



Universitetet  
i Stavanger

DET HUMANISTISKE FAKULTET

## MASTEROPPGAVE

|   |   |
|---|---|
| Studieprogram:<br>Utdanningsvitenskap, profil<br>Matematikdidaktikk   | vårsemesteret, 2015<br><br>Åpen   |
| Forfatter: Johanne Hovig  | .....<br>(signatur forfatter)   |
| Veileder: Elin Reikerås, Mary Billington  |   |
| Tittel på masteroppgaven: "Det var en glimrende feil!"<br>Implikasjoner fra feilmønstre i oppstilte og uoppstilte regneoppgaver fra<br><i>Kartleggingsprøve i regning for 2. trinn.</i><br><br>Engelsk tittel: "Two wrong don't make a right, but three lefts do"<br>Implications from 2.grades pupils incorrect responses in a diagnostic test in<br>arithmetic. |   |
| Emneord:<br>Kartleggingsprøven 2. trinn, telle- og<br>regnestrategier, feilmønstre i addisjon og<br>subtraksjonsoppgaver, kartlegging,<br>matematikdidaktikk, grunnskolen   | Antall ord: .....26 980.....<br>+ vedlegg/annet: ...3 842....<br><br>Stavanger, ...12/06/2015...<br>dato/år |

## Forord

Arbeidet med masteroppgaven i Utdanningsvitenskap profil Matematikdidaktikk har vært en spennende og særdeles lærerik prosess. Det er et privilegium å få lov å fordype seg faglig over et så langt tidsrom. Det å skrive en masteroppgave medfører også prosessen med å gå inn i 'masterbobla', og i den anledning er det flere som skal berømmes og takkes for tålmodighet og innsats. Jeg vil benytte anledning her i forordet til å rette oppmerksomheten mot noen av dere.

Først av alt takk til mine veiledere Elin Reikerås og Mary R. Billington, ved Nasjonalt senter for Leseopplæring og Leseforskning. Deres oppmuntring, konstruktive kritikk og engasjement har vært uvurderlig i dette arbeidet. Elin takkes spesielt for entusiasme ved oppstart og igangsetting av prosjektet. Mary som avløste Elin i april takkes for tålmodigheten og innsatsen som ligger bak å sette seg inn i et prosjektet på kort tid. Nasjonalt senter for Leseopplæring og Leseforskning ved Sentersjef Åse Kari Hansen Wagner takkes også for at det faktisk ble mulig for meg å bli veiledet av disse. Fagmiljøet på Lesesenteret har vært en berikelse i arbeidet med oppgaven. En takk rettes også til Stavangerprosjektet for tilgang til, og bruk av deres datamateriale.

Arbeidet med oppgaven har til tider vært utfordrende, frustrerende og seigt. Jeg har selvinnsett nok til se at det også har gjeldt meg som person, derfor rettes en stor takk til venner, medstudenter og familie for tålmodighet og gode ord på veien. Nå når dette arbeidet er avsluttet skal det igjen bli tid til AlleMannAlle og andre, og alt det som er satt på vent på grunn av oppgaveskriving. Læringsutbytte fra oppgaveskrivingen har vært stort, og selv om det ligger utallige timer med arbeid bak er jeg glad for at jeg valgte å skrive om et relativt nytt felt for meg.

Jeg avslutter med Sokrates sine ord som har aldri vært sannere for meg:

*Det eneste jeg vet er at jeg intet vet.*

Johanne Hovig, 12. Juni, Stavanger.

## Sammendrag

Studien hadde fokus på indikasjoner feilmønster i oppstilte og ikke oppstilte regneoppgaver gir. Datamaterialet besto av 440 elevers *Kartleggingsprøve i regning for 2. trinn*, i Stavanger kommune. En sammenligning ble gjort mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen for å beskrive feilmønstre opp mot gruppetilhørighet. Feilmønstrene ble sett i sammenheng med oppgavekategori og kompetanse som kreves fra eleven. Oppgavekategoriene var sentrert rundt plassering av den ukjente og regnetegnet. I de ikke oppstilte regneoppgavene ble illustrasjoner kategorisert etter funksjon i oppgaven. Hovedfunn fra studien indikerer at plassering av den ukjente og regnetegn påvirker elevenes svar. Regneoppgaver som hadde regnetegnet til høyre for likhetstegnet, samt regneoppgaver hvor den første eller andre addenden var ukjent ( $5 = x + 3$  eller  $3 + x = 5$ ), gav flest feilsvar. Feilsvarene fra disse oppgavene indikerte en operasjonell forståelse for likhetstegnet. Studien finner også at informative og representative illustrasjoner er utfordrende for aldersgruppen, der informative illustrasjoner gav høyest utslag. Det var indikasjoner på at mistolkning eller feilavlesning fra illustrasjonene var hovedutfordringen, da feilmønster forbundet tellefeil var lite utslagsgivende i disse oppgavene. Oppsummert var funnene at avkoding av oppgavene som skal løses er betingende for svarene som blir gitt.

# Innholdsfortegnelse

---

|  |             |
|--|-------------|
| <b>Forord</b> .....  | <b>ii</b>   |
| <b>Sammendrag</b> .....  | <b>iii</b>  |
| Oversikt over tabeller .....   | <b>vii</b>  |
| Oversikt over figurer .....  | <b>viii</b> |
| <b>1 Innledning</b> .....  | <b>1</b>    |
| <b>1.1 Bakgrunn for valg av tema</b> .....   | <b>1</b>    |
| <b>1.2 Relevans</b> .....  | <b>2</b>    |
| <b>1.3 Forskningsbehov</b> .....   | <b>4</b>    |
| <b>1.4 Forskningsspørsmål</b> .....  | <b>5</b>    |
| <b>1.5 Oppgavens oppbygning</b> .....  | <b>6</b>    |
| <b>2 Tidligere forskning og teorier</b> .....  | <b>7</b>    |
| <b>2.1 Ulike oppgavetyper innen addisjon og subtraksjon</b> .....  | <b>7</b>    |
| 2.1.1 Oppstilte regneoppgaver .....  | 8           |
| 2.1.2 Uoppstilte regneoppgaver - tekstopp-gaver.....   | 9           |
| 2.1.3 Illustrerte tekstopp-gaver.....  | 13          |
| 2.1.4 ”Problem-size” effekten.....   | 16          |
| <b>2.2 Feilmønstre i addisjon og subtraksjonsoppgaver, i oppstilte regneoppgaver og tekstopp-gaver</b> ..... | <b>17</b>   |
| <b>2.3 Likhetstegnet (=) og elevers forståelse for tegnet</b> .....  | <b>18</b>   |
| <b>2.4 Kompetanse</b> .....  | <b>20</b>   |
| 2.4.1 Et fleksibelt tallbegrep.....  | 21          |
| 2.4.2 Strategier.....  | 22          |
| <b>2.5 Oppsummering tidligere forskning og teori</b> .....   | <b>23</b>   |
| <b>3 Metode og analyse</b> .....   | <b>24</b>   |
| <b>3.1 Valg av metode og forskningsdesign</b> .....  | <b>24</b>   |
| <b>3.2 Datakilde - Stavangerprosjektet</b> .....   | <b>24</b>   |
| 3.2.1 Bruk av allerede eksisterende data .....   | 25          |
| <b>3.3 Utdanningsdirektoratet sin Kartlegging i regning for 2. trinn</b> .....                               | <b>25</b>   |
| 3.3.1 Mål for <i>Kartleggingsprøve i regning for 2. trinn</i> .....  | 25          |
| 3.3.2 Design av prøven.....  | 26          |
| 3.3.3 Gjennomføring av <i>Kartleggingsprøve i regning for 2. trinn</i> .....                                 | 28          |

|             |   |           |
|-------------|---|-----------|
| <b>3.4</b>  | <b>Ulike oppgavetyper .....</b>               | <b>29</b> |
| 3.4.1       | Tekstoppgaver .....                           | 30        |
| 3.4.2       | Åpne svar eller svaralternativer .....        | 31        |
| <b>3.5</b>  | <b>Analyseverktøy SPSS.....</b>               | <b>32</b> |
| 3.5.1       | Metode.....                                   | 32        |
| <b>3.6</b>  | <b>Datamaterialet og utvalg .....</b>         | <b>34</b> |
| <b>3.7</b>  | <b>Kategorisering av svar .....</b>           | <b>41</b> |
| <b>3.8</b>  | <b>Validitet.....</b>                         | <b>44</b> |
| <b>3.9</b>  | <b>Reliabilitet .....</b>                     | <b>44</b> |
| <b>3.10</b> | <b>Forskningsetiske hensyn.....</b>           | <b>45</b> |
| <b>4</b>    | <b>Resultater og analyse.....</b>             | <b>46</b> |
| <b>4.1</b>  | <b>Deskriptiv analyse .....</b>               | <b>46</b> |
| <b>4.2</b>  | <b>Side A.....</b>                            | <b>49</b> |
| <b>4.3</b>  | <b>Analyse og drøfting .....</b>              | <b>54</b> |
| 4.3.1       | Oppsummering Side A.....                      | 56        |
| <b>4.4</b>  | <b>Side B – Tekstoppgaver.....</b>            | <b>57</b> |
| 4.4.1       | Oppgave 1 .....                               | 59        |
| 4.4.2       | Oppgave 2.....                                | 61        |
| 4.4.3       | Oppgave 3.....                                | 62        |
| 4.4.4       | Oppsummering for Side B.....                  | 63        |
| <b>4.5</b>  | <b>Side C – Oppstilte regneoppgaver .....</b> | <b>64</b> |
| <b>4.6</b>  | <b>Side D – Tekstoppgaver .....</b>           | <b>68</b> |
| 4.6.1       | Oppgave 1.....                                | 71        |
| 4.6.2       | Oppgave 2.....                                | 73        |
| 4.6.3       | Oppgave 3.....                                | 75        |
| 4.6.4       | Oppgave 4.....                                | 76        |
| 4.6.5       | Oppsummering Side D .....                     | 78        |
| <b>4.7</b>  | <b>Side E – Oppstilte regneoppgaver .....</b> | <b>79</b> |
| <b>4.8</b>  | <b>Side F – Tekstoppgaver.....</b>            | <b>83</b> |
| <b>4.9</b>  | <b>Side G – Oppstilte regneoppgaver.....</b>  | <b>86</b> |
| 4.9.1       | Oppgave 4, RR5 .....                          | 90        |
| 4.9.2       | Oppgave 6, RR5 .....                          | 90        |
| 4.9.3       | Oppsummering Side G .....                     | 92        |
| <b>4.10</b> | <b>Oppsummering resultat og analyse.....</b>  | <b>93</b> |
| <b>5</b>    | <b>Diskusjon.....</b>                         | <b>94</b> |

|                            |            |
|----------------------------|------------|
| <b>5.1 Konklusjon.....</b> | <b>101</b> |
| <b>Referanseliste.....</b> | <b>104</b> |
| <b>Vedlegg.....</b>        | <b>109</b> |

## Oversikt over tabeller

|  |    |
|--|----|
| Tabell 1 Kategorisering av oppstilte regneoppgaver .....   | 8  |
| Tabell 2 Oversikt over tekstoppgaver med eksempler.....  | 12 |
| Tabell 3 Strategier i addisjon og subtraksjon.....   | 22 |
| Tabell 4 Oppgavetyper i <i>Kartleggingsprøve i regning for 2. trinn</i> – studiens utvalg .            | 30 |
| Tabell 5 Sentralmål for utvalget .....   | 36 |
| Tabell 6 Gruppeoversikt med Utdanningsdirektoratet sin scoring .....                                   | 36 |
| Tabell 7 Gruppeoversikt med Stavangerprosjektet sin scoring .....                                      | 36 |
| Tabell 8 Cohens standardforslag.....   | 37 |
| Tabell 9 Independent samples test for gruppenes totalscore .....                                       | 38 |
| Tabell 10 Score gruppevis .....  | 50 |
| Tabell 11 Svar på Side A fordelt i kategoriene.....  | 52 |
| Tabell 12 Prosentvis fordeling knyttet til årsak for feilmønsteret .....                               | 52 |
| Tabell 13 Effektstørrelser* for oppgavene Side A .....   | 52 |
| Tabell 14 Feilmønster tilknyttet årsak for gruppen Under bekymringsgrensen for oppgave 1-6 Side A..... | 53 |
| Tabell 15 feilmønster tilknyttet årsak i gruppen Over bekymringsgrensen for oppgave 1-6 Side A. ....   | 53 |
| Tabell 16 Sidescore Side B prosent og antall i gruppene .....  | 59 |
| Tabell 17 Poengscore Side C fordelt i gruppene .....   | 66 |
| Tabell 18 Effektstørrelser Side C.....   | 66 |
| Tabell 19 Feilmønstre etter oppgavenummer Side C .....   | 67 |
| Tabell 20 Poengfordeling Side D.....   | 69 |
| Tabell 21 Effektstørrelser Side D, $p < .001$ .....  | 71 |
| Tabell 22 Oversikt over feilmønster T1 og T2 på oppgave 2 Side D .....                                 | 74 |
| Tabell 23 effektstørrelse Side E, $p < .001$ .....   | 81 |
| Tabell 24 Svar fordelt i kategoriene for Side E .....  | 81 |
| Tabell 25 Prosentvis fordeling knyttet til årsak .....   | 82 |
| Tabell 26 Effektstørrelser Side F, $p < .001$ .....  | 85 |
| Tabell 27 Prosentvis fordeling av svar.....  | 85 |
| Tabell 28 Svar fordelt i kategoriene for Side G .....  | 88 |
| Tabell 29 Prosentvis fordeling tilknyttet årsak for feilmønster .....                                  | 88 |
| Tabell 30 Effektstørrelser for Side G, $p < .001$ .....  | 89 |

## Oversikt over figurer

|   |    |
|---|----|
| Figur 1 Eksempel på informativ illustrasjon.....  | 15 |
| Figur 2 Fordeling av poengscore ved de to ulike scoringene .....  | 35 |
| Figur 3 Mann-Whitney U Test på totalscore .....   | 39 |
| Figur 4 Mann-Whitney U test med utvalgte sider .....  | 40 |
| Figur 5 Elevenes score på utvalgte sider.....   | 47 |
| Figur 6 Poengscore på Side A prosentvis i gruppene .....  | 49 |
| Figur 7 Mann-Whitney U test sidescore Side A .....  | 50 |
| Figur 8 Mann-Whitney U test resultat for Side B.....  | 57 |
| Figur 9 poengscore på Side B prosentvis i gruppene.....   | 58 |
| Figur 10 Stolpediagram for svar på oppg.1 Side B.....   | 59 |
| Figur 11 Stolpediagram for svar oppg.2 Side B .....   | 61 |
| Figur 12 Stolpediagram med svar for oppg. 3 Side B.....   | 62 |
| Figur 13 Mann-Whitney U-test for Side C .....   | 64 |
| Figur 14 Poengscore på Side C prosentvis i gruppene .....   | 65 |
| Figur 15 Stolpediagram for poengscore prosentvis i gruppene .....   | 68 |
| Figur 16 Mann-Whitney U-test for Side D.....  | 70 |
| Figur 17 Feilmønstre oppgave 1 Side D .....   | 72 |
| Figur 18 Feilmønstre oppgave 2 - Side D.....  | 73 |
| Figur 19 Feilmønstre oppgave 3 - Side D.....  | 75 |
| Figur 20 Feilmønstre oppgave 4 Side D .....   | 77 |
| Figur 21 Mann-Whitney U test for Side E.....  | 79 |
| Figur 22 Poengscore for Side E prosentvis i gruppene.....   | 80 |
| Figur 23 Poengscore på Side F prosentvis i gruppene.....  | 83 |
| Figur 24 Mann-Whitney U test for Side F .....   | 84 |
| Figur 25 Mann-Whitney U test for Side G .....   | 86 |
| Figur 26 Poengscore Side G prosentvis i gruppene .....  | 87 |
| Figur 27 Feilmønstre tilknyttet årsak og oppgavene 1-6 Side G, for gruppen Under<br>bekymringsgrensen ..... | 88 |
| Figur 28 Feilmønstre tilknyttet årsak og oppgavene 1-6 Side G, for gruppen Over<br>bekymringsgrensen .....  | 89 |
| Figur 29 Oversikt over feilsvar .....   | 91 |



# 1 Innledning

Denne masteroppgaven handler om en obligatorisk kartleggingsprøve i matematikk, som blir brukt på 2. trinn i barneskolen. Fokuset er på svarene til elevene med lav score. Prøven er utviklet for å kartlegge **regning som grunnleggende ferdighet**, og er forankret i både kompetansemålene for 2. trinn fra læreplanverket for Kunnskapsløftet 2006 (LK06) og regning som grunnleggende ferdighet fra rammeverket for grunnleggende ferdigheter. Hensikten med kartleggingen er å identifisere "...elever som har svake begreper og svake telle- og regneferdigheter" (Utdanningsdirektoratet, 2014b, s. 5). Et godt utviklet begrepsapparat og gode telle- og regneferdigheter ansees som grunnleggende for videre læring. Denne studien ser nærmere på hvilken kompetanse elevene som blir identifisert av kartleggingsprøven kan se ut til å mangle.

## 1.1 Bakgrunn for valg av tema

Jeg gikk inn i dette prosjektet med et ønske om å lære mer om regning som grunnleggende ferdighet og faget matematikk i småskolen. *Kartleggingsprøve i regning for 2. trinn* ble et naturlig valg. Jeg hadde ikke kjennskap til denne prøven, annet enn at jeg hadde jobbet med rådataene til forgjengeren Regneprøven (Alseth, Throndsen & Turmo, 2007). Gjennom hele mitt studie har jeg vært opptatt av forståelsesaspektet i matematikk. Både den forståelsen som foretrekkes, men også det som omtales som misforståelse eller misoppfattelser. Det ble derfor interessant for meg å se om jeg kunne prøve å analysere hva som lå bak elevenes feilsvar, med tanke på forståelse. Feilmønstre og ulike årsaker til de ulike feilmønstrene i *Kartleggingsprøve for regning 2. trinn* ble derfor fokus for studien.

"Matematikk er nyttig." Slik begynner rapporten om *Matematikk i norsk skole anno 2014* (Borge et al., 2014). Solid kunnskap i grunnleggende regning og matematikk er et verktøy for å bedre kunne forstå verden rundt oss. Matematikk og regning innebærer mye mer enn å mestre de fire regningsartene. Logikk, bevisformer, kausalitet, overslag, sannsynlighet, sammenhenger og gyldighet i påstander er eksempler på uttrykk og tenkemåter som vi omgir oss med til daglig. Ferdigheter som

i den ”moderne verden” er nært knyttet til matematikkfaget. Grunnleggende ferdigheter i matematikk og regning kan ses på som et bidrag til det som omtales som allmenndannelse. Sjøberg (1998, s. 38) skriver om allmenndannelse:

...at skolen og skolens fag skal fremme dannelsen, eller være allmenndannende, går ut på at skolen skal bidra til at elevene utvikler seg til individer som er i stand til å delta på en selvstendig, reflektert og kritisk måte i vårt demokratiske samfunn.

Forståelse her er et nøkkelord. Richard Skemp var en pionér innen matematikdidaktikk, og klassikeren hans ”Relational and instrumental understanding” (1976) kastet lys over dette sentrale aspektet forståelse i matematikk. Hovedsakelig beskriver Skemp (1976) at forståelse har to ulike konnotasjoner. Der instrumentell forståelse er å vite hva du skal gjøre i en matematikkoppgave, eller med hans ord ”*rules without reasons*”. Relasjonell forståelse derimot er ikke bare å vite hva du skal gjøre, men også *hvorfor*. Det kan se ut til å være samsvar mellom dem som anser matematikk som et virkelighetsfjernt skolefag og instrumentell forståelse, og mellom dem som ser matematikk i alt og relasjonell forståelse, nettopp på grunn av synet på selve matematikken. Ved en relasjonell forståelse er det sammenhenger og relasjoner mellom ulike matematiske idéer og konsept, som er fokus. Men en relasjonell forståelse som grunnlag blir det da lett å få øye på matematikk i omgivelsene, og også se nytten vi har av matematikk i hverdagen.

## 1.2 Relevans

Fokuset på matematikk og realfag har vært stort også i Norge den siste tiden, eksempelvis omtalen av PISA-sjokket (Sivesind & Elstad, 2010), TIMMS-resultatene (Sjøberg, 2007), regjeringens satsing på lærerløftet, NyGiv-satsingen, ny lov om hvor mange studiepoeng en må ha for å undervise i matematikk på de ulike trinnene, antall elever som går ut med laveste karakter fra ungdomsskolen og så videre. Matematikk i skolen er et diskutert tema, og en gjenganger ser ut til å være et ønske om økt kompetanse. Det er likefremt tydelig at noe ikke er som det skal. Resultatene fra PISA-undersøkelsen 2012 viser at andelen av norske elever på de laveste prestasjonsnivåene økte fra 2009, og er nå på samme nivå som i tidligere undersøkelser. PISA-undersøkelsen tar ikke utgangspunkt i landenes læreplaner, men tar i hovedsak sikte på å måle elevens evne til å aktivt bruke kunnskaper og erfaringer

i aktuelle situasjoner. (Kjærnsli & Olsen, 2013, s. 14). Et slikt fokus kan sees i sammenheng med grunnleggende ferdigheter i regning. Begrepsforståelse, tallforståelse, og telle- og regneferdigheter er selve grunnlaget for videre læring i matematikk, og således et fundament som er nødvendig å få på plass tidlig i skoleløpet. Regjeringen rettet derfor også fokuset mot de laveste trinnene og elever i risikogruppen da opplæringsloven i 2008 fikk dette tillegget:

På 1. til 4. årstrinn skal kommunen sørge for at den tilpassa opplæringa i norsk eller samisk og matematikk mellom anna inneber særleg høg lærartettleik, og er særleg retta mot elevar med svak dugleik i lesing og rekning (Opplæringslova, §1.3).

Det var dette tillegget i opplæringsloven som førte til at vi fikk obligatorisk kartleggingsprøve i regning på 2. trinn (Borge, et al., 2014, s. 13). Hensikten med kartleggingsprøven er å identifisere elever som har svake begreper<sup>1</sup> og svake telle- og regneferdigheter, da disse elevene anses som å ikke ha tilstrekkelig ferdigheter som grunnlag for videre læring (Utdanningsdirektoratet, 2014b). Prøven består av mange enkle oppgaver, der de fleste elevene skal få til alt eller mye. Det vil si at datamaterialet har en intensjonell takeffekt, der flesteparten av elevene ligger i øvre ende av poengskalaen og få elever i den nedre. En av de fremste predikatorene for matematikkvansker er stagnering i utvikling av telle- og regnestrategier (Ostad, 2013). Derfor har prøven også tidsbegrensninger per side, som gjør at elever med uhensiktsmessige strategier ikke vil rekke å fullføre siden. For å identifisere elevene med svake begreper og svake telle- og regneferdigheter er det satt en poenggrense på prøven, der en anser elever som scorer over denne til å ha tilstrekkelig ferdigheter som grunnlag for videre læring. Denne grensen kalles bekymringsgrensen, og er satt utfra et nasjonalt representativt utvalg, der en anså de laveste 20% som å være under bekymringsgrensen. Poengsummen fra de laveste 20% ved det nasjonale utvalget, brukes nå som bekymringsgrense.

Det *Kartleggingsprøve i regning for 2. trinn* ikke gjør, er å spesifisere hvilket av områdene elevene har utfordringer med, det er overlatt til den enkelte læreren å kartlegge elever med lav score videre. ”Dersom prøveresultatene til en elev viser at

---

<sup>1</sup> Utdanningsdirektoratet bruker begrepet ”svake begreper” i omtale av formålet med *Kartleggingsprøve i regning 2. trinn*. Begrepet benyttes derfor videre i oppgaven.

eleven har manglende begrepsforståelse og/eller ferdigheter, kan det være behov for videre kartlegging før eleven får ekstra oppfølging” (Utdanningsdirektoratet, 2014b, s. 3). Læreren vil i de fleste tilfeller ha en idé om hvilke elever som kan komme til å score under bekymringsgrensen i forkant av kartleggingen. Spørsmålet er ikke om *hvilke* elever som scorer lavt, men *hva* som gjør at de scorer lavt. I mitt arbeid som vitenskapelig assistent ved Nasjonalt senter for leseopplæring og leseforskning har registrering av ulike prøve- og kartleggingsdata vært hovedoppgaven min. Noe av det jeg har syntes var ekstra spennende er å se på feilmønstrene som kommer frem i dataene, og undre meg over hvorfor elevene velger å svare akkurat som de gjør. I noen forbindelser der prøvematerialet har vært fra piloteringer, har dette vært sentrale diskusjoner for å løfte validiteten til den endelige prøven. I den forbindelse synes jeg det var interessant å se nærmere på feilsvarene i *Kartleggingsprøve i regning for 2. trinn*, i første omgang for å identifisere ulike feilmønstre, men også se om det er mulig å identifisere noen spesifikke områder eller kompetanser elevene ser ut til å ha utfordringer med. For eksempel ser utfordringene ut til å være knyttet til avkodingen av oppgaven, eller er det selve regningen eleven ser ut til å streve med? Avkoding av oppgaven kan omhandle elevens forståelse og bruk av ulike matematiske symbol, eller elevens evne til å tolke informasjon gitt i en illustrasjon. Avkoding av oppgaven er betingende for at oppgaven skal kunne løses av eleven og er derfor et interessant felt å studere.

### **1.3 Forskningsbehov**

Resultatet fra *Kartleggingsprøve i regning for 2. trinn* brukes kun i klasserommet, og indikasjoner som feilmønstrene kan gi er derfor viktig kunnskap i lærerens arbeid med å følge opp elevene som scorer lavt på prøven. Oppgavene som blir analysert i studien er også oppgaver som er brukt i de fleste læreverk og klasserom.

Implikasjoner fra denne studien er således relevant for alle som er interessert i matematikk og regning på de laveste trinnene. Jeg har fått mine data fra Stavangerprosjektet som er en tverrfaglig, longitudinell studie som gjennomføres i samarbeid med Stavanger kommune. Dataene jeg bruker er fra første nasjonale kartlegging med den reviderte prøven som ble gjennomført våren 2014, og er gjort i et avgrenset geografisk område.

Resultatene fra denne studien kan bidra i den videre oppfølgingen av elever som scorer under bekymringsgrensen. Denne kunnskapen er aktuell for både enkeltlærere og samarbeidende tjenester som Pedagogisk psykologisk-tjenesten, PPT. Kunnskap som blir konstruert ut fra denne studien er også relevant for lærerutdannere og for lærere under utdanning.

*Kartleggingsprøve i regning for 2. trinn* er pilotert, men rapporten er ikke offentliggjort.

## 1.4 Forskningsspørsmål

1. Hvilke indikasjoner gir feilmønstre i elevsvar, fra oppstilte og uoppstilte regneoppgaver i Utdanningsdirektoratet sin *Kartleggingsprøve i regning for 2. trinn*?

For å kunne svare på dette spørsmålet har jeg valgt å sammenligne responsene til elevene under bekymringsgrensen med responsene til elevene over bekymringsgrensen. Fokuset blir hovedsakelig på elevene under bekymringsgrensen, da disse står for flesteparten av feilsvarene, og således også eventuelle feilmønstre.

- a) Finnes det mønstre i elevens feilsvar og hva kan eventuelt årsaken til disse mønstrene være?
- b) Hva kjennetegner oppgavene som elevene ikke scorer på?
- c) Hvilken kompetanse, eventuelt kompetanser er det disse elevene tilsynelatende mangler?
- d) Hvordan skiller Stavangerprosjektet sin scoring seg fra Utdanningsdirektoratet sin?

## 1.5 Oppgavens oppbygning

Denne masteroppgaven er delt opp i fem hoveddeler. I oppgavens første del er tema, motivasjon for valg av tema og studiens målsetting presentert. Videre tar del 2 for seg tidligere studier og teori som synes relevant for å belyse studiens tema. Del 3 beskriver analyseprosessen og de metodiske valg som er gjort. Her omtales også *Kartleggingsprøve i regning for 2. trinn* nærmere, som er verktøyet for konstruering av data i denne studien. Metodedelen avsluttes med omtale av validitet og reliabilitet, samt forskningsetiske hensyn som er tatt i forbindelse med arbeidet. Del 4 presenterer resultater fra de statistiske metodene som ble brukt, samt gir en beskrivelse og analyse av feilmønstre som fremkom på de ulike oppgavesidene som er studert. Denne delen er utgjør en stor del av oppgaven, med bakgrunn i hvilke type data som er studert, og at resultatene for hver oppgaveside drøftes fortløpende. I del 5 av oppgaven løftes diskusjonen av funn fra studien til en større kontekst og sees i sammenheng med tidligere studier og teori som er presentert i del 2. Avslutningsvis i del 5 oppsummeres svarene på forskningsspørsmålene som ble stilt innledningsvis.

## 2 Tidligere forskning og teorier

Her presenteres tidligere forskning og teorier som er relevante for å studere forskningsspørsmålene. Denne studien ser nærmere på feilsvarene gitt i Utdanningsdirektoratet sin *Kartleggingsprøve i regning for 2. trinn* i oppstilte regneoppgaver og tekstoppgaver. Har oppgavene med lav score fellestrekk? Hvilke feilmønstre kommer til syne? Kan disse to faktorene til sammen belyse hvilke/en kompetanse elevene eventuelt mangler?

Sentrale tema i den teoretiske innramming er derfor hvordan oppgaver er kategorisert og vansker tilknyttet forskjellige typer av oppgaver. Først presenteres ulike oppgavetyper i addisjon og subtraksjon, og aspekter knyttet til oppgavetyper. Videre omtales hva tidligere studier har indikert med tanke på vanskegrad og feilmønstre i de ulike oppgavetyperne. Avslutningsvis kommenteres kompetanse som kreves for å besvare de ulike oppgavetyperne, samt hva som er forventet av elever i 2. trinn i henhold til Kunnskapsløftet (2006) sine kompetansemål.

### 2.1 Ulike oppgavetyper innen addisjon og subtraksjon

Utformingen av oppgaver er avgjørende for svarene en får, og dermed også for informasjonen en kan få kartlagt. Ulike oppgaver krever ulik kompetanse og kunnskap. Som i lesing er avkoding en viktig faktor. Elevene må kunne avkode regneoppgavene for å kunne løse dem korrekt. Reikerås (2005) bygget på Cornoldi og Lucangeli (2004) da hun presenterte denne formelen for regning:

Regning = basis regnefakta x aritmetisk oppgave- og problemløsning

I faktoren 'basis regnefakta' er regne- og tellestrategier sentrale, som hovedsakelig deles i to hovedkategorier, backup- og retrieval-varianter (S. Ostad, 1999; Ostad, 2008). Backup- strategier er primitive tellestrategier mens retrieval- strategier brukes først når regnefakta er automatisert. På denne måten er regne- og tellestrategier betingende for telle- og regneferdighet. Den andre faktoren omhandler i stor grad avkoding av oppgaven. Forståelse av tallbegreper, ord og ulike matematiske symboler er sentralt for riktig avkoding av oppgaven. Formelen ovenfor tydeliggjør at en må ha begge disse faktorene for å få produktet regning, og med det deres gjensidige

avhengighet av hverandre. Jeg vil nå presentere **en del ulike typer oppgaver, som krever ulikt av både faktoren 'basis regnefakta' og faktoren 'aritmetisk oppgave- og problemløsning'**. Det er flere måter å dele inn oppgaver på. Noen forskere har lagt vekt på vanskegrad i sin kategorisering (Carpenter & Moser, 1982; Rosenthal, Resnick & Williams, 1974), andre har forsøkt å tydeliggjør **hvilke ferdigheter som kreves for å løse de ulike oppgavene** (Riley, 1983). Jeg vil først omtale oppstilte regneoppgaver for deretter omtale tekstoppgaver

### 2.1.1 Oppstilte regneoppgaver

Ostad (2013) bygger på Rosenthal, et al. (1974) når han gjengir en klassifikasjon av oppstilte regneoppgaver, basert på hvilken av mengdene som var ukjent. Kategoriene er kalt RR- etter forfatterne av artikkelen, Rosenthal og Resnick.

Tabell 1 Kategorisering av oppstilte regneoppgaver

| Kategori | Generalisering     | Eksempel                      |
|----------|--------------------|-------------------------------|
| RR1      | $a + b = ukjent$   | $6 + 2 = ?$                   |
| RR2      | $a - b = ukjent$   | $6 - 2 = ?$                   |
| RR3      | $a \pm ukjent = c$ | $6 + ? = 8$ eller $6 - ? = 4$ |
| RR4      | $ukjent \pm b = c$ | $? + 2 = 8$ eller $? - 2 = 4$ |
|          |                    |                               |
| RR5*     | $c = ukjent \pm b$ | $8 = ? + 2$ eller $8 = ? - 2$ |
| RR6*     | $c = a \pm ukjent$ | $8 = 6 + ?$ eller $8 = 6 - ?$ |

Hovedforskjellen på kategoriene er hvilken av de tre mengdene som er den ukjente. En kategori som Rosenthal, et al. (1974) ikke tok med, men som Weaver (1973) så nærmere på, er kategorien hvor regnetegnet er plassert til høyre for likhetstegnet, altså,  $c = a \mp b$ . Fra nå av omtalt som RR5 når  $a$  er ukjent og RR6 når  $b$  er ukjent.

Når det kommer til vanskegrad for de ulike kategoriene tyder empiriske funn på at RR1 og RR2 er de kategoriene som har minst feilsvar, etterfulgt av RR3 og RR4



(Elia, Gagatsis & Demetriou, 2007; Nesher, Greeno & Riley, 1982). Weaver (1973) fant at elever hadde større vanskeligheter med å korrekt løse oppgaver med regnetegnet til høyre for likhetstegnet (for eksempel  $6 = \_?_ + 4$ ) enn for oppgaver med regnetegnet til venstre for likhetstegnet (for eksempel  $\_?_ + 4 = 6$ ). Dette blir også støttet av senere forskning på området (Gilmore, 2006; Ibarra & Lindvall, 1982). Dette tyder på at oppgaver av kategori RR5 og RR6 kan synes å være de mest utfordrende for elevene. Hvorfor disse oppgavene skiller i vanskegrad har sammenheng med hvilken kompetanse de ulike oppgavene krever av elevene. Oppgaver av kategori RR1 og RR2 er de vanligste oppgavene i de laveste trinnene, følgelig er elevene mest vant til denne. RR1/2 oppgaver egner seg også godt til bruk av backup- strategier, altså tellestrategier, som tilsier at eleven ikke trenger like god regneferdighet for å lykkes med disse oppgavene. For kategorien RR3 fungerer backup- strategien å telle videre, forutsatt at eleven avkoder oppgaven riktig med tanke på likhetstegnet. Backup- strategien 'å telle videre' vil ikke fungere for oppgaver av kategori RR4, fordi startmengden ikke er gitt, og således er det ikke noe å telle videre fra. Denne oppgavetyper krever derfor mer av elevene, fordi færre backup- strategier lar seg anvende. Felles for RR3, RR4, RR5 og RR6 er at de krever mer relasjonell forståelse for likhetstegnet for å avkode oppgaven riktig.

### **2.1.2 Uoppstilte regneoppgaver - tekstopp-gaver**

Tekstoppgaver er mer omfattende enn oppstilte regnestykker. I tekstopp-gaver får elevene et regnestykke plassert i en kontekst, og tallene de skal regne med er ikke nødvendigvis gitt eksplisitt eller i samme rekkefølge som regningen skal gjøres. Flere ulike aspekter ved tekstopp-gaver har vært fokus for tidligere studier, (se Daroczy, Wolska, Meurers og Nuerk (2015) for en oversikt). Carpenter og Moser (1982) med Riley (1983) samlet seg om en firedeling av tekstopp-gaver basert på semantikk (Landau & Lesh, 1983). Flere har brukt delingen etter dem (Ostad, 2013; Verschaffel & De Corte, 1993) Delingen deres var: Change, Combine, Compare og Equalize, oversatt til norsk med: Endring, Sammensetning, Sammenligning og Jamstilling (Ostad, 2013). En Endringsoppgave har en utgangsmengde og en eksplisitt eller implisitt handling som endrer utgangsmengden, enten positivt eller negativt.

Eksempel Endringsoppgave:

*Katrine hadde 2 karameller, så fikk hun 3 til av lillebror, hvor mange har hun nå?*

En Sammensetningsoppgave er statistisk, i den forstand at en skal ikke endre noe, men finne en total av noe som er.

Eksempel Sammensetningsoppgave:

*Katrine har 2 store karameller og 3 små karameller, hvor mange karameller har hun til sammen?*

En Sammenligningsoppgave er også av en statistisk art. Her sammenligner en to ulike mengder og oppgir differansen.

Eksempel Sammenligningsoppgave:

*Katrine har 3 karameller og lillebror har 5. Hvor mange flere karameller har lillebror?*

En Jamstillingsoppgave er derimot ikke statistisk. Således ligner den mer på en Endringsoppgave, men her likestilles to ulike mengder.

Eksempel Jamstillingsoppgave:

*Katrine har 3 karameller og lillebror 5. Hvor mange karameller trenger Katrine for å få like mange som lillebror?*

Carpenter og Moser (1982) med (Riley, 1983) forslo altså de fire kategoriene som sier noe om de ulike kontekstene til tekstoppgavene. Videre ble det fremholdt at det alltid ville være tre typer informasjon i enhver av disse oppgavene (Landau & Lesh, 1983). For eksempel vil en *endringsoppgave* inneholde en startmengde, en endring og et resultat. Det medfører også at det er tre ulike mengder en kan finne, dette gir da en total av 6 ulike typer Endringsoppgaver. Endringsoppgavene kategoriseres utfra hvilken mengde som er den ukjente (3 muligheter) og om endring er positiv eller negativ (2 muligheter). Det samme er tilfellet for Sammenligningsoppgaver. En kan

enten finne hvem som har mer eller hvem som har mindre (2 muligheter), og den ukjente kan være en av de tre mengdene (3 muligheter). Det gir 6 ulike typer *sammenligningsoppgaver*, avhengig av retning av sammenligningen og hvilken mengde som er ukjent. *Jamstillingsoppgaver* har også 6 ulike former, avhengig av hvilken mengde som er ukjent og retning på likestillingen (mer eller mindre). *Sammensetningsoppgaver* har færre muligheter, enten er den ukjente summen eller en av delmengdene det gir 2 muligheter. Ser en disse i sammenheng med RR-kategoriene får vi en god oversikt over de ulike oppgavetyperne. Der RR-kategoriene vil representere det rent algebraiske og de fire kategoriene (Endring, Jamstilling, Sammensetning og Sammenligning) vil gi konteksten til oppgaven. Tabell 2 viser en oversikt over ulike tekstopp-gaver.

Tabell 2 Oversikt over tekstoppgaver med eksempler

| Kontekst                             | Eksempel  | Varianter  |
|--------------------------------------|---|--|
| <i>Endring</i><br>(Ikke-statisk)     | Katrine hadde 2 karameller, så fikk hun 3 til av lillebror.<br>Hvor mange har hun nå?                             | Ukjent resultat:<br>RR1/2: $a \mp b = Ukjent$<br>Ukjent endring:<br>RR3: $a \mp Ukjent = c -$<br>Ukjent start:<br>RR4: $Ukjent \mp b = c$  |
| <i>Sammenligning</i><br>(Statisk)    | Katrine har 3 karameller og lillebror har 5.<br>Hvor mange flere karameller har lillebror?                        | Ukjent differanse:<br>RR3: $a + ukjent = c$<br>RR6: $c = b + ukjent$<br>Ukjent sammenligningsverdi:<br>RR1/2: $a \mp b = ukjent$<br>Ukjent referent:<br>RR1/2:<br>$ukjent = a \mp b$ |
| <i>Sammensetning</i><br>(Statisk)    | Katrine har 2 store karameller og 3 små karameller, hvor mange karameller har hun til sammen?                     | RR1: $a + b = Ukjent$<br>RR3: $a - Ukjent = c$   |
| <i>Jamstilling</i><br>(Ikke-statisk) | Katrine har 3 karameller og lillebror 5. Hvor mange karameller trenger Katrine for å få like mange som lillebror? | RR6: $c = b \mp Ukjent$<br>RR5: $c = ukjent \pm a$<br>RR1/2: $ukjent = a \pm b$  |

I Rosenthal, et al. (1974) var det hovedsakelig tekstoppgaver som var fokus, og flere hypoteser ble fremstilt for vanskegrad. De så for eksempel på semantisk linearitet og fant støtte for hypotesen om at tekstoppgaver der rekkefølgen var motsatt kronologisk (sum, endring, start) ville få flere feilsvar. Videre argumenterte de med at ukjent mengde som startpunkt (RR4) ville ta lengre tid for elevene å løse, fordi backup-strategier som å telle videre ikke ville fungere. Fordi det er ingen mengde å telle videre fra. Ett forslag til hvorfor det ville ta lengre tid med disse RR4 oppgavene, var at elevene omgjorde ligningen til en kanonisert ligning. Det vil si at i stedet for å løse

ligningen ukjent  $+ b = c$ , manipulerte de ligningen til kanonisk form:  $c - b =$  ukjent. Dette innebærer å måtte ta hensyn til hvilken tall som skal trekkes fra. Dette unngår en ved manglende minuend som startpunkt, som gir en kanonisert ligning med addisjon som jo er kommutativt. Det er andre alternativer til løsningen på disse ligningene enn å gjøre dem om til kanonisert form. Ved å ha automatisert regnefakta, vil en kunne benytte retrieval- strategier. Det vil si at en henter frem regnefakta for å løse ulike oppgaver (for eksempel at eleven vet at det som mangler på 3 for å få 5 er 2). En annen løsningsmetode er prøv og feil metoden. Eleven prøver med et tilfeldig tall, summere opp, for så å sammenligne. Er det ikke samsvar, prøves det med et nytt tall.

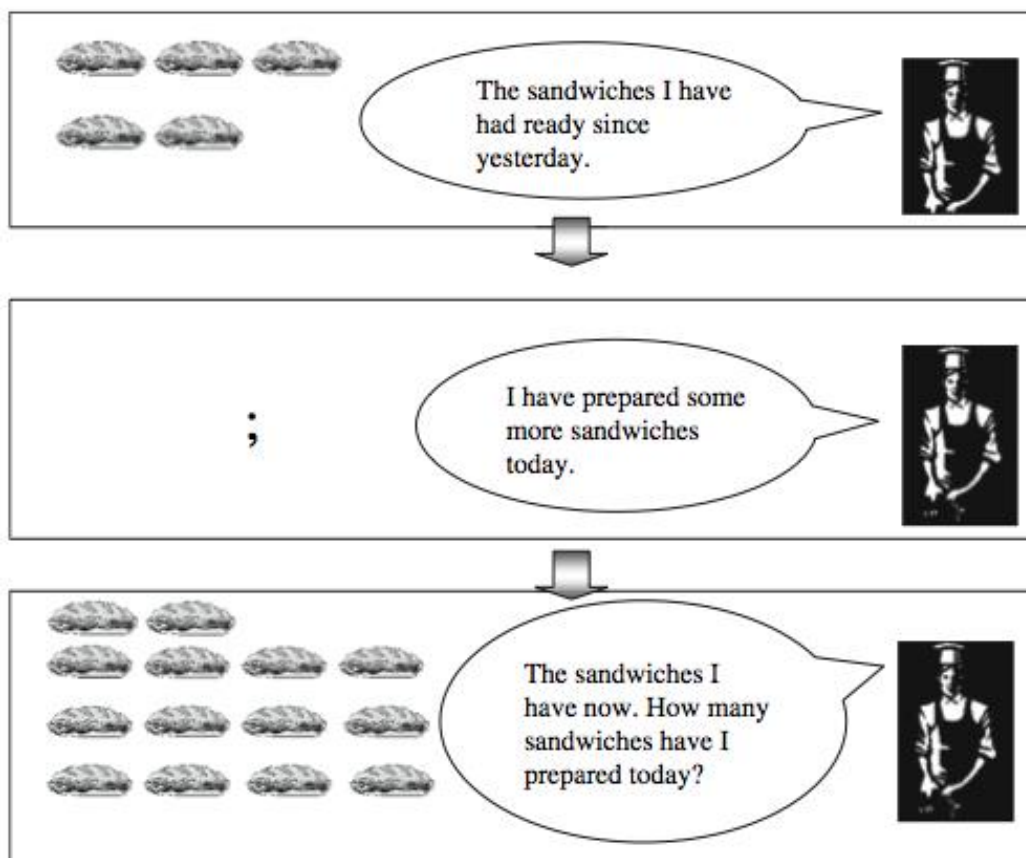
### **2.1.3 Illustrerte tekstopp-gaver**

Tekstoppgaver i de lavere trinnene inneholder oftest ikke mye skrevet tekst, men konteksten er ofte gitt ved illustrasjoner. Det eksisterer en del forskning på bruk av illustrasjoner i sammenheng med matematikk. Resultatene er delvis sprikende, og det kan tyde på at bruk av illustrasjoner i matematikk er et komplekst tema å studere, med flere aspekter å ta hensyn til. Ett av aspektene er hvilken type illustrasjon som er brukt, hva illustrasjonen er tenkt å bidra med i oppgaven. Elia og Philippou (2004) bygget på Levin (1981) sine funksjoner for bilder i tekst da de forslo følgende fire kategorier for illustrasjoner i matematikk: dekorative; representative; organisatoriske; og informative. Dekorative illustrasjoner gir ingen direkte informasjon om problemet. Representative illustrasjoner, representerer deler av eller hele innholdet i problemet. Organisatoriske illustrasjoner gir retningslinjer for tegning eller skriving som støtter løsningsprosessen. Endelig gir informative illustrasjoner essensiell informasjon for å kunne løse problemet, det vil si at problemet er basert på illustrasjonen. Videre undersøkte Elia og Philippou (2004) hvordan de ulike illustrasjonene påvirket elevens løsninger på de matematiske problemene/opp-gavene. Hovedsakelig fant de at organisatoriske, informative og representative illustrasjoner hadde effekt, men ikke de dekorative. Elia og Philippou (2004) begrunner dette med at de første tre kategoriene førte til diskusjon i elevgruppene som igjen førte til løsningen av problemet, mens de dekorative illustrasjonene ikke en gang ble nevnt. Videre fant Elia og Philippou (2004) at den informative illustrasjonen ble brukt som representativ og den mest brukte strategien her var prøv og feil. Det vil si at elevene brukte illustrasjonen til å

tegne på og se om løsningen de foreslo var rimelig, og på den måten kom de frem til den korrekte løsningen. Elia, et al. (2007) gjorde en oppfølgingsstudie hvor de kombinerte ulike typer illustrasjoner med ulike typer oppgaver (RR1-RR4), i 1. -3. trinn. De så på verbal beskrivelse, dekorative illustrasjoner, informative illustrasjoner og illustrasjon av tallinjen. Hovedsakelig fant de at informative illustrasjoner var de mest utfordrende for elevene, og indikasjoner på at mestring av denne typen oppgaver er mer tidkrevende enn for de andre representasjonstypene. Elia, et al. (2007) argumenterer med at i slike oppgaver må elevene trekke ut informasjon fra flere representasjoner samtidig, for å så sette dem i relasjon, og at dette er mer kognitivt krevende enn de andre representasjonstypene. Andre studier bekrefter dette (Ainsworth, 2006; Berends & Van Lieshout, 2009; Demetriou, 2002). Denne effekten er veldokumentert, Berends og Van Lieshout (2009, s. 346) oppsummerer det slik:

In short, these studies (Bobis, Sweller & Cooper, 1994; Chandler & Sweller, 1996; Mayer & Anderson, 1991) reveal that a necessity for switching between physically separated sources of information places a demand on working memory resources, leading to lower performance when the two sources are located further away from each other.

Eksempler på slike informative illustrasjoner er: Sammenligne flere illustrasjoner, trekke ut informasjon av en graf eller en prisliste og lignende. Eksempelet nedenfor er hentet fra Elia, et al. (2007, s. 663) og brukt på 5. trinn i Nederland.



Problem represented as an informational picture (Ijb).

Figur 1 Eksempel på informativ illustrasjon

Det er tre typer informasjon i informativ illustrasjon: kvantitativ informasjon som gir startantall – endring i antall – og sluttantall; kvalitativ informasjon som gir typen endring (+/-); og informasjon om tid som gir i rekkefølgen i oppgaven. Slike illustrasjoner for 2. trinn vil være langt mindre komplekse enn den vist ovenfor, men informasjonen er den samme. Det er ofte brukt illustrasjoner for å gi kontekst til oppgaver i de lavere trinnene, med en tanke om å støtte opp om læring- og løsningsprosessen. Studier (Ainsworth, 2006; Colin, Chauvet & Viennot, 2002; Demetriou, 2002; Duval, 2006) viser at dette ikke alltid har ønsket effekt og en bør derfor være bevisst bruken.

...it is suggested that representations (such as informational pictures) do not always facilitate problem solving or they may even hinder it because they need more complex mental processes relative to other modes of representation (Elia, et al., 2007, s. 668).

#### 2.1.4 "Problem-size" effekten

Til nå har jeg omtalt ulike oppgavetyper, plassering av den ukjente og konteksten oppgaven har. Det er enda en faktor som har vist seg å ha påvirkning på oppgaveløsningen, i første omgang på strategivalget.

Baroody and Ginsburg (1986) noted that three factors influence strategy choice: semantic structure of problem, cognitive economy, and problem's size (Baroody & Dowker, 2003, s. 47).

Sifrene som blir brukt i oppgaven påvirker også vanskegraden. Flere har interessert seg for hvordan denne sammenhengen arter seg (se Ashcraft (1992) for en oversikt). Fra et matematisk standpunkt kan en argumentere med at bruk av høyere siffer eller tall ofte gir flere trinn mot løsningen. Eksempelvis vil et addisjonsstykke med ensifret tall med sum under 10 være enklere enn et med sum over 10, fordi en unngår tierovergang, som er et ekstra trinn. Dette gjelder for både sterke og svake regnere. Eksempelvis vil en elev med svake regneferdigheter benytte seg av en form for telling (backup- strategi) i stykket  $6 + 7 = \_? \_$ . Høyere siffer gir mer å telle. Mens en elev med god regneflyt gjerne ville delt opp stykket til deler hvor regnefaktaene var kjent for henne. Eksempelvis:  $6 + 7 = (6 + 4) + (7 - 4) = 10 + 3 = 13$ . På denne måten påvirker størrelsen på sifrene vanskegraden og således tiden som blir brukt på oppgaven. Det er også blitt diskutert en analogi til høyfrekvente ord i lesing. Der høyfrekvente ord blir erstattet med høyfrekvente siffer og regnefakta i matematikk (Ashcraft, 1992). Denne analogien baserer seg på assosiasjonsstyrken til oppgavene. Ashcraft og Stazyk (1981) undersøkte sammenhengen mellom løsningsstid og størrelsen på addendene, den såkalte "problem-size – effekten". De fant støtte for at effekten ikke bare kunne bero på mental telling, altså backup- strategier, men heller at dette kunne reflektere forskjeller i assosiasjonsstyrke.

Determinants of associative strength are frequency of exposure and order of acquisition. Sums with smaller addends are solved faster by adults because they are more familiar with them, rather than because they are easier to compute (Cowan, 2003, s. 48).

Studiens funn er gjort med bakgrunn i et utvalg av voksne, men det er ingen grunn til å tro at det ikke er gjeldende for barn også.



## 2.2 Feilmønstre i addisjon og subtraksjonsoppgaver, i oppstilte regneoppgaver og tekstopp-gaver

Med utgangspunkt i RR-kategoriene kan en kategorisere en del feilsvar. Et av de vanligste feilsvarene er tellefeil, denne finner man i alle RR-kategoriene. Denne gjenkjennes som +/- 1 fra riktig svar. I noen tilfeller er det blitt brukt +/- 2 for denne kategorien, men da i sammenheng med oppgaver som har flere ledd eller store tall. Tellefeil kommer ofte av at eleven har brukt en backup- strategi, men kan også stamme fra svak assosiasjonsstyrke til oppgaven eller feil i faktagjenhenting. Eksempelvis kan en oppgave som  $2 + 3 = \_? \_$  få tallet 4 som assosiasjon fordi det er det neste tallet i telleremsen. Et annet feilmønster som har tilknytting til tellefeil, er svar som avviker med 10 fra rett svar. Det kan skyldes reell tellefeil, eller det kan grunne i svak forståelse for algoritmen som blir brukt. Disse to feilene er generelle for alle RR-kategoriene og har sammenheng med svake telle- og regneferdigheter.

Feil som er mer fremtredende for oppgaver av RR3 og RR4 kategoriene (se Tabell 1, s. 7), er tilknyttet avkoding av oppgaven. Hovedsakelig er det forståelsen for likhetstegnet som er avgjørende for om oppgaven blir avkodet riktig. Mønster som kan være fremtredende her er for eksempel addisjon av de gitte tallene, eller gjengivelse av ett av tallene. Ibarra og Lindvall (1982) bygget på Weaver (1973) og fant at det vanligste feilmønsteret i oppgaver med regnetegn til høyre for likhetstegnet (RR5/6) var at elevene adderte de to oppgitte tallene. Videre fant Lindvall og Ibarra (1982) at elevene leste disse oppgavene 'baklengs'. Mer presist, når elevene ble bedt om å lese stykket  $6 = \_? \_ + 4$  høyt, leste elevene "4 pluss noe er lik 6". Der elevene i andre oppgaver leste fra venstre til høyre, skiftet de leseretning til høyre til venstre når oppgaver av denne strukturen ble presentert. Disse lese- og feilmønstrene er knyttet til avkoding av oppgavene, og kan ikke på samme måte som tellefeil knyttes direkte til telle- og regneferdigheter. Det er ikke telling- eller addisjonsferdighetene her som er svake, men heller forståelsen for likhetstegnet og forståelsen av hva oppgaven ber dem om å gjøre.

Når vi legger til prefikset mis- til forståelse eller oppfattelse, er det for å signalisere at den type forståelse/oppfattelse er i klar konflikt med den forståelse/oppfattelse som det er konsensus om i feltet. Men det er viktig å poengtere at for eleven det gjelder er

forståelsen logisk og fullstendig. Det er noe av det som gjør det utfordrende, at eleven ikke selv er klar over konflikten. Av samme årsak har begrep som *alternativ forståelse* blitt brukt, for å understreke at forståelsen er fullstendig, men bare ikke den vi har entes om i feltet. Misforståelser eller misoppfattelser blir ofte til fra hverdagskonteksten, og således er de ofte etablert før formell instruksjon som skolegang og tilsvarende. Dette faktum gjør det enda viktigere for en lærer å være klar over mulige misoppfattelser som kan finnes i klasserommet. Kjennskap til mulige misoppfattelser er første steg til å avdekke dem, for så i andre omgang forsøke å luke dem vekk. Slike misoppfattelser kan eksempelvis omfatte bruk av algoritmer, eller være knyttet til posisjonssystemet.

Erfaring har vist oss at elever med misoppfatninger er akkurat like tenksomme – og ubetenksomme – som andre elever, forskjellen ligger i at de tenker på en annen måte omkring et begrep enn det som var meningen (Zernichow & Nygaard, 2006, s. 38).

### **2.3 Likhetstegnet (=) og elevers forståelse for tegnet**

Det er gjort et betraktelig antall studier på likhetstegnet og elevers forståelse for tegnet (Baroody & Ginsburg, 1983; Carpenter, Franke, Levi, Bass & Ball, 2003; Kieran, 1981; E. Knuth, Alibali, Mcneil, Weinberg & Stephens, 2011; N. M. Mcneil & M. W. Alibali, 2005; Prediger, 2010; Rittle-Johnson & Alibali, 1999; Stephens et al., 2013). I hovedsak er det fokusert på to ulike typer forståelse, relasjonell forståelse og operasjonell forståelse (Carpenter, et al., 2003). Relasjonell forståelse er forbundet med at likhetstegnet representerer en relasjon mellom høyre og venstre side i ligningen, og følgelig inkluderes ekvivalens. Med relasjonell forståelse blir likhetstegnet oppfattet som et relasjonstegn. Operasjonell forståelse for likhetstegnet er basert rundt at det blir oppfattet som et resultattegn, 'nå kommer svaret', og vil således ikke nødvendigvis ta høyde for ekvivalens. Det er hovedsakelig tre feilmønstre som er funnet hos elever med operasjonell forståelse for likhetstegnet (N. M. Mcneil, Rittle-Johnson, Hattikudur & Petersen, 2010):

- a) Antakelsen om at likhetstegnet alltid er til slutt i en likning (Alibali, Phillips & Fischer, 2009; Cobb, 1987; E. Knuth, et al., 2011; Nicole M. Mcneil & Alibali, 2004)

- b) Forståelse av likhetstegnet som et regnetegn som signaliserer ”summer opp tallene” eller ”svaret” (Baroody & Ginsburg, 1983; Kieran, 1981; Matthews & Rittle-Johnson, 2009; N. Mcneil & M. Alibali, 2005).
- c) Løse likninger med regnetegn på begge sider ved å summere opp alle tallene (Matthews & Rittle-Johnson, 2009; Perry, Church & Goldin-Meadow, 1988; Rittle-Johnson & Alibali, 1999).

Hovedforskjellen på disse tre mønstrene er antall mengder det tas hensyn til, til felles for alle tre er at ekvivalensen ikke opprettholdes. For regnestykker av kategori RR1 og RR2 ( $a \mp b = \_? \_$ ) vil ikke noen av de tre feilmønstrene komme til syne. For regnestykker av kategori RR3 ( $a \mp \_? \_ = c$ ) og RR4 ( $\_? \_ \mp b = c$ ) vil en derimot kunne få svar som  $a + c$ , med oppfatning a) og eller b) (Ibarra & Lindvall, 1982). For RR5 og RR6 vil det bli spesielt tydelig, fordi en i RR5 får den ukjente rett etter likhetstegnet. Ved oppfatning b) vil en da konsekvent få  $a + c$  som svar på slike oppgaver. Dette samsvarer med funnene i Lindvall og Ibarra (1980) sin studie som fant en betraktelig økning av svar av type  $a + c$  fra kategori RR3/4 til RR5/6. Et annet interessant funn i deres studie var at flere elever lot være å svare på oppgaver av Weaver sin kategori (RR5/6) enn oppgaver av RR3 og RR4.

We observed in the children's behavior an extreme rigidity about written sentences, insistence that statements be written in a particular form, and a tendency to perform actions (e.g. add) rather than to reflect, make judgments, and infer meanings (Behr, 1980, s. 15).

Med operasjonell forståelse for likhetstegnet vil en likevel få riktig svar på oppgaver fra kategori RR1 og RR2. Viktigheten av en relasjonell forståelse for likhetstegnet presiseres likevel av Carpenter, et al. (2003, s. 22)

A limited conception of what the equal sign means is one of the major stumbling blocks in learning algebra. Virtually all manipulations on equations require understanding that the equal sign represents a relation.

E. J. Knuth, Stephens, Mcneil og Alibali (2006) sin studie underbygger denne påstanden. Studien viser til en sterk positiv korrelasjon mellom elevers forståelse for likhetstegnet og prestasjonen deres på løsning av enkle ligninger. Uavhengig av om

elevene løste ligningen aritmetisk eller algebraisk, hadde elever med relasjonell forståelse for likhetstegnet større sannsynlighet for å løse ligningen korrekt, enn elever uten relasjonell forståelse for likhetstegnet.

This finding, coupled with the fact that, overall, far too few middle school students viewed the equal sign as representing a relation, clearly illustrate the need to give more explicit attention to the equal sign in middle school mathematics (E. J. Knuth, et al., 2006, s. 310).

Den operasjonelle forståelsen for likhetstegnet ser ut til å ha sitt opphav fra elevenes første møte med likninger, det vil si ett-stegsoppgaver i addisjon og subtraksjon. Ofte blir elevene møtt med kun kanoniserte likninger og utvikler ideen derifra om at likhetstegnet betyr ”svaret er” (Baroudi, 2006; Falkner, Levi & Carpenter, 1999; Kieran, 1981; Nicole M. Mcneil, 2008).

## 2.4 Kompetanse

Telling, tallbegrep og strategier er fundamentalt i all regning. De første oppgavene en løser er som regel ved telling. Først lærer en å telle, videre vil arbeid med telling utvikle tallbegrepet som danner grunnlag for videreutvikling av strategier. Gelman og Gallistel (1978) har utarbeidet fem prinsipp for telling:

- Parkobling – en til en korrespondanse; et objekt i en mengde danner par med et annet objekt fra en annen mengde. Eksempelvis en stol til hver person rundt middagsbordet.
- Tallordene kommer alltid i samme rekkefølge, selv om dette for barn kan være en ukonvensjonell rekkefølge: 1, 2, 3, d, e, f.
- Abstraksjonsprinsippet: Alt er mulig å telle; objekter, lyder, ideer osv.
- Kardinaltallsbegrep: siste tallord angir også mengden som er telt, og har antallskonservering (det blir fem fingre hver gang vi teller dem).
- Rekkefølgen en teller objektene i spiller ingen rolle, så lenge de andre prinsippene blir fulgt.

De første tre punktene er telleprosedyren, mens punkt fire omhandler hvilke sett som er tellbare og punkt fem skiller telling fra kategorisering (Gelman & Meck, 1983).

Telling og tallbegrep er komponenter i tallforståelse og således en betingelse for utvikling av i første omgang tellestrategier og senere regneferdigheter.

Det å telle er en av de første aritmetiske aktivitetene som et lite barn tar del i, men selv det å telle opp til tjue er ikke så enkelt som det kan virke for en voksen (Mcintosh, Settemsdal, Stedøy-Johansen, Arntsen & Opplæringen, 2007, s. 15).

Fremovertelling har nær tilknytting til addisjon: hva er 1 mer enn 2? 3! Hvordan vet du det? Jo, fordi 3 kommer etter 2 i telleremsen. På samme måte har bakovertelling nær tilknytting til subtraksjon. Teller vi frem eller tilbake med mer enn 1 er vi inn på multiplikasjon eller divisjon. Dette er noen av årsakene til at det er interessant å se om elever mestrer å telle fremlengs og baklengs ikke bare med 1 men også med mer enn 1. ”Utviklingen av mer sofistikerte, aritmetiske framgangsmåter kan betraktes som forsøk på å gjøre telleprosessen enklere og mer effektiv” (Mcintosh, et al., 2007, s. 49). Etter 2. trinn forventer en at elevene kan sette sammen og bygge mengder opp til 100. Elever som teller sikkert og fleksibelt vil finne dette relativt enkelt i forhold til elever som ser tallene som en lang rekke vilkårlige tall og lyder. Hovedforskjellen på disse to grupperingene er deres tallbegrep og tallforståelse, i tillegg til hvor flinke de er til å telle. Tallbegrep og tallforståelse henger sammen, en trenger et godt og fleksibelt tallbegrep for å forstå hvordan tallene henger sammen, og det er enklere å telle fleksibelt når en har forstått systemet tallene er i.

#### **2.4.1 Et fleksibelt tallbegrep**

Vi omgir oss med tall hele tiden i hverdagen, og som voksen er det lett å ta for gitt alle de ulike sammenhengene tall dukker opp i. Vi stusser ikke over at husnummer 5 ikke nødvendigvis er det femte huset, eller at buss nummer 7 alltid kommer før buss nummer 6, ei heller at klokken kan være både 13 og ett samtidig. Det samme kan sies om begreper vi bruker for å klassifisere, vi vet at selv om Olav er størst er han fremdeles yngst i gruppa. Eller at tidsbegreper kan være relative, men i midten og mellom er en fast plass i rekka. Vi fyller enkelt ut huller på en tallinje, men vet også at vi aldri vil klare å fylle den helt opp. Alt dette vi tar for gitt er resultatet av vårt rike tallbegrep, vi forstår hvorfor det er slik, vi har lært det. Vi vil gjerne at barn også skal utvikle et slikt rikt tallbegrep. For å gjøre dette er det viktig at elevene møter tall i ulike situasjoner og former. Det å gå fra en representasjon til en annen kobles til

hvordan mengder kan representeres med tall. For eksempel kan dette innebære å forstå at en mynt kan ha verdi 1, 5 og 10 selv den fremdeles bare er én mynt. Ulike representasjoner hjelper elevene til å oppdage at tall kan være både verdi, navn eller si noe om rekkefølge. Bruk av mengder og grupperinger og sammenligninger dem imellom er med å utvikle tallbegrep om hva som er større enn, mindre enn eller likt. I tillegg til grupperinger som kan være nyttig i tallregning, eksempelvis tiere og femmere. Etter endt 2. trinn er målet at elevene skal kunne bygge mengder opp til 10, sette sammen og dele opp tiergrupper opp til 100 og dele tosifrede tall i tiere og enere. Videre skal de også mestre å gjøre overslag over mengder, telle opp, sammenligne tall og uttrykke tallstørrelser på varierte måter (Kunnskapsløftet, 2006).

#### **2.4.2 Strategier**

Strategier er måter å komme frem til svaret på en regneoppgave. Stagnering i utvikling av telle- og regnestrategier regnes som en viktig predikator for matematikkvansker (Ostad, 2013), og er således også en viktig byggestein i regning.

Siegler og Jenkins (1989) som er av de mest betydningsfulle forskerne når det gjelder strategier i matematikk, benytter uttrykket obligatorisk for å skille strategi fra prosedyre generelt, og definerer strategi som enhver ikke-obligatorisk og målrettet prosedyre (Ostad, 2013, s. 11).

Ikke- obligatorisk prosedyre spiller på at en selv velger hvilken strategi en vil bruke og målrettet peker på at hensikten er å komme frem til en løsning på oppgaven. Videre deles strategier i Generelle og Oppgavespesifikke, hvor den siste er mest aktuell for denne studien. Innen Oppgavespesifikke strategier deler en videre inn i Retrieval- og Backup- strategier. ”Uttrykkene har vokst frem fra teorier hvor forskerne tenker seg elevenes matematikkunnskaper som et lager av kunnskapsenheter”(Ostad, 2013, s. 12). Retrievalbegrepet peker på at en henter frem kunnskap fra dette lageret, eksempelvis automatiserte regnefakta, en vet svaret, eller et annet svar som kan være grunnlag for videre arbeid. Backup- strategier er alle de andre strategiene, som ulike tellinger og ’omveier’ til oppgavens løsning. Begge typer strategier leder frem til oppgavens løsning, men backup- strategier er mer tidkrevende og de gir rom for flere steder å gjøre feil enn retrievalstrategiene. Regne- og tellestrategier kan sees som to sider av samme sak (S. A. Ostad, 1999), og i Tabell 3

under gjengir jeg eksempler på backup- og retrievalvarianter innen addisjon og subtraksjon hentet fra Ostad (2013).

**Tabell 3 Strategier i addisjon og subtraksjon**

| Addisjon |  |           |  | Subtraksjon |   |           |  |
|----------|--|-----------|--|-------------|---|-----------|--|
| Backup   |  | Retrieval |  | Backup      |   | Retrieval |  |
| A1a      | Telle alt og forfra igjen                        | A2a       | Vet svaret   | S1a         | Telle alt og forfra igjen   | S2a       | Vet svaret   |
| A1b      | Telle alt  | A2b       | Avledet variant 1<br>(vet svaret på addisjons-kombinasjoner og benytter dette som grunnlag for videre telling) | S1b         | Tilvekst-varianten<br>(Teller videre på det som skal trekkes fra)                 | S2b       | Avledet variant 1<br>(vet svaret på addisjons-kombinasjoner og benytter dette som grunnlag for videre telling) |
| A1c      | Telle videre                                     | A2c       | Avledet variant 2<br>(lik A2b, men uten å telle)   | S1c         | Minkings-varianten<br>(teller baklengs)   | S2c       | Avledet variant 2<br>(lik S2b, men uten å telle)   |
| A1d      | Minimums-varianten<br>(minimum antall tellesteg) |           |  | S1d         | Tilvekst minkings-varianten<br>S1a eller S1b, velger den som krever minst telling |           |  |
| A1e      | Tegne-varianten                                  |           |  | S1e         | Tegne-varianten   |           |  |
| A1f      | Telle-punkter i tallsymbol                       |           |  | S1f         | Telle-punkter i tallsymbol  |           |  |
| A1g      | Andre telle-varianter                            |           |  | S1g         | Andre telle-varianter   |           |  |
| A1h      | Verbal telling                                   |           |  | S1h         | Verbal telling  |           |  |



En forutsetning for å ta i bruk disse strategiene er at elevene har forstått hva oppgaven ber dem om å gjøre, altså faktoren *aritmetisk oppgave- og problemløsning* fra regningsformelen. Det vil si at de må kunne bruk og betydning av symboler som +, - og =. Elevens type forståelse av aritmetiske uttrykk påvirker elevenes strategibruk. Elever som forstår uttrykket både som et uttrykk for et konsept og en prosedyre kan bruke regnestrategier mer fleksibelt, enn elever som forstår uttrykket kun som et uttrykk for en prosedyre (Edward M. Gray, 1991; Eddie M. Gray & Tall, 1994). Rigid strategibruk er en av varsellampene når det kommer til utvikling av regneferdigheter (Ostad, 2013). Dowker (2005b, s. 90) eksemplifiserer med at hvis et barn tolker alle spørsmål av typen ”hvor mange/mye?” kun som en telleoppgave (prosedyre), kan barnet overse fakta og sammenhenger som gjør tellingen overflødig i utgangspunktet. For eksempel antallskonservering eller faktagjenhenting. Etter endt 2. trinn er målet at eleven skal kunne ”utvikle, bruke og samtale om varierte reknestrategiar for addisjon og subtraksjon av tosifra tal og vurdere kor rimelege svara er” (Kunnskapsløftet, 2006).

Overall, research on addition strategies demonstrates that they evolve adaptively as children progress toward increasingly rapid efficient strategies. Although an earlier generation considered all methods apart from retrieval as crutches and teachers were likely to forbid the use of fingers, we now see these as a necessary part of children’s development. From the very beginning of children’s learning to solve addition problems, the methods they invent are valid. (Baroody & Dowker, 2003, s. 56).

## **2.5 Oppsummering tidligere forskning og teori**

Det er hovedsakelig tre aspekt som gjør seg gjeldende for oppstilte oppgavetyper; Plassering av den ukjente og regnetegnet, og problem- size effekten med assosiasjonsstyrke. For tekstopp-gaver er det mer komplekst, men plassering av den ukjente gjør seg gjeldende her også. I tillegg kommer eventuelle illustrasjoner som både kan fremme og hindre løsningsprosessen. Felles for begge typer oppgaver er avkodning og forståelse av matematiske symbol som +, - og =.

### **3 Metode og analyse**

I dette kapittelet presenterer jeg de metodiske valgene som er gjort i studien. Jeg presenterer også Stavangerprosjektet som er min datakilde, Utdanningsdirektoratet sin *Kartleggingsprøve i regning for 2. trinn* som er verktøyet for datakonstrueringen og de spesifikke oppgavene jeg har sett nærmere på. Videre omtales de ulike statistiske analysene jeg har brukt, og grunnlaget for å velge disse. Utvalg, validitet og reliabilitet blir også omtalt, samt forskningsetiske hensyn som er tatt i forbindelse med studien.

#### **3.1 Valg av metode og forskningsdesign**

Denne studien har en empirisk kvantitativ tilnærming. Dette valget er tatt med bakgrunn i forskningsspørsmålene som var aktuelle for studien og datamaterialet studien er gjort på. Undersøkelsen er en tverrsnittsundersøkelse, i den forstand at den er foretatt på ett tidspunkt med et stort utvalg bestående av elever på ett klassetrinn (Johannessen, Christoffersen & Tuft, 2010). Selve materialet inngår i en i større longitudinell studie, mens denne delstudien ikke er longitudinell.

Tverrsnittsundersøkelser er gunstige når en ønsker å undersøke karakteristiske trekk ved en gruppe på et gitt tidspunkt, som jo er hensikten med en kartleggingsprøve som Utdanningsdirektoratet sin *Kartlegging i regning for 2. trinn*. I og med at materialet mitt bare består av data fra ett tidspunkt, skal jeg være forsiktig med å trekke konklusjoner om utvikling over tid. Resultatet fra denne studien kan derimot være med å forklare hvordan ulike feilmønstre kan henge sammen med gruppetilhørighet, da i betydningen om elever som scorer over eller under bekymringsgrensen.

#### **3.2 Datakilde - Stavangerprosjektet**

Stavangerprosjektet "Det lærende barnet" startet i 2007 og er en tverrfaglig, longitudinell studie som gjennomføres i samarbeid med Stavanger kommune. Prosjektet har fokus på barns utvikling fra 2-10 år, er tilrådd av Datafaglig sekretariat og skal pågå til utgangen av 2018. (Lesenteret, 2015). Prosjektet er en studie av barns utvikling innen språk, matematikk, motorikk og sosial kompetanse. Utvalget besto av to og et halvt årskull barn i Stavanger kommune (født perioden 1.juli 2005 til 31.

desember 2007. For mer informasjon om Stavangerprosjektet se Vedlegg I), og data ble samlet inn når barna var 2,9 år, 4,9 år, og i 1.klasse og 2.klasse på barneskolen. Mitt datamateriell består av ett av årskullene fra dette prosjektet og er på totalt 440 elever på 2. trinn i Stavanger kommune. Det er et eget materiell for registrering av data i barnehagefasen, mens datamaterialet for barneskolen hovedsakelig er nasjonale kartleggingsprøver i lesing og regning. I tillegg kommer noen tester utviklet for Stavangerprosjektet spesielt. Ved oppstart av prosjektet var alle kommunale barnehager i Stavanger kommune med, og omtrent halvparten av de private barnehagene.

### **3.2.1 Bruk av allerede eksisterende data**

Det å ta i bruk allerede eksisterende data er et viktig bidrag til forskning. Dette fordi at flere prosjekter samler inn store mengder data, og belyser ulike problemstillinger utfra dem. Disse dataene er så rike, og følgelig er det omfattende mengder problemstillinger å undersøke. Dermed er det gunstig at flere får anledning til å bruke dataene, både for å muligens avdekke eller komplimentere forhold og sammenhenger. Å bruke allerede innhentede data er også med å skåne feltet en studerer, elever slipper å bli belastet og lærere slipper å organisere. Ikke minst er det tidsbesparende for den som utfører studien, ved at en får mer tid til eksempelvis selve analysen. Det å gå inn i et allerede eksisterende prosjektet er allikevel ikke problemfritt.

### **3.3 Utdanningsdirektoratet sin Kartlegging i regning for 2. trinn**

Mitt datamaterialet består av kartleggingsheftene, *Kartlegging i regning for 2. trinn*, til ett av årskullene som er med i Stavanger Prosjektet. Denne kartleggingsprøven er obligatorisk nasjonalt, og blir revidert hvert 5.år. Materialet jeg brukte er fra første gjennomføring av den reviderte prøven, det vil si den som i utgangspunktet skal brukes til og med 2018. Å gjengi konkrete oppgaver er derfor vanskelig med tanke på konfidensialitet. Grunnlaget for konfidensialitet er at en spredning av kartleggingsprøven på forhånd kan ødelegge kartleggingseffekten.

#### **3.3.1 Mål for Kartleggingsprøve i regning for 2. trinn**

Kartleggingsprøvens formål tydeliggjøres i veiledningen til prøven:

Kartleggingsprøven skal brukes til å kartlegge regning som grunnleggende ferdighet, avgrenset til elevenes tallforståelse og regneferdighet. Formålet med prøven er å identifisere elever som har svake begreper og telle- og regneferdigheter. (Utdanningsdirektoratet, 2014b, s. 5)

### 3.3.2 Design av prøven

Videre beskrives prøvens forankring ”Prøven er forankret i definisjonen og progresjonsbeskrivelsen for regning i *Rammeverk for grunnleggende ferdigheter* og kompetansemål i læreplanen” (Utdanningsdirektoratet, 2014b, s. 5). Når det refereres til kompetansemål i læreplanen presiseres det at det gjelder kompetansemål i matematikk 2. trinn (Utdanningsdirektoratet, 2014b). Kompetansemålene i matematikk i LK06 etter 2. trinn er delt inn i fire tema: tall, geometri, måling og statistikk (Utdanningsdirektoratet, 2013). *Kartlegging i regning for 2. trinn* er avgrenset til elevenes tallforståelse og regneferdighet, og følgelig er kompetansemålene for temaet tall mest aktuelle.

Mål for opplæringen er at eleven skal kunne:

- Telle til 100, dele opp og bygge mengder opp til 10, sette sammen og dele opp tiergrupper
- Bruke tallinja til beregninger og til å vise tallstørrelser
- Gjøre overslag over mengder, telle opp, sammenligne tall og uttrykke tallstørrelser på varierte måter
- Utvikle og bruke varierte regnestrategier for addisjon og subtraksjon av tosifrede tall
- Doble og halvere
- Kjenne igjen, samtale om og videreføre strukturer i enkle tallmønstre

(Kunnskapsløftet, 2006).

Den andre delen av *Kartleggingsprøve i regning for 2. trinn* sin forankring er i definisjonen og progresjonsbeskrivelsen fra rammeverket for grunnleggende ferdigheter.

Å kunne regne er å bruke matematikk på en rekke livsområder. Å kunne regne innebærer å resonnerer og bruke matematiske begreper, fremgangsmåter, fakta og verktøy for å løse problemer og for å beskrive, forklare og forutse hva som skjer. Det innebærer å gjenkjenne regning i ulike kontekster, stille spørsmål av matematisk karakter, velge holdbare metoder når problemene skal løses, være i stand til å gjennomføre dem og tolke gyldigheten og rekkevidden av resultatene. Videre innebærer det å kunne gå tilbake i prosessen for å gjøre

nye valg. Å kunne regne innebærer å kommunisere og argumentere for valg som er foretatt ved å tolke konteksten og arbeide med problemstillingen til en ferdig løsning (Utdanningsdirektoratet, 2012, s. 12).

Progresjonsbeskrivelsen for regning er satt opp i matrise (se vedlegg II). Det er utviklet oppgaver til to av ferdighetsområdene i *Kartleggingsprøve i regning for 2. trinn, Bruke og bearbeide* og *Gjenkjenne og beskrive*. ”Hovedtyngden av oppgavene er utviklet til området Bruke og bearbeide. Bruke og bearbeide handler om å velge strategier for problemløsning” (Utdanningsdirektoratet, 2014b, s. 5). I tillegg måles ferdighetsområdet *Kommunikasjon* indirekte, og det er også utviklet enkelte oppgaver som tester mer grunnleggende begrepsforståelse (Utdanningsdirektoratet, 2014b). Til sammen er det 15 sider med oppgaver i heftet, og hver side har en overskrift med instruksjon. I veiledningen presiseres det at hovedvekten av hva som kartlegges er henholdsvis begrepsforståelse og telleferdigheter før pausen, og tallregning (addisjon og subtraksjon) etter pausen (Utdanningsdirektoratet, 2014b).

Oppgavene måler kompetanse på det laveste nivå fra progresjonsbeskrivelsen:

Gjenkjenne og beskrive, nivå 1:

Gjenkjenner konkrete situasjoner som kan løses ved regning, og formulerer spørsmål til dem.

Bruke og bearbeide, nivå 1:

Bruker enkle strategier for opptelling og klassifisering av mengder og geometriske former. Utfører enkle beregninger på ulike måter.

(Utdanningsdirektoratet, 2012, s. 13).

I veilederen til *Kartleggingsprøve i regning for 2. trinn* presiseres det hvordan ovennevnte punkter er tolket for 2. trinn. Under punktet *Gjenkjenne og beskrive* står det:

For 2. trinn vil det si å kunne formulere korrekt regneuttrykk til situasjoner som beskrives ved illustrasjoner, som for eksempel å regne ut hva to gjenstander koster til sammen. Det kan også handle om noe så enkelt som å forstå at sju epler (halvkonkreter/tegning) kan representeres med (tall)symbolet 7, eller kunne sammenlikne to representasjoner for samme tall (for eksempel tallsymbol og konkreter) (Utdanningsdirektoratet, 2014b, s. 5).

Videre under punktet *Bruke og bearbeide* er presisjonen gitt slik:

For prøven for 2. trinn er det naturlig å avgrense dette til å bruke strategier for oppstilling og klassifisering av mengder samt å kunne gjennomføre aritmetiske beregninger (både oppstilte og uoppstilte oppgaver) (Utdanningsdirektoratet, 2014b, s. 5).

*Kartleggingsprøve i regning for 2. trinn* deles inn i fire temaer: Telleferdigheter; Tallbegrep; Tallrekka; og Regneferdigheter. Denne studien ser nærmere på sistnevnte, som omtales slik i veilederen til *Kartleggingsprøve i regning for 2. trinn*:

*Regneferdigheter* handler om å kunne bruke kunnskap om tall og mengder sammen med strategier til å løse oppstilte og uoppstilte oppgaver (i aritmetikk) (Utdanningsdirektoratet, 2014b, s. 6).

### **3.3.3 Gjennomføring av *Kartleggingsprøve i regning for 2. trinn***

Prøven har to deler, og det er anbefalt med en pause på ti minutter mellom de to delene. Det er beregnet 60 minutter til gjennomføringen ekskludert pause. Videre er det anslått at elevene bruker 16 minutter på oppgavene før pausen og 12 minutter på oppgavene etter pausen. Det resterende av de 60 minuttene brukes av lærer til instruksjon. Det er mulig å oppnå 55 poeng på prøven, og bekymringsgrensen er satt til 41 poeng våren 2014, og gjelder så lenge prøven blir brukt. Instruksjon for gjennomføring av kartleggingsprøven er gitt av Utdanningsdirektoratet, og lærere bes sette seg nøye inn i denne før gjennomføringen (Utdanningsdirektoratet, 2014a). Alle oppgaver i *Kartleggingsprøve i regning for 2. trinn* har en egen instruksjon som leses høyt av lærer før elevene skal løse oppgavene. Etter at prøven er gjennomført, samler læreren inn heftene og retter dem etter fasiten som står i instruksjonen. Kun elevens score på hver oppgave legges inn i PAS (Prøveadministrasjonssystemet). Elevenes konkrete svar på prøven er det kun læreren som ser under retting.

#### **3.3.3.1 Data**

Under lister jeg sidene jeg har studert med antall oppgaver. Hver riktig oppgave gir ett poeng. Sidene med oppstilte regneoppgaver består av 4 eller 6 oppgaver som står i to kolonner, oppgavene er gitt nummer etter leseretning venstre til høyre. For de to sidene med uoppstilte oppgaver med svaralternativer er oppgavene listet vertikalt, og

gitt nummer der etter. Tallene som benyttes i både i de oppstilte og uoppstilte oppgavene øker med oppgavenummeret. Det vil si at vanskegrad også kan tenkes økes med oppgavenummeret jf. ”problem-size” effekten.

Side A: Oppstilte regneoppgaver (6 oppgaver av RR5 og RR6)

Pause

Side B: Tekstoppgaver (3 oppgaver – tekstoppgave med svaralternativ)

Side C: Oppstilte regneoppgaver (4 oppgaver av RR1)

Side D: Tekstoppgaver (4 oppgaver - tekstoppgave)

Side E: Oppstilte regneoppgaver (4 oppgaver av RR2)

Side F: Tekstoppgaver (3 oppgaver – tekstoppgaver med svaralternativ)

Side G: Oppstilte regneoppgaver (6 oppgaver – 3 av RR3, 1 av RR4, 2 av RR5)

Valget om å studere side for side og ikke oppgavene etter kategori er tatt med bakgrunn i tidsbegrensningen som er gitt elevene per oppgaveside. Refleksjonen rundt dette valget var at tidsbegrensningen for siden kunne føre til flere ubesvart på de siste oppgave på siden, og at dette ville få konsekvenser for utfallet av feilmønstrene for oppgavekategorien. Det ble også reflektert rundt hvordan problem-size effekten kunne spille inn på dette, da oppgavene konsekvent bruker høyere tall mot slutten av siden. Det er mulig at en ren analyse av kategoriene alene kunne vært hensiktsmessig, men det er ikke gjort i denne studien.

### **3.4 Ulike oppgavetyper**

For å avgrense studien min måtte jeg velge ut noen sider eller spesifisere hvilken type oppgaver jeg ville se på. Med bakgrunn i ønsket om å se på feilmønstre, ble det gjort noen innledende analyser for å se hvilken av oppgavesidene som skilte best mellom de to gruppene. Det ble sett på median, gjennomsnittsscore og standardavvik for begge gruppene på alle oppgavesidene, samt laget scatterplott for de to delene i *Kartleggingsprøve i regning for 2. Trinn* (se vedlegg III). Del 2 av *Kartleggingsprøve i regning for 2. trinn* viste seg å skille mest mellom gruppen Under bekymringsgrensen og gruppen Over bekymringsgrensen. Denne delen bygger på kunnskap som kartlegges i del 1, og er derfor antatt noe vanskeligere for elevene enn del 1. Del 2 består hovedsakelig av regneoppgaver, både oppstilte og tekstoppgaver.

Valget ble derfor regneoppgaver, både oppstilte og tekstoppgaver. En side ble tatt med fra del 1, da den består av 6 oppstilte regneoppgaver og er vektet med 6 poeng. Den tunge poengvektingen gjør at siden skiller godt mellom de to gruppene og den består av oppstilte regneoppgaver, på den bakgrunn ble den med i utvalget av sider jeg har studert. Til sammen står sidene jeg studerte for 30 av de 55 mulige poengene på *Kartleggingsprøve i regning for 2. trinn*. Det vil si at elevene må ha minst 17 poeng på de utvalgte sidene, forutsatt at de har alt annet riktig, for å komme over bekymringsgrensen. Feilmønstre i disse oppgavene, kan derfor være en god indikator på hva elevene ikke mestrer.

Innenfor oppstilte regneoppgaver er det flere ulike varianter på oppgavesidene i *Kartleggingsprøve i regning for 2. trinn* denne studien tar for seg. De er gjengitt i tabellen nedenfor.

**Tabell 4 Oppgavetyper i *Kartleggingsprøve i regning for 2. trinn* – studiens utvalg**

| Kategori | Generalisering             | Antall tekstoppgaver | Antall oppstilte regneoppgaver | Totalt |
|----------|----------------------------|----------------------|--------------------------------|--------|
| RR1      | $a + b = \text{elevsvar}$  | 7                    | 4                              | 11     |
| RR2      | $a - b = \text{elevsvar}$  | 3                    | 4                              | 7      |
| RR3      | $a + \text{elevsvar} = c$  |                      | 3                              | 3      |
| RR4      | $\text{elevsvar} + b = c$  |                      | 1                              | 1      |
| RR5      | $c = \text{elevsvar} + b$  |                      | 5                              | 5      |
| RR6      | $c = a + \text{elevssvar}$ |                      | 3                              | 3      |

### 3.4.1 Tekstoppgaver

Tekstoppgavene jeg har sett på i *Kartleggingsprøve i regning for 2. trinn* er basert på illustrasjoner med en muntlig instruksjon, og omhandler kjøp og salg. To av oppgavesidene bruker en kombinasjon av halvkonkreter og tallsymbol, og én bruker bare tallsymbol. De tre variantene av tekstoppgaver jeg har sett på i *Kartleggingsprøve i regning for 2. trinn* ser slik ut:



Tekstoppgave variant 1: Summer halvkonkretene og sammenlign med tallsymbol, verbal informasjon: Hvem har kjøpt hva?

Tekstoppgave variant 2: Summer halvkonkretene, trekk fra kjøpet og sammenlign med tallsymbol. Verbal informasjon: Hvor mye har du igjen?

Tekstoppgave variant 3: Sammenlign illustrasjonene, finn prisen, summer og skriv summen. Verbal informasjon: Hvor mye til sammen?

Alle disse tre variantene innebærer en sammenligningskomponent. Variant 1 og 2 har svaralternativer, mens i variant 3 bes elevene skrive svaret selv. Alle oppgavesidene blir gitt en kort beskrivelse ved presentasjon av resultatet for leservennlighet.

### **3.4.2 Åpne svar eller svaralternativer**

I *Kartleggingsprøve i regning for 2. trinn* er det to måter elevene avgir svar på. Den ene er åpne svar, hvor eleven skriver svaret sitt selv. Den andre varianten er svaralternativer, hvor eleven bes om å sette et kryss, trekke en linje eller lignende til riktig svaralternativ. Slike flervalgsoppgaver er vanlig i norsk skole, men bruken av dem i kartlegging kan diskuteres. Grunnlaget for dette er at vi kan ikke garantere at eleven svarer på bakgrunn av kunnskap, det kan like gjerne være ren gjetting. Et annet poeng er at *Kartleggingsprøve i regning for 2. trinn* har få oppgaver per side, og således øker sannsynligheten for å gjette riktig. En kan argumentere med at elever kan gjette i åpne oppgaver også, men der vil en eventuell feilgjetting kunne fortelle noe mer om elevens tankegang. Ved å ha ferdige svarkategorier mister man denne muligheten, som kan være med å belyse elevens tankegang i problemløsningen. Hvis eleven ikke gjetter, og det ligger noen misoppfattelser til grunn i elevens forståelse, kan disse også komme til syne ved åpne svar. Mulige misoppfattelser er ofte representert i ett eller flere av svaralternativene. Ved analyse av slike svar og mulig årsaker til feilmønstre, må en ta høyde for at svaralternativet representerer misoppfattelsen og at dette svaret ikke nødvendigvis hadde vært elevens svar hvis oppgaven hadde et åpent svar. Disse forholdene er viktig å ta høyde for når en vektlegger resultatet av analysen i oppgaver med svaralternativer.

### 3.5 Analyseverktøy SPSS

Råmaterialet som besto av elevenes konkrete svar på de ulike oppgavene var allerede registrert i SPSS da jeg mottok datafilen, og jeg valgte å fortsette med SPSS for de videre analysene. Både av hensyn til enkelheten, men også fordi at SPSS er et program jeg hadde noe kjennskap til fra før. Datamaterialet ble behandlet både beskrivende og analytisk. Den beskrivende statistikken er med å klargjøre hovedkonturene i materialet, mens den analytiske statistikken er med å påvise sammenhenger, teste hypoteser og finne generaliseringsverdi. Det vil i hovedsak si hvor sannsynlig det er at resultatene skyldes tilfeldigheter eller realitet.

I analytisk statistikk går det et viktig skille mellom parametrisk og ikke-parametrisk statistikk. Mitt datamaterialet har en intensjonell takeffekt, det vil si negativ skewness. Der et normalt fordelt materiale ville vært tilnærmet symmetrisk og lett gjenkjennbart som Bellskurve med gjennomsnittet midt på poengskalaen, er mitt materiale forskjøvet mot høyre og mangler store deler av den siste halen. Det tilsier at de fleste elevene i materialet har en score som er i øvre ende av poengskalaen, og at gjennomsnittet ligger over bekymringsgrensen. Det tilsier i utgangspunktet at ikke-parametrisk statistikk vil være å foretrekke, med grunnlag i at materialet ikke er normalfordelt. Men, i og med at utvalget er så stort som det er, kan en lene seg på sentralgrenseteoremet som grovt sier: med stort nok utvalg vil et hvert materiale bli normalfordelt til slutt (Field, 2013). Det som er risikoen med å bruke parametrisk statistikk på ikke-normalfordelt materiale, er at den som regel er sensitiv for ekstremverdier. Dette fordi store deler av denne statistikken er bygget opp rundt gjennomsnittsverdier, og en gjennomsnittsverdi kan endres mye ved en eller få ekstremverdier.

#### 3.5.1 Metode

For å avgjøre hvilken statistisk metode jeg ville bruke, måtte jeg undersøke de to gruppene først med tanke på varians og størrelse. I mitt materiale er det gruppen Under bekymringsgrensen som står for ekstremverdiene, på score på hver enkelt side og totalscore på *Kartleggingsprøve i regning for 2. trinn*. Dette fordi totalscore har den potensielle variasjonsbredden 41 (0-41) for gruppen Under bekymringsgrensen og kun 13 (42-55) for gruppen Over bekymringsgrensen. Det vil si at gruppen Under

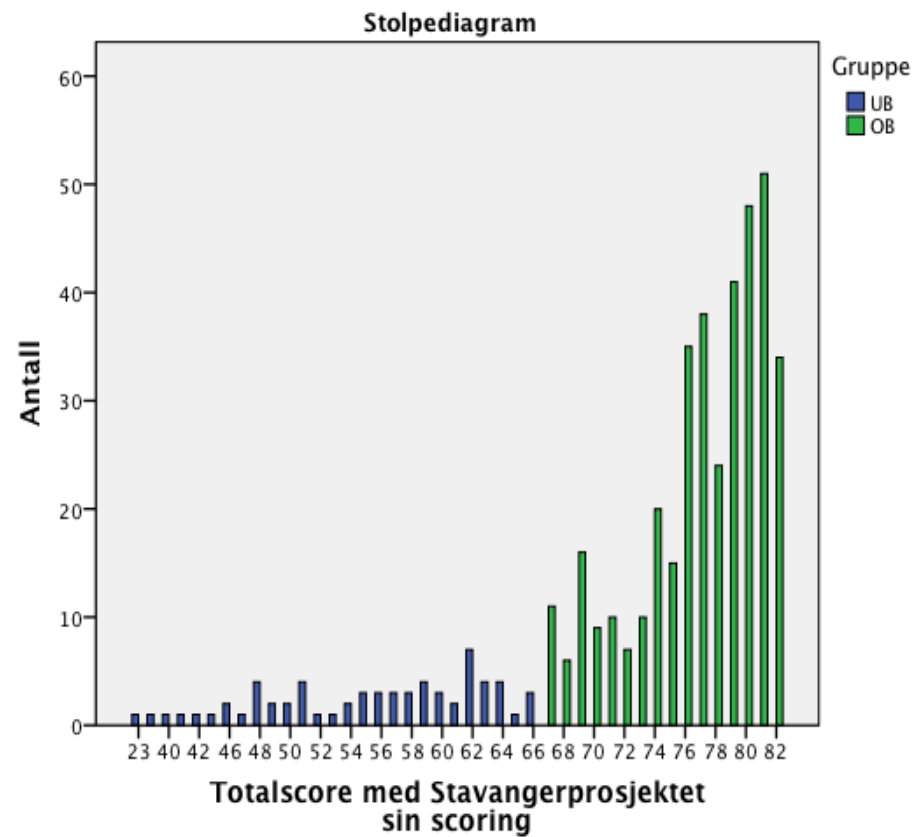
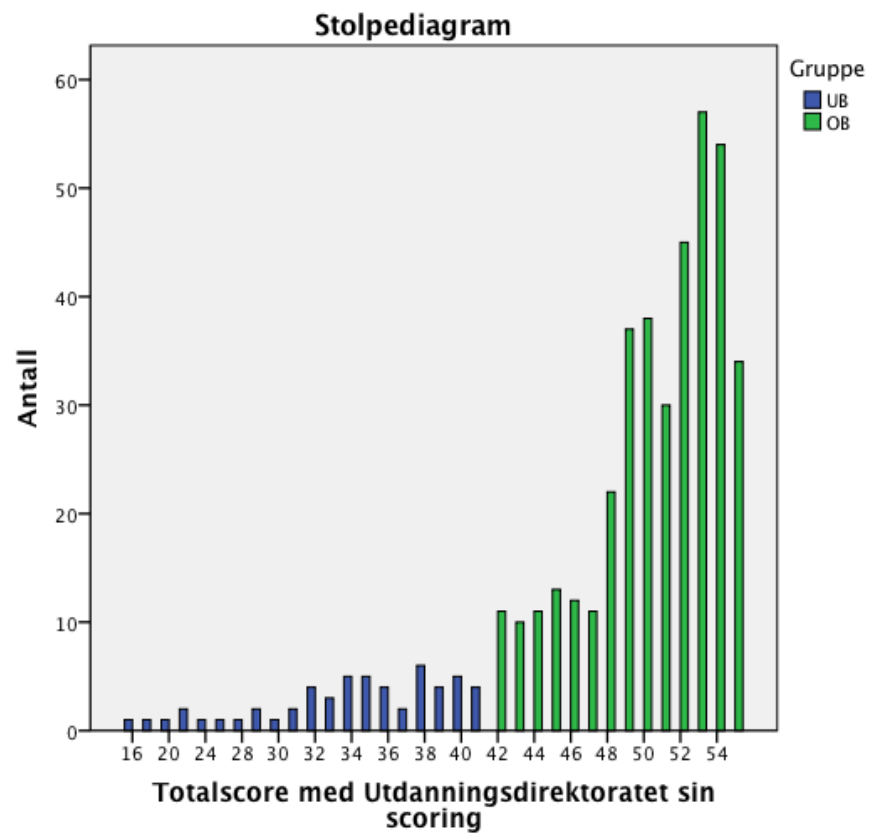
bekymringsgrensen risikerer å få gjennomsnittsverdier som ikke nødvendigvis gir det riktige bilde, i og med at gjennomsnittet er sensitivt for ekstremverdier og gruppen er såpass liten i størrelse. Gruppen Over bekymringsgrensen har færre slike ekstremverdier og de er også flere i antall, slik at gjennomsnittsverdien her ikke blir like mye påvirket som for gruppen Under bekymringsgrensen. For sidescore er variasjonsbredden lik for gruppene, men størrelsen på gruppene er ulik og det vil i sin tur påvirke effekten av ekstremverdier. Likhet innad i gruppen kalles i statistikk varians. En av betingelsene for å gjøre parametrisk statistikk er at det skal være lik varians innad i gruppene en ønsker å sammenligne. Mine to grupper oppfyller denne betingelsen på noen oppgaver, men for det meste er det ulik varians mellom de to. For å undersøke varians i gruppene kan en gjøre en *Levene's test for equality of variance* (Pallant, 2013). SPSS gjør automatisk en *Levene's test for equality of variance* når du ber den om å gjøre en *Student's t-test*. Av resultatet vil det komme frem om variansen er lik, eller ulik og tilhørende t- og p-verdi. Kort fortalt vil den si om gruppenes gjennomsnittsverdi er statistisk signifikant forskjellig, og om variansen mellom gruppene er lik eller ulik. Det er stilt spørsmålsteget ved bruk av Levene's test ved store utvalg, fordi selv små forskjeller mellom gruppene kan kunne gi signifikante utslag, uten at det i realiteten er stor forskjell (Field, 2013). Derfor anbefales det at materialet heller undersøkes deskriptivt først for å bedømme likhet i varians, og på bakgrunn av det bestemme om det er hensiktsmessig å benytte Welch's test (Welch, 1938) eller en student t-test.

Den største utfordringen er likevel at gruppene jeg sammenligner er så ulike i størrelse. Det er frarådet å bruke parametrisk statistikk når gruppestørrelsen avviker med mer en 1,5 og det i tillegg er ulik varians i gruppene, eksempelvis at standardavviket i den ene er under halvparten til den andre (Pallant, 2013). SPSS gjør automatisk en Welch's (Welch, 1938) t-test når gruppene en sammenligner har ulik varians. En Welch's t-test har sin styrke i at den korrigerer frihetsgradene brukt i sammenligningen mellom de to gruppene, når gruppene har ulik varians og størrelse. I SPSS kommer dette frem som sig. og t-verdier for *Equality of variance not assumed* under *Levene's test for equality of variance*, men er i realiteten en Welch's t-test. Frihetsgradene som blir presentert i resultatdelen vil derfor varierer fra oppgave til oppgave, selv om det er nøyaktig de samme 440 svarene som er analysert.

I og med at materialet mitt er utsatt for ekstremverdier, har jeg valgt og også ta med den ikke-parametriske varianten av t-test, som er Mann-Whitney U test. Denne testen rangerer hele materialet, altså begge gruppene. Laveste score får rang 1, neste rang 2 og så videre. Etterpå summeres rangene for hver av gruppene og et gruppegjennomsnitt for rang blir beregnet. Det vil si at denne testen ikke er like sensitiv for ekstremverdier. En Mann-Whitney U test vil på samme måte som en Welch's T- test gi en z -verdi og en p-verdi. Disse verdiene kan en regne ut effekt fra, henholdsvis eta squared for t-testen og r-verdi for Mann-Whitney U test. Jeg har valgt å foreta begge testene på mitt materiale og ber leseren være oppmerksom på ekstremverdier og deres betydning for resultatet i tolkning av verdiene. For leservennlighet har jeg valgt å beskrive resultatet for poengscore for oppgavesiden i detalj, mens testene som er gjort på de enkelte oppgavene kun blir presentert med effektstørrelse og p-verdi. Jeg gjør oppmerksom på at forkortelsene UB for gruppen Under bekymringsgrensen og OB for gruppen Over bekymringsgrensen blir brukt i tabeller og figurer av plasshensyn.

### **3.6 Datamaterialet og utvalg**

Da jeg mottok datafilen var det kun råmaterialet som var registrert og kontrollert, det vil si de nøyaktige svarene til elevene på de ulike oppgavene. Det var brukt kode 999 for ubesvart og kode 666 for uleselig/uregistrerbart svar. Filen var også sjekket for registreringsfeil ved vanlig metode (Pallant, 2013). Mitt datamateriale besto i utgangspunktet av 451 kartleggingsprøver fra elever på 2. trinn i Stavanger kommune. Jeg valgte å ta bort 11 elevers prøver hvor vi kun hadde Utdanningsdirektoratet sin sidscore og ikke mulighet for å score på oppgavenivå grunnet makulering av prøveheftet av lærer. Det vil si at mitt endelige materiale består av 440 elever på 2. trinn. Totalsum som er mulig å oppnå ved Utdanningsdirektoratet sin scoring er 55 poeng, ved Stavangerprosjektet sin scoring 82 poeng. Årsaken til at Stavangerprosjektet har høyere totalscore er at scoringen deres er gjort på nedbrutt itemnivå, mens Utdanningsdirektoratet scorer på oppgavenivå. Nærmere forklart hvis en oppgave består av 5 underoppgaver, vil eleven kunne score 5 poeng ved Stavangerprosjektet sin scoring og kun 1 poeng ved Utdanningsdirektoratet sin scoring. Det er fire sider i *Kartleggingsprøve i regning for 2. trinn* hvor dette er gjeldende. Sentraltendenser i materialet er gjengitt på de neste sidene.



Figur 2 Fordeling av poengscore ved de to ulike scoringene

**Tabell 5 Sentralmål for utvalget**

|                        |         | Utdannings-<br>direktoratet<br>Totalscore | Stavanger-<br>prosjektet<br>Totalscore |
|------------------------|---------|---|--|
| N                      | Valid   | 440                                       | 440                                    |
|                        | Missing | 0   | 0                                      |
| Mean                   |         | 48.50                                     | 73.79                                  |
| Median                 |         | 50.50                                     | 77.00                                  |
| Mode                   |         | 53  | 81                                     |
| Std. Deviation         |         | 6.755                                     | 9.265                                  |
| Skewness               |         | -1.879                                    | -2.007                                 |
| Std. Error of Skewness |         | .116                                      | .116                                   |
| Minimum                |         | 16  | 23                                     |
| Maximum                |         | 55  | 82                                     |
| Percentiles            | 25      | 46.00                                     | 71.00                                  |
|                        | 50      | 50.50                                     | 77.00                                  |
|                        | 75      | 53.00                                     | 80.00                                  |

Gruppeoversikt etter materialet er delt i to, etter henholdsvis Utdanningsdirektoratet sin bekymringsgrense og Stavangerprosjektet sin bekymringsgrense, vises i Tabell 6 og 7.

**Tabell 6 Gruppeoversikt med Utdanningsdirektoratet sin scoring**

| Udir.      | N   | Prosent | Mean  | Median | Std. Deviation | Minimum | Maximum |
|------------|-----|---------|-------|--------|----------------|---------|---------|
| TotalScore |     |         |       |        |                |         |         |
| UB         | 55  | 12,5    | 33.89 | 35.00  | 6.057          | 16      | 41      |
| OB         | 385 | 87,5    | 50.59 | 51.00  | 3.474          | 42      | 55      |
| Total      | 440 | 100     | 48.50 | 50.50  | 6.755          | 16      | 55      |

**Tabell 7 Gruppeoversikt med Stavangerprosjektet sin scoring**

| SP.        | N   | Prosent | Mean  | Median | Std. Deviation | Minimum | Maximum |
|------------|-----|---------|-------|--------|----------------|---------|---------|
| TotalScore |     |         |       |        |                |         |         |
| UB         | 65  | 14,8    | 55.12 | 57.00  | 8.663          | 23      | 66      |
| OB         | 375 | 85,2    | 77.02 | 78.00  | 4.106          | 67      | 82      |
| Total      | 440 | 100     | 73.79 | 77.00  | 9.265          | 23      | 82      |

Av Tabell 6 og 7 fremgår det at Stavangerprosjektet sin scoring gjør at en får noen flere (+10 stk) elever under bekymringsgrensen kontra Utdanningsdirektoratet sin scoring. Jeg valgte å forholde meg til Utdanningsdirektoratet sin scoringsguide og bekymringsgrense, med bakgrunn i at oppgavesidene jeg så nærmere på ble scoret likt fra henholdsvis Utdanningsdirektoratet og Stavangerprosjektet. Eventuelle forskjeller mellom disse to delingene på de oppgavesidene jeg så på, ville i så fall ikke skyldtes oppgavesidene jeg så på, men oppgavesidene jeg ikke så på.

For å sikre at oppgavesidene jeg hadde valgt ut var representative for hele *Kartleggingsprøve i regning for 2. trinn* gjennomførte jeg en Welch's test og en Mann Whitney U test på **totalscoren** for de to gruppene, Over- og Under bekymringsgrensen. Disse testene ble gjentatt for oppgavesidene som var valgt ut slik at effektstørrelsene kunne sammenlignes. Resultatene vises i Tabell 9, samt i Figur 3 og Figur 4. Cohen (1988) sine standardforslag for vurdering av effekt er et av de mest brukte i feltet. Jeg fulgte hans forslag når effekter omtales som liten, medium/moderat og stor. Cohen (1988) presiserte at disse forslagene er gitt for å kunne sammenligne effekter mellom studier uavhengig av målekategori og variabler, og at dette medfører at noe av egenheten i studiene faller bort. Vurderingen var likevel at gevinsten overgikk tapet ved å ha en standard for alle. Han understreker likevel at standarden bare bør brukes når ikke bedre alternativer foreligger (Cohen, 1988, s. 25).

The terms "small", "medium", and "large" are relative, not only to each other, but to the area for behavioral science or even more particularly to the specific content and research method being employed in any given investigation (Cohen, 1988, s. 25).

**Tabell 8 Cohens standardforslag**

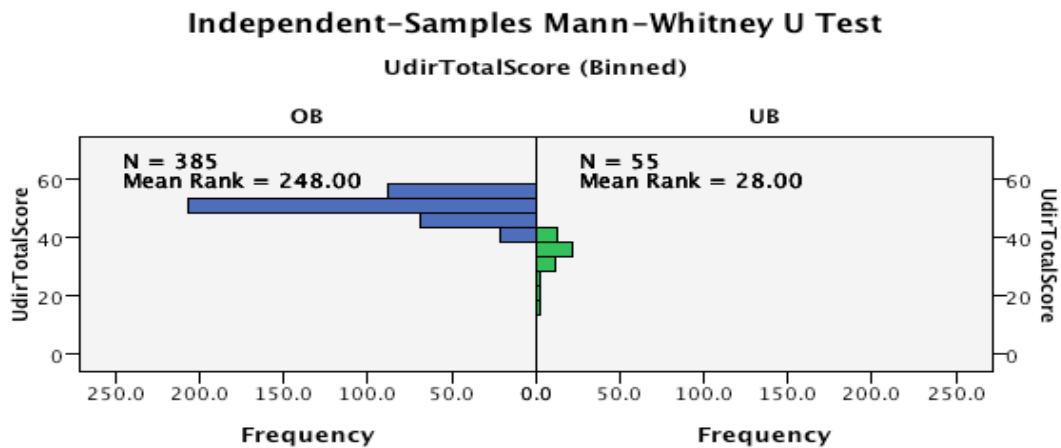
|                        | R-verdier | Eta squared |
|------------------------|-----------|-------------|
| Liten effekt           | 0.1       | 0.01        |
| Middels/moderat effekt | 0.3       | 0.06        |
| Stor effekt            | 0.5       | 0.14        |

**Tabell 9 Independent samples test for gruppenes totalscore**

|                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |        |                 |                 |                       |   |         |
|-----------------------------|---|------|------------------------------|--------|-----------------|-----------------|-----------------------|---|---------|
|                             | F                                       | Sig. | t                            | df     | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |         |
|                             |   |      |                              |        |                 |                 |                       | Lower                                     | Upper   |
| Equal variances assumed     | 25.539                                  | .000 | -29.800                      | 438    | .000            | -16.696         | .560                  | -17.797                                   | -15.595 |
| Equal variances not assumed |   |      | -19.977                      | 59.177 | .000            | -16.696         | .836                  | -18.368                                   | -15.024 |

En Welch's test viste statistisk signifikant forskjell mellom gruppen Under bekymringsgrensen ( $M = 33.89$ ,  $SD = 6.057$ ) og gruppen Over bekymringsgrensen ( $M = 50.59$ ,  $SD = 3.474$ ;  $t(59) = -19.977$ ,  $p < .001$  tohalet). Med bakgrunn i antatt ulik varians ble frihetsgradene justert fra 440 til 59. Størrelsen på forskjellen i gjennomsnittene (Mean difference =  $-16.696$ , 95% CI:  $-18.368$  til  $-15.024$ ) var stor ( $\eta^2 = 0.477 = 47,7\%$ ).



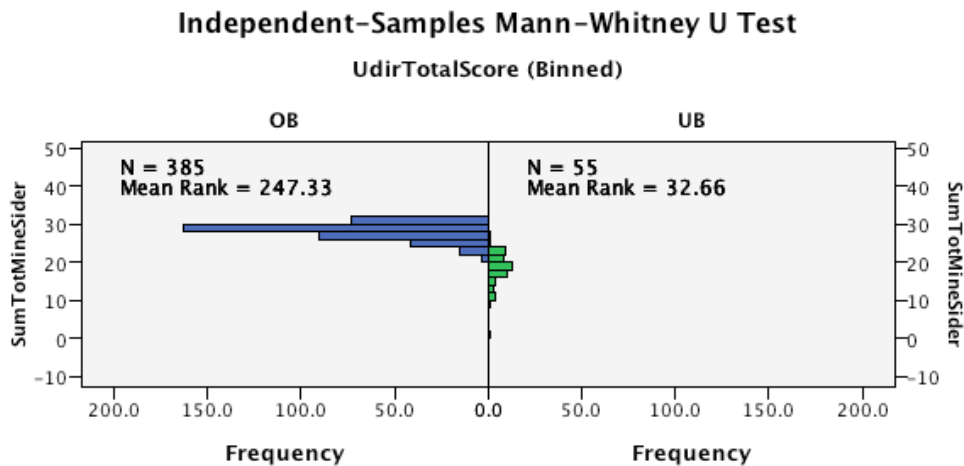


|                                       |            |
|---------------------------------------|------------|
| <b>Total N</b>                        | 440        |
| <b>Mann-Whitney U</b>                 | 21,175.000 |
| <b>Wilcoxon W</b>                     | 95,480.000 |
| <b>Test Statistic</b>                 | 21,175.000 |
| <b>Standard Error</b>                 | 878.900    |
| <b>Standardized Test Statistic</b>    | 12.046     |
| <b>Asymptotic Sig. (2-sided test)</b> | .000       |

Figur 3 Mann-Whitney U Test på totalscore

I og med at datamaterialet som ble undersøkt ikke er normalfordelt foretok jeg også en Mann-Whitney U test som indikerte en statistisk signifikant forskjell i total sumscore mellom gruppen Under bekymringsgrensen ( $Md = 35$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $Md = 51$ ,  $n = 385$ ),  $U = 21\,175$ ,  $z = 12.046$ ,  $p < .001$ , effekten var stor,  $r = 0.57$ .

Resultatet av testene ovenfor var forventet, i og med at bekymringsgrensen er satt utfra totalscore. Jeg repeterte disse testene for å sammenligne med oppgavesidene jeg selv har valgt ut. Poengsum som er mulig å oppnå på de utvalgte sidene til sammen er 30. Under er resultatet av Mann-Whitney gjort på kun sumscore av de oppgavesidene jeg så nærmere på.



|                                       |            |
|---------------------------------------|------------|
| <b>Total N</b>                        | 440        |
| <b>Mann-Whitney U</b>                 | 20,918.500 |
| <b>Wilcoxon W</b>                     | 95,223.500 |
| <b>Test Statistic</b>                 | 20,918.500 |
| <b>Standard Error</b>                 | 872.602    |
| <b>Standardized Test Statistic</b>    | 11.839     |
| <b>Asymptotic Sig. (2-sided test)</b> | .000       |

**Figur 4** Mann-Whitney U test med utvalgte sider

Mann-Whitney U test indikerer statistisk signifikant forskjell i sumscore på utvalgte oppgavesider mellom gruppen Under bekymringsgrensen ( $Md = 18$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $Md = 28$ ,  $n = 385$ ),  $U = 20\,918.5$ ,  $z = 11.84$ ,  $p < .001$ , effekten var stor,  $r = 0.56$ .

En Welch's test ble også gjennomført og indikerte statistisk signifikant forskjell mellom de to gruppene på sumscore på mine oppgavesider. Gruppen Under bekymringsgrensen ( $M = 17.73$ ,  $SD = 4.548$ ) og gruppen Over bekymringsgrensen ( $M = 27.68$ ,  $SD = 2.059$ ;  $t(57) = -15.994$ ,  $p < .001$  tohalet). Størrelsen på forskjellen i gjennomsnittene (Mean difference =  $-9.951$ , 95% CI:  $-11.196$  til  $-8.705$ ) var stor (eta squared =  $0.3687 = 36,9\%$  av variansen forklares av gruppetilhørigheten).

Resultatet fra testene viser at sidene jeg har valgt å studere har sammenfallende effekt for gruppetilhørighet som totalscoren. En kan med andre ord anta at fordelingen av score mellom gruppene er lik på sidene jeg studerer og *Kartleggingsprøve i regning for 2. trinn* totalt. Resultatet antyder at oppgavesidene som er utvalgt står for en stor del av oppgavesidene elvene i gruppen Under bekymringsgrensen scorer lavt på.

### 3.7 Kategorisering av svar

I de **oppstilte regneoppgavene** sorterte jeg svarene til elevene i 8 ulike kategorier. Jeg har brukt kategoriene til Lindvall og Ibarra som utgangspunkt og utvidet med en kategori (Lindvall & Ibarra, 1980). I tillegg til ”oppgir a” har jeg også med en kategori som er ”oppgir et siffer fra a”. Denne kategorien gjør seg hovedsakelig gjeldende i oppgaver av kategori RR5/6. Eksempelvis i en regneoppgave av kategori RR5, der  $a = 12$  og  $b = 2$ :

$$a = \text{elevsvar} + b$$

$$12 = \text{elevsvar} + 2$$

Mulige svar som havnet i den nye kategorien er: 1 og 20. Sifferet 2 blir ikke med da det regnes som ”gi b”. Jeg er innforstått med at 20 ikke er et siffer i 12, men det er en fortolkning av eleven som kan tyde på mangelfull forståelse for posisjonssystemet, ved at han/hun har byttet verdi på tier og enerplassen. De 8 kategoriene svarene er blitt sortert i er da følgende:

M1 = Riktig svar

M2 = Ubesvart

M3 = +/- 1 fra riktig svar (tellefeil)

M4 = gjengir a

M5 = gjengir b

M6 = adderer gitte tall

M7 = annet (svar som ikke passer i noen av kategoriene)

M8 = gir siffer fra a

Der variabel  $a$  var det sifferet/tallet som sto først i leseretning venstre til høyre og variabel  $b$  det neste. I noen oppgaver kunne enkelte svar passe både i kategori M4/5 og M3. Da ble M4/5 valgt, fordi en tellefeil ved addisjon av 1 er mindre sannsynlig

enn ren kopi av tallet som stod der. Kategoriene over er tilknyttet årsaker for feilmønstre. Der en tenker at *ubesvart* (M2) hovedsakelig dreier seg om tidsfaktoren og eventuell avkoding av oppgaven hvis den er lik for begge gruppene. I den forstand at uhensiktsmessige strategier vil medføre at eleven scorer under bekymringsgrensen, så hvis *ubesvart* (M2) er like fremtredende i begge grupper, kan det indikere at årsaken er noe annet enn uhensiktsmessige strategier. Eksempelvis avkoding av oppgaven, eller at oppgaven var utfordrende for begge gruppene. M3 er tellefeil, og er således knyttet til svake regne- og tellestrategier. I den forstand at tellestrategier hovedsakelig er backup- strategier som tyder på at regnefaktaene ikke er automatisert. *Tellefeil* (M3) i kombinasjon med *ubesvart* (M2), vil være med å styrke en antakelse om en elev med svake telle- og regneferdigheter basert på svake telle- og regnestrategier. *Gir a* (M4), *gir b* (M5), *adderer gitte tall* (M6) er alle knyttet til forståelse av likhetstegnet. En slik feil kommer ikke av tellefeil, men heller av en grunnleggende svak eller mangelfull forståelse av likhetstegnet. *Gir siffer fra a* (M8) har nærmere tilknytting til forståelse av posisjonssystemet, da en slik feil antyder feil i avlesning av et siffer eller tall, *gir siffer fra a* (M8) kan også tolkes som svak eller mangelfull forståelse av likhetstegnet. Hvilken av de to årsakene som gjør seg gjeldende for ulike oppgaver, vil kunne belyses ved analyse av instruksjonen gitt av lærer på de ulike oppgavesidene. Har instruksjonen fokus på å fylle inn eller finne et tall, kan manglende forståelse for posisjonssystemet være en sannsynlig årsak. Har instruksjonen derimot fokus på å gjøre regnestykket riktig, eller opprettholde ekvivalensen, vil svak eller manglende forståelse for likhetstegnet være en mer sannsynlig årsak for feilmønstret. *Gir siffer fra a* (M8) er med andre kontekstavhengig med tanke på årsak for feilmønstret, det er begrunnelsen for at feilmønstret blir egen kategori og ikke slått sammen med noe av de andre feilmønstrene. Kategori *annet* (M7) er svar som ikke har kvalifisert til noen av kategoriene. Alle elevsvarene på de oppstilte regneoppgavene ble sortert i de åtte kategoriene, og kategorier med mer enn 4 svar ble omtalt som mønster.

I **tekstoppgavene** ble arbeidet gjort mer suksessivt ved å se på hva elevenes faktiske svar. Tekstoppgavene jeg har studert har ikke oppstilte regneoppgaver, derfor ble kategorier som *gir a* og *gir b* lite hensiktsmessig, da jeg ikke med sikkerhet kunne si hvilken *a* eller *b* elevene benyttet. Jeg satte også en grense her på minimum 4 elever

til sammen som måtte ha avgitt svaret for at jeg kunne kalle det et mønster. Det vil si at kategoriene her er:

M1 = Riktig

M2 = Ubesvart

M3 = Tellefeil

M7 = Annet

Tn = Mønstre som fremkom gjennom analysen

Jeg er klar over at det kan virke noe søkt å kategorisere etter hva analysene viser, men hensikten har vært å forsøke å finne ut hva elevene kan ha tenkt på sin vei til svaret, eventuelt hvor gikk de feil. Det er derfor et metodisk poeng å se på hva de faktisk har svart, og ikke alle feilmønstrene de har unngått å gå i. Jeg understreker at min tolkning av feilmønstre er tentativ, som alle andres, men anser den for den beste tilgjengelige hypotese så fremt motstridende bevis ikke finnes. Det viste seg å være hovedsakelig to mønstre som gikk igjen i tekstoppavene. Eller det vil si at det kunne se ut til at det var to hovedårsaker til ulike feil som ble gjort, og at det følgelig ble to mønstre. Den ene årsaken kan indikere utfordringer tilknyttet til tieroverganger, eventuelt bruken av algoritmen. Denne kategorien rommer svar som avviker med 10 fra riktig svar og er navngitt *Tier*. Den andre årsaken så ut til å ha tettere tilknytning til selve illustrasjonen, og innebærer å lese av feil eller bruke feil informasjon fra illustrasjonene. Disse svarene er ulike fra oppgave til oppgave, men essensen er at eleven mest sannsynlig har lest av feil i illustrasjonen. Denne kategorien fikk navnet *Till*. Eksempel på svar som er i denne kategorien kan være summen av feil tall i en illustrasjon. Disse to kategoriene ble resultatet etter den innledende analysen, i tillegg er de vante kategoriene riktig, tellefeil (+/- 1 fra riktig svar) og ubesvart med.

Det har vært utfordrende å skulle presentere de ulike svarene og mønstrene uten å bryte konfidensialiteten med tanke på selve oppgavene i kartleggingsprøven. Jeg har forsøkt etter beste evne å beskrive mønstrene metodisk, og ikke konkret, samtidig som jeg ønsket å presentere mønstrene tydelig. For leservennlighet vil kategoriene bli repetert ved presentasjon av resultatet.

Et viktig aspekt ved kartleggingsprøver er at den skal identifisere de elevene som trenger særlig oppfølging. For å kunne avgjøre om feilmønsteret grunnes svake

begreper og svake telle- og regneferdigheter eller avkoding av selve oppgaven, har jeg sett på utvalgets totalscore på hver oppgaveside. P-verdi brukes i dette tilfellet om prosent elever som greier en oppgave. ”Når det gjelder kartleggingsprøver, bør de fleste oppgavene ha p-verdier over 0,70, mens p-verdiene ikke bør ligge under 0,50” (Hopfenbeck, Kjærnsli & Olsen, 2012, s. 194). Det vil si at en ønsker at 70% eller flere av elevene svarer korrekt på oppgavene, og man ønsker helst ikke at færre enn 50% skal avgi korrekt svar. Det vil si at hvis andelen elever som har fått full poengscore for en side er lav, kan dette tolkes som at selve oppgavesiden er en utfordring for hele utvalget, og ikke nødvendigvis forbundet med svake begreper og svake telle- og regneferdigheter. Derfor er også deskriptive fremstillinger av score med i analysen min.

### **3.8 Validitet**

Validitet kan tenkes på som gyldighet, og peker på hvor gyldig resultatet er. Statistisk validitet er et spørsmål om generalisering fra utvalg til populasjon (Johannessen, et al., 2010, s. 357). Mitt utvalg er ett årskull fra Stavangerprosjektet og populasjonen er alle 2. klassinger i Norge som tar *Kartleggingsprøven i regning for 2. trinn*. For å kunne generalisere til populasjonen er det visse krav som må være oppfylt. Utvalget jeg benytter møter ikke alle disse kravene, eksempelvis er den geografiske plasseringen Stavanger. Det er likevel grunn til å tror at funn som gjøres i denne studien kan ha relevans for den øvrige populasjonen.

### **3.9 Reliabilitet**

Reliabilitet går på troverdighet, eller målestyrke til analysen. Mer konkret kan en si om det ville blitt samme resultat hvis noen gjorde analysen etter meg. Hvis det blir det samme resultat uavhengig av hvem som gjør analysene er reliabiliteten god.

”Reliabilitet knytter seg til nøyaktigheten av undersøkelsens data, hvilke data som brukes, den måten de samles inn på, og hvordan de bearbeides” (Johannessen, et al., 2010, s. 40). Denne studien bruker data fra *Kartleggingsprøve i regning for 2. trinn* som er del av Stavangerprosjektet sitt datamateriell. Stavangerprosjektet registrerer sine egne data i SPSS direkte fra elevenes prøvehefter, og det er utarbeidet gode kontrollrutiner for registreringen. Rettingen og scoringen er gjort i henhold til

instruksjonen gitt fra Utdanningsdirektoratet. Reliabiliteten til dataene må således kunne sies å være god.

Når det kommer til bearbeidelsen av dataene er dette først om fremst relevant angående kategoriseringen av svarene. Jeg har etterstrebet å gjøre dette så gjennomiktig som mulig, slik at det er mulighet for andre å følge opp studien.

### **3.10 Forskningsetiske hensyn**

Det er et ufravikelig krav i forskningsetikken at man ikke skal betrakte et annet menneske bare som et middel, men alltid også som et mål i seg selv (Nesh, 2006). Dette er noe av grunnlaget til at jeg ikke bare ser på scoren til elevene, men også hva som ligger bak svarene, for å på den måten kunne bidra til videreutvikling av kompetanse. Å forske på svakstilte grupper krever særskilt etisk vurdering, for å unngå stigmatisering. Gruppen denne studien har fokus på er elever med lav score på *Kartleggingsprøve i regning for 2. trinn*. Det har vært min hensikt i arbeidet å bidra til et konstruktivt løft for denne gruppen, og ikke videre stigmatisering. Anonymiteten til elevene er bevart gjennom at materialet allerede var anonymisert før jeg fikk det. I Stavangerprosjektet er det kun to som har tilgang til denne omkodingsnøkkelen. Stavangerprosjektet er tilrådd av Datafaglig sekretariat og følger De nasjonale forskningsetiske komiteene sine retningslinjer for samfunnsvitenskap, humaniora, juss og teologi (NESH). Når det kommer til *Kartleggingsprøve i regning for 2. trinn* er det lagt ned betydelig arbeid av flere for å lage denne. Derfor søker jeg å presentere mine funn på en respektfull og samvittighetsfull måte. Åpenhet rundt mine egne refleksjoner og begrensningene av dem er en del av dette. Åpenhet rundt funn og refleksjoner, men også gjennomsiktighet i analyseprosessen er viktig, med tanke på hvordan mine funn kan brukes videre av fagfeltet. Når datamaterialet har blitt behandlet er det lagt vekt på å fremstille data nøyaktig og samvittighetsfullt og i tråd med metodisk teori, slik at grunnlaget for analysene reflekterer de innsamlede data.

## 4 Resultater og analyse

I denne delen vil resultatene fra de analyserte data presenteres. Tabeller og fremstillinger som vises er hentet fra statistikkprogrammet SPSS. Der SPSS sine modeller ble for omfattende laget jeg dem i Word. Først kommenteres funnene fra den deskriptive analysen, før funnene fra de syv oppgavesidene blir beskrevet.

Opgavesidene blir presentert under egne delkapitler. Delkapitlene starter med en presentasjon av siden som helhet, før det blir kommentert på oppgavene siden besto av. Avslutningsvis i hvert delkapittel oppsummeres analysen for siden. Resultatene fra Welchs´s test og Mann Whitney U test blir rapportert både samlet for siden og for hver oppgave på siden. Statistiske begreper som blir benyttet er gjort rede for i metodekapittelet.

### 4.1 Deskriptiv analyse

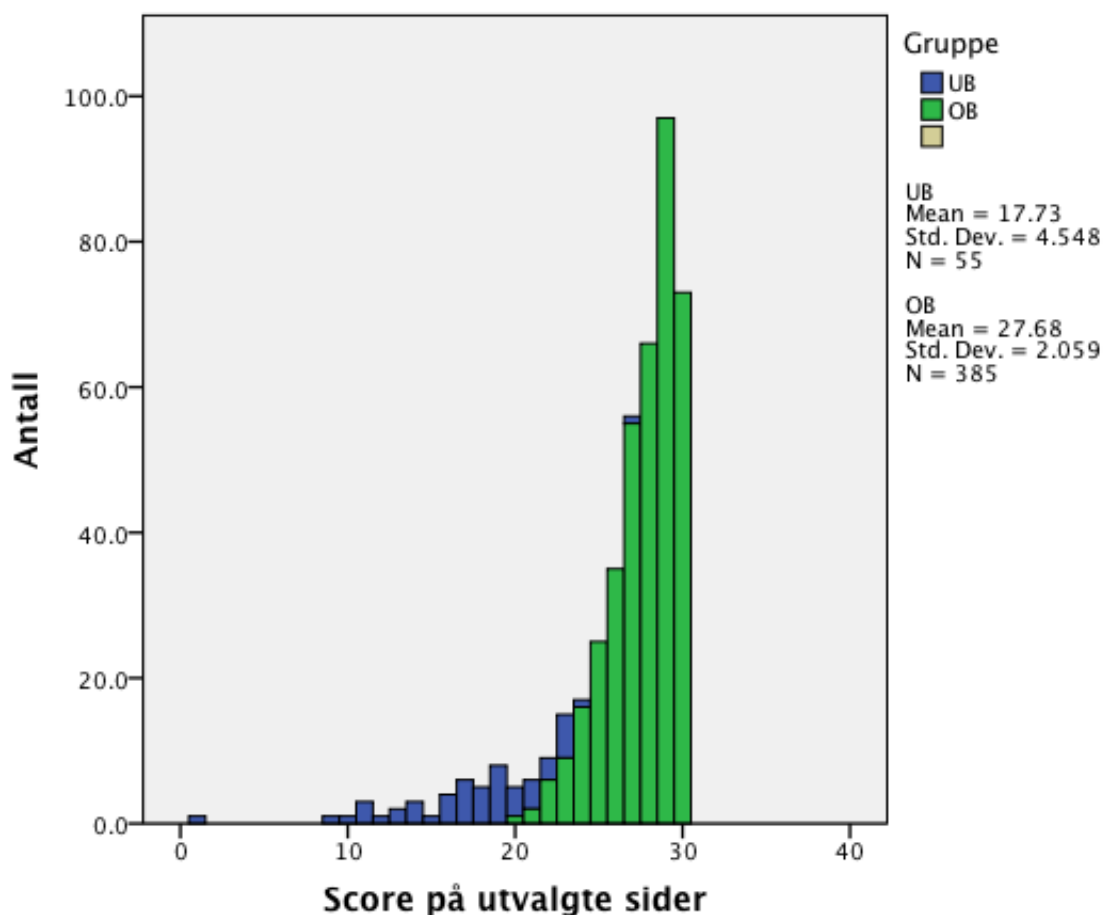
De deskriptive analysene viser at *Kartleggingsprøve i regning for 2. trinn* har den ønskede takeffekten med en negativ skewness på 1.879 for Utdanningsdirektoratet sin totalscore på prøven. Det vil si at flesteparten av elevene scorer i den øvre enden av poengskalaen, alt eller det meste rett, og at gjennomsnittet er over bekymringsgrensen. Når det kommer til prosentandel under bekymringsgrensen er denne lavere i studiens utvalg enn hva den nasjonale cutoffen på 20% var.

Henholdsvis 12,5% av elevene er under bekymringsgrensen med Utdanningsdirektoratet sin scoring og 14,8% av elevene er under bekymringsgrensen ved Stavangerprosjektet sin scoring. Det er vanskelig å konkludere noe på bakgrunn av dette, annet enn at utvalget denne studien har brukt scorer høyere enn utvalget som bekymringsgrensen er satt ut fra. Dette kan skyldes flere årsaker, for eksempel at utvalget er fra en større by, men det kan ikke trekkes noen konklusjoner rundt dette da studiens utvalg ikke er representativt for populasjonen.

Forskjellen som kommer frem mellom de to scoringene, Utdanningsdirektoratet og Stavangerprosjektet, viser at en får noen flere elever under bekymringsgrensen ved Stavangerprosjektet sin scoring. En får 10 flere elever i gruppen Under bekymringsgrensen med bruk av Stavangerprosjektet sin scoring. 3 elever som scorer over bekymringsgrensen med Stavangerprosjektet sin scoring, har henholdsvis 40, 41 og 41 poeng med Utdanningsdirektoratet sin scoring. Utdanningsdirektoratet har



bekymringsgrensen sin på 41 poeng. 4 elever som scorer under bekymringsgrensen i Stavangerprosjektet scorer 42 poeng med Utdanningsdirektoratet sin scoring. Hovedvekten av diskrepansen finner vi på Utdanningsdirektoratets score 45, med 7 elever som scorer under bekymringsgrensen i Stavangerprosjektet. Det var 2 elever til som skilte seg ut ved Stavangerprosjektet sin scoring, henholdsvis en elev med score 45 og en med score 46 med Utdanningsdirektoratet sin scoring. Oppsummert for disse elevene kan en si de får poeng fra Stavangerprosjektet der de har gjort riktig, mens Utdanningsdirektoratet sin scoring er noe strengere og gir først poeng hvis hele oppgaven er riktig. Oppgavesidene som scores ulikt er hovedsakelig basert rundt telling, tallinjen og begreper. En oversikt over elevens totalscore på de utvalgte sidene vises i Figur 5.



Figur 5 Elevenes score på utvalgte sider

Av standardavviket og fordelingen i histogrammet blir det tydelig at variansen er større hos gruppen Under bekymringsgrensen enn gruppen Over bekymringsgrensen også på sidene jeg har sett nærmere på. I tillegg er det verdt å merke seg at noen elever fra gruppen Under bekymringsgrensen scorer relativt høyt med tanke på at de kun trenger 17 poeng for å komme over bekymringsgrensen, hvis de har alt annet rett på *Kartleggingsprøve i regning for 2. trinn*. Oppgavesidene som er analysert antas å være de vanskeligste i *Kartleggingsprøve i regning for 2. trinn*. Med bakgrunn i at disse oppgavene krever mer av elevene, kan det fremstå som noe motstridende at flere elever i gruppen Under bekymringsgrensen scorer relativt høyt på de utvalgte sidene. De utvalgte oppgavesidene og analysen av disse presenteres nå fortløpende. Hver oppgaveside er kort beskrevet for leservennlighet. De konkrete regneoppgavene er ikke oppgitt, grunnet konfidensialitet, oppgavene er derfor beskrevet metodisk.

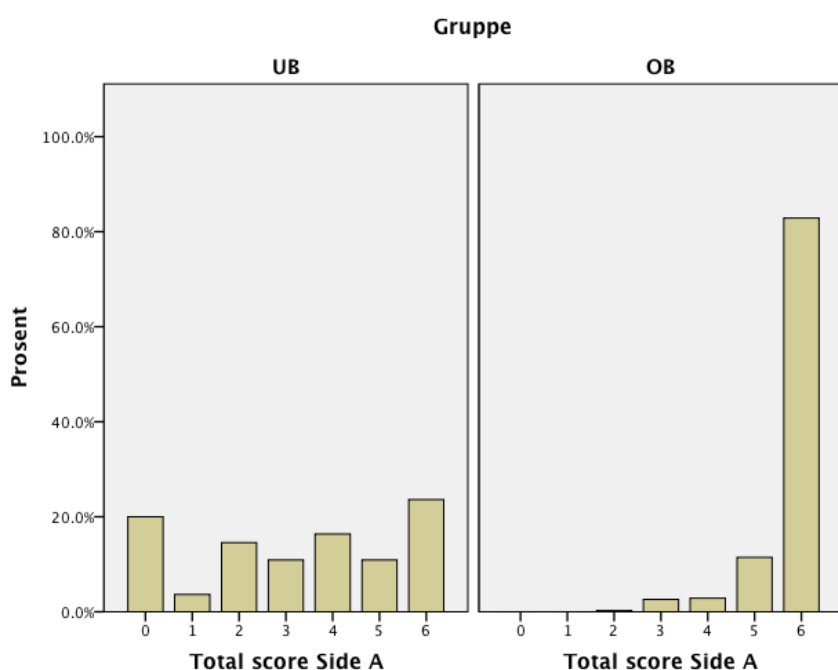
## 4.2 Side A

Overskriften indikerer at denne siden omhandler posisjonssystemet og oppgavene er 6 addisjonsoppgaver av både variant RR5 og RR6:

RR5:  $a = \text{elevsvar} + b$

RR6:  $a = b + \text{elevsvar}$

Konsekvent gjennom siden har oppgaver av variant RR5 hele tiersvar og oppgaver av variant RR6 enersvar. Oppgavenummer 1, 3 og 5 er av kategori RR6 og oppgavenummer 2, 4 og 6 er av kategori RR5. Fordelingen av poengscore på siden for de to gruppene er vist i Figur 6.



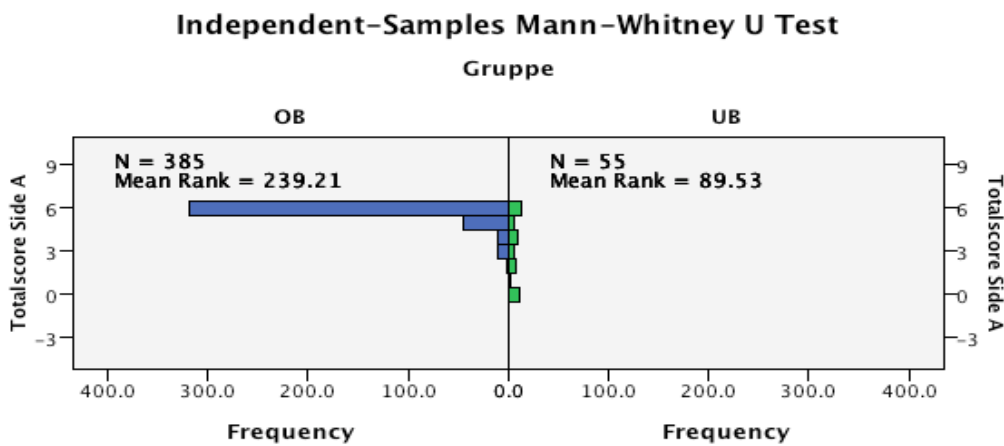
Figur 6 Poengscore på Side A prosentvis i gruppene

En får den forventet takeffekten i gruppen Over bekymringsgrensen, mens bildet er noe mer utydelig for gruppen Under bekymringsgrensen. Ved å omgjøre diagrammet til en tabell med kumulativ prosent blir forskjellen mellom de to gruppene tydeligere.

**Tabell 10 Score gruppevis**

| Poeng | Kumulativ prosent |       |
|-------|-------------------|-------|
|       | UB                | OB    |
| 0     | 20%               | 0%    |
| 1     | 23,6%             | 0%    |
| 2     | 38,1%             | 0,3%  |
| 3     | 49%               | 2,9%  |
| 4     | 65,4%             | 5,8%  |
| 5     | 76,3%             | 17,2% |
| 6     | 100 %             | 100%  |

Fra Tabell 10 blir det tydelig at over 80% av elevene fra gruppen Over bekymringsgrensen fikk full score. Knappe 24% av elevene fra gruppen Under bekymringsgrensen fikk full score. Tabellen viser også at minimum sidescore for gruppen Over bekymringsgrensen var 2 poeng. Situasjonen for gruppen Under bekymringsgrensen er at nesten halvparten (49%) av elevene fikk 3 poeng eller mindre på Side A. Resultatet fra Mann-Whitney U test vises i Figur 7.



|                                       |            |
|---------------------------------------|------------|
| <b>Total N</b>                        | 440        |
| <b>Mann-Whitney U</b>                 | 17,791.000 |
| <b>Wilcoxon W</b>                     | 92,096.000 |
| <b>Test Statistic</b>                 | 17,791.000 |
| <b>Standard Error</b>                 | 665.291    |
| <b>Standardized Test Statistic</b>    | 10.828     |
| <b>Asymptotic Sig. (2-sided test)</b> | .000       |

**Figur 7 Mann-Whitney U test sidescore Side A**

En Mann-Whitney U test indikerte en statistisk signifikant forskjell i score på Side A, mellom gruppen Under bekymringsgrensen ( $Md = 4$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $Md = 6$ ,  $n = 385$ ),  $U = 3384$ ,  $z = 10.828$ ,  $p < .001$ , effekten var stor  $r = 0.52$ .

En Welch's test ble også gjennomført for å sammenligne sidescore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Den indikerte statistisk signifikant forskjell i score mellom gruppen Under bekymringsgrensen ( $M = 3.27$ ,  $SD = 2.198$ ) og gruppen Over bekymringsgrensen ( $M = 5.74$ ,  $SD = .661$ ;  $t(55) = -8.272$ ,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference =  $-2.468$ , 95% CI =  $-3.065$  til  $-1.870$ ) var opp mot stor (eta squared =  $0.1351 = 13,5\%$ ).

Det er tydelig at denne oppgavesiden skiller godt mellom de to gruppene, med  $r = 0.52$  og en eta squared på 13,5%. Videre presenteres svarene fordelt i de ulike kategoriene som fremkom på Side A, resultatet vises i Tabell 11. Tabell 12 viser prosentvis fordeling tilknyttet årsak til feilmønsteret, som beskrevet i metodekapittelet. M2+M3 er tilknyttet svake telle- og regneferdigheter, M4, M5 og M6 er tilknyttet svak eller manglende forståelse for likhetstegnet, M8 er kontekstavhengig og blir diskutert for hver oppgaveside, M7 er kategorien for andre svar.

Totalt antall svar på siden er 2640 (6 oppgaver x 440 elever), fordelt på 330 svar fra gruppen Under bekymringsgrensen (UB) og 2310 svar fra gruppen Over bekymringsgrensen (OB).

**Tabell 11 Svar på Side A fordelt i kategoriene**

| Gruppe                        | M1 - Riktig svar | Sum feilmønster | M2 - Ubesvart | M3 - Tellefeil | M4 - Gjengir a | M5 - Gjengir b | M6 - Adderer gitte tall | M7 - Annet | M8 - Gir siffer fra a |
|-------------------------------|------------------|-----------------|---------------|----------------|----------------|----------------|-------------------------|------------|-----------------------|
| <b>UB</b><br>Prosent (antall) | 54,5% (180)      | 45,6% (150)     | 16,7% (55)    | 1,2% (4)       | 1,8% (6)       | 3,3% (11)      | 13,3% (44)              | 5,5% (18)  | 3,6% (12)             |
| <b>OB</b><br>Prosent (antall) | 95,7% (2210)     | 4,3% (100)      | 4,4% (17)     | < 1% (1)       | 1,3% (5)       | 3,1% (12)      | 7,5% (29)               | 4,2% (16)  | 0,8% (20)             |
| <b>Totalt*</b>                | 90,5% (2390)     | 9,5% (250)      | 28,8% (72)    | 2% (5)         | 4,4% (11)      | 9,2% (23)      | 29,2% (73)              | 13,6% (34) | 12,8% (32)            |

\*Riktig svar og sum feilmønster er gitt prosentseter utfra antall total svar. Feilmønstrene er gitt prosentseter utfra antall feilsvar.

**Tabell 12 Prosentvis fordeling knyttet til årsak for feilmønsteret**

|    | Riktig svar | Svake regnetelleferdigheter | Forståelse for likhetstegnet | M8    | Andre | Totalt |
|----|-------------|-----------------------------|------------------------------|-------|-------|--------|
| UB | 54,5 %      | 17,9 %                      | 18,5 %                       | 3,6 % | 5,5 % | 100%   |
| OB | 95,7 %      | 0,8 %                       | 2 %                          | 0,9 % | 0,6 % | 100%   |

Utfallet av de innledende analysene viser at den største konsentrasjon av elevsvarene faller inn under kategoriene *ubesvart* (M2) og *adderer gitte tall* (M6).

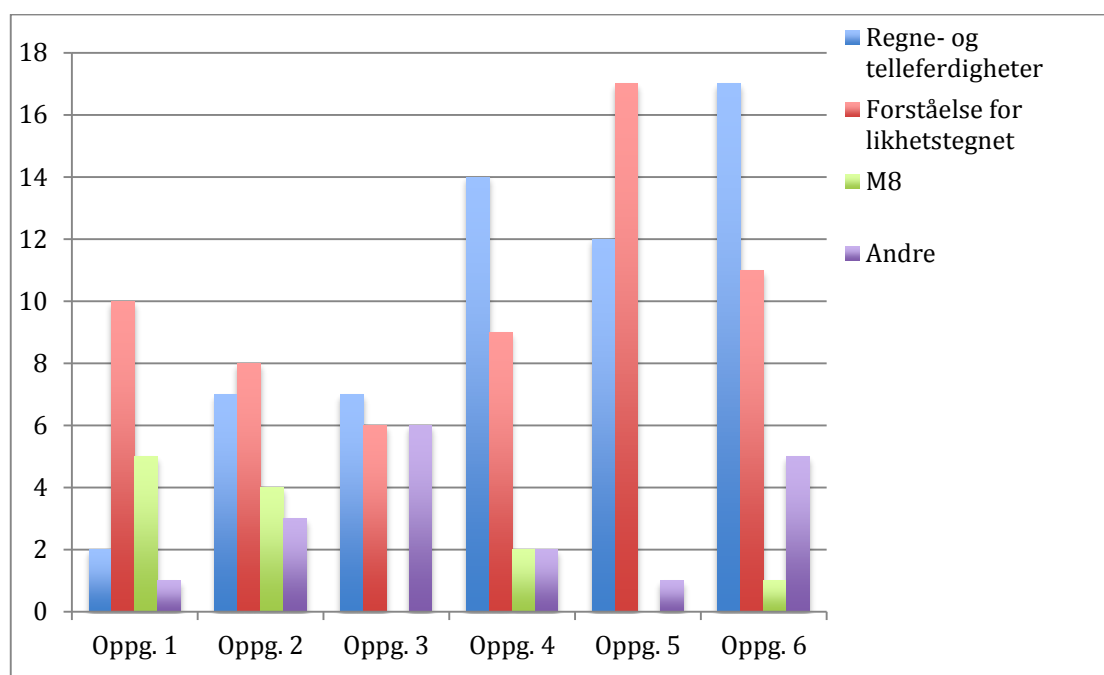
Videre presenteres feilmønstrene slik de fremkom på de ulike oppgavene. Effekten for hver av oppgavene kan indikere løsningsretning, samt opplevd vanskegrad for hele utvalget. En stor effektstørrelse tilsier en god diskriminering mellom sterke og svake regnere. Er effekten liten og scoreopptakelsen totalt lav, kan det indikere at oppgaven i seg selv var utfordrende for hele utvalget, og ikke nødvendigvis bare for svake regnere. I tabell 13 er effekten for oppgavene på Side A oppsummert. På neste side presenteres Figur 15 og Figur 16, som viser hvordan de ulike årsakene til feilmønstrene gjorde seg gjeldende gjennom oppgavesiden for de to grupperingene.

**Tabell 13 Effektstørrelser\* for oppgavene Side A**

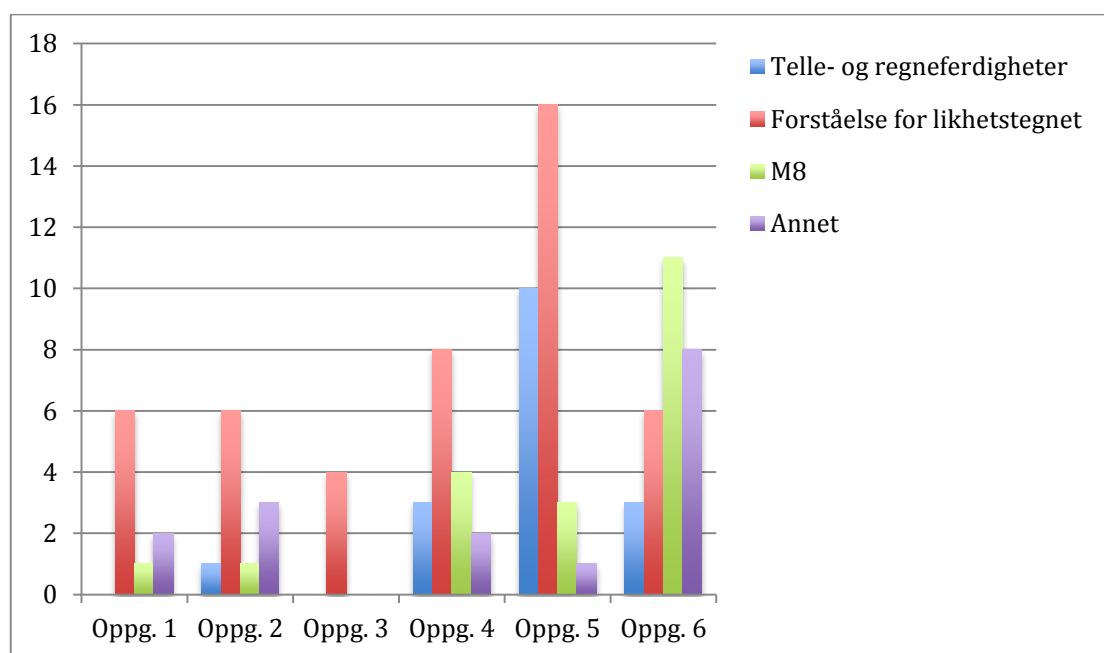
|             | Oppg.1 | Oppg.2 | Oppg.3 | Oppg.4 | Oppg.5 | Oppg.6 |
|-------------|--------|--------|--------|--------|--------|--------|
| r- verdi    | .42    | .48    | .5     | 4.9    | 4.4    | .52    |
| Eta squared | 4.8%   | 6.5%   | 5.7%   | 8.8%   | 9.4%   | 13%    |

\*r-verdien er kalkulert fra Mann Whitney U test og eta squared fra Welch's test  $p < .001$  for alle seks oppgavene.

**Tabell 14 Feilmønster tilknyttet årsak for gruppen Under bekymringsgrensen for oppgave 1-6 Side A.**



**Tabell 15 feilmønster tilknyttet årsak i gruppen Over bekymringsgrensen for oppgave 1-6 Side A.**



### 4.3 Analyse og drøfting

Det er hovedsakelig to løsningsretning for siden, enten kolonne for kolonne eller i leseretning venstre til høyre. Det er mulig å si noe om dette utfra forklart varians/*r*-verdi. *Kartleggingsprøve i regning for 2. trinn* er laget slik at uhensiktsmessige strategier gjør at eleven ikke rekker å fullføre alle oppgavene på siden (Utdanningsdirektoratet, 2014b). Det vil si at en indikasjon på løsningsretning gis ved å følge prosent forklart varians/*r*-verdiens stigning. Resultatet indikerer at elevene fra gruppen Under bekymringsgrensen jobber med oppgavene kolonnevis, se Tabell 13. Det gir utslag både i effektstørrelsen fra oppgave til oppgave, men også utfra antall elever som ikke besvarer oppgavene. Effektstørrelsen varierer dog ikke mye gjennom siden (fra  $r = .042$  til  $r = 0,52$ ), noe som indikerer at tidsbruken kanskje ikke var en så stor faktor. Den endringen som er i effekt kan like plausibelt tilskrives oppgavens økning av tallverdier og at en slik endring påvirker vanskelighetsgraden i oppgaven (problem- size effekten). Det er med andre ord uvisst om det er oppgavens vanskegrad eller begrenset tidsbruk som står for den svake økning av *r*-verdier gjennom siden. Begge årsakene har likevel rot i svake regne- og telleferdigheter.

Kategorien *ubesvart* (M2) er en årsak for lav scoreoppnåelse, men det kan diskuteres om den er representativ som feilmønster. Dette med bakgrunn i at det kan være flere årsaker til manglende svar på oppgaven, og alle årsakene er ikke nødvendigvis forbundet med regneferdigheter. Kategorien *ubesvart* (M2) er mer fremtredende for elevene i gruppen Under bekymringsgrensen enn for elevene i gruppen Over bekymringsgrensen (16,7% av svarene i gruppen Under bekymringsgrensen og 4,4 % av svarene i gruppen Over bekymringsgrensen, se Tabell 11). Resultatet viser også at oppgaver av kategori RR6 gir flere feilsvar tilknyttet forståelse for likhetstegnet enn oppgaver av kategori RR5. Dette kan skyldes selve oppgavekategorien, men også at oppgaver av kategori RR5 på denne siden har enersvar noe som gjør backup-strategier mer plausible, fordi det er mindre å telle. Dette siste blir motbevist når en ser på kategorien for tellefeil (Tabell 11, s.51), som er minimal på denne siden. Konklusjonen blir derfor at det er oppgavekategorien som er betingende for feilmønsteret.



Oppgave nummer 5 skiller seg ut, når en sammenligner med gruppen Over bekymringsgrensen, ved at antall svar av kategoriene *ubesvart* (M2) og *adderer gitte tall* (M6) stiger for akkurat denne oppgaven. En plausibel årsak her kan være at denne oppgaven krever forståelse for tallsymbolet 0. Denne spesifikke oppgaven representerer en form for identitetslikning. I den forstand at elevene skal skrive inn enere fra en hel tier, som er 0. Et eksempel på en slik regneoppgave kan være  $10 = 10 + \_? \_$ . Muligens lar flere elever være å svare på denne oppgaven på bakgrunn av at det ikke er noen enere som skal skrives og fortsetter på neste oppgave. Kategorien *adderer gitte tall* (M6) har også en økning her, noe som indikerer at oppgaven muligens krever mer av elevene i selve avkodingsprosessen.

Ved å undersøke Tabell 11 (s. 51), blir det tydelig at i kategorien *Regne- og telleferdigheter* er det *ubesvart* som er dominerende, etterfulgt av *gir siffer fra a* (M8) og *tellefeil* (M3). At kategorien *tellefeil* (M3) er så lite fremtredende som den er på Side A er interessant, og indikerer at enten har elevene svart i tråd med en av de andre kategoriene eller latt være å svare på oppgaven. Det indikerer også at det er brukt lite backup- strategier på siden. Studien til Lindvall og Ibarra (1980) fant at elever oftere lot være å svare på regneoppgaver av kategori RR5 og RR6, og med bakgrunn i den studien er det vanskelig å si om elevene lar være å svare på grunn av tidsfaktoren eller avkodning av oppgaven. Mønsteret som har sammenheng med forståelse for likhetstegnet er tydeligere i den forstand at det er knyttet til en konkret årsak. Det er da interessant å se at mønsteret med forståelse for likhetstegnet er så fremtredende som det er på Side A. Tabell 12 indikerer at denne kategorien har minimum lik om ikke større betydning for score på siden enn kategorien for svake *telle- og regneferdigheter*. Der svake telle- og regneferdigheter står bak 17,9 % av svarene og forståelse for likhetstegnet 18,5%.

Kategorien *gir siffer fra a* (M8) med 3,6% for gruppen Under bekymringsgrensen og 0,8% for gruppen Over bekymringsgrensen er vanskelig å avgjøre om skyldes forståelse for likhetstegnet eller kunnskap om posisjonssystemet, derfor er det valgt å la denne stå alene. Mulig årsak til dette mønsteret på Side A kan ligge i instruksjonen til elevene, som har fokus på ”fyll inn” og ikke eksplisitt fokus på å gjøre ekvivalensen riktig. Kategorien *Andre* (M7) er relativt liten, noe som antyder at de andre kategoriene identifiserer mønstrene i svarene godt.

### 4.3.1 Oppsummering Side A

Utfra analysene som er gjort kan det se ut som elevene som svarer feil her hovedsakelig har utfordringer med forståelse for likhetstegnet. Dette med bakgrunn i at feilmønstrene som er fremtredende alle er forbundet med operasjonell forståelse for likhetstegnet. Feilsvar som er forbundet med forståelse for likhetstegnet står for 18,5% av svarene i gruppen Under bekymringsgrensen og 2% av svarene i gruppen Over bekymringsgrensen. Feilsvar som er forbundet med svake telle- og regneferdigheter står for 17,9% av svarene for gruppen Under bekymringsgrensen og 0,8% av svarene for gruppen Over bekymringsgrensen. Det vil si at denne siden først og fremst kan si noe om elevens forståelse for likhetstegnet, sekundært kommer telle- og regneferdigheter. I hovedsak er det altså avkoding av oppgaven som er utfordringen her, det samme mønsteret gjør seg gjeldende for gruppen Over bekymringsgrensen. En kan da stille spørsmål ved om det er hensiktsmessig å bruke oppgaver av kategori RR5 og RR6, når *Kartleggingsprøve i regning for 2. trinn* skal bestå av mange enkle oppgaver (Utdanningsdirektoratet, 2014b). Oppgaver av disse kategoriene har vist seg å være de vanskeligste innen ett-trinnsoppgaver (Lindvall & Ibarra, 1980; Powell, 2014; Weaver, 1973). En annen utfordring er at oppgavene på denne siden består av flere lag. For det første har en tidsbegrensningen som skal identifiserer elever med uhensiktsmessige strategier. For det andre øker oppgavens vanskelighetsgrad også utover siden. En elev som ikke svarer på de siste oppgavene kan da ha minimum to ulike årsaker til dette, enten svake strategier eller ikke tilstrekkelig kunnskap om algoritmen. For eksempel når en går fra ensifret til tosfifret addisjon. Det gjør at oppgavesiden her ikke er finmasket nok til å avgjøre årsak for feilmønstre. Resultatet kan kun indikere noe om at eleven ikke kan nok, men ikke spesifikt hva eleven har utfordringer med.

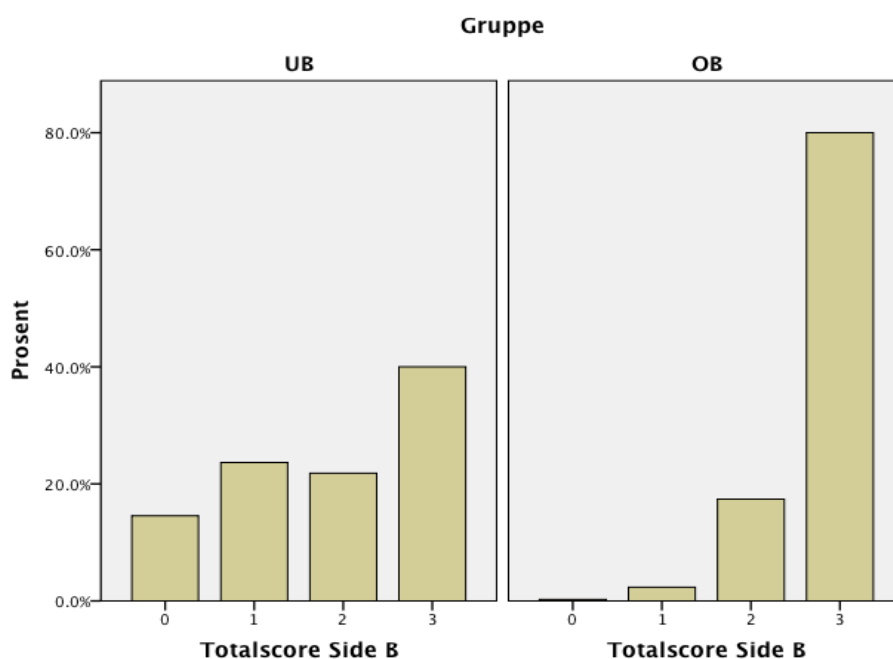
For Side A er feilene som går igjen forbundet hovedsakelig med forståelse for likhetstegnet. Oppgaver av kategori RR6 gir flere slike feilsvar enn oppgaver av kategori RR5. Antall ubesvarte oppgaver øker mot slutten av siden, men det er ikke mulig å avgjøre om dette skyldes tidsfaktoren eller avkoding av oppgavene. Kompetansen elevene tilsynelatende mangler her er hovedsakelig knyttet til likhetstegnet og avkoding av ikke-kanoniserte oppgaver.



bekymringsgrensen (Md = 3, n = 385),  $U = 15\,495.5$ ,  $z = 7.355$ ,  $p < .001$ , effekten var moderat  $r = 0.35$ .

En Welch's test ble også gjennomført for å sammenligne sidescore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Den indikerte statistisk signifikant forskjell i score for gruppen Under bekymringsgrensen ( $M = 1.87$ ,  $SD = 1.106$ ) og gruppen Over bekymringsgrensen ( $M = 2.77$ ,  $SD = .489$ ;  $t(57) = -5.941$ ,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference =  $-.899$ , 95% CI =  $-1.202$  til  $-.596$ ) var moderat (eta squared =  $0.075 = 7,5\%$ ).

Effekten totalt for siden er moderat, og noe lavere enn for Side A. Fordelingen av score, Figur 9, følger samme mønster som for Side A, med tydelig takeffekt for gruppen Over bekymringsgrensen og et noe mer utydelig bilde for gruppen Under bekymringsgrensen. Det skal bemerkes at en større prosentandel fra gruppen Under bekymringsgrensen scorer fullt på denne siden sammenlignet med Side A.



Figur 9 poengscore på Side B prosentvis i gruppene

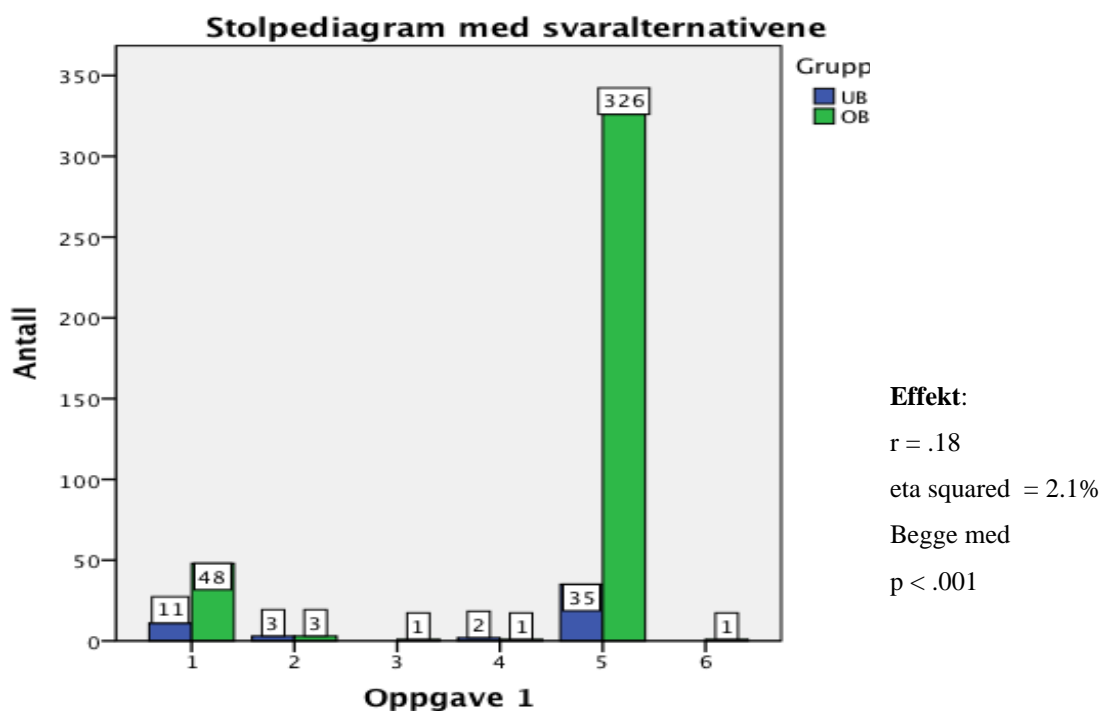
Tabell 16 med de eksakte prosentverdiene tydeliggjør forskjellen mellom de to gruppene. Denne oppgaven har svaralternativer, og det blir derfor sett nærmere på de svaralternativene som utpeker seg utenom det korrekte for hver oppgave. Svaralternativene er gitt nummer etter hvordan de er presentert på oppgavesiden, i stigende orden. Det blir presentert

Tabell 16 Sidescore Side B prosent og antall i gruppene

| Totalscore Side B |             | Gruppe |        | Totalt |
|-------------------|-------------|--------|--------|--------|
|                   |             | UB     | OB     |        |
| 0                 | Antall      | 8      | 1      | 9      |
|                   | % av Gruppe | 14.5%  | 0.3%   | 2.0%   |
| 1                 | Antall      | 13     | 9      | 22     |
|                   | % av Gruppe | 23.6%  | 2.3%   | 5.0%   |
| 2                 | Antall      | 12     | 67     | 79     |
|                   | % av Gruppe | 21.8%  | 17.4%  | 18.0%  |
| 3                 | Antall      | 22     | 308    | 330    |
|                   | % av Gruppe | 40.0%  | 80.0%  | 75.0%  |
| Totalt            | Antall      | 55     | 385    | 440    |
|                   | % av Gruppe | 100.0% | 100.0% | 100.0% |

et stolpediagram som viser fordeling mellom de ulike svaralternativene (1-7) for hver oppgave.

#### 4.4.1 Oppgave 1



Figur 10 Stolpediagram for svar på oppg.1 Side B

Det var 4 elever fra gruppen Under bekymringsgrensen og 5 elever fra gruppen Over

bekymringsgrensen som ikke hadde avgitt svar på denne oppgaven. Ingen av elevene hadde svart alternativ 7.

Det riktige svaret her er svaralternativ 5. I Figur 11 ser en at det er 326 (84.7%) fra gruppen Over bekymringsgrensen og 35 (63.6%) fra gruppen Under bekymringsgrensen som gir dette svaret. Det er to feilmønstrene for denne oppgaven: svaralternativ 1 og 2, begge tilhører kategorien *Till*, disse blir kommentert.

#### 4.4.1.1 Svaralternativ 1

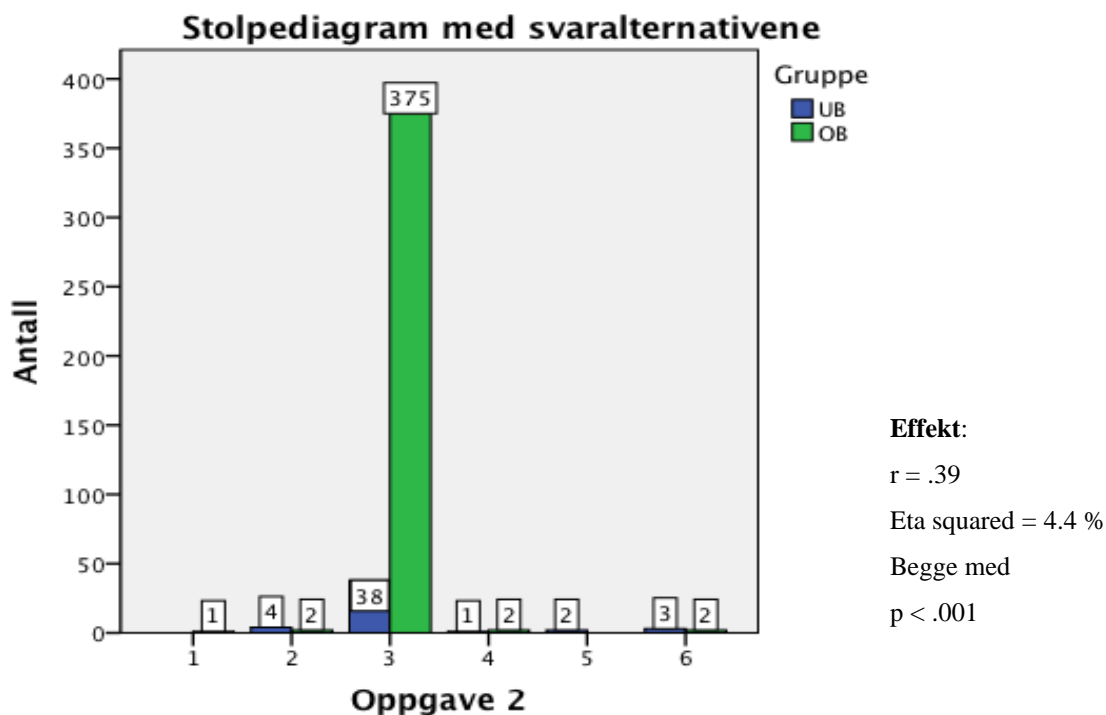
Dette svaralternativet samsvarer med å telle halvkonkretene uten å ta hensyn til den enkelte halvkonkretens verdi. Her har elevene mest sannsynlig talt halvkonkretene som likeverdige enheter. Et slikt feilmønster kan sees i sammenheng med elevenes evne til å oppfatte og telle abstrakte mengder (jf. kap. 2.4). I den forstand at en halvkonkret kan ha verdi 7, 12 eller 20, selv om den bare er én halvkonkret. Dette kan la seg eksemplifisere med piler på en blink, der pilene er plassert på forskjellig score. Dette feilmønsteret samsvarer da med å telle antall piler i stedet for å summere poengene pilene representerer.

#### 4.4.1.2 Svaralternativ 2

Dette svaralternativet er en variant av feilmønsteret som kommer frem i svaralternativ 1. Her har elevene mest sannsynlig tatt hensyn til enkeltverdien til hvert halvkonkret, men feiltolket eller misforstått et av halvkonkretens verdi. Om dette skyldes lite kjennskap til denne typen halvkonkreter, eller om det skyldes svak kunnskap om hvordan skrive mengder med tall er ikke mulig å si. Fra eksempelet ovenfor tilsvarer dette å summere scorene til pilene, men å feiltolke en av dem for eksempel regne med 10 for 12 scoren.

Oppsummert for denne oppgaven viser resultatet at feilmønstrene fra *Till* kategorien står for 25,5% av svarene fra gruppen Under bekymringsgrensen og 13,2 % av svarene fra gruppen Over bekymringsgrensen. Primært er feilmønsteret knyttet til elevenes evne til å oppfatte abstrakte mengder, som er tilknyttet tallbegrep og telleferdigheter.

#### 4.4.2 Oppgave 2



Figur 11 Stolpediagram for svar oppg.2 Side B

Det var 7 elever fra gruppen Under bekymringsgrensen og 3 fra gruppen Over bekymringsgrensen, som ikke hadde svart på denne oppgaven. Ingen elever hadde svart alternativ 7.

Det riktige svaret her er svaralternativ 3. 97,4 % av gruppen Over bekymringsgrensen og 69,1% av gruppen Under bekymringsgrensen svarte svaralternativ nr. 3. Det utpeker seg ikke ett spesifikt feilmønster i denne oppgaven slik det gjorde for oppgave 1, svarene er mer fordelt på de ulike alternativene. Svaralternativ 2 og 6 tilfredsstiller kravet om minimum 4 svar, begge tilhører kategorien *Till*. Mønstrene kommenteres nedenfor.

##### 4.4.2.1 Svaralternativ 2

Mulig tolkning her er at elevene kun har sett på enerplassen, dette alternativet er det eneste med denne eneren, og eneren er også representert i halvkonkretene som skal summeres opp. En annen plausibel tolkning er at elevene har oversett en av halvkonkretene, eller glemt å telle den. Mønsteret kan knyttes til evnen til å oppfatte

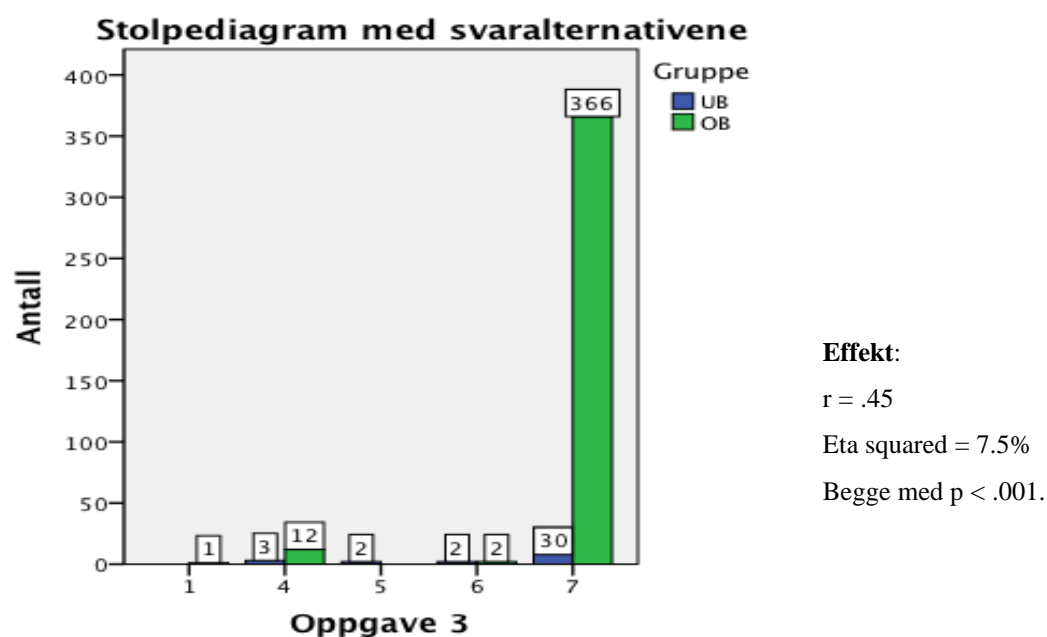
abstrakte mengder, samt svake telle- og regneferdigheter.

#### 4.4.2.2 Svaralternativ 6

Her er det stor sannsynlighet for at elevene ikke har tatt hensyn til den ene halvkonkretens verdi. Dette mønsteret er det samme som for oppgave 1 (eksempelet med pilkasting), og er knyttet til evnen til å oppfatte abstrakte mengder.

Oppsummert for oppgaven er det relativt likt fordelt mellom feilmønstrene. For gruppen Under bekymringsgrensen ser tiden ut til å ha vært den største faktoren kombinert med svake telle- og regneferdigheter. Feilmønstre tilknyttet kategorien *Till* står for 10,9% av svarene fra gruppen Under bekymringsgrensen og 1% av svarene fra gruppen Over bekymringsgrensen.

#### 4.4.3 Oppgave 3



Figur 12 Stolpediagram med svar for oppg. 3 Side B

Det var henholdsvis 18 elever fra gruppen Under bekymringsgrensen og 4 elever fra gruppen Over bekymringsgrensen, som ikke hadde svart på denne oppgaven. Ingen av elevene hadde svart alternativ 2 eller 3.



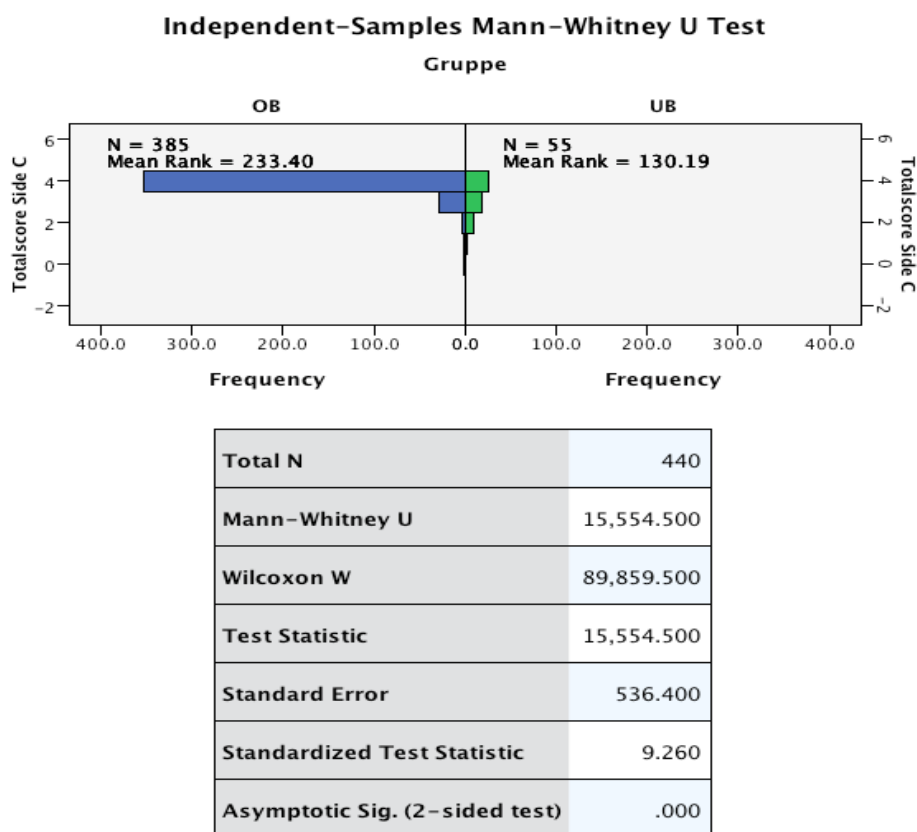
Oppgaven har to feilmønstre, som har to mulige tolkninger. Elevene kan ha fokusert på eneren, da dette svaralternativet er det eneste som har denne verdien på enerplassen. Den andre tolkningen er at elevene ikke tar hensyn til et av halvkonkretens verdi, og er knyttet til evne til å oppfatte abstrakte mengder. Begge mønstrene hører til kategorien *Till* og kan knyttes til svake telle- og regneferdigheter. Oppsummert ser det ut til at tidsfaktoren absolutt har vært den mest utslagsgivende for gruppen Under bekymringsgrensen, mens feilmønsteret gjør seg mest gjeldende for gruppen Over bekymringsgrensen.

#### **4.4.4 Oppsummering for Side B**

Tidsfaktoren ser ut til å være mest avgjørende for elevens score på Side B. Dette kan begrunnes med økning i kategorien *ubesvart* gjennom siden, spesielt for gruppen Under bekymringsgrensen. Dette kan eventuelt stamme fra utfordringer med avkodingen av oppgaven, og kan diskuteres om det da kan ansees som et feilmønster. Feilmønsteret som er mest fremtredende på Side B har tilknytning til evnen til å oppfatte abstrakte mengder, og kan sees i sammenheng med tallbegrep og telling.

## 4.5 Side C – Oppstilte regneoppgaver

Denne siden består av 4 regneoppgaver av kategori RR1. Resultatet fra Mann-Whitney U – og Welch´s test for siden samlet presenteres først, før feilmønstre og effekt av hver enkelt oppgave kommenteres.

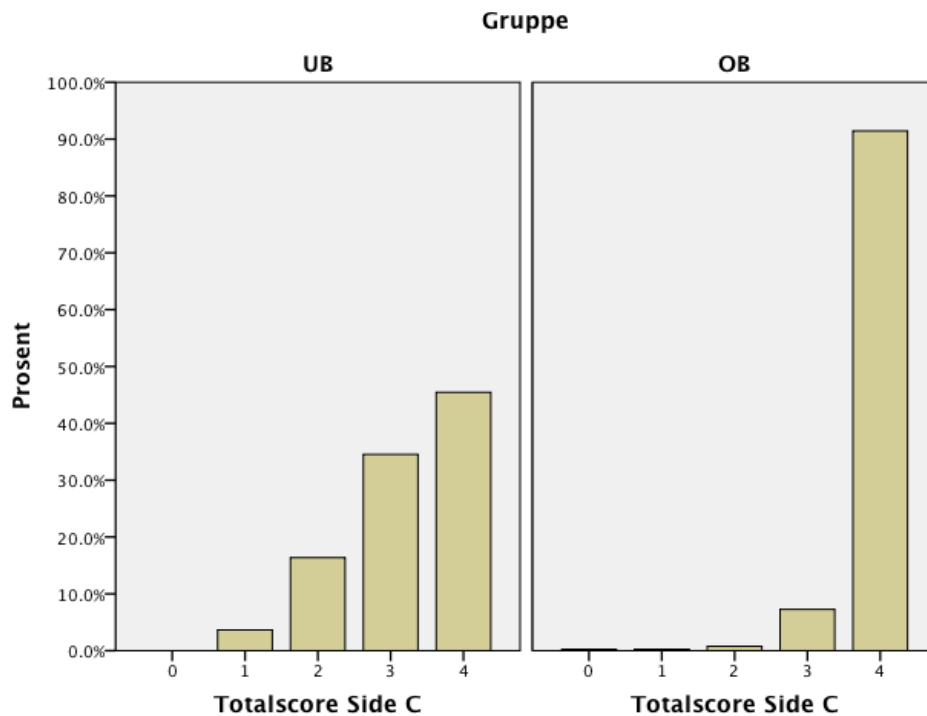


Figur 13 Mann-Whitney U-test for Side C

Mann-Whitney U test indikerte statistisk signifikant forskjell i score på Side C mellom gruppen Under bekymringsgrensen ( $M_d = 3$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $M_d = 4$ ,  $n = 385$ ),  $U = 15\,554.5$ ,  $z = 9.26$ ,  $p < .001$ , effekten var moderat mot stor  $r = .44$ .

En Welch´s test ble også gjennomført for å sammenligne sidescore mellom gruppene. Resultatet indikerte statistisk signifikant forskjell i score mellom gruppen Under bekymringsgrensen ( $M = 3.22$ ,  $SD = .854$ ) og gruppen Over bekymringsgrensen ( $M = 3.89$ ,  $SD = .397$ ;  $t(58) = -5.776$ ,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference =  $-0.675$ , 95% CI =  $-0.909$  til  $-0.441$ ) var moderat (eta squared =  $0.071 = 7,1\%$ ).

Effekten av Side C er moderat, litt større enn for Side B, men mindre enn effekten for Side A. Figur 14 viser et stolpediagram med oversikt over poengscoren for de to gruppene.



Figur 14 Poengscore på Side C prosentvis i gruppene

Figur 14 viser tydelig takeffekten for gruppen Over bekymringsgrensen. Gruppen Under bekymringsgrensen har også samme tendens, men mye svakere stigningstall. For å bedre vise hvordan dette arter seg legger er det lagt ved en tabell med prosentvis score i de to gruppene.

Tabell 17 viser tydelig at stigningstallet **Tabell 17 Poengscore Side C fordelt i gruppene**

til gruppen Over bekymringsgrensen er høyere enn stigningstallet for gruppen Under bekymringsgrensen. Den viser også at ingen elever fra gruppen Under bekymringsgrensen hadde 0 i score på Side C, men at flesteparten scoret 3 eller 4 poeng på denne siden.

| Totalscore Side C |             | Gruppe |        | Totalt |
|-------------------|-------------|--------|--------|--------|
|                   |             | UB     | OB     |        |
| 0                 | Antall      | 0      | 1      | 1      |
|                   | % av gruppe | 0.0%   | 0.3%   | 0.2%   |
| 1                 | Antall      | 2      | 1      | 3      |
|                   | % av gruppe | 3.6%   | 0.3%   | 0.7%   |
| 2                 | Antall      | 9      | 3      | 12     |
|                   | % av gruppe | 16.4%  | 0.8%   | 2.7%   |
| 3                 | Antall      | 19     | 28     | 47     |
|                   | % av gruppe | 34.5%  | 7.3%   | 10.7%  |
| 4                 | Antall      | 25     | 352    | 377    |
|                   | % av gruppe | 45.5%  | 91.4%  | 85.7%  |
| Totalt            | Antall      | 55     | 385    | 440    |
|                   | % av gruppe | 100.0% | 100.0% | 100.0% |

I Tabell 18 vises effekten til de ulike oppgavene. Feilsvarene til elevene på denne oppgavesiden var mer ensartet enn for Side A. Det er derfor valgt å presentere dem samlet i Tabell 19.

Alle oppgavene er av RR1 kategori, noe som ser ut til å ha hatt effekt for feilmønstrene, i og med at kategoriene som har tilknytting til forståelse for likhetstegnet faller vekk. Dette har noe effekt på kategorien *Annet* (M7), som derfor ble analyserte videre for å undersøke om det var flere mønster. Resultatene fra Mann-Whitney U – og Welch´s test for hver av oppgavene blir presentert før feilmønstre knyttet til oppgavesiden blir diskutert.

Effekten av de forskjellige oppgavene kan indikere at gruppen Under bekymringsgrensen løser oppgavene kolonne for kolonne, hvis en antar at det

**Tabell 18 Effektstørrelser Side C**

|             | Oppg. 1 | Oppg. 2 | Oppg. 3 | Oppg. 4 |
|-------------|---------|---------|---------|---------|
| r-verdi     | .04*    | .25     | .26     | .45     |
| Eta squared | 1.3%**  | 2.2%    | 1.9%    | 6.7%    |

\*p = .352 ikke statistisk signifikant forskjell

\*\* p= .014

kun er tidsbruk som skiller på de to gruppene. Effektstørrelsen på de ulike oppgavene kan også være klagjørende vedrørende opplevd vanskegrad hos elevene i gruppen Under bekymringsgrensen. For oppgave 1 er gjennomsnittet til elevene i gruppen Under bekymringsgrensen faktisk høyere enn gjennomsnittet til elevene fra gruppen

Over bekymringsgrensen, effektstørrelsen for oppgave 1 er dog ikke basert på en statistisk signifikant forskjell mellom gruppene. På denne siden er det RR1 med tierovergang som har størst effekt, både ensifrede og tosifrede. Dette funnet kan indikere at gruppen Under bekymringsgrensen har utfordringer med tieroverganger i RR1 oppgaver.

Av *Annet* kategorien (M7)

**Tabell 19 Feilmønstre etter oppgavenummer Side C**

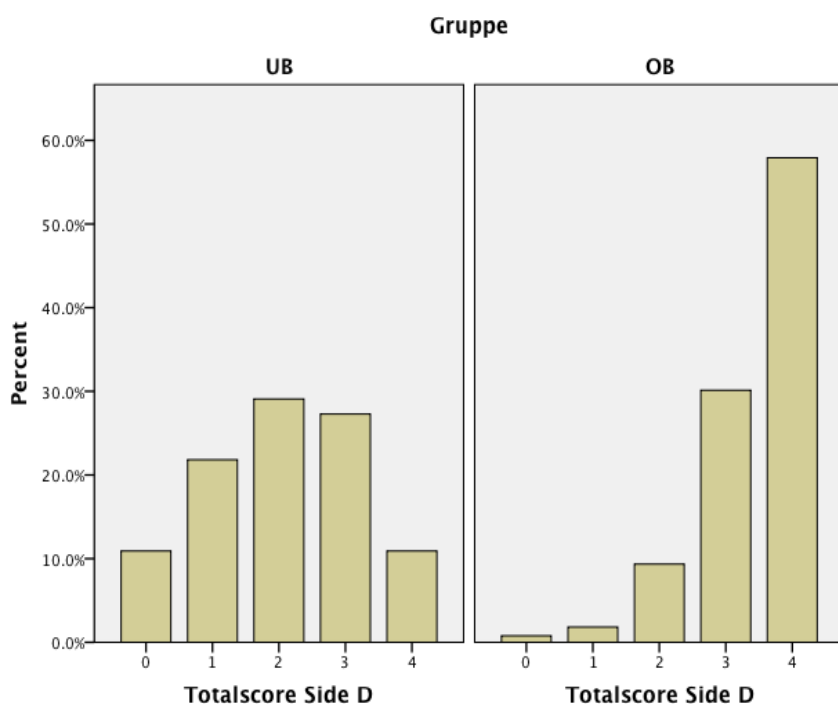
var det ingen tydelige mønstre med unntak på oppgave 4, der et mønster var +/- 10 fra riktig svar. Dette kan skyldes at tallene i denne oppgaven er forholdsvis store (sum mellom 50-60). Av dem

| Oppgave | M2 |    | M3 |    | M7 |    |
|---------|----|----|----|----|----|----|
|         | UB | OB | UB | OB | UB | OB |
| 1.      | -  | -  | -  | -  | -  | 6  |
| 2.      | 3  | 1  | 3  | 7  | 6  | 7  |
| 3.      | 6  | 1  | -  | 1  | 2  | 3  |
| 4.      | 9  | 1  | 2  | 1  | 12 | 13 |
| Sum     | 18 | 3  | 5  | 9  | 20 | 29 |

som oppga -10 fra riktig svar var 7 fra gruppen Under bekymringsgrensen og 6 fra gruppen Over bekymringsgrensen. Av dem som oppga + 10 fra riktig svar var ingen fra gruppen Under bekymringsgrensen og 3 fra gruppen Over bekymringsgrensen. Det er interessant at ingen fra gruppen Under bekymringsgrensen gir svar som er +10. Det kan indikere at backup- strategier er brukt, at det har blitt talt fra første addend til neste, og at på den måten er den vanskelig å svare 10 for mye. Mønsteret med +/- 10 fra riktig svar kan knyttes til svake telle- og regnestrategier. Antall manglende svar på oppgave 4 på denne siden kan igjen skyldes to hovedårsaker, den ene er begrenset tidsbruk som i så fall kan tyde i retning av at elevene har uhensiktsmessige telle- og regnestrategier. Den andre årsaken er ”problem-size” -effekten, ved at tallene er de høyeste i denne oppgaven (sum mellom 50 – 60), noe som gjenspeiler seg i feilmønsteret med å gi svar +/- 10. Begge feilmønstrene peker i retning av svake telleferdigheter, som igjen bidrar til svake regneferdigheter. Det faktum at denne siden består av rene RR1 oppgaver bidra til å luke vekk ulike årsaker til feilmønstrene, og letter arbeidet med å avgjøre hvilken kompetanse det er elevene har utfordringer med.

## 4.6 Side D – Tekstoppgaver

Denne oppgavesiden består av 4 tekstoppgaver gitt som *Sammensetningsoppgaver* med ukjent sum (se Tabell 2, side 9). Illustrasjonene kan anses som både informative og dekorative. Illustrasjonen elevene skal sammenligne med er plassert øverst på siden og de fire oppgavene nederst, etter eksempeloppgaven, det gjør avstanden mellom dem forholdsvis stor. Ved innledende analyser av denne oppgavesiden, ble det tydelig at siden er utfordrende for alle elevene, Figur 15 viser dette.



Figur 15 Stolpediagram for poengscore prosentvis i gruppene

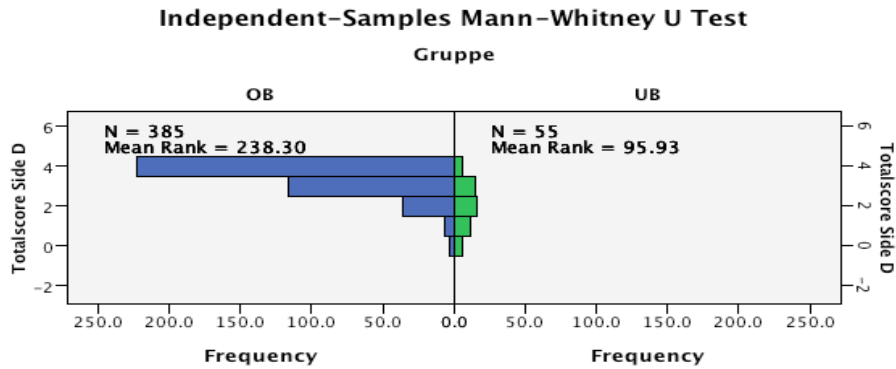
Poengscoren er tilnærmet normalfordelt for elevene i gruppen Under bekymringsgrensen. Elevene i gruppen Over bekymringsgrensen har tendens til takeffekt, men har langt lavere prosentandel på fullscore enn de foregående sidene. Totalt for siden er det en p-verdi på 52% noe som er i ytterkanten av hva som er ønsket for kartleggingsoppgaver.

For å tydeliggjøre hvordan poengscoren fordeler seg, presenteres Tabell 20 med antall og prosent elever for hver poengscore fordelt i de to gruppene. Med dette utgangspunktet er det interessant å undersøke effekten av Side D ved hjelp av Mann-Whitney U- og Welch´s test. Denne siden er spesielt interessant av flere årsaker. For det første er det en side med tekstoppgaver, men den er også den eneste i utvalget med åpent svar. Resultatene fra Mann-Whitney U- og

**Tabell 20 Poengfordeling Side D**

| Score Side D |             | Gruppe |        | Totalt |
|--------------|-------------|--------|--------|--------|
|              |             | UB     | OB     |        |
| 0            | Antall      | 6      | 3      | 9      |
|              | % av gruppe | 10.9%  | 0.8%   | 2.0%   |
| 1            | Antall      | 12     | 7      | 19     |
|              | % av gruppe | 21.8%  | 1.8%   | 4.3%   |
| 2            | Antall      | 16     | 36     | 52     |
|              | % av gruppe | 29.1%  | 9.4%   | 11.8%  |
| 3            | Antall      | 15     | 116    | 131    |
|              | % av gruppe | 27.3%  | 30.1%  | 29.8%  |
| 4            | Antall      | 6      | 223    | 229    |
|              | % av gruppe | 10.9%  | 57.9%  | 52.0%  |
| Totalt       | Antall      | 55     | 385    | 440    |
|              | % av gruppe | 100.0% | 100.0% | 100.0% |

Welch´s test i Figur 16 og kan indikere hvordan åpne svar i tekstoppgaver påvirker elevenes scoreoppgjør.



|                                |            |
|--------------------------------|------------|
| Total N                        | 440        |
| Mann-Whitney U                 | 17,439.000 |
| Wilcoxon W                     | 91,744.000 |
| Test Statistic                 | 17,439.000 |
| Standard Error                 | 804.107    |
| Standardized Test Statistic    | 8.521      |
| Asymptotic Sig. (2-sided test) | .000       |

**Figur 16 Mann-Whitney U-test for Side D**

Mann-Whitney U test indikerte statistisk signifikant forskjell i poengscore på Side D mellom gruppen Under bekymringsgrensen ( $M_d = 2$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $M_d = 4$ ,  $n = 385$ ),  $U = 17\,439$ ,  $z = 8.521$ ,  $p < .001$ , effekten var moderat mot stor,  $r = 0.41$ .

En Welch's test ble også gjennomført for å sammenligne sidescore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Resultatet indikerte statistisk signifikant forskjell i score mellom gruppen Under bekymringsgrensen ( $M = 2.05$ ,  $SD = 1.177$ ) og gruppen Over bekymringsgrensen ( $M = 3.43$ ,  $SD = .797$ ;  $t(61) = -8.370$ ,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference =  $-1.371$ , 95% CI =  $-1.699$  til  $-1.044$ ) var opp mot stor (eta squared =  $0.138 = 13,8\%$ ).

Effekten for Side D skiller seg ikke nevneverdig fra de andre sidene. Dette kan skyldes at begge gruppene presterer jevnt svakere her. Utfra effektscore (Se Tabell 21), kan en indikere noe om løsningsretning. Det ser ut til å følge mønsteret med å



løse oppgavene kolonne for kolonne, men oppgave 2 skiller seg ut med lave 2,8% i effekt. Ved nærmere undersøkelser viser det seg at det er et mulig

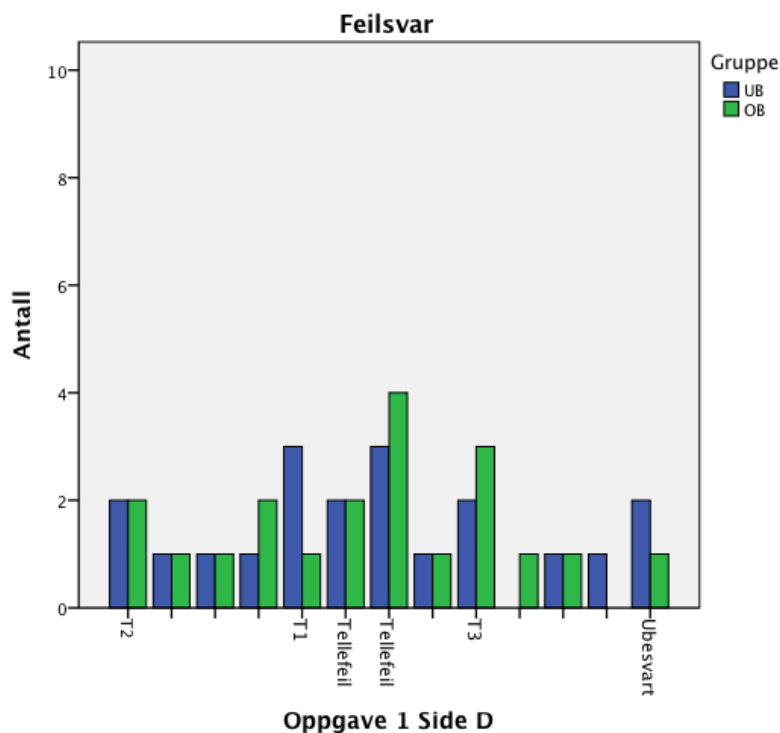
**Tabell 21 Effektstørrelser Side D,  $p < .001$**

|             | Oppg.1 | Oppg.2 | Oppg. 3 | Oppg.4 |
|-------------|--------|--------|---------|--------|
| r-verdi     | .36    | .19    | .29     | .37    |
| Eta squared | 4.8%   | 2.8%   | 5.2%    | 9.3%   |

falskt positivt funn på oppgave 2. Det er to kombinasjoner av ulike tall i illustrasjonene som gir riktig svar, og følgelig reduseres effekten for denne oppgaven. Ved å se på hva elevene faktisk har svart, som er fordelingen med åpne svar, får en mer informasjon om hva eleven kan ha tenkt. På denne siden skal elevene finne frem til en sum via tall i illustrasjonen og skrive summen i oppgaveruten. Ved nærmere undersøkelser av illustrasjonen i sammenheng med elevens svar blir det tydelig at flere elever kan ha mistolket illustrasjonen, det vil si at det kan se ut som at elevene har koblet feil gjenstand til tallet, eller motsatt, feil tall til gjenstanden. Bakgrunnen for denne påstanden er at summene elevene oppgir samsvarer med andre tall som er nært plassert til det tallet eller gjenstanden de skulle brukt. For å kategorisere feilsvarene til elevene her måtte elevenes faktiske svar analyseres. Svar som ble avgitt av 4 eller flere elever blir omtalt som feilmønstre. For å ikke bryte med konfidensialiteten oppgis ikke de eksakte svarene, men omtales som ulike mønstre og mulig logikk som kan ligge bak dem blir forklart.

#### **4.6.1 Oppgave 1**

I denne oppgaven er illustrasjonen av to like objekter, det vil si at elevene skal finne et tall og doble det. 63,6 % (35 elever) fra gruppen Under bekymringsgrensen og 94.8% (365 elever) fra gruppen Over bekymringsgrensen har riktig svar.

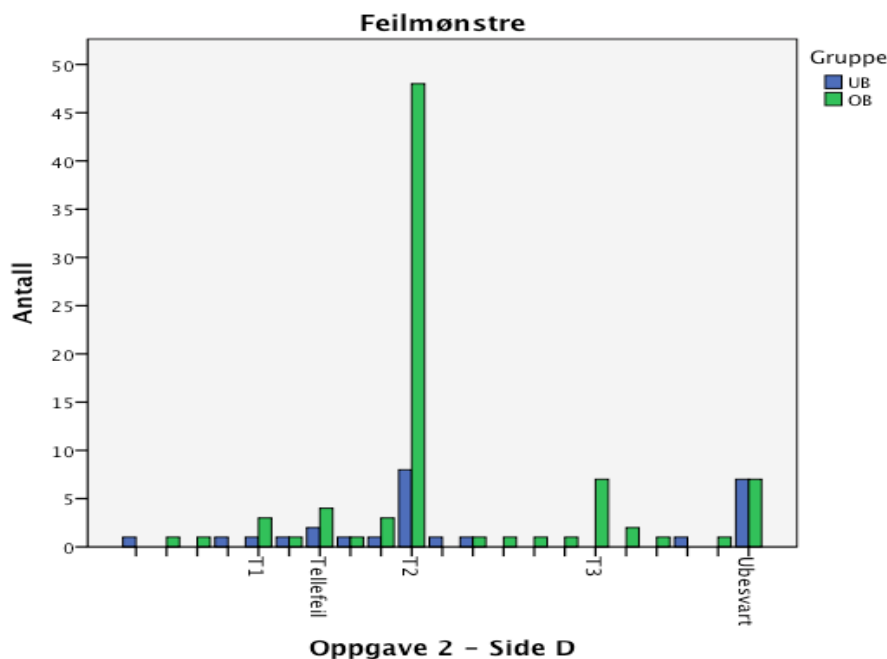


Figur 17 Feilmønstre oppgave 1 Side D

Av feilsvarene på oppgaven kommer det frem et ganske interessant mønster, summen av tvillingtall. Grunnen til at dette er et spennende mønster ligger i illustrasjonen til oppgaven. Illustrasjonen har flere tellbare detaljer, og de to mest fremtredende mønstrene utenom tellefeil stemmer overens med disse detaljene. T1 er summen av tvillingtallene før tallet de skal bruke, med 3 fra gruppen Under bekymringsgrensen og 1 fra gruppen Over bekymringsgrensen. Det er vanskelig å bedømme om dette grunnes illustrasjonen, eller feil i gjenhenting av regnefakta. Det andre feilsvaret (T2) derimot er såpass mye mindre (se Figur 17) enn tallet oppgaven ber elevene om å doble at det synes usannsynlig at det er problemer knyttet til gjenhenting av regnefakta. Her finner vi 2 fra gruppen Under bekymringsgrensen og 2 fra gruppen Over bekymringsgrensen. Det var ett svar (T3) til som flere enn 4 elever ga. Det lar seg ikke forklare ved hjelp av detaljene på illustrasjonene, men det kan se ut til å ha sammenheng med utfordringer knyttet til tierovergang, ved at elevene gjør feil i bruken av algoritmen når regnestykket inneholder tierovergang. Henholdsvis 2 svar fra gruppen Under bekymringsgrensen og 3 svar fra gruppen Over bekymringsgrensen er i denne kategorien. Oppsummert kan en si at den mest avgjørende faktoren for denne oppgaven ser ut til å ha tilknytning til svake telle- og regneferdigheter. Sekundært kommer evnen til å tolke illustrasjonen og instruksjon.

## 4.6.2 Oppgave 2

I denne oppgaven er det to ulike objekter elevene skal finne den sammenlagte summen til. Den innledende analysen avdekket 3 mønstre spesifikke for oppgaven, i tillegg til de faste kategoriene. Henholdsvis 78,2% fra gruppen Over bekymringsgrensen og 52,7% fra gruppen Under bekymringsgrensen fikk poeng på denne oppgaven. Oversikten over svar vises først, så kommenteres de ulike mønstrene.



Figur 18 Feilmønstre oppgave 2 - Side D

### 4.6.2.1 Till kategori, T1 og t2

Elevene har her mest sannsynlig lagt sammen feil tall for å komme til dette svaret. Det er plassert et tall til hvert objekt i illustrasjonen, men det varierer om tallet er plassert under eller over objektet. Elever som leser illustrasjonen i leseretning risikerer å lese av feil tall til objektet. Mønster T1 og T2 kan være resultat av en slik tolkning av illustrasjonen. Tabell 22 viser antall elever som svarer ett av disse to mønstrene i sine respektive grupper. Andelen elever som gir svar tilhørende *Till*

kategorien antyder at illustrasjonen her skaper betydelig utfordring for elevene, og kan i verste fall være misvisende. Et annet

Tabell 22 Oversikt over feilmønstre T1 og T2 på oppgave 2 Side D

|    | TI | T2 | Sum              |
|----|----|----|------------------|
| UB | 1  | 8  | 9 (16.4% av UB)  |
| OB | 3  | 48 | 51 (13.2% av OB) |

tankekors er at addisjon er gjort riktig, men med feil tall. I den forstand skulle elevene fått poeng for addisjon, men ikke poeng for tolkning av illustrasjon. Det kan tenkes at en bedre løsning her er at eleven bes om skrive regnestykket de svarer på. Da kunne læreren hatt en mulighet til å se om det var illustrasjonen eller addisjonen eleven hadde utfordring med.

#### 4.6.2.2 Tier kategori, T3

Dette mønsteret er kun fremtredende hos gruppen Over bekymringsgrensen, og er forbundet med feil gjort i tierovergang. Dette svaret er 10 mer enn det riktige, og det kan derfor tenkes at det er tierovergangen her som er utfordrende. 8 elever fra gruppen Over bekymringsgrensen har svar i denne kategorien.

#### 4.6.2.3 Tellefeil og Ubesvart

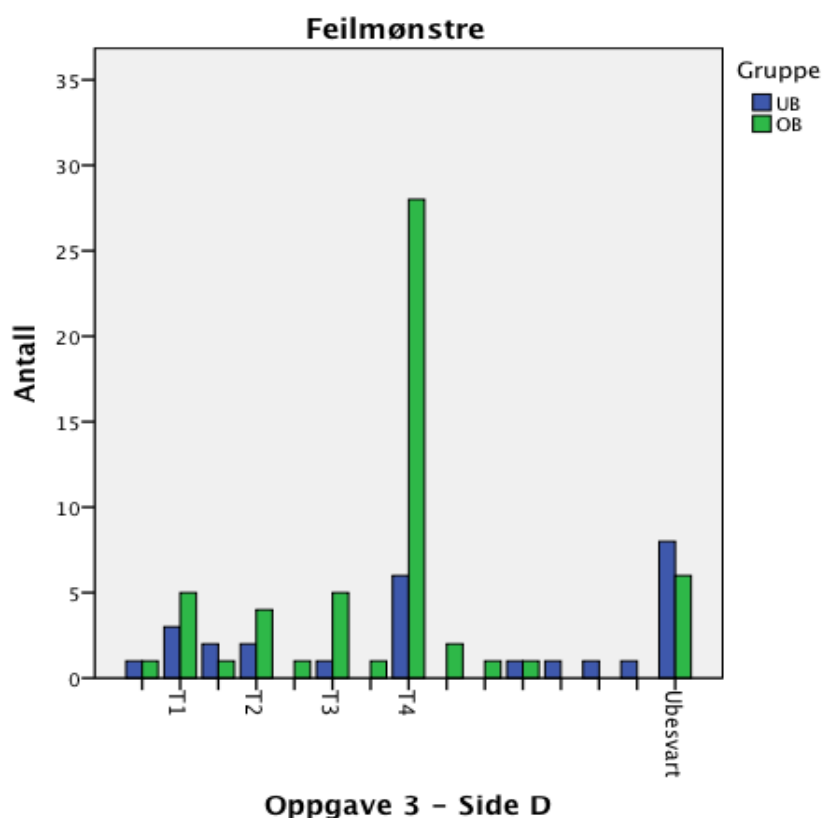
Denne kategorien er større for denne oppgaven enn de foregående, fordi tellefeil til de andre feilmønstrene også må ansees som tellefeil. Til sammen får kategorien tellefeil 7 svar fra gruppen Under bekymringsgrensen og 9 svar fra gruppen Over bekymringsgrensen. Tellefeil forbindes med svake telle- og regneferdigheter, som igjen knyttes til backup- strategier. Backup- strategier er mer tidkrevende, og derfor sees denne kategorien i sammenheng med kategorien *ubesvart*. Til sammen har disse kategoriene 14 elever fra gruppen Under bekymringsgrensen og 16 elever fra gruppen Over bekymringsgrensen.

Oppsummert for denne oppgaven er det tydelig at tolkningen av illustrasjonen er en sterkt påvirkende faktor for score, sekundært kommer tidsfaktoren og mestring av tieroverganger.

### 4.6.3 Oppgave 3

Her skal elevene addere tre tall fra illustrasjonen, to av dem er like.

Av elevene i gruppen Over bekymringsgrensen var det 85,5% som fikk poeng, og for elevene i gruppen Under bekymringsgrensen var prosentandelen 50,9%. I kategorien *ubesvart* var det 8 svar fra gruppen Under bekymringsgrensen og 6 svar fra gruppen Over bekymringsgrensen. Under presenteres 4 mønstre som kom frem gjennom analysen.



Figur 19 Feilmønstre oppgave 3 - Side D

#### 4.6.3.1 Till, T1, T2 og T4

T1 samsvarer med å kun legge sammen tallet til to av objektene. 3 elever fra gruppen Under bekymringsgrensen gir dette svaret, og 5 fra gruppen Over bekymringsgrensen. Det er vanskelig å avgjøre om et slikt feilmønster stammer fra utfordringer med avkodingen av oppgaven eller uoppmerksomhet fra elevenes side. T2 og T4 er mest sannsynlig resultat av å lese illustrasjonen i leseretning i stedet for vertikalt. Da bruker elevene feil tall i sin regning. Totalt var det 8 svar fra gruppen Under bekymringsgrensen og 32 svar fra gruppen Over bekymringsgrensen som var i disse

kategoriene.

#### 4.6.3.2 Tier, T3

Dette mønsteret ser ut til å ha sammenheng med utfordringer i tieroverganger, 1 elev fra gruppen Under Bekymringsgrensen og 5 elever fra gruppen Over bekymringsgrensen svarer dette.

Oppsummert kan en si at tolkning av illustrasjonen (*Till*) her er særs avgjørende for scoren. Av feilene som er gjort står *Till* for 51,9% av dem for gruppen Under Bekymringsgrensen og 76,9% av dem for gruppen Over bekymringsgrensen. Det er også et tankekors at disse svarene for så vidt er riktig med tanke på addisjon, altså at feilen mest sannsynlig ligger i eller kan knyttes til avkoding av illustrasjonen.

#### 4.6.4 Oppgave 4

I denne oppgaven er det to ulike objekter elevene skal finne samlet sum for. Effekten for denne oppgaven er ikke veldig høy med tanke på at det er den siste oppgaven på denne siden. En slik effekt kan indikere at siden har vært utfordrende for begge gruppene, eller at tidsfaktoren ikke har vært den mest avgjørende for scoren på siden. Mistanken styrkes ved at det er like mange fra begge gruppene i kategorien *ubesvart* (19). Andel elever fra gruppen Over bekymringsgrensen som fikk poeng på denne oppgaven var 88,5% og for elevene i gruppen Under bekymringsgrensen var prosentandelen 58,3%. Under presenteres mønstrene som fremkom gjennom analysen.



Oppsummert for denne oppgaven ser utfordringer tilknyttet tieroverganger, og bruk av addisjonsalgoritmen ut til å være en avgjørende faktor, spesielt for gruppen Over bekymringsgrensen. Dette har sammenheng med regnestrategier og således også regneferdighet. Tellefeil har sammenheng med tellestrategier og således telle- og regneferdighet. Disse mønstrene står for hovedvekten av feilene i denne oppgaven, samt *ubesvart* som kan sies å ha samme årsak, tidsfaktor. Feiltolkning av illustrasjonen kommer sekundært i denne oppgaven.

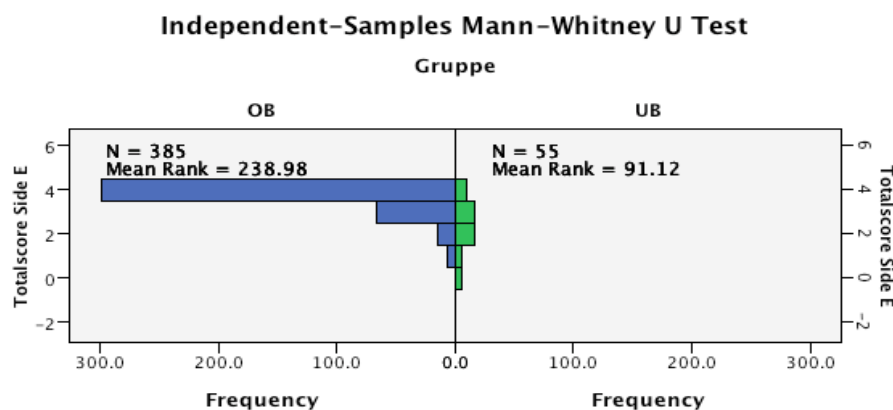
#### **4.6.5 Oppsummering Side D**

Siden er utfordrende for begge gruppene. Oppgave 2 og 3 skiller seg ut ved at det er svært fremtredende mønstre som har sammenheng med feiltolkning av illustrasjonen. Med en så betydelig andel av elevene som gjør denne feilen, er det åpning for å stille spørsmål ved om selve illustrasjonen er hensiktsmessig. For de to andre oppgavene er utfordringer knyttet til tieroverganger en avgjørende faktor for gruppen Over bekymringsgrensen, mens gruppen Under bekymringsgrensen sine feil hovedsakelig er forbundet med svake telle- og regnestrategier.



## 4.7 Side E – Oppstilte regneoppgaver

Denne siden består av 4 regneoppgaver av kategori RR2. Resultatet fra Mann-Whitney U – og Welch´s test for siden samlet presenteres først. Før feilmønstre og effekt i hver av oppgavene blir omtalt.



|                                |            |
|--------------------------------|------------|
| Total N                        | 440        |
| Mann-Whitney U                 | 17,703.500 |
| Wilcoxon W                     | 92,008.500 |
| Test Statistic                 | 17,703.500 |
| Standard Error                 | 711.276    |
| Standardized Test Statistic    | 10.005     |
| Asymptotic Sig. (2-sided test) | .000       |

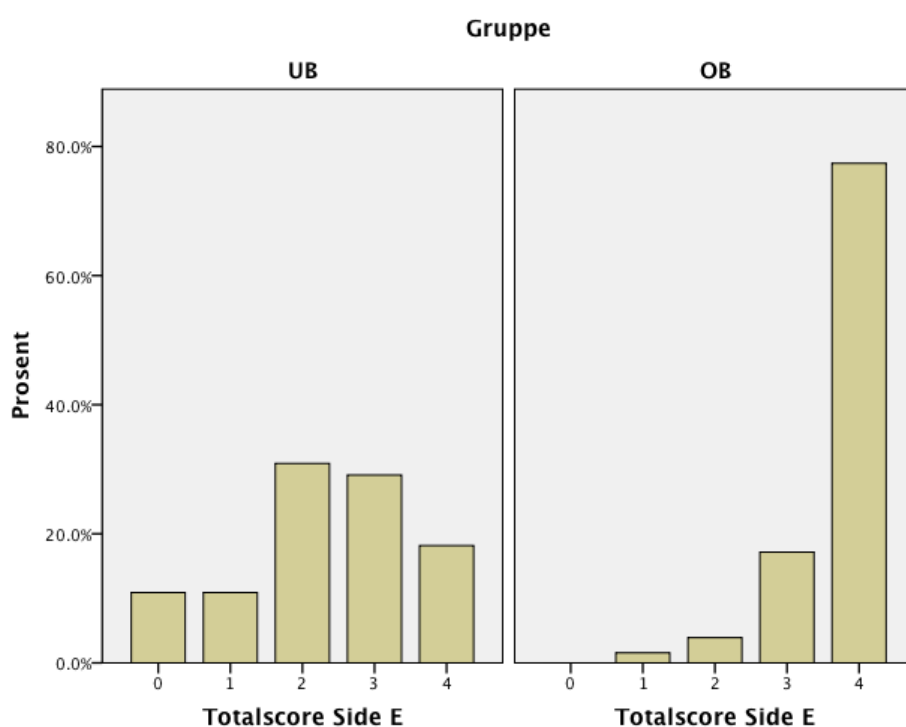
Figur 21 Mann-Whitney U test for Side E

Mann-Whitney U test indikerte statistisk signifikant forskjell i score på Side E mellom gruppen Under bekymringsgrensen ( $Md = 2$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $Md = 4$ ,  $n = 385$ ),  $U = 17\,703.5$ ,  $z = 10.005$ ,  $p < .001$ , effekten er opp mot stor,  $r = 0.48$ .

En Welch´s test ble også gjennomført for å sammenligne sidscore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Resultatet indikerte statistisk signifikant forskjell i score mellom gruppen Under bekymringsgrensen ( $M = 2.33$ ,  $SD = 1.218$ ) og gruppen Over bekymringsgrensen ( $M = 3.70$ ,  $SD = .617$ ;  $t(58)$

= -8.232,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference = -1.377, 95% CI = -1.711 til -1.042) var opp mot stor (eta squared = 0.134 = 13,4% ).

Effekten for siden er relativt stor, med tanke på at den kun består av 4 RR2 oppgaver. Muligens kan det gi indikasjoner på elevene i gruppen Under bekymringsgrensen sin kompetanse angående subtraksjon. Poengfordeling for siden er vist i Figur 22 under.



Figur 22 Poengscore for Side E prosentvis i gruppene

Fra Figur 22 kan det se ut som det er minst én oppgave elevene har utfordringer med, dette med bakgrunn i den store andelen elever fra gruppen Over bekymringsgrensen som scorer 3 poeng. Det er også interessant at en såpass lav prosentandel fra gruppen Under bekymringsgrensen scorer fullt på denne siden. Totalt for siden for hele utvalget er p-verdien 70%. Effekten fra de ulike oppgavene presenteres i Tabell 23. Oppgavene på denne siden er utelukkende av kategori RR2 derfor presenteres feilmønstrene samlet i Tabell 25, etter presentasjon av effekten av hver av oppgavene.

**Tabell 23 effektstørrelse Side E,  $p < .001$**

|             | Oppg. 1 | Oppg. 2 | Oppg. 3 | Oppg. 4 |
|-------------|---------|---------|---------|---------|
| r-verdi     | .26     | .28     | .40     | .46     |
| Eta squared | 2.3%    | 3.7%    | 7.4%    | 11.2%   |

Effekten øker jevnt med oppgavenummeret – det kan antyde at tidsbruk her har vært avgjørende for scoreoppnåelse. Samtidig øker også vanskegraden på oppgavene, fra ensifret subtraksjon, ensifret fra tosifret, tosifret fra tosifret. Den siste oppgaven er ensifret fra tosifret, men har til gjengjeld det høyeste tallet og innebærer å subtrahere fra 0 på enerplass. Tidligere i resultatdelen (se Side A) kom det fram at 0 var en utfordring for begge gruppene, men da i sammenheng med oppgave i kategori RR6. Tabell 24 viser hvordan svarene fordelte seg i de 7 kategoriene, oppgavenummeret står til venstre i tabellen. Nederste rad i tabellen viser prosentandel kategorien utgjorde for gruppen. Det var totalt 1760 svar på siden (4 oppgaver x 440 elever), fordelt på 220 svar fra gruppen Under bekymringsgrensen (55 elever) og 1540 svar fra gruppen Over bekymringsgrensen (385 elever).

**Tabell 24 Svar fordelt i kategoriene for Side E**

|        | M1<br>(Riktig svar) |       | M2<br>(Ubesvart) |      | M3<br>(Tellefeil) |     | M4<br>(Gir a) |     | M5<br>(Gir b) |     | M6<br>(adderer gitte tall) |      | M7<br>(Annet) |      |
|--------|---------------------|-------|------------------|------|-------------------|-----|---------------|-----|---------------|-----|----------------------------|------|---------------|------|
|        | UB                  | OB    | UB               | OB   | UB                | OB  | UB            | OB  | UB            | OB  | UB                         | OB   | Ub            | OB   |
| 1      | 43                  | 372   | 6                | -    | 3                 | 7   |               |     | -             | 1   | 0                          | 1    | 3             | 4    |
| 2      | 35                  | 351   | 6                | 5    | 5                 | 12  |               |     | 1             | 3   | 1                          | 1    | 7             | 13   |
| 3      | 28                  | 354   | 17               | 6    | 4                 | 5   |               |     |               |     | -                          | 3    | 6             | 17   |
| 4      | 22                  | 349   | 20               | 15   | 2                 | 9   | 0             | 1   | 1             | 0   | 4                          | 5    | 6             | 6    |
| Totalt | 58.2%               | 92.6% | 22.3%            | 1.9% | 6.4%              | 15% | -             | <1% | <1%           | <1% | <1%                        | 4.5% | 10.5%         | 2.6% |

Ved første øyekast ser en at antall *ubesvart* øker med oppgavenummer spesielt for gruppen Under bekymringsgrensen, men også gruppen Over bekymringsgrensen har en del *ubesvart* på siste oppgave. Det er vanskelig å avgjøre om dette skyldes tidsfaktoren eller subtraksjon fra 0. Tabellen viser også at *tellefeil* varierer lite for gruppen Under bekymringsgrensen mens det er litt mer varians hos gruppen Over bekymringsgrensen. Det er 10 svar fra gruppen Over bekymringsgrensen som tilsvarende å addere i stedet for å subtrahere (M6), det er 5 svar fra gruppen Under

bekymringsgrensen i den kategorien. Kategorien *Annet* har to oppgaver den skiller seg ut på hovedsakelig for gruppen Over bekymringsgrensen, oppgave nr. 2 og 3. Dette blir kommentert.

#### 4.7.1.1 Oppgave 2 – kategori annet

Et fremtredende mønster her er å svare -10 fra riktig svar, med 6 svar fra gruppen Over bekymringsgrense. Det kan tenkes at elevene her har utfordringer med hvordan algoritmen for subtraksjon skal utføres når en subtraherer ensifret tall fra tosifret tall. Begrunnelsen for tanken om at det må ha tilknytting til bruk av algoritmen er at ved de ulike backup- strategiene vil en vanskelig få et resultat som avviker med 10. Det er derimot lettere å forstille seg en elev som bruker algoritmen, men som overser tierplassen i og med at det er ensifret subtraksjon fra tosifret. Det kan også tenkes at eleven har en misoppfattelse angående hvordan en skal behandle tierplassen i en slik oppgave.

#### 4.7.1.2 Oppgave 3 – kategori annet

Denne oppgaven har svar 1, som tilsier at tallene oppgaven bruker er ved siden av hverandre i telleremsen. Et fremtredende mønster, ved henholdsvis 4 elever fra gruppen Under bekymringsgrensen og 10 elever fra gruppen Over bekymringsgrensen, er neste tall i telleremsen. I og med at et det er snakk om subtraksjon og 2. trinn blir dette telleremsen baklengs, da negative tall ikke er introdusert på 2. trinn. Tidligere i prøveheftet har det vært fokus på tallrekker og rekkefølge, som en del av kartleggingen innen tallforståelse. Det er mulig at dette fokuset har påvirket

assosiasjonsstyrken til disse tallene og at det derfor er rekkefølge som blir fokus for oppgaven og ikke subtraksjon. Dette mønsteret kan sees i sammenheng med kategori *adderer gitte tall* (M6) som i subtraksjons-

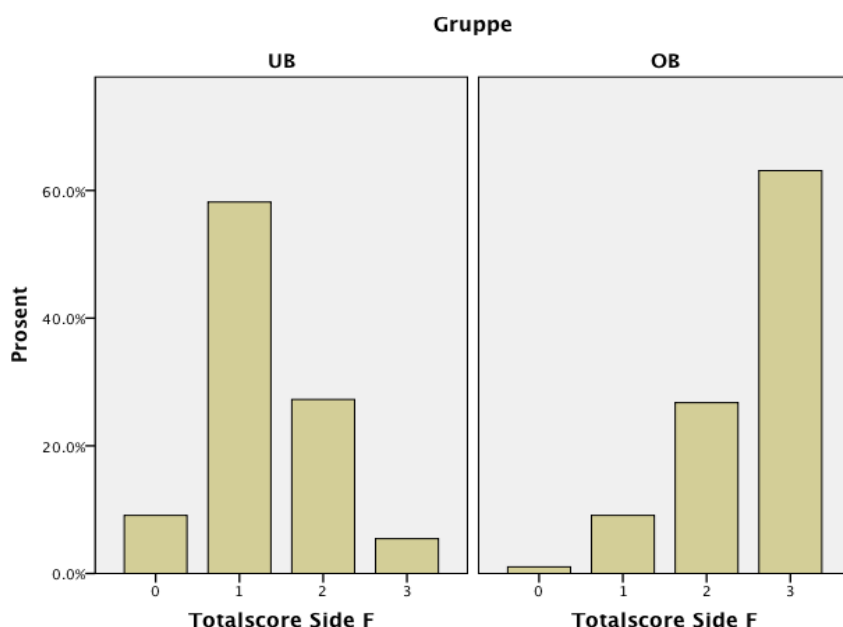
**Tabell 25 Prosentvis fordeling knyttet til årsak**

|        | Tidsfaktor | Avkoding av symbol | Annet  |
|--------|------------|--------------------|--------|
| UB     | 67,3 %     | 11,9 %             | 20,8 % |
| OB     | 51,8 %     | 21,9 %             | 26,3 % |
| Totalt | 58,7 %     | 17,5 %             | 23,8 % |

oppgaver grunner i feiltolkning av matematiske symboler. Oppsummert for Side E er det hovedsakelig to årsaker til feilmønster det er mulig å trekke frem. Tabell 25 viser prosentfordelingen av disse i gruppene og totalt for utvalget.

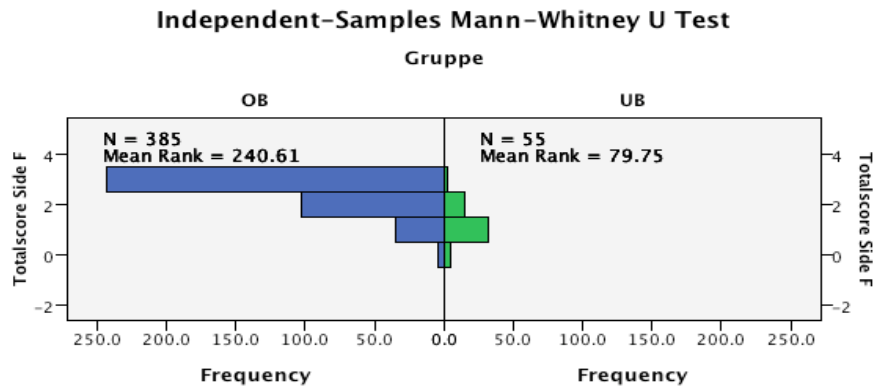
## 4.8 Side F – Tekstoppgaver

Denne oppgavesiden består av 3 tekstoppgaver gitt som *endringsoppgaver* med subtraksjon. Det er brukt ulike halvkonkreter som tallrepresentasjoner, og det er 4 svaralternativer til hver oppgave. Oppgavene er utformet med informative og dekorative illustrasjoner. Sidens poengscore fordelt på gruppene vises i Figur 23.



Figur 23 Poengscore på Side F prosentvis i gruppene

Det er tydelig at denne siden er utfordrende for begge gruppene. Opp mot 60% av elevene i gruppen Under bekymringsgrensen scorer kun 1 poeng på denne siden, mens 60% av elevene i gruppen Over bekymringsgrensen scorer fullt. Med tanke på at kartleggingsoppgaver søker en p-verdi mellom 70% og 90% og helst ikke mindre enn 50% er Side F i ytterkant med en total p-verdi for hele utvalget på 55,9%. Resultatet fra Mann-Whitney U- og Welch's test vises i Figur 24.



|                                |            |
|--------------------------------|------------|
| Total N                        | 440        |
| Mann-Whitney U                 | 18,329.000 |
| Wilcoxon W                     | 92,634.000 |
| Test Statistic                 | 18,329.000 |
| Standard Error                 | 790.205    |
| Standardized Test Statistic    | 9.797      |
| Asymptotic Sig. (2-sided test) | .000       |

**Figur 24 Mann-Whitney U test for Side F**

Mann-Whitney U test indikerte statistisk signifikant forskjell i score på Side F mellom gruppen Under bekymringsgrensen ( $M_d = 1$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $M_d = 3$ ,  $n = 385$ ),  $U = 18\,329$ ,  $z = 9.797$ ,  $p < .001$ , effekten var opp mot stor,  $r = 0.47$ .

En Welch's test ble også gjennomført for å sammenligne sidscore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Resultatet indikerte statistisk signifikant forskjell i score mellom gruppen Under bekymringsgrensen ( $M = 1.29$ ,  $SD = .712$ ) og gruppen Over bekymringsgrensen ( $M = 2.52$ ,  $SD = .704$ ;  $t(70) = -11.994$ ,  $p < .001$  tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference =  $-1.229$ , 95% CI =  $-1.433$  til  $-1.024$ ) var stor (eta squared =  $0.25 = 25\%$ ).

Side F har moderat mot stor effekt for gruppetilhørigheten, det vil si at for elevene i gruppen Under bekymringsgrensen er dette en av sidene hvor deres gjennomsnittsscore er betraktelig lavere enn for elevene i gruppen Over

bekymringsgrensen. Denne siden har svaralternativer, og effekten av de ulike oppgavene presenteres i Tabell 26 før omtale av hvordan eventuelle feilmønstre i oppgavene arter seg.

**Tabell 26 Effektstørrelser Side F,  $p < .001$**

|             | Oppg. 1 | Oppg. 2 | Oppg. 3 |
|-------------|---------|---------|---------|
| r-verdi     | .26     | .43     | .35     |
| Eta squared | 2.2%    | 14.2%   | 14.3%   |

Fra testene gjort ovenfor er det bemerkelsesverdig stor forskjell i effekt mellom oppgave 1 og oppgave 2 og 3. Ved nærmere undersøkelser kan dette tenkes å skyldes at oppgave 1 er subtraksjon med 'tiervenn' og differanse under 10, mens oppgave 2 og 3 er subtraksjon med flersifret tall og tieroverganger. Tidligere i *Kartleggingsprøve i regning for 2. trinn* kom det frem at spesielt tieroverganger var utfordrende. Oppgave 2 innebærer også å subtrahere fra 0 på enerplass. For de to siste oppgavene er mønsteret med svar som avviker med 10 også representert i svaralternativene. Fordeling av svarene mellom riktig, ubesvart og avviker med 10 (*Tier*) er gitt i Tabell 27.

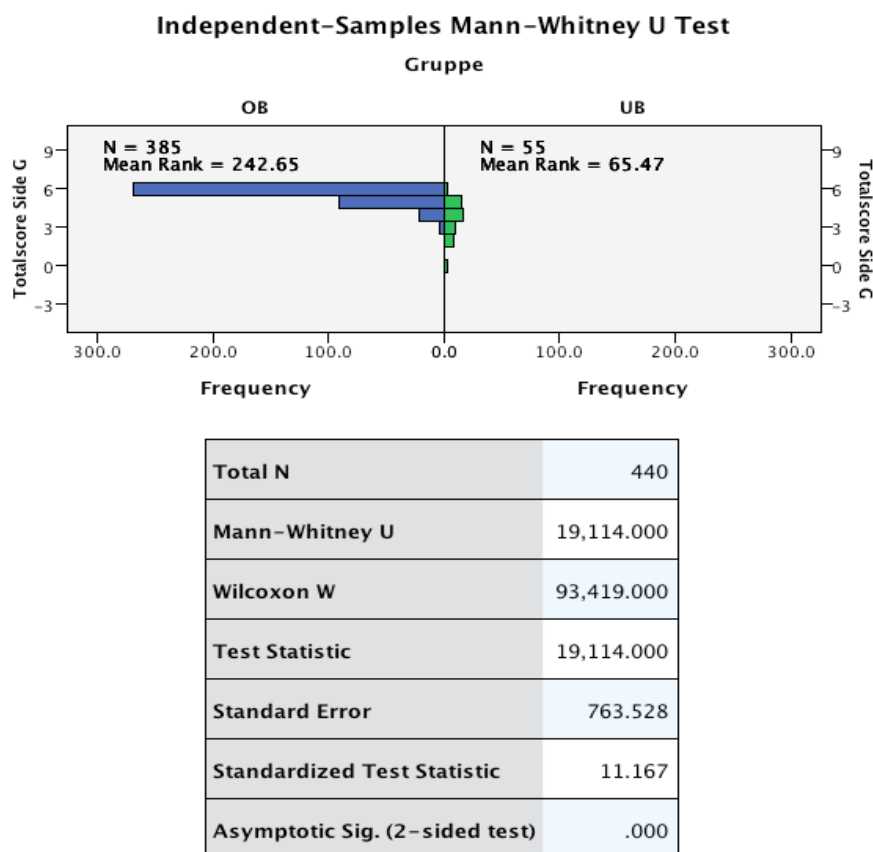
**Tabell 27 Prosentvis fordeling av svar**

|           | Riktig |       | Ubesvart |      | Tier  |       |
|-----------|--------|-------|----------|------|-------|-------|
|           | UB     | OB    | UB       | OB   | UB    | OB    |
| Oppgave 2 | 29,1%  | 84,2% | 21,8%    | 3,1% | 36,4% | 7,8%  |
| Oppgave 3 | 20%    | 70,6% | 21,8%    | 3,1% | 31 %  | 20,5% |

Tabellen tydeliggjør hvordan utfordringer knyttet til tieroverganger er sentrale for begge grupperingene. Det er også interessant å merke seg at antall ubesvart er likt for begge gruppene i begge oppgavene. Det gjør det mulig å spekulere i om det er vanskegraden på oppgavene elevene har utfordring med, og ikke tidsaspektet. Oppsummert for siden ser en at svaralternativet som representere feilmønsteret med svar som avviker med 10 har stor effekt for scoreoppnåelse for begge gruppene.

## 4.9 Side G – Oppstilte regneoppgaver

Siden består av 6 regneoppgaver, 1 av kategori RR4, 2 av kategori RR5 og 3 av kategori RR3. Effekten for siden målt med Mann-Whitney U – og Welch´s test presenteres i Figur 25. Etterfulgt av poengscore fordelt på de to gruppene i Figur 26. Deretter omtales effekten av hver enkelt oppgavene og eventuelle feilmønstre som fremkom. Avslutningsvis oppsummeres funnene for siden.



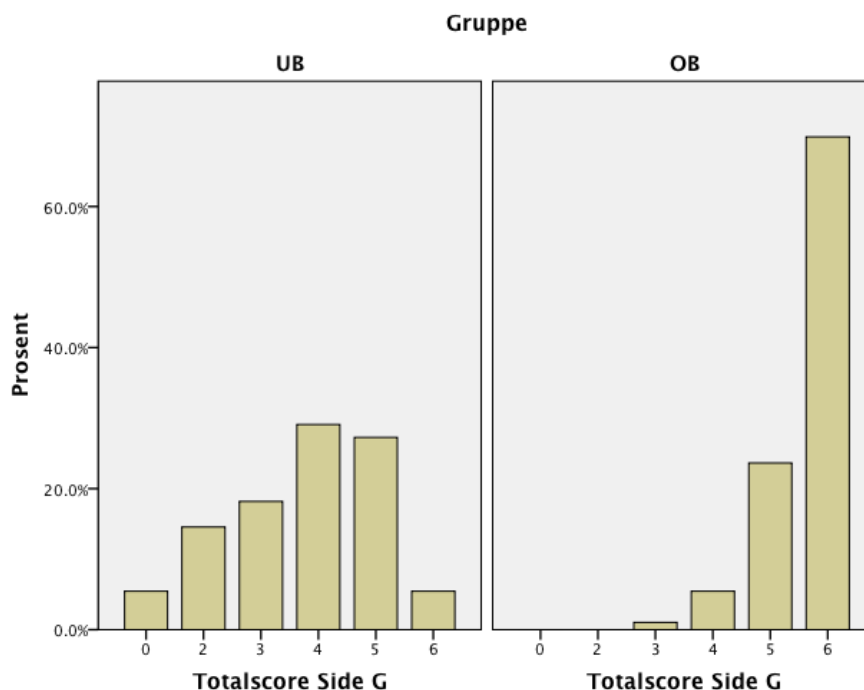
Figur 25 Mann-Whitney U test for Side G

Mann-Whitney U test indikerte statistisk signifikant forskjell i score på Side G mellom gruppen Under bekymringsgrensen ( $Md = 4$ ,  $n = 55$ ) og gruppen Over bekymringsgrensen ( $Md = 6$ ,  $n = 385$ ),  $U = 19\,114$ ,  $z = 11.167$ ,  $p < .001$ , effekten var stor,  $r = 0.53$ .

En Welch´s test ble også gjennomført for å sammenligne sidscore mellom gruppen Over bekymringsgrensen og gruppen Under bekymringsgrensen. Resultatet indikerte statistisk signifikant forskjell i scoren mellom gruppen Under bekymringsgrensen (M



= 3.69, SD = 1.438) og gruppen Over bekymringsgrensen (M = 5.62, SD = .638 ; t (57) = -9.825, p < .001 tohalet). Størrelsen av forskjellen på gjennomsnittene (mean difference = -1.932, 95% CI = -2.326 til -1.539) var stor (eta squared = 0.1805 = 18.1% ).



Figur 26 Poengscore Side G prosentvis i gruppene

Figuren 26 viser tydelig takeffekten for elevene i gruppen Over bekymringsgrensen. For elevene i gruppen Under bekymringsgrensen er situasjonen mer tilnærmet normalfordelt. Kun 5,5% av elevene i gruppen Under bekymringsgrensen scorer fullt på denne siden mot 69,9% av elevene i gruppen Over bekymringsgrensen. Totalt for siden for utvalget er p-verdien 61,8%, som er i ytterkant av hva kartleggingsprøver søker. Det er tydelig at denne oppgavesiden skiller godt på de to gruppene, med  $r = 0.53$  og en eta squared på 18,1%. Videre beskrives feilmønstre som fremkom på siden og mulige årsaker til oppgavesidens effekt for gruppetilhørighet. Svarene fordelt i de ulike kategoriene som fremkom på Side G presenteres i Tabell 28. Prosentvis fordeling tilknyttet årsak til feilmønsteret er presentert i Tabell 29.

Totalt antall svar på siden er 2640 (6 oppgaver x 440 elever), fordelt på 330 svar fra gruppen Under bekymringsgrensen og 2310 svar fra gruppen Over bekymringsgrensen.

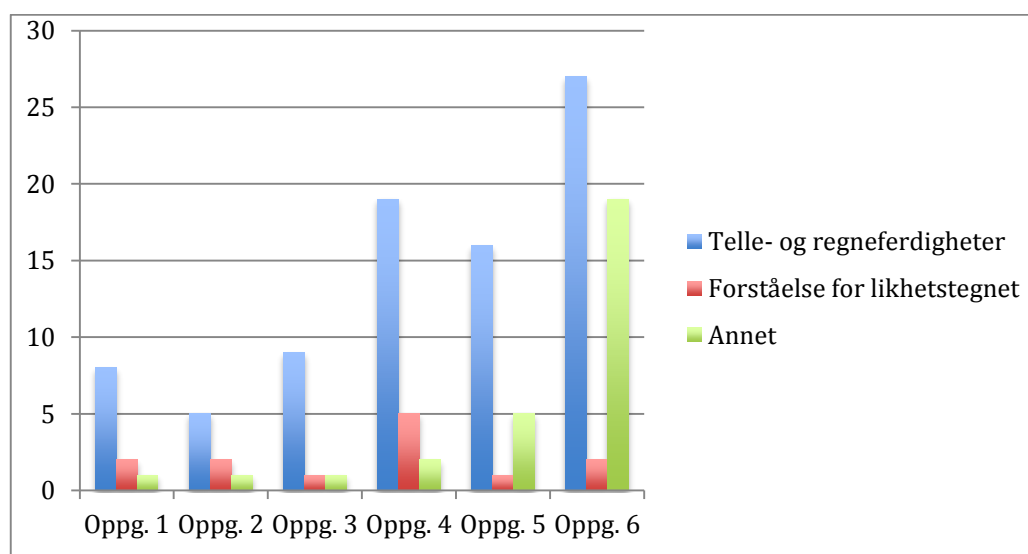
**Tabell 28 Svar fordelt i kategoriene for Side G**

|                 | M1              | M2           | M3           | M4          | M5           | M6           | M7            | Sum feil       |
|-----------------|-----------------|--------------|--------------|-------------|--------------|--------------|---------------|----------------|
| UB<br>(antall)  | 61.5%<br>(203)  | 19%<br>(63)  | 6.4%<br>(21) | 0.3%<br>(1) | 1.8%<br>(6)  | 2.1%<br>(7)  | 8.8%<br>(29)  | 38.5%<br>(127) |
| OB<br>(antall)  | 93.7%<br>(2165) | 1.4%<br>(33) | 0.6%<br>(15) | 0.1%<br>(2) | 0.3%<br>(6)  | 0.8%<br>(17) | 3.1%<br>(72)  | 6.3%<br>(145)  |
| Sum<br>(antall) | 90%<br>(2368)   | 3.6%<br>(96) | 1.4%<br>(36) | < 1%<br>(3) | < 1%<br>(12) | 1%<br>(24)   | 4.9%<br>(129) | 10%<br>(272)   |

**Tabell 29 Prosentvis fordeling tilknyttet årsak for feilmønster**

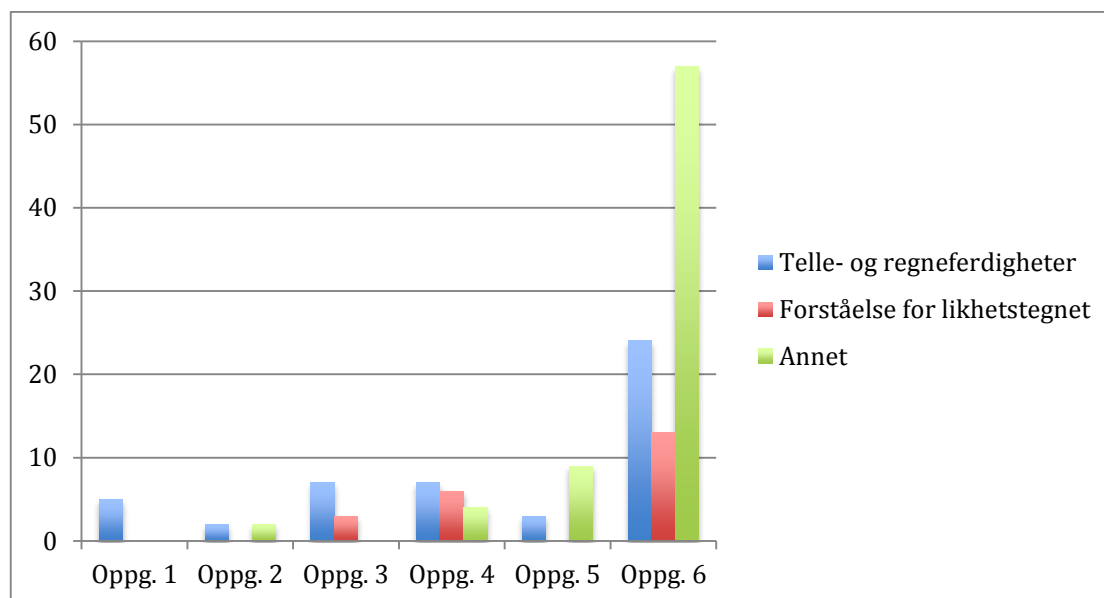
|    | Riktig svar | Svake regne-<br>telleferdigheter | Forståelse for<br>likhetstegnet | Andre | Totalt |
|----|-------------|----------------------------------|---------------------------------|-------|--------|
| UB | 61,5%       | 25,5%                            | 4,2 %                           | 8,8 % | 100%   |
| OB | 93,7 %      | 2,1%                             | 1,1 %                           | 3,1 % | 100%   |

Utfallet fra de innledende analysene viser at den største konsentrasjonen av feilsvar for elevene i gruppen Under bekymringsgrensen er i kategoriene *ubesvart* (M2), *annet* (M7) og *tellefeil* (M3). Figur 27 viser hvordan de ulike årsakene tilknyttet feilmønstrene gjør seg gjeldende for gruppen på de ulike oppgavene på Side G.



**Figur 27 Feilmønstre tilknyttet årsak og oppgavene 1-6 Side G, for gruppen Under bekymringsgrensen**

Telle- og regneferdigheter ser ut til å være den mest avgjørende faktoren, oppgave 4 og 6 skiller seg noe ut. Figur 28 viser samme oversikt for gruppen Over bekymringsgrensen.



Figur 28 Feilmønstre tilknyttet årsak og oppgavene 1-6 Side G, for gruppen Over bekymringsgrensen

Fra de to figurene over blir det tydelig at oppgave 6 skiller seg ut for begge grupperingene. Oppgave 4 ser også ut til å stå i en særstilling. Felles for disse oppgavene er at de er av kategori RR5. Figurene viser også at oppgavekategorien gjør utslag på feilmønstre tilknyttet forståelse for likhetstegnet. Effekten for de ulike oppgavene er presentert i Tabell 30, oppgave 4 og 6 blir omtalt etter den.

Tabell 30 Effektstørrelser for Side G,  $p < .001$

|             | Oppg. 1 | Oppg. 2 | Oppg. 3 | Oppg. 4 | Oppg.5 | Oppg. 6 |
|-------------|---------|---------|---------|---------|--------|---------|
|             | RR3     | RR3     | RR4     | RR5     | RR3    | RR5     |
| r-verdi     | .33     | .27     | .27     | .50     | .45    | .44     |
| Eta squared | 2.6%    | 1.8*%   | 2.2**%  | 8.5%    | 6.7%   | 25.7%   |

\*  $p = .007$

\*\* $p = .003$

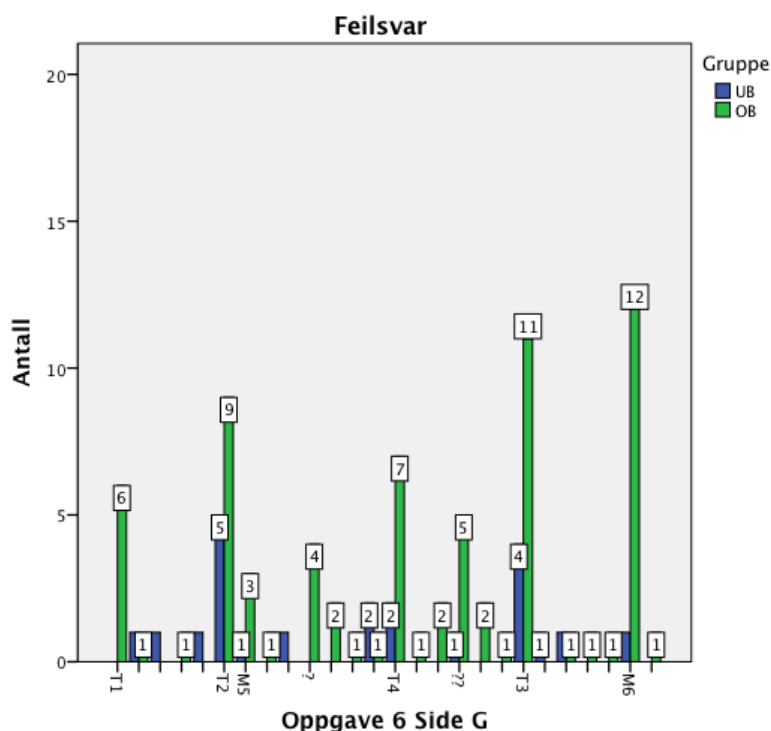
#### 4.9.1 Oppgave 4, RR5

Denne oppgaven skiller seg ut og har moderat mot stor effekt. Den er også interessant fordi den er brukt på Side A, da gitt med en annen instruksjon. To feilmønstre gjør seg gjeldende henholdsvis *tellefeil* (M3) og *gir a* (M4). For *tellefeil* er tallene 3 svar fra gruppen Under bekymringsgrensen og 2 svar fra gruppen Over bekymringsgrensen. *Gir a* (M4) har tallene 5 svar fra gruppen Under bekymringsgrensen og 2 svar fra gruppen Over bekymringsgrensen. Sammenlignes disse resultatene med resultatene for oppgaven når den er gitt på Side A, dukker det opp et interessant funn. Når oppgaven ble gitt på Side A som oppgave 2, var det 5 elever fra gruppen Under bekymringsgrensen som adderte de gitte tallene (M6). Når oppgaven blir gitt igjen som oppgavenummer 4 på Side G, er det ingen av elevene i gruppen Under bekymringsgrensen som adderer de gitte tallene (M6). Dette var interessant og det ble undersøkt videre for å se om muligens de 5 elevene fra Side A hadde latt være å svare på Side G. Det hadde de ikke, men noe overraskende viste analysen at disse 5 elevene hadde svart riktig på Side G. Hva kunne dette skyldes? Det er et par aspekt en må ta hensyn til for å avdekke årsaken her. Oppgavens nummer på siden, med tanke på tidsfaktoren, og instruksjonen gitt i forkant av oppgaven. Oppgaven er gitt som nummer 2 på Side A, og nummer 4 på Side G, hvis tidsfaktoren var innvirkende skulle elevene ha gjort det dårligere på Side G enn A. En står da igjen med instruksjonen for siden. Det ble kommentert under resultatet på Side A, at instruksjonen fokuserer på å *finne* og *fylle* inn. Instruksjonen gitt for Side G har fokus på å gjøre regnestykket riktig, med tanke på likhetstegnet. Dette kan tyde på at instruksjonen på Side G er mer hensiktsmessig enn instruksjonen på Side A, om ikke annet så for de elevene i gruppen Under bekymringsgrensen. Dette kan igjen være tilknyttet begrenset begrepsforståelse.

#### 4.9.2 Oppgave 6, RR5

Effekten for oppgaven er moderat mot stor, *ubesvart* (M2) er høy for begge grupperingene. Henholdsvis 26 elever fra gruppen Under bekymringsgrensen og 23 elever fra gruppen Over bekymringsgrensen, det bidrar til at oppgaven ikke får større effekt enn hva den har. Fra feilsvarene elevene gir er det tydelig at det har vært utfordrende for dem, med en relativ stor spredning. Oppgaven er siste oppgave i *Kartlegging i regning for 2. trinn* og resultatet indikerer at det kan ha hatt påvirkning.

Med bakgrunn i variansen er det valgt å legge ved et stolpediagram som viser spredningen av elevenes feilsvar.



Figur 29 Oversikt over feilsvar

Figur 29 viser spredningen tydelig. Svarene som ble avgitt av 4 eller flere elever er navngitt i figuren og blir omtalt.

#### 4.9.2.1 T1

Dette mønsteret tilsvarer å skrive tier som ener. Eksempelvis i stykket  $25 = \_ + 15$  ville svaret i denne kategorien være 1. 6 elever fra gruppen Under bekymringsgrensen og 1 elev fra gruppen Over bekymringsgrensen avgir dette svaret. Feilmønsteret kan knyttes til begrenset kunnskap om posisjonssystemet.

#### 4.9.2.2 Tier, T2 og T3

Disse mønstrene er kjent, T2 er svar som er 10 mindre enn riktig og T3 er svar som er 10 mer enn riktig. De to mønstrene er representert med henholdsvis 5 og 4 elever fra gruppen Under bekymringsgrensen og 9 og 11 fra gruppen Over bekymringsgrensen. Mønsteret knyttes til utfordringer med tieroverganger.

#### 4.9.2.3 Forståelse for likhetstegnet, M5, M6 og T4

M5 er kategorien for *gjengir b* og er representert med 1 elev fra gruppen Under bekymringsgrensen og 3 elever fra gruppen Over bekymringsgrensen. M6 er kategorien for *addere gitte tall* og er representert med 1 elev fra gruppen Under bekymringsgrensen og 12 elever fra gruppen Over bekymringsgrensen. T4 er i familie med *gjengir b* og er 10 mer enn b. Det vil si at det kan tenkes knyttet til utfordringer i tieroverganger, men det ser også ut til å være knyttet til forståelse av likhetstegnet.

#### 4.9.2.4 ? og ??

Dette er et tydelig mønster som hovedsakelig er representert med svar fra gruppen Over bekymringsgrensen (Se Figur 29). Det er ingen umiddelbar plausibel forklaring på dette, og det samsvarer ikke med noen kategorier som er brukt i tidligere studier. Derav er det gitt det symbolet ? og ?? i stolpediagrammet. Det er en differanse på 10 mellom ? og ??, som kan antyde at mønsteret har tilknytning til tieroverganger.

### 4.9.3 Oppsummering Side G

Funnene og analysen av denne siden viser at oppgaver av kategori RR5 gir en økning av feilsvar som er tilknyttet svak eller manglende forståelse for likhetstegnet. Side G ga også mulighet for å sammenligne påvirkningen instruksjonen fra lærer har i sammenheng med elevens løsninger på oppgaven. Det blir tydelig at et fokus på å gjøre regnestykkene riktig gir færre feilsvar enn fokus på å *finne* og *fylle inn*. Funnet understreker viktigheten av hva som vektlegges i instruksjonen til oppgaver.

#### **4.10 Oppsummering resultat og analyse**

Oppgavekategori er avgjørende for hvilke feilmønstre som er fremtredende. Oppgaver av kategori RR1/2 har færrest feilmønstre, og er således lettere å tolke for læreren som skal følge opp elevene og tydeligere i hva oppgaven kan kartlegge. Oppgaver av de andre kategoriene introduserer feilmønstre som er forbundet med likhetstegnet. Dette gjør det vanskeligere å avgjøre hvilken kompetanse elevene har angående gjenhenting av regnefakta, som igjen peker på regne- og tellestrategier. Dette på grunnlag av at disse oppgavene først og fremst krever en relasjonell forståelse for likhetstegnet for å bli avkodet riktig, for å så kunne løses korrekt.

Tekstoppgavene viser seg utfordrende for begge grupperingene, der oppgavesiden med åpne svar hadde lavest score. Resultatene for siden med åpne svar indikerer at feilavlesning eller feiltolkning av illustrasjonen er en av årsakene til den lave scoren. Dette gjør det vanskeligere for læreren som skal følge opp elevene, med tanke på at det er uklart hvilken kompetanse oppgavesiden egentlig kartlegger. For tekstoppgavene med svaralternativer indikerer resultatet at utfordringer i tieroverganger er en av årsakene til lav score.

## 5 Diskusjon

Indikasjoner av feilmønstre i kartleggingsprøven gir et viktig bidrag til matematikkfeltet i skolesammenheng, spesielt når kartleggingen er obligatorisk nasjonalt og blir foretatt så tidlig i skoleløpet som på 2. trinn. Det blir særlig relevant når det er hver enkelt lærer som står for oppfølging av elever som scorer lavt på disse prøvene. Kartleggingsprøven som denne studien har sett på bruker totalscoren som indikator på om elevene har svake begreper og svake telle- og regneferdigheter. En slik tolkning på totalscore gjør arbeidet med å identifisere hvilken kompetanse elevene egentlig mangler vanskeligere for læreren som skal følge opp eleven.

Lærerhverdagen er travel og det er ikke sannsynlig at flertallet av lærere vil gjennomføre en dyptgående analyse av feilmønstre i kartleggingsprøven. Det er også vanskeligere for en enkelt lærer å identifiserer feilmønstre. For det første, rettingen av prøvene gjøres manuelt ved at læreren går gjennom elevheftene enkeltvis. Det er kun poengscore på de ulike oppgavene som blir registrert digitalt. Siden de faktiske feilsvarene ikke blir registrert, blir det også vanskeligere for læreren å se feilmønstrene. For det andre, materialet læreren har tilgang til består som oftest bare av én klasse, og mønstrene blir således også mindre fremtredende. Et feilsvar som er gitt av noen få elever i klassen, vil gjerne ikke bli identifisert som et feilmønster med rot i misoppfattelser, men gjerne heller tilskrives tilfeldigheter.

Et alternativ, som kunne bedret lærerens mulighet til å identifisere feilmønstrene, er å endre fra registrering av poeng på oppgavene til registrering av elevenes faktiske svar på oppgavene. Et slikt grep vil kunne tydeliggjøre eventuelle feilmønstre, ved å registrere elevenes svar vil en også åpne for muligheten til å sammenligne elevsvar nasjonalt. En slik sammenligning vil bedre forutsetningen for å avdekke feilmønstre betraktelig. Dette åpner også for muligheten til at feilsvar kan diskuteres mellom flere lærere. En diskusjonsarena kan bidra til kunnskapsutvikling blant lærerne og også lette arbeidet med oppfølging av enkeltelever som scorer lavt på *Kartleggingsprøve i regning for 2. trinn*. En annen mulighet kunne vært at veiledningen til *Kartleggingsprøve i regning for 2. trinn* også inneholdt en oversikt med 'vanlige' feilsvar til de ulike oppgavesidene, og hva de ulike feilsvarene kunne indikere. I sær er det relevant når oppgavene som blir brukt består av flere lag, og at feilsvar eller



ubesvart da kan ha flere ulike årsaker, som igjen henger sammen med ulike mangelfulle kompetanser. I veiledningen til *Kartleggingsprøve i regning for 2. trinn* presiseres det at videre kartlegging er nødvendig i oppfølgingsarbeidet av lavt scorende elever (Utdanningsdirektoratet, 2014b), men her er det også en mulighet til å få mer informasjon ut av elevens svar på kartleggingsprøven hvis læreren vet hva de ulike feilsvarene indikerer med tanke på kompetanse.

En av de fremste indikatorene for matematikkvansker er stagnering eller rigiditet i strategibruk og -utvikling innen regning (Dowker, 2005a; Ostad, 2013; Utdanningsdirektoratet, 2014b). Et essensielt utviklingstrinn innen strategier er automatisering av regnefakta, som er grunnlaget for alle retrieval- strategier. For å kartlegge om elevene har automatisert regnefakta, eller hvor langt de er kommet mot å automatisere regnefakta må oppgavene konstrueres til å kartlegge akkurat dette. Resultatet av denne studien viser at flere av sidene med oppstilte regneoppgaver hovedsakelig kartlegger forståelse for likhetstegnet, og at kategorien *ubesvart* på disse oppgavesidene like plausibelt kan ha sammenheng med utfordringer i avkodingen av oppgavene. Dette gjelder for oppgaver av kategori RR3 til RR6, funnet er sammenfallende med tidligere forskning (Lindvall & Ibarra, 1980; Powell, 2014; Weaver, 1973).

Kronometrisk måling slik som *Kartleggingsprøve i regning for 2. trinn* benytter seg av med tidsbegrensning for å kartlegge strategier er også noe omdiskutert. Dette med begrunnelse i at ved tidsbegrensning vil også valget av strategier påvirkes (Dowker, 2005b; Ostad, 2013). Et annet temaet for kronometrisk måling av strategiutvikling er at det trengs en mer longitudinell tilnærming for å se utviklingen til elevene. Det er utviklingen av strategiene som er sentral, ikke at elevene på 2. trinn bruker backup-strategier. De fleste av elevene på 2. trinn bruker backup- strategier, det er den naturlige utviklingen av strategiene. Problemet er først hvis elevene kun fortsetter med backup- strategier, noe en vil karakterisere som strategifattigdom eller strategirigiditet. Skal en bruke kronometrisk måling for å kartlegge strategiutvikling blir det sentralt at oppgavene treffer på strategi og ikke avkoding eller andre kompetanser.

Flere studier har sett på viktigheten av relasjonell forståelse for likhetstegnet i matematikk (se for eksempel Byrd, Mcneil, Chesney & Matthews, 2015; Carpenter, et al., 2003; Falkner, et al., 1999), men få har sett på plassering av regnetegnet i forhold til likhetstegnet og hvilket utslag det gir (Ibarra & Lindvall, 1982; Weaver, 1973). Det er interessant at studiene som påpekte effekten av plassering av regnetegn i forhold til likhetstegn hovedsakelig er gjort en god tid tilbake. Det kan gi en indikasjon om at denne problemstillingen ikke har hatt fokus den senere tid. En kan da også sette spørsmålsteget ved om den gjennomsnittlige 2. trinns læreren er klar over de ulike feilmønstrene som indikerer svak eller manglende forståelse for likhetstegnet i disse oppgavekategoriene. Etablering av en god forståelse for likhetstegnet tidlig i skoleløpet har det lenge vært enighet om, men trenden som gjør seg gjeldende i denne studien er at likhetstegnet fremdeles er en betydelig utfordring for elever på 2. trinn. For å bedre forståelsen, eller legge til rette for relasjonell forståelse for likhetstegnet, er bruk av ukanoniserte likninger (RR3-RR6) tidlig i skoleløpet sentralt (N. M. Mcneil & M. W. Alibali, 2005), gjerne i sammenheng med klasseromsdiskusjoner hvor ulike oppfatninger av likhetstegnet kan løftes frem i plenum (Van De Walle, 2007). Kartleggingsprøven sitt fokus på bruk av ukanoniserte likninger kan i den forstand sees på som en positiv utvikling, men resultatet fra studien tyder på at den samme utviklingen ikke nødvendigvis har funnet sted i skolene, i det minste ikke for studiens utvalg.

Bruken av ukanoniserte likninger i kartlegging blir sånn sett tosidig. På den ene siden er det positivt at elevene får møte disse likningene tidlig i skoleløpet, på den andre siden kan det diskuteres om elevenes forståelse for likhetstegnet er det som skal kartlegges når forankringen til *Kartleggingsprøve i regning for 2. trinn* er i kompetansemålene for 2. trinn (Kunnskapsløftet, 2006) og regning som grunnleggende ferdighet (Utdanningsdirektoratet, 2012). Flere av oppgavesidene som er tungt vektet med poeng krever relasjonell forståelse for likhetstegnet for å avkode oppgavene riktig. Det vil si at forståelse for likhetstegnet blir vektlagt i prøven. Elever som ikke viser relasjonell forståelse for likhetstegnet, viser sine telle- og addisjonsferdigheter ved å legge sammen de gitte tallene. Vi har med andre ord flere sider i *Kartleggingsprøve i regning for 2. trinn* hvor elever som regner riktig med tanke på addisjon, ikke får poeng fordi de adderer feil tall. Årsaken til dette er ikke feilaktig addisjon, men forståelse for likhetstegnet. Relasjonell forståelse for

likhetstegnet er viktig, og sentralt med tanke på videre læring i matematikk, eksempelvis i algebra og likninger (Carpenter, et al., 2003; E. J. Knuth, et al., 2006). Men, i 2. klasse der fokuset er på å komme seg videre fra backup- strategier og bli stødige og fleksible tellere og regnere, kan det tenkes det hadde vært mer hensiktsmessig å kartlegge disse kompetansen hver for seg.

Funnene som er gjort vedrørende oppgaver av kategori RR1/2 var noe uventede. Forventningen til feilsvarene i disse oppgavene, var en tiltakende økning i kategoriene *ubesvart* og *tellefeil* gjennom siden, denne effekten var tilstede i hovedsak for gruppen Under bekymringsgrensen. Et annet like fremtredende mønster var feil gjort i forbindelse med tieroverganger. Mønsteret knyttet til tieroverganger var også representert i tekstopp-gavene med svaralternativ, men kan ikke tillegges like mye vekt nettopp på grunn av svaralternativene. Mønsteret kan indikere svake tellestrategier, men kan også være et tegn på at elevene har begynt å bruke algoritmen, men har utfordringer med hvordan tieroverganger skal behandles der. Trenden indikerte også at det var hovedsakelig elevene i gruppen Under bekymringsgrensen som avga svar som var 10 mindre enn riktig, mens elevene i gruppen Over bekymringsgrensen avga svar som var 10 mer enn riktig. Denne forskjellen innen samme feilmønster, indikerer bruk av tellestrategier for elevene i gruppen Under bekymringsgrensen og utfordringer med algoritmen for elevene i gruppen Over bekymringsgrensen. Ved å løfte frem slike nyanseringer innen samme feilmønster, kan en bedre forutsetningen for en hensiktsmessig oppfølging av enkeltelever. Det kunne vært interessant med en større studie knyttet til norske elevers kompetanse til å løse oppgaver i addisjon og subtraksjon i kategoriene RR1-RR6, og feilmønstre knyttet til de ulike kategoriene. Spesielt hvis studien avgrenset tallområdet som blir brukt slik at en enda tydeligere får mulighet til å se hva de ulike kategoriene krever av ulik kompetanse for elevene.

Bruk av halvkonkreter i tekstopp-gaver er et annet tema som kan diskuteres. Er det fremdeles relevant å ha mynter som halvkonkreter i kartlegging? Penger har lenge vært brukt som halvkonkreter og konkrete i regning, men er dette like relevant i 2015? Flertallet bruker nå kort og mobil som betalingsmiddel, ikke kontanter. Spørsmålet blir da om penger som halvkonkreter skaper større distansene mellom elevene og oppgavene, i stedet for å sette oppgavene i en hverdagskontekst som er kjent for elevene. En kan argumentere med at enkel aritmetikk i egen valuta faller inn

under regning som grunnleggende ferdighet, og således er en ferdighet elevene bør beherske. Men da blir også spørsmålet om hvorfor en bruker tegninger av 'lekepenger' i stedet for bilder av norske penger relevant. Hvis tanken er å kartlegge elevens kompetanse i håndtering av egen valuta, ville bilder av reelle penger vært nærmere konteksten en søker informasjon om. Halvkonkretene som er brukt i *Kartleggingsprøve i regning for 2. trinn* er ikke oppgitt med benevnning, og således distanserer de seg enda mer fra realiteten. Studier (Nicole M Mcneil, Uttal, Jarvin & Sternberg, 2009) har også påpekt at perseptuelt rike konkreter (fargerike og mye detaljer) gir flere feilsvar enn konkreter som er mindre perseptuelt rike. Det kan tenkes at valget er tatt med bakgrunn i et ønske om å redusere belastningen av arbeidsminnet (jf. Kap 2.1.3). Men, Nicole M Mcneil, et al. (2009) fant også at feilsvar gjort med perseptuelt rike konkreter (penger) hadde mindre sannsynlighet for å være basert på misoppfattelser enn feilsvarene som var gjort med mindre perseptuelt rike konkreter, og i den forstand bidro de perseptuelt rike konkretene i elevenes oppgaveløsning. Studien (Nicole M Mcneil, et al., 2009) ble foretatt i en amerikansk kontekst, og det er ikke sikkert at samme tendens gjør seg gjeldende i norsk kontekst. Implikasjoner til videre forskning kan være studier som omhandler norske elever og bruken av penger som konkreter og halvkonkreter, med tanke på hvilke halvkonkreter som er best egner seg for å nærme seg elevens hverdagskontekst.

Konkreter og halvkonkreter brukes gjerne for å støtte elever i oppgaveløsningen. De halvkonkretene *Kartleggingsprøve i regning for 2. trinn* benytter seg av har skjult tierstruktur. Det vil si at elevene ikke kan peketelle halvkonkretene, fordi det er gjerne én halvkonkret som symboliserer 5 eller 10 enheter. I avklaringen for hvordan grunnleggende ferdigheter er tolket for 2. trinn pekes det på at det kan handle om noe så enkelt som å forstå at syv epler (halvkonkreter/tegning) kan representeres med tallet 7 (Utdanningsdirektoratet, 2014b). Den store forskjellen her er at eplene kan telles en for en, mens det ved halvkonkretene som er brukt ikke lar seg telle én for én. En kan da stille spørsmål til om oppgavene er enkle nok til at de faktisk får kartlagt det som er ønsket. Totalt i *Kartleggingsprøve i regning for 2. trinn* er det kun én siden med reell telling én for én. De andre sidene som innebærer telling krever kompetanse i mer enn bare telling for å lykkes, for eksempel bruk av og forståelse for tallinjen. Det er klart at bruk og forståelse av tallinjen også er et viktig område å få kartlagt,

men ved å kombinere de ulike kompetansene inn i oppgavene blir det vanskelig å forstå i etterkant hvilken kompetanse elevene faktisk har utfordringer med.

Kategorien *ubesvart* er interessant og kan diskuteres med tanke på reliabilitet for årsakstilknytning. I denne studien er den slått sammen med *tellefeil*, og representerer feilmønstre som har bakgrunn i tidsbegrensingen på *Kartleggingsprøve i regning for 2. trinn*. Er denne kategorien kun representativ for tidsfaktoren på en oppgaveside? Resultatet fra studien viser at den ikke kan tolkes så entydig. Avkoding av oppgaven og forståelse av instruksjonen som er gitt av lærer kan ansees som like plausible forklaringer for ubesvarte oppgaver. Dette viser seg i analysen ved at det gjerne er en oppgave midt på siden som flere elever velger å la stå ubesvart. Det er vanskelig å argumentere for at alle disse elevene valgte samme tilfeldige løsningsrekkefølge, og at akkurat denne oppgaven ble den siste og dermed forble ubesvart på grunn av for liten tid. Det er derimot mer sannsynlig at oppgaven opplevdes vanskelig slik at eleven lot den stå ubesvart og gikk videre, muligens med en tanke om å komme tilbake til den senere. Uansett blir kategorien *ubesvart* vanskelig å tolke med tanke på årsakstilknytning. Analysene viser at oppgavesider med RR1 har færre ubesvarte oppgaver enn sider med oppgaver av de andre kategoriene, trass i problem-size effekten. Det peker i retning av at avkodingen av oppgaver er en faktor som bør sees i sammenheng med kategorien *ubesvart*. Når det er sagt, er sidene med oppgaver av kategori RR1 også sidene med færrest oppgaver, og det kan være en feilkilde. Tiden som elevene blir gitt på de ulike sidene er justert etter oppgaveantall, så konklusjonen anses fremdeles valid.

Tolkning av illustrasjoner (kategorien *Till*) viste seg å være en betydelig årsaksfaktor for lav scoreoppnåelse på sidene med tekstoppgaver. Bruken av halvkonkreter i disse oppgavene er allerede diskutert, men Side D som viste seg å gi særdeles fremtredende mønstre, som kunne sees i sammenheng med feiltolkning av illustrasjonen, har ikke blitt drøftet. Hovedårsaken til feilmønstrene på Side D var feiltolkning av illustrasjonen, ved at elevene brukte feil tall fra illustrasjonen i sine regnestykker. Årsaken her var mest sannsynlig at tallenes plassering i forhold til objektet varierte, noe som resulterte i at flere elever leste tallet til naboobjektet (tallet som var i leseretning til objektet). Feilmønsteret kan også sees i sammenheng med at illustrasjonen som er brukt er informativ, som krever mer av arbeidsminnet enn andre

illustrasjoner (Elia & Philippou, 2004). Dette mønsteret kunne kanskje vært unngått hvis tallene var plassert konsekvent i forhold til objektet det tilhørte. Men burde de vært det? I en hverdagskontekst er ikke alltid systemene like synlig, og evnen til å se mønstre og sammenhenger i sammensatte situasjoner vil gagne elevene både i regningen og i hverdagen. Det er dog mulig at denne oppgaven kunne åpnet for at elevene skrev regnestykket de svarte på, slik at læreren kunne se om det var addisjonen eller evnen til å se sammenhenger i illustrasjonen elevene hadde utfordringer med.

Implikasjoner til videre forskning basert på denne studien er bruk av kartleggingsoppgaver med færre og flere lag av vanskegrad, for å undersøke effekt med tanke på identifisering av elever med svake begreper og svake telle- og regneferdigheter.

## 5.1 Konklusjon

Svar på forskningsspørsmålene denne studien er basert på kom frem gjennom resultatene og analysen, her oppsummeres de kort, med unntak av forskningsspørsmål d) som ble besvart under kapittel 4.1.

**Hvilke indikasjoner gir feilmønstre i elevsvar fra oppstilte og uoppstilte regneoppgaver i Utdanningsdirektoratet sin *Kartleggingsprøve i regning for 2. trinn*?**

- a) Finnes det mønstre i elevens feilsvar og hva kan eventuelt årsaken til disse mønstrene være?
- b) Hva kjennetegner oppgavene som elevene ikke scorer på?
- c) Hvilken kompetanse, eventuelt kompetanser er det disse elevene tilsynelatende mangler?
- d) Hvordan skiller Stavangerprosjektet sin scoring seg fra Utdanningsdirektoratet sin?

Det viste seg å være mønstre som gikk igjen i feilsvar på oppstilte og uoppstilte regneoppgaver i *Kartleggingsprøve i regning for 2. trinn*. Hovedsakelig var disse forbundet med ubesvart, forståelse for likhetstegnet, tellefeil og tieroverganger, der det siste mønsteret var mest fremtredende for gruppen Over bekymringsgrensen. Årsakene til mønstrene er hovedsakelig utfordringer knyttet til avkoding av ikke-kanoniserte oppgaver, svak eller mangelfull forståelse for likhetstegnet og svake telle- og regneferdigheter. Funnet om svak eller mangelfull forståelse av likhetstegnet er sammenfallende med tidligere studier (Lindvall & Ibarra, 1980; Weaver, 1973). Implikasjonene av dette resultatet er at arbeidet med å tidlig etablere en relasjonell forståelse for likhetstegnet bør prioriteres i småskolen. Spesielt med bakgrunn i de hindringer en operasjonell forståelse for likhetstegnet kan utgjøre for videre læring i matematikk (Byrd, et al., 2015; Nicole M. Mcneil, 2008; N. M. Mcneil & M. W. Alibali, 2005).

Det skal bemerkes at det er vanskelig å avgjøre om kategorien *ubesvart* tilhører svake telle- og regneferdigheter eller utfordringer i avkoding av oppgavene. Dette kan ansees som en svakhet ved studien, og det er mulig at en mer kvalitativ vinkling kunne vært oppklarende. For eksempel ved gjennomgang av prøveheftene til elevene

for se om elevene hadde skrevet noe, eller avsatt blyantmerker som kunne bidratt til å kaste lys over hvorfor oppgavene er ubesvart.

Oppgaveoppstilling spiller en stor rolle for svarene elevene gir, noe som indikerer at det ikke er strategivalg som i første omgang er avgjørende, men avkoding av oppgaven. Kjennetegn for oppgaver elevene scoret lavt på viste seg å kunne forklares ved hjelp av RR-kategoriene. For de oppstilte oppgavene var det hovedsakelig oppgaver av kategori RR3 til RR6 som var utfordrende, der RR5 og RR6 var oppgavene som fikk lavest score. Det vil si oppgavene som enten hadde en ukjent addend (RR3 og RR4), eller oppgaver hvor regnetegnet var til høyre for likhetstegnet (RR5 og RR6). Et annet gjennomgående trekk var at oppgaver som var sist på siden fikk lavere score, noe som kan indikere at tidsbruk var en påvirkende faktor for score, men i og med at oppgavens vanskegrad også økte utover siden, var ikke dette mulig å avgjøre. Når det kommer til hvilken kompetanse feilmønstrene indikerer at elevene mangler, viser studien at elevene i gruppen Under bekymringsgrensen har begrenset kompetanse i avkoding av oppgaver som ikke er kanoniserte, trolig grunnet operasjonell forståelse for likhetstegnet. Avkoding av illustrasjoner er også en begrenset kompetanse hos denne gruppen, dette kan grunne i støy i illustrasjonene eller være forårsaket av avstanden mellom illustrasjonene, som belaster arbeidsminnet (Elia & Philippou, 2004). Oppgavesidene med tekstoppgaver var sammensatte, noe som gjør det vanskelig å peke på hvilket aspekt som er årsak til feilmønstrene. Implikasjoner av dette er å være oppmerksom på hvordan bruk og plassering av illustrasjoner påvirker elevens svar, med tanke på hvilken informasjon en ønsker å kartlegge.

Studien viser også at elevene i gruppen Under bekymringsgrensen har svake telle- og regneferdigheter. Funnet baseres på feilmønstre som er knyttet til disse ferdighetene, og den andelen mønstrene utgjorde av elevens svar. Funnet er sammenfallende med formålet til *Kartleggingsprøve i regning for 2. trinn* (Utdanningsdirektoratet, 2014b).

Angående generalisering av funnene fra studien til den øvrige populasjonen, møter ikke utvalget alle krav til dette punktet. Det er likevel sannsynlig at funnene fra denne studien også har relevans for den øvrige populasjonen, da det er rimelig å anta at



feilsvar fra andre elever vil variere innenfor samme spenn som i studien, mens prosentandel og effekt for hver av sidene kan være ulik fra klasse til klasse.

Med tanke på veien videre kunne det vært interessant med en studie om hvilken nytteverdi lærerne mener *Kartleggingsprøve i regning i 2. trinn* har og hvordan resultater fra den blir brukt i oppfølging av enkeltelever. Blir prøven ansett som et supplement eller en plikt, eller noe midt i mellom? Resultater fra en slik studie vil være et verdifullt bidrag inn mot utvikling av neste kartleggingsprøve.

## Referanseliste

- Ainsworth, S. (2006). Deft: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.
- Alibali, M. W., Phillips, K. M. O. & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development, 24*(2), 89-101.
- Alseth, B., Throndsen, I. & Turmo, A. (2007). Rapport fra utvikling og pilotering av "regneprøven" *Acta Didactica Oslo: Institutt for Lærerutdanning og Skoleutvikling*
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory *Cognition, 44*(1-2), 75-106.
- Ashcraft, M. H. & Stazyk, E. H. (1981). Mental addition: A test of three verification models. *Memory & cognition, 9*(2), 185-196.
- Baroody, A. J. & Dowker, A. (2003). *The development of arithmetic concepts and skills : Constructing adaptive expertise*. Mahwah, N.J: Lawrence Erlbaum.
- Baroody, A. J. & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *Elementary School Journal, 84*(2), 199-212.
- Baroudi, Z. (2006). Easing students' transition to algebra. *Australian Mathematics Teacher, 62*(2), 28-33.
- Behr, M. (1980). How children view the equals sign. *Mathematics Teaching, 92*, 13-15.
- Berends, I. & Van Lieshout, E. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction, 19*(4), 345-353.
- Bobis, J., Sweller, J. & Cooper, M. (1994). Demands imposed on primary-school students by geometric-models. *Contemporary Educational Psychology, 19*(1), 108-117.
- Borge, I. C., Sanne, A., Nortvedt, G. A., Meistad, J. A., Skrindo, K., Ranestad, K., . . . Kristensen, T. E. (2014). Matematikk i norsk skole anno 2014 *Fagjennomgang av matematikkfagene - Rapport fra ekstern arbeidsgruppe oppnevnt av Utdanningsdirektoratet*.
- Byrd, C. E., Mcneil, N. M., Chesney, D. L. & Matthews, P. G. (2015). A specific misconception of the equal sign acts as a barrier to children's learning of early algebra. *Learning and Individual Differences, 38*, 61-67.
- Carpenter, T. P., Franke, M. L., Levi, L., Bass, H. & Ball, D. L. (2003). *Thinking mathematically : Integrating arithmetic and algebra in elementary school*. Portsmouth, N.H: Heinemann.
- Carpenter, T. P. & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. I T. P. Carpenter, T. A. Romberg & J. M. Moser (red.), *Addition and subtraction : A cognitive perspective*. Hillsdale, N.J: L. Erlbaum Associates.
- Chandler, P. & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10*(2), 151-170.
- Cobb, P. (1987). An investigation of young children's academic arithmetic contexts. *Educational Studies in Mathematics, 18*(2), 109-124.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. utg.). Hillsdale, N. J: Laurence Erlbaum.

- Colin, P., Chauvet, F. O. & Viennot, L. (2002). Reading images in optics: Students' difficulties and teachers' views. *International Journal of Science Education*, 24(3), 313-332.
- Cornoldi, C. & Lucangeli, D. (2004). Arithmetic education and learning disabilities in Italy. *Journal of Learning Disabilities*, 37(1), 42-49.
- Cowan, R. (2003). Does it all add up? Changes in children's knowledge of addition combinations, strategies, and principles. In A. J. Baroody & A. Dowker (red.), *The development of arithmetic concepts and skills : Constructing adaptive expertise*. Mahwah, N.J: Lawrence Erlbaum.
- Daroczy, G., Wolska, M., Meurers, W. D. & Nuerk, H.-C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. [Review]. *Frontiers in Psychology*, 6, 348.
- Demetriou, A. (2002). *The development of mental processing : Efficiency, working memory, and thinking* (vol. 268). Boston, Mass: Blackwell.
- Dowker, A. (2005a). Early identification and intervention for students with mathematics difficulties. *Journal of Learning Disabilities*, 38(4), 324-332.
- Dowker, A. (2005b). *Individual differences in arithmetic : Implications for psychology, neuroscience and education*. Hove: Psychology Press.
- Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61(1-2), 103-131.
- Elia, I., Gagatsis, A. & Demetriou, A. (2007). The effects of different modes of representation on the solution of one-step additive problems. *Learning and Instruction*, 17(6), 658-672.
- Elia, I. & Philippou, G. (2004). *The function of pictures in problem solving*. Paper presented at the Conference of the International Group for the Psychology of Mathematics Education, Bergen.
- Falkner, K. P., Levi, L. & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching children mathematics*, 6(4), 232-236.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics : And sex and drugs and rock 'n' roll* (4. utg.). Los Angeles: SAGE.
- Gelman, R. & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gelman, R. & Meck, E. (1983). Preschoolers' counting: Principles before skill. *Cognition*, 13(3), 343-359.
- Gilmore, C. K. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development*, 21(3), 301-316.
- Gray, E. M. (1991). An analysis of diverging approaches to simple arithmetic: Preference and its consequences. *Educational Studies in Mathematics*, 22(6), 551-574.
- Gray, E. M. & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic. *Journal for research in Mathematics Education*, 116-140.
- Hopfenbeck, T. N., Kjærnsli, M. & Olsen, R. V. (2012). *Kvalitet i norsk skole : Internasjonale og nasjonale undersøkelser av læringsutbytte og undervisning*. Oslo: Universitetsforlaget.
- Ibarra, C. G. & Lindvall, C. M. (1982). Factors associated with the ability of kindergarten children to solve simple arithmetic story problems. *Journal of Educational Research*, 75(3), 149-155.

- Johannessen, A., Christoffersen, L. & Tufte, P. A. (2010). *Introduksjon til samfunnsvitenskapelig metode* (4. utg.). Oslo: Abstrakt.
- Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, 12(3), 317-326.
- Kjærnsli, M. & Olsen, R. V. (2013). Pisa 2012 - sentrale funn. I M. Kjærnsli & R. V. Olsen (red.), *Fortsatt en vei å gå. Norske elevers kompetanse i matematikk, naturfag og lesing i pisa 2012* (s. 13-43). Oslo: Universitetsforlaget.
- Knuth, E., Alibali, M., Mcneil, N., Weinberg, A. & Stephens, A. (2011). Middle school students' understanding of core algebraic concepts: Equivalence & variable. I J. Cai & E. Knuth (red.), *Early algebraization* (s. 259-276): Springer Berlin Heidelberg.
- Knuth, E. J., Stephens, A. C., Mcneil, N. M. & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, 37(4), 297-312.
- Kunnskapsløftet. (2006). *Læreplan i matematikk fellesfag*.
- Landau, M. & Lesh, R. (1983). *Acquisition of mathematics concepts and processes*. New York: Academic Press.
- Lesenteret. (2015, 06.03.2015). Stavangerprosjektet - det lærende barnet Lastet, 2015, fra <http://lesesenteret.uis.no/stavangerprosjektet/>
- Levin, J. R. (1981). On functions of pictures in prose I. I F. J. Pirozzolo & M. C. Wittrock (red.), *Neuropsychological and cognitive processes in reading* (s. 203-228). San Diego, CA: Academic Press.
- Lindvall, C. M. & Ibarra, C. G. (1980). Incorrect procedures used by primary grade pupils in solving open addition and subtraction sentences. *Journal for Research in Mathematics Education*, 50-62.
- Matthews, P. & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, 104(1), 1-21.
- Mayer, R. E. & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83(4), 484-490.
- Mcintosh, A., Settemsdal, M. R., Stedøy-Johansen, I., Arntsen, T. J. & Opplæringen, N. S. F. M. I. (2007). *Alle teller!: Håndbok for lærere som underviser i matematikk i grunnskolen : Kartleggingstester og veiledning om misoppfatninger og misforståelser på området : Tall og tallforståelse*. Trondheim: Matematikksenteret.
- Mcneil, N. & Alibali, M. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development*, 6(2), 285-306.
- Mcneil, N. M. (2008). Limitations to teaching children  $2 + 2 = 4$ : Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, 79(5), 1524-1537.
- Mcneil, N. M. & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, 28(3), 451-466.
- Mcneil, N. M. & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, 76(4), 883-899.
- Mcneil, N. M., Rittle-Johnson, B., Hattikudur, S. & Petersen, L. (2010). Continuity in representation between children and adults: Arithmetic knowledge hinders

- undergraduates' algebraic problem solving. *Journal of Cognition and Development*, 11(4), 437-457.
- Mcneil, N. M., Uttal, D. H., Jarvin, L. & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learning and Instruction*, 19(2), 171-184.
- Nesh, D. N. F. K. (2006). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi. Lastet ned fra <https://http://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/forskningsetiske-retningslinjer-for-samfunnsvitenskap-humaniora-juss-og-teologi-2006.pdf>
- Nesher, P., Greeno, J. G. & Riley, M. S. (1982). The development of semantic categories for addition and subtraction. *Educational Studies in Mathematics*, 13(4), 373-394.
- Opplæringslova. Lov av 7. Juni 1998 nr. 61 om grunnskolen og den vidaregåande opplæringa.: Lastet ned fra [https://lovdata.no/dokument/NL/lov/1998-07-17-61/KAPITTEL\\_1#KAPITTEL\\_1](https://lovdata.no/dokument/NL/lov/1998-07-17-61/KAPITTEL_1#KAPITTEL_1).
- Ostad, S. (1999). Developmental progression of subtraction strategies: A comparison of mathematically normal and mathematically disabled children. *European Journal of Special Needs Education*, 14(1), 21-36.
- Ostad, S. A. (1999). *Elever med matematikkvansker: Studier av kunnskapsutviklingen i strategisk perspektiv*. Oslo: Unipub.
- Ostad, S. A. (2008). *Strategier, strategiobservasjon og strategiopplæring: Med fokus på elever med matematikkvansker : Ressurshefte til boken strategier, strategiobservasjon og strategiopplæring*. Trondheim: Læreboka forlag.
- Ostad, S. A. (2013). *Strategier, strategiobservasjon og strategiopplæring: Med fokus på elever med matematikkvansker* (2. utg. rev. utg.). Trondheim: Læreboka forl.
- Pallant, J. (2013). *Spss survival manual : A step by step guide to data analysis using ibm spss* (5. utg.). Maidenhead: McGraw-Hill.
- Perry, M., Church, R. B. & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, 3(4), 359-400.
- Powell, S. R. (2014). The influence of symbols and equations on understanding mathematical equivalence. *Intervention in School and Clinic*, 50(5), 266-272.
- Prediger, S. (2010). How to develop mathematics-for-teaching and for understanding: The case of meanings of the equal sign. *Journal of Mathematics Teacher Education*, 13(1), 73-93.
- Reikerås, E. (2005). Skriftspråkvansker i norsk og matematikk: To sider av samme sak? I S. Skjong (red.), *Glsm: Grunnleggjande lese-, skrive- og matematikkopplæring* (s. 202-215). Oslo: Samlaget.
- Riley, M. S., Greeno, J. G., & Heller, J. H. (1983). Development of children's problem-solving ability in arithmetic. I H. Ginsburg (red.), *The development of mathematical thinking* (s. 153-196). Orlando, Florida: Academic Press.
- Rittle-Johnson, B. & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, 91(1), 175-189.
- Rosenthal, D. J. A., Resnick, L. B. & Williams, J. (1974). Children's solution processes in arithmetic word problems. *Journal of Educational Psychology*, 66(6), 817-825.
- Siegler, R. S. & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, N.J: Erlbaum.

- Sivesind, K. & Elstad, E. (2010). Oecd setter dagsorden *Pisa : Sannheten om skolen?* (s. 20-37). Oslo: Universitetsforlag.
- Sjøberg, S. (1998). *Naturfag som allmenndannelse : En kritisk fagdidaktikk*. Oslo: Ad notam Gyldendal.
- Sjøberg, S. (2007). Internasjonale undersøkelser: Grunnlaget for norsk utdanningspolitikk? I H. Hølleland (red.), *På vei mot kunnskapsløftet. Begrunnelser, løsninger og utfordringer* (s. 112-134). Valdres Trykkeri Fagernes: Cappelen Akademisk.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics teaching*, 77, 20-26.
- Stephens, A. C., Isler, I., Knuth, E. J., Blanton, M. L., Gardiner, A. M. & Marum, T. (2013). Equation structure and the meaning of the equal sign: The impact of task selection in eliciting elementary students' understandings. *Journal of Mathematical Behavior*, 32(2), 173-182.
- Utdanningsdirektoratet. (2012). *Rammeverk for grunnleggende ferdigheter til bruk for læreplangrupper oppnevnt av utdanningsdirektoratet*. Lastet ned fra [http://www.udir.no/Upload/larerplaner/lareplangrupper/RAMMEVERK\\_grf\\_2012.pdf?epslanguage=no](http://www.udir.no/Upload/larerplaner/lareplangrupper/RAMMEVERK_grf_2012.pdf?epslanguage=no).
- Utdanningsdirektoratet. (2013). *Læreplan i matematikk fellesfag*. <http://www.udir.no:Utdanningsdirektoratet> Lastet ned fra <http://data.udir.no/k106/MAT1-04.pdf?lang=nno>.
- Utdanningsdirektoratet. (2014a). *Kartleggingsprøve i regning for 2. Trinn - instruksjon til gjennomføring og vurderingsveiledning*.
- Utdanningsdirektoratet. (2014b). *Kartleggingsprøve i regning for 2. Trinn. Veiledning til lærere 2014*. <http://www.udir.no>: Utdanningsdirektoratet Lastet ned fra [http://www.udir.no/PageFiles/82502/140904\\_veiledning\\_2trinn\\_BM.pdf](http://www.udir.no/PageFiles/82502/140904_veiledning_2trinn_BM.pdf).
- Van De Walle, J. A. (2007). Developing understanding in mathematics. I J. A. Van de Walle, K. S. Karp, J. M. Bay-Williams & J. Wray (red.), *Elementary and middle school mathematics: Teaching developmentally* (s. 22-36). Boston: Alley and Bacon.
- Verschaffel, L. & De Corte, E. (1993). A decade of research on word problem solving in leuven: Theoretical, methodological, and practical outcomes. *Educational Psychology Review*, 5(3), 239-256.
- Weaver, J. F. (1973). The symmetric property of the equality relation and young children's ability to solve open addition and subtraction sentences. *Journal for Research in Mathematics Education*, 4(1), 45-56.
- Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29(3/4), 350-362.
- Zernichow, A. G. & Nygaard, O. (2006). Den blokkerende misoppfatning. *Spesialpedagogikk : tema: Matematikkmestring*, 4.

## **Vedlegg**

- I. Stavangerprosjektet
- II. Progresjonsbeskrivelsen for regning er satt opp i matrise.
- III. Scatterplot for poengsum på Del1 og Del2

Stavangerprosjektet har som mål å bidra til større kunnskap om barns utvikling fra 2,5 år til 10 år. Stavangerprosjektet er organisert som et programområde med fire prosjekter: Barnehagebarns utvikling; Skolebarns utvikling; Barn med observerte risikofaktorer; og Flerspråklige barn. Denne studien går inn under Skolebarns utvikling. Stavangerprosjektet har to hovedfaser, barnehagealder og skolealder. Innsamling av data fra barnehagealder ble avsluttet i 2012 og innhenting av data fra skolealder pågår ennå. Det er de samme barna som følges i begge fasene. Deltakelse i prosjektet er basert på foresattes skriftlige samtykke og de står fritt til å trekke sine barn fra prosjektet når de måtte ønske. Det er drøyt 1350 barn påmeldt til prosjektet, av disse er det 1130 som går på skole i Stavanger og fremdeles deltar i prosjektet på skolenivå. Siden prosjektets oppstart er det to elever som har trukket seg, resten av frafallet er grunnet fraflytting.



**Å kunne regne som grunnleggende ferdighet**

|  | Nivå 1  | Nivå 2  | Nivå 3   | Nivå 4  | Nivå 5  |
|--|---|---|--|---|---|
| <b>Ferdighets-områder</b><br><b>Gjenkjenne og beskrive</b> | Gjenkjenner konkrete situasjoner som kan løses ved regning, og formulerer spørsmål til dem.   | Gjenkjenner mønstre og formulerer problemer som kan løses i ett trinn. Analyserer innholdet i tekster og situasjoner der regning inngår.  | Analyserer tekster og situasjoner som forutsetter matematisk problemløsning i flere trinn. Forenkler situasjonen eller problemet slik at det blir håndterbart. | Analyserer mer sammensatte matematiske problemstillinger i dagligdage og faglige sammenhenger. Identifiserer størrelser som varierer.   | Analyserer et vidt spekter av matematiske problemstillinger som kan beskrives med en modell. Omformer og formulerer modellen for videre arbeid. |
| <b>Bruke og bearbeide</b>                                  | Bruker enkle strategier for optelling og klassifisering av mengder og geometriske former. Utfører enkle beregninger på ulike måter. | Velger hensiktsmessig regneart og bruker ulike metoder for å finne svaret. Bruker blant annet geometriske former, måleenheter, tabeller og grafiske framstillinger i prosessen. | Sammenlikner størrelser og uttrykker sammenhengen mellom dem. Velger hensiktsmessige måleenheter og gjennomfører egne undersøkelser.                           | Vurderer sjånse og utfører enkle statistiske beregninger. Bruker symboler for å uttrykke størrelser som er ukjente eller varierer. Kobler sammensatte problemstillinger til kjente løsningsmetoder. | Bruker et variert utvalg problemløsningsstrategier og kan begrunne metodevalg. Uttrykker sammenhenger med ord og bokstavuttrykk.                |
| <b>Kommunisere</b>   | Bruker ulike virkemidler for å uttrykke enkle beregninger.  | Beskriver resultatene ved blant annet å bruke geometriske former, måleenheter, tabeller og grafiske framstillinger.   | Presenterer resultater fra regneprosesser på en egnet måte ut fra problemstillingen.   | Setter sammen ulike måter å presentere resultater fra regneprosesser på.  | Presenterer resultater fra regneprosesser i tekster i ulike fag og i egen hverdag.  |
| <b>Reflektere og vurdere</b>                               | Avgjør om et resultat er svar på spørsmålet som ble stilt.  | Vurderer om resultatet er rimelig og tar beslutninger ut fra foreliggende fakta.  | Vurderer prosessen og vurderer om andre fremgangsmåter er mer effektive og enklere å kommunisere.  | Vurderer betydningen resultatet har for den situasjonen det skal brukes i. Drøfter feltkilder og gjør eventuelle justeringer.   | Sammenlikner ulike modeller og vurderer dem i lys av forholdet de beskriver.  |

Scatterplott med poengsum for del 1 og for del 2 på *Kartleggingsprøven i regning for 2. trinn* sett opp imot totalsum, den vertikale linjen er satt på bekymringsgrensen.

