar per elev, med en median på 0. Her viser gruppen til at det de ikke klarte å generere særlig mange originale svar, spesielt siden halvparten ikke genererte noen originale svar. Totalt genererte de 17 originale problemer av en total på 88 problemer, en andel originale problemer på 19,32%. Gruppen med eksempler genererte gjennomsnittlig 2,364 originale svar per elev, med en median på 1. Totalt genererte de 29 originale problemer, av 122 genererte problemer, en andel på 23,77%. Dette tyder på at hoveddelen av elevene genererte få antall originale svar, og noen elever genererte flere som drar opp gjennomsnittet. Her ser vi tydelig at elevgruppen med eksempler genererte flere (nesten dobbelt så mange) originale problemer, enn gruppen uten eksempler. Det at gruppen uten eksempler er noe større enn gruppen med eksempler, kan også tyde på at gruppen med eksempler presterte bedre enn den uten. Om dette er grunnet tiltaket i form av eksempel eller ikke, er vanskelig å si. Det kan også være verdt å nevne at det ikke kan sies at gruppen med eksempler hadde prestert bedre, eller dårligere, om de var like mange. Dette får stå som spekulasjoner.

Ser en på gruppene ut fra måloppnåelse får vi følgende resultater. Elevgruppen med høy måloppnåelse uten eksempler genererte i snitt 3 originale problemer per elev, med en median på 2. Elevene med middels og lav måloppnåelse genererte i snitt henholdsvis 1 og 0, med en median på 0,5 og 0. Her ser en igjen at elevgruppen med høy måloppnåelse presterer bedre, og genererer flere originale problemer, tre ganger så mange som gruppen med middels måloppnåelse og en median som er fire ganger så høy. Elevgruppen med lav måloppnåelse genererte ingen originale problemer. Om dette er grunnet måloppnåelsen i faget, vanskelighetsgraden på testen eller fordi de rett og slett ikke «orket» kan en bare spekulere i.

Gruppen med eksempler har litt høyere tall. Elevene med høy måloppnåelse genererte i snitt 6 originale problemer, med en median på 6. De med middels og lav måloppnåelse genererte i snitt henholdsvis 1 og 0,333, og median på 1 og 0. Her ser en også en tydelig sammenheng mellom måloppnåelse og antall originale problemer som er generert. Gruppen med høy måloppnåelse genererte seks ganger så mange originale problemer som gruppen med middels måloppnåelse, og atten ganger så mange problemer som gruppen med lav måloppnåelse. Vi ser også tydelig at medianen ikke avviker særlig fra gjennomsnittet.

Ser en gruppene i lys av hverandre med hensyn på måloppnåelse, genererte gruppen med høy måloppnåelse, med eksempler, dobbelt så mange originale problemer, som samme måloppnåelse uten eksempler. Elevene med middels og lav måloppnåelse genererte noenlunde likt antall originale problemer i gruppene, dog genererte gruppen med lav måloppnåelse og med eksempler flere originale problemer enn samme måloppnåelse uten eksempler. Gruppen uten eksempler i forhold til den andre gruppen genererte nesten halvparten så mange originale problemer, og har en median på 0 (én lavere enn gruppen med eksempler). Dette kan tyde på at det er et fåtall elever som står for hoveddelen av de originale problemene i gruppen uten eksempler.

Gruppen med høy måloppnåelse uten eksempler, sto for 9 av 17 originale problemer, som tilsvarer 52,94%. Sett i forhold til antallet problemer denne gruppen genererte, har de en andel originale problemer på 9 av 33, som er 27,27%. Elevgruppen med middels måloppnåelse genererte 8 av 17 originale problemer, 47,06%. Ut fra det totale antall problemer denne undergruppen genererte, har den en andel på 8 av 68, som er 11,765%. Elevene med lav måloppnåelse genererte totalt 0 originale problemer. Her ser vi en tydelig større andel av originale problemer i gruppen med høy måloppnåelse, både for det totale antall originale problemer, samt sammenliknet med de to andre gruppene.

Elever med høy måloppnåelse med eksempler genererte totalt 24 av 29 originale problemer, en andel på 82,76%. Innad i gruppen har den en andel på 24 av 52, som tilsvarer 46,15%. Middels måloppnåelse sto for 4 av 29 originale problemer, en andel på 13,79%. Gruppen har generert totalt 29 problemer, hvor 4 av dem er originale, en andel som tilsvarer 13,79%. Ser en til elevene med lav måloppnåelse har de generert 1 originalt svar, av totalt 29, en andel på 3,45%. Til sammen genererte de 7 problemer totalt, og har en andel originale svar på 14,29%. Her ser en at elevene med høy måloppnåelse har generert desidert flest antall originale svar, både i den store gruppen, og hvis en skiller ut fra måloppnåelse. Elever med middels måloppnåelse har generert tydelig flere antall originale problemer, men sett i lys av andelen har gruppen generert noe mindre enn gruppen med lav måloppnåelse.

***Tabell 4.4 1***: *Antall svar på oppgave 1, 2 og 3 elever med eksempel*

| | | | | |
|---|---|---|---|---|
| Elevnr. | Måloppnåelse | Oppgave 1 | Oppgave 2 | Oppgave 3 |
| 15 | Høy | 3 | 3 | 3 |
| 16 | Høy | 8 | 3 | 4 |
| 17 | Høy | 5 | 5 | 5 |
| 29 | Høy | 4 | 1 | 0 |
| 5 | Middels | 3 | 1 | 0 |
| 6 | Middels | 0 | 0 | 0 |
| 7 | Middels | 2 | 0 | 2 |
| 34 | Middels | 5 | 0 | 5 |
| 21 | Lav | 0 | 0 | 0 |
| 24 | Lav | 1 | 0 | 1 |
| 1 | Lav | 3 | 0 | 0 |
| Sum svar | | 34 | 13 | 20 |
| Sum høy | 44 | 20 | 12 | 12 |
| Sum middels | 18 | 10 | 1 | 7 |
| Sum lav | 5 | 4 | 0 | 1 |

***Tabell 4.4 2**: Antall svar på oppgave 1, 2 og 3 elever uten eksempel*

| Elevnr. | Måloppnåelse | Oppgave 1 | Oppgave 2 | Oppgave 3 |
|---|---|---|---|---|
| 19 | Høy | 3 | 1 | 2 |
| 27 | Høy | 6 | 4 | 6 |
| 30 | Høy | 0 | 1 | 2 |
| 8 | Middels | 4 | 1 | 0 |
| 9 | Middels | 4 | 3 | 4 |
| 11 | Middels | 3 | 1 | 1 |
| 12 | Middels | 0 | 0 | 0 |
| 14 | Middels | 1 | 0 | 0 |
| 31 | Middels | 3 | 0 | 0 |
| 33 | Middels | 3 | 0 | 3 |
| 35 | Middels | 5 | 4 | 0 |
| 22 | Lav | 0 | 0 | 0 |
| 23 | Lav | 1 | 1 | 0 |
| 3 | Lav | 4 | 4 | 0 |
| Sum svar | | 37 | 20 | 18 |
| Sum høy | 25 | 9 | 6 | 10 |
| Sum middels | 40 | 23 | 9 | 8 |
| Sum lav | 10 | 5 | 5 | 0 |

Gruppene sammenlignet med hverandre viser at det er samme trend for begge gruppene. De leverer klart flest på oppgave 1. Antallet går betydelig ned til oppgave 2 og 3. For begge gruppene er det elevene med høy måloppnåelse som har flest oppgaver per elev (flytskåre). Det er en viss forskjell mellom gruppene, hvor eksperimentgruppen leverer flere oppgaver på oppgave 3 enn oppgave 2 (henholdsvis 20 og 13), i motsetning til kontrollgruppen som leverer flere oppgaver på oppgave 2, enn oppgave 3 (henholdsvis 20 og 18).

For begge gruppene er det elevene med lav måloppnåelse som leverer desidert minst, dog leverer elevene uten eksempeloppgave flere oppgaver enn elevene med eksempeloppgave på oppgave 1 og 2, henholdsvis 5 mot 4 og 5 mot 0. Det er bare på oppgave 3 elevene i eksperimentgruppen leverer flere enn kontrollgruppen, med 1 mot 0 oppgaver.

Én kan se i både eksperiment og kontrollgruppen, at det er elever som utmerker seg. Fra tabell 4.3 11 ser en at elev 16 og 17 fra eksperimentgruppen har levert 15 løselige oppgaver hver. På den andre siden har elev 6 og 21 levert 0 oppgaver. Fra tabell 4.3.12 kan en se at elev nummer 27 har levert 16 oppgaver, og elev nummer 9 har levert 11 oppgaver, mens elev nummer 12 og 22 har levert 0.

# 5 Drøfting

I dette kapittelet skal vi ta for oss resultatene av vår studie, og drøfte de opp mot relevant teori. Her forekommer det drøfting i ulike deler, ettersom hvert kjerneelement innenfor kreativitet har interessante resultater. Deretter foretar vi en drøfting av resultatene sett i lys av elevgruppens måloppnåelse, ettersom det kan synes å være en forskjell der. Neste underkapittel tar for seg resultatene våre i lys av utforskende undervisning, som er hovedtemaet i studien vår. Her forekommer problemløsing og – generering. Til slutt forekommer et kapittel som tar for seg de feilkildene vi anser som relevante til studien, som også kan ha vært med på å påvirke resultatene våre.

## 5.1 Elevenes flyt

Vår studie viser en gjennomsnittlig flytskåre på 5,36 for elevgruppen uten eksempler, og 6,09 for elevgruppen med eksempler. Dette er noe som kan dras i retning at et eksempel til oppgaven bidrar til at elevene skårer høyere på flyt, altså antall oppgaver som er løselige og ikke trivielle. Sammenliknet med det Van Harpen og Sriraman (2013, s. 212), var skåren 4.6, 2.0 og 4.9, henholdsvis til amerikanske elever, elever fra Shanghai og Jiaozhou, ser en at våre elevgrupper skårer tydelig høyere. Deres median var henholdsvis 4, 1.5 og 5. Vår studie viser til en median på 4 for elever med eksempel og 5.5 for elevene uten eksempel. Her kan en se at elevgruppene våre i større grad hadde ulikheter fra median til gjennomsnitt, spesielt for gruppen med eksempler. Det kan diskuteres hvorvidt sammenlikningsgrunnlaget er valid, ettersom Van Harpen og Sriraman (2013) foretok studien på tydelig eldre deltakere, som en kan tenke seg har større kompetanse innenfor matematikk. Dog kan dette være med på å styrke våre resultater, ettersom elevenes problemer, både i vår og deres studie, ble analysert ut fra kompetansenivået til elevene. Ettersom problemene blir analysert ut fra flyt, fleksibilitet, trivialitet og uløselighet, opp mot elevenes forkunnskaper og forventet kompetanse, viser studien vår høyere skåre på flyt. Her kan en diskutere og problematisere studien vår, på ulike grunnlag.

En kan dra frem Kokinov et al. (1997, s. 4-5) som fant at elever som får presentert et hint, og relevansen til oppgaven, presterte dårligere enn kontrollgruppen i problemløsning. De fant også at elever som får presentert et hint, i form av en illustrasjon, uten noen form for forklaring, presterte bedre enn kontrollgruppen. I den grad vår studie kan kategorisere problemgenerering som problemløsning, i den forstand at et løselig og relevant problem er korrekt, kan dette sammenliknes. Her fant vår studie at elever som får presentert et hint, i form av eksempel, genererte flere problemer i snitt per elev. Dog kan dette problematiseres ettersom vår studie er direkte myntet til generering av problemer, med og uten eksempel, og Kokinov et al. (1997) sin studie omhandler problemløsning. Deres deltakere er studenter på universitetet, og våre er på barneskolen, så hvorvidt dette kan direkte sammenliknes er vanskelig å si. Ettersom vår studie har noe sammenfallende med deres, anså vi det som relevant, men graden av relevans er vanskelig å si. Samtidig gjennomførte Kokinov et al. (1997, s. 4) studien med klart flere deltakere, som igjen kan gjøre det vanskelig å sammenlikne studiene.

Ser en på antall trivielle problemer generert, viser dette at elevgruppen med eksempler produserte 8, av totalt 67, trivielle problemer. Dette tilsvarer en prosentandel på 11,94%. Elevgruppen uten eksempler produserte en klart større andel, med 35 trivielle av 75 løselige, en andel på 46,67%. Her kan en dra frem hvorvidt eksemplet har vært med på å påvirke antall trivielle problemer, eller ikke. Det er nærliggende å tro at det har hatt en viss effekt, ettersom spriket mellom kontroll og eksperimentgruppen er så stort. Elevgruppene ble ikke intervjuet, og kan derfor ikke besvare selv om dette er tilfellet. Sett i lys av Van Harpen og Sriraman (2013, s. 211-212), produserte våre elever med eksempel en noe høyere andel enn deltakerne deres fra USA. De amerikanske elevene genererte en andel trivielle problemer på 9%, Shanghai hadde en andel på 8% og Jiaozhou hadde 6%. Elevgruppen vår uten eksempel ligger nært de amerikanske elevene, men samtidig noe over. Her viser vår studie at det er kun gruppen med eksempel som produserer en andel trivielle problemer i nærheten av deltakerne til Van Harpen og Sriraman (2013). Samtidig kan en problematisere hvorvidt antallet trivielle problemer kunne vært annerledes, hvis det hadde vært samme antall deltakere i vår studie. Det er nærliggende å tenke at vår elevgruppe hadde generert en noenlunde lik andel trivielle problemer, men det kan ikke vites. En kan argumentere for at elevgruppen hadde generert en mindre andel trivielle problemer hadde det vært en større gruppe, ettersom vår studie viser at det er et fåtall elever som har generert hoveddelen av de trivielle problemene. Dette er mest gjeldende for elevene uten eksempler.

Våre elevgrupper genererte en total andel uløselige problemer på 9,84% for elever uten eksempel, og 14,77% for gruppen med eksempel. Her ser vi at gruppen uten eksempel genererte tydelig færre uløselige problemer, og har dermed «prestert bedre». Vår tanke på vei inn i studiet var at elevene med eksempel kunne generere færre uløselige problemer, enn kontrollgruppen, men her tok vi åpenbart feil. Vi så for oss at elevene i større grad skulle bruke eksemplet på nytt, bare med nye tall, som hadde resultert i et løselig problem. Det kan tenkes at elevene har blitt påvirket av eksemplet, og vårt forsøk på å motivere til å generere så vanskelige problemer som mulig, i den grad at de har generert oppgaver som ikke kan løses. Om dette er tilfellet eller ikke, er vanskelig å si. Det kan være flere faktorer som spiller inn på hvordan elevene valgte å generere problemer, og hvordan de har kommet frem til uløselige problemer. Kan eksempelet ha vært med på å hindre elevene i å generere løselige problemer, eller kan det ha vært en faktor som har hjulpet elevene? Dataene våre tilsier at elevene med eksempel har generert flere problemer i snitt per elev, samtidig som de har generert en større andel uløselige problemer. Verdt å nevne her er at en enkelt elev står for 6 av 13 uløselige problemer for eksperimentgruppen, som åpenbart påvirker datasettet når utvalget er så lite.

Ser en andelen uløselige problemer, opp mot Van Harpen og Sriraman (2013, s. 211-212), viser vår studie at begge gruppene produserte en mindre andel uløselige problemer enn deres tre grupper. Her har gruppen vår med eksempel produsert en noenlunde lik andel som gruppen fra Jiaozhou, som hadde 15%. Videre understreker Van Harpen og Sriraman (2013, s. 212) at flere elever hadde problemer med å gi nok informasjon til å løse oppgaven, og derfor ble problemet kategorisert som uløselig. Dette har også vært tilfellet for flere av våre oppgaver, samt oppgaver som ikke er relevante til teksten. Her kommer problematiseringen av hvorvidt en egentlig kan sammenlikne vår studie, med deres. Hvorvidt våre resultater har noen sammenfallende likheter eller ikke, kan diskuteres, mens vår subjektive oppfatning og vurdering av problemene i henhold til elevenes kunnskaper og kompetanse i matematikk, fører til at vi anser resultatene som sammenliknbare. Det at vår studie har en tydelig mindre andel deltakere, kan antas å ha påvirket dataene og analysen vi har gjort. Hadde det vært flere deltakere i vår studie, kunne resultatene vært annerledes, men det er vanskelig å vite.

Studien vår tar for seg samme rammeverk og oppgaver, med tilpasninger på oppgave 2, som Van Harpen og Sriraman (2013, s. 207-208). I likhet med dem analyserer og beregner vi dataene på samme måte, og bruker samme teori, blant annen teori, for å underbygge valgene våre. Det er klart poengtert, både i vår og Van Harpen og Sriraman (2013, s. 209-210), at de genererte problemene er vurdert av forskerne for hvorvidt de er løselige eller trivielle. Her forekommer det åpenbart en subjektiv vurdering av problemene. Hvorvidt våre vurderinger kan være, og er, sammenfallende med Van Harpen og Sriraman (2013) sine vurderinger, kan ikke sies. Hadde vi vurdert de genererte problemene i deres studie, og motsatt, kan det være seg at resultatene hadde vært helt annerledes. Våre krav og standarder til våre deltakere, er ikke å anse som en form for fasit, men en subjektiv og diskutabel vurdering av kompetanse og kunnskap i matematikk. Her er det også verdt å nevne at elevene våre har hatt lite, til ingen, erfaring med problemgenerering og -løsning. Noe som igjen kan være med på å påvirke både oss og deltakerne. Problematiseringen her går på grunnlag av hvorvidt vi har vært litt mer føyelig til å ta en diskutabel oppgave som triviell eller ikke, ettersom vi har vært klar over elevenes tidligere kunnskap innenfor temaet.

## 5.2 Elevenes fleksibilitet

I dette underkapittelet tar vi for oss drøfting av elevenes fleksibilitetsskåre. Herunder skal vi se på elevenes fleksibilitetsskåre opp mot teori, og drøfte hvorvidt dette er relevant og case-studien vår har hatt noe effekt. Igjen kan det være fornuftig å påminne at fleksibiliteten er beregnet som antall kategorier de genererte problemene faller inn under. Altså; hvis en elev genererer flere oppgaver i samme kategori får eleven en lav fleksibilitetsskåre, og en elev som genererer flere problemer i flere kategorier får eleven en høy fleksibilitetsskåre.

Ser en fleksibilitet i forhold til antall kategorier hver enkelt gruppe har presentert i, får vi følgende. Elevene uten eksempler hadde én elev som genererte to problemer som falt inn under kategorien «algebra», og for elever med eksempler var det ingen problemer i denne kategorien. Gruppen uten eksempler hadde to elever som genererte ett problem hver som havnet i kategorien «likninger», hvorpå gruppen med eksempler hadde like mange. På generell basis viser fleksibiliteten til elevene til at de hovedsakelig genererte problemer som havnet i kategorien «fire regnearter». Ser én dette i lys av LK20 (2020), skal elevene ha noe forståelse for bruken av algebra i matematikk, samt generell kunnskap om bruken av de fire regneartene. Her kan det tenkes at elevene i studien er tryggere på bruken av regneartene, og derfor velger å bruke dette for å besvare testen. Det at elevene, som presentert tidligere i oppgaven, har hatt lite til ingen erfaring med problemløsning og -generering, kan få en til å tenke at det er nye territorier de er uvante og usikre på, og derfor har problemer med å bruke mer avansert matematikk for å besvare testen. Læreplanen (2020) skal kunne noen punkter som går inn under geometri, dog omhandler dette i stor grad bygge, analysere, beskrive og gjennomføre. Med henhold til elevenes fleksibilitet i studien, og læreplanen, kan det tenkes at de ikke har vært gjennom det å selv generere problemer som kan knyttes til geometri. Elevgruppen med eksempler genererte totalt 6 problemer innenfor kategorien «geometri», alle av elever med høy måloppnåelse. Gruppen uten eksempler genererte totalt 10 problemer som faller i samme kategori, hvor fordelingen på høy, middels og lav måloppnåelse henholdsvis er 3, 6 og 1. Dette kan tyde på at elevene i denne gruppen har noe mer kunnskap angående bruken av geometri for å generere problemer. Om dette er grunnet høyere kunnskap, eller et større fokus på geometriske beregninger hos de spesifikke elevene, er vanskelig å si. Ettersom gruppene er randomiserte kan en tenke seg til at denne elevgruppen har noe bedre kunnskap. Dog kan en også dra frem at gruppene kan være i det minste laget for å få et svar som kan repliseres av andre studier. Det er verdt å nevne at elevgruppen uten eksempler er noe større enn den med eksempler, og dette kan også være grunnen til flere svar innenfor denne kategorien.

Vår studie viser til at elever genererer problemer i hovedsakelig én kategori, «fire regnearter», og innad i gruppene er det også «sammensatte oppgaver» for elever med eksempler, og «prosent/brøk» i gruppen uten eksempler. Dette kan tyde på at elevene gjerne genererer problemer som er i tråd med hva de føler at de behersker matematisk sett. Her har elever uten eksempler generert 2,64 problemer i snitt innenfor «fire regnearter», og elever med eksempler generert 3,18 i den samme kategorien. Det at elevene hovedsakelig genererer problemer i to kategorier er i tråd med hva Van Harpen og Sriraman (2013) skriver om. De fant at deltakerne hovedsakelig fokuserte på to kategorier, areal og lengde. Her må det poengteres at utvalget er noe ulikt ettersom deres studie foregår på eldre studenter, hvor de også har valgt ut deltakere ut fra en skåre på en matematisk kunnskapstest. Studien vår viser et fokus på to kategorier, «fire regnearter» og «sammensatte oppgaver». Her kan det være vanskelig å begrunne hvorfor det er slik, ettersom randomiseringen av gruppene skal være med på å utelukke at en gruppe har mer fokus på en kategori i matematikkundervisningen sin.

## 5.3 Elevenes originalitet

Når en ser på originalitetsskåren for kontroll og eksperimentgruppen, kan det se ut som en ser et paradoks. Eksperimentgruppen, som alle hadde samme eksempel og utgangspunkt, produserte klart flere originale oppgaver enn kontrollgruppen, som ikke hadde noe annet utgangspunkt enn sin egen kreativitet for å produsere oppgaver. I forkant av gjennomføringen av datainnsamlingen hadde vi sett for oss at et eksempel kunne samlet eksperimentgruppen rundt en felles tankemåte for å generere oppgaver, slik at kontrollgruppen ville hatt et fortrinn her. Dog går dette inn i en spekulativ argumentasjon, hvor en bare kan tenke rundt hvorvidt dette kan skape en sammenfallende tankegang, eller åpner opp for flere tankemåter.

Vår studie viser til at elever som får presentert eksempel, genererer tydelig flere originale problemer, sammenliknet med gruppen uten eksempel. Her er det tilnærmet dobbelt så mange problemer. Dog er vurderingen av originalitet foretatt som en subjektiv vurdering av oss, og derfor kan det diskuteres hvorvidt et problem er originalt eller ikke. Dømmingen av originalitet er foretatt i likhet med Van Harpen og Sriraman (2013, s. 204) sine vurderinger, og kan muligens dermed anses som noe korrekt. De sammenlikner dog ikke gruppenes originalitet mot hverandre, noe vi gjør.

Ettersom det er manglende forskning på sammenlikning av originale svar på tvers av grupper, med og uten tiltak, i studier, er det vanskelig for oss å sette våre resultater i perspektiv av annen forskning. Studien vår viser en klar forskjell, og kan muligens dermed argumentere for at ved presentasjon av eksempel genererer elever flere originale svar. Det kan også argumenteres i motsatt retning, ettersom gruppene ikke er helt homogene når det kommer til måloppnåelse.

Med dette sagt viste Bonotto og Santo (2015, s. 115-117) sin studie på elever på 5. trinn (10-11 år), med totalt 63 deltakere, at 3 av 58 problemer var originale på en skole og 10 av 131 på den andre skolen. For ordens skyld; var den første skolen litt dårligere i matematikk, sammenliknet med den andre skolen. Deres studie viser en liten økning i originalitet, når det kommer til prestasjoner i matematikk. Vår studie viser en mer tydelig sammenheng mellom måloppnåelse i matematikk, og evnen til å generere originale problemer. Dog går deres studie ut på å sammenlikne to relativt like gruppers kreativitet i problemløsning, når det kommer til virkelige artefakter, så sammenlikningen her er noe diskuterbar, ettersom vi undersøker hvorvidt et eksempel påvirker elevers kreativitet.

Studien vår står litt i stil med studien til Bonotto og Santo (2015), ettersom deres studie viser liten økning i originalitet, mens vår viser en tydelig større økning. Her kan en tenke seg til at elevenes originale svar kan ha vært påvirket av elevenes kompetanse i faget. Studien vår viser en tydelig sammenheng her, og har noe sammenfallende resultater, i originale svar, med

Bonotto og Santo (2015). Dog kan en problematisere vår studie med tanke på antall deltakere, og hvorvidt det faktisk er måloppnåelsen som er grunnen til et økt resultat. Ettersom både middels og lav måloppnåelse har generert færre originale svar, er det nærliggende å tenke at det er grunnet høy måloppnåelse at denne gruppen har generert flere originale svar.

Hvorvidt gruppestørrelsen også har en påvirkning, kan vi ikke vite, ettersom det ikke er foretatt noen etter-test, eller test med helt homogene grupper. Det kan være nærliggende å tenke at antall originale svar er noe annerledes, om gruppene hadde vært helt like. Dette er noe vi anser som en svakhet i oppgaven, og kan igjen gjenspeile hvorvidt resultatene er direkte valide eller ikke. Dog er det viktig å trekke frem at antallet originale svar, er et gjennomsnitt fordelt på hver enkelt elev, som viser at elever som fikk eksempel presenterte flere originale svar i gjennomsnitt.

## 5.4 Måloppnåelse mot hverandre

Våre data og analyser, viser at elever med høy måloppnåelse presterer tydelig bedre, enn andre elever i sine respektive grupper. Dette kan være nærliggende å tenke, og er ikke så overraskende, ettersom de presterer bedre i matematikk generelt. Vår studie viser at elever med høy måloppnåelse i gruppen uten eksempel skårer bedre på flyt, fleksibilitet og originalitet, med gjennomsnitt på henholdsvis 8,33, 4 og 3. Middels og lav måloppnåelse har et gjennomsnitt på henholdsvis 5, 2 og 1, og 3,33, 1 og 0. Undergruppene med lav og høy måloppnåelse har like mange deltakere, 3, mens middels har 8 deltakere. En kan argumentere for at dataene til middels måloppnåelse er noe mer korrekt, ettersom det er flere som deltar. Dette er åpenbart diskutabelt, ettersom at flere, og like mange, deltakere i hver undergruppe kunne endret resultatene. Elever med høy måloppnåelse og med eksempel fikk skåren 11, 4 og 6, på henholdsvis flyt, fleksibilitet og originalitet. Dette er høyere enn middels og lav måloppnåelse i samme gruppe, med henholdsvis 4,5, 1,75 og 1, og 1,67, 1,33 og 0,33.

Som sagt viser vår studie at elever med høy måloppnåelse, presterer tydelig bedre enn elever med lavere måloppnåelse, til tross for at samtlige elever i utgangspunktet ikke skulle ha erfaring med tematikken fra tidligere. Bonotto og Santo (2015, s. 115-117) fant at elevene fra skole nummer to, som hadde bedre akademiske resultater i matematikk, genererte tydelig flere problemer, og hadde flere originale problemer, enn den andre skolen. Her ser vi likheter med vår studie, som også viser at elever med høyere måloppnåelse i matematikk, presterer bedre i både antall problemer og antall originale problemer, dette står i tråd med Krutetskii (1986). Her er det verdt å merke seg at Bonotto og Santo (2015, s. 115) sin studie hadde klart flere deltakere, noe som kan argumenteres for at gir mer korrekte resultater. Sett i lys av vår oppgave, kan det tenkes at våre resultater hadde vært noe annerledes hadde det vært samme antall deltakere.

Studien vår har noe sammenfallende resultater med Bonotto og Santo (2015), ettersom begge studiene viser flere genererte problemer for høyere måloppnåelse og akademiske resultater. Her kan en tenke seg til at det nok er grunnet kompetansen i faget, at deltakerne har generert flere problemer, dog kan ikke dette nødvendigvis sies med sikkerhet. Vår studie viser en tydelig sammenheng mellom måloppnåelse og antall genererte problemer, og det kan derfor argumenteres at dette er grunnen. Dog kan en også problematisere dette ved å dra frem antall deltakere, noe som kan vise til en svakhet i vår studie. Her kan en argumentere for at studien vår ikke har nok valide resultater, ettersom at det er relativt få deltakere. Dog viser vår studie en sammenheng mellom måloppnåelse og antall problemer, ut fra et gjennomsnitt per deltaker. Studien har ikke et stort gap i deltakere i hver gruppe, og det kan derfor tenkes at resultatene viser noe reliabilitet ut fra deltakere og resultater per deltaker i gjennomsnitt.

Bonotto og Santo (2015, s. 115) fant også at elevene fra skole en genererte flere sannsynlige problemer med nok informasjon, sammenliknet med skole to. De fant også at elevene fra skole en genererte færre sannsynlige problemer uten nok informasjon, enn skole to. Dette er ulikt våre funn, som tilsier at høyere måloppnåelse generer flere problemer med nok informasjon. Om dette er grunnet kompetansen i matematikk, eller hvorvidt elevene var fokuserte på å tilby nok informasjon, er vanskelig å vite ettersom det ikke er foretatt intervjuer i vår studie. Bonotto og Santo (2015, s. 115) sin studie fant at elevene fra skole en genererte en mindre andel usannsynlige problemer, noe vi forstår som uløselige problemer, sammenliknet med skole to. Her har skole en 19% uløselige, og skole to 29%. Noe som igjen er ulikt vår studie, hvorpå elevene med høyere måloppnåelse har generert færre uløselige problemer i snitt, sammenliknet med middels og lav. Dog er dette gapet noe større for elevene med eksempel, sammenliknet med gruppen uten eksempel. 0,75 uløselige oppgaver i snitt for elever med høy måloppnåelse i gruppen med eksempel, mot 2,25 for elever med middels måloppnåelse. Et interessant funn er at elevene med lav måloppnåelse med eksempel genererte i snitt 0,33 uløselige problemer, tydelig færre enn middels og høy, som er sammenfallende med Bonotto og Santo (2015, s. 115) sine resultater. Dette er ikke tilfellet i vår studie for gruppen uten eksempel, hvor lav måloppnåelse genererte flest uløselige problemer, middels med nest flest og høy med færrest.

Ser en på resultatene nevnt i avsnittet over viser vår studie noe sammenfallende med Bonotto og Santo (2015) sine resultater. Hvorvidt elevene genererer flere problemer med nok informasjon eller ikke, kan være grunnet kompetansen i faget, og at de faktisk generer problemer de tenker at de selv kan generere. Det kan være nærliggende å tenke at dette er grunnen til at de genererte flere sannsynlige problemer med nok informasjon. Drar en frem gruppen med eksempel, med tanke på antall uløselige problemer, er dette også ulikt Bonotto og Santo (2015) sine resultater. Her ser en at elevens måloppnåelse kan ha en sammenheng med antall uløselige problemer. I gruppen med eksempel og lav måloppnåelse, er det sammenfallende resultater med Bonotto og Santo (2015). Om dette er grunnet eksemplet eller ikke, er vanskelig å vite. Det kan tenkes at eksemplet har vært med på å påvirke elevene i den grad at de har generert færre uløselige problemer. I gruppen uten eksempler, har alle tre måloppnåelsene resultater som ikke sammenfaller med Bonotto og Santo (2015). Her tenker vi at dette kan være grunnet måloppnåelsen i faget, og at elevene har, med stigende måloppnåelse, kompetanse til å generere problemer som kan anses som løselige. Om dette er grunnet måloppnåelsen eller noen andre faktorer, er det vanskelig å argumentere for.

Sett i sammenheng med studien til Bonotto og Santo (2015) er det noen sammenfallende likheter, som kan settes opp mot vår studie. Et argument mot dette, er hvorvidt antall deltakere i deres og vår studie, har hatt innvirkning på resultatene. Deres data kunne sett annerledes ut hadde de hatt samme antall deltakere som i vår studie, og motsatt. Her er det også verdt å nevne at Bonotto og Santo (2015) sammenliknet to skoler mot hverandre, uten noen tiltak, mens vi gjennomførte en studie med en kontrollgruppe og en eksperimentgruppe. Det at eksperimentgruppen hadde mer sammenliknbare resultater med Bonotto og Santo (2015), sammenliknet med kontrollgruppen, er interessant. Dette kunne gjerne vært undersøkt med samme gruppestørrelse, for å se om det blir noenlunde samme resultater, eller om det blir helt ulikt.

## 5.5 Resultatene i lys av utforskende undervisning

Van Harpen og Sriraman (2013) fant, gjennom sin studie, at det ser ut til at problemgenerering ikke er et etablert element i klasserommet. I likhet med vår studie, hvorpå vi har fått informasjon direkte fra lærerne i klassene om at problemgenerering ikke er noe de har gjort tidligere, viser resultatene til at dette ikke er et område som er brukt mye. Elevene har generert og presentert problemer, dog befinner hoveddelen av dette seg i kategorien «fire regnearter», en kategori elevene skal være godt kjent med i den alderen de er i. Flere elever har generert ingen problemer, eller veldig få, og noen elever har generert problemer i en tydelig større grad. Her kan vi også se at elevene som presenterte flest problemer, hovedsakelig hadde høy måloppnåelse i matematikk, samt at hoveddelen av problemene befinner seg i kategorien «fire regnearter». Ettersom problemgenerering er et relativt ukjent område i undervisningen for elevene, kan en tenke seg til at det er derfor de genererer såpass «få» problemer, og hovedsakelig problemer innenfor «fire regnearter». Samtidig er det nærliggende å tenke at elevene ikke presterer best i møte med en ukjent undervisningsform. Van Harpen og Sriraman (2013) fant også at elevene slet med å generere vanskelige problemer, selv med oppfordring i forkant. Vår studie viser noen likheter til dette, ettersom hoveddelen av problemene kan ansees som relativt greie å løse for noen med deres kunnskaper i matematikk. Med dette sagt, er det flere problemer som også kan ansees som relativt vanskelige å løse, ettersom vi selv måtte diskutere frem og tilbake for å komme frem til en entydig løsning. Elevenes nivå og kunnskaper i matematikk kan muligens forklare hvorfor det er slik, men møtet med problemgenerering som et annerledes element kan også forklare det. Studien viser at elever med lav måloppnåelse ikke genererer like mange problemer generelt, som elevene med høyere måloppnåelse. Dette kan knyttes til hvorvidt de forstår hva de skal gjøre, hvordan de skal gjøre det og hvorvidt noe nytt kan virke avskrekkende. Om elevene har hatt mulighet til å generere flere, og vanskeligere, problemer, etter noe tid med problemgenerering i undervisning, hadde vært interessant og funnet ut. Van Harpen og Sriraman (2013) poengterer at forskjellen mellom deres grupper delvis kan forklares ved at de har lært om ulike temaer i matematikk. I vår studie har det vært en randomisering av elever fra to byskoler innenfor samme kommune, som bruker samme læreplan (LK20 for å være konkret). Dette kan være med på å styrke vårt resultat, ettersom at ulikheter i resultatene ikke nødvendigvis har en direkte link til ulikt innhold i undervisningen. Med dette sagt, kan vi ikke garantere at hver enkelt elev har hatt nøyaktig samme undervisning, fått med seg nøyaktig det samme i undervisningen og sitter med samme forkunnskaper i forkant av testen.

Kokinov et al. (1997) sin studie fant at elever som fikk presentert et hint, og fikk konkretisert relevansen av hintet til oppgaven, presterte dårligere enn kontrollgruppen. Gruppen som fikk presentert et hint, uten noen form for forklaring til relevansen til oppgaven, presterte derimot bedre enn kontrollgruppen. Vår studie har et eksempel som er presentert, uten noen form for konkretisering av relevansen, dog kan det være nærliggende å se sammenhengen med oppgaven. Studien vår viser også at elever som har fått presentert et eksempel presterer bedre, på generell basis, når det kommer til kreativitet i genereringen av problemer. Her har eksperimentgruppen en høyere gjennomsnittlig skåre, på de tre kjerneelementene i kreativitet, sammenliknet med kontrollgruppen. Ser en eksemplet vårt som en form for hint, uten en eksplisitt forklaring til relevansen, er våre resultater sammenfallende med Kokinov et al. (1997, s. 5) sine resultater. En kan argumentere for at vår studie har et noe mer selvforklarende hint (i form av eksempel), sammenliknet med Kokinov et al. (1997), og elevene derfor presterte bedre. Ettersom det aldri ble spesifisert at hintet skulle brukes, var noe relevant eller inneholdt noe informasjon som kunne hjelpe elevene med generering av problemer, står vi ved sammenlikningen. En kan dra frem hvorvidt det er forskjeller ettersom Kokinov et al. (1997) foretar en studie på problemløsning, og vår studie omhandler problemgenerering, og at det derfor ikke er sammenliknbart. Elevenes generering av problemer, med eksempel, kan ha blitt påvirket av andre faktorer også, men studien har vist at noen problemer er en indirekte kopi av hintet, med noen endringer slik at det ikke skulle bli helt likt. Ut fra dette, kan det tenkes at noen elever har brukt eksempelet for å generere problemer, men det kan ikke definitivt sies at det er blitt gjort hos alle som fikk det presentert. Elevene ble heller ikke begrenset til å kun bruke det som står på arket, samt at det ikke forekom noen beskjed angående hva som ikke var lov til å bruke. Dette kan også ha påvirket elevene til å generere problemer, uten at de spesifikt har gått for bruken av hintet. Dog er det samme gjennomføring i begge gruppene, for å få et mest mulig likt scenario, slik at hintet skulle være det eneste som skilte gruppene fra hverandre. Ut fra dette, kan en tenke at hintet har påvirket elevenes problemgenereringsevne, men vi kan ikke si spesifikt at det har gjort det. Ved bruk av intervju av elevene, kunne det kanskje vært mer tydelig om det ble brukt eller ikke.

Problemgenerering som en del av utforskende undervisning kan, alt etter hvordan en planlegger og legger opp undervisningsøkta være en sentral del av en økt. Elevenes evne til å stille, men ikke minst løse matematiske problemer kan dermed bli svært viktig. Som Silver (1994, s. 23) skriver, henger disse to sammen. Derfor kan det være nærliggende å anta at resultatene for studien vi har gjennomført, angående problemgenerering, i alle fall i noen grad kan overføres til problemløsning, som også er en mulig sentral del av utforskende undervisning.

Ettersom en tydelig kan se at det er elevene som jevnt over presterer høyest, og har høyest måloppnåelse, som gjør det desidert best (per gjennomsnitt) er dette noe en må tenke på når en planlegger et opplegg med utforskende undervisning. Gruppesammensetninger kan dermed bli viktig. En kan argumentere for å lage både homogene og heterogene grupper basert på måloppnåelse. Homogene grupper, hvor elevene plasseres med andre elever som presterer på tilnærmet samme nivå vil muligens kunne utfordre flere på gruppen. På den andre siden vil dette kreve større tilpasning av undervisningsopplegget, slik at alle grupper blir utfordret på sitt nivå. Det er ikke gitt at oppgaven som utfordrer høyt presterende elever i det hele tatt vil være løselig for elevene med lav og middels måloppnåelse, slik som vi ser det fra vår data. Ved å lage heterogene grupper kan en gjøre tilpasninger, og sørge for å lage oppgaver som kan fange hele spekteret av gruppen. I tillegg kan høyt presterende elever på gruppen hjelpe elever med middels og lav måloppnåelse, slik at de kan strekke seg lenger. Dette er i tråd med hva Johnson et al. (1996, s. 66) skriver. Det vil naturligvis være fordeler og ulemper ved begge måter å dele inn grupper for utforskende undervisning. Hvorvidt en gruppesammensetning, både homogen og heterogen, er det beste for de fleste, kan diskuteres. En elevs evne til å samarbeide med noen på sitt eget nivå, kan tenkes å både være en hindring og en mulighet for å utfordre hverandre. Elevers evne til å samarbeide med noen som er på et annet nivå, kan også muligens få elevene til å føle seg for dårlige sammenliknet med de med høyere måloppnåelse, og for de med høy måloppnåelse kan det føles ut som de er satt til å gjøre lærerens jobb. Elevene kan kanskje også føle en form for mestring av å lære bort, og lære av, andre med ulik måloppnåelse, og derfor dra nytte av en heterogen gruppesammensetning. Dog går dette igjen over i spekulasjoner, og en kan nok argumentere både for og imot begge delene. Grupper er selvfølgelig ikke et kriterium for utforskende undervisning, og en kan naturligvis gjøre de samme praktiske betraktningene for individuelt utforskende arbeid.

I utforskende undervisning kan det være et poeng at elever skal finne sitt eget tema, og lage egne oppgaver. Resultatene til denne studien viser at det kan være hensiktsmessig å gi eksempeloppgaver til elevene, slik at de har denne som veiledning.

Når en ser på originalitetsskåren for kontroll og eksperimentgruppen, så ser en et paradoks. Eksperimentgruppen, som alle hadde samme eksempel og utgangspunkt, produserte klart flere originale oppgaver enn kontrollgruppen, som ikke hadde noe annet utgangspunkt enn sin egen kreativitet for å produsere oppgaver. Uten å se på data for studiet hadde vi i forkant sett for oss at et eksempel ville samle gruppen rundt en felles tankemåte å generere oppgaver på, slik at dette kunne påvirket eksperimentgruppen til å produsere like oppgaver, som ville ført til en lavere originalitetsskåre. Dermed tenkte vi at dette var en av kategoriene hvor kontrollgruppen ville hatt en fordel, ettersom de kun hadde sin egen kreativitet å belage seg på. Dette viste seg å ikke stemme, og realiteten var snarere tvert imot.

## 5.6 Innvendinger mot studien

Her skal vi ta for oss innvendinger mot studien, som kan ha vært med på å påvirke utfallet og resultatene våre. Dette er for å belyse vår forståelse av effekten de kan ha hatt, samt for å understreke at det er flere faktorer som kan ha spilt inn. Her forekommer det noen faktorer vi tenker er mest essensielle, dog betyr ikke det at det er alle faktorene som har hatt påvirkning. De åpenbare faktorene som dagsform, kompetanse, nye ansikter i klasserommet, kjønnsfordeling og diagnoser tar vi ikke stilling til her. Dette er grunnet at de kan ha en tydelig forskjell, men studiens natur gjør det ikke hensiktsmessig å drøfte rundt dette.

Vårt valg av eksempel til hver enkelt oppgave kan ha hatt en påvirkning, dog er dette vanskelig å si. Det kan også argumenteres for at vi burde brukt en annen form for eksempel, men etter litt diskusjon falt vi på det eksempelet vi har med. Elevene kan ha blitt påvirket både i positiv og negativ retning her, dog kan det ikke sies hvilken form for eksempel som hadde vært lurest og valgt. En idé hadde vært og gjennomført studien med ulike grupper med ulike eksempler, sammenliknet opp mot en kontrollgruppe, for å se hvilke eksempler som eventuelt har hvilken effekt. Grunnet størrelsen på utvalget lot ikke dette seg gjøre, og vi måtte velge et eksempel. Den store innvendingen her, er hvorvidt vi valgte den beste formen for eksempel, og i hvilken grad dette har påvirket dataene. Ved bruk av intervju kunne dette kanskje latt seg belyses, men dette ble ikke gjort.

Som nevnt flere ganger i studien, har deltakerantallet vært noe lavt og kan ha påvirket utfallet, noe vi er klar over. Studien hadde originalt tenkt å ha 35 elever totalt, 17 i en gruppe og 18 i den andre. Dette er som nevnt i tråd med hva Borg og Gall (1979) skriver angående gruppestørrelser. Her hadde vi håpet at alle elever kunne deltatt, og at et frafall på noen få elever fortsatt kunne brukes. Ettersom frafallet ble større enn vi så for oss, ble gruppene på 11 og 14 elever. Siden gruppestørrelsene er noe mindre enn hva som anses som nok, kan det argumenteres for at studiens resultater ikke er helt korrekte. Etter diskusjoner oss mellom, ble vi enige om å gjennomføre studien på den tenkte måten, og argumentere for resultatene samt belyse at vi er klar over påvirkningen dette kan ha. Studien drar frem resultatene ut fra gjennomsnittlig skåre per elev i hver gruppe, samt ut fra måloppnåelse i hver gruppe. Her drar vi frem dataene i lys av teori, dog har teorien noe annerledes gruppestørrelse, noe vi spesifiserer at kan gjøre resultatene noe annerledes. Det hadde vært ønskelig og gjennomført studien med en tydelig større gruppe, men dette lot seg ikke gjøre siden dette var utvalget vi hadde mulighet til å gjennomføre studien på. Vi er fult klar over at resultatene at dataene er diskutable ut fra gruppestørrelsen, men står fortsatt ved at validiteten og reliabiliteten til studiet gjør funnene relevante.

Gruppene har noe ulik fordeling av måloppnåelse, hvorpå middels er den som har flest, og det kan derfor også argumenteres for at dette har hatt påvirkning på dataene våre. Her ønsker vi igjen å dra frem at skåren er et gjennomsnitt for hele gruppen, og at det derfor kan anses som relevant. Grunnet ulik fordeling av måloppnåelse, har vi også sett resultatene i lys av dette, og understreker at gruppestørrelsene er noe ulike. I fordelingen av gruppene tok vi utgangspunkt i at det skulle være homogene grupper, ut fra måloppnåelse, og tok derfor en randomisert fordeling med dette i bakhodet. Her var tanken at fordelingen skulle sørge for cirka like mange av hver måloppnåelse skulle være i hver gruppe. Ettersom at det var frafall av flere elever, har dette også påvirket andelen av de ulike måloppnåelsene i gruppene. Dette er en faktor vi er klar over at kan ha hatt en påvirkning på dataene, og tenker samtidig at resultatene kunne vært noe annerledes om det var som planlagt. Resultatene er drøftet i lys av både gruppene, og måloppnåelse, hvorpå vi understreker at størrelsene på gruppene kan ha hatt en påvirkning. Det hadde vært ønskelig og gjennomført samme studie, med større grupper og mer homogene fordelinger av måloppnåelse i hver gruppe, men dette lot seg dessverre ikke gjøre.

Oppgavene, og forskningsmetoden, er hentet direkte fra Van Harpen og Sriraman (2013) sin studie, sett bort fra oppgave 2 som ble modifisert. Deres studie ble foretatt på eldre elever, som kan ha mer kompetanse i matematikk, og det kan derfor argumenteres at dette er en eventuell feilkilde i vår studie. Her har vi ansett oppgavene og forskningsmetoden som anvendelig til flere aldre, og derfor gjennomført studien på yngre elever. Ettersom oppgavenes natur ikke krever at elevene skal komme frem til et svar med spesifikk regning, men heller skape problemer selv, tenkte vi at dette fint kunne brukes. Det er verdt å nevne at oppgave 2 ble modifisert, for å tilpasses elever på 7. års-trinn, ettersom den originale oppgaven ble ansett som for vanskelig. Dette er en til faktor som kan ha hatt påvirkning på resultatene, dog ble den modifisert for å tilpasses kompetansen til elevene, samt at den, etter vårt skjønn, fortsatt passer inn i frihetsgraden som den originale oppgaven var i.

Det siste punktet vi ønsker å belyse, er elevenes erfaringer med utforskende undervisning. Her fikk vi belyst at begge elevgruppene hadde lite til ingen erfaring med dette. Hvorvidt dette er positivt eller negativt med henhold til vår studie, er vanskelig å si. Hadde elevene hatt omfattende erfaring med dette, kunne resultatene våre vært annerledes. Vi anså dette som en positiv faktor, ettersom deltakerne hadde samme erfaring med dette, og at det ikke er et åpenbart skille mellom noen elever. Hvorvidt elevene opplevde testen som vanskelig grunnet dette, vet vi ikke, men det kan være. Her kan det tenkes at resultatene er påvirket av denne faktoren, men siden det er likt for alle elevene, ser vi på dette som en positiv faktor.

# 6 Avslutning og veien videre

Mastergradsavhandlingen vår søkte etter å besvare en problemstilling og fire forskningsspørsmål, ved gjennomføring av en studie. Dataene våre har blitt analysert og drøftet i henhold til forskningsspørsmålene, og vi vil forsøke å oppsummere avhandlingen med utgangspunkt i problemstillingen vår. Grunnet dette gjentar vi de her:

*I hvilken grad påvirkes elevers kreativitet i problemgenerering ved bruk av eksempel?*

1) *I hvilken grad skårer elever på kategoriene fleksibilitet, flyt og originalitet, når de får presentert et eksempel i en oppgave?*
2) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de får presentert et eksempel?*
3) *I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de ikke får presentert et eksempel?*

### I hvilken grad påvirkes elevers kreativitet i problemgenerering ved bruk av eksempel?

Gjennom case-studien ønsket vi å besvare denne problemstillingen. Hvordan elever presterer i problemgenerering kan nok påvirkes av forskjellige faktorer. Om dette er ved bruk av eksempel, er noe som dannet grunnlaget for både problemstillingen og det første forskningsspørsmålet. Hvorvidt elevenes måloppnåelse i matematikk er en slik vesentlig faktor, var med på å danne de to siste forskningsspørsmålene. Gjennom analyse og drøfting, har vi forsøkt å tydeliggjøre at elevenes skåre innenfor kjerneelementene i kreativitet er høyere når de har fått presentert et eksempel. Hvorvidt dette faktisk er grunnen, vet vi ikke helt sikkert, dog er det nærliggende å tenke det. Gruppene ble fordelt noenlunde homogent, men grunnet frafall av deltakere ble gruppene påvirket. Størrelsen på eksperimentgruppen og kontrollgruppen kan også ha spilt en faktor i resultatene våre.

Via drøfting og analyse, kan en se at elevene med eksempel presterer klart bedre innenfor hvert kjerneelement, her skårer de til og med dobbelt så bra på to punkter. Kjerneelementene er brukt for å belyse hvorvidt hver enkelt elevs besvarelser kan ansees som kreative, og hvorvidt hvert enkelt svar er originalt eller ikke. Studien tar, som sagt, utgangspunkt i tre hovedelementer av kreativitet, dog kan ikke dette avskrives som all form for kreativitet i matematikk. En kan tenke seg at kreativitet kan innebære andre ting også. Ut fra rammeverket studien tar for seg, er det nærliggende å tenke at det stilles andre krav til kreativitet når deltakerne er eldre og har mer matematisk kunnskap. Eksempelvis hvorvidt et generert problem er trivielt eller ikke, her har både vi og rammeverket tatt en subjektiv vurdering ut fra hva som kan anses som «for enkelt» for deltakerne.

Vi har ønsket å belyse hvorvidt elevers kreativitet blir påvirket av et eksempel til oppgavene. Vår tanke gående inn i gjennomføringen, var at elevene ble å bruke eksempelet på ulike måter relativt hyppig. Gjennom analysen viste det seg at fåtallet bevisst brukte eksempelet med en liten endring. Vi har drøftet hvorvidt det er eksempelet som har gitt gruppen høyere skåre, eller om det er grunnet et ulikt deltakerantall. Drøftingen ønsker å belyse hvorvidt det er eksempelet, eller ikke, samtidig som vi drar frem gruppestørrelsene og at vi er klare over at dette kan ha en påvirkning. Analysen viser, som nevnt tidligere, at elevgruppen med eksempel presterte bedre innenfor alle tre elementene. Deltakerne med eksempel viser en større grad av kreativitet innenfor problemgenerering. Hvorvidt dette er grunnet eksempelet, eller ulik fordeling av måloppnåelse er vanskelig å si.

## I hvilken grad skårer elever på kategoriene fleksibilitet, flyt og originalitet, når de får presentert et eksempel i en oppgave?

Elevenes fleksibilitetsskåre er beregnet ut fra rammeverket, og hvor mange kategorier de genererte problemer i. Her telles ikke oppgavene som ansees som uløselige. Elevene som fikk presentert et eksempel til hver oppgave, fikk en gjennomsnittlig fleksibilitetsskåre per elev på 2,455, samt en median på 2. Gruppen viser her at de genererer problemer som fordeles på litt over 2 grupper. Sammenliknet med kontrollgruppen, som fikk 2,071 i fleksibilitetsskåre, er det ikke en stor forskjell. Eksperimentgruppen viser noe høyere skåre, selv med 3 færre deltakere. Analysen og drøftingen vår tar for seg dette, og forsøker å tydeliggjøre at det kan være grunnet eksempelet, samt kanskje grunnet gruppestørrelsen.

Flytskåre er beregnet som antall løselige problemer som er generert. Her er det beregnet flytskåre i gruppen, og så et gjennomsnitt per elev. Analysen viser at eksperimentgruppen fikk en gjennomsnittlig flytskåre per elev på 6,091, sammenliknet med kontrollgruppen på 5,357, viser dette en tydelig forskjell. Gjennom drøftingen ønsker vi å belyse hvorvidt gruppenes forskjellige flytskåre er grunnet eksempelet eller ikke, og som vi understreker er dette vanskelig å vite. Elevenes skåre kan ha hatt flere ulike faktorer som spiller inn.

Studien tar til slutt for seg originalitet, hvor hvert problem blir kategorisert som originalt hvis det er færre enn 10% som har generert et liknende problem. Gjennom analysen fant vi at eksperimentgruppen genererte 2,364 originale problemer per elev, og kontrollgruppen genererte 1,214. Videre analyse viser at gruppen som fikk presentert et eksempel også genererte en større andel originale problemer, 23,77% i eksperimentgruppen mot 19,32% i kontrollgruppen. I drøftingen ønsker vi å fremheve at selv om eksperimentgruppen genererte nært dobbelt så mange originale problemer, er ikke andelen originale problemer dobbelt så høy. Dette er med på å vise at eksperimentgruppen har generert tydelig flere problemer totalt, sammenliknet med kontrollgruppen.

Studien ønsker å besvare forskningsspørsmålet gjennom analyse og drøfting. Vi har her kort oppsummert hvordan eksperimentgruppen har skåret, sammenliknet med kontrollgruppen. Ut fra dette viser resultatene at elevgruppen som fikk presentert et eksempel skårer høyere på samtlige kategorier. Presentasjonen av eksempel, ut fra resultater, analyse og drøfting, kan vise til at elevene skårer høyere innenfor kategoriene. Elevgruppen har ikke en mye høyere skåre på hver enkelt kategori, dog er den høyere til tross for færre antall deltakere.

## I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse, når de får presentert et eksempel?

Analysen og drøftingen tar for seg skårene til elevene, ut fra hvilken måloppnåelse de har i matematikk. Her har vi ønsket å undersøke hvorvidt måloppnåelsen kan være en faktor for hvordan elevene skårer, samt kanskje forklarer skårene. Her har analysen tatt for seg både gruppen med eksempel, og gruppen uten eksempel.

I henhold til flytskåre har analysen vist at elevene med høy måloppnåelse skårer tydelig høyere, enn middels og lav måloppnåelse. Innenfor fleksibilitet har elevene med høy måloppnåelse igjen skåret tydelig høyere, hvorpå middels og lav måloppnåelse skårer ganske så likt. Til slutt analyserte vi originalitet, og hvor mange originale svar per elevgruppe. Her ser vi at elevene med høy måloppnåelse har generert betraktelig flere originale svar, 6 ganger så mange som middels og lav.

I studiens analyse og drøfting ønsker vi å belyse forskjellene innad i gruppene, sett ut fra elevenes måloppnåelse. Videre drar vi frem hvorvidt høy måloppnåelse har prestert bedre eller dårligere. Resultatene våre viser at elevene med høy måloppnåelse, i gruppen med eksempel, skårer tydelig bedre på samtlige kategorier. Som nevnt over, har gruppen generert flerdobbelt antall originale svar, noe som tyder på mer kreativitet hos elevene. Drøftigen ønsker å tydeliggjøre hvorfor det er en så stor forskjell, dog er dette noe som vanskelig kan stadfestes til en spesifikk grunn. Det er nærliggende å tenke at elevene skårer bedre, grunnet deres måloppnåelse i faget. Vår tanke på tur inn i studien var at dette kunne komme til å bli resultatet, men var åpen for å ta feil.

## I hvilken grad skårer elever med høy måloppnåelse i matematikk i kategoriene flyt, fleksibilitet og originalitet, sammenliknet med middels og lav måloppnåelse når de <u>ikke</u> får presentert et eksempel?

I tillegg til å undersøke elevenes skåre i kategoriene ut fra måloppnåelse når de fikk presentert eksempel, ønsket vi å se på det samme for elever som ikke fikk presentert eksempel. Her ønsket vi å se hvorvidt måloppnåelsen i faget, hadde noe å si for skårene, selv om elevene ikke fikk eksempel. Analysen og drøftingen tar for seg de ulike gruppenes skåre, og setter dem opp mot hverandre.

Innenfor flytskåre presterer elevene med høy måloppnåelse noe bedre, sammenliknet med middels og lav måloppnåelse. Analysen viser at spranget fra lav til høy måloppnåelse er størst, noe som kan være nærliggende å tenke. Ser en på fleksibilitetsskåren til gruppene, viser elevene med høy måloppnåelse en høyere skåre her også. Analysen vår viser at fleksibilitetsskåren til elevene ut fra måloppnåelse, er noe sammenfallende mellom eksperiment- og kontrollgruppen. Originale svar per elev i snitt, ut fra måloppnåelse, viser også at høy måloppnåelse genererte flest. Her med 3 ganger så mange som middels måloppnåelse, og 6 ganger så mange som lav måloppnåelse.

Analysen og drøftingen i studien har som hensikt å ta for seg skårene fordelt på måloppnåelse, og hvordan de skårer i forhold til hverandre. Det forsøkes å tydeliggjøre forskjellene innad i gruppene, og om måloppnåelsen har noe å si for hvordan elevene skårer. Resultatene våre viser også her at elever med høy måloppnåelse skårer bedre, sammenliknet med middels og lav, men at det ikke er like stor forskjell i gruppen uten eksempel som i gruppen med eksempel. I drøftingen ønsker vi å dra frem forskjellene, og forsøke å begrunne hvorfor de finner sted. Det er nærliggende å tenke at elevenes måloppnåelse er grunnen, noe vi selv tenker. Dette er, som sagt under forrige forskningsspørsmål, noe vi hadde tenkt i forkant av gjennomføringen av studien. Analysen og drøftingen viser at det er en forskjell på skårene til elevene, ut fra måloppnåelse, hvorpå høy måloppnåelse skårer bedre enn middels og lav. Dette kan dermed gi grunnlag for å tenke at høy måloppnåelse viser frem elever som er mer kreative innenfor matematikk.

## Konklusjon

Gjennom studien vår ønsker vi å belyse ulike forhold som har innvirkning på elevers kreativitet innenfor problemgenerering. Argumentasjonen vår går på det grunnlag at elevers kreativitet er noe mer fremtredende ved presentasjonen av et eksempel til oppgaven, uten å gi direkte instruksjon om forholdet til eksempelet og oppgaveteksten. Dog kan dette være ganske innlysende, ettersom eksempelet er presentert rett etter teksten, og bruker informasjon fra teksten. Vi har ønsket å undersøke eventuelle forskjeller når det kommer til elevers måloppnåelse og kreativitet, som er presentert i drøfting og analysen. Studien viser til at elever med høy måloppnåelse presterer jevnt over bedre innenfor samtlige kategorier, noe vi har forsøkt å understreke. En kan argumentere for at dette er grunnet kompetansen i faget, ikke nødvendigvis direkte grunnet mer kreativitet blant disse elevene. Studien ønsker også å fremheve at problemgenerering er et noe neglisjert område innenfor matematikk, og understreker dette ved å gjenta at elevene har lite til ingen erfaring med utforskende undervisning fra før av. Dette kan dog argumenteres mot ved å vise til elevenes alder, og fremtidig skolegang.

Vi ønsker å konkludere med at elever som har fått presentert et eksempel til hver oppgave, presterer noe bedre når det kommer til kjerneelementene i kreativitet, samt at ut fra måloppnåelse er det elever med høy måloppnåelse som presterer best. Ved mer erfaring innenfor problemgenerering kan resultatene dog være noe annerledes. Vi vil også hevde at gruppestørrelsene har hatt en faktor når det kommer til resultatene våre, og at det ideelt sett skulle vært mer homogene grupper. Med dette sagt ønsker vi å understreke at studiens natur fremhever gjennomsnittlig skåre fordelt på hver enkelt elev, og at det derfor kan hevdes at resultatene er valide.

## Veien videre

Gjennom vårt arbeid med denne studien har vi kommet over flere elementer vi anser som interessante å forske videre på. Det første vi vil dra frem, er en noenlunde lik studie som tar for seg større grupper for å analysere eventuelle forskjeller som fremkommer. Vår studie har tatt utgangspunkt i en spesifikk form for eksempel, men en interessant studie kunne vært å undersøke effekten av ulike former for eksempler som kan brukes. Det ville også vært interessant å undersøke problemgenerering ved bruk av andre former for oppgaver, og hvorvidt ulike oppgaver har ulik effekt på kreativiteten hos elever. Analysen vår inkluderer kun tre oppgaver, likt for begge grupper, som tidligere er testet på eldre elever i andre land. Det kunne videre være interessant og studert kreativiteten hos eldre elever innunder samme læreplan. Studien vår foretar analyser av genererte problemer hos elever med lite til ingen erfaring med utforskende undervisning. Studier av elever med erfaring innenfor utforskende undervisning, kunne derfor vært interessant å undersøke.

I vår konklusjon kommer det frem at elever som fikk presentert eksempel presterte bedre, sammenliknet med elever som ikke fikk presentert eksempel. Ved å legge dette til grunn kan det argumenteres for viktigheten av eksempler i utforskende undervisning. Vi tenker at lærere bør være klare over effekten av eksempler på undervisning, og hvorfor dette skal og ikke skal brukes. Det kunne vært interessant og undersøkt læreres egne tanker og meninger angående bruken av eksempler i utforskende undervisning. Om dette er noe de er klare over effekten av, eller er ikke.

Vår mastergradsavhandling ønsker å belyse hvorvidt eksempler har noen effekt på elevers kreativitet, både grupper mot hverandre, men også når det kommer til elevers måloppnåelse. Det er ønskelig at avhandlingen vår kan fungere som et utgangspunkt for nevnte forslag til videre forskning.

Avslutningsvis ønsker vi å trekke frem viktigheten av utforskende undervisning, særlig problemløsning og problemgenerering. Som nevnt tidligere er dette implementert i flere læreplaner, inkludert LK20 i gitt grad, og viktigheten av dette innenfor læring.

# 7 Referanseliste


Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM, 45*(6), 797-810.

Balka, D. S. (1974). Using research in teaching: Creative ability in mathematics. *The Arithmetic Teacher, 21*(7), 633-636.

Bassok, M. (2003). Analogical transfer in problem solving. *The psychology of problem solving*, 343-369.

Blomhøj, M. (2016). *Fagdidaktik i matematik*: Frydenlund Academic.

Bonotto, C., & Santo, L. D. (2015). On the relationship between problem posing, problem solving, and creativity in the primary school. In *Mathematical problem posing* (pp. 103-123): Springer.

Borg, W. R., & Gall, J. P. (1979). *Educational research: An introduction* (Vol. 3.). London: Longman Publishing.

Cai, J. (2000). Mathematical thinking involved in US and Chinese students' solving of process-constrained and process-open problems. *Mathematical thinking and learning, 2*(4), 309-340.

Cai, J., & Hwang, S. (2002). Generalized and generative thinking in US and Chinese students' mathematical problem solving and problem posing. *The Journal of mathematical behavior, 21*(4), 401-421.

Cai, J., Hwang, S., Jiang, C., & Silber, S. (2015). Problem-posing research in mathematics education: Some answered and unanswered questions. In *Mathematical problem posing* (pp. 3-34): Springer.

Christoffersen, L., & Johannessen, A. (2012). *Forskningsmetode for lærerutdanningene*. Oslo: Abstrakt forl.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education: Eighth edition*. New York,London: New York: Routledge.

Creswell, J. W. (2014). *Research design : qualitative, quantitative, and mixed methods approaches* (4th ed.; International student ed. ed.). Los Angeles, Calif: SAGE.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological monographs, 58*(5), i.

Elgrably, H., & Leikin, R. (2021). Creativity as a function of problem-solving expertise: Posing new problems through investigations. *ZDM–Mathematics Education, 53*(4), 891-904.

Ellerton, N. F. (1986). Children's made-up mathematics problems—a new perspective on talented mathematicians. *Educational Studies in Mathematics, 17*(3), 261-271.

English, L., & Sriraman, B. (2010). Problem solving for the 21 st century. In *Theories of mathematics education* (pp. 263-290): Springer.

English, L. D. (1998). Children's problem posing within formal and informal contexts. *Journal for research in mathematics education, 29*(1), 83-106.

Ervynck, G., & Tall, D. (1991). Advanced mathematical thinking. In: Hingham, MA: Kluwer, Academic Publishers.

Hayes, J. R. (1989). *The complete problem solver* (2nd ed. ed.). Hillsdale, N.J: Erlbaum Associates.

Imsen, G. (2020). *Lærerens verden : innføring i generell didaktikk* (6. utgave. ed.). Oslo: Universitetsforlaget.

Jay, E. S., & Perkins, D. N. (1997). *Problem finding: The search for mechanism* (Vol. 1). Cresskill, NJ: Hampton Press.

Johnson, D. W., Johnson, R. T., Haugaløkken, O. K., & Aakervik, A. O. (1996). *Samarbeid i skolen* (3. rev. utg. ed.). Namsos: Pedagogisk psykologisk forl.

Kokinov, B., Hadjiilieva, K., & Yoveva, M. (1997). *Is a hint always useful in problem solving? The influence of pragmatic distance on context effects.* Paper presented at the Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children*. Chicago: The University
of Chicago Press.

Leikin, R. (2013). Evaluating mathematical creativity: The interplay between multiplicity and insight1. *Psychological Test and Assessment Modeling, 55*(4), 385.

Leikin, R., & Elgrably, H. (2020). Problem posing through investigations for the development and evaluation of proof-related skills and creativity skills of prospective high school mathematics teachers. *International Journal of Educational Research, 102*, 101424.

Leung, S. S., & Silver, E. A. (1997). The role of task format, mathematics knowledge, and creative thinking on the arithmetic problem posing of prospective elementary school teachers. *Mathematics Education Research Journal, 9*(1), 5-24.

Liljedahl, P., Santos-Trigo, M., Malaspina, U., & Bruder, R. (2016). *Problem solving in mathematics education*: Springer Nature.

Matsko, V. J., & Thomas, J. (2015). Beyond routine: Fostering creativity in mathematics classrooms. In *Mathematical problem posing* (pp. 125-139): Springer.

NESH. (2021). Forskningsetiske retningslinjer for samfunnsvitenskap og humaniora. 5.

Novick, L. R., & Bassok, M. (2005). *Problem Solving*: Cambridge University Press.

Piirto, J. (1999). A different approach to creativity enhancement. *Tempo, XIX, 3*(1).

Pólya, G. (1957). *How to solve it : a new aspect of mathematical method* (2nd ed. ed.). Garden City, N.Y: Doubleday.

Postholm, M. B., Jacobsen, D. I., & Søbstad, R. (2018). *Forskningsmetode for masterstudenter i lærerutdanningen*. Oslo: Cappelen Damm akademisk.

Reese, H. W. (1994). Cognitive and behavioral approaches to problem solving.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education, 14*(4), 477-488.

Robertson, S. I. (2016). *Problem solving: Perspectives from cognition and neuroscience*: Psychology Press.

Silver, E. A. (1994). On mathematical problem posing. *For the learning of mathematics, 14*(1), 19-28.

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM, 29*(3), 75-80.

Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for research in mathematics education, 27*(5), 521-539.

Silver, E. A., & Cai, J. (2005). Assessing students' mathematical problem posing. *Teaching children mathematics, 12*(3), 129-135.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics teaching, 77*(1), 20-26.

Stoyanova, E. N. (1997). Extending and exploring students' problem solving via problem posing.

Stoyanova, E. N., & Ellerton, N. F. (1996). A framework into students´ problem posing in school mathematics. In P. C. Clarkson (Ed.), *Technology in mathematics education (Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 518-525). Melbourne: Mathematics Education Research Group of Australasia.

Usiskin, Z. (2000). The development into the mathematically talented. *Journal of Secondary Gifted Education, 11*(3), 152-162.

Utdanningsdirektoratet. (2020). *Læreplan i matematikk 1.-10. trinn*. (MAT01-05). Retrieved from https://www.udir.no/lk20/mat01-05/kompetansemaal-og-vurdering/kv14?lang=nob

Van Harpen, X. Y., & Sriraman, B. (2013). Creativity and mathematical problem posing: an analysis of high school students' mathematical problem posing in China and the USA. *Educational Studies in Mathematics, 82*(2), 201-221.

Wæhle, E., Dahlum, s., & Grønmo, S. (2020). Case-studie. *Store Norske Leksikon*. doi:https://snl.no/case-studie

Zhang, L., Cai, J., Song, N., Zhang, H., Chen, T., Zhang, Z., & Guo, F. (2022). Mathematical problem posing of elementary school students: the impact of task format and its relationship to problem solving. *ZDM–Mathematics Education*, 1-16.

# Vedlegg

Under forekommer vedlegg som er brukt i studien og avhandlingen.

## Vedlegg 1: Oppgaver uten eksempel

### Oppgave 1

Det står 10 gutter og 10 jenter i kø. Lag så mange oppgaver du kan, der du bruker denne informasjonen på en eller annen måte.

**Oppgave 2**

Alle vinklene er like store, og én side i trekanten er 5 cm.
Lag så mange oppgaver/spørsmål du kan, som kan knyttes til figuren eller teksten, på en eller annen måte.

## Oppgave 3

I går hadde Kåre en bursdagsfest, og det ringte på døra ti ganger. Første gang dørklokka ringte kom det én gjest. Hver gang dørklokka ringte etter det, kom det tre gjester mer enn det gjorde forrige gang dørklokka ringte.

1- Hvor mange gjester kom det da dørklokka ringte for tiende gang? Forklar hvordan du fant svaret ditt?
2- Lag så mange oppgaver/spørsmål som du klarer, som kan knyttes til situasjonen over, på en eller annen måte.

## Vedlegg 2 oppgaver med eksempel

### Oppgave 1

Det står 10 gutter og 10 jenter i kø. Lag så mange oppgaver du kan, der du bruker denne informasjonen på en eller annen måte.

### Eksempel

Guttene og jentene står i kø for å komme inn på en restaurant. Det er plass til 3 stykker rundt et bord, og alle skal ha et bord å sitte med, men ingen skal sitte alene. Hvor mange bord trenger du?

## Oppgave 2

Alle vinklene er like store, og én side i trekanten er 5 cm.
Lag så mange oppgaver/spørsmål du kan, som kan knyttes til figuren eller teksten, på en eller annen måte.

## Eksempel

Hvis trekanten over er en av sidene til en pyramide, hvor stor er grunnflata?

## Oppgave 3

I går hadde Kåre en bursdagsfest, og det ringte på døra ti ganger. Første gang dørklokka ringte kom det én gjest. Hver gang dørklokka ringte etter det, kom det tre gjester mer enn det gjorde forrige gang dørklokka ringte.

1- Hvor mange gjester kom det da dørklokka ringte for tiende gang? Forklar hvordan du fant svaret ditt?
2- Lag så mange oppgaver/spørsmål som du klarer, som kan knyttes til situasjonen over, på en eller annen måte.

### Eksempel

Én femtedel av gjestene gikk rett inn med sko på, hvor mange par sko ble satt i gangen?

## Vedlegg 3: Vurdering fra NSD

# Vurdering

**Referansenummer**

897490

**Prosjekttittel**

Mastergrad i matematikkdidaktikk ved UIT

**Behandlingsansvarlig institusjon**

UiT Norges Arktiske Universitet / Fakultet for humaniora, samfunnsvitenskap og lærerutdanning / Institutt for lærerutdanning og pedagogikk

**Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)**

Per Øystein Haavold, per.oystein.haavold@uit.no, tlf: 77645587

**Type prosjekt**

Studentprosjekt, masterstudium

**Kontaktinformasjon, student**

Ole Kvikstad, okv005@uit.no, tlf: 48258107

**Prosjektperiode**

01.01.2022 - 15.05.2022

**Vurdering (1)**

---

**09.02.2022 - Vurdert**

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet 09.02.2022 med vedlegg, samt i meldingsdialogen mellom innmelder og Personverntjenester. Behandlingen kan starte.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 15.05.2022.
LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra foresatte til behandlingen av personopplysninger om barna. Vår

vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte/foresatte kan trekke tilbake.

Lovlig grunnlag for behandlingen vil dermed være foresattes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

PERSONVERNPRINSIPPER
Personverntjenester vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at foresatte får tilfredsstillende informasjon om og samtykker til behandlingen

formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål

dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet

lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Personverntjenester vurderer at informasjonen om behandlingen som de registrerte og deres foresatte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18) og dataportabilitet (art. 20).

Vi minner om at hvis en registrert/foresatt tar kontakt om sine/barnets rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
Personverntjenester legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til oss ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å melde:

https://www.nsd.no/personverntjenester/fylle-ut-meldeskjema-for-personopplysninger/melde-endringer-i-meldeskjema. Du må vente på svar fra oss før endringen gjennomføres.

OPPFØLGING AV PROSJEKTET
Personverntjenester vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Kontaktperson hos oss: Belinda Gloppen Helle

Lykke til med prosjektet!