Finnes det én oppskrift på god undervisningskvalitet i matematikk?

En kvantitativ undersøkelse av sammenhengen mellom elevers opplevelse av god undervisning i matematikk og hvordan de presterer i faget

Marie Vollen

Masteroppgave i matematikkdidaktikk
ved det Utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

3. juni 2019

Finnes det én oppskrift på god undervisningskvalitet i matematikk?

En kvantitativ undersøkelse av sammenhengen mellom elevers opplevelse av god undervisning i matematikk og hvordan de presterer i faget.

© Marie Vollen

2019

Finnes det én oppskrift på god undervisningskvalitet i matematikk?

Marie Vollen

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

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Denne oppgaven er en kvantitativ undersøkelse hvor fokuset er på hvordan elever opplever det som skal kjennetegner god undervisningskvalitet i matematikk, og hvilke sammenhenger dette har med hvordan de presterer i matematikkfaget. Jeg ønsket å undersøke hvilken effekt god undervisningskvalitet faktisk har på elevenes prestasjoner i matematikk. Problemstillingen så derfor på hvilke sammenhenger det er mellom elevenes syn på god undervisningskvalitet og deres prestasjoner. Finnes det en god oppskrift på hvordan undervisningen burde være for at elevene skal prestere godt?

Utvalget i denne studien er hentet fra et tidligere forskningsstudie, "Linking instruction and student achievement (LISA)". Det er totalt 46 deltakende matematikklasser med til sammen 991 elever. Datamaterialet består av et spørreskjema og klasseresultater på to nasjonale prøver, en fra elevene gikk på 8. trinn og en fra de gikk på 9. trinn. På bakgrunn av hvilket datamateriale jeg fikk tilgang til, valgte jeg i hovedsak å fokusere på tre undergrupper; høyt presterende vs. lavt presterende klasser, stor fremgang vs. liten fremgang i prestasjoner, stor spredning vs. liten spredning i prestasjoner. En av mine hypoteser var at høyt presterende klasser ville være mer fornøyde med lærerens undervisning i forhold til de lavt presterende klassene. Tilsvarende for de klassene med stor fremgang, at de var mer fornøyde enn de med liten fremgang. Alle analysene i studien har blitt utført i programvaren SPSS, gjennom en deskriptiv analyse. Mine funn antyder at alle disse 991 elevene i stor grad var fornøyde med lærerens undervisning. Det var ingen tydelige sammenhenger mellom hvor fornøyde de var med undervisningen og hvor godt en klasse presterte. Likevel fant jeg enkelte indikasjoner på at forskjellige grupper av klasser fremhever ulike aspekter av undervisningskvalitet. Faktisk var de lavt presterende klassene i litt større grad mer fornøyde med lærerens undervisning, enn de høyt presterende. Et usikkert moment i denne studien er hvordan elevene oppfattet spørsmålene som ble stilt i spørreskjemaet, og om de hadde samme oppfatning om god undervisning som hva teorien beskriver.

Forord

Denne oppgaven er skrevet som en avsluttende del av lektorutdanningen ved Universitetet i Oslo. Forskningsfeltet er knyttet til matematikkdidaktikk. Prosessen med å skrive denne oppgaven har vært lang og lærerik. Jeg har lært hvor viktig det er å være strukturert, tålmodig og målrettet. Det har vært spennende og interessant å få mer kunnskap om god undervisningskvalitet i matematikk. Samtidig som jeg har fått et større bilde av hvordan gode prestasjoner og hvordan elever opplever undervisningen, henger sammen.

Jeg vil først og fremst takke veilederen min, Nils Buchholtz, for ditt engasjement rundt oppgaven og alle gode råd. Takk for at du alltid satte av god tid til veiledning slik at vi fikk diskutert interessante funn sammen. Vil også takke det matematiske institutt for tilgang til egen leseplass.

Jeg vil takke familie og venner for god støtte, oppmuntrende ord og korrekturlesing. I en så tidkrevende prosess har jeg kjent godt på hvor viktig det er å ha en god familie og gode venner for å holde motivasjonen oppe. Til slutt vil jeg min kjære Martin, for at du klarer å holde humøret og motet mitt oppe. Du er en viktig støttespiller og jeg setter stor pris på de gode rådene dine.

Nå er jeg klar for arbeidslivet og nye utfordringer.

Marie Vollen,

Oslo, juni 2019

Innholdsfortegnelse

1 Innledning 1
1.1 Forskningsspørsmål 1
1.2 Linking instruction and student achievement 2
1.3 Oppgavens oppbygging 3
2 Teori 4
2.1 God undervisningskvalitet i matematikk 4
2.2 Hva vet vi om god undervisningskvalitet i matematikk i dag? 4
2.2.1 Hva er god klasseromsledelse i matematikk og hvordan påvirker dette elevens læring? 7
2.2.2 Hva er støttende undervisning i matematikk og hvordan påvirker dette elevenes læring? 10
2.2.3 Hva er kognitiv aktivering i matematikk og hvordan påvirker dette elevenes læring? 12
2.3 Hensikten med nasjonale prøver i regning 13
2.3.1 Hva kan resultatene fra nasjonale prøver i regning fortelle oss? 15
2.4 Empiriske funn fra tidligere studier 16
3 Metode 20
3.1 Valg av metode 20
3.2 Datamaterialet 20
3.2.1 Utvalg 21
3.2.2 Nasjonale prøver i regning 22
3.2.3 Spørreskjemaet 22
3.3 Analysestrategi 23
3.3.1 Konfirmerende faktoranalyse i Mplus 24
3.3.2 Deskriptiv analyse i SPSS 26
3.3.3 Multippel regresjonsanalyse i SPSS 26
3.4 Validitet 26
3.5 Forskningsetikk 28
4 Analyse av resultater 29
4.1 Resultater fra den konfirmerende faktoranalysen 29
4.1.1 Statistiske modelltester 31
4.1.2 Deskriptiv analyse av de syv aspektene for undervisningskvalitet 32

4.1.3 Studiens resultater i forhold til andre empiriske resultater 33
4.2 Resultater fra de deskriptive analysene på klassenivå 34
4.2.1 Høyt presterende vs. lavpresterende klasser 35
4.2.2 Stor fremgang vs. liten fremgang 40
4.2.3 Stor spredning vs. liten spredning i prestasjoner 42
4.3 Resultater fra den multiple regresjonsanalysen 44
4.3.1 Regresjonsanalyse av de høyt og lavt presterende klassene 46
4.3.2 Regresjonsanalyse av de med størst og minst fremgang i prestasjoner 47
4.3.3 Regresjonsanalyse av de med størst og minst spredning i prestasjoner 49
4.3.4 Oppsummering av resultatene fra den multiple regresjonsanalysen 50
5 Drøfting av forskningsfunn 51
5.1 Måling av undervisningskvalitet 51
5.2 Sammenhenger mellom undervisningskvalitet og prestasjoner 53
5.2.1 Høyt presterende klasser vs. lavt presterende klasser 54
5.2.2 Stor fremgang vs. liten fremgang på regneprøven 59
5.2.3 Stor spredning vs. liten spredning i prestasjoner i klassene 61
5.3 Oppsummering av de tre *casene* 63
5.4 Sammenligning av mine funn og tidligere empiriske funn 66
6 Avslutning 68
6.1 Svar på problemstillingen 68
6.2 Videre profesjonsutvikling 70
6.3 Forslag til videre forskning 70
Litteraturliste 72

1 Innledning

Flere forskere påpeker at lærerens undervisning er den enkeltfaktoren som har størst påvirkning på elevens læring (Kane og Cantrell, 2009; Klette, 2013; Nilsen et al., 2016), og at det er en sammenheng mellom elevenes prestasjoner i matematikk og hvor god undervisning de får (Blömeke et al., 2016). Vi vet at god undervisning skal ha en positiv påvirkning på elevenes prestasjoner, men hva vil egentlig en god undervisning innebære? Som nyutdannet lærer er det spennende å utdype seg i et tema som burde være interessant for alle matematikklærere. Finnes det én oppskrift på god undervisning som skal føre til bedre prestasjoner hos elevene? At det finnes én god oppskrift på hva som kjennetegner god undervisningskvalitet i matematikk? Og om det er slik at god undervisningskvalitet alltid bidrar til at elevene presterer bedre?

1.1 Forskningsspørsmål

Det finnes mye teori om hva som kjennetegner god kvalitet på undervisningen i matematikk og hva det er viktig at læreren gjør og ikke gjør. Men det er lite teori på hvordan elevene opplever undervisningen til læreren. Klette (2013) mener at god undervisning ikke handler om hvem læreren er, men mer om hva læreren gjør. Jeg ønsket å undersøke undervisningskvaliteten i matematikk fra elevenes perspektiv og kom frem til følgende problemstilling:

> Hvilke sammenhenger er det mellom elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget?

Formålet med denne problemstillingen er å få en dypere forståelse og et større innblikk i hva teorien beskriver som god undervisningskvalitet og i hvor stor grad elevene opplever dette i praksis. Jeg ønsket å se etter eventuelle sammenhenger mellom disse elevoppfatningene og hvordan elevene faktisk presterer i matematikk. I denne studien fokuserer jeg derfor hovedsakelig på tre typer av sammenligninger av elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget. Disse tre typene av sammenligninger vil videre i oppgaven omtales som tre *caser*.

- Hvordan opplever klasser som presterer høyt lærerens undervisning, sammenlignet med de som presterer lavt?

- Hvordan opplever klasser som har stor fremgang på den nasjonale prøven i regning fra 8. til 9. trinn, lærerens undervisning, sammenlignet med de med liten fremgang?

- Hvordan opplever klasser med stor variasjon av mestringsnivåer i klassen, lærerens undervisning, sammenlignet med de med liten variasjon av mestringsnivåer?

Ved å undersøke disse forskningsspørsmålene vil jeg få en større forståelse av sammenhengen mellom elevenes prestasjoner og hvordan de opplever undervisningen. I tillegg til disse tre spørsmålene ønsker jeg også å se nærmere på;

- Hvordan kan elevvurderinger av læreren brukes for å tolke undervisningskvaliteten?

- Hvordan vurderer norske elver lærerens undervisningskvalitet i matematikk i forhold til amerikanske elever i tidligere empiriske studier?

1.2 Linking instruction and student achievement

Datamaterialet til denne oppgaven er hentet fra den eksisterende studien, "Linking Instruction and Student Achievement», som videre forkortes til LISA-studien. Denne studien ble utført i perioden 2012 og 2013. Formålet med studien var å undersøke sammenhengen mellom undervisning og elevprestasjoner i matematikk og lesing på ungdomstrinnet. Denne studien er en videostudie hvor totalt 49 norske klasserom på 8. trinn har blitt filmet i to klassetimer i både matematikk og norsk, hvor 46 av dem deltok i matematikk. De deltakende elevene har også utført to nasjonale prøver, en på 8. trinn og en på 9. trinn. Dette var for å få en oversikt over elevenes fremgang i de to fagene. I tillegg til at alle klassene har blitt filmet og utført to nasjonale prøver, har de deltakende elevene svart på et spørreskjema mens de gikk på 9. trinn, som handler om sammenhengen mellom undervisning og elevprestasjoner i matematikk og lesing.

1.3 Oppgavens oppbygging

I kapittel 2 presenteres teoretiske perspektiver rundt god undervisningskvalitet i matematikk. Deretter hensikten med nasjonale prøver og hva disse resultatene kan fortelle. Det vil også bli redegjort for forskningsfunn fra tidligere empiriske studier av det aktuelle temaet.

I kapittel 3 presenteres valget av metode for denne oppgaven. Videre blir oppgavens datamateriale presentert og vurdert. Her kommer det også en oversikt over alle analysestrategiene som er blitt brukt for å analysere resultatene. Til slutt trekkes frem oppgavens validitet og forskningsetikk.

I kapittel 4 presenteres og analyseres oppgavens resultater. Først presenteres resultatene fra den konfirmerende faktoranalysen. Deretter flere deskriptive analyser av de tre *casene* jeg ønsket å undersøke. Til slutt resultatene fra en multippel regresjonsanalyse.

I kapittel 5 drøftes mine funn opp mot teori. Først drøftes funnene fra den konfirmerende faktoranalysen. Deretter drøftes funnene fra de tre *casene*. Til slutt drøftes mine funn opp mot tidligere empiriske forskningsfunn.

I kapittel 6 kommer en kort oppsummering av mine funn, hva jeg har lært fra denne oppgaven og forslag til videre forskning.

2 Teori

I dette kapittelet vil aktuell teori bli presentert. Først trekker jeg frem relevant teori som beskriver hva som kjennetegner og hva vi i dag vet om god undervisningskvalitet i matematikk. Videre trekkes inn hensikten med nasjonale prøver i regning og hva disse resultatene kan fortelle. Til slutt beskrives empiriske studier som har undersøkt ulike problemstillinger knyttet til måling av undervisningskvalitet i matematikk og deres funn.

2.1 God undervisningskvalitet i matematikk

Over en lang periode har forskere innen utdanning forsket på hva som definerer god undervisning, og svaret varierer etter tilnærmingene fra de ulike forskningsprosjektene (Schlesinger og Jentsch, 2016). Variasjonen avhenger blant annet av ytre faktorer som sosiale, økonomiske og kulturelle forhold (Nosrati og Wæge, 2015). Å definere effektiv undervisning er ikke lett (Coe et al., 2014) og i flere av definisjonene som er kommet frem trekkes elevenes progresjon og utfall frem som sentrale faktorer for hva som kjennetegner god undervisning. Nilsen et al (2016) definerer undervisningskvalitet som en forklaring på egenskapene til lærerens undervisningspraksis som er kjent for å være positivt knyttet til elevenes utfall. God undervisning defineres også som det som fører til forbedret progresjon hos elevene (Coe et al., 2014). Det trekkes også frem at lærerens undervisning er den enkeltfaktoren som har størst påvirkning på elevenes læring (Kane og Cantrell, 2009; Klette, 2013; Nilsen, Gustafsson og Blömeke, 2016). Det er en relevans mellom lærerens undervisningskvalitet og elevenes prestasjoner (Blömeke, Olsen og Suhl, 2016).

2.2 Hva vet vi om god undervisningskvalitet i matematikk i dag?

Klette (2013) trekker frem følgende hypoteser om god undervisning: en undervisning som balanserer tilegnelsessituasjoner, utprøvingssituasjoner og konsolideringssituasjoner, samt en lærer som balanserer ulike former for støtte i klasserommet: emosjonell, organisatorisk og undervisningsrettet støtte. Tilegnelsessituasjoner handler om hvordan læreren underviser bort nytt fagstoff til elevene. Utprøvingssitasjoner handler om hvordan elevene arbeider for å beherske dette nye fagstoffet. Hvordan læreren følger opp enkeltelevene og tilpasser

undervisningen. Fokuset i slike situasjoner er at elevene skal prøve seg frem på egenhånd eller i grupper, mens læreren fungerer som en veileder. Konsolideringssituasjoner handler om at elevene skal sette ord på egen tenkning gjennom blant annet metakognisjon. Det handler om at elevene skal bli mer bevisste på egen læring, de må lære seg å planlegge hvordan de skal løse en oppgave og reflektere rundt egne valg. Det viser seg at når undervisningen fokuserer på å reflektere og knytte sammen eksisterende kunnskap med ny kunnskap så er dette med på å påvirke elevenes læring (Klette, 2013). Klette (2013) trekker også frem at studier fra flere norske klasserom har vist at det er noen undervisningsmetoder som forekommer oftere enn andre. Det har vist seg at det er mye mer bruk av utprøvingssituasjoner i undervisningen enn konsolideringssituasjoner. Det ideelle for god læring er en variasjon av disse tre metodene (Klette, 2013).

Videre trekker Klette (2013) frem hvor nødvendig det er at læreren balanserer emosjonell, organisatorisk og undervisningsrettet støtte i undervisningen. I en god undervisning er det viktig at læreren tilbyr støtte til elevene. Emosjonell støtte handler om at læreren viser elevene at de bli verdsatt og sett underveis i læringsarbeidet deres. Tilbakemeldinger fra læreren i form av ros er former for støtte, som kan være avgjørende for kvaliteten på elevenes læring. For elevene er det viktig at læreren skaper en god relasjon med dem. I følge Klette (2013) bidrar gode relasjoner til et bedre læringsklima, noe som viser seg også å være viktig for elevenes læring. Organisatorisk støtte handler mer om lærerens organisering av undervisningen og lærerens klasseledelse. Det vil si at læreren må ha klare prosedyrer i timen, etablere rutiner og regler i klassen og hva som forventes av elevene. Også hvordan læreren håndterer ulike situasjoner i klassen som uro, forstyrrelser og bytte av aktiviteter. (Klette, 2013). Forskning har vist at organisatorisk støtte som klasseledelse mer har en indirekte påvirkning på elevenes læring. Det vil si at klasseledelse er en faktor som er med på å påvirke elevenes læring, men ikke en faktor som alene kan relateres til elevenes læring (Klette, 2013). Undervisningsmessig støtte handler om hva læreren gjør i klasserommet. Det er flere faktorer som spiller sammen for at undervisningsmessig støtte skal bidra til læring: kvaliteten på elevenes oppgaver, hvor mye tid elevene får på en oppgave, kvaliteten på samtalene som skjer i klassen, hvordan læreren introduserer det nye fagstoffer og hvordan læreren bruker elevinnspill osv. (Klette, 2013).

Coe et al. (2014) trekker frem seks faktorer som er viktige for god undervisning i matematikk: pedagogisk innholdskunnskap, kvaliteten på instruksjonene, klasseromsklima, klasseledelse, lærerens *beliefs* og profesjonell oppførsel. Av disse seks er det kun fire av dem som viser seg å ha en påvirkning på elevenes læring i matematikk. Pedagogisk innholdskunnskap og kvaliteten på instruksjonene, er de to faktorene som ser ut til å ha størst påvirkning. Pedagogisk innholdskunnskap vil si at læreren må ha god forståelse av innholdet som skal undervises bort og læreren må forstå og kjenne til hvordan elevene tenker om innholdet, kunne gjenkjenne deres valg av metoder og deres misoppfatninger (Coe et al., 2014). God kvalitet på hvordan læreren presenterer fagstoffet vil si at læreren er flink til å gå gjennom tidligere kunnskap med elevene for så å knytte det opp mot ny kunnskap. Det er viktig å gi elevene god tid til å praktisere kunnskapen og for så gradvis introdusere elevene for ny kunnskap (Coe et al., 2014). De to neste faktorene, klasseromsklima og klasseledelse, viser å ha en moderat påvirkning på elevenes læring. Dette stemmer godt overens med hva Klette (2013) trekker frem om at klasseledelse indirekte påvirker elevenes læring. Et godt klasseromsklima vil si at det er et godt samspill mellom lærer og elever. Læreren har tydelige forventninger og krav til elevene. Det handler om å få frem suksess hos elevene (Coe et al., 2014). God klasseromsledelse vil si at læreren benytter tiden i klasserommet effektiv, kontrollerer og styrer elevenes oppførsel med klare regler. Disse faktorene er nødvendige for god læring (Coe et al., 2014). Mens Coe et al. (2014) og Klette (2013) skriver mer om generelle aspekter læreren må ha i matematikkundervisningen, skriver matematikksenteret om mer faglige aspekter ved læreren. God matematikkundervisning som skal fremme læring hos elevene er ifølge dem å skape en undersøkende matematikkundervisning, skape en relasjonell forståelse hos elevene, elevene må vise selvinnsikt og bevissthet, de må oppleve motivasjon og undervisningen må være tilpasset (Nosrati og Wæge, 2015). De trekker også frem at det ikke finnes en bestemt måte for å oppnå dette, men at fokuset burde generelt være mer på undersøkende matematikkundervisning fremfor en undervisning med mye fokus på regler og algoritmer som elevene må huske (Nosrati og Wæge, 2015). En undervisning som består av at elevene får en større rolle og læreren er lærende, vil føre til bedre utfall for elevgruppen (Hattie, 2009). I følge elevvurdering av læreren er det en lærer som utfordrer, en som har høye forventninger og oppmuntrer elevene til faget og som har en god forståelse for faget sitt, det som kjennetegner en høyt kvalifisert lærer (Hattie, 2009).

"The seven C´s"

Tidligere studier som har forsket på hva som kjennetegner god undervisningskvalitet i matematikk, bruker forskjellige aspekter som skal forklare og måle undervisningskvaliteten. De fleste studiene som er blitt utført i USA, har benyttet seg av Tripod Survey laget av Ronald F. Ferguson ved Universitetet i Harvard. Tripod Survey er et spørreskjema som består av syv aspekter som sammen skal beskrive god undervisningskvalitet. Dette spørreskjemaet kalles ofte for "the seven C´s", fordi alle de syv aspektene starter med forbokstaven "C". Spørreskjemaet er blitt benyttet i det amerikanske prosjektet, "Measures of effective teaching", ofte kalt for MET, som er stiftet av Bill og Melinda Gates foundation. Også LISA-studien har valgt å bruke dette spørreskjemaet for å se nærmere på sammenhengen mellom undervisningskvaliteten i matematikk og norsk opp mot elevenes prestasjoner. De syv aspektene er; *care, control, clarify, challenge, captivate, confer* og *consolidate*. I europeiske studier er det særlig tre aspekter som er blitt benyttet for å undersøke undervisningskvaliteten; klasseromsledelse, støttende undervisning og kognitiv aktivering (Bergem et al., 2016; Nilsen et al., 2016; Schlesinger og Jentsch, 2016, Schlesinger et al., 2018).

Videre i oppgaven skal jeg beskrive hva som kjennetegner de tre aspektene; klasseromsledelse, støttende undervisning og kognitiv aktivering. Før jeg beskriver hva som kjennetegner disse tre aspektene har jeg valgt å trekke de syv aspektene som er brukt i LISA-studien inn i hver av disse tre, siden jeg opplever at de beskriver mye av det samme.

Tabell 2.1: De syv aspektene brukt i LISA-studien knyttet til tre mer overordna aspekter, som skal kjennetegne god undervisningskvalitet i matematikk.

Klasseromsledelse	Støttende undervisning	Kognitiv aktivering
Control	Care	Challenge
Consolidate	Confer	Captivate
Clarify		

2.2.1 Hva er god klasseromsledelse i matematikk og hvordan påvirker dette elevens læring?

Til dimensjonen klasseromsledelse har jeg valgt å trekke inn de tre aspektene *control*, *consolidate* og *clarify,* ettersom det disse står for inngår i hva jeg definerer som klasseromsledelse. For at elever skal lære, handler det om å skape et klasserom som er fritt for

forstyrrelser og som består av god orden og et godt miljø. En lærer med god klasseromsledelse har god evne til å holde god struktur og orden i klasserommet (Baumert et al., 2010; Bergem et al., 2016; Schlesinger og Jentsch, 2016; Schlesinger et al., 2018). God struktur og orden innebærer blant annet å ha en strukturert undervisning med strukturerte læringsaktiviteter (Ogden, 2012; Schlesinger et al., 2018). Det kan bety at undervisningen er tydelig og klar og består av aktiviteter som gir elevene et bredt spekter av læringsformer (Nilsen et al., 2016). God struktur og orden i undervisningen består av glidende overganger mellom de ulike læringsaktivitetene (Kunter et al., 2007). Til god klasseromsledelse kan vi trekke inn aspektet *control* av Ferguson (2012), som handler om at læreren skal ha god kontroll på klassen. Det vil si at læreren har ferdigheter til å håndtere klasserommet, forstyrrende atferder fra elevene, for å fremme et godt læringsmiljø hvor det er rom for klasseromsdiskusjoner (Ferguson, 2012). God kontroll i klasserommet bidrar til å gjøre klasserommet til et sted som oppleves som rolig og trygt sted å være for eleven, som føles som støttende og et sted hvor det er rom for samarbeid (Ferguson, 2012; Bergem et al., 2015). *Control* handler også om å håndtere oppførselen til elevene i undervisningen (Kuhfeld, 2017).

Det er viktig å skape god orden i klassen for at elevene skal lære best (Postholm, 2013). En god klasseromsleder klarer å utvikle og opprettholde en ordnet atmosfære i klasserommet hvor læreren klarer å forebygge uro og forstyrrelser (Lipowsky et al., 2009). For å få til gode læringsprosesser i timen er det viktig at læreren har en produktiv og forebyggende klasseromsledelse hvor læreren håndterer forstyrrelser og konflikter som kan oppstå i timen (Baumert et al, 2010; Ogden, 2012; Schlesinger og Jentsch, 2016; Schlesinger et al, 2018). Læreren har som oppgave å styre elevenes oppførsel i timen (Coe et al., 2014) og en god klasseromsleder gjør dette ved å ha klare regler, prosedyrer og forventninger til elevene i undervisningen (Kunter et al., 2007; Lipowsky et al., 2009; Ogden, 2012; Coe et al., 2014; Schlesinger og Jentsch, 2016; Schlesinger et at., 2018). En god klasseromsleder som stiller realistiske forventninger til elevene fører til at elevene kan klarer å strekke seg litt lenger og prestere litt bedre (Ogden, 2012). Når elevene følger disse reglene, prosedyrene og forventningene til læreren, bidrar det til å forbedre miljøet i klassen, og et godt arbeidsmiljø er med på å forebygge uro i timen (Ogden, 2012). Gode klasseromledere klarer å skape et inkluderende læringsmiljø og opprettholder en god arbeidsro blant elevene, noe som igjen bidrar til mer effektiv læringstid hos elevene (Postholm, 2013).

Det er viktig å ha en effektiv tidsbruk i undervisningen for å gi mest effektiv læringstid hos elevene og for at elevene skal få størst læringsutbytte av aktivitetene (Kunter et al., 2007; Baumert et al., 2010; Coe et al, 2014; Bergem et al., 2016; Schlesinger og Jentsch, 2016; Schlesinger et al., 2018). I et klasserom hvor det er god klasseromsledelse og effektiv bruk av tid, kan elevene bruke mer tid på oppgavene og derfor ha større mulighet til å sette seg inn i stoffet og forstå innholdet (Lipowsky et al., 2009). For at elever skal lære mest mulig må elevene forstå innholdet, og da er det viktig å få tid og anledning til å bearbeide den nye kunnskapen (Ogden, 2012). For at læreren skal klare å få mest ut av undervisningen på best mulig tid, er det viktig at læreren planlegger og forbereder seg godt til undervisningen (Lipowsky et al., 2009). Det er viktig at læreren har en tydelig undervisning med tydelige intensjoner. Det vil si at læreren må ha klare læringsmål til elevene slik at de forstå hva som forventes av dem (Nilsen et al., 2016; Bergem et al., 2016).

Læreren må også klare å tilpasse undervingen ved å koble eksisterende kunnskap hos elevene opp med ny kunnskap (Bergem et al., 2016; Nilsen et al., 2016). Dette stemmer godt overens med hva aspektet *consolidate* trekker frem som viktig i en god undervisning. *Consolidate* handler om at læreren knytter ideer og materiale sammen. Med andre ord, knytter fagstoffet sammen (Ferguson, 2012). Hvordan læreren knytter fagstoffet sammen, gir elevene tilbakemeldinger og sjekker om de forstår (Kuhfeld, 2017). Det innebærer at læreren ser etter forståelse hos elevene og hjelper dem med å organisere den tillærte informasjonen for å bidra til mer effektiv resonnering (Ferguson, 2012). Det er også viktig å sette av tid til gjennomgang og oppsummering i slutten av timen for å integrere og knytte ideene sammen (Ferguson, 2012; Bergem et al., 2016; Nilsen et al., 2016). Når læreren legger ned arbeid i å planlegge undervisningen, vil det være lettere å gjøre elevene mer motiverte og oppmerksomme i timen ved at de kan få en mer aktiv rolle i timen (Ogden, 2012). Aspektet *clarify* fra Ferguson (2012) handler om at læreren klarer å skape en tydelig undervisning, som vil si at læreren klaree å fremme forståelse hos elevene ved blant annet å fjerne forvirring og misforståelser hos elevene (Ferguson, 2012). Det handler om at læreren er i stand til å hjelpe elevene til å forstå det de blir lært i undervisningen (Kuhfeld, 2017). Ved å klargjøre undervisningen vil det bidra til å øke elevenes utholdenhet i timen. Alle elever er forskjellige og alle lærer og forstår ulikt, og det er derfor viktig at læreren har en evne til å kunne klargjøre undervisningen ved å kunne klare å diagnostisere elevenes ferdigheter og kunnskaper, samt kunne variere metodene å forklare fagstoffet på (Ferguson, 2012).

Læreren må ha evnen til å kunne avgjøre og bedømme hvor mye informasjon elevene kan tilegne seg til enhver tid.

I tillegg til generelt god orden og effektiv tidsbruk, innebærer rollen som klasseromsleder å skape gode relasjoner med klassen. En forutsetning for at undervisningen skal fungere er at det er god relasjon mellom lærer og elever (Ogden, 2012). Ogden (2012) trekker også frem at elever lærer best i miljøer hvor de føler seg trygge på lærer og klassekamerater. "Et godt læringsmiljø legger grunnlaget for læring" (Postholm, 2013, s. 138). Som aspektet *clarify* trekker frem, er det viktig at læreren kjenner til elevgruppen for å kunne vurdere innholdet slik at det kan oppleves som interessant for elevene og at elevene føler mestring (Ferguson, 2012; Bergem et al., 2015). Flere studier trekker frem en god klasseromsleder sin påvirkning på elevenes læring. Lipowsky et al. (2009) fremhever at elevenes mulighet til å lære er avgjørende av kvaliteten på lærerens klasseromsledelse. God klasseromsledelse er en forutsetning for god kvalitet på undervisningen og elevenes læring (Baumert et al., 2010).

2.2.2 Hva er støttende undervisning i matematikk og hvordan påvirker dette elevenes læring?

Til dimensjonen støttende undervisning har jeg valgt å trekke inn to av Ferguson (2012) sine aspekter, *care* og *confer*, ettersom at disse beskriver hva jeg definerer som støttende undervisning. Støttende undervisning vil si at elevene opplever undervisningen som et sted hvor det er trygt å være og et sted hvor det er rom for å dele sine tanker og ideer. Det er derfor viktig at læreren fokuserer på å bygge et trygt og godt miljø i klassen (Schlesinger og Jentsch, 2016; Schlesinger et al., 2018). Et slikt støttende klassemiljø kan være med på å påvirke i hvilken grad elevene engasjerer seg i læringsarbeidet (Lipowsky et al., 2009). Aspektet *confer* handler om å skape et trygt og godt klassemiljø hvor det er rom for å snakke høyt foran klassen (Ferguson, 2012). For å klare å skape et slikt klassemiljø er det viktig med en støttende lærer som behandler alle elever likt. Elevene skal oppleve å bli sett, respektert og behandlet rettferdig av læreren (Bø og Hovdenak, 2011; Ferguson, 2012). Det handler i stor grad om at læreren skal få frem elevenes ideer (Kuhfeld, 2017). Aspektet *care*, handler om at læreren bryr seg, oppmuntrer og støtter elevene i undervisningen (Ferguson, 2012; Bergem et al., 2016). *Care* handler også om at læreren har et støttende forhold til elevene (Kuhfeld, 2017). I en undervisningssituasjon kan læreren både gi personlig støtte og mer faglig støtte til elevene (Bø og Hovdenak, 2011).

Personlig støtte handler om at læreren støtter hver enkelt elev og deres rolle i klasserommet (Lipowsky et al., 2009). Det vil si at læreren sørger for å se hver enkelt elev og behandle alle rettferdig. Elevene skal ikke føle at de må være noen andre enn seg selv i klasserommet og det er læreren sin rolle å sørge for at alle føler seg respektert i undervisningen (Bø og Hovdenak, 2011). Den personlige lærerstøtten går ut på å bry seg om elevene og vise dem det (Nilsen et al., 2016) og det handler om å skape et godt forhold og god relasjon mellom lærer og elev (Lipowsky et al., 2009; Schlesinger og Jentsch, 2016; Schlesinger et al., 2018). Aspektet *care* innebærer også lærerens rolle ovenfor elevene,. Det er viktig at læreren viser god omsorg ovenfor elevene for at de skal oppleve undervisningen som en positiv opplevelse (Ferguson, 2012). For å skape slike gode lærer-elev relasjoner, er det læreren sin oppgave å skape positive holdninger med eleven, slik at begge får positive oppfattelser og tydelige forventninger til hverandre (Drugli, 2012). Faglig støtte innebærer at læreren skal oppmuntre elevene i undervisningen og gi dem ekstra hjelp når det er nødvendig (Nilsen et al., 2016; Schlesinger et al., 2018). For å få elevene opp og frem i faget handler det blant annet om at læreren må gi elevene en støtte som kan være med på å bidra til faglig fremgang (Bø og Hovdenak, 2011). Det er derfor viktig at læreren gir elevene realistiske forventninger og sørger for å følge opp elevene underveis med støtte (Bø og Hovdenak, 2011). Faglig støtte handler også om de tilbakemeldingene læreren gir underveis i undervisningen og hvordan læreren håndterer elevenes feil og misoppfatninger (Lipowsky et al., 2009; Bergem et al., 2016; Schlesinger og Jentsch, 2016).

Det er viktig for elevene å oppleve at de får støtte fra læreren, både personlig og faglig. Elever som opplever at de har en god relasjon med læreren vil ofte engasjere seg litt ekstra i undervisningen (Bø og Hovdenak, 2011). Elever som opplever å bli respektert og anerkjent vil mest sannsynlig også anerkjenne læreren og tar lærerens undervisning mer på alvor (Drugli, 2012). Noen viktige forutsetninger for vellykket læring hos elevene er at de må oppleve undervisningen som et varmt og respektfullt læringsmiljø og en omsorgsfull atmosfære (Lipowsky et al., 2009). Læreren har en stor betydning i elevenes faglige utbytte og forholdet mellom lærer og elever viser å ha en stor påvirkning elevenes faglige interesse og motivasjon (Bø og Hovdenak, 2011; Schlesinger et al., 2018). Gjennom både personlig og faglig støtte vil læreren komme enda mer i kontakt med elevene og få bedre innsikt i hvor de befinner seg faglig. Det vil derfor være enklere å kunne tilrettelegge undervisningen og differensiere når læreren kjenner elevgruppen godt (Drugli, 2012; Schlesinger og Jentsch, 2016).

2.2.3 Hva er kognitiv aktivering i matematikk og hvordan påvirker dette elevenes læring?

Til kognitiv aktivering har jeg valgt å trekke inn Ferguson (2012) sine to aspekter *captivate* og *challenge*. Dette er nok den dimensjonen som kan knyttes mest faglig spesifikt til matematikk (Schlesinger og Jentsch, 2016). Kognitiv aktivering vil si at læreren må engasjere elevene mer i undervisningen. Det betyr at elever ikke kan lære alt ved at lærer fyller på med kunnskap fra utsiden (Schlesinger et al., 2018). Dette er noe som Lipowsky et al. (2009) trekker frem som lite sannsynlig for at elevene skal utvikle en dypere forståelse. Elevene kan engasjeres mer i undervisningen i form av at de aktiviseres med mer utfordrende oppgaver og blir mer oppfordret til å muntlig uttrykke seg om sine tanker og ideer (Schlesinger og Jentsch, 2016). Disse oppgavene burde, som aspektet *captivate* trekker frem, oppleves som interessante og relevante for elevene for å unngå at elevene skal oppleve aktivitetene som kjedelige (Ferguson, 2012; Bergem et al., 2015). Det innebærer også i hvilken grad læreren klarer å holde på elevenes oppmerksomhet i undervisningen (Kuhfeld, 2017). Aspektet *challenge* handler om at læreren skal utfordre elevene i undervisningen (Ferguson, 2012; Bergem et al., 2016). Læreren må oppfordre elevene til å engasjere seg på et høyere nivå av tenkning, noe som bidrar til at elevene utvikler en dypere kunnskap (Lipowsky et al., 2009). Utfordre elevene til å anvende eksisterende kunnskap til å forklare deres løsningsmetoder (Nilsen et al., 2016; Schlesinger og Jentsch, 2016; Schlesinger et al., 2018) og integrere denne kunnskapen i sammenheng med utfordrende oppgaver som problemløsningsoppgaver (Nilsen et al., 2016). Når elevene utfordres er det viktig at læreren krever god innsats fra elevene, noe som medfører at læreren må ha evnen til å være streng med elevene og konfrontere dem hvis deres innsats ikke er tilfredsstillende (Ferguson, 2012). Kognitiv aktivering innebærer også hvilke problemer som blir valgt og måten de blir brukt i undervisningen (Baumert et al., 2010). Det er oppgaver som er mer utfordrende for elevene hvor de må strekke seg det lille ekstra for å mestre oppgaven (Baumert et al., 2010; Nilsen et al., 2016; Schlesinger og Jentsch, 2016; Schlesinger et al., 2018). Oppgaver som stiller et høyere krav til elevene (Lipowsky et al., 2009). Det er oppgaver som ikke har en tydelig løsningsmetode, men oppgaver hvor elevene må gjøre en ekstra innsats for å forstå hva de skal gjøre (Nosrati og Wæge, 2015). Elevene kan også engasjeres i undervisningen i form av klasseromsdiskusjoner hvor læreren for eksempel kan bygge videre på elevenes ideer og få elevene til å forklare deres løsningsmetoder på en oppgave (Schlesinger og Jentsch, 2016). Elevene må resonnere og reflektere over innholdet og tenke kritisk gjennom oppgavene (Lipowsky et al., 2009;

Nosrati og Wæge, 2015). I slike klasseromsdiskusjoner er det viktig at læreren ikke bare utrykker overfor elevene om deres svar var ”riktig” eller ”galt”, men heller gir elevene en anledning til å vurdere deres egne og andre løsninger og prøve å finne ut flere løsningsstrategier (Baumert et al., 2010). Elevene må lære seg å arbeide selvstendig og tenke mer over egen tenkning, også kalt for metakognisjon (Nosrati og Wæge, 2015). Elevene må lære seg å bruke selvinnsikt og bli bevisste over egne tankeprosesser. Oppgaver som kan løses kun ved kjente prosedyrer lært fra læreren bidrar ikke til å øke elevenes selvinnsikt og bevissthet (Nosrati og Wæge, 2015).

Gjennom utfordringer er det meningen at elevene skal fremkalle en dypere og grundigere forståelse av faget (Ferguson, 2012). Ved bruk av kognitive aktiviteter og utfordringer i undervisningen legges det mer vekt på å danne en relasjonell forståelse hos elevene, som viser seg å ha en positiv effekt på elevenes prestasjoner i matematikk (Nosrati og Wæge, 2015). Relasjonell forståelse handler både om å forstå hvordan oppgaven skal løses og om å forstå hvorfor det er slik (Nosrati og Wæge, 2015). Læreren kan fremme elevenes relasjonelle forståelse ved å legge mer vekt på å forstå og se sammenhengen mellom matematiske ideer og prosedyrer. Læreren må få elevene til å reflektere og diskutere meningen bak de matematiske løsningsmetodene og valg av strategier, samt gjøre elevene bevisste over hvordan ting henger sammen (Nosrati og Wæge, 2015). Sannsynligheten for at elevenes dypere forståelse øker skjer når elevene reflekterer, ser etter sammenhenger mellom begreper og ideer, og knytter det nye innholdet opp mot eksisterende kunnskap (Lipowsky et al., 2009). Tilsvarende er sannsynligheten for at elevenes dypere forståelse minker når de kun blir bedt om å løse oppgaver etter metoder som er demonstrert av læreren (Lipowsky et al., 2009). Matematikksenteres trekker frem følgende for å fremheve viktigheten av kognitiv aktivering i undervisningen:

> *Når elevene får lov til å utforske et felt og diskutere hvordan de tenker med hverandre, oppdager de at matematikk slett ikke er et fag som kun består av å huske hva læreren har sagt. I stedet blir det et spennende og aktivt fag som består av utforskning på elevenes egne premisser. (Nosrati og Wæge, 2015, s. 12).*

2.3 Hensikten med nasjonale prøver i regning

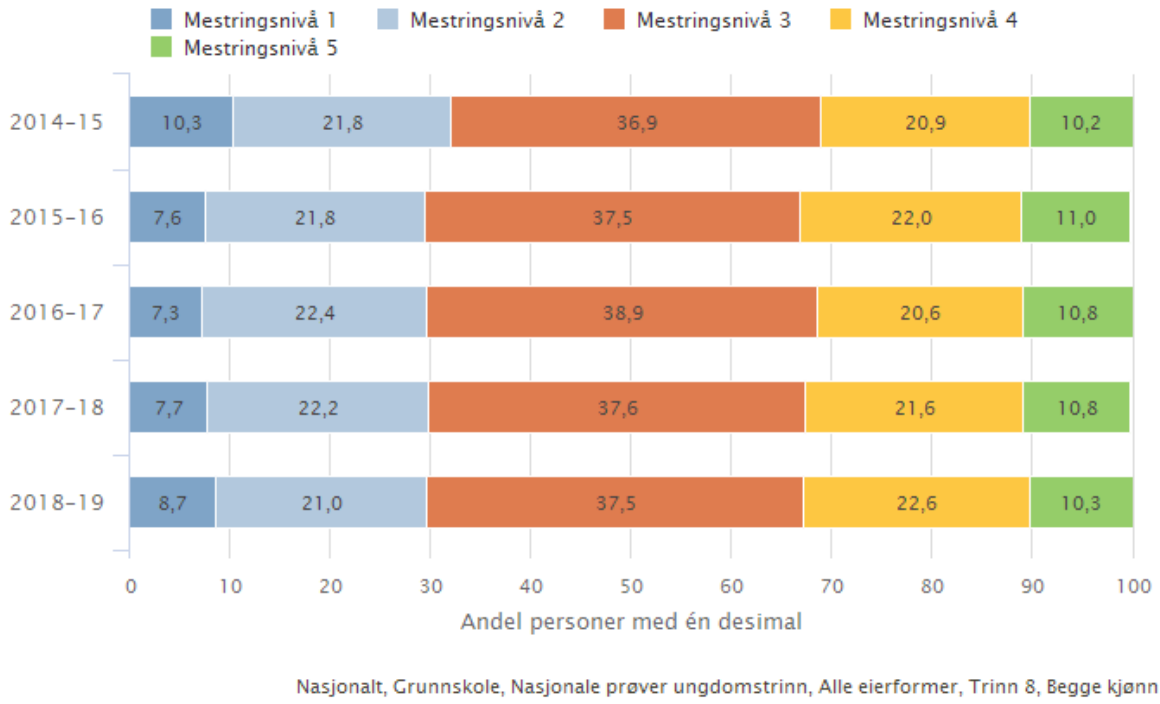
Nasjonale prøver i matematikk er prøver om grunnleggende ferdigheter i å kunne regne. Disse prøvene gjennomføres både på 5., 8. og 9. trinn over hele landet. Det er

Utdanningsdirektoratet som har i oppdrag å utføre disse prøvene nasjonalt, og det er matematikksenteret som har fått oppdraget av utdanningsdirektoratet til å utvikle disse prøvene (Nosrati og Wæge, 2015). Poenget med den nasjonale prøven er å kartlegge i hvilken grad elevenes ferdigheter samsvarer med kompetansemålene i Kunnskapsløftet for matematikk. Siden prøven som holdes utføres både på 8. og 9. trinn samtidig, tar prøven utgangspunkt i kompetansemålene elevene skal kunne etter 7. trinn i matematikk. Grunnleggende ferdigheter i regning innebærer tallforståelse, måleferdighet og tallbehandling knyttet til et bredt spekter av oppgaver og utfordringer i faglige og dagligdagse sammenhenger. Å kunne regne vil si at eleven må ha ferdigheter til å kunne gjenkjenne og beskrive situasjoner fra virkeligheten der matematikk er involvert. De må også kunne bruke og bearbeide matematiske begreper, prosedyrer, fakta og verktøy for å finne løsninger på problemer. Videre må de kunne reflektere over rimeligheten av egne svar og svaralternativer i flervalgsoppgaver og vurdere om dette er gode svar på de problemene som skal løses. Ved bruk av disse ferdighetene skal elevene klare å løse ulike oppgaver innenfor områdene tall og algebra, måling og geometri, og statistikk og sannsynlighet. Problemstillingene i oppgavene er situasjoner hvor elevene skal kunne kjenne seg igjen i (Nosrati og Wæge, 2015). Det er en lengre prosess å lage de nasjonale prøvene. Det er flere som er med på å samarbeide om å forme prøven, og det følges et rammeverk for hvordan prøvene skal se ut og hva de skal inneholde (Utdanningsdirektoratet, 2018). Oppgavene som blir gitt på de nasjonale prøvene, blir på forhånd prøvd ut flere ganger for å kvalitetssikre oppgavene og prøvens hensikt (Utdanningsdirektoratet, 2018).

Hensikten med nasjonale prøver er å få informasjon om enkeltelever, grupper, trinn og skoler som er nødvendig for lærere og skoleledere for å utvikle skolen faglig sterkere. Lærerne kan bruke resultatene til å følge opp elevene gjennom underveisvurdering og tilpasset opplæring. Kommuner og skoler kan bruke resultatene som grunnlag for kvalitetsutvikling av opplæringen. Forskere kan bruke resultatene til videre forskning om matematikk i skolen (Utdanningsdirektoratet, 2018). Det skal sies at disse prøvene ikke gir et helhetlig bilde av elevenes ferdigheter og forståelser. Resultatene fra de nasjonale prøvene gir bare et begrenset bilde av elevenes kompetanse og ferdigheter. Derfor er det viktig å se på resultatene i sammenheng med annen relevant informasjon (Utdanningsdirektoratet, 2018). At disse prøvene har blitt utført over flere år bidrar til at man får informasjon om utviklingen som skjer i faget.

2.3.1 Hva kan resultatene fra nasjonale prøver i regning fortelle oss?

I dag kan vi sammenligne resultatene for regning fra året 2014 til året 2019. Resultatene kan fortelle oss at det er små endringer i elevenes prestasjoner over denne tidsperioden. LISA-studien ble utført i 2012 og 2013, som betyr at vi ikke får sett resultatene fra dette året på Utdanningsdirektoratet sin nettside. Elevene kan befinne seg på et mestringsnivå fra 1 til 5.



Figur 2.1: En oversikt fordelingen av mestringsnivåene på de nasjonale prøvene som har blitt utført fra 2014 til 2019 (Utdanningsdirektoratet, 2019).

Figur 2.1 viser resultatene på 8. trinn fra perioden 2014 til 2018. Den viser at det er en minimal endring i elevenes prestasjoner. Altså at det er liten endring i hvor stor andel av elever på 8. trinn som befinner seg på de ulike mestringsnivåene i regning (Utdanningsdirektoratet, 2018). Det er flere elever som befinner seg på mestringsnivå 4 og 5 og færre elever på mestringsnivå 1 og 2. Disse resultatene beskriver Utdanningsdirektoratet (2017) som forventet, ettersom at elevene på 9. trinn er eldre og har gått ett år lenger på skolen.

Det som kjennetegner en elev som befinner seg på mestringsnivå 1 er at eleven klarer å gjenkjenne konkrete situasjoner som man klarer å løse ved bruk av enkle matematiske strategier. En slik elev har fått en skalapoeng fra 0 til 36. En elev på mestringsnivå 2 klarer å

velge den mest hensiktsmessige regnearten og klarer å bruke ulike metoder for å løse oppgaver som krever flere enn ett trinn. En slik elev har fått en skalapoeng fra 37 til 44. En elev på mestringsnivå 3, klarer å løse enkle sammensatte problemstillinger der tallene er enkle å regne med. En slik elev har fått en skalapoeng fra 45 til 54. En elev på mestringsnivå 4 klarer å se sammenhenger mellom sammensatte problemer og klarer å benytte kjente løsningsmetoder for å løse disse. En slik elev har fått en skalapoeng fra 55 til 62. En elev på mestringsnivå 5 mestrer flere problemløsningsstrategier og klarer å begrunne valg av metode. En slik elev klarer også å finne løsninger på kognitivt krevende og utfordrende oppgaver med tall. En slik elev har fått en skalapoeng fra 63 og høyere på den nasjonale prøven i regning (Utdanningsdirektoratet, 2017).

2.4 Empiriske funn fra tidligere studier

Det ser ut til at elevenes vurdering av lærerens undervisningskvalitet er relatert til elevenes prestasjoner, selv om disse tilbakemeldingene fra elevene sjeldent fører til forbedret undervisning av lærerne (Hattie, 2009). Som nevnt tidligere i kapittelet, baseres flere tidligere forskninger på undervisningskvalitet i matematikk seg på MET-prosjektet. Dette er hovedsakelig et prosjekt for å teste nye tilnærminger for å måle effektiv undervisning i blant annet matematikk (Kane og Cantrell, 2009). Flere forskere har bekreftet at det er lærerens evne til å undervise som elevenes akademiske utvikling avhenger av (Kane og Cantrell, 2009). Målet til MET-prosjektet var derfor å forbedre kvaliteten på den informasjonen vi kjenner til om effektiv undervisning for å skape flere effektive lærere. I denne oppgaven er det tre amerikanske studier som trekkes frem. Disse er de empiriske studiene Kuhfeld (2017), Wallace et al. (2016) og Schweig (2016) som alle har benyttet seg av MET-data i ulik tilknytning for å måle undervisningskvaliteten i matematikk. I tillegg til disse tre studiene, trekkes også frem en norsk studie fra 2015.

Først ut er Kuhfeld (2017), som er en empirisk studie fra Amerika som forsket på validiteten av å bruke elevers vurderinger av læreren til måle undervisningskvaliteten, med utgangspunkt i spørreskjemaet til Ferguson. Hun ønsket å se nærmere på hvor godt hver av de syv aspektene faktisk måler undervisningskvaliteten og hvilke biaser som må tas hensyn til for at resultatene skal være troverdige og pålitelige (Kuhfeld, 2017). En bias eksiterer når vurderingen i et spørreskjema blir påvirket, enten positivt eller negativt, av irrelevante faktorer (Kuhfeld, 2017). For eksempel kan en elev velge å vurdere undervisningskvaliteten

til læreren mer positivt, i håp om at læreren vil belønne eleven for god vurdering. Eller at en elev vurderer lærerens undervisning mer negativt, på grunn av dårlig relasjon med læreren. Gjennom en eksplorerende faktoranalyse fant hun tre ulike modeller for å måle undervisningskvaliteten. Den ene faktoranalysen foreslo at aspektet *control* var en separat faktor fra de andre seks aspektene som ble slått sammen til et felles aspekt. Videre undersøkte hun en todimensjonal modell som bestod av *support* og *control*. Den siste modellen bestod av den originale versjonen med alle de syv aspektene. Hun fant ut at alle disse modellene viste en tilsvarende påvirkning på elevenes prestasjoner, men særlig aspektet *control* hadde en signifikant påvirkning (Kuhfeld, 2017). Disse funnene støtter påliteligheten i å bruke Ferguson (2010) sin spørreundersøkelse blant elever for å vurdere lærerens undervisningskvalitet i matematikk (Kuhfeld, 2017). Hun trekker også frem viktigheten av å ta hensyn til eventuelle biaser som elevene kan ha når de vurderer lærerens undervisningskvalitet. Som nevnt kan elever ha vurdert læreren mer positivt eller negativt på grunn av personlige forhold til læreren. Det kan også være en mulighet for at elever som ikke har hatt en konkret mening til påstandene som blir stilt, velger å svare den midterste verdien (Kuhfeld, 2017). Hun trekke frem at gyldigheten av en studie bestående av dette spørreskjemaet hadde økt dersom datainnsamlingen av spørreskjemaet hadde foregått over flere anledninger i løpet av et skoleår, ettersom at det ville bidratt til å gjøre en grundigere diskusjon av reliabiliteten av poengene fra spørreundersøkelsen (Kuhfeld, 2017).

Studien til Wallace et al. (2016) undersøkte om de begrepene som finnes om effektiv undervisning fra teorien, er reflektert i elevenes svar på spørreundersøkelsen til Ferguson (2010). Denne studien var opptatt av å undersøke hvor godt elevundersøkelser faktisk måler undervisningskvaliteten til læreren, hvor godt hver av spørsmålene måler det de skal og om elevene har samme oppfatning som forskerne bak teorien har om god undervisningskvalitet (Wallace et al., 2016). De benyttet seg av tre ulike faktormodeller som alle skulle beskrive undervisningskvaliteten til læreren. En bifaktormodell hvor de syv aspektene ble fordelt på to faktorer; *support* og *press*, en trifaktormodell bestående av undervisningsrettet støtte, emosjonell støtte og klasseromsledelse, og en bifaktormodell bestående av alle de syv aspektene hver for seg. Dette gjorde de for å se hva som best kjennetegner god undervisningskvalitet i samsvar med elevenes svar på spørsmålene i spørreskjemaet. Wallace et al. (2016) trekker frem at dette spørreskjemaet er en god undersøkelse som trekker frem flere aspekter av undervisningskvaliteten. Deres funn var at elevene fungerer godt til å vurdere lærerens undervisningskvalitet (Wallace et al., 2016).

Schweig (2016) forsket på den generelle enigheten blant elevene – konsensus - i hvordan de vurderte undervisningskvaliteten til læreren. Om dette kunne gi en forståelse av læringsmiljøet i klassen, og om dette potensielt kunne brukes for å forstå fremtidig undervisningskvalitet. Han mente at det å bruke elevundersøkelser for å undersøke lærerens undervisningskvalitet kan være med på å gi verdifull informasjon om undervisningen (Schweig, 2016). Å kjenne til de generelle enighetene i klassen vil derfor være viktig informasjon om undervisningskvaliteten til læreren og læringsmiljøet. Resultatene fra spørreskjemaet til Ferguson (2010) gir et gjennomsnittlig svar fra de ulike klassene, og et gjennomsnitt kan bli det samme i tre ulike grader av enighet innad i klassene. Schweig (2016) fant dermed ut at konsensusen i klassen har direkte implikasjoner for hvordan elevundersøkelser blir brukt i evaluering av undervisningen. Han kom derfor frem til at det er viktig å se nærmere på hvor stor grad av enighet det er innad i klassene og ikke bare på den gjennomsnittlige verdien av undervisningskvaliteten. Han fant ut at i de klassene som hadde høyere konsensus om lærerens undervisning viste til å oppleve mer effektiv undervisning (Schweig, 2016).

Kuhfeld (2017) og Wallace et al. (2016) presenterte også de gjennomsnittlige resultatene elevene hadde besvart på spørreskjemaet. Resultatene fra disse empiriske studiene kan fortelle oss at det er aspektet *challenge* de amerikanske ungdomsskoleelevene gjennomsnittlig var mest enige i. Det er ingen av de resterende seks aspektene som skiller seg veldig fra hverandre i hvilket aspekt som fikk den laveste gjennomsnittlige verdien. Det var kun et spørsmål som skulle måle aspektet *confer,* som fikk en betydelig lavere gjennomsnittlig verdi enn alle de andre spørsmålene i begge disse to studiene.

MET-studien viser til funn om at lærere med klasser som får høye verdier på statlige tester ser ut til fremheve en dypere konseptuell forståelse hos elevene enn klasser som får lave verdier på statlige tester (Kane og Cantrell, 2009). De fant også ut at det er en jevn fordeling i flere klasser hvor enige og uenige de er i de ulike aspektene av undervisningskvalitet (Kane og Cantrell, 2009). Videre påpeker de at bedre prestasjoner hos elevene vil kreve en bedre undervisning fra læreren (Kane og Cantrell, 2009).

Tidligere forskning i Norge om undervisningskvaliteten i matematikk er blitt utført av Bergem et al. (2016). De undersøkte undervisningskvaliteten i matematikk gjennom fire dimensjoner; god klasseromsledelse, støttende lærer, tydelige intensjoner og kognitive utfordringer. Sammenligner vi disse fire dimensjonene med de syv aspektene fra Ferguson

(2012) sitt spørreskjema ser vi at de i teorien innebærer de samme elementene. Bergem et al. (2016) brukte TIMSS 2015-studien sitt spørreskjema og resultater for å undersøke sammenhengen mellom lærerens undervisningskvalitet i matematikk og dens påvirkning på elevenes læringsutbytte i matematikk. Deres funn viser at det er en signifikant sammenheng mellom lærerens undervisningskvalitet og prestasjoner i matematikk på 9. Trinn med en regresjonskoeffisient 0,28. Resultatene fra spørreskjemaene viser også at den gjennomsnittlige *scoren* på alle spørsmålene er høyere enn den gjennomsnittlige verdien elevene kan svare (Bergem et al., 2016). De norske elevene i denne studien var gjennomsnittlig mest enige og fornøyde med dimensjonen ”støttende lærer” (Bergem et al., 2016).

3 Metode

I dette kapittelet presenterer jeg valget av metode for å belyse min problemstilling. Videre en beskrivelse av det brukte datamaterialet og fordeler og ulemper med dette. Deretter kommer en oversikt over analysestrategiene som er blitt brukt for å analysere resultatene. Kapittelet avsluttes med en innføring om studiens validitet og etiske rammer.

3.1 Valg av metode

For å belyse problemstillingen har jeg valgt å bruke en kvantitativ metode. Typisk for kvantitativ metode er at man gjennom en deduktiv tilnærming ønsker å gå i bredden og finne det som er typisk og generelt (Postholm og Jacobsen, 2014). I denne oppgaven har jeg valgt å fordype meg mer i sammenhengen mellom elevers syn på god undervisningskvalitet og prestasjoner, ved å gjøre flere statistiske analyser knyttet til de tre *casene,* nevnt innledningsvis. Basert på dette og teoretiske perspektiver presentert i kapittel 2, har jeg kommet frem til følgende hypoteser;

- At høyt presterende klasser er mer fornøyde med kvaliteten på undervisningen enn lavt presterende klasser.
- At klasser med stor progresjon i prestasjoner er mer fornøyde med kvaliteten på undervisningen enn klasser med liten progresjon.
- At klasser med stor variasjon i prestasjoner i klassen er mindre fornøyde med kvaliteten på undervisningen enn de med liten variasjon.

Gjennom en kvantitativ tilnærming får jeg også tilgang til et større utvalg, som dermed vil bidra til en dypere forståelse av min forskning.

3.2 Datamaterialet

Datamaterialet består av resultatene fra de nasjonale prøvene i regning på 8. og 9. trinn på aggregert klassenivå, det vil si at jeg ikke får sett hvordan hver enkelt elev har prestert, men får kun en oversikt over det gjennomsnittlige resultatet til de 46 klassene som har deltatt. At

alle de individuelle resultatene er den samme for alle elevene i en klasse. I tillegg består datamaterialet av individuelle elevevalueringer fra spørreskjemaet.

3.2.1 Utvalg

For å få fornuftige analyser gjennom en kvantitativ tilnærming, egner det seg best med et stort utvalg på minst 100 til 150 enheter (Postholm og Jacobsen, 2014). Utvalget i denne studien består totalt av 46 matematikklasserom som tilsvarer 991 elever, hvor 500 er jenter og 456 er gutter. Av de 991 elevene som er med i utvalget er det 35 som ikke har oppgitt om de er en gutt eller ei jente. Det er ønskelig at utvalget skal være representativt for en større populasjon enn de som har deltatt i studien (Postholm og Jacobsen, 2014). At LISA-studien er blitt utført i tilfeldige klasser rundt om i Norge bidrar til å styrke dette ønsket. Størrelsene på klassene varierer og disse 46 klassene består av 11 til 29 elever i hver klasse.

I denne studien varierer størrelsen på utvalget til de tre *casene* som blir undersøkt. Alle *casene* er delt opp i to hovedgrupper basert på gjennomsnittet fra alle de 46 klassene som har deltatt. For den første *casen*, som ser på de klassene som presterer over og under gjennomsnittet, vil de to hovedgruppene være høyt presterende klasser (22 klasser med til sammen 470 elever) og lavt presterende klasser (24 klasser med til sammen 521 elever). Neste *case* ser på de klassene som har stor fremgang (23 klasser med 494 elever) og de med liten fremgang (23 klasser med 497 elever). Den tredje og siste *casen* ser på de klassene med stor spredning i prestasjoner (22 klasser med 462 elever) og de med liten spredning i prestasjoner (24 klasser med 529 elever).

Gjennomsnittet i prestasjoner på den nasjonale prøven i regning på 9. trinn var 54,53. Det vil si at mestringsnivåene i de klassene som presterer over gjennomsnittet mest sannsynlig består av flest elever på mestringsnivå 3, 4 og 5. For de klassene som presterer under gjennomsnittet vil mest sannsynlig flest elever befinne seg på mestringsnivå 1, 2 og 3. At en klasse presterer under gjennomsnittet trenger ikke å bety at alle elevene i denne klassen presterer under dette gjennomsnittet. Det er helt naturlig at en klasse består av elever på flere mestringsnivåer. Gjennomsnittet baserer seg på hva flertallet har prestert.

3.2.2 Nasjonale prøver i regning

Fra LISA-studien har jeg tilgang til hva hver av de 46 klassene i gjennomsnitt har prestert på både den nasjonale prøven i regning på 8. trinn, som ble utført i 2012. Har også tilgang til resultatene den nasjonale prøven på 9. trinn, som ble utført i 2013. Denne informasjonen kan brukes til å gi en oversikt over hvordan disse klassene har prestert i forhold til hverandre og i forhold til resten av Norge i perioden 2012 til 2013. Gjennom deskriptive analyser er det enkelt å få oversikt over hvilke klasser som har prestert høyt og lavt, hvilken utvikling de har fra den første prøven til den andre, og hvor stor variasjon det er i hvordan hver enkelt elev i en klasse har prestert. Det vil si at selv om jeg ikke får sett hva hver elev i en klasse har fått som resultat på de nasjonale prøvene, så får jeg en indikasjon over prestasjonsnivået i klassene.

3.2.3 Spørreskjemaet

Spørreskjemaet som er blitt brukt i LISA-studien er som nevnt basert på Ronald Ferguson (2011) sin Tripod Student Perception Survey, som er et spørreskjema som er blitt brukt i flere amerikanske studier. Spørreskjemaet fokuserer som vi allerede vet, på undervisningskvaliteten i klasserommet med fokus på innhold, pedagogikk og relasjoner. I LISA-studien består spørreskjemaet av, i tillegg til de 35 spørsmålene til Ferguson (2011), tre ekstra spørsmål (Klette et al., 2017). Spørsmålene vil bli presentert senere i kapittel 4, som spørsmål 36, 37 og 38. Spørsmål 36 og 37 ønsker å måle aspektet *care*, og spørsmål 38 ønsker å måle *clarify*.

Ettersom at jeg bruker et datamateriale som allerede er blitt samlet inn gjennom LISA-studien, har jeg derfor ikke laget spørsmålene eller spørreskjemaet selv. I dette spørreskjemaet er det faste svaralternativer som består av en fem-punkts-skala, hvor 1=aldri, 2=sjeldent, 3=noen ganger, 4=ofte, 5=alltid. Faste svaralternativer betyr at deltakeren ikke kan svare med egne ord men må krysse av på kategorier som er bestemt på forhånd (Postholm og Jacobsen, 2014).

Styrker og svakheter med spørreskjema

Den største styrken ved spørreundersøkelser er at det kan inkludere mange deltakere som kan gi en indikator på hva som kan ses på som det generelle for den populasjonen, noe som gjør det mulig å kunne beskrive hvor ofte eller sjeldent noe finner sted (Postholm og Jacobsen,

2014). Når man ønsker å måle noe generelt for en populasjon er det dermed fint å ha muligheten til et stort utvalg. At spørsmålene består av faste svaralternativer er med på å sikre at samtlige svar blir avgitt på samme presisjonsnivå og det vil være større mulighet til å sammenligne svar fra mange personer (Kleven, 2014). Et spørreskjema er veldig effektivt når det gjelder å få mange enheters svar på klart avgrensede spørsmål (Postholm og Jacobsen, 2014).

Den største svakheten ved spørreskjemaer er at det ikke gir svar på mer enn det som blir spurt om. Når det brukes faste svaralternativer er det lite egnet for å få frem individuelle variasjoner og nyanser (Postholm og Jacobsen, 2014). Det er derfor viktig for meg som forsker å være oppmerksom på at det finnes flere tolkninger av de svarene som blitt utført (Kleven, 2014). Kvaliteten på slike undersøkelser er svært avhengige av hvordan spørsmålene og svarene er utformet. ”Hvis spørsmålene og svaralternativer er uklare, vage, upresise og lite relevante, vil svarene man får, også være nokså ubrukelige” (Postholm og Jacobsen, 2014, s. 87).

3.3 Analysestrategi

For å analysere resultatene brukte jeg et analyseverktøy som heter ”Statistical Package for the Social Sciences” som videre forkortes til SPSS. Forskerne som er med i LISA-studien bruker dette programmet, dermed fikk jeg tilgang til deres datamateriale som allerede var blitt overført til SPSS. Det betyr at i de tilfellene hvor det er ”missing values” som vil si at en elev ikke har svart på et spørsmål, har LISA-studien allerede ryddet opp i dette slik at analysene som blir utført ikke blir påvirket av manglende informasjon. I tillegg er 6 av de 38 spørsmålene som er brukt i studien formulert med en ”negativ” betydning, som vil si at verdien 1 vil være mer positiv enn verdien 5. Disse spørsmålene er også allerede blitt redigert i SPSS slik at verdien 5 vil være mer positiv enn verdien 1, som er tilfellet for alle de andre ”positivt” stilte spørsmålene. Et eksempel på et ”negativt” spørsmål er; ”Jeg misliker måten elevene oppfører seg på i timene”. Et eksempel på et ”positivt” spørsmål er; ”Elevene oppfører seg bra i mattetimene”.

I denne oppgaven har jeg valgt å bruke én avhengig variabel og totalt ni uavhengige variabler, som blir brukt til ulike formål. Alle variablene er hentet fra LISA-studien sin datafil i SPSS. Denne datafilen gir meg tilgang til hva alle de 991 elevene har svart på spørreskjemaet, samt hvordan de 46 klassene har prestert på regneprøven på 8. og 9. trinn. Den avhengige

variabelen jeg er interessert i er resultatene fra den nasjonale prøven i regning fra 9. trinn. Denne variabelen kommer jeg videre til å kalle for ”regne9”. Dette er en variabel som er aggregert på klassenivå, hvor jeg kun kjenner til klassegjennomsnittet. At alle de individuelle elevene i en klasse har fått den samme individuelle prestasjonsnivået er en ulempene med LISA-studien, siden det setter begrensninger i hvordan resultatene kan tolkes. De uavhengige variablene som blir brukt er de syv aspektene som måles gjennom spørreskjemaet: *care, control, clarify, challenge, captivate, confer* og *consolidate*. Dette for å undersøke undervisningskvaliteten i forhold til prestasjoner. Verdiene for disse syv variablene er manifeste verdier. Den åttende variabelen er resultatene fra den nasjonale prøven fra 8 trinn som jeg videre vil kalle for ”regne8”. Dette for å undersøke hvilken påvirkning elevenes forkunnskaper har på resultatene på regne9. Den siste uavhengige variabelen er fremgangen klassene har fått fra 8. til 9. trinn, for å se elevenes fremgang i forhold til hvordan de presterer.

For å analysere dataene brukte jeg flere metoder for å få et større grunnlag for videre antagelser. Først utførte jeg en konfirmerende faktoranalyse i Mplus, før jeg utførte ulike deskriptive analyser og til slutt en multippel regresjonsanalyse i SPSS.

3.3.1 Konfirmerende faktoranalyse i Mplus

Faktoranalyse benyttes for å lage et færre antall variabler (Clausen, 2009; Johannesen, 2009) ved at en undersøker om et spørsmål måler en eller flere dimensjoner (Ringdal, 2018). Det finnes to former for faktoranalyse; eksplorerende og konfirmerende (Johannesen, 2009). Til denne studien ble det utført en konfirmerende faktoranalyse, som benyttes som en deduktiv metode hvor man tester hypoteser og teorier som ligger til grunn for studien (Clausen, 2009). Denne analysen ble i denne studien brukt for å måle i hvor stor grad de 38 spørsmålene faktisk beskriver de syv aspektene for undervisningskvalitet. Spørsmålene i spørreskjemaet i LISA-studien er allerede knyttet opp mot hver av de syv aspektene, og gjennom en konfirmerende faktoranalyse kan vi måle hvor godt de tilhørende spørsmålene passer til hvert aspekt.

I samarbeid med veilederen min, Nils Buchholtz, importerte vi dataene fra spørreskjemaet fra SPSS til programvaren Mplus. Vi spesifiserte to ulike modeller for å prøve å finne ut hvilken kombinasjon som finner den best passende modellen for å måle undervisningskvalitet. Basert på Wallace et al. (2016) og Kuhfeld (2017), prøvde vi en syvdimensjonal faktormodell og

basert på min tolkning av aspektene, prøvde vi en tredimensjonal faktormodell, hvor vi knyttet de 38 spørsmålene til hvert sitt aspekt som vist i tabell 2.1 (side 15).

Modellkriterier

For å teste hvor godt disse to modellene passer til de empiriske dataene, er det nødvendig å undersøke visse modellkriterier. Fra Mplus-filen får vi informasjon om disse modellkriteriene; forholdet mellom khikvadrat og frihetsgrader, CFI og RMSEA, som er statistiske størrelser for å granske kvaliteten av modellen, i forhold til dataene.

Forholdet mellom khikvadrat og frihetsgrader – for å teste nullhypotesen og hvor godt faktormodellen passer til hypotesen (Geiser, 2013). I våre to tilfeller er nullhypotesene hvordan vi har valgt å fordele de 38 spørsmålene for å best måle undervisningskvalitet. For at modellen skal være god burde verdien være omtrent 3,5.

CFI (*comparative fit index*) – for å sammenligne hvor godt modellen passer til dataene, og om den spesifiserte modellen passer bedre enn en såkalt uavhengig *baseline* modell – som antar at det er ingen sammenheng mellom variablene. For en god modell burde CFI-verdien være større enn 0,95 (Geiser, 2013)

RMSEA (*root mean square error of approximation*) - for å måle tilnærmingene av dataene og hvor godt de passer til modellen. En god modell burde ha en RMSEA-verdi som er mindre enn 0,05 (Geiser, 2013).

Korrelasjonsanalyse

Fra Mplus-datafilen, finner vi også en korrelasjonsanalyse. Siden denne korrelasjonsanalysen er med latente verdier, som er verdier som kan observeres direkte, og resten av oppgaven er basert på manifeste verdier, som ikke kan observeres direkte, valgte jeg å utføre en egen korrelasjonsanalyse i SPSS. Gjennom en bivariat korrelasjonsanalyse med manifeste variabler, som gir en oversikt over hvordan de syv aspektene er i relasjon med hverandre (Clausen, 2009). Korrelasjonsnivået går fra 0 til 1, hvor 1 tilsvarer perfekt korrelasjon, som vil si at de måler det samme.

3.3.2 Deskriptiv analyse i SPSS

Bruker en deskriptiv analyse for å analysere de gjennomsnittlige verdiene fra de 46 klassene. Dette for å enkelt kunne velge ut de mest aktuelle klassene for en videre analyse. ”Ved beskrivende statistikk, eller deskriptiv statistikk, tar vi utgangspunkt i et datasett og ser hvordan enheter fordeler seg, enten i utvalg eller i populasjoner” (Johannesen, 2009, s. 48).

3.3.3 Multippel regresjonsanalyse i SPSS

Regresjonsanalysen ble brukt for å undersøke hvilke sammenhenger det er mellom den avhengige variabelen (regne9) og de uavhengige variablene (regne8, *care, control, clarify, captivate, challenge, confer, consolidate)*. Når vi ser på én avhengig variabel opp mot flere uavhengige variabler kaller vi det for en multippel regresjonsanalyse (Johannesen, 2009; Christophersen, 2012). Hensikten er å undersøke hvordan hver av de uavhengige variablene påvirker den avhengige variabelen (Skog, 2005; Johannesen, 2009). Det finnes to former for regresjonsanalyse, lineær og logistisk regresjonsanalyse (Johannessen, 2009). Dersom den avhengige variabelen har mange verdier benyttes det som i vårt tilfelle en lineær regresjonsanalyse (Johannesen, 2009). Det ble blant annet brukt en hierarkisk regresjonsanalyse på de syv aspektene og resultatene fra den nasjonale prøven på 8. trinn. En hierarkisk regresjonsanalyse vil si at de uavhengige variablene ble fordelt i blokker (Johannsen, 2009), som gir en oversikt over de ulike blokkenes påvirkning på den avhengige variabelen.

Det ble utført en multippel regresjonsanalyse til alle de tre *casene*, for å undersøke sammenhengene mellom dem og elevenes prestasjoner på regne9. Dette ble gjort for å se hvilken påvirkning de uavhengige variablene i de tre tilfellene har på prestasjonene. Som nevnt tidligere består hver *case* av to hovedgrupper. Disse hovedgruppene blir også brukt i regresjonsanalysene.

3.4 Validitet

Forskningen er valid når de samme resultatene ville kommet frem om studien ble gjort på nytt (Johnson og Christensen, 2013). Validitet i kvantitative studier handler om at de funnene man finner og de slutningene man tar er gyldige (Postholm og Jacobsen, 2014; Everett og Furuseth, 2012). Siden jeg bruker datamaterialet til LISA-studiet er det ikke noe jeg kunne

gjort annerledes for å bedre validiteten på spørreundersøkelsen. Validiteten i undersøkelsen ville kanskje økt dersom ikke all datamaterialet var på aggregat klassenivå. Dersom jeg også hadde hatt tilgang på individuelle data fra de nasjonale prøvene, slik at jeg også kunne ha trukket en mer gyldig slutning i sammenhengen mellom elevenes prestasjoner og deres syn på god undervisningskvalitet.

Begrepsvaliditet

Begrepsvaliditeten i denne studien handler om i hvilken grad de spørsmålene som er blitt stilt i spørreskjemaet faktisk måler disse tre aspektene av undervisningskvalitet; klasseromsledelse, støttende undervisning og kognitiv aktivering (Kleven, 2007; Johnson og Christensen, 2013). Jeg målte undervisningskvaliteten gjennom tre aspekter, mens spørreskjemaet måler gjennom syv aspekter, og den konfirmerende faktoranalysen bidrar til å styrke validiteten på oppgaven ved at den foreslår den beste modellen for å måle undervisningskvalitet.

Reliabilitet

Oppgavens reliabilitet handler om det er et godt nok grunnlag for å kunne gjøre en tolkning. Det vil si om vi har tilgang til nok datamateriale for å gjøre en tolkning og om disse tolkningene også gjelder for et større utvalg (Kleven, 2007; Johnson og Christensen, 2014). Reliabiliteten er god dersom de samme resultatene oppstår med flere målinger med det samme instrumentet.

Ytre og indre validitet

Den ytre validiteten handler om i hvilken grad de slutningene som tas kan overføres til en større populasjon (Johnson og Christensen, 2014). Siden jeg har et datamateriale fra tilfeldige ungdomsskoler rundt om i Norge, kan man anta at mine resultater også holder for andre ungdomsskoler i landet. Siden LISA-studien ikke har et representativt utvalg for hele landet, kan jeg derfor ikke si at resultatene i denne studien er generaliserbare for de andre ungdomsskolene som ikke har deltatt i denne studien, og kan derfor ikke konkludere for at dette stemmer for dem også. Studiens indre validitet handler om vi har gyldige bevis for å trekke de sammenhenger som tas mellom de avhengige og uavhengige variablene (Johnson og Christensen, 2014).

Troverdighet

For å sikre oppgavens validitet, må også troverdigheten rundt elevenes vurderinger i spørreskjemaet vurderes. Det hadde vært gunstig å måle hvor objektive elevene er med læreren sin (Kuhfeld, 2017), det vil si om de vurderer læreren sin riktig eller om de vurderer læreren ut i fra eventuelle biaser. Dessverre er ikke dette mulig å undersøke. Likevel har tidligere studier funnet ut at både elevvurderinger og lærervurderinger kan brukes til å forske på lærerens undervisningskvalitet (Kunter et al., 2007) og at elevenes oppfatning ofte viser til større sammenheng med prestasjoner enn lærerens oppfatning (Schweig, 2016). Wallace et al. (2016) fant i sin studie ut at elever og lærere også kan vurdere undervisningskvaliteten ganske ulikt. I tillegg er alle elevvurderingene oppgitt i gjennomsnitt, så til de gjennomsnittene som er oppgitt på klassenivå må også konsensusen i klassen vurderes (Schweig, 2016). Jeg var interessert i hvordan klassene i gjennomsnitt har vurdert de syv aspektene, ut i fra hvilke klasser som er høyt presterende, lavt presterende, stor fremgang, liten fremgang, stor spredning i prestasjoner og liten spredning i prestasjoner. Det er derfor viktig å undersøke konsensusen i de aktuelle klassene for å få et mer korrekt bilde av klassens oppfatning. Lavt standardavvik vil derfor forklare at det er stor enighet blant elevene i klassen. Det er også nødvendig å ta i betrakting hvordan elevene oppfatter spørsmålene som blir stilt. Det kan være at de har en annen tolkning av spørsmålene og opplever noen spørsmål som mer relevante enn andre, noe som kan ha en påvirkning på resultatene.

3.5 Forskningsetikk

Siden jeg er en del av et eksisterende forskningsprosjekt var det flere etiske problemstillinger jeg ikke trengte å ta stilling til. Det var derfor ikke et behov til å melde i fra til NSD (Norsk senter for forskningsdata) og sende ut forespørsel til de som ønsket å delta. Jeg har ikke tilgang til hvilke skoler som har deltatt i studien og jeg har heller ikke mulighet til å gi ut sensitiv informasjon om de deltakende elevene og skolene. Uansett er det viktig at jeg som forsker ikke misbruker datamaterialet og behandler det med verdighet.

4 Analyse av resultater

I dette kapittelet presenteres mine funn fra analysen. Kapittelet er delt opp i tre hoveddeler; konfirmerende faktoranalyse av de syv aspektene for god undervisningskvalitet, deskriptive analyser av de tre *casene* og en multippel regresjonsanalyse for å analysere påvirkningen fra de syv aspektene på regne9.

4.1 Resultater fra den konfirmerende faktoranalysen

Den konfirmerende faktoranalysen i Mplus kom frem til en faktormodell beskrevet i tabell 4.1. Der ser vi hvilke spørsmål Mplus mente passet best til hver av de syv aspektene for undervisningskvalitet. Faktorladningene som er oppgitt i tabellen forteller om hvor sterk sammenheng det er mellom variablene, som i dette tilfellet er spørsmålene i spørreskjemaet, og de faktorene de er blitt knyttet til (Johannesen, 2009). Sammenhengen er sterk dersom faktorladningen er over 0,71, veldig bra dersom den ligger mellom 0,63-0,70, god mellom 0,55-0,62, middels mellom 0,45-0,54, svak mellom 0,32-0,44 og dårlig fra 0,31 og lavere (Tabachnick og Fidell, 2007). I de tilfellene hvor et spørsmål har fått en svak faktorladning betyr det at det spørsmålet som er blitt stilt ikke måler den oppgitte faktoren i like stor grad som et annen spørsmål med en sterk faktorladning. Det vil si at det er noen av spørsmålene som i større grad måler de oppgitte faktorene enn andre.

Tabell 4.1: En oversikt over hvilke spørsmål som måler hvilket aspekt med tilhørende faktorladning. De cellene som er markert blå viser spørsmål som er silt på en "negativ" måte, slik at verdiene til disse spørsmålene er blitt snudd om på i resultatene.

Care	Spørsmål	Faktorladning
1	Mattelæreren får meg til å føle at han/hun virkelig bryr seg om meg	0,75
2	Hvis det er noe som plager meg, tror jeg mattelæreren forstår det	0,76
3	Mattelæreren prøver virkelig å forstå hvordan elevene har det	0,76
37	Mattelæreren har forståelse for at elevene er slitne eller har hatt en lang dag	0,64
38	Mattelæreren bruker tid på å hjelpe alle elevene	0,67
Control		
4	Elevene oppfører seg bra i mattetimene	0,69
5	Jeg misliker måten elevene oppfører seg på i timene	0,54
6	Elevene i denne klassen oppfører seg på en måte som gjør mattelæreren irritert	0,66

7	Elevenes oppførsel i mattetimene er et problem	0,64
8	I mattetimene oppfører elevene seg slik læreren vil	0,69
9	Elevene behandler mattelæreren med respekt	0,73
10	I mattetimene jobber vi bra og kaster ikke bort tiden	0,64
Clarify		
11	Hvis det er noe vi ikke forstår, forklarer mattelæreren det på en annen måte	0,78
12	Mattelæreren skjønner om elevene har forstått, eller om vi ikke har forstått	0,74
13	Når mattelæreren underviser, tror han/hun at vi har forstått selv om vi ikke har det	0,48
14	Mattelæreren har mange gode måter å forklare det vi skal lære på	0,82
15	Mattelæreren forklarer ting på en tydelig måte	0,78
36	Mattelæreren gjennomgår stoffet for fort	0,47
Challenge		
16	Mattelæreren stiller spørsmål for å forsikre seg om at vi følger med når hun/han underviser	0,60
17	Når elevene svarer på spørsmål fra mattelæreren, ber han/hun dem om å begrunne svarene sine eller forklare dem bedre	0,49
18	Mattelæreren godtar bare full innsats fra elevene i denne klassen	0,36
19	Mattelæreren lar ingen få lov til å gi opp selv om oppgavene er vanskelige	0,38
20	Mattelæreren vil at jeg skal begrunne svarene mine – hvorfor jeg mener det jeg gjør	0,53
21	I mattetimene lærer vi mye nesten hver gang	0,69
22	I mattetimene lærer vi å rette feilene våre	0,68
Captivate		
23	Jeg klarer ikke å følge med i mattetimene, for jeg kjeder meg	0,60
24	Mattelæreren gjør det morsomt å lære	0,88
25	Mattelæreren gjør timene interessante	0,93
26	Jeg liker måten vi lærer på i mattetimene	0,83
Confer		
27	Mattelæreren vil at vi skal dele tanker med hverandre	0,56
28	Elevene får være med å bestemme hvordan de forskjellige aktivitetene skal gjøres i mattetimene	0,56
29	Mattelæreren gir oss tid til å forklare tanker og ideer som vi har	0,70
30	Elevene tar gjerne ordet og sier hva de mener om arbeidet i mattetimene	0,51
31	Mattelæreren respekterer mine ideer og forslag	0,74

Consolidate		
32	Mattelæreren tar seg tid til å oppsummere det vi lærer hver dag	0,64
33	Mattelæreren sjekker om vi har forstått det han/hun har lært oss	0,78
34	Vi får nyttige kommentarer på tekstene våre, slik at vi forstår hva vi har gjort feil	0,73
35	Kommentarene jeg får på arbeidet mitt i mattetimene, hjelper meg til å forstå hvordan jeg skal forbedre meg	0,73

Flertallet av faktorladningene i tabell 4.1 har fått en høy verdi, som betyr at spørsmålene i spørreskjemaet faktisk måler de syv aspektene. De cellene som har fått en oransje farge viser til faktorladninger som har en sterk sammenheng med faktoren de beskriver. De cellene som har fått en grønn farge viser til faktorladninger som har en svak sammenheng. Vi ser at det kun er spørsmål 18 og 19 som har fått en svak faktorladning. Særlig faktorene *care, clarify, captivate* og *consolidate,* har hovedsakelig fått faktorladninger som viser en sterk sammenheng mellom flere av spørsmålene og faktoren de skal måle. Vi kan ut ifra tabellen si at de spørsmålene som er blitt stilt i stor grad måler de syv faktorene.

4.1.1 Statistiske modelltester

Ulike tester som blir utført i Mplus kan fortelle oss om hvor godt en modell passer i forbindelse med å måle undervisningskvaliteten til læreren.

Resultatene for den syvdimensjonale faktormodellen:

Forholdet mellom Khikvadrat og frihetsgrader = $\frac{\chi^2}{df} = 3.55$, CFI = 0.905, RMSEA = 0.051

og resultatene fra tredimensjonale faktormodellen:

Forholdet mellom Khikvadrat og frihetsgrader = $\frac{\chi^2}{df} = 6.91$, CFI = 0.775, RMSEA = 0.077

I denne studien ble det ikke laget en test for sammenligning av modeller ettersom at det ville være for komplekst for denne oppgaven. Dette er noe som burde gjøres for å best måle de to modellene opp mot hverandre. Likevel gir de verdiene fra disse tre testene en god nok indikasjon på hvor godt de passer til hva vi ønsker å undersøke i denne oppgaven. Disse testene synliggjør at det er den syvdimensjonale faktormodellen som best måler lærerens undervisningskvalitet i matematikk. Vi ser at forholdet mellom khikvadratverdien og

frithetsgrader viser en veldig god verdi på 3.55, som vil si at faktormodellen passer godt. Videre er det ønskelig med en CFI-verdi større enn 0.95, og resultatene viser at den syvdimensjonale faktormodellen har den høyeste verdien. Ønskelige verdier på RMSEA er mindre enn 0.05, og også her har den syvdimensjonale faktormodellen den beste verdien i forhold til den tredimensjonale faktormodellen.

Korrelasjonsanalysen

Korrelasjonsanalysen er vist i tabell 4.2. Denne tabellen viser at de fleste aspektene samsvarer ganske høyt med hverandre, utenom *control*. At korrelasjonsverdiene har to stjerner (**) bak forteller at disse korrelasjonsverdiene er veldig presise med en usikkerhet på 1%. Det betyr at vi kan være ganske sikre på at disse sammenhengene ikke er tilfeldige.

Tabell 4.2: En oversikt over hvordan de syv aspektene korrelerer med hverandre. *Den oppgitte verdien er Pearson r.*

Aspekt	Care	Control	Clarify	Challenge	Captivate	Confer	Consolidate
Care	1	0,39**	0,69**	0,55**	0,63**	0,64**	0,65**
Control	0,39**	1	0,45**	0,38**	0,39**	0,30**	0,32**
Clarify	0,69**	0,45**	1	0,58**	0,69**	0,57**	0,65**
Challenge	0,55**	0,38**	0,58**	1	0,53**	0,56**	0,64**
Captivate	0,63**	0,39**	0,69**	0,53**	1	0,57**	0,59**
Confer	0,64**	0,30**	0,57**	0,56**	0,57**	1	0,63**
Consolidate	0,65**	0,32**	0,65**	0,64**	0,59**	0,63**	1

4.1.2 Deskriptiv analyse av de syv aspektene for undervisningskvalitet

Videre ble det gjort en deskriptiv analyse av de syv aspektene på hele utvalget for å få en oversikt over hvordan elevene har vurdert de syv faktorene. Tabell 4.2 består av manifeste data for en bedre tolkning av verdiene.

Tabell 4.3: En deskriptiv analyse av de syv aspektene på klassenivå.

Aspekt	N	Minimum	Maximum	Gjennomsnitt	Standardavvik
Care	990	1,00	5,00	3,88	0,75
Control	991	1,43	5,00	3,97	0,57
Clarify	991	1,00	5,00	3,92	0,69
Challenge	991	1,71	5,00	3,90	0,58
Captivate	990	1,00	5,00	3,52	0,92
Confer	989	1,00	5,00	3,38	0,72
Consolidate	985	1,00	5,00	3,87	0,78

Tabell 4.3 beskriver hvor mange elever (N) som har vurdert de ulike spørsmålene som er knyttet til de syv aspektene for undervisningskvalitet. Vi ser at antallet varierer fra aspekt til aspekt, noe som betyr at ikke alle elevene har svart på alle spørsmålene i spørreskjemaet. SPSS registrerer disse som *missing values*, og blir dermed ikke tatt med i beregningen av gjennomsnittene. De gjennomsnittlige verdiene er alle relativt høye og like, hvilket betyr at elevene i gjennomsnitt har svart ganske likt på alle spørsmålene. Standardavvikene har også litt høye verdier, som betyr at det er litt variasjon i hva elevene har svart på spørsmålene. Noe vi ser tydelig fra tabellen, med minimumsverdier fra 1 til maksimumsverdier 5. Den gjennomsnittlige verdien viser at det er *confer* som fikk minst enighet blant elevene og *control* som fikk mest enighet. Ser vi på standardavvikene til disse to aspektene viser deg seg at *confer* fikk det høyeste standardavviket og *control* det minste. Det betyr at *control* er det aspektet, hvor flest elever er enige om påstandene og enige med hverandre, mens det for *confer* ser ut til å være motsatt.

4.1.3 Studiens resultater i forhold til andre empiriske resultater

Tabell 4.4: En oversikt over de gjennomsnittlige verdiene fra LISA-studien, Kuhfeld (2017) og Wallace et al. (2016).

Aspekt	LISA-studien	Kuhfeld (2017)	Wallace et al. (2016)
Care	3,88	3,64	3,40
Control	3,97	3,32	3,12
Clarify	3,92	3,83	3,62
Challenge	3,90	4,10	4,10
Captivate	3,52	3,54	3,36
Confer	3,38	3,38	3,39
Consolidate	3,87	3,74	3,75

For en bedre tolkning av de manifeste verdiene fra tabell 4.2, sammenligner jeg i tabell 4.4

gjennomsnittsverdiene med manifeste data fra to tidligere empiriske studier utført i USA. Fra denne tabellen (4.4) vises de gjennomsnittlige verdiene fra den deskriptive analysen over alle klassene i LISA-studien, sammenlignet med de empiriske resultatene fra Kuhfeld (2017) og Wallace et al. (2016). Det er viktig å merke seg at de gjennomsnittlige verdiene i LISA-studien fra *care* og *clarify* består av flere spørsmål som måler dette aspektet enn for de to empiriske studiene Kuhfeld (2017) og Wallace et al. (2016). Poenget med disse resultatene er å få en indikasjon over hvilket aspekt de deltakende elevene i gjennomsnitt har vurdert seg mest fornøyde med. Tabell 4.4 viser også at de elevene i LISA-studien som representerer Norge, er mest fornøyde med *control*, i motsetning til Kuhfeld (2017) og Wallace et al. (2016) sine studier i USA, som er mest fornøyde med *challange*. Vi ser også at elevene i LISA-studien er minst fornøyde med *confer* og i Kuhfeld (2017) og Wallace et al. (2016) sine studier er man minst fornøyde med *control*. Verken *control* eller *confer* har fått noen ekstra spørsmål i LISA-studien, noe som vil si at de verdiene vi har i tabell 4.4 på de to aspektene består av de samme spørsmålene som er brukt i studiene til Kuhfeld (2017) og Wallace et al. (2016). En mulig hypotese for at Norge skiller seg fra USA kan være at det er ulik fokus på hva som definerer god undervisning i de to landene. For eksempel kan det bety at undervisningen i USA kanskje fokuserer mer på lærerstyrt undervisning, mens det i Norge er mer fokus på elevsentrert undervisning.

4.2 Resultater fra de deskriptive analysene på klassenivå

Jeg valgte å gjøre flere deskriptive analyser ut ifra de tre *casene* som ble nevnt innledningsvis. I denne delen viser jeg kun tabeller til noen klasser for å unngå for mange tabeller. Først gjorde jeg en deskriptiv analyse for å kunne sammenligne de klassene som presterte best og dårligst på den nasjonale prøven på 9. trinn. Deretter for å kunne sammenligne de klassene som hadde størst og minst fremgang fra den første til den andre nasjonale prøven, og til slutt for å sammenlikne de klassene som hadde størst og minst spredning i prestasjoner innad i klassene.

4.2.1 Høyt presterende vs. lavpresterende klasser

Først valgte jeg å fordele de 46 klassene opp i to grupper. Den ene gruppen består av de klassene som på klassenivå presterte over gjennomsnittet og den andre av de som presterte under gjennomsnittet. Dette gjorde jeg for å få en oversikt over hvilke klasser som presterte best og dårligst i forhold til de andre klassene som deltok i studien.

De resultatene jeg har fra de nasjonale prøvene på 8. og 9. trinn er bare et utvalg av alle resultatene på den nasjonale prøven utført i 2012 og 2013 fra alle de norske ungdomsskolene i Norge. Informasjonen fra de nasjonale prøvene på Utdanningsdirektoratet sin nettside, viser kun resultater fra perioden 2014-2019. Disse resultatene viser at den gjennomsnittlige verdien på den nasjonale prøven er på 54.00. Siden gjennomsnittene fra 2012-2013 ikke lenger ligger offentlig publisert, kan det tenkes at også den ligger på rundt 54 (Utdanningsdirektoratet, 2019). Mine resultater viser at den gjennomsnittlige *scoren* fra de deltakende klassene på klassenivå fra den nasjonale prøven i regning på 9. trinn som jeg har tilgang til var 54,53. Det vil si at vårt utvalg i gjennomsnitt presterte ganske likt med gjennomsnittet fra alle ungdomsskolene i Norge. Gjennom en gjennomsnittsanalyse i SPSS viste det seg at det er 22 klasser som på klassenivå presterte over dette gjennomsnittet og 24 klasser som presterte under.

I denne studien har jeg valgt å kalle de klassene som presterer over gjennomsnittet for høyt presterende klasser og de som presterer under for lavtpresterende klasser. Dette for å få en indikasjon over hvordan de presterer i forhold til hverandre. Henviser vi til Utdanningsforbundet (2017) sin fordeling av mestringsnivåer på de nasjonale prøvene, består en gjennomsnittlig sum på 54,53 av gjennomsnittlige mestringsnivåer mellom tre og fire (se kapittel 2.2).

Analyse av de høyt presterende klassene

For å få den beste oversikten over hvordan de klassene som presterte aller best i denne studien har vurdert lærerens undervisningskvalitet, valgte jeg først å gjøre en deskriptiv analyse av den klassen som presterte aller best av alle de 46 klassene, det var klasse2. Deretter analyserte jeg de fem klassene som presterte best, for å se på eventuelle sammenhenger mellom hvordan disse fem opplevde undervisningen i forhold til hvor godt de presterte.

Deskriptiv analyse av den aller beste klassen (klasse2).

Klasse2 er den klassen som presterte aller best på regne9 av alle de 46 klassene som var med i studien og nest best på regne8. De er også den klassen som har den aller største fremgangen fra regne8 til regne9.

Tabell 4.5: En deskriptiv analyse av klasse2, hvor vi ser hvordan de i gjennomsnitt har svart på spørreskjemaet.

Aspekt	Antall	Minimum	Maksimum	Gjennomsnitt	Standardavvik
Care	20	2,80	4,60	4,03	0,67
Control	20	3,14	5,00	4,21	0,47
Clarify	20	3,33	4,67	4,05	0,43
Challenge	20	3,43	4,86	4,05	0,47
Captivate	20	3,00	5,00	4,20	0,59
Confer	20	3,00	4,60	3,57	0,45
Consolidate	20	3,25	5,00	4,21	0,53

Tabell 4.5 viser hvor mange elever som svarte på de ulike spørsmålene i spørreskjemaet. Vi ser også hva elevene i gjennomsnitt svarte som minste og største verdi til de ulike aspektene. Gjennomsnittene viser at denne klassen på en skala fra 1 til 5, rangerte ganske høyt. Standardavvikene forteller at det ikke var store uenighetene innad i klassen om de ulike aspektene. Tabellen viser likevel at det var litt mer uenighet til *care* og *captivate*. Gjennomsnittet viser at klasse2 har ganske jevne oppfatninger til seks av de syv aspektene. Det er særlig gjennomsnittet til *confer* som skiller seg ut, som er det aspektet de var minst enige i påstandene om. Tabellen forteller at de var mest fornøyde med aspektene *consolidate*, *control* og *captivate*. Det er ikke noe aspekt som peker seg mer positivt ut enn andre. Ut ifra standardavvikene ser det ut til at det var *control* elevene var mest enige om, også innad i klassen.

Deskriptiv analyse av de fem sterkeste klassene

Videre har jeg valgt å ikke presentere en like dyptgående analyse av de fem beste klassene. Likevel er det viktig å nevne at tilsvarende analyser er blitt gjort og kun den viktigste informasjonen er blitt samlet i neste tabell 4.6. De fem klassene som presterte aller best er i avtagende rekkefølge; klasse2, klasse7, klasse15, klasse25 og klasse44. Der klasse2 presterte aller best og klasse44 femte best. For å kunne sammenligne de klassene som presterte best og for å prøve å finne noen sammenhenger mellom dem, viser tabell 4.6 hva disse klassene har

vurdert høyest og lavest i gjennomsnitt, av de syv aspektene. Slik får vi en oversikt over hva disse klassene har vært mest enige i av de syv aspektene og hva de har vært minst enige i.

Tabell 4.6: En oversikt over de fem klassene som presterte høyest på den nasjonale prøven i regning på 9. trinn og hvilke aspekter de vurderte seg mest og minst enige i.

Klasse	*Score*	Std.	Mest enig i		Minst enig i	
2	67,06	6,37	control	4,21	confer	3,57
7	62,92	9,06	control	3,91	captivate	2,78
15	60,77	8,54	consolidate	3,93	confer	3,46
25	60,18	9,86	consolidate	4,21	confer	3,66
44	59,47	9,59	control	4,45	confer	3,28

Standardavvikene i tabell 4.6 viser hvor stor variasjon av prestasjoner det er blant disse fem klassene. Klasse2 skiller seg ut med den laveste verdien. De andre fire klassene ser ut til å bestå av elever på relativt like mestringsnivåer. Det betyr at disse resultatene kan sammenlignes med hverandre i relativt stor grad. Videre viser tabellen at disse fem klassene hadde en ulik grad av enighet i de syv aspektene for god undervisningskvalitet. Tre av fem klasser hadde gjennomsnittlige verdier over fire, mens to under fire. Det vil si at selv om det er en høyt presterende klasse trenger det ikke å bety at de var veldig fornøyde med kvaliteten på undervisningen. Som nevnt varierer graden av enighet blant klassene, likevel var det *control* og *consolidate* som komm best ut av vurderingen. Fire av fem klasser opplevde *confer* minst i undervisningen, og en klasse opplevde *captivate* minst. Det er viktig å legge merke til at de gjennomsnittlige verdiene på de lavest vurderte aspektene, ikke er spesielt lave i forhold de høyest vurderte aspektene. Det betyr at disse elevene generelt var positive til lærerens undervisning.

Analyse av de lavt presterende klassene

Tilsvarende for de høyt presterende klassen utførte jeg de samme statistiske analysene i SPSS for de lavt presterende klassene. Først presenteres derfor resultatene fra den klassen som presterte aller dårligst, som er klasse1. Videre presenteres resultatene fra de fem klassene som presterte dårligst.

Deskriptiv analyse av den dårligste klassen (klasse1)

Klasse1 er den klassen som presterte aller dårligst på regne9 og regne8 av alle de 46 deltakende klassene. Som for klasse2 viser jeg her den deskriptive analysen av klasse1, for å få en oversikt over hvordan den svakeste klassen vurderte i spørreskjemaet.

Tabell 4.7: Deskriptiv analyse av klasse1.

Aspekt	Antall	Minimum	Maksimum	Gjennomsnitt	Standardavvik
Care	22	3,00	5,00	4,06	0,62
Control	22	3,14	5,00	3,94	0,42
Clarify	22	3,00	4,50	3,87	0,47
Challenge	22	3,29	4,93	4,21	0,55
Captivate	22	2,50	5,00	3,58	0,75
Confer	22	2,60	4,70	3,73	0,66
Consolidate	22	2,00	5,00	4,05	0,91

Standardavviket i tabell 4.7 forteller oss at det var enkelte aspekter elevene var mer enige i innad i klassen enn på andre aspekter. De var mest enige med hverandre i aspektet *control,* og minst enige i *consolidate*. De gjennomsnittlige verdiene viser at elevene i denne klassen på en skala fra 1 til 5 generelt var fornøyde med deres lærer sin undervisning. Likevel var det *challenge* de var mest enige i påstandene om og *captivate* de var minst enige i.

Deskriptiv analyse av de fem svakeste klassene

Tilsvarende som for de fem klassene som presterte best har jeg gjort det samme med de fem klassene som presterte dårligst. Disse fem klassene er i avtagende rekkefølge; klasse1, klasse14, klasse37, klasse38 og klasse24, hvor klasse1 presterte dårligst.

Tabell 4.8: En oversikt over de klassene som presterte dårligst på den nasjonale prøven i regning på 9. trinn og hvilke aspekter de var mest og minst enige i.

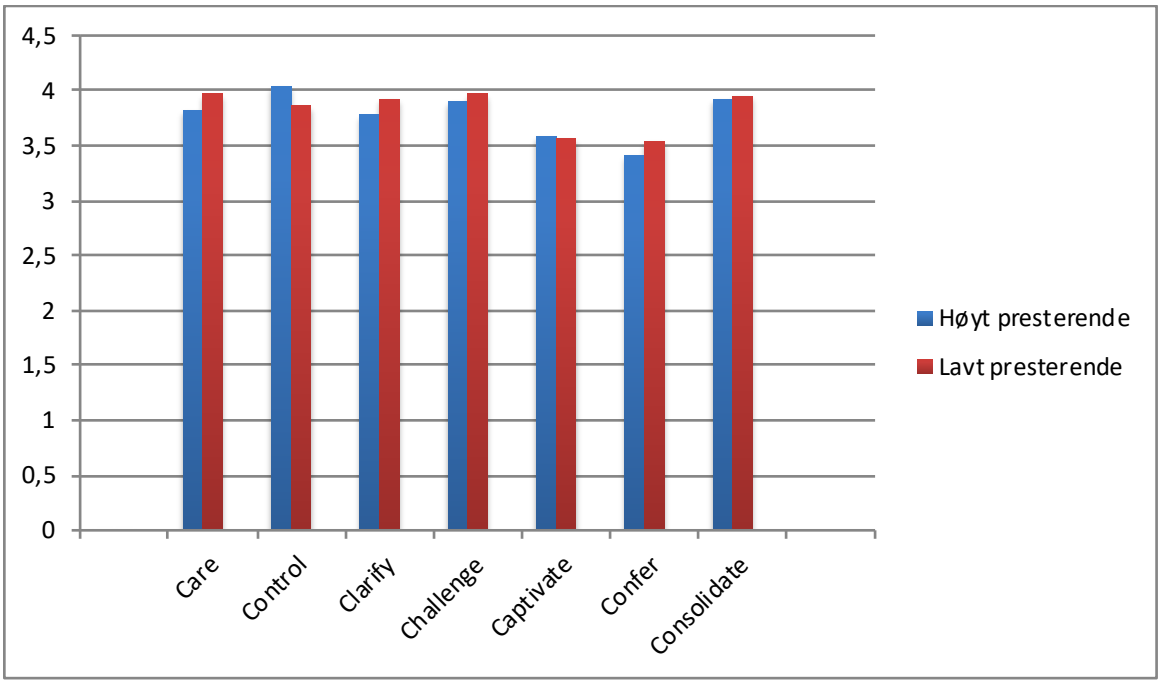
Klasse	*Score*	Std.	Mest enig i		Minst enig i	
1	45,35	6,62	challenge	4,21	captivate	3,58
14	49,42	9,47	control	4,20	confer	3,00
37	49,65	11,39	consolidate	4,05	challenge	3,48
38	50,05	10,35	clarify	4,29	confer	3,72
24	50,43	9,72	challenge	3,88	captivate	3,33

Tabell 4.8 viser de klassene som presterte under gjennomsnittet på den nasjonale prøven. Standardavvikene viser at det var større spredning i prestasjoner i disse klassene i forhold til

hverandre. Som vi ser i tabellen, har klasse1 et veldig lavt standardavvik, noe som betyr at dette var en klasse som bestod av flere faglig svakere elever. For de fire andre klassene er standardavvikene mye høyere, det vil si at disse klassene hadde mer variasjon av mestringsnivåer innad i klassen. Tabellen viser også at det er mer variert i hvilke aspekt som ble vurdert høyest og lavest. De gjennomsnittlige verdiene på de aspektene som ble høyest vurdert viser til høye verdier. For de aspektene som ble lavest vurdert viser også til relativt høye verdier. Dette indikerer at de fem faglig svakeste klassene også var relativt fornøyde med undervisningen.

Sammenligning av de som presterte best og de som presterte dårligst

For å få en større oversikt over hvordan de høyt presterende vurderte i forhold til de lavt presterende klassene, utførte jeg en gjennomsnittsanalyse av disse to gruppene. Resultatene er presentert i figur 4.1.



Figur 4.1: Et stolpediagram som viser hvordan gjennomsnittet til de fem mest høyt presterende og lavt presterende klassene har vurdert de syv aspektene i forhold til hverandre.

Figur 4.1 viser de gjennomsnittlige verdiene fra alle de klassene som presterte best og de som presterte lavest. Figuren viser at disse to gruppene hadde ganske like verdier, noe som vil si at det ikke var store forskjeller mellom de som presterte best og dårligst i denne studien i hvordan de vurderte i spørreskjemaet om undervisningskvalitet. Figuren viser faktisk at de

lavt presterende hadde vurdert seg litt mer enige i fem av syv aspekter; *care, clarify, challenge, confer, consolidate.* De høyt presterende var litt mer enige i *control* og *captivate.* Uavhengig av hvem gruppe som var mest eller minst enige, så var det for begge gruppene *confer* og *captivate* som ble vurdert aller dårligst.

4.2.2 Stor fremgang vs. liten fremgang

Gjennom den deskriptive analysen i SPSS fikk jeg også en oversikt over hvilke klasser som hadde den største fremgangen fra den første nasjonale prøven i regning på 8. trinn, og til den andre på 9. trinn. Denne analysen viste at den gjennomsnittlige fremgangen på de to nasjonale prøvene fra de deltakende klassende var 3,57 poeng. Dermed fant jeg ut hvilke klasser som hadde bedre fremgang enn gjennomsnittet og hvilke som hadde dårligere.

Deskriptiv analyse av de fem klassene med størst fremgang.

Det er blitt gjort tilsvarende analyser av disse fem klassene som er blitt vist tidligere for klasse2 og klasse1, men i dette tilfellet for å kunne gjøre en sammenligning ser jeg kun behovet for å presentere hva de er mest og minst enige i.

Tabell 4.9: En oversikt over de klassene som hadde størst forbedring på de to nasjonale prøvene i regning enn gjennomsnittet.

Klasse	8. trinn	9. trinn	Fremgang	Mest enig i		Minst enig i	
2	57,92	67,06	9,15	control	4,21	confer	3,57
15	53,88	60,77	6,89	consolidate	3,93	confer	3,46
36	49,46	55,65	6,20	care	3,74	confer	3,27
14	43,28	49,42	6,14	control	4,20	confer	3,00
4	50,45	56,27	5,81	care	4,39	confer	3,83

Tabell 4.9 viser at fire av fem klasser med størst forbedring presterte over gjennomsnittet på regne9. Det var kun klasse14 som presterte under gjennomsnittet. Tabellen viser også at klasse2 som var den aller beste klassen på regne9, hadde en betydelig større fremgang enn de andre fem klassene. Sammenligner vi resultatene på regne9 med regne8 ser vi at flere av klassene har gått fra å prestere under gjennomsnittet til å prestere over. Ser vi på hvilket aspekt disse fem klassene har vurdert seg mest enige i, så ser vi ingen tydelig sammenheng. Derimot opplevde alle fem klassene at *confer* oppstod minst i undervisningen. De

gjennomsnittlige verdiene fra spørreskjemaet viser til variasjon av enighet blant disse fem klassene. På en skala fra 1 til 5 hadde likevel alle en relativt høy verdi.

Deskriptiv analyse av de fem klassene med lavest fremgang

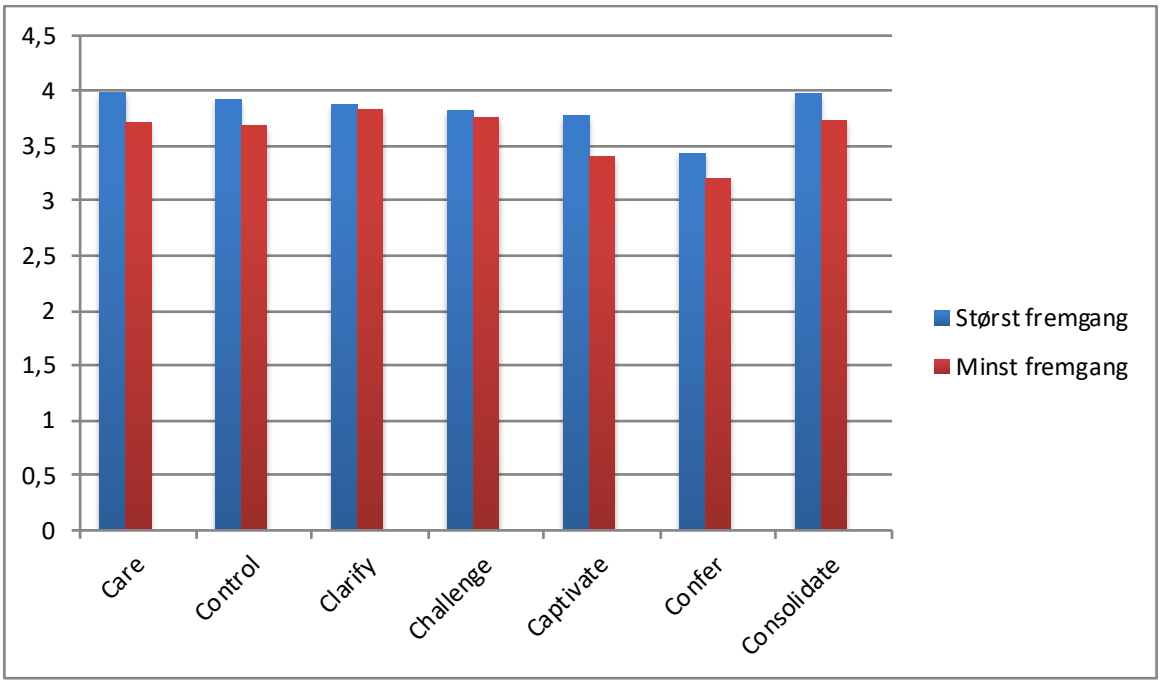
Tabell 4.10: En oversikt over de klassene som hadde mindre fremgang enn gjennomsnittet.

Klasse	8. trinn	9. trinn	Fremgang	Mest enig i		Minst enig i	
28	50,49	51,01	0,52	consolidate	4,00	confer	3,19
5	53,28	54,56	1,28	control	4,14	confer	3,24
18	50,98	52,26	1,29	clarify	3,86	control	3,12
16	54,58	55,88	1,30	consolidate	3,66	confer	3,07
6	53,44	54,93	1,49	challenge	4,17	confer	3,40

Tabell 4.10 viser at alle de fem klassene med lavest fremgang presterte over gjennomsnittet på regne8, men kun tre av fire presterte over gjennomsnittet på regne9. Klasse28 har en betydelig mindre fremgang enn alle de andre klassene. I tillegg til at de har såpass liten fremgang presterte de også under gjennomsnittet på regne9. Likevel viser gjennomsnittsverdiene fra spørreskjemaet at også disse fem klassene var relativt enige med de påstandene som ble stilt. Det er ingen tydelig sammenheng mellom hvilket aspekt de fem klassene med lavest fremgang opplevde oftest i undervisningen. Derimot opplevde fire av fem klasser *confer* minst i undervisningen.

Sammenligning av de med størst og minst fremgang

Hvis vi sammenligner de fem klassene som hadde størst fremgang opp mot de fem som hadde minst, ser vi fra figur 4.2, ingen tydelige mønstre annet enn at det var *confer* som kommer dårligst ut av vurderingen i så å si alle klassene. Hva klassene var mest enige om ser ikke ut til å ha hatt en stor påvirkning av fremgangen i prestasjoner. Ser fra figuren at det ikke var store forskjellene i gjennomsnittsverdiene mellom disse to gruppene, men at de klassene med størst fremgang generelt vurderte seg litt mer enige i alle aspektene enn de med minst fremgang.



Figur 4.2: En oversikt over hvordan de fem klassene mest størst fremgang i gjennomsnitt vurderte de syv aspektene i forhold til de fem klassene med minst fremgang.

4.2.3 Stor spredning vs. liten spredning i prestasjoner

LISA-data består kun av resultatene fra nasjonale prøver på klassenivå. Det vil si at jeg ikke får sett hva hver enkelt elev har fått på prøven. Dermed vet jeg heller ikke om en klasse som presterer høyt bare består av elever som presterer høyt, eller hvordan hver enkelt elev presterer i en klasse. Tilsvarende er det for de klassene som presterte lavt. Ut ifra standardavvikene for resultatene fra regne9, som var gitt sammen med prestasjonsdataene på klassenivå, får jeg en indikasjon av hvordan prestasjonsevnene i en klasse varierer. Jo større standardavvik det er, desto større spredning er det i prestasjoner innad i klassen, og motsatt for lavt standardavvik. Videre ønsker jeg å sammenligne hvordan de fem klassene som hadde størst og minst spredning i prestasjoner i klassen vurderte undervisningskvaliteten. Deretter en sammenligning av de to gruppene.

Deskriptiv analyse av de fem klassene med størst spredning

Tabell 4.11: En oversikt over de fem klassene som hadde størst spredning i prestasjoner i klassen på regne9.

Klasse	*Score*	Spredning	Mest enig i		Minst enig i	
5	54,56	11,58	control	4,14	confer	3,24
6	54,93	11,42	challenge	4,17	confer	3,40
29	55,99	11,33	control	4,21	captivate	3,12
37	49,65	11,39	consolidate	4,05	captivate	3,48
13	56,64	10,93	clarify	4,02	confer	3,40

Tabell 4.11 består av de fem klassene som hadde den største spredningen i prestasjoner på regne9, og vi ser at fire av fem klasser presterte over gjennomsnittet på regne9. Eneste måten disse klassene som presterte over gjennomsnittet kan ha fått en så høy *score*, er at den består av flere elever som presterte veldig høyt enn lavt. Dersom det hadde vært liten spredning med en gjennomsnittlig verdi rundt gjennomsnittet, ville flertallet i klassen også har prestert rundt gjennomsnittet. Tabellen viser også at alle disse fem klassene hadde en høy vurdering av det aspektet de var mest enige i. Derimot er det ingen tydelig sammenheng mellom hvilket aspekt de var mest enige i og hvor stor spredning de hadde. Verdiene på hva klassene vurderte seg minst enige i var også relativt høye. For de fem klassene med størst spredning opplevde de *confer* og *captivate* minst i undervisningen.

Deskriptiv analyse av de fem klassene med minst spredning

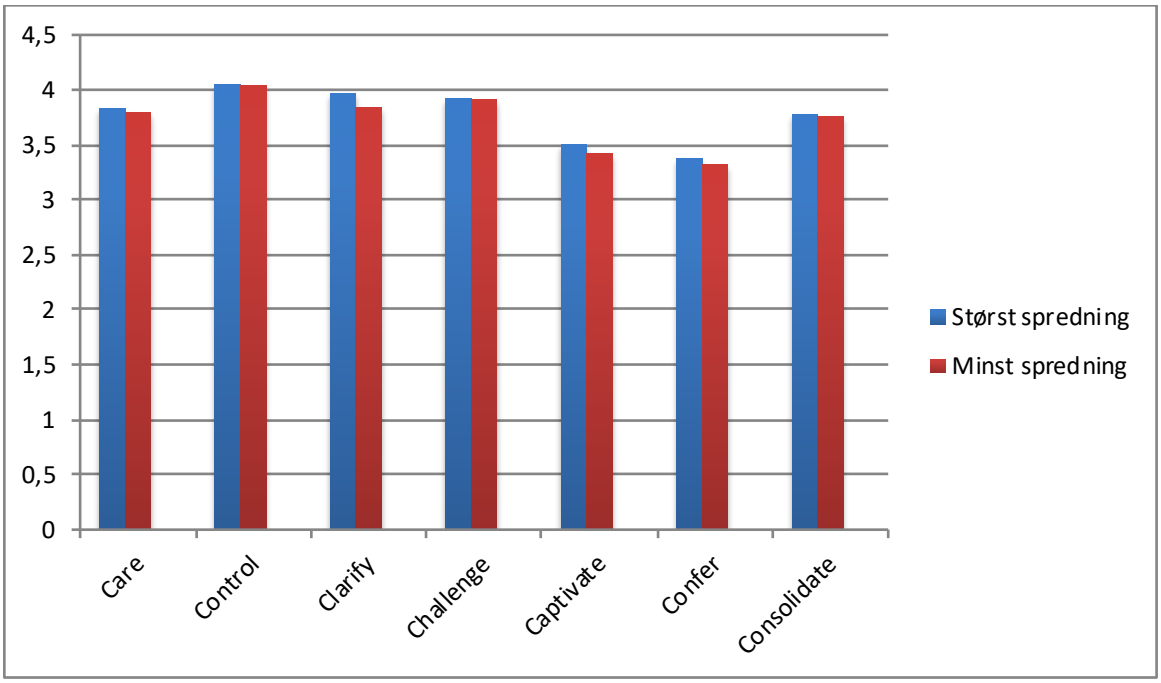
Tabell 4.12: En oversikt over de fem klassene som hadde minst spredning i prestasjoner i klassen på regne9.

Klasse	*Score*	Spredning	Mest enig i		Minst enig i	
10	51,78	5,37	control	3,74	captivate	2,69
2	67,06	6,37	control	4,21	confer	3,57
12	50,48	6,54	care	4,26	confer	3,72
1	45,35	6,62	challenge	4,21	captivate	3,58
31	56,31	7,11	control	4,01	captivate	2,81

Tabell 4.12 viser de fem klassene med minst spredning, som vil si at dette var de klassene som presterte jevnest i klassen. Tabellen viser at kun to av fem klasser presterte over gjennomsnittet på regne9. Vi ser at både klasse1 og klasse2 inngår i de fem med minst spredning. Vi vet at klasse2 var en veldig sterk klasse, og det at de hadde så liten spredning

betyr at klassen bestod av veldig mange sterke elever. Vi vet også at klasse1 var den svakeste klassen, og det at de hadde så liten spredning betyr at klassen bestod av flere svake elever enn sterke. Klasse10 skiller seg litt ut i hvor stor grad av enighet de var i vurderingen av aspektene og det viser seg at de var litt mindre enige i forhold til de andre fire klassene. Det varierer i hva klassene har vurdert seg mest og minst enige i. Igjen kommer *confer* og *captivate* dårligst ut. *Control* var det aspektet flertallet av disse fem klassene var mest enige i.

Sammenligning av de med størst og minst spredning i prestasjoner



Figur 4.3: En oversikt over de fem klassene med størst spredning og de med minst spredning og deres gjennomsnittsverdi til de syv aspektene.

Sammenligner vi de klassene med størst spredning i prestasjoner og de med minst spredning i prestasjoner, ser vi fra figur 4.3 at de vurderte aspektene ganske likt i forhold til hverandre. Det var ingen aspekter som skiller seg mye mer positivt eller negativt ut for noen av de to gruppene. Figuren viser at de klassene med størst spredning i liten grad var mer enige i alle de syv aspektene i forhold til de klassene med liten spredning. For begge gruppene var det *control* som kom best ut og *confer* og *captivate,* som kom dårligst ut av vurderingen.

4.3 Resultater fra den multiple regresjonsanalysen

I denne delen presenteres resultatene fra den multivariate regresjonsanalysen som ble utført i SPSS. Regresjonsanalysen forteller hvor god sammenheng det er mellom de uavhengige

variablene og den avhengige variabelen. Det første vi må se på er signifikansnivået som forteller hvor sikkert det er at det er en sammenheng mellom de variablene vi ønsker å se en sammenheng mellom. Lave signifikansverdier viser til høye signifikansnivåer, og desto mer sikker er sammenhengen. Vanligvis brukes bare signifikansnivåer på 0,05, som vil si 5% usikkerhet. Men for å bli litt mer åpen, godtas det i dette tilfellet også signifikansnivå på 0,10, som vil si 10%. Det betyr at signifikansnivå på 0,10 og lavere, betyr at vi kan tro på de sammenhengene vi finner. Videre må vi se på den standardiserte koeffisienten (regresjonskoeffisienten) som forklarer styrken på sammenhengen mellom variablene. Sammenhengen kan enten være positiv eller negativ. Det siste som er interessant å se på er "adjustet R square" (R^2) som viser hvor stor forklaringskraft regresjonsmodellen har på den avhengige variabelen. Høye verdier viser til god forklaring. Det er totalt blitt utført åtte regresjonsanalyser.

Regresjonsanalyse av de syv aspektene opp mot resultatene på regne9

Dette er en regresjonsanalyse for å undersøke forklaringen de uavhengige variablene (c*are, control, clarify, challenge, captivate, confer, consolidate*) har på den avhengige variabelen (elevenes prestasjoner på regne9). Denne regresjonsanalysen ble gjennomført på hele utvalget (N=991).

Tabell 4.13: Regresjonsanalyse av de syv aspektene opp mot resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	-0,056	0,022 = 2%
Control	0,000	**0,165**	
Clarify	0,072	**-0,094**	
Challenge	i.s.	-0,039	
Captivate	0,017	**0,113**	
Confer	i.s.	-0,029	
Consolidate	i.s	0,018	

Tabell 4.13 viser at det kun var *control*, *clarify* og *captivate* som hadde en signifikant sammenheng med resultatene på regne9. *Control* og *captivate* hadde en positiv sammenheng, mens *clarify* hadde en negativ sammenheng. Denne modellen forklarer imidlertid kun 2% av variansen på regne9. Det betyr at 98% av variansen på regne9 forklares av andre faktorer.

Regresjonsanalyse av de syv aspektene + regne8 opp mot resultatene på regne9

En regresjonsanalyse for å undersøke forklaringen mellom de uavhengige variablene (regne8, *care, control, clarify, challenge, captivate, confer, consolidate*) og den avhengige variabelen (regne9). Denne regresjonsanalysen ble også gjennomført på hele utvalget (N=991).

Tabell 4.14: Regresjonsanalyse av de syv aspektene og resultatene fra regne8 opp mot resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	0,011	0,808 =81%
Control	i.s.	0,014	
Clarify	0,042	**-0,047**	
Challenge	i.s.	-0,036	
Captivate	0,009	**0,055**	
Confer	i.s.	0,018	
Consolidate	i.s.	0,011	
Regne8	0,000	**0,896**	

Tabell 4.14 viser at regne8 har en stor signifikant sammenheng med prestasjonene på regne9 og at denne modellen kan forklare hele 81% av elevenes prestasjoner. I forhold til modellen vist i tabell 4.13, viser tabell 4.14 nå at det kun var *clarify, captivate* og *regne8* som hadde en signifikant sammenheng med resultatene på regne9. En kan derfor anta at elevenes forkunnskaper er det som har størst påvirkning på prestasjonene på regne9.

4.3.1 Regresjonsanalyse av de høyt og lavt presterende klassene

To regresjonsanalysre av den første *casen*, som består av de høyt og lavt presterende klassene. Utvalget (N) i disse to gruppene varierer. Først presenteres regresjonsanalysen av de høyt presterende klassene, deretter av de lavt presterende klassene.

De som presterte over gjennomsnittet (Regne9 >= 54,53, N=470)

For den første gruppa som består av de som presterte høyt, viser tabell 4.15 at både *control*, *clarify*, *captivate* og consolidate hadde en signifikant sammenheng på regne9. Men også her hadde clarify en negativ sammenheng, som vil si at jo bedre elevene presterte desto mindre fornøyde var de med clarify. Igjen forklarer denne modellen i underkant av 3% av variansen i resultatene på regne9.

Tabell 4.15: Regresjonsanalyse med fokus på de klassene som presterte over gjennomsnittet på regne9 og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	-0,089	0,026 = 3%
Control	0,018	**0,125**	
Clarify	0,001	**-0,249**	
Challenge	i.s.	0,019	
Captivate	0,055	**0,131**	
Confer	i.s.	0,001	
Consolidate	0,059	**0,139**	

De som presterte under gjennomsnittet (Regne9 <= 54,53, N=521)

For den gruppa som består av de klassene som presterte lavt viser tabell 4.16 at *care* og *clarify* hadde en signifikant sammenheng med resultatene på regne9. Hvor *care* hadde en negativ sammenheng og *clarify* en positiv. I de klassene som presterte lavt opplevde de det som positivt at læreren brukte tid på å forklare. Modellen har kun 2% forklaringskraft på resultatene på regne9.

Tabell 4.16: Regresjonsanalyse med fokus på de klassene som presterte under gjennomsnittet på regne9 og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	0,068	**-0,129**	0,017 = 2%
Control	i.s.	0,011	
Clarify	0,004	**0,204**	
Challenge	i.s.	-0,023	
Captivate	i.s.	0,039	
Confer	i.s.	-0,090	
Consolidate	i.s.	-0,057	

4.3.2 Regresjonsanalyse av de med størst og minst fremgang i prestasjoner.

To regresjonsanalyser av den andre *casen*; de klassene med størst og minst fremgang. Først presenteres regresjonsanalysen av de klassene med størst fremgang, deretter for de med minst fremgang.

De som hadde størst fremgang (FRG89 >= 3,57, N=494)

Tabell 4.17: Regresjonsanalyse med fokus på de klassene med større fremgang enn gjennomsnittet og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	-0,016	0,021= 2%
Control	0,004	**0,158**	
Clarify	0,041	**-0,154**	
Challenge	i.s.	-0,091	
Captivate	0,021	**0,153**	
Confer	i.s.	-0,063	
Consolidate	i.s.	0,064	

For de som hadde størst fremgang, viser tabell 4.17 at det var *control*, *clarify* og *captivate* som hadde en signifikant sammenheng på regne9, hvor det kun var *control* og *captivate* som hadde en positiv sammenheng. Denne modellen forklarer imidlertid igjen kun 2% av variansen i resultatene på regne9.

De som hadde minst fremgang (FRG89 <= 3,57, N=497)

Tabell 4.18: Regresjonsanalyse med fokus på de klassene med mindre fremgang enn gjennomsnittet og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	0,013	**-0,175**	0,057 = 6%
Control	0,000	**0,246**	
Clarify	i.s.	0,004	
Challenge	i.s.	0,074	
Captivate	i.s.	0,057	
Confer	i.s.	0,006	
Consolidate	i.s.	-0,077	

I den gruppen som hadde minst fremgang så viser tabell 4.18 at *care* og *control* hadde en signifikant sammenheng på regne9. *Control* hadde en høyest signifikant sammenheng på 0%, som vil si at vi med stor sikkerhet kan si at *control* hadde en sammenheng med regne9 for de elevene som hadde liten fremgang på regne9. *Care* hadde en negativ sammenheng, det vil i dette tilfellet si, desto større fremgang elevene fikk jo mindre opplevde de behovet for at læreren skulle bry seg om dem. Denne modellen har derimot en litt større forklaringskraft på 6%.

4.3.3 Regresjonsanalyse av de med størst og minst spredning i prestasjoner

To regresjonsanalyser av den tredje *casen*; de klassene med størst og minst spredning. Først presenteres regresjonsanalysen av de klassene med størst spredning. Deretter regresjonsanalysen av de med minst spredning. Også her varierer utvalget i de to gruppene.

De med størst spredning (Reg9SD >= 9,13, N=462)

Tabell 4.19: Regresjonsanalyse med fokus på de klassene med større spredning i prestasjoner enn gjennomsnittet og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	-0,138	0,026 = 3%
Control	0,001	**0,171**	
Clarify	0,089	**-0,130**	
Challenge	i.s.	0,059	
Captivate	i.s.	0,107	
Confer	i.s.	-0,015	
Consolidate	i.s.	0,043	

Tabell 4.19 viser at *care*, *control* og *clarify* hadde en signifikans sammenheng med resultatene på regne9 for de klassene med stor spredning i prestasjoner. Det var kun *control* som hadde en positiv sammenheng. Modellen forklarer kun 3% av resultatene på regne9.

De med minst spredning (Reg9SD <= 9,13, N=529)

Tabell 4.20: Regresjonsanalyse med fokus på de klassene med mindre spredning i prestasjoner enn gjennomsnittet og hvilken sammenheng det var mellom de syv aspektene og resultatene på regne9.

Aspekt	Signifikans	Regresjonskoeffisient	*Adjustet R square*
Care	i.s.	-0,028	0,023 = 2%
Control	0,003	**0,153**	
Clarify	i.s.	-0,067	
Challenge	0,020	**-0,145**	
Captivate	0,022	**0,151**	
Confer	i.s.	-0,057	
Consolidate	i.s.	0,049	

For de klassene med minst spredning viser tabell 4.20 at *control*, *challenge* og *captivate* hadde en signifikant sammenheng, hvor både *control* og *captivate* hadde en positiv

sammenheng, mens *Challenge* hadde en negativ sammenheng. Modellen forklarer igjen kun 2%.

4.3.4 Oppsummering av resultatene fra den multiple regresjonsanalysen

En oppsummering av disse regresjonsanalysene er at ingen av disse modellene hadde en veldig stor forklaringskraft på hvordan elevene presterte på regne9. Det var kun modellen vist i tabell 4.14, hvor elevenes forkunnskaper ble lagt til, at modellen forklarte mye av årsaken til resultatene på regne9. Det betyr at elevenes forkunnskaper hadde en stor forklaringskraft på hvordan elevene presterer på regne9. Hvis man fokuserer på hvor mye aspektene forklarer så var det kun *control* som oftest hadde en signifikant sammenheng med regne9. Det betyr at vi kan tro på de sammenhenger vi finner mellom *control* og regne9. Tidligere studier har også funnet ut at det kun var *control* som hadde en signifikant sammenheng med elevenes prestasjoner. Videre er det *captivate* som ser ut til å ha hatt en signifikans sammenheng med elevenes prestasjoner i de tilfellene hvor det var en høyt presterende klasse, en klasse med stor fremgang og en klasse med et jevnt prestasjonsnivå. I alle disse tilfellene var det en positiv sterk sammenheng.

5 Drøfting av forskningsfunn

I dette kapittelet drøftes mine funn opp mot teori for å svare på problemstillingen "hvilke sammenhenger er det mellom elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget?". Kapittelet starter med en vurdering av hvor godt de syv aspektene for god undervisningskvalitet (*care, control, clarify, challenge, captivate, confer og consolidate*), faktisk måler god undervisningskvalitet i matematikk, i kontekst av LISA-studien. Videre drøftes oppgavens problemstilling med fokus på høyt og lavt presterende klasser, klasser med stor og liten fremgang på regneprøven, og klasser med stor og liten spredning i prestasjoner. Til slutt drøftes mine funn opp mot funnene til tidligere empiriske studier.

5.1 Måling av undervisningskvalitet

På bakgrunn av at de teoretiske tilnærmingene om at god undervisningskvalitet i hovedsak handler om god klasseromsledelse, støttende undervisning og kognitiv aktivering, valgte jeg å knytte de syv aspektene fra spørreskjemaet, Tripod Survey, inn i disse tre. Teorien trekker frem betydningen av disse og de er alle åpne begreper som forklarer mer enn bare én ting ved undervisningen. Klasseromsledelse handler ikke bare om hvor flink en lærer er til å lede og styre klassen, men også om at læreren klarer å skape god orden og et godt miljø i klassen. Dette vil være en lærer som har god struktur i hvordan undervisningen utføres og er bygd opp, en som både har kontroll på elevene og faget. Handler også om en som er flink til å utnytte tiden for å gi elevene best læringstid og er flink til å håndtere forstyrrelser. LISA-studiet har helt klart blitt påvirket av hva tidligere teori sier om god undervisningskvalitet. Dette har hatt en påvirkning på hvordan de ønsket å forske videre på god undervisningskvalitet i matematikk og hvordan dette påvirker elvenes prestasjoner i faget. Likevel valgte de å bruke et spørreskjema som bestod av syv aspekter som skulle måle undervisningskvaliteten, istedenfor de tre teoriene trekker frem. På bakgrunn av dette valgte jeg i kapittel 2.2 å trekke disse syv aspektene som brukes i spørreskjemaet, inn i de tre teoriene trekker frem. Siden jeg valgte å trekke disse sammen og ikke bare fokusere på hver av de syv aspektene hver for seg, gjorde jeg en konfirmerende faktoranalyse. Jeg ville vurdere om den beste metoden for å måle god undervisningskvalitet i matematikk var å forkorte modellen til en tredimensjonal faktormodell med klasseromsledelse, støttende undervisning og kognitiv aktivering, eller å gå

for en syvdimensjonal faktormodell som bestod av de syv aspektene, *care, control, clarify, challenge, captivate, confer* og *consolidate*.

Mine funn fra den konfirmerende faktoranalysen bekrefter hva Kuhfeld (2017) og Wallace et al. (2016) fant ut i sine faktoranalyser, med data fra amerikanske elever. De fant ut at den modellen som best måler god undervisningskvalitet i matematikk var en faktormodell bestående av alle disse syv aspektene, istedenfor å trekke de sammen til færre aspekter. Desto flere aspekter faktormodellen består av, desto grundigere målinger får vi. De faktorladningene vi fikk gjennom den konfirmerende faktoranalysen i Mplus, viste både sterke og gode sammenhenger mellom de fleste spørsmålene og det tilhørende aspektet. Det betyr at de spørsmålene som er blitt stilt i LISA-studien, i stor grad faktisk måler det de skulle måle. Med andre ord, vi kan ha tiltro til svarene vi får fra spørreskjemaet og elevenes oppfatning av undervisningen og hvordan denne kan knyttes opp mot gode prestasjoner.

Likevel viste korrelasjonsanalysen at korrelasjonene mellom alle disse syv aspektene var ganske høye. Det var kun aspektet *control* som hadde en litt mindre korrelasjon med de andre aspektene. Dette fant også Kuhfeld (2017) ut, hun gjorde derfor en videre eksplorativ faktoranalyse av en bifaktormodell bestående av *control* og de andre seks aspektene som en egen gruppe. Korrelasjonsanalysen hennes viser også at de andre aspektene korrelerte i større grad enn de gjorde med *control*. Desto sterkere aspektene korrelerer, desto mer måler de det samme. I denne masteroppgaven har jeg ikke gjennomført tilsvarende analyser som Kuhfeld (2017) gjorde, siden det ville blitt for omfattende. Men en forsiktig sammenligning av den syvdimensjonale faktormodellen og den tredimensjonale faktormodellen, viste at den syvdimensjonale faktormodellen passet bedre med dataene våre, enn den tredimensjonale.

I kapittel 2.2 trakk jeg som nevnt en sammenheng mellom *control*, *consolidate* og *clarify* for å beskrive god klasseromsledelse. Korrelasjonsmodellen viste at *consolidate* og *clarify* hadde en relativ høy korrelasjon på 0,64. Modellen viste også at de begge korrelerte like sterkt med *care*, *challenge* og *confer*, som skulle etter min tolkning beskrive andre faktorer enn klasseledelse. Derimot korrelerte de lavt med *control* korrelerte lavt. Det vil si at aspektene korrelerer høyt med andre enn de aspektene som jeg knyttet sammen i teorien. Likevel er det ikke ønskelig at aspektene skal måle det samme, siden de skal beskrive hver sin del av klasseromsledelse. Dette bekrefter mine funn som sier at *control* har en litt større påvirkning på prestasjoner. Dersom jeg ikke hadde valgt å trekke noen av aspektene sammen, ville jeg kanskje ikke funnet denne påvirkningen. For at aspektene ikke skal måle det samme er det

derfor mer positivt med lave korrelerte verdier. Faktoranalysen var basert på hvordan alle disse 46 matematikklassene svarte på spørreskjemaet om god undervisningskvalitet. Det kan derfor tenkes at elevene opplever viktigheten av de syv aspektene ulikt i forhold til hverandre og hva forskerne, som benytter seg av dettes spørreskjemaet, ønsker å måle. For eksempel kan det være at elevene opplever det som enklere å vurdere hva som inngår i *control,* istedenfor de syv andre aspektene. Kanskje elevene synes det er enklere å relatere seg til *control*. Det at disse seks andre aspektene korrelerer såpass høyt med hverandre, kan også tyde på at det vil være vanskeligere for elevene å skille de fra hverandre i like stor grad som de klarer med *control*. *Control* handler mer om hvordan klassen fungerer i klasserommet og hvordan de føler at læreren håndterer dette. De seks andre aspektene handler i større grad om egenskaper ved læreren og hvordan læreren underviser bort.

5.2 Sammenhenger mellom undervisningskvalitet og prestasjoner

For å svare på min problemstilling ”hvilke sammenhenger er det mellom elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget” valgte jeg å se nærmere på hvordan de klassene som presterte høyt og de som presterte lavt, hvordan de klassene som hadde stor og liten fremgang på den nasjonale prøven i regning og hvordan de klassene som hadde et jevnt og ujevnt prestasjonsnivå innad i klassen vurderte undervisningskvaliteten til deres lærer i matematikk. Med denne informasjonen vil jeg få en dypere forståelse for de slutningene jeg trekker rundt de sammenhengene jeg ser.

Før jeg begynner å diskutere mine funn, er det en ting som må avklares for deg som leser. Det er at videre i kapittelet, så diskuteres blant annet hvor fornøyde en klasse var med lærerens undervisning. Med mer fornøyde elever mener jeg at en klasse fikk en høyere gjennomsnittsverdi enn 3 på de syv aspektene, og svært fornøyde elever med en gjennomsnittsverdi høyere enn 4. Desto høyere gjennomsnittsverdi klassene fikk desto mer enige og fornøyde var elevene med undervisningen.

5.2.1 Høyt presterende klasser vs. lavt presterende klasser

Generell analyse

I denne studien består de høyt presterende klassene av totalt 22 klasser (470 elever) som presterte over gjennomsnittet på den nasjonale prøven i regning på 9. trinn. De andre 24 klassene (521 elever) som presterte under gjennomsnittet, vil videre bli kalt for de lavt presterende klassene. Utvalget er litt større for de klassene som har prestert under gjennomsnittet, og vurderingene fra denne gruppen vil derfor i noe grad være påvirket av flere elevstemmer . Jeg kan derfor ikke måle disse sammenhengene like nøyaktig som om det hadde vært like mange elever i hver gruppe.

Teorien sier at god undervisning blant annet er knyttet til gode prestasjoner hos elevene (kap. 2.1). Desto bedre undervisning, jo bedre utfall er det for elevene. Derfor kan en tenke seg at de klassene som presterer høyt er mer fornøyde med undervisningen til læreren enn de klassene som presterer lavt. Mine funn fra den deskriptive analysen skal bidra til å gi en større forståelse for eventuelle sammenhenger mellom god undervisning og gode prestasjoner. Denne analysen viser at det ikke var store forskjeller i hvordan de klassene som presterte høyt vurderte undervisningskvaliteten til læreren i forhold til de klassene som har prestert lavt (se figur. 4.1, side 39). Figuren viser at det var svært liten forskjell. Faktisk hadde disse to gruppene ganske lik oppfatning av lærerens undervisning. Både de høyt presterende og de lavt presterende klassene var enige om at det var *captivate* og *confer* som forekom minst i undervisningen, i forhold til de fem andre aspektene. At de forekom minst trenger ikke å bety at de ikke forekom i det hele tatt. Faktisk så viser figur 4.1 (side…) at de gjennomsnittlige verdiene til alle disse syv aspektene fra begge disse to gruppene, har en verdi som alltid var over 3. Dette er positivt, siden skalaen går fra verdien 1 til 5. Som nevnt tidligere tilsvarer dette elever som i gjennomsnitt generelt var fornøyde med sin lærers undervisning i matematikk. Figur 4.1 (side 39), viser at det faktisk var de lavt presterende klassene som var mest fornøyde med undervisning enn de høyt presterende. De lavt presterende klassene hadde en litt høyere gjennomsnittsverdi på fem av de syv aspektene, i forhold til de høyt presterende klassene. Dette stemmer ikke med min hypotese om at de klassene som presterte høyt, var mer fornøyde med undervisningen i forhold til de som presterte lavt. Kanskje det er fordi de høyt presterende klassene stiller litt høyere krav til læreren enn de som presterer lavt. Trolig er det faglige nivået litt svakere i de klassene som presterte lavt. Dermed har kanskje læreren

en bedre forståelse for faget enn elevene, mens det for sterke elever krever en mer faglig sterk lærer for å tilfredsstille deres forventninger til en god matematikklærer.

De lavt presterende klassene vurderte høyere enn de høyt presterende på disse fem aspektene; *care, clarify, challenge, confer* og *consolidate.* For de klassene som presterte best var det kun *control* og *captivate* de hadde vurdert høyest. Grunnen til at de lavt presterende klassene i større grad opplevde *care* i undervisningen, kan være at de har et større behov for en lærer som er flink til å se og høre hver enkelt elev, enn de sterke elevene har. Svake elever har kanskje en tendens til å føle at de ikke forstår noen ting, og dermed lettere gir opp. I slike tilfeller vil det være ekstra viktig med en lærer som er flink til å se eleven og få de opp og frem. Sterke elever har kanskje færre slike tanker, og læreren i disse klassene fokuserer derfor kanskje mer på det faglige, fremfor det personlige. Likevel betyr ikke det at de sterke elevene har et mindre behov for å bli sett av læreren. Kanskje dette er grunnen til at de har lavere verdier på *care*. Kanskje læreren "glemte" å gi elevene denne oppmerksomheten siden undervisningen kanskje var preget av et større faglig engasjement fra elevene. At de sterke klassene var mer fornøyde med *control* enn de svake klassene, kan kanskje forklares ved at de sterke elevene ofte var mer motiverte i faget og hadde et større ønske om å følge med i timen og lære. De svake elevene faller nok lettere ut ved felles gjennomgang av nytt stoff i undervisningen. Dette kan kanskje bidra til å skape uro i timen og som læreren må bruke tid på å få ordnet opp i. Tilsvarende for aspektet *captivate*. De høyt presterende klassene bestod nok av flere sterke elever som mest sannsynlig oftere kjente på en mestringsfølelse og dermed opplevde faget som mer interessant og morsomt. Som nevnt er ikke forskjellen stor mellom de høyt presterende og de lavt presterende klassene når det gjelder hvordan de vurderte lærerens undervisningskvalitet. Mye tyder derfor på at det er andre forklaringer enn motivasjon som forklarer hvorfor de presterte som de gjorde.

At de høyt presterende klassene i større grad opplevde *challenge* i undervisnignen kan kanskje komme av at disse var mer faglig sterke. Det kan tenke seg at de oftere ble utfordret på et høyere nivå, siden de hadde en større forståelse for faget enn de svake klassene. Figur 4.1 (side 39), viser likevel at begge gruppene i stor grad opplevde at de ble utfordret i undervisningen. Samtidig skal det kanskje mindre til for at svake elever opplever en oppgave som utfordrende. En oppgave som oppleves som utfordrende for en sterk elev, vil da mest sannsynlig også oppleves som utfordrende for en svak elev. Det motsatte trenger ikke å være tilfellet. Kanskje dette er en av grunnene til at de høyt presterende klassene i mindre grad enn

de lavt presterende klassene opplevde at læreren utfordret dem i undervisningen. Videre opplevde de lavt presterende i større grad *clarify,* enn de høyt presterende. De lavt presterende klassene hadde kanskje et større behov for en tydelig og strukturert undervisning, med klare mål for timen. Et større behov for å forstå hva som skulle læres og hvordan dette skulle gjøres. *Clarify* handler i stor grad om at læreren klarer å fjerne elevenes forvirring, og svake elever opplever kanskje oftere enn sterke elever å være forvirret i faget. Det er derfor kanskje et større behov for at læreren må være grundig og nøye i hvordan faget blir undervist på. Videre har sterke elever et mindre behov for at læreren bruker mye tid på å koble sammen forkunnskap med ny kunnskap, siden de mest sannsynlig tar til seg ny kunnskap raskere enn svake elever. Læreren trenger derfor kanskje ikke å bruke like lang tid på dette, i forhold til elever som har større vansker med å forstå og er mer avhengig av god tid. Likevel antyder figur 4.1 (side 39), at lærerne i begge to gruppene, generelt var flinke til å trekke fagstoffet sammen. Uansett hvilket nivå eleven ligger på, er det viktig med en dyp forståelse for å være bedre rustet til å møte nye problemer senere på ungdomsskolen.

For begge gruppene var det *confer* som kom dårligst ut. Likevel var begge gruppene relativ fornøyde med også dette aspektet. *Confer* handler blant annet om at elevenes ideer blir hørt av læreren og at det er et trygt og godt klassemiljø, hvor det er rom for å dele ideer og meninger med hverandre. Man kan kanskje tenke seg at elever som er litt mer faglig sterke, har større selvtillit i faget og dermed er kanskje litt mer aktive i undervisningen. For at elever skal komme med innspill, må de ha en viss forståelse av innholdet og hva som egner seg og ikke i undervisningen. Det kan muligens være at svakere elever i større grad enn sterke elever, oftere kommer med forslag til lærerens undervisning som ikke er faglig relevante. Forslagene blir derfor kanskje ikke hørt på og tatt i bruk av læreren. Det kan også være at elevene ikke opplever dette aspektet som like relevant for hva som skal inngå i god undervisning fra læreren. Kanskje flere har et syn på at undervisningen er noe læreren skal ha ansvaret for og at elevene skal følge med og ikke komme med egne forslag. De forventer at læreren vet bedre enn elevene. Særlig svake elever vil kanskje oppleve at læreren vet og kan best, mens sterke elever kanskje har et større behov for å utfordre læreren. Dette kan være en av forklaringene på hvorfor de høyt presterende klassene vurderte *confer* lavest, men ikke en like god forklaring på hvorfor de svake klassene også gjorde det.

Den deskriptive analysen viser som nevnt ingen tydelige sammenhenger mellom hvordan klassene vurderte lærerens undervisningskvalitet og hvordan de selv presterte. Det tyder på at

grad av enighet rundt lærerens undervisningskvalitet ikke kan kobles opp mot hvordan de presterte. Derimot kan regresjonsanalysen fortelle mer om hvordan hver av de syv aspektene kan kobles opp mot prestasjonsresultatene på regneprøven i 9. trinn. Det ble utført to regresjonsanalyser for å undersøke denne sammenhengen, både for de høyt presterende klassene og de lavt presterende klassene. Hos de høyt presterende klassene var det kun *control* (0.125), *captivate* (0.131) og *consolidate* (0.139) som hadde en signifikant positiv sammenheng med elevenes prestasjoner på regne9. *Clarify* (-0.249) hadde en signifikant negativ sammenheng. Denne regresjonsmodellen av de høyt presterende klassene kan faktisk bare forklare 3% av sammenhengen mellom aspektene for god undervisningskvalitet og gode prestasjoner. Regresjonsanalysen for de lavt presterende klassene viser at både *care* og *clarify* hadde en signifikant sammenheng med resultatene på regne9. Et interessant funn fra denne regresjonsanalysen presentert i tabell 4.16 (side 47), er at denne modellen er den eneste som viser at *clarify* hadde en signifikant positiv sammenheng med resultatene på regne9. For de høyt presterende elevene hadde nemlig *clarify* en signifikant negativ sammenheng. Det skal bety at jo mer enige elevene var i opplevelsen av at læreren hadde en tydelig undervisning, desto svakere presterte de. Det vil si at læreren i disse klassene mest sannsynlig brukte mer tid på å lære bort på en enkel og forståelig måte, i forhold til de klassene som presterte høyt. En kan tenke seg at elever som er mer faglig sterke har et større ønske om å jobbe på egenhånd og bli utfordret med oppgaver. Elever som er litt faglig svakere har kanskje et større behov for å bli vist av lærer hvordan de skal bruke ulike prosedyrer for å løse oppgaver. Som nevnt i kapittel 2.2, vet vi at *control* har en moderat sammenheng med prestasjoner, og regresjonsanalysen for de lavt presterende klassene bekrefter denne sammenhengen. Dette tilfellet viser nemlig ikke at *control* hadde en sammenheng med svake prestasjoner. Regresjonsanalysen for de høyt presterende klassene viser som sagt en signifikant positiv sammenheng. Siden modellen kun kan forklare 3%, bekrefter det hva vi allerede vet fra Klette (2013), nemlig at klasseromsledelse – som jeg har knyttet *control* til - alene ikke kan forklare elevenes prestasjoner.

Dybdeanalyse - de fem sterkeste og svakeste klassene

Jeg valgte som nevnt å gå mer i dybden av de klassene som presterte aller best (5 stk.) og de som presterte aller dårligst (5 stk.), av mitt utvalg på 46 klasser. For å få et overblikk over hva som kjennetegner hvordan disse klassene opplevde undervisningen. Gjennom denne analysen fant jeg ut at det ikke var noen systematisk sammenheng i graden av enighet de fem beste

klassene hadde vurdert i spørreskjemaet. De hadde nemlig svært ulik grad av enighet. For eksempel hadde den klassen som presterte aller best (klasse2), vurdert relativt mer positivt enn den klassen som presterte nest best (klasse7). Det viste seg at begge klassene vurderte *control* høyest, der klasse2 hadde en gjennomsnittsverdi på 4,21, mens klasse7 en verdi på 3,91. Altså viste det seg at de var enige i hvilket aspekt de følte inntraff oftest i undervisningen. Elevene i klasse2 var i større grad var mer enige, enn elevene i klasse7. Med andre ord var det ingen tydelig sammenheng mellom grad av enighet om undervisningskvaliteten og hvordan de presterte. Tilsvarende for de som presterte lavt, hvor de så ut til å være like fornøyde som de høyt presterende klassene.

Hos de fem klassene som presterte aller best var det heller ingen tydelige sammenhenger i hvordan de presterte og hva de vurderte seg mest og minst enige i. Resultatene fra den deskriptive analysen viser at disse fem klassene hadde en ulik oppfatning. Noen av dem hadde vurdert *control* høyest, og andre vurderte *consolidate* høyest. Dette viser godt hva vi allerede har diskutert, at det ikke var noen aspekter som i større grad skilte seg ut, basert på hvor stor grad av enighet elevene hadde. Det ser ut som at disse elevene opplevde at læreren oppfylte påstandene om god undervisning. Hos de fem klassene som presterte aller dårligst var det mer ujevnt hvilket aspekt som var blitt vurdert høyest. Hos disse fem klassene var det både *challenge, control, consolidate* og *clarify.* Verken hos de fem sterkeste eller de fem svakeste klassene var det noe tydelig mønster og sammenheng med hvordan de vurderte lærerens undervisning og hvordan de presterte. Selv om det var flere aspekter som ble vurdert høyest av de fem svakeste klassene, trenger ikke det å bety at de var mindre fornøyde med undervisningen enn de fem sterkeste klassene. Selv om det er ulike oppfatninger av hvilket aspekt som forekom oftest, var det likevel de gjennomsnittlige verdiene som var det mest avgjørende for å fortelle hvordan disse elevene opplevde undervisningen. De fem sterkeste og de fem svakeste klassene skiller seg ikke noe mer ut i fra hvordan de to gruppene har vurdert. Dette tyder på at alle disse klassene hadde en generell lik oppfatning av sin lærers undervisning, og at det ikke var et aspekt som bedre kunne forklare gode prestasjoner enn dårlige prestasjoner.

5.2.2 Stor fremgang vs. liten fremgang på regneprøven

Generell analyse

I denne studien var det 23 klasser (494 elever) som hadde større fremgang enn gjennomsnittet og 23 klasser (497 elever) som hadde mindre. Utvalget i disse to gruppene er nesten like store og resultatene fra denne analysen er derfor mer pålitelige. Siden god undervisning defineres blant annet som det som fører til forbedret progresjon hos elevene, er det interessant å se hvordan de klassene som hadde størst progresjon vurderte lærerens undervisningskvalitet. Siden disse klassene i teorien skal oppleve god undervisning, ønsket jeg å se hvordan de i praksis opplevede undervisningen. Er det en sammenheng mellom disse klassene, i forhold til de klassene som hadde liten progresjon?

Den deskriptive analysen av de klassene som hadde større fremgang og mindre fremgang enn gjennomsnittet vises blant annet i figur 4.2 (side 42). Denne figuren viser at elevene i de klassene med størst fremgang, generelt var litt mer enige i påstandene om lærerens undervisning i forhold til de klassene med minst fremgang. Dette kan tyde på at de med stor fremgang også var mer fornøyde med lærerens undervisning, i forhold til de med liten fremgang. Noe som er naturlig, siden man kan tenke seg at bedre prestasjoner fører til mer motiverte elever, som igjen fører til at de opplever at de får en god undervisning. Men samtidig er denne forskjellen så minimal, at man kan ikke bekrefte at god undervisningskvalitet kun fører til økt progresjon hos elevene. Et interessant funn var at de klassene med stor fremgang, hadde vurdert seg mye mer enige i *captivate* enn de med liten fremgang. Kanskje dette var en av årsakene til at disse elevene fikk så stor progresjon. At flere av disse elevene opplevde undervisningen som interessant og morsom, og dermed fikk et større ønske om å lære. Man kan tenke seg at det for klasser som i mindre grad opplever faget som morsomt, også yter mindre innsats og dermed utvikler seg i mindre grad enn de som er mer motiverte i faget. Videre ser man at *care* og *consolidate* var de to aspektene som kom aller best ut av analysen. De gjennomsnittlige verdiene her var betydelig høyere hos de klassene med stor fremgang, fremfor de med liten fremgang. Kanskje dette betyr at lærerne i de klassene med stor fremgang var litt flinkere til å se hver enkelt elev og oppmuntre dem. Samtidig må de ha vært flinke til å knytte fagstoffet sammen slik at flere har forstått mer enn hva de forstod før. Det ser også ut som at klassene med størst fremgang, i større grad opplevde at læreren ba dem om å reflektere rundt egen forståelse. Gjennom refleksjon får

kanskje eleven et tydeligere bilde over hva de kan. Dersom elevene blir bedt om å reflektere sammen med hverandre, kan det føre til at elevene får utvekslet kunnskap og ideer som kanskje kan bidra til at elever også lærer av hverandre.

I regresjonsanalysen av de klassene med størst fremgang (se tabell 4.17, side 48), kan kun 2% av sammenhengen mellom de syv aspektene og hvordan det påvirket prestasjonene på regne9 forklares. Analysen viser at *control* (0.158) og *captivate* (0.153) hadde en signifikant positiv sammenheng med prestasjonene på regne9. Og *clarify* (-0.154) hadde en signifikant negativ sammenheng. I dette tilfellet betyr det altså at jo mindre disse elevene opplevde at læreren brukte mye tid på å forklare og klargjøre mål for timen, desto bedre presterte de. De klassene som opplevde at læreren brukte mye unødvendig tid, følte kanskje ikke at de fikk god nok tid til å utvikle seg. Derimot forklarer regresjonsanalysen av klassene med minst fremgang, hele 6% av sammenhengen mellom de syv aspektene og hvordan disse klassene har prestert på regne9. Modellen viser at det kun var *control* (0.246) som hadde en signifikant positiv sammenheng med prestasjonene til disse klassene. I dette tilfellet betyr det at jo mer elevene var enige i påstandene om *control*, desto dårligere fremgang fikk de. Dette gir liten mening, siden *control* viser seg å være det aspektet som skal ha størst sammenheng med prestasjoner. Man kunne derfor tenke seg at lærere som har god kontroll på klassen får til mer effektiv læringstid, slik at elevene får mulighet til å lære mer.

Dybdeanalyse - de fem klassene med størst og minst fremgang

Den dypere analysen av de fem klassene som hadde den største fremgangen, viste ingen tydelige sammenhenger i hvordan de vurderte undervisningen og hvor stor fremgang de fikk. Den deskriptive analysen viser at det var *control*, *consolidate* og *care* disse fem klassene vurderte høyest. Også blant disse klassene var det ulik grad av enighet, som også ble eksemplifisert over. Dette var fem klasser som hadde stor fremgang på den nasjonale prøven fra 8. til 9. trinn. Fire av disse fem klassene presterte over gjennomsnittet på den siste regneprøven. Det vil si at fire av disse fem klassene var høyt presterende klasser. Resultatene fra 8. trinn viser at det kun var to av disse fem klassene som presterte over gjennomsnittet på den første regneprøven. Det betyr at de har klart å forbedre seg såpass mye, at flesteparten året etter presterte høyere enn gjennomsnittet. De fem klassene som hadde minst fremgang var mest enige i *consolidate, control, clarify* og *challenge.* En sammenligning med de fem klassene som hadde størst fremgang, viser at ingen av disse klassene med minst fremgang

vurderte *care* høyest. Kanskje noe av grunnen til den lave fremgangen var at disse elevene i mindre grad opplevde en lærer som brydde seg om dem og som hjalp dem fremover faglig. Dette viser godt hvor viktig det er med en lærer som ønsker å motivere elevene. For å klare å forbedre seg, er det viktig å føle seg sett og hørt. Som nevnt i kapittel 2.2, bidrar gode relasjoner blant annet til bedre læringsklima. Hva som inngår i *care,* handler i stor grad om relasjonen mellom lærer og elev. Igjen fører et godt læringsklima til gode prestasjoner. Videre var både de fem som hadde størst fremgang og de fem som hadde minst fremgang, like enige i at det var *confer* som forekom minst i undervisningen. En kan tenke seg at de klassene med liten forbedring ikke har hatt en like aktiv rolle i undervisningen, siden de trolig ikke har forbedret seg like mye. Men det forklarer ikke hvorfor de med stor forbedring opplevde det samme. Det kan tyde på at *confer* i liten grad faktisk påvirker prestasjoner.

5.2.3 Stor spredning vs. liten spredning i prestasjoner i klassene

Generell analyse

I denne studien var det 22 klasser (462 elever) som hadde større spredning i prestasjoner på regneprøven på 9. trinn, enn gjennomsnittet. Og det var 24 klasser (529 elever) som hadde mindre spredning. Utvalget her er også litt større hos de klassene med liten spredning, og vurderingene kan derfor være påvirket av antall elever som har deltatt i hver av de to gruppene.

Resultatene fra den deskriptive analysen av disse to gruppene, viser ingen store forskjeller i hvordan de vurderte undervisningskvaliteten til læreren i matematikk (se figur 4.3, side 44). De klassene med flest forskjellige prestasjonsnivåer innad i klasser, var i liten grad mer fornøyde med undervisningen enn de klassene med mer jevne prestasjoner. Dette er et artig funn, siden en kanskje kunne tenke seg at en klasse som består av elever med mange forskjellige mestringsnivåer ville hatt flere ulike oppfatninger av undervisningen. At en klasse hadde stor spredning i prestasjoner, behøver altså ikke å bety at det var en jevn fordeling av de fem mestringsnivåene. Faktisk kan en klasse med stor spredning ha bestått av elever hvor flertallet hadde høye prestasjoner, mens kun et fåtall hadde lave prestasjoner, og motsatt. For spredningen indikerer kun at det kan ha fantes en elev som hadde prestert veldig høyt og en elev som presterte veldig lavt. Siden disse klassene i litt større grad var mer fornøyde med undervisningen enn de klassene med liten spredning, kan det se ut som at det ikke

nødvendigvis var klasser som var veldig splittet i prestasjoner, men at det var mer jevnt fordelt. Videre kan klasser med liten spredning ha bestått av elever som enten presterte lavt, middels eller høyt, mens de med stor spredning kan ha bestått av litt forskjellige elevnivåer. Uavhengig av hvilke mestringsnivåer disse klassene med størst spredning bestod av, var det uansett interessant at disse klassene var litt mer enige i påstandene fra spørreskjemaet, enn de med liten spredning. De klassene med stor spredning kan ha prestert så ulikt og likevel opplevd undervisningen relativt likt innad i klassen. Dette stemmer godt overens med hva som ble diskutert i 5.2.1, hvor vi fant ut at elevenes prestasjoner ikke hadde så stor påvirkning i hvordan de opplevde lærerens undervisning.

For klassene med liten spredning forklarte regresjonsanalysen kun 2% av sammenhengen, og i dette tilfellet hadde både *control* (. 0153) og *captivate* (0.151) en signifikant positiv sammenheng, og *clarify* (-0.145) en negativ sammenheng. Regresjonsanalysen av de klassene med størst spredning kunne derimot kun forklare 3% av sammenhengen mellom de syv aspektene og hvorfor det var så stor variasjon i prestasjoner på regne9. Det vil si at ingen av disse modellen kan forklare så mye om hvilken sammenheng det er mellom variasjon i prestasjoner og hvordan de presterte. For det klassene med stor spredning var det kun *control* (0.171) som hadde en signifikant positiv sammenheng, og *clarify* (-0.130) hadde en signifikant negativ sammenheng. Dette kan kanskje forklares ved at noen elever, som kanskje var litt sterkere, hadde et litt mindre behov for at læreren brukte mye tid på å forklare, enn andre elever som kanskje var litt svakere. Jo mer de opplevde at læreren brukte unødvendig tid på å forklare, jo mindre utviklet de seg prestasjonsmessig. Samtidig så lærte muligens de svakere elevene i dette tilfellet, mer av at læreren tok seg god tid til å forklare slik at alle til slutt forstod prosedyrene. Men hvorfor de klassene med liten fremgang også fikk en negativ signifikant sammenheng mellom *clarify* og regne9, må ha en annen vinkling. Vi vet at det er en god blanding av klasser med liten variasjon i prestasjoner som presterte over gjennomsnittet. Tilsvarende var det en del som presterte under gjennomsnittet. Det vil si at uavhengig hvor sterk klassen er, så kan det virke som at elever i mer jevne klasser heller ikke ønsker en lærer som bruker mye tid ved tavlen.

Dybdeanalyse - de fem klassene med størst og minst spredning i prestasjoner.

En dypere analyse av de fem klassene som hadde størst spredning i prestasjoner viser at alle hadde en gjennomsnittlig prestasjonsnivå rundt gjennomsnittet på den nasjonale prøven på 9.

trinn. Det betyr at disse klassene må ha bestått av elever med ganske stor variasjon i mestringsnivåer. De eneste måtene disse klassene kan ha fått en gjennomsnittlig verdi rundt gjennomsnittet, er at klassen enten bestod av likt fordelt elever som presterte svært høyt og svært lavt, eller en jevn fordeling på alle de fem mestringsnivåene. Dersom en klasse kun hadde en svak elev og resten sterke, så ville fortsatt spredningen vært stor, men det ville vært umulig å få en så svak gjennomsnittlig verdi. Det betyr at disse klassene må ha bestått av elever som presterte veldig ulikt fra hverandre. Altså en god blanding av prestasjoner. Derfor er det interessant å se at disse fem klassene viser til å har vært generelt fornøyde med lærerens undervisning.

Dybdeanalysen av de fem klassene med minst spredning viser at den klassen med aller minst spredning faktisk presterte under gjennomsnittet på regneprøven. Faktisk var det kun to av dem som presterte over, og det var klasse2 og klasse31. Vi vet at klasse2 var den klassen som presterte aller best, så dette viser bare det faktum at dette var en klasse som bestod av veldig mange sterke elever. Videre viser 4.12 (side 43), at det var klasse10 som hadde den minste variasjonen av prestasjoner. Likevel viste det seg at klasse10 presterte relativt lavt under gjennomsnittet. Det vil si at dette må ha vært en klasse hvor flertallet var litt svakere elever. Det betyr at dette er fem klasser som enten har prestert veldig godt, veldig dårlig eller helt middels. Vi ser at også disse fem klassene generelt var enige om lærerens undervisning, men at klasse10 som hadde minst spredning, var i litt mindre enige enn de fire andre. Dette kunne kanskje forklares ved at de tross alt ikke presterte like høyt, som for eksempel med klasse2. Men sammenligner vi med klasse1, som presterte dårligst av alle klassene, hadde de en mer positiv vurdering enn klasse10. Med andre ord ser det ikke ut til å være en sammenheng mellom hvor svak en klasse er og hvordan de opplever lærerens undervisning.

5.3 Oppsummering av de tre *casene*

Hva kan disse funnene jeg har diskutert nå, si noe om ”hvilke sammenhenger det er mellom elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget?”. Som nevnt viser resultatene ingen tydelige sammenhenger mellom gode prestasjoner og hvor fornøyde elevene var med lærerens undervisning. Basert på regresjonsanalysen vist i tabell 4.14 (side 46), virker det som om elevenes prestasjoner i størst grad forklares av tidligere kunnskap. Det kan se ut som at den nye kunnskapen elevene har tillært seg på det ene året på ungdomsskolen, ikke kan forklare hvor stor sammenheng det er mellom prestasjoner og hvor

fornøyde elevene var med undervisningen. Siden regneprøven på 9. trinn består av den samme vanskelighetsgraden som på 8. trinn, bare med nye oppgaver, får ikke elevene vist den nye kunnskapen de har lært på det ene året. De får kun vist den kunnskapen som ønskes å måles, nemlig hva de kan etter syv år på barneskolen. Resultatene viser at alle klassene har hatt en fremgang fra den første til den andre prøven, men som sagt er dette forventet siden elevene er blitt ett år eldre (Utdanningsdirektoratet, 2017).

I tillegg varierte det i hvilket aspekt som ble vurdert høyest. Funnene indikerer derfor at lærerne som var med i denne studien har hatt et noe ulikt fokus i undervisningen, men den er minimal, siden det ikke var store forskjeller i hvilket aspekt som ble vurdert høyest. *Confer* var det aspektet som i alle tre *casene* ble vurdert lavest. Man kan kanskje tenke seg at undervisningen var litt mer lærerstyrt enn elevstyrt, siden elevene i mindre grad opplevde at deres ideer blir hørt, eller at de ble bedt om å dele tanker og ideer med hverandre. På den ene siden kan det kanskje være at elevene opplevde det som ubehagelig og skummelt å snakke høyt foran klassen. På den andre siden kan det være at læreren ofte fulgte en planlagt undervisning, som innebærer lite innspill fra elevene underveis. At elevene opplevde minst betydning i forhold til *confer,* kan muligens også forklares ved at spørsmålene som er blitt stilt, er litt mindre tydelige for elevene. De har kanskje ikke helt forstått betydningen bak spørsmålene og kan derfor ha valgt en middelverdi, som er verken enig eller uenig. Likevel er det viktig å påpeke at selv om *confer* kom dårligst ut i vurderingen, så trenger ikke det å bety at det ikke inntreffer i undervisningen. Faktisk så fikk aldri *confer* en gjennomsnittsverdi som var lavere enn 3 i de seks gruppene som er blitt analysert. Det betyr at flertallet har vurdert dette aspektet høyt og at det inntreffer i undervisningen.

At så å si alle de 46 klassene i studien vurderte aspektet *confer* lavest kan reflektere hva som har vært typisk for undervisningen i Norge. Matematikksenteret trekker frem viktigheten av å skape en større relasjonell forståelse hos elevene og gå mer bort i fra den instrumentelle undervisningen. Som Hattie (2009) også skriver, en undervisning som innebærer at elevene får en større rolle, fører til bedre utfall for hele elevgruppen og ikke bare for enkeltelever. Elever utvikler en dypere kunnskap når de blir mer engasjerte i undervisningen og når læreren klarer å vekke en interesse hos dem. For at elevene skal få en dypere forståelse, er det viktig at de inkluderes mer og oppfordres til å resonnere og reflektere over innholdet. Sannsynligheten for at elevenes forståelse øker skjer når elevene får tid til å reflektere og se sammenhenger. Tilsvarende er sannsynligheten for at elevenes dypere forståelse minker når

de ikke får denne tiden til å reflektere, men kun blir bedt om å løse oppgaver etter metoder som er demonstrert av læreren. Derfor er det viktig for meg i framtiden å fokusere på å skape en mer relasjonell forståelse hos elevene og inkludere dem i stor grad i undervisningen slik at de føler at de også har en stor rolle. Selvfølgelig er det også viktig at jeg som lærer kan fagstoffet, men en god undervisning handler også om å se elevene. Elever får det største læringsutbytte i en undervisning hvor læreren er flink til å styre timen og har god kontroll på innholdet og hva som skal gjennomgås den timen. De får også det største læringsutbytte når de får en effektiv læringstid. Større læringsutbytte vil føre til en bedre forståelse hos elevene og dermed bedre prestasjoner. Det er klart at det er andre faktorer som også påvirker elevenes prestasjoner og deres ønske om å lære. Elever som opplever at de har en god relasjon med læreren vil ofte engasjere seg litt ekstra i undervisningen og dermed få et større læringsutbytte den timen. En viktig forutsetning for at elevene skal få en vellykket læring er at de må oppleve undervisningen som et trygt sted å være. I tillegg så har elevenes faglige interesse og motivasjon for faget en stor påvirkning på deres utbytte. Mine funn kan tyde på at det er andre faktorer som også påvirker elevenes prestasjoner, siden alle klassene synes å ha vært fornøyde med lærerens undervisning. Det var ingen tydelig sammenheng mellom hvor stor grad av enighet klassene hadde på enkelte av aspektene, og hvordan de presterte. For eksempel kan både en klasse som presterte lavt og høyt, vurdert *control* høyest.

Regresjonsanalysene av disse seks *casene* viser også at *captivate* hadde en signifikant positiv sammenheng med prestasjonene på regne9 i de klassene som presterte over gjennomsnittet, de klassene med stor fremgang og i de klassene med liten spredning. For de to førstnevnte gruppene, var dette to grupper hvor elevene presterte veldig godt eller bedre i løpet av dette ene året på ungdomsskolen. For den siste gruppen, de med liten spredning, vil det være naturlig å tenke seg at flere elever hadde den samme oppfattelsen av undervisningen, siden de presterte jevnt med hverandre. Siden *captivate* hadde en positiv sterk sammenheng med gode prestasjoner, betyr det at når læreren ga elevene oppgaver som opplevdes som interessante og morsomme, så hadde det en positiv sammenheng med hvordan de presterte. Dessverre var dette et aspekt, som i alle tre *casene,* i gjennomsnitt ble vurdert lavt. Det ser ut til at elevene i mindre grad opplevde at undervisningen faktisk var interessant og morsom, selv om det basert på mine regresjonsanalyser skal vise til å ha en positiv effekt på prestasjonene. At elevene i mindre grad opplevde at dette skjedde i undervisningen stemmer godt overens med hva Klette (2013) trekker frem om typisk undervisning i norske klasserom. Hun trekker frem at det er veldig vanlig at det er noen undervisningsformer som forekommer oftere enn andre.

5.4 Sammenligning av mine funn og tidligere empiriske funn

Mine funn viser at det ikke var noen store forskjeller i hvordan klassene som presterte høyt estimerte lærerens undervisningskvalitet, i forhold til de klassene som presterte lavt. Uavhengig av hvor godt de presterte, så var det ikke mye forskjell i vurderingen av lærerens undervisning. Mer fornøyde elever viste derfor heller ikke til bedre prestasjoner, eller mer misfornøyde elever til dårlige prestasjoner.

Funnene fra regresjonsanalysen på hele gruppa viser også at det kun var *control* og *captivate* som hadde en signifikant påvirkning på elevenes prestasjoner på regne9, med en standardisert regresjonskoeffisient på 0,17 og 0,12. Dette viser at det var en positiv sammenheng, men som ikke hadde en særlig stor forklaringskraft på elevenes prestasjoner. Funnene fra Bergem et al. (2016) sin studie av TIMSS data, som også tok for seg norske matematikklasserom, viser at faktoren undervisningskvalitet, (klasseromsledelse, støttende lærer, tydelige intensjoner og kognitiv aktivering), viste en signifikant påvirkning til elevprestasjoner på 9. trinn med en regresjonskoeffisient 0,26. Vi ser at de kan vise til en litt større sammenheng enn våre funn, men både deres og våre regresjonskoeffisienter er moderate koeffisienter, som begge viser en positiv sammenheng. Regresjonsanalysen i denne studien viser også blant annet at ikke alle syv aspekter kan forklare elevenes prestasjoner, så med de lave regresjonskoeffisientene betyr det at det må være andre faktorer som også påvirker elevenes prestasjoner i faget. Bergem et al. (2016) undersøkte også sammenhengen mellom god undervisningskvalitet og elevenes motivasjon for faget, og disse resultatene viser at god undervisning har en større påvirkning på elevenes selvtillit og motivasjon enn det har for gode prestasjoner.

Ser vi på hva våre elever estimerte seg mest enige i av alle de syv aspektene som skal måle god undervisning, ser vi at alle klassene vurderte *control* høyest. Det de vurderte lavest var *confer*. Disse vurderingene stemmer godt overens med tidligere funn i denne oppgaven. Bergem et al. (2016) fant i sine studier ut at elevene rapporterte mest positivt til aspektet støttende lærer og minst på god klasseromsledelse. Våre funn varierer derfor fra Bergem et al. (2016) sine funn, siden *control* i mitt tilfelle måler hva noe av det som beskriver klasseromsledelse i denne studien.

Sammenligner vi våre funn med de amerikanske studiene Kuhfeld (2017) og Wallace et al. (2016), opplevde deres elever i større grad *challenge* i undervisningen, og *control* minst.

Dette skiller seg fra våre elever, da det var *control* elevene vurderte høyes og *confer* lavest. Som nevnt i kapittel 4.1.3, fikk jeg en hypotese om at undervisningen i USA har andre fokusområder enn undervisningen i Norge. En kan kanskje tenke seg at undervisningen i USA er mer preget av lærerstyrt undervisning med mye fokus på klasseromsledelse. Mens det i Norge ser ut til å være mer fokus på å utfordre elevene. I følge PISA sin studie utført i 2012, så presterte norske elever litt bedre enn de amerikanske elevene i denne perioden (Nortvedt, 2013). Resultatene fra de tidligere empiriske studiene viser mye av det samme. Mine funn tyder på at det ikke var store forskjeller i hvordan de norske elevene vurdert lærerens undervisning i forhold til de amerikanske elevene. Det kan se ut som at elevene var like fornøyde med lærernes undervisning, men at lærerne opptrer forskjellig i klasserommene. Dette styrker teorien til Klette (2013) om at god undervisning ikke handler om hvem læreren er, men hva læreren gjør.

6 Avslutning

Avslutningsvis gis en kort redegjørelse for mine hovedfunn. Videre trekker jeg frem videre profesjonsutvikling basert på lærdom fra denne oppgaven. Til slutt trekkes frem noen forslag for videre forskning.

6.1 Svar på problemstillingen

Innledningsvis stilte jeg spørsmålet om hva god undervisning innebærer. Finnes det én god oppskrift på hva som kjennetegner god undervisningskvalitet i matematikk? Basert på oppgavens forskningsspørsmål ønsket jeg å undersøke om det var store forskjeller i hvordan elevene opplevde undervisningen, basert på hvordan de presterte. Temaet for denne oppgaven var elevenes vurdering av lærerens matematikkundervisning. Problemstillingen var, "hvilke sammenhenger er det mellom elevenes syn på god undervisningskvalitet i matematikk og deres prestasjoner i faget?". Gjennom flere deskriptive analyser, viser mine hovedfunn at disse 991 elevene, uavhengig av hvordan de presterte, generelt var fornøyde med deres matematikklærer og undervisning. Forskningen ble delt opp i tre undergrupper som i oppgavene ble omtalt som *caser*; høyt presterende vs. lavt presterende klasser, klasser med stor vs. liten fremgang i prestasjoner og klasser med stor vs. liten spredning i prestasjoner.

I den første undersøkelsen som var av hvordan klassene som presterte høyt og hvordan de vurderte undervisningskvaliteten i forhold til de klassene som presterer lavt, viste det seg å være ingen store forskjeller. Begge disse to gruppene viste til høye gjennomsnittlige vurderinger i spørreskjemaet. I starten av oppgaven hadde jeg en hypotese om at de høyt presterende klassene var mer fornøyde med undervisningen enn de lavt presterende. Dette viste seg å ikke stemme. Faktisk så vurderte de lavt presterende i litt større grad mer positivt enn de høyt presterende. Likevel er det viktig å påpeke at denne forskjellen var minimal.

I den andre undersøkelsen som var av hvordan de klassene med størst fremgang i prestasjoner og de med minst fremgang, ser vi en litt tydeligere forskjell i hvordan disse to gruppene vurderte undervisningskvaliteten. Også her var det positive vurderinger, men klassene med størst fremgang var litt mer enige om lærerens undervisning enn de med liten fremgang. Resultatene kan likevel ikke bekrefte hypotesen om at de med størst fremgang var mer fornøyde enn de med liten fremgang. Mine funn tyder på at både de klassene med stor

fremgang og de med liten fremgang opplevde at læreren oppfyller de kravene som kreves for en god undervisning i matematikk.

I den tredje undersøkelsen som var av klassene med stor og liten spredning i prestasjoner, var det heller ingen store forskjeller i hvordan disse to gruppene vurderte undervisningskvaliteten. Derfor stemmer heller ikke hypotesen i dette tilfellet, om at klassene med minst spredning var mer fornøyde enn de klassene med stor spredning.

Alle mine tre analyser viser at det var *confer* som var det aspektet som i gjennomsnitt ble vurdert lavest. Ikke alle klassene vurderer dette lavest, men gjennomsnittet til alle de 46 klassene gjorde det. Det tyder på at ikke alle elever opplevde det som like trygt å snakke høyt foran klassen og at ikke alle følte de fikk være like delaktige i timen som de ønsket. Dette behøver ikke nødvendigvis å være læreren som ikke var flink til å inkludere elevene, det kan også være at elevene ikke ønsket å være like aktive og derfor har unngått slike situasjoner. Videre samsvarer mine funn fra regresjonsanalysene godt med at *control* var det aspektet som hadde størst sammenheng med elevenes prestasjoner. Likevel viser de deskriptive analysene at både sterke og svake klasser opplevde at læreren utførte aspektet *control* i stor grad i undervisningen. At *control* hadde en sammenheng med gode prestasjoner kan derfor ikke bekreftes i denne undersøkelsen.

Hvilke sammenhenger er det mellom elevers syn på god undervisningskvalitet og deres prestasjoner i matematikk? Det ser ut til at det i dette tilfellet ikke var en klar sammenheng. Gode prestasjoner medførte ikke til mer fornøyde elever, enn dårlige prestasjoner. Elever med stor fremgang var i dette tilfellet litt mer fornøyde enn de med liten fremgang. Om en klasse hadde stor variasjon av prestasjoner eller liten variasjon, påvirket i liten grad elevenes syn på undervisningen. Det betyr at det i en klasse som bestod av elever på flere ulike mestringsnivåer, vurderte relativt enige med hverandre om lærerens undervisning. Alle mine analyser viser at det uansett hvilken vinkling vi ser på sammenhenger mellom elevenes syn på undervisningskvalitet og hvordan de har prestert, så viser alle klassene til å være relativt enige og fornøyde i sin undervisning. Så vi kan nok ikke si at det finnes én god oppskrift på en undervisning som skal føre til mange gode prestasjoner. Det virker som at disse lærerne oppfyller forventningene som finnes for god undervisning, men at det må være andre faktorer som også påvirker prestasjonene. For eksempel kan det se ut til at prestasjonene i stor grad forklares av elevenes forkunnskaper.

Alle disse analysene er basert på elevvurderinger av lærerens undervisning. Hvor godt måler egentlig disse vurderingene fra elevene lærerens undervisningskvalitet? Analysene viser til et flertall av positive elever, både faglig sterke og svake. Basert på mine forskningsfunn, ser det ut til at man burde gå mer i dybden for å undersøke sammenhengen mellom elevenes syn på god undervisningskvalitet i matematikk og hvordan de presterer, som for eksempel intervju. Basert på mine funn sammenlignet med de amerikanske studiene, viser det til å være like positive elever til undervisningen. Eneste forskjellen var at de har et litt ulikt syn på hvilket aspekt læreren utfører mest og minst i undervisningen.

6.2 Videre profesjonsutvikling

Hva har jeg lært fra denne studien? Jeg har lært mer om hva som kjennetegner god undervisningskvalitet. For min egen profesjonsutvikling ser jeg verdien av å ha større kunnskap over sentrale faktorer ved lærerens undervisning som skal ha en positiv betydning for elevenes prestasjoner. Fra min egen skoleoppvekst, har jeg erfart en mer tradisjonell undervisning, og fra denne oppgaven har jeg fått et tydeligere bilde over hvor viktig det også er med relasjonell undervisning. I mine fremtidige undervisninger, ønsker jeg å ta med meg hvor viktig det er å se hver enkelt elev, utfordre dem og inkludere dem mer i undervisningen. Basert på mine hovedfunn, overrasket det meg litt hvor liten sammenheng det var mellom prestasjoner og elevenes syn på undervisningen. Jeg har derfor også sett at elevenes prestasjoner ikke kun avhenger av lærerens undervisning, og at det må være andre faktorer som påvirker. For videre profesjonsutvikling er det derfor også viktig å få en bedre forståelse over hva som påvirker elevenes prestasjoner, slik at jeg kan hjelpe flest elever frem mot god forståelse av matematikkfaget.

6.3 Forslag til videre forskning

Mine funn tyder på at de 46 lærerne i denne studien var flinke til å undervise, slik som teorien beskriver som god undervisning. Det ser ut til at lærerens måte å undervise på ikke hadde den største påvirkningen på hvordan elevene presterte. Derfor ville det være spennende og forske videre på hvor store forskjeller det er i lærernes kompetanse i matematikk. Er det slik at faglig sterke lærere fører til mer faglig sterke elever? Eller klarer en mindre faglig sterk lærer, som er veldig flink til å motivere elever i undervisningen, å fremheve like faglig sterke elever?

Som nevnt, viser mine funn liten variasjon mellom hvordan elevene har vurdert undervisningen. Det ville vært interessant å undersøke hvorfor det er så liten variasjon blant disse norske elevene. Er elevvurderinger virkelig en god måte å granske undervisningskvalitet på? Mine funn tyder på at dette muligens ikke var det beste tilfellet i norsk sammenheng. Elevene var veldig enige i at undervisningen var bra, likevel presterer Norge under det internasjonale gjennomsnittet. I denne studien var det kun tilgang på elevenes resultater fra spørreskjemaet. Det ville derfor vært interessant å intervjuet flere elever om dette temaet, for å finne ut hvilke andre faktorer som påvirker elevenes prestasjoner.

Jeg synes også det ville vært mer interessant å undersøke hvorfor noen teoretiske perspektiver definerer god undervisningskvalitet gjennom tre aspekter, mens andre gjennom syv. I denne studien ble det som nevnt ikke laget en test for sammenligning av modeller som best skal måle god undervisningskvalitet. Dette fordi at det ville være for komplekst for denne oppgaven. Dette er noe som burde gjøres for å best måle de to modellene (syvdimensjonal og tredimensjonal faktormodell) opp mot hverandre. Ved å gjøre slike tester vil det kanskje oppstå en større oversikt over hva som beskriver god undervisningskvalitet. Samtidig vil det være interessant å se nærmere på hva som er mer særegent for matematikk i forhold til andre fag. Er det slik at definisjonen av god undervisningskvalitet i matematikk også er god i norsk?

Litteraturliste

Bergem, O. K., Nilsen, T., & Scherer, R. (2016). Undervisningskvalitet i matematikk, I: Bergem, O. K., Kaarstein, H., & Nilsen, T. (red), *Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015* (s. 120-136). Oslo: Universitetsforlaget.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal,* 47(1), s. 133-180.

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. I: Nilsen, T. & Gustafsson, J-E. (red), *Teacher quality, instructional quality and student outcomes, Relationship across countries cohort and time* (s. 21-50). Sveits: Springer international publishing

Bø, A-K., & Hovdenak, S. S. (2011). Faglig og personlig støtte: Om betydningen av en god relasjon mellom lærer og elev sett fra elevens ståsted, I: *Tidsskrift for Ungdomsforskning*, 11(1), s, 69-85

Christophersen, K-A. (2012). IBM SPSS/AMOS: *Databehandling og statistisk analyse* (utg. 5). Oslo: Akademika forlag.

Clausen, S. E. (2009). *Multivariate analysemetoder for samfunnsvitere: Med eksempler i SPSS.* Oslo: Universitetsforlaget.

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). What makes great teaching? Review of the underpinning research. London: Sutton Trust.

Drugli, M. B. (2012). *Relasjonen lærer og elev. Avgjørende for elevenes læring og trivsel.* Oslo: Cappelen Damm Akademisk. (s. 45-65).

Everett, E. L., & Furuseth, I. (2012). Lettere sagt enn gjort – å utforme et metodisk opplegg for oppgaven. I Everett., E. L., & Furuseth, I. (red), *masteroppgaven. Hvordan begynne og fullføre* (s. 127-144). Oslo: Universitetsforlaget.

Ferguson, R. F. (2010). *Student perceptions of teaching effectiveness. Discussion brief.* Cambridge, MA: National Center for Teacher Effectiveness and the Achievement Gap Initiative, Harvard University

Ferguson, R. F. (2012). Can student surveys measure teaching quality? Kappan. 94(3), s. 24 28.

Geiser, C. (2013). *Data Analysis with Mplus.* New York: The Guilford Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analysis relating to achievement.* New York: Routledge.

Johannesen, A. (2009). *Introduksjon til SPSS: Versjon 17* (utg. 4). Oslo: Abstrakt forlag.

Johnson, B. R., & Christensen, L. (2013). Validity of research results in quantitative, qualitative and mixed research. (kap. 11), I Johnson, B. R. & Christensen, L. (red). Educational research: quantitative, qualitative and mixed approaches (s. 277-316). Sage: Los Angeles.

Kane, T. J. & Cantrell, S. (2009). Learning about teaching: Initial findings from the Measures of Effective Teaching Project. Bill and Melinda Gates Foundation.

Klette, K. (2013). Hva vet vi om god undervisning? Rapport fra klasseromsforskningen. I: Krumsvik, R. J., & Säljö, R. (red), *Praktisk pedagogisk utdanning: En antologi* (s. 173-200). Bergen: Fagbokforlaget.

Klette, K., Blikstad-Balas, M., & Roe, A. (2017). Linking Instruction and Student Achievement: *Research design for a new generation of classroom studies.* Acta didactica Norge. 11(3), s. 1-19.

Kleven, T. A. (2014). Data og datainnsamlingsmetoder, I Kleven, T. A. (red): *Innføring I pedagogisk forskningsmetode* (s. 27-47). Bergen: Fagbokforlaget.

Kuhfeld, M. (2017). When students grade their teachers: a validity analysis of the tripod student survey. Educational Assessment, 22(4), s. 253-274.

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. Learning and Instruction, 17, s. 494-509.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. Learning and Instruction, 19(6), s. 527 537.

Nilsen, T., Gustafsson, J-E., & Blömeke, S. (2016). Conseptual framework and methodology of this report. I: Nilsen, T., & Gustafsson, J-E. (red), *Teacher quality, instructional quality and student outcomes. Relationships across countries, cohorts and time* (s. 1 20). Sveits: Springer international publishing.

Nortvedt, G. A. (2013). *Resultater i matematikk.* I Kjærnsli, M., & Olsen, R. V. (red) fortsatt en vei å gå: norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012. Oslo: Universitetsforlaget. (s. 67-95).

Nosrati, M., & Wæge, K. (2015). Sentrale kjennetegn på god læring og undervisning I matematikk. Hentet 24.01.19 fra: https://www.matematikksenteret.no/sites/default/files/attachments/product/150629kor.%20Sentrale%20kjennetegn%20på%20god%20læring%20og%20undervisning%20%20matematikk.pdf

Ogden, T. (2012). *Klasseledelse: Praksis, teori og forskning.* Oslo: Gyldendal Akademisk.

Postholm, M. B. (2013). Klasseledelse i ungdomsskolen. Fire læreres meninger og opplevelser. (s. 127-140). Trondheim: Universitetsforlaget

Postholm, M. B., & Jacobsen, D. I. (2014). *Læreren med forskerblikk: Innføring i vitenskapelig metode for lærerstudenter.* Kristiansand: Høgskoleforlaget.

Ringdal, K. (2018). *Enhet og mangfold: Samfunnsvitenskapelig forskning og kvantitativ metode* (utg. 4). Bergen: Fagbokforlaget.

Schlesinger, L. & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observation. Springer Berlin Heidelberg, 48(1-2), s. 29-40.

Schlesinger, L., Jentsch, A., Kaiser, G., König., J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics. Springer. 50, s. 475-490.

Schweig, J. (2016. Moving beyond means: revealing features of the learning environment by investigating the consensus among student ratings. Springer Science. 19, s. 441-462.

Skog, O-J. (2005). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming* (utg. 2). Oslo: Gyldendal Norsk Forlag

Utdanningsdirektoratet. (2017). Mestringsbeskrivelser for nasjonale prøver i regning. Hentet fra:https://www.udir.no/eksamen-og-prover/prover/nasjonale prover/mestringsbeskrivelser-og-hva-provene-maler/mestringsbeskrivelser-for nasjonale prover-i-regning/

Utdanningsdirektoratet. (2018). Hva er nasjonale prøver? Hentet fra: https://www.udir.no/eksamen-og-prover/prover/nasjonale-prover/om-nasjonale prover/

Utdanningsdirektoratet. (2019). Nasjonale prøver ungdomstrinn. Hentet fra: https://skoleporten.udir.no/rapportvisning/grunnskole/laeringsresultater/nasjonale proever ungdomstrinn/nasjonalt?enhetsid=00&vurderingsomrade=11&underomrade=51&skoetype=0&skoletypemenuid=0&sammenstilling=1

Wallace, T. L., Kelcey, B., & Ruzek, E. (2016). What can student perception surveys tell us about teaching? Empirical testing their underlying structure of the tripod student perception survey. American Educational. 53(06), s. 1834-1868.