

Trine Smisethjell

Karakterfri underveisvurdering i ungdomsskolen

En kvantitativ studie om forskjeller mellom elevers opplevelse av skolens målstruktur og av seg selv i ungdomsskoler med og uten karakterfri underveisvurdering

Masteroppgave i spesialpedagogikk

Veileder: Marit Uthus og Vegard Johansen

August 2019

Trine Smisethjell

Karakterfri underveisvurdering i ungdomsskolen

En kvantitativ studie om forskjeller mellom elevers opplevelse av skolens målstruktur og av seg selv i ungdomsskoler med og uten karakterfri underveisvurdering

Masteroppgave i spesialpedagogikk
Veileder: Marit Uthus og Vegard Johansen
August 2019

Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for pedagogikk og livslang læring

Sammendrag

Hensikten med denne studien er å undersøke forskjeller i elevenes opplevelse av skolens målstruktur og av seg selv i ungdomsskoler med karakterfri undervisvurdering og ungdomsskoler med undervisvurdering med karakterer. Ytterligere forklart dreier det seg om elevenes opplevelse av prestasjonsorientert målstruktur, læringsorientert målstruktur, indre motivasjon, innsats, selvvurdering av kompetanse og skolestress. Det er i denne studien anvendt en kvantitativ metode og for å besvare studiens problemstilling er det benyttet spørreundersøkelser for å samle inn data fra fem ungdomsskoler. Ved to av ungdomsskolene mottar elevene undervisvurdering med karakterer, mens i de tre andre ungdomsskolene mottar elevene karakterfri undervisvurdering. Spørreundersøkelsene ble besvart av elevene, og det endelige utvalget består av 1071 elever.

Studien bygger på et kvasiekperimentelt design, og for å undersøke forskjellene mellom elevenes opplevelse av skolens målstruktur og av seg selv i skolen er det gjennomført t-tester, enveis variansanalyse (ANOVA) og kovariansanalyser (ANCOVA). Det er i forbindelse med denne studien og gjennomført beskrivende analyser, prinsipale komponentanalyser og reliabilitetsanalyser.

Studiens resultater viser at det ikke er en signifikant forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Disse funnene stemmer i svært liten grad overens med forventningene til denne studien. Da det i denne studien er forventet at erfaring med karakterfri undervisvurdering kan ha en betydning for elevenes opplevelse av skolens målstruktur og av seg selv, er det også gjennomført analyser for å undersøke forskjeller mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering, samt elever fra skoler med undervisvurdering med karakterer. Resultatene fra disse analysene viser å stemme i liten grad overens med forventningene til studien. Funnene tyder derimot å stemme overens med forventningene om at det ikke er en signifikant forskjell mellom gruppegjennomsnittene for innsats. Samtidig tyder de på å stemme til en viss grad med forventningene knyttet til skolens målstruktur og skolestress.

Førord

Det å skrive masteroppgave har både vært en spennende og lærerik utfordring. Jeg skal derimot ikke legge skjul på at det til tider også har vært krevende. Nå er jeg endelig i mål og med det er studiet gjennomført og allerede neste uke starter et nytt kapittel i livet når jeg da begynner i læreryrket.

Når masteroppgaven nå er ferdig er det flere som fortjener en takk. Jeg vil først og fremst takke min hovedveileder Marit Uthus og min biveileder Vegard Johansen for meget verdifull veiledning gjennom hele dette masterprosjektet. Jeg vil og takke Kyrre Svarva for hjelp med utforming og skanning av spørreskjema, samt overføring av data inn SPSS. Jeg vil også takke Institutt for pedagogikk og livslang læring (IPL) for støtte til å gjennomføre dette masterprosjektet. Og så klart, takk til alle deltagere. Til slutt vil jeg takke min familie, kjæreste og venninne for all støtte og oppmuntring.

NTNU, Trondheim, 2019

Trine Smisethjell

Innholdsfortegnelse

Sammendrag	i
Forord	iii
1.0 Innledning	1
1.1 Aktualitet.....	1
1.2 Begrepsavklaring og problemstilling	3
1.3 Studiens oppbygning	4
2.0 Teori og forskning	5
2.1 Underveisvurdering i ungdomsskolen.....	5
2.2 Skolens målstruktur.....	7
2.2.1 Prestasjonsorientert målstruktur.....	7
2.2.2 Læringsorientert målstruktur	9
2.3 Indre motivasjon.....	10
2.4 Innsats.....	11
2.5 Selvvurdering av kompetanse	11
2.6 Stress	12
2.7 Bakgrunnsvariabler	14
2.8 Hypoteser.....	15
2.9 Andre relevante forhold som påvirker resultatene	18
3.0 Metode	21
3.1 Forskningsmetode og design.....	21
3.2 Datainnsamlingsprosess	22
3.3 Utvalg	24
3.4 Måleinstrument.....	25
3.4.1 Prestasjonsorientert målstruktur (avhengig variabel)	26
3.4.2 Læringsorientert målstruktur (avhengig variabel).....	26

3.4.3	<i>Indre motivasjon (avhengig variabel)</i>	26
3.4.4	<i>Innsats (avhengig variabel)</i>	27
3.4.5	<i>Selvvurdering av kompetanse (avhengig variabel)</i>	27
3.4.6	<i>Skolestress (avhengig variabel)</i>	28
3.4.7	<i>Bakgrunnsvariabler (uavhengig variabler)</i>	28
3.4.8	<i>Test- og kontrollgruppe (uavhengig variabel)</i>	28
3.5	Statistiske analyser	29
3.5.1	<i>Prinsipale komponentanalyser og reliabilitetsanalyser</i>	29
3.5.2	<i>Deskriptiv statistikk</i>	30
3.5.3	<i>T-test og ANOVA</i>	30
3.5.4	<i>ANCOVA</i>	30
3.6	Kvalitetssikring av studiens sammensatte mål	31
3.7	Etiske betraktninger	32
4.0	Resultat	35
4.1	Beskrivelse av variablene	35
4.1.1	<i>Studiens avhengige variabler</i>	35
4.1.2	<i>Studiens uavhengige variabler</i>	38
4.2	Sammenligning av elever med og uten karakterer	40
4.2.1	<i>Forskjell mellom testgruppa og kontrollgruppa</i>	40
4.2.2	<i>Har erfaring betydning for resultatene?</i>	41
4.2.3	<i>ANCOVA</i>	45
5.0	Diskusjon	49
5.1	Skolens målstruktur	49
5.2	Indre motivasjon	53
5.3	Innsats	56
5.4	Selvvurdering av kompetanse	58
5.5	Stress	60
6.0	Avslutning	65
6.1	Oppsummering av studiens hovedfunn	65

6.2 Studiens begrensninger.....	66
6.3 Videre forskning.....	68
Litteraturliste.....	69
Vedlegg 1: Spørreskjema.....	77
Vedlegg 2: Informasjonsskriv til skolene.....	78
Vedlegg 3: Informasjonsskriv til foreldre.....	80
Vedlegg 4: Informasjonsskriv til elevene.....	81
Vedlegg 5: Prinsipale komponentanalyser (PKA) og Cronbachs alfa.....	82
Vedlegg 6: Skjevhet.....	85
Vedlegg 7: ANCOVA og Cohens' d.....	86

1.0 Innledning

1.1 Aktualitet

«Frå 8. årstrinn og i vidaregåande opplæring skal vurdering også givast med talkarakterar.» (Endr. i forskrift til opplæringslova, 2009, § 3-4). Ifølge regelverkstolkningene til Utdanningsdirektoratet (2016) betyr dette at elevene skal få karakterer ved halvårsvurderinger og sluttvurderinger og at det utenfor dette ikke stilles krav om at lærere skal bruke karakterer i ungdomsskolen og i videregående skole. Med andre ord bestemmer læreren selv om vurderingen inkluderer karakterer eller ikke underveis i opplæringen i ungdomsskolen og i videregående skole. På flere ungdomsskoler velger lærere å begrense karakterbruken, hvor å styrke fokuset på læringen og redusere stress og karakterpress nevnes som sentrale årsaker til denne endringen (Jansen, 2018; Mellingsæter, 2018; Mordt, Archer & Løken, 2016; Ruud, 2018; Simensen, 2018). En rapport av Danmarks Evalueringsinstitut (2017) viser til at omtrent 70 prosent av elever opplever å være helt eller delvis enig i at det er mindre press uten karakter i videregående skole. Forsker Eriksen har forsket på stress og press blant elever (Eriksen, Sletten, Bakken & von Soest, 2017) og uttrykker en begeistring over å kutte karakterer i underveisvurderingen på ungdomsskolen (Jansen, 2018).

Klinger et al. (2015) viser i sin studie at prestasjonspresset i skolen har økt i Norge i en periode der det ble klart at Norge kom dårligere ut på PISA-undersøkelsene sammenlignet med flere andre nasjoner. Prestasjonspress er en form for press, som ifølge undersøkelsen til Skaalvik og Federici (2015) oppleves av mange elever. Samtidig finner Skaalvik og Federici (2015) at dette kan resultere i nedstemthet, utmattelse og lavere selvværd. De mener derfor at det er viktig å rette et fokus mot hva som skyldes elevenes prestasjonspress i skolen. Basert på forskning mener Uthus (2017) at prestasjonspress som utvikler seg til psykiske helseplager, kan bidra til en økning i plager i elevenes liv. Ifølge Uthus (sitert i Uthus, 2017) og Wendelborg (sitert i Uthus, 2017) er det en fare for at barn med særskilte opplæringsbehov vil kunne rammes spesielt hardt av prestasjonspresset i skolen. Det blir dermed viktig å aktualisere Skaalvik og Federici (2015) sin mening, nettopp for å kunne dempe prestasjonspresset i skolen, samt bidra til å utvikle en helsefremmende skole.

Om nettopp en karakterfri underveisvurdering vil bidra til mindre stress og press i skolen er det uenigheter om blant fagforeninger og fagforbund (Ruud, 2018). Det er derfor interessant å studere hvilken innvirkning en karakterfri underveisvurdering har på press og stress. Den

tidligere nevnte rapporten til Danmarks Evalueringsinstitut (2017) som antyder at karakterer bidrar til press, viser og til at 80 prosent av elevene rapporterer at de er helt eller delvis enig i at vurdering uten karakterer gjør dem usikre på hvilket nivå de ligger på. For noen bidrar det sistnevnte også til press, og virker derfor som et motstridende funn. 70 prosent av medlemmene i Norsk lektorlag opplever og at elevene virker usikre om hvordan de ligger an faglig når de mottar vurdering uten karakter, noe som videre gjør at de må bruke mer tid på de faglige vurderingene (Helgesen, 2018). Professor Gamlem uttaler i Aftenposten og professor Burner uttaler i podcasten Lærerrommet at man kan ikke forvente forbedring kun ved å kutte ut karakterene, det helt nødvendig å jobbe med gode tilbakemeldinger (Mellingsæter, 2018; Utdanningsforbundet, 2019). En rektor antyder i et intervju med Utdanning at det var arbeidskrevende ved overgangen til karakterfri undervisvurdering, men at skolen nå har funnet en god løsning (Ruud, 2018). Leder av Utdanningsforbundet nevner i Utdanning 11/2018 at det tar tid å endre en vurderingspraksis (Ruud, 2018). Og, det skal poengteres at forsøksskolene som deltar i rapporten av Danmarks Evalueringsinstitut (2017) kun har gjennomført karakterfri vurdering i ett år. På bakgrunn av dette kan det være en mulighet for at erfaringene skolene og elevene har med vurderingspraksisen har en betydning for blant annet elevenes pressnivå i skolen.

I tillegg til stress og press er det også betydningsfullt å undersøke eventuelle andre virkninger undervisvurderinger uten karakterer kan ha for elevene i skolen. Deriblant mener flere medlemmer av Norsk lektorlag at karakterer fungerer som en motivasjonsfaktor for elevene (Helgesen, 2018). Én ungdomsskole støtte også dette, ved at de antyder at for noen elever fungerer karakterene som ytre motivasjon (Mellingsæter, 2018). Hvordan er så motivasjonen til elever som ikke mottar karakterer? Videre nevner noen elever fra samme skole at en karakterfri vurdering har bidratt til mindre sammenligning (Mellingsæter, 2018). Professor Prøitz uttaler i Utdanning 11/2018 at det er lite forskning på karakterfri vurdering i Norge (Ruud, 2018). Videre antyder Eriksen og Bakken (2018) for å være interessert i hvilken effekt et slikt tiltak har på stress. Det vurderes derfor som aktuelt å forske nettopp på karakterfri undervisvurdering, og da finne ut av hvilken innvirkning undervisvurdering uten karakterer har på elevene i ungdomsskolen.

1.2 Begrepsavklaring og problemstilling

Det er som nevnt begrenset forskning på karakterfri underveisvurdering i Norge. Det vurderes derfor som viktig å undersøke temaet, og det er nettopp dette denne studien vil omfatte. Karakterfri underveisvurdering er underveisvurdering som ekskluderer karakterer. Med underveisvurdering menes all vurdering som gis løpende underveis i opplæringen i skolen (Utdanningsdirektoratet, 2015). Halvårsvurderingen er en del av underveisvurdering (Endr. i forskrift til opplæringslova, 2009, § 3-13), men en viktig bemerkning er at i denne studien vil begrepet underveisvurdering ekskludere halvårsvurderingen.

I denne studien vil jeg sammenligne elever fra ungdomsskoler med karakterfri underveisvurdering og elever fra ungdomsskoler med underveisvurdering med karakterer. Og derav er det et ønske om å få noen indikasjoner på hvilken betydning karakterfri underveisvurdering har for elevenes opplevelse av seg selv og skolens målstruktur. Forrige delkapittel antydte at karakterfri underveisvurdering kan ha en betydning for elevenes press og stress i skolen og denne studien kan være et bidrag til å finne ut av dette. Med utgangspunkt i aktualitetskapitlet, samt teori og forskning som presenteres i kapittel 2, er det vurdert som relevant å studere elevens opplevelse av skolestress, samt andre faktorer. Disse faktorene er elevenes opplevelse av indre motivasjon, innsats, selvvurdering av kompetanse og skolens målstruktur. På bakgrunn av de omtalte kapitlene, samt studiens metodevalg, er det utviklet en problemstilling som denne oppgaven vil ta utgangspunkt i. Problemstillingen er som følgende:

«Hvilke forskjeller er det mellom elevers opplevelse av skolens målstruktur og av seg selv i ungdomsskoler med karakterfri underveisvurdering og ungdomsskoler med underveisvurdering med karakterer?»

Opgaven vil forsøke å undersøke disse forskjellene gjennom to forskningsspørsmål:

1. Hvilke forskjeller er det mellom elevers opplevelse av prestasjonsorientert målstruktur og læringsorientert målstruktur i ungdomsskoler med karakterfri underveisvurdering og ungdomsskoler med underveisvurdering med karakterer?
2. Hvilke forskjeller er det mellom elevers opplevelse av indre motivasjon, innsats, selvvurdering av kompetanse og skolestress i ungdomsskoler med karakterfri underveisvurdering og ungdomsskoler med underveisvurdering med karakterer?

Studiens problemstilling og forskningsspørsmål konkretiseres ytterligere ved at det formuleres flere hypoteser. Disse hypotesene presenteres i delkapittel 2.8. For å kunne teste hypotesene, samt svare på de presenterte spørsmålene er det i denne studien innsamlet eget datamateriale via et spørreskjema som elever fra fem ulike skoler har svart på. I to av skolene mottar elevene undervisvurdering med karakterer, mens i de tre resterende skolene mottar elevene undervisvurdering uten karakterer. En mer detaljert forklaring av utvalget presenteres i oppgavens metodekapittel.

1.3 Studiens oppbygning

Denne studien består av seks hovedkapitler, samt vedlegg. Innledningen er oppgavens første kapittel og presenterer studiens aktualitet, begrepsavklaring og oppgavens problemstilling. Sammen med introduksjonen av problemstillingen presenteres studiens forskningsspørsmål. I det påfølgende kapitlet utdypes relevant teori og forskning, som er utgangspunktet for utviklingen av hypotesene som presenteres i den nest siste delen av kapitlet. Avslutningsvis i kapittel to introduseres noen relevante forhold som kan påvirke resultatene i denne studien. I kapittel tre vil jeg redegjøre for valg av metode og forskningsdesign, samt presentere studiens datainnsamlingsprosess, utvalg, måleinstrumenter og de valgte statistiske analysemetodene. Videre inneholder kapitlet en empirisk kvalitetssikring av studiens sammensatte mål, før kapitlet avsluttes med et delkapittel om forskningsetiske betraktninger. I studiens fjerde kapittel vil jeg presentere resultatene fra de statiske analysene i SPSS. Deretter vil jeg i kapittel fem drøfte hypotesene fra kapittel to i lys av resultatene fra kapittel fire, samt teori og forskning fra kapittel to. Studiens sjettede og siste kapittel inneholder en oppsummering av studiens hovedfunn og svar på studiens problemstilling. Samtidig vil studiens begrensninger og forslag til videre forskning presenteres i dette kapitlet.

2.0 Teori og forskning

I dette kapitlet utdypes teorier og tidligere forskning som vurderes som relevante for å kunne besvare oppgavens problemstilling og forskningsspørsmål. Et overordnet tema for denne studien er undervisvurdering uten og med karakterer, og det vil derfor til å begynne med gjøres rede for begrepet undervisvurdering i ungdomsskolen. Videre vil jeg redegjøre for begrepene prestasjonsorientert og læringsorientert målstruktur, indre motivasjon, innsats, samt selv vurdering av kompetanse og skolestress. Deretter vil studiens bakgrunnsvariabler bli presentert, med vekt på deres aktualitet i denne studien. Videre vil det formuleres seks hypoteser for denne studien, som er basert på de teorier og forskninger som presenteres i dette kapitlet. Avslutningsvis presenteres andre relevante forhold som kan påvirke de avhengige variablene som inngår i denne studien.

2.1 Undervisvurdering i ungdomsskolen

Studiens overordnede tema er undervisvurdering. Med undervisvurdering menes all vurdering som gis løpende underveis i opplæringen i skolen (Utdanningsdirektoratet, 2015), hvor det i denne studien vil dreie seg om ungdomsskolen. Det ble i forrige kapittel bemerket at begrepet undervisvurdering ekskluderer halvårsvurderinger i denne studien. Både formativ og summativ vurdering kan gis som undervisvurdering. Formativ vurdering, også omtalt som vurdering for læring eller undervisvurdering, har blant annet som hensikt å fremme læring og motivasjon. Samtidig skal den bidra til en positiv selv vurdering og skape mestringsforventninger. Mens, summativ vurdering eller vurdering av læring, har som formål å gi et uttrykk for elevens kompetansenivå. Sluttvurdering blir ofte kalt summativ vurdering, men som nevnt, summativ vurdering kan foregå også underveis i opplæringen, eksempelvis ved at eleven mottar en karakter på en prøve (Kunnskapsdepartementet, 2016; Skaalvik & Skaalvik, 2018).

Undervisvurdering er ikke én metode, den inneholder flere, men basert på studiens problemstilling vil denne studien hovedsakelig dreie seg om de to vurderingsmetodene: undervisvurdering med og uten karakterer. Disse vil videre omfatte flere ulike vurderingsmetoder. Selv om en undervisvurdering har som hensikt å bidra til motivasjon og læring, kan den også i noen tilfeller bidra til at elevens motivasjon svekkes (Skaalvik & Skaalvik, 2018). Vurdering har altså en innvirkning på elevene, og tyder på å kunne ha både en positiv og/eller negativ betydning. Dette er i tråd med Hattie og Timperley (2007), som

nevner at tilbakemelding har en innvirkning på læring, men at tilbakemeldingens effekt kan være positiv eller negativ. På bakgrunn av dette, kan det være at én vurderingsmetode er mer formålstjenlig enn andre?

En rapport fra det svenske vitenskapsrådet finner at summativ vurdering påvirker elevers prestasjoner, læring og motivasjon for læring negativt på tvers av alder. Videre nevnes det at yngre og lavtpresterende elever tyder på å være mer negativt påvirket sammenlignet med eldre og høytpresterende elever, og at voksne studenter virker å ikke være negativt påvirket av tilbakemelding i form av en karakter (Lundahl, Hultén, Klapp & Mickwitz, 2015). Butler (1987) studerer også effekten av karakterer, samt andre tilbakemeldingsformer. Elevene i studien fikk utdelt oppgaver i tre økter, hvor de etter første og andre økt mottok enten en individuell kommentar, karakter, ingen tilbakemelding eller standardisert ros, som for eksempel «veldig bra». Studiens funn viser at elevene som mottok kommentarer hadde økt interesse. Interessen til elevene som fikk ros, karakterer og ingen tilbakemeldinger hadde derimot avtatt fra første til siste økt. Samtidig etterspurte de sistnevnte elevene færre ekstraoppgaver enn de elevene som mottok kommentarer. Videre antyder studien at individuelle kommentarer vil fremme oppgaveorienterte elever, mens karakterer og ros fremmer egoorienterte elever. Begrepene oppgaveorientering og egoorientering vil bli næyere gjennomgått i delkapitlene prestasjonsorientert målstruktur og læringsorientert målstruktur.

I en senere studie av Butler (1988) rapporteres det også om effekter av ulike tilbakemeldinger. Studien er relativ lik den tidligere presenterte studien, hvor én av forskjellene er at elevene fikk utdelt enten en karakter, en kommentar eller begge deler. Resultatene viser at både høytpresterende og lavtpresterende elever som mottok kommentarer hadde økt interesse. Studien finner også forskjeller mellom høytpresterende og lavtpresterende elever. Etter andre økt, som var den siste økten hvor elevene mottok tilbakemelding, uttrykker de høytpresterende elevene lignende interesse ved alle tilbakemeldingstypene. De lavtpresterende elevene som fikk karakterer og karakterer med kommentarer uttrykte derimot en mindre interesse etter andre økt. Wiliam (2006) nevner at de fleste elever som mottar karakter og kommentar har en tendens til å sjekke karakteren først, og deretter undersøke naboens karakter. Han påpeker derfor at kommentaren mister sin hensikt dersom den gis sammen med en karakter.

Lauvås (2018) nevner også at feedback mister mye av sin verdi når det gis med karakter, da det er karakteren som blir det viktige for eleven. Karakterer tilhører den summative vurderingen, og kombinasjonen av formativ og summativ vurdering er ofte brukt i skolen (Lauvås, 2018). Basert på blant annet Butlers studie (1988), konkluderer Lauvås (2018) med at det er en ren formativ vurdering som er det beste, og ikke en kombinasjon. Formativ vurdering gir betydelige læringsgevinster, og er særlig heldig for de lavtpresterende elevene (Black & Wiliam, 1998). NOVA rapport 4/11 nevner derimot at elevene selv rapporterer at de liker å få karakterer på ungdomsskolen, og knytter sin egen innsats til karakterene (Dæhlen, Smette & Strandbu, 2011). Det er derfor interessant å studere hvilken betydning en karakterfri undervisvurdering har på elevens opplevelse av seg selv i skolen og av skolens målstruktur.

2.2 Skolens målstruktur

Læringsmiljøet betegnes som alt som har innvirkning på elevens læring i skolen og skoleopplevelse (Skaalvik & Skaalvik, 2018). For å kunne forandre opplevelsene elevene har av læringsmiljøet, må forholdene som påvirker opplevelsene endres. Et av forholdene, som både påvirker og er en del av læringsmiljøet, er skolens målstruktur (Skaalvik & Skaalvik, 2018). Dette defineres som de signalene som elevene mottar fra skolen om hva som verdsettes og hva som er viktig i skolen (Skaalvik & Skaalvik, 2005). Signalene kan sendes til elevene på mange måter og fra forskjellige kilder både utenfor og i skolen. Utenfor skolen blir det sendt signaler gjennom blant annet planer og forskrifter. Disse signalene har antakelig en indirekte effekt på elevene, altså at signalene sendes til elevene gjennom skolen og lærerne. Signaler som har direkte effekt på elevene er de som sendes fra den enkelte skolen og de enkelte lærerne. Noen signaler vil være felles for alle skoler, mens andre signaler vil variere mellom skolene og i hver enkelt skole. Signalene elevene mottar er som nevnt med å definere hvilken type målstruktur som læringsmiljøet har. De to mest nevnte hovedtypene innen forskning er prestasjonsorientert og læringsorientert målstruktur (Skaalvik & Skaalvik, 2018). Forskning viser at disse hovedtypene har flere ulike betydninger for elevene i skolen.

2.2.1 Prestasjonsorientert målstruktur

Et læringsmiljø med prestasjonsorientert målstruktur er et læringsmiljø hvor elevene mottar signaler om at deres resultater er det som er viktig for skolen og lærerne, mens innsats får en mindre verdi. I tillegg opplever elevene at deres resultater blir sammenlignet med andre elever, samt at miljøet er preget av konkurranse om oppnå bedre resultater enn de andre

elevene (Skaalvik & Skaalvik, 2005, 2017). Graden av sosial sammenligning og konkurranse i et klasserom er sammensatt av flere faktorer, en av disse er skolens vurderingsformer (Skaalvik & Skaalvik, 2005). Det er derfor naturlig å tenke at undervisvurderinger har betydning for elevenes opplevelse av læringsmiljøets målstruktur. Spørsmålet blir dermed om det foreligger en forskjell i opplevelsen av prestasjonsorientert målstruktur mellom elevene fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer.

Sistnevnte er særlig aktuelt å studere, når Skaalvik og Skaalvik (2013a) finner i sin studie at elevene opplever skolen som mindre læringsorientert og mer prestasjonsorientert fra barneskolen og over i ungdomsskolen. Sett i sammenheng med de nevnte kjennetegnene ved den prestasjonsorienterte målstrukturen, er dette et urovekkende funn. Mer bekymringsverdig blir det når Skaalvik og Skaalvik (2017) sine analyser viser til at prestasjonsorientert målstruktur har positiv betydning for elevenes opplevelse av prestasjonspress i skolen. Samtidig viser resultatene fra disse analysene at prestasjonsorientert målstruktur bidrar til lavere mental helse. Indikatorene for den mentale helsen er i denne studien nedstemthet, utmattelse, selvvverd, angst, og psykosomatiske symptomer. En rekke andre effekter ved denne målstrukturen vil presenteres i det følgende avsnittet.

Skolens målstruktur viser seg også å ha betydning for elevenes målorientering. Studie finner at prestasjonsorientert målstruktur er positivt relatert til egoorientering (Anderman & Anderman, 1999). Flere forskere antyder at egoorientering inneholder to dimensjoner (Elliot & Harackiewicz, 1996; Skaalvik, 1997; Skaalvik, Valåns & Sletta, 1994). Indikatorene i Anderman og Anderman (1999) sin studie indikerer derimot til å kun måle det Skaalvik og Skaalvik (2018) velger å kalle offensiv egoorientering. Federici, Skaalvik og Tangen (2015) inkluderer derimot både defensiv og offensiv egoorientering (performance-avoidance og performance-approach orientation) i sin studie, og finner at prestasjonsorientert målstruktur kun er signifikant relatert til defensiv egoorientering. Dette betegner elever som unngår å fremstå som inkompetente, og har som mål å unngå å få et negativt omdømme. Mens offensiv egoorientering betegner elever som ønsker å vise frem at de er bedre enn de andre elevene, samt bli oppfattet som best (Skaalvik & Skaalvik, 2018; Skaalvik et al., 1994). Forskning viser at defensiv egoorientering er negativt relatert til prestasjoner, selvoppfatning og indre motivasjon (Elliot & Harackiewicz, 1996; Sideridis & Kaplan, 2011; Skaalvik, 1997).

Offensive egoorientering viser seg derimot å være positivt relatert til prestasjoner, selvoppfatning og indre motivasjon (Skaalvik, 1997). Til tross for de ulike funnene, er det ikke anbefalt å frembringe offensiv egoorientering, da den kan utvikle seg til defensiv egoorientering når elevene møter krevende oppgaver (Federici et al., 2015). Disse funnene peker dermed i en retning av at prestasjonsorientert målstruktur er en uønsket målstruktur.

2.2.2 Læringsorientert målstruktur

I den andre hovedtypen av målstruktur, læringsorientert målstruktur, opplever elevene at det betydningsfulle i skolen er å lære og forstå. Samtidig mottar de signaler om at innsats og personlig forbedring verdsettes. Videre blir vurderingen karakterisert av å ha fokus på elevenes individuelle mål, tidligere resultater, samt selve læringsprosessen og deres læringsstrategier (Skaalvik & Skaalvik, 2005, 2017).

Som presentert litteratur viser, har vurderingen i et læringsorientert miljø flere fellestrekk med typen formativ vurdering. En masterstudie finner en samvariasjon mellom læringsorientert målstruktur og vurdering for læring, hvor målstrukturen har positiv betydning på vurdering for læring og vice versa (Eileraas, 2014). Det er dermed mye som tyder på at vurdering og elevenes opplevelse av læringsmiljøets målstruktur henger tett sammen, og det kan stilles spørsmål ved om et læringsmiljø kan være læringsorientert dersom formativ vurdering ikke praktiseres. Læringsmiljøet målstruktur tyder og på å ha betydning for elevenes opplevelse av seg selv i skolen. Som nevnt i forrige delkapittel, viser en studie at jo eldre elevene er jo mindre opplever elevene skolens læringsmiljø som læringsorientert. Dette er uheldig, da denne studien finner at læringsorientert målstruktur er positivt relatert til innsats, indre motivasjonen, skolefaglig selvverurdering og den hjelpesøkende atferden til elevene i skolen (Skaalvik & Skaalvik, 2013a).

Videre viser Skaalvik og Skaalvik (2017) sin analyse til en korrelasjon mellom læringsorientert målstruktur og bedre mental helse. Sistnevnte variabel måles som nevnt av indikatorene nedstemthet, utmattelse, selvverd, angst og psykosomatiske symptomer. Videre viser analysen at den læringsorienterte målstrukturen skiller seg fra den prestasjonsorienterte målstrukturen ved at den i tillegg har mindre innvirkning på prestasjonspresset elever opplever, selv om denne tendensen er svak. Læringsorientert målstruktur har imidlertid flere andre heldige effekter, deriblant viser Federici et al. (2015) til at læringsorientert målstruktur er signifikant positivt relatert til oppgaveorientering. Elever som er oppgaveorientert er

opptatt av å lære, samt har et ønske om å mestre oppgaver og oppnå mer innsikt (Skaalvik & Skaalvik, 2018). På grunnlag av funnene i dette delkapitlet er det mye som tyder på at den ideelle målstrukturen er en læringsorientert målstruktur.

2.3 Indre motivasjon

Oppgaveorientering kan ligne motivasjonstypen indre motivasjon (Deci & Ryan, 1985). Atferd som springer ut av interesse karakteriseres som indre motivert atferd. Det vises til to komplementære tilnærminger til definisjonen av indre motivasjon. Den ene definerer indre motivasjon som noe som bygger på menneskets grunnleggende psykologiske behov, som da er behovet for kompetanse, autonomi og tilhørighet. Mens den andre tilnærmingen til definisjonen forklarer at indre motivert atferd er uavhengig av belønning og straff, på den måten at det å utføre en interessant aktivitet er i seg selv givende (Deci & Ryan, 2000).

En ytre motivert atferd derimot drives ikke av en interesse eller glede for aktiviteten (Skaalvik & Skaalvik, 2018). Ryan og Deci (2009) presenterer fire ulike typer av ytre motivasjon, som svarer til ulike grader av autonom regulering. Videre antyder de at typene kan klassifiseres inn i kontrollert ytre og autonom ytre motivasjon. Kontrollert ytre motivert atferd utføres i situasjoner hvor eleven unngår straff eller mottar belønning ved utførelse av en aktivitet. Elever som jobber for å unngå skam for å gjøre feil er også en kontrollert ytre motivert atferd. Aktivitetene utføres altså ufrivillig. Mens autonom ytre motivert atferd baseres på at eleven har internalisert verdiene skolen har om å arbeide med fag, og aktiviteten utføres dermed frivillig. Eleven vil da arbeide med skolefag fordi denne handlingen er verdsatt, eller sett på som viktig for å kunne oppnå selvvalgte mål, og utføres ikke fordi eleven finner interessen ved fagene (Ryan & Deci, 2009; Skaalvik & Skaalvik, 2018). Skaalvik og Skaalvik (2015) nevner at formålet må være at læreren foster indre motivasjon hos hver enkelt elev, men diskuterer videre at det ikke er realistisk at alle elever utføre alle aktiviteter på bakgrunn av interesse. Å fremme autonom ytre motivasjon blir derfor også presentert som en løsning.

Indre motivasjon, er som tidligere definert, uavhengig av belønning og straff (Deci & Ryan, 2000), men belønning og straff blir allikevel brukt i skolen, da antakelig for å motivere til blant annet innsats (Dæhlen et al., 2011). Typiske eksempler på motivasjonsvirkemidler lærere bruker er karakterer og ros (Skaalvik & Skaalvik, 2018). Ifølge Deci, Koestner og Ryan (1999) underminere flere typer belønning den indre motivasjonen. Cameron (2001)

konkludere derimot til at belønning både kan ha positive, negative og ingen effekter på indre motivasjon. Hva som påvirker indre motivasjon tyder altså på å være et aktuelt forskningstema. Blant annet antyder én studie at ulike tilbakemeldingstyper har ulik effekt på indre motivasjon, ved at elever som mottar kommentarer uttrykker høyere interesse for oppgaven enn de som mottar karakterer og ingen kommentar (Butler & Nisan, 1986).

2.4 Innsats

«Innsats øker sannsynlighet for å mestre en oppgave» (Skaalvik & Skaalvik, 2005, s. 156). Det kan videre forklares som en type atferd som dreier seg om at elevene gjør sitt aller beste og forsøker å gjøre oppgavene sine (Skaalvik & Skaalvik, 2013a). Som tidligere nevnt, viser læringsorientert målstruktur å være positivt relatert til innsats. Når samme forskning viser at elevene opplever at læringsmiljøet blir mindre læringsorientert og mer prestasjonsorientert med høyere klassetrinn, så kan dette forklare funn som viser at elevenes innsatsen faller på høyere klassetrinn (Skaalvik & Skaalvik, 2013a).

Flere forskere finner at indre motivasjon har en positiv sammenheng med innsats (Skaalvik & Skaalvik, 2013a; Wendelborg, Røe & Skaalvik, 2011). Funnene fra studien til Butler og Nisan (1986) viser at omtrent halvparten av elevene i studien attribuerer deres innsats til interesse, noe som og er en indikasjon på at indre motivasjon har en sammenheng med innsats. Videre finner Butler og Nisan (1986) i sin studie at svært få elever tilskriver deres innsats til tanken om at deres resultat ville påvirke skolekarakteren. Mens i NOVA rapport 4/11 er nettopp et av hovedfunnene at flere elever sin innsats og motivasjon for skolearbeidet er et resultat av ønsket om å oppnå gode karakterer. De fleste elevene mener at karakterene motiverer til innsats. Mens for enkelte elever som ga inntrykk av å få dårligere karakterer enn forventet, uttrykker at de reduserer deres innsats på grunn av karakteren. Karakterer og karakterdrevet innsats har imidlertid sine ulemper, på den måten at den ikke er indre motivert. Flere elever legger ned en innsats fordi de vet det har relevans for dem, og det nevnes at det kan tyde på at elevene jobber kun for å oppnå gode karakterer (Dæhlen et al., 2011).

2.5 Selvvurdering av kompetanse

Skolefaglig selvvurdering nevnes som et av tre perspektiver knyttet til selvoppfatning (Skaalvik & Skaalvik, 2018). Shavelson, Hubner og Stanton (1976) presenterer en hierarkisk organisering av selvvurdering, som viser at generell selvvurdering, eller som Skaalvik og

Skaalvik (2018) velger å kalle selvvurd, formes av fire hovedområder. En av disse hovedområdene er skolefaglig selvvurdering, som videre bygger på selvvurderingen i ulike fagområder. Det er nettopp dette hovedområdet denne studien vil rette fokus mot, og den benevnes her som selvvurdering av kompetanse.

Elevenes selvvurdering av kompetanse bygges opp av deres erfaringer i skolehverdagen, og lærernes underveisvurderinger er en del av disse erfaringene (Skaalvik & Skaalvik, 2018). Ifølge Skaalvik og Skaalvik (2015) vil lærerens vurderinger også være utgangspunktet i en eventuell sosial sammenligning. Dette vil da særlig forekomme når vurderingsformen er enkel å sammenligne, eksempelvis om den inneholder en karakter (Skaalvik & Skaalvik, 2015). Festinger (1954) hevder i sin artikkel om sosial sammenligning at individet har en trang til å vurdere egne meninger og evner. Videre understreke han at i tilfeller hvor det ikke eksisterer objektive standarder, vil individet sammenligne sine meninger og evner med andres meninger og evner. Dersom slike standarder eller kriterier mangler i skolen, noe det ofte gjør, vil elevene kunne sammenligne de vurderingene de får fra læreren med vurderingene som andre elever får (Skaalvik & Skaalvik, 2015, 2018). Dette antyder Marsh og Parker (1984) sin referanserammeteori for å ha en viktig betydning for utformingen av elevenes selvvurdering av kompetanse.

Studien til Skaalvik og Skaalvik (2013a) tyder på at læringsorientert målstruktur også har betydning for selvvurdering av kompetanse, da de finner at målstrukturen både er positivt direkte og indirekte relatert til den generelle skolefaglige selvvurderingen. Samtidig finner de at prestasjonsnivået (karaktergjennomsnittet) er sterkt relatert til generell skolefaglig selvvurdering. Også en annen undersøkelse av Skaalvik og Skaalvik (2004) støtter dette, på den måten at elevenes karakter i norsk og matematikk viser seg å ha betydning for dere selvvurdering i de respektive fagene. Samtidig finner de at selvvurderingen i de fagene har en sterk betydning for senere karakterer.

2.6 Stress

At det eksisterer stress blant ungdom, støttes av flere forskere (Bakken, 2018; Eriksen et al., 2017; Lillejord, Børte, Ruud & Morgan, 2017). NOVA rapport 6/2017 viser blant annet til at omtrent halvparten av jentene og en fjerdedel av guttene i 10.trinn i Oslo plages svært mye av psykiske helseplager, og i all hovedsak knytter tiendeklassingene symptomene på disse

plagene til stress (Eriksen et al., 2017). Psykiske plager blant ungdom ble og undersøkt i fjorårets Ungdata-rapport, som viser til et høyt nivå av psykiske plager. Også her knyttes ofte symptomene på plagene til stress. Samtidig nevnes det at de unge rapporterer om å være ofte eller svært ofte stresset av skolearbeidet, dette gjelder da særlig jentene (Bakken, 2018). Basert på 33 artikler, forklarer Lillejord et al. (2017) at «(...) stress oppstår (...) fordi den enkelte opplever diskrepans eller ubalanse mellom krav og forventninger og egne ressurser» (s. 37). Flere ungdommer mener primærårsaken til stresset ligger i kravene og forventningene som kommer fra skolen eller fra individet selv (Eriksen et al., 2017). Det kan derfor tenkes at en ubalanse mellom forventninger til karakterer eller læringsresultater på prøver, samt kunnskapstester, og elevenes egne prestasjoner kan ha en innvirkning på elevens skolestress.

NOVA rapport 6/2017 viser til at ungdom knytter skolestress til blant annet prestasjonspress (Eriksen et al., 2017). Skaalvik og Federici (2015) finner at omtrent 40 prosent av guttene og omtrent halvparten av jentene i undersøkelsene rapporterer om et sterkt prestasjonspress. Videre viser annen forskning til at flere ungdomsskoleelever i Oslo er opptatt av deres egen framtid knyttet til skole og karriere, og er opptatt av å komme inn på den riktige videregående skolen eller linjen (Eriksen et al., 2017). Utsagn som: «Det å få 3 på en prøve kan ødelegge for den totale karakteren på karakterkortet, og da er *alt* ødelagt» (Eriksen et al., 2017, s. 9), og «Du kan ødelegge livet ditt hvis du får dårlige karakterer, og forbedre det hvis du får gode karakterer» (Eriksen et al., 2017, s. 64) er en indikasjon på nettopp dette. Samtidig er uttalelsene en indikasjon på karakterenes betydning for elevene, og dette sammen kan tenkes å være en medvirkende årsak til prestasjonspresset ungdom opplever i skolen. Dårlige karakterer og stress blir fremstilt som noe som ødelegger deres liv, og da elevenes budskap er ingen konkurrerende krav fra lærerne og et redusert karakterpress (Eriksen et al., 2017) er det viktig å bidra til å finne ut av hvordan presset og stresset kan reduseres.

Noe stress er derimot behøvelig for å kunne prestere godt. Stress er derfor ikke kun negativt, men kan også være positivt (Lillejord et al., 2017). Lillejord et al. (2017) viser til en sammenheng mellom prestasjoner og stress i skolen, hvor engasjement og trivsel presenteres som mulig utfall ved positivt stress, mens mulige konsekvenser av negativt stress er utmattelse, utbrenthet og helseplager. De viser og til at positivt stress øker prestasjonene, mens negativt stress gir nedgang i prestasjoner. Videre finner en annen studie at skolerelatert stress har en negativ innvirkning på elevenes indre motivasjon (Liu, 2015). Skolens

læringsmiljø skal være helsefremmende, dermed blir kunnskap om stress svært betydningsfullt for å kunne forhindre eller redusere det.

Eriksen et al. (2017) rapporterer, som tidligere nevnt, til at elever knytter stress til prestasjonspress. Samtidig viser rapporten til at elevene knytter skolestresset til press rundt det å få en god utdanning. I denne studien måles skolestress blant annet av utsagnet «Jeg følger et sterkt press for å gjøre det godt på skolen». Med utgangspunkt i dette vurderes skolestress som en konsekvens av press i skolen. Det ble innledningsvis presentert en rapport av Danmarks Evalueringsinstitut (2017), som viser til at omtrent 70 prosent av elevene fra videregående skole opplever å være helt eller delvis enig i at presset har blitt mindre med en karakterfri vurdering. Samme studie viser derimot også at noen elever opplever press, som et følge av at elevene ikke forstår ikke hvilket nivå de ligger på når de mottar vurdering uten karakterer. Det er derfor aktuelt å finne ut av om karakterfri undervisningsvurdering kan bidra til å redusere elevenes skolestress.

2.7 Bakgrunnsvariabler

I analysen vil det inkluderes fire bakgrunnsvariabler som er forventet å ha en betydning for studiens avhengige variabler. Disse bakgrunnsvariablene er kjønn, klassetrinn, foreldres utdanning og tidligere skoleprestasjoner. Dette kapitlet vil derfor inneholde en redegjørelse av hvorfor nettopp disse variablene betraktes som aktuelle bakgrunnsvariabler.

Både Bakken (2018) og Eriksen et al. (2017) viser til at flere jenter opplever mer stress knyttet til skolearbeidet enn gutter. Forskning finner og at jenter opplever mer indre motivasjon enn gutter (Skaalvik & Skaalvik, 2013a; Wendelborg et al., 2011). Skaalvik og Skaalvik (2013a) finner derimot ingen signifikant korrelasjon mellom kjønn og generell skolefaglig selv vurdering. Det er derimot, på bakgrunn av de nevnte kjønnsforskjellene når det gjelder stress og indre motivasjon, vurdert som aktuelt å ha med kjønn som bakgrunnsvariabel.

Den andre bakgrunnsvariabelen er klassetrinn. Dette valget skyldes blant annet en studie som viser at klassetrinn er signifikant korrelert med skolens målstruktur, og som da fremstiller at elever opplever læringsmiljøet som mer prestasjonsorientert og mindre læringsorientert jo eldre de blir (Skaalvik & Skaalvik, 2013a). Samme studie finner og en signifikant negativ

korrelasjon mellom indre motivasjon og klasstrinn, samt generell skolefaglig selvvurdering og klasstrinn. Funnene antyder at elevenes opplevelse av indre motivasjon og skolefaglig selvvurdering reduseres når elevene blir eldre (Skaalvik & Skaalvik, 2013a). I tillegg rapporteres det om en økning av elevers opplevelse av å være ofte eller svært ofte stresset av skolearbeidet fra 8. til 10.trinn (Bakken, 2018).

Foreldres utdanning er den neste bakgrunnsvariabelen i denne studien. Bakken (2010) rapporterer om at foreldres utdanningsnivå har økende betydning for elevenes prestasjoner i ungdomsskolen. Ettersom det er naturlig å tenke at prestasjoner henger sammen med innsats, vurderes foreldres utdanning som en aktuell bakgrunnsvariabel. Det skal også nevnes at Wendelborg et al. (2011) antyde at sosioøkonomisk bakgrunn trolig kan ha en betydning for elevenes motivasjon.

Tidligere skoleprestasjoner er den siste bakgrunnsvariabelen. Bakgrunn for dette er at Skaalvik og Skaalvik (2013a) finner i sin studie at prestasjonsnivået (gjennomsnittlig karaktersnitt) er sterkt relatert til elevenes generelle skolefaglige selvvurdering. Samtidig viser studien til at prestasjonsnivået er moderat relatert til elevenes indre motivasjon. Med utgangspunkt i Skaalvik og Skaalvik (2013a) sin studie, vurderes tidligere skoleprestasjoner som en aktuell bakgrunnsvariabel i denne studien.

2.8 Hypoteser

Med utgangspunkt i teori og tidligere forskning, samt egen intuisjon, er det utviklet forventninger til hva man vil finne i denne studien. Det er tatt hensyn til studiens problemstilling ved formulering av forventningene. Disse forventningene eller hypotesene formuleres under, og vil i kapittel fem bli bekreftet eller avkreftet.

H1- Skolens målstruktur

En skole som oppleves som prestasjonsorientert kjennetegnes av å vektlegge elevenes læringsresultater, mens en læringsorientert målstruktur verdsetter elevenes læring (Skaalvik & Skaalvik, 2005). Det kan tenkes at et læringsmiljø som velge å fjerne karakterer fra underveisvurderingen, ønsker å rette fokuset bort fra resultater og heller mot elevenes læring. Det er derfor forventet at elever fra skoler med karakterfri underveisvurdering vil oppleve målstrukturen mer læringsorientert og mindre prestasjonsorientert enn elever fra skoler med

undervisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for skolens målstruktur mellom de to gruppene.

H2- Indre motivasjon

Studien til Butler og Nisan (1986) antyder at elever som mottar tilbakemelding i form av en kommentar har høyere indre motivasjon enn elever som mottar en karakterer eller ingen kommentar som tilbakemelding. Videre finner Lundahl et al. (2015) at summativ vurdering har negativ innvirkning på elevenes motivasjon for læring, uavhengig av elevenes alder. På bakgrunn av dette utvikles hypotesen om at elever fra skoler med karakterfri undervisvurdering opplever mer indre motivasjon enn elever fra skoler med undervisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for indre motivasjon mellom de to gruppene.

H3- Innsats

En rekke studier viser til at indre motivasjon har sammenheng med innsats (Butler & Nisan, 1986; Skaalvik & Skaalvik, 2013a; Wendelborg et al., 2011). Samtidig viser forskning til at læringsorientert målstruktur er positivt relatert til innsats (Skaalvik & Skaalvik, 2013a). På bakgrunn av de tidligere hypotesene kan man forvente at karakterfri undervisvurdering har en indirekte positiv betydning for elevenes innsats.

Videre viser et av hovedfunnene i en rapport til at karakterer fremmer innsats (Dæhlen et al., 2011). En annen studie derimot finner at svært få elever attribuere deres innsats til tanken om at resultatet ville påvirke skolekarakteren. Dette er kanskje ikke uventet, da undersøkelsen elevene deltok i ble gjennomført av psykologistudenter, og det ble forklart at forskere hadde utviklet oppgavene som de skulle utføre (Butler & Nisan, 1986). Den endelige hypotesen er derfor: Elever fra skoler med karakterfri undervisvurdering opplever å ha en høy innsats, og den vil ikke skille seg mye fra den opplevde innsatsen til elever fra skoler med undervisvurdering med karakterer. Det forventes altså ingen signifikant forskjell i gjennomsnittene for innsats mellom de to gruppene.

H4- Selvvurdering av kompetanse

Lærerens vurderinger presenteres som en kilde til elevens skolefaglige selvvurdering. Vurderingene vil også være utgangspunktet i en eventuell sosial sammenligning (Skaalvik & Skaalvik, 2018). Ifølge Festinger (1954) vil sosial sammenligning forekomme i situasjoner

hvor eleven ikke har kriterier å forholde seg til. Videre antyder Marsh og Parker (1984) at sosial sammenligning har innvirkning på elevenes selv vurdering av kompetanse. Det er derfor antatt at både underveisvurdering med og uten karakterer har en effekt på elevenes selv vurdering av kompetanse.

Generell skolefaglig selv vurdering viser seg å være positivt relatert til læringsorientert målstruktur (Skaalvik & Skaalvik, 2013a). Videre viser forskning til at prestasjonsorientert målstruktur er relatert til defensiv egoorientering (Federici et al., 2015), som Skaalvik og Skaalvik (2015) hevder vil kunne true elevenes skolefaglige selv vurdering. Med utgangspunkt i den tidligere hypotesen om skolens målstruktur, samt at det kan tenkes at det er lettere å sammenligne karakterer enn en tilbakemelding uten karakterer, er det forventet at elever fra skoler med karakterfri underveisvurdering vil oppleve en mer positiv selv vurdering av kompetanse enn elever fra skoler med underveisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for selv vurdering av kompetanse mellom de to gruppene.

H5- Stress

Forskning viser at ungdomsskoleelever knytter stress i skolen til blant annet prestasjonspress (Eriksen et al., 2017). Karakterpress er en form for prestasjonspress, og ifølge Eriksen et al. (2017) er det flere som opplever at dårlige karakterer og stress ødelegger deres liv. Eriksen og Bakken (2018) nevner at skoler i dag utfører tiltak som har som hensikt å redusere stress, hvor andre vurderingsformer enn karakterer nevnes som et slikt tiltak. Ut fra dette kan man tolke at karakterfri underveisvurdering kan være et slikt tiltak. Det er derfor forventet at elever fra skoler med karakterfri underveisvurdering opplever mindre stress enn elever fra skoler med underveisvurdering med karakterer. Det er derfor forventet en signifikant forskjell i gjennomsnittene for skolestress mellom de to gruppene.

H6- Erfaringer med karakterfri skolehverdag

Professor Gamlem og professor Burner uttaler at man ikke vil få noen forbedringer kun ved å kutte karakterene og at det må jobbes med gode tilbakemeldinger (Mellingsæter, 2018; Utdanningsforbundet, 2019). Ifølge Handals uttalelse tar det tid å gjøre endringer i vurderingspraksisen (Ruud, 2018). Det kan derfor forventes at erfaringene lærerne har med tiltaket har en betydning for elevenes opplevelser. Da det er naturlig å tenke at elevenes indre motivasjon, innsats, skolestress og selv vurdering av kompetanse ikke vil endres over natten,

vil erfaringene elevene har med tiltaket også ha en betydning. I denne studien har alle elever fra samme klassetrinn lik erfaring, mens skolene derimot har ulike erfaringer. På bakgrunn av dette er det forventet at resultatene hos elever fra skole med lang erfaring med karakterfri underveisvurdering vil skille seg noe fra resultatene hos elever med kort erfaring med karakterfri underveisvurdering, samt i en større grad skille seg fra resultatene hos elever fra skoler med underveisvurdering med karakterer. Det betyr at det forventes en signifikant forskjell i gjennomsnittene for alle de avhengige variablene mellom gruppene, med unntak av gjennomsnittene for innsats. Det forventes ingen signifikant forskjell i gjennomsnittene for innsats mellom gruppene.

2.9 Andre relevante forhold som påvirker resultatene

Forskjellen mellom ungdomsskolene i denne studien er underveisvurderingen, som da enten ekskluderer eller inkluderer karakterer. Ved å studere forskjellene i datamaterialet kan det være at man finner noen indikasjoner på betydningen av karakterfri underveisvurdering. Det kan derimot være andre forhold som kan påvirke resultatene, og som dermed bidrar til å svekke studiens indre validitet. Erfaringene skolene og elevene har med karakterfri underveisvurdering, som det ble vist til i forrige delkapittel, er blant annet et forhold som kan forventes å påvirke resultatene. I denne oppgaven vil det derfor bli tatt hensyn til erfaringene skolene har med karakterfri underveisvurdering.

På bakgrunn av studiens omfang og tidsmessige begrensninger inkluderer denne studien ingen grundig undersøkelse av lærernes underveisvurderinger. Dersom lærernes tilbakemeldinger ikke bidrar til at elevene forstår hvilket nivå de ligger på, kan det med bakgrunn i rapporten til Danmarks Evalueringsinstitut (2017) ha følger for elevenes opplevelse av press. Som tidligere nevnt er skolestress vurdert som en konsekvens av press, og det kan derfor forventes at dette samtidig vil ha betydning for elevenes opplevelse av skolestress. Videre kan det være at en skole med karakterfri underveisvurderinger har erstattet karakterene med andre former for skalaer, da eksempelvis høy, middels og lav måloppnåelse. En rektor nevner i et intervju med Utdanning 11/2018 å ha brukt en slik vurderingsmåte, og at elevene da klarte å finne ut av hvilken karakterer som tilsvarte det nivået de var på (Ruud, 2018). Disse forholdene kan derfor påvirke resultatene i denne studien, og dermed påvirke studiens indre validitet.

Til slutt skal det nevnes at det er flere andre elementer enn undervisvurdering som påvirker elevenes opplevelser i skolen. Ifølge Bandura (2012) er mennesket et produkt av det gjensidige samspillet mellom personlige faktorer, atferdsfaktorer og miljøfaktorer. I likhet med Bandura, formidler Bronfenbrenner (1979) at individet er et produkt av det gjensidige forholdet mellom individet og miljø. Videre antyder han at miljøene henger sammen og påvirker hverandre (Bronfenbrenner, 1979). Eksempelvis vil lokalpolitikken påvirke skolen, som videre vil påvirke eleven (Helsedirektoratet, 2015). Samtidig kan det være andre miljøforhold på skolen som påvirker elevenes opplevelse av seg selv i skolen og av skolen. Eksempler her er lærerens undervisningspraksis og læringsmiljøet på skolen. Det er allikevel et håp om å få noen målbare indikasjoner på betydningen av karakterfri undervisvurdering, ved å forsøke å velge ut skoler som i utgangspunktet er like.

3.0 Metode

I det følgende delkapitlet vil det bli forklart hvilken forskningsmetode og forskningsdesign som er benyttet i denne studien. Prosessen knyttet til innsamling av data vil så beskrives i det neste delkapitlet. Deretter vil studiens utvalg, måleinstrumenter og statistiske analyser presenteres. Videre vil kapitlet også inneholde empirisk kvalitetssikring av studiens sammensatte mål, før jeg til slutt presenterer noen forskningsetiske betraktninger relatert til denne studien.

3.1 Forskningsmetode og design

I denne studien er det et ønske å få noen indikasjoner på hvilken betydning karakterfri undervisvurdering har for elevenes opplevelse av seg selv i skolen og av skolens målstruktur. En karakterfri undervisvurdering betyr at elevene ikke blir gitt karakterer på prøver eller oppgaver, men at de får karakter ved halvårsvurdering og sluttvurdering. I et forsøk på å måle om en karakterfri undervisvurdering har betydning for elevers opplevelser knyttet til skolestress, innsats, motivasjon med mer, ble det i utviklingen av mastergradsprosjektet bestemt at studien skulle omfatte elever ved to typer skoler: elever fra ungdomsskoler med karakterfri undervisvurdering (testgruppe) og elever fra ungdomsskoler med undervisvurdering med karakterer (kontrollgruppe). Da denne studien ønsker å søke etter målbare indikasjoner på betydningen av den karakterfrie undervisvurderingen er det vurdert som mest hensiktsmessig å benytte kvantitativ metode som forskningsmetode i denne studien. For å kunne gjennomføre dette ble det videre ansett som nødvendig å samle inn et stort utvalg. Dermed ble spørreundersøkelse beregnet som en nyttig datainnsamlingsteknikk for denne studien. Spørreskjemaet vil bli presentert i de neste delkapitlene.

I kvantitative forskningsstrategier kan det benyttes flere ulike typer forskningsdesign. Et forskningsdesign er en plan for hvordan en undersøkelse skal utarbeides. I denne studien har jeg brukt et kvasiekperimentelt design, og dette kjennetegnes av at det foregår en sammenligning mellom eksperiment- og kontrollgruppe (Ringdal, 2018). Eksperimentelt design er noe brukt innen pedagogisk forskning og benyttes for å studere årsakssammenhenger. Designet inkluderer to hovedtyper; ekte eksperiment og kvasiekperiment. Kvasiekperiment omfatter to varianter: én som kun undersøker en eksperimentgruppe med bare post-test, og en annen variant som både studerer en eksperimentgruppe og kontrollgruppe med kun post-test. Ekte eksperimenter er den andre

hovedtypen, som også inneholder to varianter. Den første inkluderer to randomiserte grupper, da både eksperiment- og kontrollgruppe, hvor det kun utføres post-test. Med randomiserte grupper menes at enhetene i utvalget fordeles tilfeldig på alle gruppene. Til forskjell fra den første varianten har den andre varianten også en pre-test (Ringdal, 2018).

I startfasen ble det bestemt å bygge studien på en test- og kontrollgruppe. Denne studien bygger derfor på et eksperimentelt design. Både på grunn av studiens omfang og tidsmessige grunner, ble det vurdert som krevende å ta i bruk ekte eksperimenter i denne studien. Det ble derimot ansett som mulig å gjennomføre et såkalt kvasieksperiment, hvor man da gjennomfører målinger på en eksperimentgruppe og en kontrollgruppe. Dette innebærer altså at man kun gjør målinger etter at tiltaket «karakterfri undervisvurdering» var igangsatt på eksperimentgruppen (post-test).

Videre hadde jeg et ønske om å inkludere elever fra flere skoler i testgruppa og kontrollgruppa. Når det gjelder fordeling av utvalget er randomisering gullstandarden, da det er nøkkelen til å sikre indre validitet (Kleven, 2014b). Indre validitet kan defineres som «Mulighetene til å gjøre sikre slutninger om årsakssammenhenger i et eksperiment» (Ringdal, 2018, s. 145). Denne studien er som nevnt ikke et ekte eksperiment, og inneholder derfor ingen tilfeldig fordeling. For å styrke validiteten i denne kvasieksperimentelle studien har jeg derfor foretatt et strategisk utvalg i et forsøk på å få til en «match» mellom elever fra testskoler med kontrollskoler. I den henseende ble det besluttet visse kriterier som skolene måtte oppfylle, og disse var at ungdomsskolene var rene ungdomsskoler, at de hadde noenlunde lik størrelse, og at prestasjonsnivået på nasjonale prøver skulle være noenlunde like. For å ytterligere styrke validiteten i denne studien bestemte jeg meg tidlig for å gjøre multivariate analyser som kunne kontrollere for andre faktorer på individnivå som kunne ha betydning for elevenes opplevelse av seg selv i skolen og av skolen (slik som elevens kjønn, foreldres utdanning, klassetrinn og tidligere skoleprestasjoner).

3.2 Datainnsamlingsprosess

Det er rimelig å anta at skolens undervisvurdering kan ha betydning for hvordan elevene opplever skolen, og spørsmålet blir derfor om det er en forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. I planleggingsfasen ble det dermed aktuelt å finne ut av hvilke faktorer som skulle undersøkes.

Redusert stress blir nevnt som ett av formålene med karakterfri undervisvurdering (Jansen, 2018; Ruud, 2018). Det ble derfor vurdert som aktuelt å undersøke hvilken effekt karakterfri undervisvurdering har for elevenes skolestress. Førsteamanuensis Gamlem nevner og at karakterfri undervisvurdering kan ha positive virkninger på skolens læringsmiljø (Ruud, 2018). Elever med en karakterfri undervisvurdering kan tenkes å unngå signaler som retter fokus på at karakterer er viktig. Det kan derfor hende at elevens opplevelse av skolens målstruktur påvirkes av den karakterfrie undervisvurderingen. Videre tyder forskning på at vurdering har effekt på elevenes indre motivasjon og innsats (Butler & Nisan, 1986; Dæhlen et al., 2011). En annen faktor, som også kan ses i sammenheng med lærers vurdering, er elevenes skolefaglige selvvurdering (Skaalvik & Skaalvik, 2018). Med dette anses det derfor også som relevant å studere temaene skolens målstruktur, indre motivasjon, innsats og selvvurdering av egen kompetanse.

For å både sikre kvalitet og spare arbeid ble det tatt i bruk sammensatte mål som allerede var utviklet og utprøvd i det longitudinelle prosjektet «Livet i skolen». Alle målene i dette store prosjektet er kvalitetssikret empirisk og er funnet å fungere tilfredsstillende. I denne studien finner man også tilfredsstillende resultater for både dimensjonalitet og intern konsistens, noe jeg vil komme tilbake til senere i kapitlet. Dette bidrar også å styrke begrepsvaliditeten, ettersom høy reliabilitet styrke validiteten. Med begrepsvaliditet menes at utsagnene faktisk måler det teoretiske begrepet det har som hensikt å måle (Ringdal, 2018). I tillegg til de seks sammensatte målene ble det lagt til fire bakgrunnsvariabler i spørreskjemaet (vedlegg 1). Studiens variabler, både de avhengige og uavhengige, vil presenteres i delkapitlet om måleinstrumenter.

Målet om å få mest mulig like grupper var styrende for hvilke skoler som kunne velges i studien, og utvalget kan derfor anses for å være trukket gjennom en strategisk utvelgelse (Ringdal, 2018). Da det ikke finnes noen oversikt over hvilke skoler som gjennomfører såkalt karakterfri undervisvurdering i Norge, kontaktet jeg Utdanningsdirektoratet for hjelp. De kunne ikke hjelpe meg, men tipset meg om å kontakte Utdanningsetaten. Ved Utdanningsetaten fikk jeg noen tips til skoler som hadde karakterfri undervisvurdering. I tillegg gjennomførte jeg egne søk og fant tips gjennom artikler og skolenes egne hjemmesider.

Etter denne kartleggingen satt jeg med navn på en rekke skoler. Deretter tok jeg kontakt med ungdomsskolene som var relevante for denne studien, og presenterte mitt prosjekt.

Rekrutteringsprosessen ble en lang prosess med flere avslag, og det skulle derfor vise seg å være vanskelig å oppfylle alle de valgte kriteriene. Jeg foretok derfor en ny kartlegging, hvor det ble lagt vekt på å få rene ungdomsskoler, samt at skolestørrelsen var omtrent lik (middels store skoler). Men, studerer man de gjennomsnittlige prestasjonsnivåene på de nasjonale prøvene i etterkant tyder resultatene på at det ikke er noen utprega forskjeller mellom gruppene.

De skolene som viste interesse fikk tilsendt et informasjonsskriv (vedlegg 2) og av disse, mottok de som ønsket å delta mer informasjon om gjennomføring og et passivt samtykke- og informasjonsskriv til foreldre (vedlegg 3). Jeg leverte spørreskjemaene selv ved alle de fem ungdomsskolene, og gjennomførte undersøkelsen sammen med klassene ved én skole. De fire andre skolene ønsket å gjennomføre undersøkelsen alene, og fikk derfor sammen med spørreskjemaene et skriv som inneholdt informasjon til elevene (vedlegg 4).

Spørreskjemaene som deltakerne mottok var utformet på den måten at de var maskinlesbare. Etter at alle data var samlet inn, var neste fase å skanne alle de utfylte skjemaene. Deretter ble all data overført inn i programmet IBM SPSS. Dette ble gjort ved hjelp av Kyrre Svarva ved SU-fakultetet ved NTNU. Da seks av spørreskjemaene inneholdt få svar, ble disse tatt ut av datasettet. Alle de andre deltakerne hadde svart for nesten hele datasettet, og det innebærer at det var få utfordringer knyttet til manglende verdier. Hvert enkelt sammensatt mål inneholder fire indikatorer eller utsagn, og i noen tilfeller var det slik at ikke alle respondentene hadde besvart alle utsagn. Hvis respondenten hadde besvart tre av fire utsagn, så ble manglende verdier erstattet ved å anvende imputasjon av gjennomsnittsverdi. Med dette menes at respondentens gjennomsnittlige verdi for de tre andre utsagn ble lagt inn som verdi for det fjerde utsagnet.

3.3 Utvalg

Det endelige utvalget i denne studien er 1071 elever, fordelt på fem ungdomsskoler. To av ungdomsskolene har undervisvurdering med karakterer. De tre andre skolene er såkalte ungdomsskoler hvor elevene har karakterfri undervisvurdering. Her har én av skolene vært *helt* karakterfri i snart tre år, mens en annen har vært snart i ett år. Med *helt* karakterfri menes

at alle trinne ved skolen har karakterfri undervisvurdering. Den siste skolen begynte også med karakterfri undervisvurdering nylig, og er spesiell da man ikke kan karakterisere hele skolen som karakterfri. Ett av trinnene får karakterer, og er derfor ikke inkludert i studien. De to andre trinnene er inkludert, men det skal merkes at et av trinnene får karakterer i ett av fagene. Denne studien vil derfor også ta hensyn til hvor lang erfaring skolene har med karakterfri undervisvurdering, da dette også kan tenkes å ha betydning for elevenes opplevelser.

For å øke muligheten for at resultatet kan generaliseres fra utvalget til populasjonen, er sannsynlighetsutvalg den beste utvalgsteknikken (Ringdal, 2018). I arbeidet med utvalget i denne studien er deltakende skoler strategisk valgt ut fra gitte kriterier. Dette påvirker studiens ytre validitet og begrenser muligheter for generalisere funn (Ringdal, 2018). Man kan med dette ikke si at elevene ved de tre testkolene er representative for alle elever fra skoler med karakterfri undervisvurdering. Det samme gjelder for kontrollkolene. Men, ved å sammenligne gruppene av elever fra skoler med ulike undervisvurderinger så kan denne studien gi noen indikasjoner på om karakterfri undervisvurdering har en betydning for elevers opplevelser av seg selv og av skolens målstruktur.

3.4 Måleinstrument

I denne studien er måleinstrumentet et spørreskjema (vedlegg 1), som består av spørsmål, påstander og svaralternativer. Del A i spørreundersøkelsen inneholder lukkede spørsmål og er studiens bakgrunnsvariabler. Egenskapene her er direkte observerbare og variablene brukes derfor for seg selv. Del B av spørreskjemaet inneholder påstander, som er variabler som er ment å måle egenskaper som ikke er direkte observerbare. S sammensatte mål benyttes for å måle slike latente variabler. Disse variablene er som nevnt indirekte observerbar, som da måles ved hjelp av et sett med indikatorer som henger sammen og har tilknytning til denne teoretiske variabelen. Det finnes to typer sammensatte mål: indekser og skalaer. I en indeks former indikatorer verdier på den latente variabelen. Mens, når indikatorer er effekter av den latente variabelen, kalles det en skala. Alle sammensatte mål i denne studien er av typen skala. Fordelen med bruk av sammensatte mål er at det dekker flere sider ved et teoretisk begrep, samtidig er de med å redusere sjansen for målefeil (Ringdal, 2013, 2018). I det følgende presenteres studiens sammensatte mål og studiens uavhengige variabler.

3.4.1 Prestasjonsorientert målstruktur (avhengig variabel)

Det nevnes at karakterfri undervisvurdering kan ha en positivt effekt på skolens læringsmiljø (Ruud, 2018). Om dette er tilfellet, kan det tenkes at elever fra skoler med karakterfri undervisvurdering opplever læringsmiljøet som mindre prestasjonsorientert. Det anses derfor som aktuelt å studere variabelen prestasjonsorientert målstruktur i denne studien. Skalaen som er ment å måle prestasjonsorientert målstruktur består av fire utsagn, som er utviklet av Midgley et al. (2000). Sammen gir utsagnene et bilde på elevenes opplevelse av det prestasjonsorienterte målstrukturen. Skalaen er gradert på en skala fra 1 til 6, hvor 1=svært usant og 6=svært sant. Påstandene er:

- (1) Det viktigste for lærerne våre er at jeg får gode karakterer
- (2) Det viktigste for lærerne våre er at jeg svarer riktig på spørsmål
- (3) Det viktigste for lærerne våre er at jeg gjør det godt på prøver
- (4) Det viktigste for lærerne våre er at jeg er flink på skolen

3.4.2 Læringsorientert målstruktur (avhengig variabel)

Det blir nevnt i forrige avsnitt at karakterfri undervisvurdering kan ha en positivt virkning på skolens læringsmiljø (Ruud, 2018). Denne studien kan dermed bidra til å finne ut om dette er tilfellet. Den læringsorienterte målstruktur måles i denne studien ved hjelp av fire utsagn. Skalaen er utviklet og testet av Midgley et al. (2000), og sammen gir utsagnene her et bilde på elevenes opplevelse av den læringsorienterte målstrukturen. Skalaen er gradert på en skala fra 1 til 6, hvor 1=svært usant og 6=svært sant. Påstandene er:

- (1) Det som betyr noe for lærerne våre er at jeg gjør så godt jeg kan i fagene
- (2) Det er viktig for lærerne våre at jeg skjønner det jeg gjør, ikke bare pugger det
- (3) Det er viktig for lærerne våre at jeg lærer nye ting
- (4) Det er helt OK for lærerne våre at jeg gjør feil i klassen, så lenge jeg lærer av det

3.4.3 Indre motivasjon (avhengig variabel)

Butler og Nisan (1986) finner at elever som mottar tilbakemelding uten karakter har høyere interesse for oppgaven enn elever som mottar karakterer og ingen kommentar. Denne studien kan bidra til å finne ut om det er en forskjell i indre motivasjon mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Variabelen indre motivasjon måles i denne studien ved hjelp av fire påstander. Denne skalaen er utviklet av Deci og Ryan (1985) og Vallerand et al. (1992), og sammen er utsagnene ment å

måle elevenes interesse for å arbeide med skolefag. Skalaen måles på en skala fra 1 til 6, der 1 =svært usant og 6=svært sant. Påstandene er:

- (1) Jeg liker å holde på med skolearbeid
- (2) Jeg synes det er morsomt å arbeide med fagene
- (3) Fagene vi har på skolen interesserer meg
- (4) Jeg har lyst til å arbeide med fagene

3.4.4 Innsats (avhengig variabel)

Indre motivasjon viser å ha en positiv sammenheng med innsats (Skaalvik & Skaalvik, 2013a; Wendelborg et al., 2011). På bakgrunn av dette og forventningene knyttet til indre motivasjon, samt at et av hovedfunnene til NOVA rapport 4/2011 er at flere elever knytter innsats til karakterer (Dæhlen et al., 2011), er det vurdert som aktuelt å studere elevenes opplevelse av innsats. Det sammensatte målet som har som hensikt å måle innsats består av fire påstander. Skalaen er utviklet av Frostad og Mjaavatn (2018) og måles i denne studien på en skala fra 1 til 6, der 1=svært usant og 6= svært sant. Påstandene er:

- (1) Jeg jobber godt med oppgavene vi får på skolen
- (2) Jeg forbereder meg så godt jeg kan til prøver
- (3) Jeg slurver mye med skolearbeidet mitt
- (4) Jeg følger godt med i timene på skolen

3.4.5 Selvvurdering av kompetanse (avhengig variabel)

Skaalvik og Skaalvik (2018) nevner lærerens vurdering som en kilde til elevenes skolefaglige selvvurdering. Det anses derfor som interessant å studere om det er en forskjell i selvvurdering av kompetanse mellom elever fra skoler med underveisvurdering med karakterer og elever fra skoler med karakterfri underveisvurdering. Skalaen selvvurdering av kompetanse er basert på fire utsagn og er hentet fra Tafarodi og Swann Jr. (1995) sin Self-Liking/Self Competence Scale (SLCS), Deci og Ryan (1985) og PISA-undersøkelsene (OECD, 2009). Skalaen er gradert på en skala fra 1 til 6, hvor 1=svært usant og 6=svært sant. Påstandene er:

- (1) Jeg har gode evner, så jeg kan få til det meste
- (2) Jeg takler utfordringen med å lære nytt stoff på skolen
- (3) Jeg gjør det bra på en rekke områder
- (4) Når jeg setter meg ned for å lære noe virkelig vanskelig, kan jeg klare det

3.4.6 Skolestress (avhengig variabel)

Ifølge NOVA rapport 6/2017 knytter elever deres skolestress til prestasjonspress. Videre uttrykker elevene i denne rapporten at dårlige karakterer og stress bidrar til å ødelegge deres liv (Eriksen et al., 2017). Redusert stress nevnes som et av hensiktene med karakterfri undervisvurdering (Jansen, 2018). Denne studien kan dermed bidra til å finne ut om nettopp dette er tilfellet. Elevenes opplevelse av skolestress er i denne studien målt ved bruk av en skala. Denne skalaen består av fire utsagn, som er basert på NOVA rapport 6/2017 (Eriksen et al., 2017). Skalaen måles på en skala fra 1 til 6, hvor 1=svært usant og 6=svært sant.

Påstandene er:

- (1) Jeg blir stresset av skolearbeidet
- (2) Jeg føler meg utslitt på grunn av skolearbeidet
- (3) Jeg har mer skolearbeid enn jeg klarer å gjøre
- (4) Jeg føler et sterkt press for å gjøre det godt på skolen

3.4.7 Bakgrunnsvariabler (uavhengig variabler)

I denne studien er kjønn, klassetrinn, tidligere skoleprestasjoner og foreldres utdanning benyttet som bakgrunnsvariabler. Den første bakgrunnsvariabelen er som nevnt kjønn, og måler om eleven er gutt eller jente. Klassetrinn er også brukt som uavhengig variabel, med verdiene 8., 9. og 10.trinn. Den tredje bakgrunnsvariabelen, tidligere skoleprestasjoner, måler hvilken karakter som er nærmest elevenes karaktersnitt ved halvårsvurderingen 2018/2019, og er derfor gradert fra 1 til 6. Foreldres utdanning er den siste bakgrunnsvariabelen, og spørsmålet knyttet til foreldres utdanning er «Har én eller begge foreldrene dine minst 3 års utdanning på universitet/høgskole?» og hvor svaralternativene er:

- (1) Nei, ingen av dem
- (2) Ja, én av dem
- (3) Ja, begge to

3.4.8 Test- og kontrollgruppe (uavhengig variabel)

På bakgrunn av studiens problemstilling er helt avgjørende å ta med test- og kontrollgruppe som uavhengige variabel i denne studien. Denne variabelen inngår ikke i selve spørreskjemaet, men er utviklet etter at alle spørreskjema var samlet inn. Verdiene på denne variabelen er «karakterfri undervisvurdering» og «undervisvurdering med karakterer». I teorikapitlet er det beskrevet at erfaringene med karakterfri undervisvurdering kan tenkes å

ha betydning for de avhengige variablene. Det er derfor også utviklet en variabel med de tre verdiene «karakterfri undervisvurdering (lang)», «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer».

3.5 Statistiske analyser

For å analysere studiens kvantitative datamateriale er programmet IBM SPSS (versjon 25) benyttet. Det er gjennomført en rekke analyser, hvor den første av dem er kvalitetssikring av studiens sammensatte mål, og resultatene fra denne analysen presenteres i det nest siste delkapitlet i dette hovedkapitlet. Deretter presenteres beskrivende analyser av studiens variabler. Videre er det utført t-tester, samt enveis variansanalyse (ANOVA), og kovariansanalyse (ANCOVA) for å kunne sammenligne gruppene i studien.

3.5.1 Prinsipale komponentanalyser og relabilitetsanalyser

For å kvalitetssikre de sammensatte målene i studien er det både gjennomført prinsipale komponentanalyser (PKA) og relabilitetsanalyser (vedlegg 5). Det er ønskelig at en skala er endimensjonal, nemlig at indikatorene måler det samme (Ringdal, 2018). Dette kan da undersøkes ved å blant annet benytte den statistiske teknikken prinsipal komponentanalyse. I denne analysen dekomponeres de målte variablene inn til sett av lineære komponenter. I analysen blir Kaiser kriterium benyttet, som betyr at man beholder komponenter med egenverdi lik eller større enn 1. Videre er det vanlig å benytte en rotasjonsteknikk, og da det antas at det kan være en sammenheng mellom komponentene benyttes oblik rotasjon (direct oblimin) (Field, 2018). Når det gjelder ladningene, tolkes alt over 0,3 eller 0,4 for å være rimelig sterk, mens ladninger over 0,7 eller 0,8 er veldig sterk (Furr, 2011).

I tillegg til dimensjonalitet brukes også reliabilitet for å vurdere kvaliteten på målene. God reliabilitet eller pålitelighet oppstår ifølge Ringdal (2018) ved «(...) at gjentatte målinger med samme måleinstrument gir samme resultat» (s. 103). For å estimere reliabiliteten kan det benyttes tre måter, hvor den mest vanlige av dem er intern konsistens. Dette måles ved bruk av Cronbachs alfa, som da estimerer reliabiliteten til en summert skala eller indeks. Alfakoeffisienten (α) vil påvirkes av den gjennomsnittlige korrelasjonen mellom indikatorene og antall indikatorer. Verdien til α vil variere mellom 0 til 1, hvor verdier over 0,7 indikerer en god indre konsistens (Furr, 2011; Ringdal, 2018). For å måle reliabilitet til de sammensatte målene er det i denne studien benyttet Cronbachs alfa.

3.5.2 Deskriptiv statistikk

Da det i denne studien er få variabelverdier og mange enheter, vil man kunne få en god oversikt over dataene gjennom en frekvensfordeling (Ringdal, 2018). Frekvensfordelingen på de avhengige variablene presenteres grafisk. Det lages seks histogram for de avhengige variablene, hvor også normalfordelingskurven er innlagt. Sammen med skjevhetsverdier kan man vurdere om fordelingen er skjev eller symmetrisk. Videre presenteres frekvensfordelingen på de uavhengige variablene i tabeller.

3.5.3 T-test og ANOVA

T-test benyttes for å sammenligne to gjennomsnitt (Field, 2018). I denne studien brukes t-test for å sammenligne gjennomsnittene mellom gruppen «karakterfri undervisvurdering» og gruppen «undervisvurdering med karakterer». Da disse to gruppene er uavhengige brukes t-test for to uavhengige utvalg. Denne analysen forteller om det er en statistisk signifikant forskjell i gjennomsnittene for de to gruppene. I denne studien er utvalget stort, og i slike tilfeller kan veldig små forskjeller mellom gruppene bli statistisk signifikante (Pallant, 2016). At noe er signifikant forteller ingenting om viktigheten av en effekt. Dette kan løses ved å måle effektstørrelsen, som her vil være et mål på hvor stor forskjell det er mellom gruppene. Det mest brukte målet her er Cohen's d, som er differansen mellom gjennomsnittene delt på standardavviket (Field, 2018; Svartdal, 2018). Når standardavvikene er ulike blir det delt på et vektet gjennomsnitt av standardavvikene i de to gruppene (Kleven, 2013). Ifølge Cohen (1969) betraktes effektstørrelsen 0,2 som liten, 0,5 som moderat, og 0,8 som stor.

Videre i analyseprosessen deles gruppen «karakterfri undervisvurdering» inn etter erfaringene med vurderingsformen, og man får dermed tre grupper. For å kunne sammenligne gjennomsnittene fra flere enn to grupper benyttes enveis-ANOVA (Pallant, 2016). Analysen vil her fortelle om det er signifikant forskjell i gjennomsnittene på variabelen på tvers av de tre gruppene. For å finne ut hvor forskjellen ligger, utføres også Post hoc test (Pallant, 2016). Også her er Cohen's d beregnet, hvor det blir brukt samme uttrykk for standardavvik.

3.5.4 ANCOVA

ANCOVA, kovariansanalyse, forteller om det er en forskjell i gjennomsnittene for studiens avhengige variabler mellom gruppene etter å ha kontrollert for en eller flere variabler. Slike variabler kalles for kovariater. I ANCOVA blir variasjonen i den avhengige variabelen som

skyldes disse kovariatene fjernet. Gruppens gjennomsnitt blir derfor justert for effekten av disse kovariatene. Videre inkluderer ANCOVA en variansanalyse, som vil fortelle om det er en statistisk signifikant forskjell i de justerte gjennomsnittene for variabelen mellom gruppene. Man kan derfor si at ANCOVA er en utvidelse av ANOVA (Field, 2018; Pallant, 2016). I IBM SPSS legges skårene på skalaene som avhengige variabler, skårene på bakgrunnsvariablene er kovariatene, og den uavhengige variabelen med verdiene «karakterfri undervisvurdering» og «undervisvurdering med karakterer» legges inn som faktor i første analyse. I neste ANCOVA-analyse legges den uavhengige variabelen med verdiene «karakterfri undervisvurdering (lang)», «karakterfri undervisvurdering (kort)», og «undervisvurdering med karakterer» inn som faktor. Også her beregnes Cohen's d, gjennom å dividere differansen mellom de justerte gjennomsnittene med det ujusterte standardavviket (Johansen, 2018). Det blir også her brukt samme uttrykk for standardavvik når Cohen's d beregnes.

3.6 Kvalitetssikring av studiens sammensatte mål

Kvalitetssikring gjennomføres for å se om studiens sammensatte mål fungerer tilfredsstillende. For å vurdere kvaliteten til målene sjekkes dimensjonalitet og reliabilitet. Dimensjonalitet på målene sjekkes i denne studien ved hjelp av to separate prinsipale komponentanalyser. I tabell 6 (vedlegg 5) vises mønstermatrisen fra prinsipal komponentanalyse av de avhengige variablene indre motivasjon, skolestress, selvvurdering av kompetanse og innsats. Analysen resulterte i fire komponenter, hvor hver enkelt avhengige variabel ladet høyest på ulike komponenter, noe som da indikerer endimensjonalitet. Videre er det et godt tegn at alle ladninger er over 0,6, da alt over 0,3 eller 0,4 vurderes som rimelig sterk (Furr, 2011). Den prinsipale komponentanalysen av de avhengige variablene prestasjonsorientert og læringsorientert målstruktur vises i tabell 7 (vedlegg 5). Dette resulterte i to komponenter, og også her kom hver enkelt avhengige variabel ut som en komponent. Alle ladningene er over 0,7, noe som tolkes som veldig sterk (Furr, 2011). Målene vurderes derfor som endimensjonale.

For å empirisk kvalitetssikre de sammensatte målene, er også den indre konsistensen i indikatorene sjekket. Indre konsistens er et mål på reliabilitet, og forteller hvilken sammenheng det er mellom indikatorene i et sammensatt mål (Clausen & Johansen, 2012). Cronbachs alfa er i denne studien benyttet for å måle den indre konsistensen. Det fremgår av

tabell 6 og 7 i vedlegg 5 at alfaverdiene til indre motivasjon ($\alpha = 0,91$), skolestress ($\alpha = 0,78$), selvvurdering av kompetanse ($\alpha = 0,84$), innsats ($\alpha = 0,81$), prestasjonsorientert målstruktur ($\alpha = 0,81$) og læringsorientert målstruktur ($\alpha = 0,78$) befinner seg alle over 0,7. Dette indikerer at skalaene har en god indre konsistens (Ringdal, 2018). Det er derimot verdt å legge merke til at indre motivasjon har en høy alfa-verdi, noe som kan indikere at indikatorene er for like. De bivariate korrelasjonene forteller hvor sterkt indikatorene korrelerer, hvor det ideelle er mellom 0,3 til 0,6 (Clausen & Johansen, 2012). I dette tilfellet er de bivariate korrelasjonene over 0,6 (0,67-0,74). Dette vurderes som sterkt, men det indikerer og at indikatorene ikke er helt like.

3.7 Etske betraktninger

Enhver forskning skal integrere forskningsetikk. Innen forskning hvor mennesker blir benyttet som deltakere, skal en forsker sørge for å ta hensyn til deltakerne i forskningen. Med dette menes at forskeren ikke skal sette menneskers frihet, selvbestemmelse og integritet til side for å oppnå mer innsikt. Det skal vises respekt for menneskeverdet (Kleven, 2014a; NESH, 2016). Videre innebærer det at forskeren har et ansvar for å ivareta krav knyttet til informert samtykke, krav på privatlivet og kravet på å bli riktig presentert av data (Postholm & Jacobsen, 2018). Prosjekter som behandler personopplysninger skal meldes inn til NSD før forskningen starter. I startfasen av prosjektet ble det derfor, på NSD (u.å.) sine nettsider, sjekket for om prosjektet måtte meldes inn til NSD. Dette gjorde jeg ved å gjennomføre NSD sin test om man har et meldepliktig prosjekt i samarbeid med mine veiledere. Det hadde jeg ikke. Det avgjørende var at denne studien kun skal behandle anonyme opplysninger, og det ble verken samlet inn direkte eller indirekte personidentifiserende opplysninger. Prosjektet ble dermed ikke meldt inn til NSD. Forskeren bør allikevel ivareta etiske prinsipper både før og underveis i forskningsprosessen, samt i teksten, da den tar utgangspunkt i forskningen (Postholm & Jacobsen, 2018). I det følgende avsnittet vil jeg presentere de etiske betraktninger knyttet til denne studien.

Med informert samtykke menes at de som undersøkes skal motta tilfredsstillende informasjon om hva det betyr å delta i forskningen. Enhver forsker skal informere sine deltakere om forskningen. Dette innebærer blant annet dens formål, hvilken nytte deltagelse har og hva opplysningene skal brukes til. Videre skal det formidles om deltakernes frivillighet (NESH, 2016; Postholm & Jacobsen, 2018). Dette, samt informasjon om gjennomføring,

spørreundersøkelsen og personvern ble inkludert i informasjonsskriv som ble sendt til skolene (vedlegg 2). Fullstendig informasjon om forskningen kan bli for mye informasjon, og medføre at deltakerne ikke får med seg innholdet (Postholm & Jacobsen, 2018). Informasjonsskrivet til foreldrene (vedlegg 3) og informasjonsskrivet til elevene (vedlegg 4) var derfor mer sammenfattet og tilpasset mottakeren. Da denne studien forsker på barn, må det tas hensyn til beskyttelse av barnene (NESH, 2016). Siden studien behandler opplysninger som verken er sensitive eller personidentifiserbare, ble passivt samtykke fra foreldre vurdert som tilstrekkelig, og derfor inkludert i informasjonsskrivet som ble sendt til skolene.

Videre er det viktig at forskeren viser respekt for privatlivets fred. Det er derfor avgjørende som forsker å sørge for å at det ikke er mulig å identifisere enkeltmennesker fra datamaterialet, samtidig å ta hensyn til at områder i menneskets liv ikke skal studeres (NESH, 2016; Postholm & Jacobsen, 2018). Gjennom denne studien vises det respekt for privatlivets fred ved at det behandles verken sensitive eller personidentifiserbare data. Deltakernes anonymitet er også ivarettatt ved at det er tatt i bruk papirbaserte spørreskjema og ved at skolene i undersøkelsen er erstattet med et nummer. For å sikre skolens anonymitet i forskningsteksten blir skolene kun kort presentert, da detaljert informasjon kan bidra til å identifisere skolene.

4.0 Resultat

I dette kapitlet vil resultatene fra de statistiske analysene i SPSS bli presentert. Det vil i første del bli presentert en beskrivende analyse, som da vil inneholde elevenes svarfordeling på hvert enkelt sammensatte mål. Samtidig presenteres beskrivende analyser av studiens uavhengige variabler. For å kunne besvare studiens problemstilling er det utviklet to ulike grupperinger som vil sammenlignes i kapitlets siste del. Denne sammenligningen forekommer ved bruk av t-test, enveis ANOVA og ANCOVA, samt Cohen's d.

4.1 Beskrivelse av variablene

Dette delkapitlet vil inneholde en beskrivende analyse av studiens variabler. Til å begynne med presenteres beskrivende analyser av studiens avhengige variabler, hvor svarfordelingen på de forskjellige variablene framstilles grafisk. Deretter presenteres studiens uavhengige variabler, og det er her laget to tabeller som viser frekvensfordelingene på de uavhengige variablene.

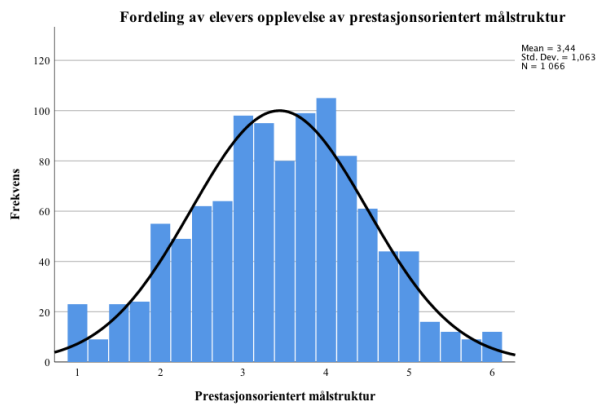
4.1.1 Studiens avhengige variabler

For å få oversikt over frekvensfordelingene er det laget seks histogram som viser svarfordelingene på hver enkelt avhengige variabel (figur 1.1-1.6). Hvert enkelt histogram inkluderer normalfordelingskurve. Ved å studere histogrammene og skjevheten (vedlegg 6) gir det en indikasjon på om variabelen er symmetrisk eller skjev. Dersom det er for mange lave verdier og halen til fordelingen er til høyre, samt at skjevhetsverdien er positiv, indikerer dette at fordelingen er positiv eller høyreskjev. En negativt eller venstreskjev fordeling har derimot for mange høye verdier, halen er til venstre og skjevhetsverdien er negativ. Jo lengre verdien er fra null, desto større sannsynlighet er det for at avhengige variabelen ikke er normalfordelt (Field, 2018; Ringdal, 2018).

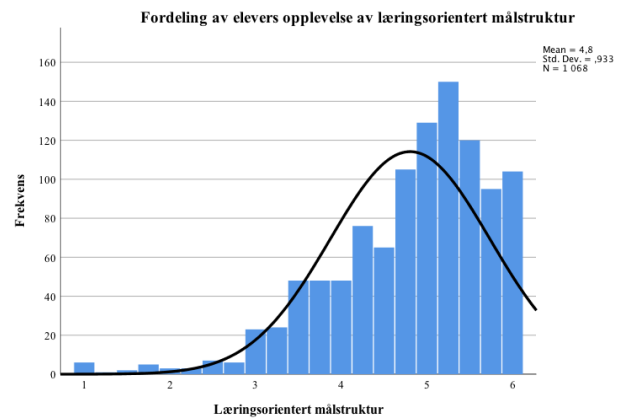
I de to første figurene, 1.1 og 1.2, vises svarfordelingen av elevers opplevelse av prestasjonsorientert målstruktur og læringsorientert målstruktur. I figur 1.1 viser skjevhetsverdien ($=-0,08$) og histogrammet at fordelingen er svakt venstreskjev og at flesteparten av elevene svarer 3 og 4. Når det gjelder fordelingen av den avhengige variabelen læringsorientert målstruktur er også denne venstreskjev, men langt tydeligere enn alle de andre variablene (skjevhet= $-1,11$). Det fremkommer her at flesteparten av elevene svarer 5 og 6.

Videre fremkommer det av skjevhetsverdien (skjevhet=-0,02) og figur 1.3 at den avhengige variabelen indre motivasjon er svakt venstreskjev. Det fremgår av figuren at det er flere elever som svarer 1 (svært usant) enn 6 (svært sant) på utsagnene knyttet til indre motivasjon. Flesteparten svarer derimot 3 og 4, noe som indikerer at de fleste elevene opplever en middels grad av indre motivasjon i skolen. En tydeligere venstreskjev fordeling vises både i figur 1.4 (skjevhet=-0,54) og figur 1.5 (skjevhet=-0,53). Det fremkommer av figurene at flesteparten av utvalget svarer 3-6 på utsagnene og at svært få svarer 1. I den siste figuren, 1.6, vises svarfordelingen av elevenes opplevelse av skolestress. Det fremgår av denne og skjevhetsverdien (= -0,06) at den avhengige variabelen skolestress er svakt venstreskjev. Det er flere elever som svarer 6 (svært sant), sammenlignet med elever som svarer 1 (svært usant). Det er derimot verdt å merke seg at flesteparten svarer 3 og 4, som indikerer at de fleste elevene opplever en middels grad av skolestress.

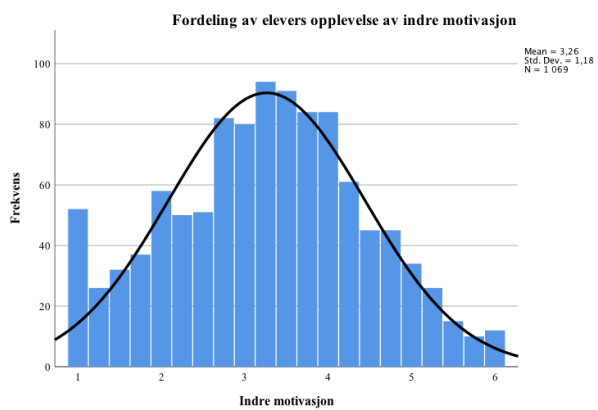
1.1 Prestasjonsorientert målstruktur



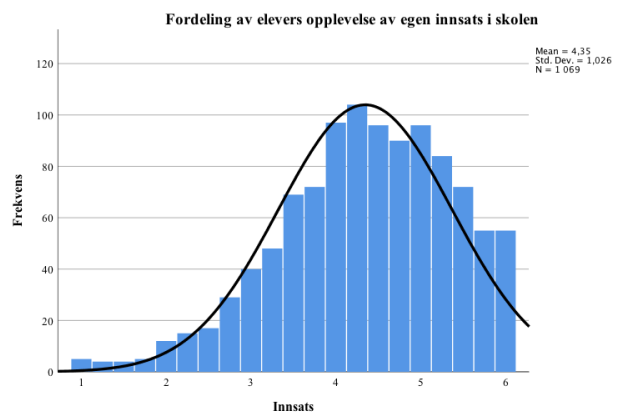
1.2 Læringsorientert målstruktur



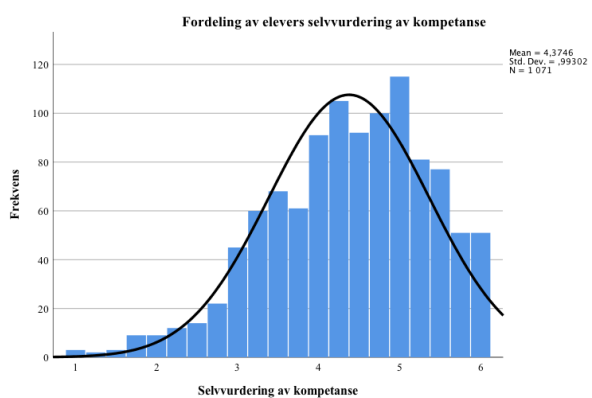
1.3 Indre motivasjon



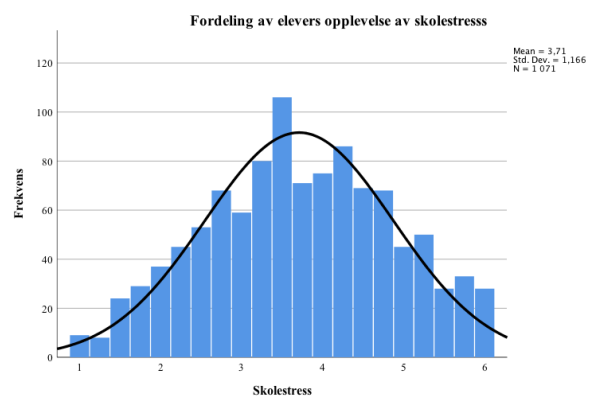
1.4 Innsats



1.5 Selvvurdering av kompetanse



1.6 Skolestress



Figur 1.1-1.6: Histogram for prestasjonsorientert målstruktur, læringsorientert målstruktur, indre motivasjon, innsats, selvvurdering av kompetanse og skolestress med inntegnet normalfordelingskurve.

4.1.2 Studiens uavhengige variabler

For å få et bilde av frekvensfordelingen på de uavhengige variablene er det laget to tabeller. Tabell 1 presenterer elevenes svarfordeling på studiens bakgrunnsvariabler, mens tabell 2 viser til fordelingen på studiens test- og kontrollgruppe. I denne studien vil det, som tidligere nevnt, foregå analyser av forskjeller mellom elever fra skoler med ulik erfaring med karakterfri undervisvurdering. Tabell 2 inkluderer derfor også en inndeling hvor det tas høyde for at skolene har kort og lang erfaring med karakterfri undervisvurdering.

Tabell 1: Elevenes svarfordelingen på kjønn, klassetrinn, tidligere skoleprestasjoner og foreldres utdanning

Bakgrunnsvariablene	Verdier	Antall	Prosent (%)
Kjønn (n=1064)	Jente	553	52
	Gutt	511	48
Klassetrinn (n=1058)	8.	356	34
	9.	411	39
	10.	291	28
Tidligere skoleprestasjoner (n=1033)	1-2	15	1
	3	139	13
	4	475	44
	5-6	404	38
Foreldres utdanning (n=1071)	Usikker/Ingen av dem	183	17
	Ja, én av dem	281	26
	Ja, begge to	607	57

Det fremgår av tabell 1 at en meget stor andel av utvalget har svart på spørsmålene knyttet til bakgrunnsopplysninger. Når det gjelder kjønn er fordelingen omtrentlig lik. Videre fremkommer det av tabell 1 at det er flest 9. klasseelever som har gjennomført spørreundersøkelsene, men fordelingen på klassetrinn tyder allikevel å være noe lik. Fordelingen på tidligere skoleprestasjoner indikerer derimot å være noe ulik, hvor flesteparten tyder å ha et karaktersnitt på 4-6. Når det gjelder svarfordelingen på foreldres utdanning fremgår det av tabell 1 at flesteparten (83%) av elevene har én eller to foreldre med høyere utdanning.

Tabell 2: Frekvensfordelingen på test- og kontrollgruppe, samt frekvensfordelingen på gruppene «karakterfri undervisvurdering (lang)», «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer».

Uavhengige variabler	Verdier	Antall	Prosent (%)
Test- og kontrollgruppe (n=1071)	Karakterfri undervisvurdering	641	60
	Undervisvurdering med karakterer	430	40
Erfaring med karakterfri undervisvurdering og undervisvurdering med karakterer (n=1071)	Karakterfri undervisvurdering (lang)	214	20
	Karakterfri undervisvurdering (kort)	427	40
	Undervisvurdering med karakterer	430	40

Med tanke på oppgavens problemstilling er det interessant å studere forskjellen mellom studiens test- og kontrollgruppe. Det fremgår av tabell 2 at antall elever med karakterfri undervisvurdering er noe høyere enn antall elever som mottar undervisvurdering med karakterer. Dette kommer av at testgruppen inkluderer elever fra tre skoler, mens kontrollgruppen inkluderer elever fra to skoler. Prosentandelen viser derimot at forskjellen ikke er så stor. Videre tas det høyde for at skolens erfaring med karakterfri undervisvurdering over kortere eller lengre tid kan ha betydning for elevene, og gruppen «karakterfri undervisvurdering» deles derfor inn etter erfaring. Det fremkommer av tabell 2 at frekvensfordelingen på «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer» er tilnærmet lik og at det er færrest av elever som kommer fra skole med lang erfaring med karakterfri undervisvurdering.

4.2 Sammenligning av elever med og uten karakterer

I det følgende delkapitlet vil det foregå en sammenligning av elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Samtidig omfatter kapitlet en sammenligning av grupperinger hvor det tas hensyn til erfaringen skolene har med karakterfri undervisvurdering. For å kunne sammenligne gruppene er t-test, enveis ANOVA, ANCOVA og Cohen's d benyttet.

4.2.1 Forskjell mellom testgruppa og kontrollgruppa

T-tester blir her benyttet for å sammenligne testgruppa og kontrollgruppa i denne studien. Tabell 3 presenterer flere t-tester, og viser til ingen statistisk signifikant forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Effektstørrelsesmålet Cohen's d er også lagt inn i tabellen.

Tabell 3: Sammenligning av studiens avhengige variabler for gruppene «Karakterfri undervisvurdering» (nr. 1) og «Undervisvurdering med karakterer» (nr. 2), ved bruk av t-tester og Cohen's d (mellom gruppe nr. 1 og nr. 2)

Avhengige variabler		Gj.snitt	Std.avvik	P-verdi	Cohen's d
Prestasjonsorientert målstruktur	1	3,43	1,05	0,92	-0,01
	2	3,44	1,09		
Læringsorientert målstruktur	1	4,81	0,92	0,78	0,02
	2	4,79	0,96		
Indre motivasjon	1	3,32	1,18	0,06	0,12
	2	3,18	1,17		
Innsats	1	4,33	1,02	0,63	-0,03
	2	4,36	1,04		
Selvvurdering av kompetanse	1	4,38	0,99	0,68	0,02
	2	4,36	1,00		
Skolestress	1	3,70	1,15	0,62	-0,03
	2	3,73	1,19		

Det framgår av tabell 3 at differansen mellom gjennomsnittverdiene mellom gruppene er svært liten. Forskjellen i gjennomsnittsverdiene er mellom 0,01-0,03, med unntak av for indre motivasjon. Når man sammenligner indre motivasjon mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer finner man en forskjell på 0,14 i gjennomsnitt. Men, i likhet med alle de andre t-testene, er det ikke en statistisk signifikant forskjell mellom disse gruppene. Effektstørrelsene er også veldig små, som da indikerer at det ikke er en forskjell mellom gruppegjennomsnittene.

4.2.2 Har erfaring betydning for resultatene?

Som tidligere presentert, kan erfaringene skolen har med karakterfri undervisvurdering tenkes å ha betydning for elevenes opplevelse. På bakgrunn av dette ble gruppene «karakterfri undervisvurdering (lang)», «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer» utviklet. Disse gruppene og andelen elever i hver gruppe er presentert i tabell 2. I tabell 4 sammenlignes disse gruppene, hvor både gruppenes gjennomsnitt og standardavvik på hver enkelt variabel presenteres. For å muliggjøre en sammenligning av gjennomsnittene mellom disse tre gruppene, er det gjennomført en enveis-ANOVA, og til forskjell fra tabell 1 finner man her en statistisk signifikant forskjell mellom de tre gruppene i flere av variablene. For å finne ut hvor forskjellen ligger, er det utført Post-hoc test. Samtidig er Cohen's d presentert i tabell 4.

Tabell 4: Sammenligning av studiens avhengige variabler for gruppene «karakterfri undervisvurdering (lang)» (nr. 1), «karakterfri undervisvurdering (kort)» (nr. 2) og «undervisvurdering med karakterer» (nr.3), ved bruk av ANOVA og Cohen's d (mellom gruppe nr.1 og nr. 2 og mellom nr. 1 og nr 3, samt mellom nr. 2 og nr. 3).

Avhengige variabler		Gj.snitt	St.avvik	P-verdi	Cohen's d		
					1	2	3
Prestasjonsorientert målstruktur	1	3,26	0,97	0,02		-0,25*	-0,17
	2	3,52	1,08				0,07
	3	3,44	1,09				
Læringsorientert målstruktur	1	4,95	0,83	0,03		0,23*	0,17
	2	4,74	0,95				-0,05
	3	4,79	0,96				
Indre motivasjon	1	3,42	1,16	0,05		0,13	0,21*
	2	3,27	1,19				0,08
	3	3,18	1,17				
Innsats	1	4,45	1,06	0,12		0,17	0,09
	2	4,28	1,00				-0,08
	3	4,36	1,04				
Selvvurdering av kompetanse	1	4,55	0,96	0,01		0,25**	0,19
	2	4,30	1,00				-0,06
	3	4,36	1,00				
Skolestress	1	3,44	1,11	0,00		-0,33**	-0,25**
	2	3,82	1,15				0,08
	3	3,73	1,19				

Merk: * = Signifikant på 0,05-nivå, **=Signifikant på 0,01-nivå

Elevenes opplevelse av skolens målstruktur

Når det kommer til elevenes opplevelse av prestasjonsorientert målstruktur og læringsorientert målstruktur, indikerer tallverdien en statistisk signifikant forskjell mellom gruppene ($p < 0,05$). Videre viser post-hoc testene til at forskjellene ligger mellom elever fra skole med lang erfaring med karakterfri undervisvurdering (nr. 1) og elever fra skoler med kort erfaring med karakterfri undervisvurdering (nr. 2). Effektstørrelsene tyder også på dette, men med hensyn til effektstørrelsen, som betraktes som liten, både for prestasjonsorientert målstruktur ($d = -0,25$) og læringsorientert målstruktur ($d = 0,23$), indikerer dette at det kun er en liten forskjell mellom de to gruppene. Forskjellen mellom gruppene «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer» tilsvarer en veldig liten effektstørrelse. Videre er Cohen's d mellom gruppe «karakterfri undervisvurdering (lang)» og «undervisvurdering med karakterer» for prestasjonsorientert målstruktur $-0,17$ og for læringsorientert målstruktur $0,17$, noe som indikerer en liten forskjell, ettersom Cohen's d er rundt $0,2$. Post-hoc testene viser derimot til ingen statistisk signifikant forskjell mellom de sistnevnte gruppene (nr. 2 og nr. 3, samt nr. 1 og nr. 3), noe som indikerer at det ikke er en forskjell mellom disse gruppene.

Elevenes opplevelse av seg selv i skolen

Når det gjelder variablene indre motivasjon og innsats, fremgår det av tabell 4 at det ikke er en statistisk signifikant forskjell mellom gruppene, da p -verdi for indre motivasjon er rundet ned til $0,05$ ($p = 0,051$), og p -verdi for innsats er $0,12$. P -verdien til indre motivasjon ligger derimot i grenseland og bør derfor ikke avskrives helt. Ved å utføre en post-hoc test finner man at det er en statistisk signifikant forskjell mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Forskjellen mellom de nevnte gruppene tilsvarer en effektstørrelse på $0,21$, noe som indikerer at det kun er en liten forskjell mellom dem. Videre tyder post-hoc test og Cohen's d på at det ikke er noen forskjell mellom gruppe nr. 1. og nr. 2, samt nr. 2 og nr. 3 i indre motivasjon. Når det gjelder innsats, tilsvarer forskjellen mellom gruppe nr. 1 og gruppe nr. 2 en effektstørrelse på $0,17$, noe som indikerer en liten forskjell, men på bakgrunn av at det ikke vises til noen signifikant forskjell indikerer dette at det ikke er en forskjell mellom noen av gruppegjennomsnittene for innsats.

Det fremkommer av tabell 4 at det også er en statistisk signifikant forskjell ($p < 0,05$) mellom gruppene når det kommer til elevenes selvsvurdering av kompetanse. P-verdien til selvsvurdering av kompetanse er rundet ned til 0,01 ($p = 0,012$), og regnes derfor ikke som signifikant på 0,01-nivået. Den er derimot i grenseland, og ved å utføre post-hoc test finner man en forskjell som er statistisk signifikant på 0,01-nivået, da mellom elever fra skole med lang erfaring med karakterfri undervisvurdering (nr. 1) og elever fra skoler med kort erfaring med karakterfri undervisvurdering (nr. 2). Cohen's d ($d = 0,25$) gir også uttrykk for at det er en forskjell mellom de to gruppene, men kun en liten forskjell. Videre tilsvarer forskjellen mellom gruppe nr. 2 og gruppe nr. 3 en effektstørrelse på $< 0,2$, noe som kan tolkes som ingen forskjell. Mens forskjellen mellom gruppe nr. 1 og nr. 2 tilsvarer en effektstørrelse på 0,19, som da indikerer en liten forskjell. Da det ikke er en statistisk signifikant forskjell mellom noen av de sistnevnte gruppene (nr. 2 og nr. 3, samt nr. 1 og nr. 2), tyder dette på at det ikke er en forskjell mellom disse.

Når det gjelder skolestress, framgår det av tabell 4 en statistisk signifikant forskjell mellom de tre gruppene ($p < 0,01$). Ved å utføre en post-hoc test finner man at forskjellen ligger mellom elever fra skole som har lang erfaring med karakterfri undervisvurdering (nr. 1) og elever fra skoler som har kort erfaring med karakterfri undervisvurdering (nr. 2), samt mellom elever fra skole med lang erfaring med karakterfri undervisvurdering (nr. 1) og elever fra skoler med undervisvurdering med karakterer (nr.3). Ved å studere gjennomsnittsverdiene og Cohen's d tyder det på at den største forskjellen ligger mellom gruppe nr. 1 og gruppe nr. 2. Effektstørrelsen ($d = 0,33$) indikerer derimot en liten forskjell. Forskjellen mellom gruppe nr. 1 og gruppe nr. 3 tilsvarer en effektstørrelse på 0,25, noe som også tyder på en liten forskjell. Det kan derfor tolkes at det er en forskjell mellom de nevnte gruppene, men at den er liten.

4.2.3 ANCOVA

ANCOVA, kovariansanalyse, brukes i denne studien til å kontrollere for flere relevante kovariater. Man kan dermed få svar på om man fortsatt finner en signifikant forskjell mellom gjennomsnittene for de avhengige variablene mellom gruppene når man kontrollere for kovariatene. I denne statistiske prosedyren blir kjønn, klassetrinn, foreldres utdanning og tidligere skoleprestasjoner benyttet som kovariater. Det ble først testet for gruppene «karakterfri undervisvurdering» og «undervisvurdering med karakterer» (vedlegg 7). Det fremkommer av tabell 3 at det ikke er noen signifikante forskjeller mellom gruppene, og sammenligning ved bruk av ANCOVA viser heller ikke til at det er noen forskjeller mellom de justerte gjennomsnittene for de avhengige variablene mellom de to gruppene. Da det er antatt at skolens erfaring med karakterfri undervisvurdering kan ha betydning for elevenes opplevelse av seg selv i skolen og av skolens målstruktur er det derfor interessant å sammenligne elever fra skoler med ulik erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Da ved bruk av ANCOVA. Denne sammenligning presenteres i tabell 5, som og inkluderer Cohen's d.

Tabell 5: Sammenligning av gruppene «karakterfri undervisvurdering (lang)» (nr.1), «karakterfri undervisvurdering (kort)» (nr. 2) og «undervisvurdering med karakterer» (nr3) og studiens avhengige variabler, ved bruk av ANCOVA og Cohen's d (mellom gruppe nr. 1 og nr. 2 og mellom nr. 1 og nr. 3, samt mellom nr. 2 og nr. 3). Kovariatene er kjønn, klasstrinn, foreldres utdanning og tidligere skoleprestasjoner.

Avhengige variabler	Gj.snitt		Cohen's d		
			1	2	3
Prestasjonsorientert målstruktur	1	3,26		-0,25*	-0,14
	2	3,52			0,10
	3	3,41			
Læringsorientert målstruktur	1	4,94		0,24*	0,14
	2	4,72			-0,09
	3	4,81			
Indre motivasjon	1	3,38		0,08	0,16
	2	3,29			0,08
	3	3,19			
Innsats	1	4,40		0,09	0,03
	2	4,31			-0,06
	3	4,37			
Selvvurdering av kompetanse	1	4,46		0,08	0,10
	2	4,38			0,02
	3	4,36			
Skolestress	1	3,48		-0,32**	-0,19*
	2	3,84			0,12
	3	3,70			

Merk: * = Signifikant på 0,05-nivå, **=Signifikant på 0,01-nivå

Det fremkommer av tabell 5 at det er en signifikant forskjell i de justerte gjennomsnittene for prestasjonsorientert læringsorientert målstruktur mellom de justerte gjennomsnittene for gruppe «karakterfri undervisvurdering (lang)» (nr. 1) og «karakterfri undervisvurdering (kort)» (nr. 2). Disse to gruppene var og signifikant forskjellige før man kontrollerte for kovariatene (se tabell 4). Cohen's d indikerer også fortsatt en liten forskjell ($d=0,25$ og $d=0,24$). Videre betraktes alle effektstørrelsene i tabell 5 for å være lignende de fra tabell 4.

Når det gjelder elevenes opplevelse av indre motivasjon finner man heller ingen signifikant forskjell i de justerte gjennomsnittene mellom gruppene. Dette er ikke i tråd med funnene fra tabell 4, som da viser til en signifikant forskjell mellom gjennomsnittene for indre motivasjon mellom gruppe nr. 1 og nr. 3. Det blir derfor tolket at kovariatene har en effekt på elevenes opplevelse av indre motivasjon. Videre fremgår det av tabell 5 at Cohen's d mellom gruppe nr. 1 og gruppe nr. 2, samt gruppe nr. 1 og gruppe nr. 3 er svært liten grad redusert fra Cohen's d i tabell 4. Mens Cohen's d mellom gruppe nr. 2 og gruppe nr. 3 i tabell 5 er identiske med mål på effektstørrelse i tabell 4. Endring i effektstørrelsene kommer av at gjennomsnittene på den avhengige variabelen er justert på grunn av effekten kovariatene har på den. Da Cohen's d for de nevnte gruppene vurderes som lignende de man finner i tabell 4, kan det tolkes som at kovariatene kun har en liten betydning.

Man finner heller ingen signifikant forskjell i de justerte gjennomsnittene for innsats mellom gruppene, men dette er i tråd med funnene i tabell 4. Cohen's d har derimot endret seg, hvor Cohen's d mellom gruppe nr. 1. og gruppe nr. 2 vurderes for å være ulik fra den man finner i tabell 4. Denne endringen skyldes, som tidligere nevnt, av at gjennomsnittene på den avhengige variabelen er justert på grunn av effekten kovariatene har på den.

Videre fremgår det av tabell 5 at det ikke er en signifikant forskjell i de justerte gjennomsnittene for selvvurdering av kompetanse mellom gruppene. Dette er en forskjell fra funnene i tabell 4, som viser til en signifikant forskjell mellom gjennomsnittene for selvvurdering av kompetanse mellom gruppe nr. 1 og nr. 2. Dette funnet indikerer at kovariatene har en effekt på elevenes selvvurdering av kompetanse. Videre fremgår det av tabell 5 at alle Cohen's d er $<0,2$, som kan tolkes som ingen effekt. Alle effektstørrelsene er redusert fra de man finner i tabell 4, hvor Cohen's d mellom gruppe nr. 1 og gruppe nr. 2,

samt mellom gruppe nr. 1 og gruppe nr. 3 vurderes for å være forskjellig fra de man finner i tabell 4.

Når det kommer til elevenes opplevelse av skolestress, finner man i tabell 5 en signifikant forskjell i de justerte gjennomsnittene mellom gruppene «karakterfri undervisvurdering (lang)» og «karakterfri undervisvurdering (kort)», samt mellom gruppe «karakterfri undervisvurdering (lang)» og «undervisvurdering med karakterer». Disse gruppene var og signifikant forskjellige før man kontrollerte for kovariatene (se tabell 4). Cohen's d mellom de nevnte gruppene (mellom gruppe nr. 1 og nr. 2, samt mellom gruppe nr. 1 og nr. 3) vurderes som tilsvarende de man finner i tabell 4, som da indikerer en liten forskjell.

5.0 Diskusjon

I det følgende kapitlet vil jeg diskutere hver enkelt hypotese fra kapittel 2, i lys av funnene fra forrige kapittel, samt presentert teori og forskning. Hvert enkelt delkapittel vil derfor ta utgangspunkt i én hypotese, med unntak av hypotesen om erfaringer med karakterfri underveisvurdering som blir forsøkt besvart innunder alle delkapitler. Delkapitlene innledes med sin tilhørende hypotese, men ettersom alle delkapitlene tar utgangspunkt i hypotesen om erfaringer med karakterfri underveisvurdering presenteres den her. Hypotesen (H6) er som følgende:

Resultatene hos elever fra skole med lang erfaring med karakterfri underveisvurdering vil skille seg noe fra resultatene hos elever med kort erfaring med karakterfri underveisvurdering, samt i en større grad skille seg fra resultatene hos elever fra skoler med underveisvurdering med karakterer. Det betyr at det forventes en signifikant forskjell i gjennomsnittene for alle de avhengige variablene mellom gruppene, med unntak av gjennomsnittene for innsats. Det forventes ingen signifikant forskjell i gjennomsnittene for innsats mellom gruppene.

5.1 Skolens målstruktur

H1: Elever fra skoler med karakterfri underveisvurdering vil oppleve målstrukturen mer læringsorientert og mindre prestasjonsorientert enn elever fra skoler med underveisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for skolens målstruktur mellom de to gruppene.

Resultatene fra analysene tyder på at flesteparten av elevene opplever en middels grad av prestasjonsorientert målstruktur og en sterk grad av læringsorientert målstruktur i skolen. Skaalvik og Skaalvik (2013a) finner i sin studie at elevenes opplevelse av læringsorientert målstruktur reduseres jo eldre elevene er, mens opplevelsen av prestasjonsorientert målstruktur øker jo eldre elevene er. Deres funn er bekymringsverdig, mens de høye verdiene knyttet til læringsorientert målstruktur i denne studien tyder å være svært betryggende. Spørsmålet er om hva som er årsaken til funnene i denne studien. Kan det til dels massive fokuset på ungdoms psykiske helseplager og deres selvrapporterte opplevelse av dette skyldes stress i skolen (Bakken, 2018; Eriksen et al., 2017), ha skapt en bevissthet hos lærere i skolen om å redusere skolestress og prestasjonspress? Og at det er dette som gir utslag i denne studien?

Når det gjelder forskjeller mellom elevenes opplevelse av skolens målstruktur i skoler med karakterfri undervisvurdering og skoler med undervisvurdering med karakterer indikerer resultatene både fra t-testen og ANCOVA til ingen forskjell. Dermed avkreftes H1. Om man derimot tar hensyn til erfaringene skolene har med karakterfri undervisvurdering indikere funnene fra ANOVA, samt Cohen's d, én liten forskjell mellom elever fra skoler med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering. Etter å ha kontrollert for effekten av kovariatene, finner man fortsatt en liten forskjell mellom disse to gruppene. Dette, samt gjennomsnittene, gir en indikasjon på at elever fra skoler med lang erfaring med karakterfri undervisvurdering opplever en liten grad mindre prestasjonsorientert målstruktur og en liten grad mer læringsorientert målstruktur i skolen sammenlignet med elever fra skoler med kort erfaring med karakterfri undervisvurdering. Dermed beholdes deler av H6 knyttet til skolens målstruktur. Både ANOVA og ANCOVA viser derimot til ingen forskjell mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer, samt mellom elever fra skoler med kort erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Den andre delen av H6 knyttet til skolens målstruktur må derfor forkastes.

En prestasjonsorientert målstruktur kjennetegnes av å være et miljø preget av blant annet konkurranse og sammenligning (Skaalvik & Skaalvik, 2005). Graden av konkurranse og sammenligning i et klassemiljø hevder Skaalvik og Skaalvik (2005) påvirkes av flere faktorer, deriblant skolens vurderingsformer. Noen elever uttrykker at innføring av karakterfri undervisvurdering har bidratt til mindre sammenligning. Dette er i tråd med Gamlem sin uttalelse, om at karakterer kan bidra til et større fokus på sosial sammenligning (Mellingsæter, 2018). Når det gjelder karakterfri undervisvurdering, uttrykker flere skoler at hensikten har vært å styrke fokuset på læringen (Jansen, 2018; Ruud, 2018). Lauvås (2018) og Wiliam (2006) antyder å være enig i denne vurderingsformen, ettersom de mener at tilbakemelding miste sin hensikt når den inkluderer en karakter. Vurdering for læring og læringsorientert målstruktur har, ikke overraskende, en positiv sammenheng (Eileraas, 2014). På bakgrunn av presentert teori og meninger i dette avsnittet er det forventet en forskjell mellom test- og kontrollgruppa, men resultatene strider imot dette. Dette kan tyde på at det har ingen tilsynelatende betydning å innføre karakterfri undervisvurdering når det gjelder elevenes opplevelse av skolens målstruktur, ettersom studien ikke finner noen forskjell mellom gruppenes gjennomsnitt.

Handal, leder i Utdanningsforbundet, uttaler i Utdanning 11/2018 at det tar tid å gjøre endringer i vurderingspraksisen (Ruud, 2018). Det ble derfor utviklet en forventning om at erfaringene skolene hadde med karakterfri undervisvurdering hadde en betydning for elevenes opplevelser om skolens målstruktur. Dette tyder på å både stemme og stride med resultatene i denne studien. Hvorfor det er slik, er vanskelig å svare på innenfor rammen av denne studien. En mulig forklaring kan være at karakterfri undervisvurdering og undervisvurdering med karakterer ikke har noen ulik effekt på elevenes opplevelse av skolens målstruktur, men dette ville i så fall være en forklaring som strider med tidligere teori og forskning. Samtidig er det viktig å huske på at funnene faktisk viser til en liten forskjell mellom elevene fra skoler med lang og kort erfaring med karakterfri undervisvurdering, som da kan tyde på at skoler må erfare karakterfri undervisvurdering over lengre tid for at dette skal ha en effekt på elevenes opplevelse av skolens målstruktur. Kanskje det er av større betydning hva skolen og lærerne faktisk gjør, når karakterene først er kuttet? Hvilke signaler sender de da ut? Kan det for eksempel hende at lærerne sender ut prestasjonsorienterte signaler selv om undervisvurderingen er karakterfri? Forskning antyder at lærere opplever en middels grad av prestasjonsorientert målstruktur i skolen, som blant annet måler om lærerne opplever at skolen er svært opptatt av elevenes læringsresultater og at elevene gjør det bedre på tester enn andre elever ved andre skoler (Skaalvik & Skaalvik, 2013b). Signalene skolen sender kan komme utenfra skolen, eksempelvis myndighetene, og dette kan da underbygge en slik fortolkning i denne studien.

En annen forklaring til forskjellen mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og de med kort erfaring med karakterfri undervisvurdering, kan komme av at det er andre faktorer i skolen som er utslagsgivende på elevenes opplevelse av skolens målstruktur. Med utgangspunkt i Bandura (2012) og Bronfenbrenner (1979) vil eleven påvirkes av miljøet. Det kan derfor være andre faktorer i skolen eller utenfor skolemiljøet som påvirker eleven. Dette støttes også av Skaalvik og Skaalvik (2018), som presenterer at signaler som sendes om hva som er viktig i skolen, sendes fra kilder både utenfor og i skolen. Utenfor skolen kan elever eksempelvis påvirkes av lokalpolitikken, samt planer og forskrifter. Disse vil enten ha en direkte effekt på elevene, eller sendes indirekte gjennom blant annet skolen og lærerne. Noen av signalene elevene mottar er felles for alle skoler, mens andre signaler vil variere mellom hver enkelt skole og i hver enkelt skole (Skaalvik & Skaalvik, 2018). Det kan derfor hende at det er de andre signalene som elevene mottar, som bidrar til den lille forskjellen mellom de to elevgruppene. Dessuten kan signaler som er felles for alle

skoler, basert på myndighetenes utdanningspolitiske føringer, også være en årsak til at elevene fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer har lignende opplevelser av skolens målstruktur. Videre er, som nevnt i kapittel 3, en av deltakerskolene i denne studien som ikke fullstendig karakterfri. Denne skolen inngår som en av skolene med kort erfaring, og det kan hende dette har en betydelig effekt på elevenes gjennomsnittlige opplevelse av skolens målstruktur. Dette kan dermed også bidra til forskjellen mellom elever fra skoler med ulik erfaring med karakterfri undervisvurdering. Det skal derimot poengteres at det kun er en liten forskjell mellom disse gruppene, og at det fremgår av tabell 4 og 9 at gjennomsnittene for de to gruppene er noenlunde like, mens gjennomsnittene for de gruppene som ikke er signifikant forskjellige vurderes som tilsvarende.

Videre kan det faktisk gi grunn til bekymring når elevene i denne studien ikke bekrefter tydeligere at det er usant at læringsmiljøet preges av en prestasjonsorientert målstruktur. Tidligere forskning viser at denne målstrukturen har en sammenheng med opplevelse av prestasjonspress, lavere mental helse, samt defensiv egoorientering (Federici et al., 2015; Skaalvik & Skaalvik, 2017). Sistnevnte viser seg også å være negativt relatert til prestasjoner, selvpoppfatning og indre motivasjon (Elliot & Harackiewicz, 1996; Sideridis & Kaplan, 2011; Skaalvik, 1997). Selv om ikke gjennomsnittene knyttet til prestasjonsorientert målstruktur er høye, er det allikevel betydningsfullt at enhver skole jobber mot å redusere nivået av den prestasjonsorienterte målstrukturen og fortsetter å jobbe for å øke den læringsorienterte målstrukturen. Gjennomsnittene for læringsorientert målstruktur er i denne studien nærmest svaralternativ 5, som indikerer at elevene opplever en sterk grad av læringsorientert målstruktur. Dette er betryggende, særlig når denne målstrukturen viser å ha positiv sammenheng med innsats, indre motivasjon og hjelpesøkende atferd (Skaalvik & Skaalvik, 2013a), bedre mental helse og oppgaveorientering, samt viser forskning at den har mindre innvirkning på prestasjonspress enn prestasjonsorientert målstruktur (Federici et al., 2015; Skaalvik & Skaalvik, 2017). Som allerede nevnt kan man velge å tolke dette til at skoler og lærere nå har utviklet en bevissthet, som følge av en dels massiv oppmerksomhet omkring unges psykiske helseplager og opplevelse av stress i skolen.

5.2 Indre motivasjon

H2: Elever fra skoler med karakterfri undervisvurdering opplever mer indre motivasjon enn elever fra skoler med undervisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for indre motivasjon mellom de to gruppene.

Svarfordelingen på variabelen indre motivasjon tyder på at flesteparten av elevene opplever en middels grad av indre motivasjon i skolen. Videre indikere t-test og ANCOVA at det ikke er noen forskjell i gjennomsnittene for indre motivasjon mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Med utgangspunkt i dette kan også H2 forkastes, noe som videre kan være en indikasjon på at det har ingen tilsynelatende betydning å innføre karakterfri undervisvurdering når det gjelder elevenes opplevelse av indre motivasjon, ettersom gjennomsnittene ikke er forskjellig. Om man derimot tar i betraktning om skolene har lang eller kort erfaring med karakterfri undervisvurdering, viser ANOVA til en liten forskjell i elevens opplevelse av indre motivasjon mellom elever fra skoler med lang erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Denne forskjellen finner man derimot ikke i de justerte gjennomsnittene for indre motivasjon. Det kan dermed ikke trekkes en slutning om at elever fra skole med lang erfaring med karakterfri undervisvurdering er forskjellig fra elever fra skoler med undervisvurdering med karakterer når det gjelder deres opplevelse av indre motivasjon. Selv om gjennomsnittet er høyest for elever fra skole lang erfaring med karakterfri undervisvurdering og nest høyest for elever fra skoler med kort erfaring med karakterfri undervisvurdering, indikerer ANCOVA at de justerte gjennomsnittene er lignende, og H6 knyttet til indre motivasjon må derfor også forkastes.

Kovariatene i denne studien er kjønn, klassetrinn, foreldres utdanning og tidligere skoleprestasjoner. I denne studien tyder disse å ha en effekt på indre motivasjon. Dette er i tråd med forskning som viser til forskjeller mellom kjønn, alder og prestasjonsnivå når det kommer til opplevd indre motivasjon (Skaalvik & Skaalvik, 2013a; Wendelborg et al., 2011). Samtidig er det antydnet at foreldres sosioøkonomiske bakgrunn også kan ha effekt på elevenes motivasjon (Wendelborg et al., 2011). Det er derfor ikke overraskende at kovariatene har en effekt på den avhengige variabelen indre motivasjon. Det er derimot uventet at det ikke er en forskjell mellom gjennomsnittene for indre motivasjon, ettersom en rekke studier antyde at kommentarer fremme indre motivasjon og interesse, mens karakterer underminere indre

motivasjon og interesse (Butler, 1987, 1988; Lundahl et al., 2015). Studien til Butler og Nisan (1986) antyder også at kommentarer har en mer positiv effekt på indre motivasjon enn karakterer, da deltakerne som mottok kommentarer i studien uttrykte en høyere interesse for oppgaven enn de som mottok karakterer. Erfaringene skolene har med tiltaket karakterfri undervisvurdering tyder heller ikke å ha en betydning for elevenes opplevelse av indre motivasjon.

Hvordan kan man så forstå dette? Ifølge Bandura (2012) og Bronfenbrenner (1979) blir individet påvirket av miljøet. Det er med utgangspunkt i dette derfor naturlig å tenke at det er andre faktorer i skolen eller utenfor skolen som har innvirkning elevenes opplevelse av indre motivasjon. Deci og Ryan (2000) presenterer to komplementære tilnærminger til definisjonen av indre motivasjon. Den første definerer indre motivasjon som noe som er basert på de grunnleggende psykologiske behovene kompetanse, autonomi og tilhørighet. Dette kan tolkes dit hen at det er helt avgjørende for elevenes indre motivasjon at skolen og lærerne har et særskilt fokus på hvordan de legger til rette for at elever får erfare kompetanse, autonomi og tilhørighet. Disse, samt andre faktorer i skolemiljøet, kan derfor være helt avgjørende for elevenes opplevelse av indre motivasjon. Om disse ikke er tilstede, kan det være at det er ubetydelig for elevenes indre motivasjon om undervisvurderingen inneholder karakterer eller ikke.

Den andre tilnærmingen definerer indre motivert atferd som noe springer ut av interesse, og er uavhengig av belønning og straff (Deci & Ryan, 2000). Belønning og straff blir allikevel brukt i skolen, antakelig for å motivere barn til innsats i skolen. Flere elever rapporterer også at karakterer motiverer til innsats (Dæhlen et al., 2011). Når karakterer er et motivasjonsvirkemiddel, virker karakterer som en belønning for elevene, og denne motivasjonen viser kjennetegn på den motivasjonen som Ryan og Deci (2009) antyder kan klassifiseres som en kontrollert ytre motivasjon. Skaalvik og Skaalvik (2015) uttrykker at det er den indre motivasjonen og den autonome ytre motivasjonen som er de ønskelige motivasjonstypene. Ifølge Deci et al. (1999) underminerer flere typer belønning elevenes indre motivasjon, og på bakgrunn av dette betraktes belønning som uheldig. En mulig årsak til hvorfor funnene antyder at karakterfri undervisvurdering og undervisvurdering med karakterer har lik effekt på elevenes indre motivasjon kan være at elevene som har karakterfri undervisvurdering mottar andre former for belønning. Anerkjennelse fra læreren er et eksempel på et motivasjonsvirkemiddel (Skaalvik & Skaalvik, 2018), og som da fungerer som

en belønning. Ifølge Cameron (2001) kan derimot belønning ha både positiv, negativ og ingen effekt på indre motivasjon. Med utgangspunkt i dette kan det være at belønningene elevene mottar medfører at elevene opplever en lignende indre motivasjon. Dette kan derfor være en medvirkende årsak til hvorfor hypotesen må forkastes.

Videre kan en ikke signifikant forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer skyldes at vurderingene uten karakterer er erstattet med en annen form for vurderingsskala. En rektor nevner i et intervju med Utdanning 11/2018 å ha brukt en vurderingsmetode hvor elevene mottok lav, middels eller høy måloppnåelse, og at elevene regnet raskt ut hvilken karakter dette tilsvarte (Ruud, 2018). I et slikt tilfelle kan det stilles spørsmål ved om det egentlig er en forskjell mellom karakterfri undervisvurdering og undervisvurdering med karakterer. For å få en forandring, er det som tidligere nevnt, ikke nok å droppe karakterene fra vurderingene. Det må samtidig jobbes med gode tilbakemeldinger (Mellingsæter, 2018; Utdanningsforbundet, 2019). Det kan derfor være at tilbakemeldingene ikke er gode nok, og at det er som gjør at resultatene indikere at karakterfri undervisvurdering og undervisvurdering med karakterer har lik innvirkning på elevenes opplevelse av indre motivasjon.

Funnene i denne studien kan også muligens knyttes til press. Liu (2015) finner i sin studie at skolestress har en negativ effekt på indre motivasjon. I denne studien er skolestress, som nevnt i kapittel 2, ansett som en konsekvens blant annet press. Det kan derfor tenkes at elevenes indre motivasjon har sammenheng med presset de opplever i skolen. Skaalvik og Federici (2015) finner at flere elever opplever prestasjonspress i skolen. Ifølge Eriksen et al. (2017) sin rapport uttrykker flere elever for å være opptatt av egen framtidig utdanning og karriere. Samme rapport antyder at dette fokuset også bidrar til at karakterene får en viktig betydning for elevene. Dersom presset da er knyttet til de avsluttende karakterene kan det være at presset oppleves både av elever som mottar karakterfri undervisvurdering og elever som mottar undervisvurdering med karakterer. Dersom prestasjonspresset elevene opplever underminerer elevenes interesse og indre motivasjon for skolearbeidet, kan dette være en medvirkende årsak til hvorfor for gruppegjennomsnittene for indre motivasjon tyder å være lignende.

5.3 Innsats

H3: Elever fra skoler med karakterfri underveisvurdering opplever å ha en høy innsats, og den vil ikke skille seg mye fra den opplevde innsatsen til elever fra skoler med underveisvurdering med karakterer. Det forventes altså ingen signifikant forskjell i gjennomsnittet for innsats mellom de to gruppene.

Frekvensfordelingen på variabelen innsats tyder på at flesteparten av elevene opplever å ha en middels til sterk innsats knyttet til skolearbeidet. Videre viser t-test og ANOVA at det ikke er forskjell i gjennomsnittene for innsats mellom elever fra skoler med karakterfri underveisvurdering og elever fra skoler med underveisvurdering med karakterer. Dette er med på å bekrefte hypotesen om at elever fra skoler med karakterfri underveisvurdering har en høy innsats, og at det ikke er signifikant forskjell i gjennomsnittene for innsats mellom test- og kontrollgruppa. De respektive analysene viser til en forskjell på 0,03 i gjennomsnittene, og indikerer at elevenes innsats er omtrent like. Dette tyder da på at det har ingen tilsynelatende betydning å innføre karakterfri underveisvurdering når det gjelder elevenes opplevelse av innsats, ettersom gjennomsnittene er lignende. Samtidig tyder det heller ikke erfaring med karakterfri underveisvurdering å ha en betydning for elevenes innsats, da både ANOVA og ANCOVA ikke viser til noen forskjeller mellom gruppegjennomsnittene. Det skal derimot nevnes at gjennomsnittet er lavest hos elever fra skoler med kort erfaring med karakterfri underveisvurdering, men forskjellene er så små at gjennomsnittene kan karakteriseres som lignende. Derfor beholdes også H6 knyttet til innsats.

Forventingene til denne studien er basert på forskning som antyder at innsats kan være relatert til indre motivasjon (Butler & Nisan, 1986; Skaalvik & Skaalvik, 2013a; Wendelborg et al., 2011), og at karakterer kan ha effekt på innsats (Dæhlen et al., 2011). På bakgrunn av at det ikke er noe forskjell i elevenes opplevelse av indre motivasjon, utvikles det derimot nye forventinger til funnene, altså at elever som mottar underveisvurdering med karakterer vil ha en høyere innsats, da deres innsats også vil kunne drives av karakterer. Dette er derimot ikke tilfellet, og en mulig forklaring kan være at elever fra skoler med karakterfri underveisvurdering mottar andre former for belønning som motiverer til innsats.

Tidligere forskning finner at læringsorientert målstruktur er positivt relatert til elevenes innsats (Skaalvik & Skaalvik, 2013a). Funnene i denne studien viser til at det ikke er en signifikant forskjell i gjennomsnittene for læringsorientert målstruktur mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Dette kan derfor også være en medvirkende årsak til at det ikke er noe forskjell mellom elevenes innsats. Resultatene viser derimot til en forskjell mellom elever fra skoler med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering når det gjelder elevenes opplevelse av læringsorientert målstruktur. Funnene tyder derimot på at det kun er en liten forskjell, noe som kan være årsaken til at det ikke er forskjell mellom gjennomsnittene for innsats mellom de respektive gruppene.

En annen forklaring kan tenkes å ha sammenheng med prestasjonspresset elever opplever i skolen. En studie av Skaalvik og Federici (2015) viser at flere elever rapporterer om et sterkt prestasjonspress. Karakterpress er en form for prestasjonspress og NOVA rapport 6/2017 antyde at karakterer er svært viktig for elevene (Eriksen et al., 2017). Ifølge Dæhlen et al. (2011) rapporterer noen elever om at karakterer legger en demper på innsatsen, mens flesteparten rapporterer at karakterene bidrar til innsats i skolen. Det er derfor naturlig å tenke at det kan være en sammenheng mellom prestasjonspresset blant unge og den innsatsen elevene legger ned. For å kunne oppnå de avsluttende karakterene som elevene selv mener de behøver for å komme inn på videregående skole, må de legge ned en innsats, uavhengig om de mottar en undervisvurdering med eller uten karakterer. En ulempe med en innsats drevet av karakter og planer for fremtiden er at den ikke drives av en indre motivasjon. Denne innsatsen kan tyde på en stressrelatert nødvendighet, som kommer av presset om å komme inn på riktig videregående skole og å få høyere utdanning. Utsagn som «Du kan ødelegge livet ditt hvis du får dårlige karakterer, og forbedre det hvis du får gode karakterer» (Eriksen et al., 2017, s. 64) kan indikere at elevene opplever at de må legge ned en innsats. Det kan da tyde på at elevenes innsats drives av motivasjonstypen Ryan og Deci (2009) kaller for kontrollert ytre motivasjon.

5.4 Selvvurdering av kompetanse

H4: Elever fra skoler med karakterfri undervisvurdering vil oppleve en mer positiv selvvurdering av kompetanse enn elever fra skoler med undervisvurdering med karakterer. Det forventes altså en signifikant forskjell i gjennomsnittene for selvvurdering av kompetanse mellom de to gruppene.

Frekvensfordelingen på den avhengige variabelen selvvurdering av kompetanse indikere at flesteparten av elevene opplever en middels til sterk grad av selvvurdering av kompetanse. Dette er et godt tegn, ettersom forskning viser at skolefaglig selvvurdering er positivt relatert til indre motivasjon, innsats og senere prestasjoner (Skaalvik & Skaalvik, 2004, 2013a). Videre viser analysemetodene t-test og ANCOVA til ingen signifikant forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Dermed forkastes H4. Om man derimot tar i betraktning at skolene har kort og lang erfaring med karakterfri undervisvurdering, tyder både ANOVA og Cohen's d på at det er liten forskjell mellom elever fra skoler med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering. Denne forskjellen finner man derimot ikke i de justerte gjennomsnittene for selvvurdering av kompetanse. Derfor må også H6 knyttet til selvvurdering av kompetanse forkastes.

I denne studien tyder kovariatene å ha effekt på selvvurdering av kompetanse. Dette er i tråd med forskning som viser til forskjeller mellom klassetrinn og prestasjonsnivå når det kommer til skolefaglig selvvurdering (Skaalvik & Skaalvik, 2013a). At karakterer har en betydning for elevenes selvvurdering, støttes også av en annen studie som viser at karakterer i norsk og matematikk har betydning for elevenes selvvurdering i de respektive fagene (Skaalvik & Skaalvik, 2004). Forskning viser derimot ikke til en signifikant sammenheng mellom kjønn og skolefaglig selvvurdering (Skaalvik & Skaalvik, 2013a). Det kan derfor ikke forventes at hvert enkelt kovariat har en effekt på selvvurdering av kompetanse. Dette kan også være tilfellet i studien, men i denne studien blir kovariatene kun benyttet til å få noen indikasjoner på betydningen av karakterfri undervisvurdering. Da kovariansanalysen ikke viser til noen forskjell mellom gjennomsnittene for selvvurdering av kompetanse mellom gruppene, indikerer dette at det har ingen tilsynelatende betydning å innføre karakterfri undervisvurdering når det gjelder selvvurdering av kompetanse. Dette strider da som nevnt med hypotesene i denne studien.

Hvordan skal man så forstå dette? Skolefaglig selvvurdering er dannet av elevenes erfaringer i skolen, og lærerens vurderinger er en del av disse. Vurderingene elevene mottar fra lærerne vil også brukes i en eventuell sosial sammenligning (Skaalvik & Skaalvik, 2015). Marsh og Parker (1984) antyde at sosial sammenligning har innvirkning på elevenes selvvurdering av kompetanse. Det er derfor forventet at undervisvurdering har en effekt på elevenes selvvurdering. Karakterer presenteres som en vurderingsform som er enkel å sammenligne (Skaalvik & Skaalvik, 2015), og det er tenkt at karakterer er enklere å sammenligne enn en tilbakemelding uten karakter. Det er derfor forventet en forskjell mellom gjennomsnittene for selvvurdering av kompetanse mellom gruppene, dette er derimot ikke tilfellet. En forklaring kan være at i en karakterfri undervisvurdering er karakterene kun erstattet med andre vurderinger som er enkle å sammenligne, deriblant lav, middels og høy måloppnåelse. Ifølge Festinger (1954) vil sosial sammenligning forekomme i situasjoner hvor elevene ikke har kriterier å forholde seg til. Sosial sammenligning vil derfor forekomme uavhengig om undervisvurderingen inneholder karakterer eller ikke, og kan derfor også være en forklaring til hvorfor funnene indikerer at det ikke er forskjell i gjennomsnittene i selvvurdering av kompetanse mellom gruppene.

Ifølge Skaalvik og Skaalvik (2013a) sin studie er læringsorientert målstruktur positivt relatert til skolefaglig selvvurdering. Videre viser Federici et al. (2015) til at prestasjonsorientert målstruktur er relatert til defensiv egoorientering, som er en målorientering Skaalvik og Skaalvik (2015) hevder true elevenes skolefaglig selvvurdering. På bakgrunn av hypotesen knyttet til skolens målstruktur, ble det derfor forventet at denne studien ville finne en forskjell i gjennomsnittene for elevenes selvvurdering av kompetanse mellom gruppene. En antagelse til hvorfor dette ikke er tilfellet, kan ha en sammenheng med at hypotesen knyttet til skolens målstruktur (H1) ble forkastet. Studien viser derimot til en forskjell i gjennomsnittene for skolens målstruktur mellom elever fra skoler med lang erfaring og kort erfaring med karakterfri undervisvurdering. Denne forskjellen er derimot veldig liten, og kan være en medvirkende årsak til hvorfor funnene viser at det ikke er en forskjell i de justerte gjennomsnittene for selvvurdering av kompetanse mellom elever fra skole med lang erfaring med karakterfri undervisvurdering, elever fra skoler med kort erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer.

I denne studien måles elevens generelle selvvurdering av kompetanse. Shavelson et al. (1976) uttrykker at denne selvvurderingen bygges opp av elevens skolefaglige selvvurdering i ulike fag, og som ifølge Skaalvik og Skaalvik (2018) i sin tur nevner bygges opp av elevenes erfaringer. Underveisvurderingene elevene mottar vil være en del av disse erfaringene, og disse vil, som nevnt, også være utgangspunktet i en sosial sammenligning (Skaalvik & Skaalvik, 2015, 2018). Erfaringen elevene da tar med seg etter å ha sammenlignet egne vurderinger med andre elevers vurderinger, vil da kunne påvirke elevens selvvurdering av kompetanse. Sistnevnte antyder også Marsh og Parker (1984). Erfaringer er derimot ikke utelukkende koblet til erfaringer fra underveisvurderinger og sosiale sammenligninger. Eleven vil ha erfaringer fra ulike fag i skolen, som sammen bygge elevenes selvvurdering av kompetanse. Det kan derfor hende at det er andre erfaringer som har en sterk betydning for elevenes selvvurdering av kompetanse, og som da er en forklaring til hvorfor hypotesen (H4) må forkastes.

5.5 Stress

H5: Elever fra skoler med karakterfri underveisvurdering opplever mindre stress enn elever fra skoler med underveisvurdering med karakterer. Det er derfor forventet en signifikant forskjell i gjennomsnittene for skolestress mellom de to gruppene.

Svarfordelingen på den avhengige variabelen skolestress indikere at de fleste elever opplever en middels grad av stress i skolen. Ifølge H5 er det forventet å finne en forskjell mellom elevenes stressnivå i skolen mellom elever fra skoler med karakterfri underveisvurdering og elever fra skoler med underveisvurdering med karakterer. I likhet med alle andre t-tester og ANCOVA analyser er det heller ingen signifikant forskjell i gjennomsnittene for skolestress mellom disse to gruppene. Dermed forkastes H5. Tar man derimot høyde for at skolene har kort og lang erfaring med karakterfri underveisvurdering, finner man noen interessante funn. Både ANOVA og ANCOVA viser til en liten forskjell mellom elever fra skole med lang erfaring med karakterfri underveisvurdering og elever fra skoler med kort erfaring med karakterfri underveisvurdering, samt en liten forskjell mellom elever fra skole med lang erfaring med karakterfri underveisvurdering og elever fra skoler med underveisvurdering med karakterer. Både de ujusterte og justerte gjennomsnittene indikerer at elevene fra skole med lang erfaring med karakterfri underveisvurdering har lavest skolestress ut av de tre gruppene. Dermed beholdes H6 knyttet til skolestress ved at elever med lang erfaring med karakterfri

undervisvurdering har mindre stress enn de andre gruppene. Men, deler av H6 forkastet også, ettersom gjennomsnittet for stress mellom elever fra skoler med kort erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer ikke er signifikant forskjellig.

Hvordan lar så dette seg forklare? Rapporter viser til at flere unge plages av psykiske helseplager og at disse plagene ofte knyttes til stress (Bakken, 2018; Eriksen et al., 2017). Dette er bekymringsverdig, særlig når fjorårets Ungdata-rapport viser til at flere unge rapporter om å være ofte eller svært ofte stresset av skolearbeidet (Bakken, 2018). Skolestress knyttes blant annet til prestasjonspress (Eriksen et al., 2017), og ifølge Skaalvik og Federici (2015) sin studie rapporterer flere unge om et sterkt prestasjonspress i skolen. Derfor mener Skaalvik og Federici (2015) at det er viktig å finne ut av hva som er årsaken til prestasjonspresset elevene opplever. Dette er aktuelt, særlig når elevenes budskap er et redusert karakterpress i skolen. Ved å finne de bakenforliggende årsakene til presset, vil det også kunne bidra til å redusere stresset elevene opplever i skolen. Flere skoler uttrykker at en av hensiktene med en karakterfri undervisvurdering er at det skal bidra til å senke elevenes stress og prestasjonspress i skolen (Jansen, 2018; Mordt et al., 2016; Simensen, 2018). Det er og naturlig å tenke at karakterpresset vil reduseres dersom karakterene fjernes, ettersom det sende signaler om at karakterene ikke er det viktige i skolen. De første analysene tyder derimot på at det ikke er en forskjell, men ved å ta høyde for at skolene har kort og lang erfaring, kan funne tyde på at lang erfaring med karakterfri undervisvurdering kan ha en positiv betydning for elevenes skolestress.

Et uventet funn er at elevene fra skoler med kortere erfaring med karakterfri undervisvurdering viser å ha det høyeste gjennomsnittet for skolestress. Denne studien anser stress som en følge av blant annet press. Dette er i tråd med Eriksen et al. (2017) som finner at elever knytter skolestress til blant annet prestasjonspresset og presset omkring det å få en god nok utdanning. En rapport av Danmarks Evalueringsinstitut (2017) viser at omtrent 70 prosent av elevene uttrykker at de opplever et mindre press uten karakterer, men samtidig viser rapporten til at omtrent 80 prosent av elever uttrykker en usikkerhet omkring hvilket nivå de ligger på, og at for noen elever bidrar dette til press. Leder av Utdanningsforbundet presentere i Utdanning 11/2018 at det tar tid å endre en vurderingspraksis (Ruud, 2018). At elever fra skoler med kort erfaring med karakterfri undervisvurdering har det høyeste gjennomsnittet for skolestress kan være et resultat av de utfordringene som lærere står ovenfor

med å utvikle en ny vurderingspraksis. Uten karakterer, og uten en god vurderingspraksis som erstatning for dette, kan elevene oppleve å være usikre på hvordan de ligger an. På bakgrunn av Danmarks Evalueringsinstitut (2017) sine funn kan dette bidra til at elevene opplever en ny form for press: «Når jeg ikke vet sikkert hvordan jeg gjør det i skolefagene, må jeg yte desto mer for å være på den sikre siden». Press anses for å være nært knyttet stress, og dermed kan det tenkes at stressnivået vil øke. Dersom elevene ikke forstår hvordan de ligger an, kan dette også forstås som et tap av kontroll. Bandura (1997) antyder at eleven har et behov for å kontrollere eget liv. Dersom elevene da ikke føler at de har kontroll på hvordan de ligger an faglig, eller ikke forstår hvilken karakter tilbakemeldingen tilsvarer, kan det tenkes at det kan utvikles et stress. Etter en lengre erfaring med vurderingspraksisen derimot, kan det hende at skolen og lærerne har funnet en bedre løsning for hvordan de skal sikre at elevene forstår de faglige vurderingene og opplever disse som tydelige. Dette kan videre bidra til at elevene opplever et redusert stressnivå, som en følge av en mindre opplevelse av press i skolen.

Et høyt stressnivå fremstilles som negativt stress, men all stress er ikke negativt. Stress er behøvelig for å prestere, og kan derfor være positivt. Et positivt stress kan da bidra til høye prestasjoner, engasjement og trivsel (Lillejord et al., 2017). Utsagnene som måler skolestress i denne studien indikerer derimot å måle et negativt stress. Gjennomsnittene tyder på at alle elever opplever en middels grad av stress. Selv om elevene ikke uttrykker en sterk grad av skolestress, er det allikevel aktuelt å undersøke elevenes stress i skolen og hva som er de bakenforliggende årsakene til dette.

Funnen i den studien kan gi en indikasjon på at en lang erfaring med karakterfri undervisningsvurdering kan ha en positiv betydning for elevene i stressnivå, noe som dermed kan være en følge av at lærerne har utviklet en bedre vurderingspraksis, og at presset er senket. Det er svært positivt om et tiltak med å fjerne karakterer fra undervisningsvurderingen bidrar til å senke elevenes press og stress i skolen. Presset og stresset vil derimot ikke fjernes helt av den grunn. Med utgangspunkt i Bandura (2012) Bronfenbrenner (1979) vil eleven påvirkes av miljøet. Skaalvik og Skaalvik (2018) støtter også dette, ved at de nevner at elevene vil motta signaler både utenfor og i skolen om hva som er viktig i skolen. Det er disse signalene som bidrar til elevenes opplevelse av skolens målstruktur (Skaalvik & Skaalvik, 2018). Skaalvik og Skaalvik (2017) sine analyser viser at prestasjonsorientert målstruktur fremmer prestasjonspress, og at læringsorientert målstruktur har en mindre innvirkning på prestasjonsstresset. Det er derfor naturlig å tenke at disse signalene og kan ha en betydning for

elevenes prestasjonspress i skolen, og dermed deres stressnivå. Det blir derfor viktig å finne ut av hvor signalene som bidrar til press kommer fra, og hvilke signaler som bidrar til å dempe dette presset. Funn fra denne studien kan være en indikasjon på at en lang erfaring med karakterfri undervisvurdering er et eksempel på det sistnevnte signalet.

6.0 Avslutning

Denne masteroppgaven er bygd opp rundt problemstillingen: «Hvilke forskjeller er det mellom elevers opplevelse av skolens målstruktur og av seg selv i ungdomsskoler med karakterfri undervisvurdering og ungdomsskoler med undervisvurdering med karakterer?».

Denne problemstillingen er ytterligere konkretisert gjennom følgende forskningsspørsmål:

1. Hvilke forskjeller er det mellom elevers opplevelse av skolens målstruktur i ungdomsskoler med karakterfri undervisvurdering og ungdomsskoler med undervisvurdering med karakterer?
2. Hvilke forskjeller er det mellom elevers opplevelse av indre motivasjon, innsats, selvvrdering av kompetanse og skolestress i ungdomsskoler med karakterfri undervisvurdering og ungdomsskoler med undervisvurdering med karakterer?

I studiens siste kapittel vil jeg oppsummere studiens hovedfunn ved å forsøke å besvare problemstillingen og forskningsspørsmålene som er vist til ovenfor, samt hypotesene som er presentert både i kapittel 2 og 5. Videre inneholder kapitlet refleksjoner omkring studiens begrensninger, før det til slutt presenteres noen ideer knyttet til videre forskning.

6.1 Oppsummering av studiens hovedfunn

Det har med denne studien vært et ønske å finne ut om det er en forskjell mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer, og dermed forhåpentligvis få noen indikasjoner på betydning av karakterfri undervisvurdering. Denne studien finner derimot ingen signifikante forskjeller mellom test- og kontrollgruppa, noe som da kan tyde på at det har ingen betydning å innføre karakterfri undervisvurdering. Disse resultatene er i all hovedsak ikke i tråd med hypotesene (H1, H2, H4, H5) med unntak av hypotesen knyttet til innsats (H3). Det er videre vanskelig å komme med en årsaksforklaring til hvorfor de fleste hypotesene ikke stemmer overens med resultatene, da det med bakgrunn i drøftingene av resultatene viser seg at årsakene er sammensatte.

I tillegg til å studere eventuelle forskjeller mellom elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer, er det antatt at det er av betydning for resultatene om skolene har kort eller lang erfaring med karakterfri undervisvurdering. Når jeg delte opp gruppen «karakterfri undervisvurdering» ut fra

erfaringene skolen hadde med karakterfri undervisvurdering, viser funnene til en liten forskjell i gjennomsnittene for læringsorientert målstruktur og prestasjonsorientert målstruktur mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering. På bakgrunn av dette beholdes deler av H6 knyttet til skolens målstruktur. Dette kan da indikere at skoler må erfare karakterfri undervisvurdering over lengre tid for at dette skal ha en betydning for elevenes opplevelse av skolens målstruktur. Det er derimot vanskelig å trekke en slutning om hvorvidt dette er tilfellet. Videre viser funnene til at det ikke er en signifikant forskjell mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer, samt mellom elever fra skoler med kort erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Dermed forkastes den andre delen av H6 knyttet til skolens målstruktur.

Når det gjelder elevenes opplevelse av innsats, indre motivasjon og selvvurdering av kompetanse tyder resultatene på at det ikke er av betydning om skolen har lang eller kort erfaring med karakterer. Da det ikke vises til noen signifikant forskjell mellom gruppene forkastes H6 knyttet til indre motivasjon og selvvurdering av kompetanse. Mens H6 knyttet til innsats beholdes, da det var ikke var forventet en signifikant forskjell mellom gruppene.

Resultatene viser derimot til en signifikant forskjell i gjennomsnittene for skolestress mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med kort erfaring med karakterfri undervisvurdering, samt mellom elever fra skole med lang erfaring med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Dermed beholdes deler av H6 knyttet til skolestress, men ettersom det ikke er en signifikant forskjell mellom gruppene «karakterfri undervisvurdering (kort)» og «undervisvurdering med karakterer» forkastes andre del av H6 knyttet til skolestress. Disse resultatene kan tyde på at lang erfaring med karakterfri undervisvurdering har en liten positiv betydning for elevenes skolestress, altså at den reduseres. Dette er svært positivt, da ambisjonene bør være å redusere unges negative skolestress.

6.2 Studiens begrensninger

Valgene som er gjort underveis i denne studien har hatt betydning for studiens validitet. Det er forsøkt å tilstrebe en høy validitet, men likevel inneholder studien noen begrensninger. I

begynnelsen av masterprosjektet ble det valgt at studien skulle baseres på et kvasiekperimentelt design. En svakhet ved dette designet er at den ikke bygger på randomisering. Dette er en teknikk som er brukt innen ekte eksperiment og som bidrar til å beskytte mot systematiske målefeil. Et ekte eksperiment ville derfor ha styrket studiens indre validitet. Dersom det ekte eksperimentet samtidig hadde en pretest, ville dette også bidratt til å skape en større trygghet om at gruppene i utgangspunktet er like, samt øke nøyaktigheten i beregningen av effekten av den karakterfrie underveisvurderingen (Ringdal, 2018). På bakgrunn av studiens omfang og tidsmessige årsaker ble dette designet ikke valgt for denne studien. For å styrke studiens validitet er det, som nevnt i kapittel 3, foretatt et strategisk utvalg hvor jeg har forsøkt å få til en «match» mellom elever fra testskoler med kontrollskoler. Samtidig er det gjennomført multivariate analyser hvor det kontrolleres for variabler som kan ha en effekt på elevenes opplevelse av seg selv i skolen og av skolen. Det kan allikevel være andre forhold som gjør at gruppene i utgangspunktet ikke er helt like, som da påvirke resultatene i denne studien.

Tilfeldig fordeling (randomisering) bidrar til å styrke indre validitet, mens tilfeldig utvelgning bidrar til å styrke studiens ytre validitet (Kleven, 2014b). At utvalget i denne studien er strategisk valgt er med å svekke studiens ytre validitet og begrenser derfor muligheten for å generalisere. Det er allikevel et håp på å få noen indikasjoner om betydningen av karakterfri underveisvurdering ved å samle inn et stort utvalg og å sammenligne gruppene i denne studien.

Når det gjelder validitet, er det også viktig å sikre høy begrepsvaliditet. Spørreskjemaet i denne studien bygger på utsagn som er ment å måle flere sammensatte mål. Dersom utsagnene ikke måler det teoretiske begrepet man ønsker å måle, svekkes studiens begrepsvaliditet (Ringdal, 2018). Men, på bakgrunn av at de utvalgte sammensatte målene er utviklet og utprøvd i det longitudinelle prosjektet «Livet i skolen» og at målene i dette prosjektet er kvalitetssikret empirisk og er funnet å fungere tilfredsstillende er det forventet en høy begrepsvaliditet.

I forbindelse med spørreskjemaene, ble disse besvart av elever, hvor fire av fem skoler valgte å utføre spørreskjemaene selv. Dette kan ha ført til misforståelser hos elevene som gjennomførte spørreundersøkelsene. Derav kan det ha oppstått en målefeil. For å forhindre misforståelser fikk alle skoler som ønsket å gjennomføre undersøkelsen selv et

informasjonsskriv, samtidig skal det poengteres at språket i spørreskjemaene vurderes som forståelig for ungdomsskoleelever.

6.3 Videre forskning

I denne studien har jeg sammenlignet elever fra skoler med karakterfri undervisvurdering og elever fra skoler med undervisvurdering med karakterer. Som en fortsettelse av denne studien har det vært interessant å intervju noen av elevene fra denne studien, som gir elevene mulighet til å selv reflektere over undervisvurderingsmetodene med og uten karakterer. I tillegg har det og vært aktuelt å studere lærere sine opplevelser og meninger knyttet de to undervisvurderingsmetodene.

Samtidig har det vært interessant å studert om eldre og yngre elever, samt høytpresterende og lavtpresterende elever påvirkes ulikt av karakterfri undervisvurdering. Dette er på bakgrunn av at forskning viser at yngre og lavtpresterende elever tyder på å være mer negativt påvirket av karakterer sammenlignet med eldre og høytpresterende elever (Lundahl et al., 2015). Som et større prosjekt hadde det også vært interessant å gjennomført et ekte eksperiment med både pre-test og post-test, som da ville gitt et mer nøyaktig svar på hvilken effekt en karakterfri undervisvurdering har på ungdomsskoleelevene.

Litteraturliste

- Anderman, L. H. & Anderman, E. M. (1999). Social Predictors of Changes in Students' Achievement Goal Orientations. *Contemporary Educational Psychology*, 24(1), 21-37. <https://doi.org/10.1006/ceps.1998.0978>
- Bakken, A. (2010). *Prestasjonsforskjeller i Kunnskapsløftets første år - kjønn, minoritetsstatus og foreldres utdanning* (NOVA rapport 9/2010). Hentet fra http://www.nova.no/asset/4069/1/4069_1.pdf
- Bakken, A. (2018). *Ungdata 2018. Nasjonale resultater* (NOVA rapport 8/2018). Hentet fra <http://www.hioa.no/Om-OsloMet/Senter-for-velferds-og-arbeidslivsforskning/NOVA/Publikasjoner/Rapporter/2018/Ungdata-2018.-Nasjonale-resultater>
- Bandura, A. (1997). *Self-efficacy. The exercise of control*. New York: Freeman.
- Bandura, A. (2012). On the Functional Properties of Percieved Self-Efficacy Revisited. *Journal of Management*, 38(1). <https://doi.org/10.1177/0149206311410606>
- Black, P. & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.
- Bronfenbrenner, U. (1979). *The ecology of human development : experiments by nature and design*. Cambridge, Mass: Harvard University Press.
- Butler, R. (1987). Task-Involving and Ego-Involving Properties of Evaluation: Effects of Different Feedback Conditions on Motivational Perceptions, Interest, and Performance. *Journal of Educational Psychology*, 79(4), 474-482. <https://doi.org/10.1037/0022-0663.79.4.474>
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *Journal of Educational Psychology*, 58(1), 1-14. <https://doi.org/10.1111/j.2044-8279.1988.tb00874.x>
- Butler, R. & Nisan, M. (1986). Effects of No Feedback, Task-Related Comments, and Grades on Intrinsic Motivation and Performance. *Journal of Educational Psychology*, 78(3), 210-216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Cameron, J. (2001). Negative Effects of Reward on Intrinsic Motivation- A Limited Phenomenon: Comment on Deci, Koestner, and Ryan (2001). *Review of Educational Research*, 71(1), 29-42. <https://doi.org/10.3102/00346543071001029>

- Clausen, T. H. & Johansen, V. (2012). Cronbachs alfa. I T. A. Eikemo & T. H. Clausen (Red.), *Kvantitativ analyse med SPSS. En praktisk innføring i kvantitative analyseteknikker*. (2. utg.). Tapir akademisk forlag.
- Cohen, J. (1969). *Statistical Power of Analysis for the Behavioral Sciences*. New York: Academic Press.
- Danmarks Evalueringsinstitut. (2017). *Forsøg med karakterfri 1. g-klasser*. Hentet fra <https://www.eva.dk/ungdomsuddannelse/forsoeg-karakterfri-1-g-klasser>
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668. <https://doi.org/10.1037/0033-2909.125.6.627>
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Press.
- Deci, E. L. & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01
- Dæhlen, M., Smette, I. & Strandbu, Å. (2011). *Ungdomskoleelevers meninger om skolemotivasjon. En fokusgruppestudie* (NOVA rapport 4/2011). Hentet fra http://www.nova.no/asset/4538/1/4538_1.pdf
- Eileraas, K. H. (2014). *Vurdering for læring: Indre motivasjon, motivert atferd og skolens målstruktur* (Masteravhandling, NTNU). Hentet fra <http://hdl.handle.net/11250/270068>
- Elliot, A. J. & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, 70(3), 461-475. <https://doi.org/10.1037/0022-3514.70.3.461>
- Endr. i forskrift til opplæringslova. (2009). Forskrift om endring i forskrift til opplæringslova (FOR-2009-07-01-964). Hentet fra <https://lovdata.no/dokument/LTI/forskrift/2009-07-01-964>
- Eriksen, I. M. & Bakken, A. (2018). Stress, press og psykisk plager blant unge - hva sier NOVA-forskningen? <https://doi.org/10.13140/RG.2.2.31512.65283>
- Eriksen, I. M., Sletten, M. A., Bakken, A. & von Soest, T. (2017). *Stress og press blant ungdom. Erfaringer, årsaker og utbredelse av psykiske helseplager* (NOVA rapport 6/2017). Hentet fra <http://www.hioa.no/Om-OsloMet/Senter-for-velferds-og-arbeidslivsforskning/NOVA/Publikasjoner/Rapporter/2017/Stress-og-press-blant-ungdom>

- Federici, R. A., Skaalvik, E. M. & Tangen, T. N. (2015). Students' Perceptions of the Goal Structure in Mathematics Classrooms: Relations with Goal Orientations, Mathematics Anxiety, and Help-Seeking Behavior. *International Education Studies*, 8(3).
<https://doi.org/10.5539/ies.v8n3p146>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-140. <https://doi.org/10.1177/001872675400700202>
- Field, A. (2018). *Discovering Statistics using IBM SPSS Statistics* (5. utg.). Los Angeles: Sage edge.
- Frostad, P. & Mjaavath, P. E. (2018). Fra ungdomsskolen til videregående skole; Faktorer som predikerer elevens intensjon om å slutte på skolen. *Psykologi i kommunen*, 53(2), 27-40.
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. London: SAGE. Hentet fra
<https://ebookcentral.proquest.com/lib/ntnu/reader.action?docID=689464>
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Helgesen, R. (2018). Lektorkonferansen 2018: Karakterstrid i skolen. Hentet fra
<https://www.norsklektorlag.no/lektorkonferansen/category271.html>
- Helsedirektoratet. (2015). Trivsel i skolen. Hentet fra
<https://utdanningsforskning.no/artikler/teoretiske-perspektiver-pa-trivsel/>
- Jansen, K. (2018, 28.01.2018). Skolene skroter karakterer. *Bergens Tidende*. Hentet fra
<https://www.bt.no/nyheter/lokalt/i/G1QVOQ/Skolene-skroter-karakterer>
- Johansen, V. (2018). Sammenhenger mellom deltakelse i Ungdomsbedrift og indre motivasjon og innsats i videregående skole. *Tidskriftet FOU i praksis*, 12(2), 29-47. Hentet fra <https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2599484/Johansen.pdf?sequence=1&isAllowed=y>
- Kleven, T. A. (2013). *Effektstørrelse*. Notat, Uio. Hentet fra
<https://www.uio.no/studier/emner/uv/iped/PED4010/h13/effektstorrelse%5B1%5D.pdf>
- Kleven, T. A. (2014a). Forskning og forskningsresultater. I T. A. Kleven (Red.), *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolking og vurdering* (2. utg., s. 9-26). Bergen: Fagbokforlaget.

- Kleven, T. A. (2014b). Hvilken kontekst er resultatene gyldige i? Spørsmål om ytre validitet. I T. A. Kleven (Red.), *Innføring i pedagogisk forskningsmetode. En hjelp til kritisk tolkning og vurdering* (2. utg., s. 123-138). Bergen: Fagbokforlaget.
- Klinger, D. A., Freeman, J. G., Bilz, L., Liiv, K., Ramelow, D., Sebok, S. S., ... Rasmussen, M. (2015). Cross-national trends in perceived school pressure by gender and age from 1994 to 2010. *The European Journal of Public Health, 25*, 51-56.
<https://doi.org/10.1093/eurpub/ckv027>
- Kunnskapsdepartementet. (2016). *Fag- Fordypning- Forståelse. En fornyelse av Kunnskapsløftet*. (Meld. St. nr. 28 (2015-2016)). Hentet fra
<https://www.regjeringen.no/contentassets/e8e1f41732ca4a64b003fca213ae663b/no/pdfs/stm201520160028000dddpdfs.pdf>
- Lauvås, P. (2018). *Vurdering i skolen*. Oslo: Cappelen Damm Akademisk.
- Lillejord, S., Børte, K., Ruud, E. & Morgan, K. (2017). *Stress i skolen - En systematisk kunnskapsoversikt*. Oslo: Kunnskapssenter for utdanning. Hentet fra
<https://utdanningsforskning.no/globalassets/stress-i-skolen---en-systematisk-kunnskapsoversikt.pdf>
- Liu, Y. (2015). The longitudinal relationship between Chinese high school students' academic stress and academic motivation. *Learning and Individual Differences, 38*, 123-126.
<https://doi.org/10.1016/j.lindif.2015.02.002>
- Lundahl, C., Hultén, M., Klapp, A. & Mickwitz, L. (2015). *Betygens geografi: Forskning om betyg och summativa bedömningar i Sverige och internationellt*. Stockholm: Vetenskapsrådet. Hentet fra
https://www.vr.se/download/18.2412c5311624176023d25a3a/1529480538522/Betygens-geografi_VR_2015.pdf
- Marsh, H. W. & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*(1), 213-231. <https://doi.org/10.1037/0022-3514.47.1.213>
- Mellingsæter, H. (2018, 12.02.2018). Flere ungdomsskoler dropper karakterer. *Aftenposten*. Hentet fra <https://www.aftenposten.no/norge/i/9jg0W/Flere-ungdomsskoler-dropper-karakterer>
- Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., ... Urdan, T. (2000). *Pattern of Adaptive Learning Scales (PALS)* University of Michigan.

- Mordt, H., Archer, E. K. & Løken, K.-M. (2016). Osloskoler satser på mindre bruk av karakterer. *Nrk*. Hentet fra <https://www.nrk.no/ostlandssendingen/osloskoler-satser-pa-mindre-bruk-av-karakterer-1.13201867>
- NESH. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Hentet fra https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf
- NSD. (u.å.). Hvilke personopplysninger skal du behandle? Hentet fra <https://meldeskjema.nsd.no/test/>
- OECD. (2009). *OECD Programme for International Student Assessment. Student Questionnaire*. Paris: OECD.
- Pallant, J. (2016). *SPSS Survival Manual. A step by step guide to data analysis using IBM SPSS*. (6. utg.). New York: McGraw-Hill Education.
- Postholm, M. B. & Jacobsen, D. I. (2018). *Forskningsmetode for masterstudenter i lærerutdanning*. Oslo: Cappelen Damm Akademisk.
- Ringdal, K. (2013). *Enhet og Mangfold. Samfunnsvitenskapelig forskning og kvantitativ metode* (3. utg.). Bergen: Fagbokforlaget.
- Ringdal, K. (2018). *Enhet og Mangfold. Samfunnsvitenskapelig forskning og kvantitativ metode* (4. utg.). Bergen: Fagbokforlaget.
- Ruud, M. (2018). På Ringshaug ungdomsskole har elevene karakterfri hverdag. *Utdanning*, (11), 10-17.
- Ryan, R. M. & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. I Kathryn R. Wentzel & Allan Wigfield (Red.), *Handbook of motivation at school* (s. 171-195). New York: Routledge.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research*, 46(3), 407-441. <https://doi.org/10.3102/00346543046003407>
- Sideridis, G. D. & Kaplan, A. (2011). Achievement Goals and Persistence Across Tasks: The Roles of Failure and Success. *The Journal of Experimental Education*, 79(4), 429-451. <https://doi.org/10.1080/00220973.2010.539634>
- Simensen, H. M. (2018). Lærerne fornøyde med karakterfri hverdag på Gressvik ungdomsskole- delte meninger blant elevene. *Fredrikstad Blad*. Hentet fra <https://www.f-b.no/nyheter/skole/gressvik-ungdomsskole/larerne-fornoyde-med-karakterfri-hverdag-pa-gressvik-ungdomsskole-delte-meninger-blant-elevne/s/5-59-1022408>

- Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology, 89*(1), 71-81. <https://doi.org/10.1037/0022-0663.89.1.71>
- Skaalvik, E. M. & Federici, R. A. (2015). Prestasjonspresset i skolen. *Bedre skole, (3)*, 11-15. Hentet fra <https://www.utdanningsnytt.no/files/2019/06/27/Bedre%20Skole%203%202015.pdf>
- Skaalvik, E. M. & Skaalvik, S. (2004). Self-Concept and Self-Efficacy: A Test of the Internal/External Frame of Reference Model and Predictions of Subsequent Motivation and Achievement. *Psychological Reports, 95*, 1187-1202. <https://doi.org/10.2466/pr0.95.3f.1187-1202>
- Skaalvik, E. M. & Skaalvik, S. (2005). *Skolen som læringsarena. Selvoppfatning, motivasjon og læring*. Oslo: Universitetsforlaget.
- Skaalvik, E. M. & Skaalvik, S. (2013a). School goal structure: Associations with students' perceptions of their teachers as emotionally supportive, academic self-concept, intrinsic motivation, effort, and help seeking behavior. *International Journal of Educational Research, 61*, 5-14. <https://doi.org/10.1016/j.ijer.2013.03.007>
- Skaalvik, E. M. & Skaalvik, S. (2013b). Teachers' perceptions of the school goal structure: Relations with teachers' goal orientations, work engagement, and job satisfaction. *International Journal of Educational Research, 62*, 199-209. <https://doi.org/10.1016/j.ijer.2013.09.004>
- Skaalvik, E. M. & Skaalvik, S. (2015). *Motivasjon for læring. Teori + praksis*. Oslo: Universitetsforlaget.
- Skaalvik, E. M. & Skaalvik, S. (2017). Prestasjonspresset i skolen: Om skolens målstruktur og elevenes mentale helse. I M. Uthus (Red.), *Elevenes psykiske helse i skolen: Utdanning til å mestre egne liv* (s. 47- 69). Oslo: Gyldendal Akademisk.
- Skaalvik, E. M. & Skaalvik, S. (2018). *Skolen som læringsarena. Selvoppfatning, motivasjon og læring* (3. utg.). Oslo: Universitetsforlaget.
- Skaalvik, E. M., Valåns, H. & Sletta, O. (1994). Task Involvement and Ego Involvement: relations with academic achievement, academic self-concept and self-esteem. *Scandinavian Journal of Educational Reserach, 38*(3-4), 231-243. <https://doi.org/10.1080/0031383940380306>
- Svartdal, F. (2018). Effektstørrelse. I *Store norske leksikon*. Hentet fra <https://snl.no/effektstørrelse>

- Tafarodi, R. W. & Swann Jr., W. B. (1995). Self-liking and self-competence as dimensions of global self-esteem: initial validation of a measure. *Journal of Personality Assessment*, 65(2), 322-342. https://doi.org/10.1207/s15327752jpa6502_8
- Utdanningsdirektoratet. (2015). Fire prinsipper for god undervisvurdering. Hentet fra <https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/undervisvurdering/>
- Utdanningsdirektoratet. (2016). *Individuell vurdering* (Udir-5-2016). Hentet fra <https://www.udir.no/regelverkstolkninger/opplaring/Vurdering/individuell-vurdering-udir-5-2016/i/#3-4>
- Utdanningsforbundet. (2019). *Lærerrømmet: Karakterer- skal vi kutte dem ut?* [Lyd podcast]. Hentet fra https://www.utdanningsforbundet.no/nyheter/2019/larerrrommet-karakterer-skal-vi-kutte-dem-ut/?utm_medium=social&utm_source=Facebook&utm_campaign=Larerrrommet+episode+21
- Uthus, M. (2017). Elevenes psykiske helse i skolen. Å mestre skolen er å mestre livet. Å mestre livet er å mestre skolen. I Marit Uthus (Red.), *Elevenes psykiske helse i skolen*. (s. 17-43). Oslo: Gyldendal Akademisk.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C. & Vallières, E. F. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic and Amotivation in Education. *Educational and Psychological Measurement*, 52(4), 1003-1017. <https://doi.org/10.1177/0013164492052004025>
- Wendelborg, C., Røe, M. & Skaalvik, E. (2011). *Elevundersøkelsen 2011. Analyse av Elevundersøkelsen 2011*. NTNU Samfunnsforskning AS. Hentet fra https://www.udir.no/globalassets/upload/forskning/2011/elevundersokelsen_2011_analyse.pdf
- Wiliam, D. (2006). Keeping learning on track: Classroom assessment and the Regulation of learning. I F. K. Lester (Red.), *Second Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics* (2. utg., s. 1053-1098). Charlotte: Information Age Publishing. Hentet fra <https://ebookcentral.proquest.com/lib/ntnu/detail.action?docID=3315201>

Vedlegg 2: Informasjonsskriv til skolene

Informasjonsskriv om masterprosjektet

Formålet med masterprosjektet

Mitt navn er Trine Smisethjell og jeg er mastergradsstudent i spesialpedagogikk ved NTNU i Trondheim. Denne våren skal jeg skrive masteroppgave, hvor formålet er å finne ut av hvilken innvirkning karakterer har på elevens skoleinnsats, indre motivasjon, skolestress, selvvurdering og læringsmiljø.

Flere ungdomsskoler i Norge velger å redusere bruken av karakterer. De mest nevnte målene er å redusere stress blant elevene, styrke fokuset på læring og fremme et positivt læringsmiljø. Vi har begrenset kunnskap om skoler med denne vurderingspraksisen i Norge, og det er derfor et behov for forskning på dette området. I prosjektet vil jeg invitere ungdomsskoler hvor elevene får karakterer gjennom hele semesteret og ungdomsskoler som kun gir karakterer ved semesterslutt til å delta. I utvelgelse av skoler vil det også legges vekt på at skolene har noenlunde lik skolestørrelse.

Deltagelse i masterprosjektet

Din deltagelse vil være et bidrag i arbeidet med å få mer kunnskap om hvilken innvirkning karakterer har for ungdomsskoleelevene. Deltagelse i prosjektet innebærer at elevene besvarer et spørreskjema i skoletiden. Det vil ta i underkant av 10 minutter for elevene å fylle ut spørreskjemaet. Jeg har mulighet til å komme til deres skole, for å dele ut og samle inn spørreskjemaene selv. Dette vil da foregå i løpet av februar og mars 2019. Det er frivillig i å delta i masterprosjektet. Dersom du ønsker å delta i prosjektet, behøver jeg samtykke fra skolen og elevenes foresatte.

Spørreskjemaet

Spørreskjemaet inneholder noen spørsmål om bakgrunnsopplysninger (kjønn, klassetrinn, foreldres utdanning og karaktersnitt fra 1. semester), og spørsmål om skolestress, elevens skoleinnsats, indre motivasjon for skolen, selvvurdering av kompetanse og læringsmiljøet. Dersom skole og/eller foresatte ønsker å se spørreskjemaet på forhånd av undersøkelsen, kan du ta kontakt med meg.

Ditt personvern

Ingen personopplysninger skal behandles i dette prosjektet, som betyr at det ikke vil være mulig å bli identifisert i datamaterialet og masteroppgaven. Skolens navn vil erstattes med en kode, og skolen vil derfor ikke kunne bli identifisert. Opplysningene om deltagerne vil kun brukes til formålene jeg har fortalt om i dette informasjonsskrivet. Opplysningene behandles konfidensielt, og det er kun jeg og mine veiledere som vil ha tilgang til datamaterialet.

Har du spørsmål om masterprosjektet kan disse rettes til:

Masterstudent Trine Smisethjell, epost trinesmi@stud.ntnu.no eller telefon 93444288.

Veileder Marit Uthus, epost marit.uthus@ntnu.no eller telefon 92650438.

Biveileder Vegard Johansen, epost vegard.johansen@ntnu.no eller telefon 90743388.

Med vennlig hilsen

Trine Smisethjell

Vedlegg 3: Informasjonsskriv til foreldre

Informasjonsskriv til foreldre og foresatte

Formålet og deltagelse

Mitt navn er Trine Smisethjell og jeg er mastergradsstudent i spesialpedagogikk ved NTNU i Trondheim. Denne våren skal jeg skrive masteroppgave, hvor formålet er å finne ut av hvilken innvirkning karakterer har på elevens skoleinnsats, indre motivasjon, skolestress, selvvurdering og læringsmiljø. I forbindelse med dette masterprosjektet vil jeg invitere ungdomsskoler hvor elevene får karakterer gjennom hele semesteret og ungdomsskoler som kun gir karakterer ved semesterslutt til å delta. Deltagelse i prosjektet innebærer at elevene besvarer et spørreskjema i skoletiden. Spørreundersøkelsen er anonym og det tar i underkant av 10 minutter å besvare spørreskjemaet.

Spørreskjemaet

Spørreskjemaet inneholder noen spørsmål om bakgrunnsopplysninger (kjønn, klassetrinn, foreldres utdanning og karaktersnitt fra 1.termin 2018/19), og spørsmål om skolestress, elevens skoleinnsats, indre motivasjon for skolen, selvvurdering av kompetanse og læringsmiljøet. Dersom du/dere ønsker å se spørreskjemaet på forhånd av undersøkelsen, kan dere ta kontakt med skolens rektor.

Ditt personvern

Ingen personopplysninger skal behandles i prosjektet, som betyr at det ikke vil være mulig å bli identifisert i datamaterialet og masteroppgaven. Skolens navn vil erstattes med en kode, og skolen vil derfor ikke kunne bli identifisert. Opplysningene om deltagerne vil kun brukes til formålene jeg har fortalt om i dette informasjonsskrivet. Opplysningene behandles konfidensielt, og det er kun jeg og mine veiledere som vil ha tilgang til datamaterialet.

Dersom du/dere ikke tillater at barnet ditt/deres besvarer spørreskjemaet, må det gis beskjed til kontaktlærer innen 2019.

Med vennlig hilsen

Trine Smisethjell

Vedlegg 4: Informasjonsskriv til elevene

Informasjon til elevene

1. Målet med denne undersøkelsen er å finne ut av hvilken påvirkning karakterer har på elevers skoleinnsats, indre motivasjon, skolestress, selvvurdering og læringsmiljø.
2. Både ungdomsskoleelever som får karakter gjennom hele semesteret og ungdomsskoleelever som kun får karakterer ved semesterslutt er invitert til å delta i denne spørreundersøkelsen.
3. Spørreundersøkelsen er frivillig, men det er veldig fint om dere ønsker å delta. For å finne ut av hvordan skolen kan bli bedre, så behøver vi informasjon fra dere.
4. Spørreundersøkelsen er anonym, så ingen får vite hva du har svart. Du skal derfor ikke skrive navn på skjemaet.
5. Spørreskjemaet er på én side. Skjemaet vil bli lest av en maskin, det er derfor viktig at dere følger disse reglene:
 - a. Bruk svart/blå kulepenn.
 - b. Sett kun ett kryss på hvert spørsmål/påstand.
 - c. Ikke tegn/skriv på arket. Kun sett kryss.
 - d. Om du svarer feil, fyller du feltet. Deretter sett kryss i rett felt.
6. Videre er det viktig at du leser spørsmålene nøye og krysser av for det svaret som passer deg best.
7. Tusen takk!

Vedlegg 5: Prinsipale komponentanalyser (PKA) og Cronbachs alfa

Tabell 6: Mønstermatrise fra PKA med oblik rotasjon (direct oblimin) og med Kaizers kriterium (egenverdi >1). PKA av indre motivasjon, stress, selvvurdering av kompetanse og innsats.

Indikatorer	Komponentladninger			
	1	2	3	4
Indre motivasjon				
Jeg liker å holde på med skolearbeid	0,87			
Jeg synes det er morsomt å arbeide med fagene	0,88			
Fagene vi har på skolen interesserer meg	0,85			
Jeg har lyst til å arbeide med fagene	0,84			
% av varians	40,2 %			
Egenverdi	6,4			
Cronbachs alfa	0,91			
Skolestress				
Jeg blir stresset av skolearbeidet		0,80		
Jeg føler meg utslitt på grunn av skolearbeidet		0,82		
Jeg føler et sterkt press for å gjøre det godt på skolen		0,75		
Jeg har mer skolearbeid enn jeg klarer å gjøre		0,68		
% av varians		14%		
Egenverdi		2,2		
Cronbachs alfa		0,78		
Selvvurdering av kompetanse				
Jeg har gode evner, så jeg kan få til det meste			-0,91	
Jeg takler utfordringen med å lære nytt stoff på skolen			-0,66	
Jeg gjør det bra på en rekke områder			-0,79	
Når jeg setter meg ned for å lære noe virkelig vanskelig, kan jeg klare det			-0,84	
% av varians			8,2%	
Egenverdi			1,3	
Cronbachs alfa			0,84	

Innsats	
Jeg jobber godt med oppgavene vi får på skolen	-0,62
Jeg forbereder meg så godt jeg kan til prøver	-0,77
Jeg slurver IKKE mye med skolearbeidet mitt	-0,84
Jeg følger godt med i timene på skolen	-0,61
<i>% av varians</i>	<i>6,8%</i>
<i>Egenverdi</i>	<i>1,1</i>
<i>Cronbachs alfa</i>	<i>0,81</i>

Kaiser-Meyer Olkin (KMO)= 0,91

Tabell 7: Mønstermatrise fra PKA med oblik rotasjon (direct oblimin) og med Kaizers kriterium (egenverdi > 1). PKA av prestasjonsorientert og læringsorientert målstruktur.

Indikatorer	Komponentladninger	
	1	2
Prestasjonsorientert målstruktur		
Det viktigste for lærerne våre er at jeg får gode karakterer	0,81	
Det viktigste for lærerne våre er at jeg svarer riktig på spørsmålene	0,78	
Det viktigste for lærerne våre er at jeg gjør det godt på prøver	0,81	
Det viktigste for lærerne våre er at jeg er flink på skolen	0,79	
<i>% av varians</i>	33,5%	
<i>Egenverdi</i>	2,7	
<i>Cronbachs alfa</i>	0,81	
Læringsorientert målstruktur		
Det som betyr noe for lærerne våre er at jeg gjør så godt jeg kan i fagene		0,77
Det er viktig for lærerne våre at jeg skjønner det jeg gjør, ikke bare puffer det		0,79
Det er viktig for lærerne våre at jeg lærer nye ting		0,79
Det er helt OK for lærerne våre at jeg gjør feil i klassen, så lenge jeg lærer av det		0,74
<i>% av varians</i>		29,2%
<i>Egenverdi</i>		2,3
<i>Cronbachs alfa</i>		0,78

Kaiser-Meyer Olkin (KMO)= 0,79

Vedlegg 6: Skjevhet

Tabell 8: Skjevhetsverdier for studiens avhengige variabler

Avhengige variabler	Skjevhet
Prestasjonsorientert målstruktur	-0,08
Læringsorientert målstruktur	-1,11
Indre motivasjon	-0,02
Innsats	-0,54
Selvvurdering av kompetanse	-0,53
Skolestress	-0,06

Vedlegg 7: ANCOVA og Cohens' d

Tabell 9: Sammenligning av gruppene «karakterfri undervisvurdering» (1) og «undervisvurdering med karakterer» (2) og studiens avhengige variabler, ved bruk av ANCOVA og Cohens' d (mellom nr. 1 og nr. 2). Kovariatene er kjønn, klassetrinn, foreldres utdanning og tidligere skoleprestasjoner

Variabler		Gj.snitt	Cohen's d
Prestasjonsorientert målstruktur	1	3,43	0,02
	2	3,41	
Læringsorientert målstruktur	1	4,80	-0,01
	2	4,81	
Indre motivasjon	1	3,32	0,11
	2	3,19	
Innsats	1	4,34	-0,03
	2	4,37	
Selvvurdering av kompetanse	1	4,40	0,04
	2	4,36	
Skolestress	1	3,71	0
	2	3,71	

