



# Hverandrevurdering med rubrikker i høyere utdanning

En casestudie av studenters og underviseres opplevelse av å delta i hverandrevurdering

Karina Rusås Olsen og Katrin Fallmyr

Masteroppgave i pedagogikk  
Kommunikasjon, design og læring

Institutt for pedagogikk  
Det utdanningsvitenskapelige fakultet

Vår 2021

© Karina Rusås Olsen & Katrin Fallmyr

2021

Tittel: Hverandrevurdering med rubrikker i høyere utdanning. En casestudie av studenters og underviseres opplevelse av å delta i hverandrevurdering.

Forfattere: Karina Rusås Olsen & Katrin Fallmyr

## SAMMENDRAG AV MASTEROPPGAVEN I PEDAGOGIKK

Tittel	Hverandrevurdering med rubrikker i høyere utdanning
Av	Karina Rusås Olsen og Katrin Fallmyr
Emnekode	PED4591
Semester	Vår 2021
<u>Stikkord</u>  Formativ vurdering  Hverandrevurdering  Læringsdesign  Læringsteknologi  Mixed methods  Sosiokulturell læringsteori  Tematisk analyse  Universitetspedagogikk  Virksomhetsteori  Vurderingsrubrikk	

# Sammendrag

Denne oppgaven er en casestudie som undersøker formativ vurdering og hverandrevurdering i høyere utdanning i et sosiokulturelt perspektiv. Formålet er å undersøke hvordan studenter og undervisere opplever å gjennomføre hverandrevurdering med bruk av rubrikker. Studien undersøker også hvilke faktorer som påvirker en hverandrevurderingsprosess.

Studien tar utgangspunkt i to ulike studiekurs som høsten 2020 arbeidet med hverandrevurdering. De to studentgruppene bestod av førsteårs- og mastersterstudenter. Datamaterialet i denne oppgaven er samlet inn ved hjelp av både kvantitative og kvalitative metoder. Det er gjennomført kvalitative intervju med undervisere og studenter, samt en spørreundersøkelse blant studenter i de to aktuelle studentgruppene. Det er brukt bivariat korrelasjonsanalyse og multipel regresjonsanalyse av kvantitative data. Videre er det anvendt tematisk analyse som metode for å analysere kvalitative data.

Resultatene viser at førsteårsstudentene synes at å både gi og å motta hverandrevurdering er omtrent like lærerikt, mens masterstudentene oppgir at det å gi hverandrevurdering er mer lærerikt enn det å motta. Førsteårsstudentene er derimot mer skeptiske til hverandrevurdering enn masterstudentene. Masterstudentene er positive til hverandrevurdering, men opplevde utfordringer knyttet til *innrammingen* høsten 2020. At masterstudentene er mer positive til hverandrevurdering forstås i sammenheng med studentenes ferdighetsnivå og deres kjennskap til normer og konvensjoner i akademia. Resultater fra studien underbygger potensialet for læring som finnes i en hverandrevurderingsprosess.

Denne studien viser at enkeltfaktorer i et læringsdesign har stor påvirkning på en hverandrevurderingsprosess. For det første er en viktig faktor utformingen av rubrikk. Studien viser viktigheten av at den må tilpasses studentenes ferdigheter, samt åpne for konstruktiv kritikk. I denne studien var rubrikken mest nyttig for førsteårsstudentene. For det andre er det viktig at aktiviteten tilpasses den enkelte studiekulturen. Med studiekultur menes faglig fellesskap hos studentgruppen samt studentenes ferdighetsnivå og erfaring fra høyere utdanning. For det tredje viser analysen hvordan de ulike aktørene vektlegger et korttids- eller langtidsperspektiv tilknyttet arbeidet med hverandrevurdering. Underviser vektlegger læring og nytthet i et langtidsperspektiv, mens studentene vektlegger læring og nytthet til nærstående arbeidskrav og eksamener. Basert på tidligere forskning og denne studiens funn

presenterer studien avslutningsvis et forslag til hvordan rubrikker for hverandrevurdering kan integreres som en del av et større læringsdesign på en hensiktsmessig måte.

# Forord

For en god følelse det er å levere denne oppgaven! Det har vært en lang og krevende prosess, men mestringsfølelsen vi sitter med nå gjør det verdt det. Vi er ferdig med to lærerike år som masterstudenter. Etter fullført studieløp har vi fått øynene opp for hvor allsidig pedagogikkfaget er, og hvor viktig det er med pedagoger i andre sektorer enn bare skolen. Arbeidet med masteroppgaven har gitt oss innsikt i hvor mange faktorer som skal klaffe for å få til et bra kursdesign i høyere utdanning. Det har også vært interessant å følge både studenters og underviseres perspektiver i en og samme prosess. Etter å ha skrevet en masteroppgave har vi igjen fått bekreftet verdien av å motta god feedback fra andre, og vi ser et stort potensiale i hverandrevurdering. Disse erfaringene tar vi med oss ut i arbeidslivet.

En stor del av de gode minnene vi tar med oss fra studietiden er takket være alle vi er blitt kjent med. Til tross for mye hjemmekontor og innetid i vinter har vi hatt mye glede av studievenner som vi fikk på Blindern. Tusen takk for hyggelige lunsjer, god humor, og ikke minst et sjeldent godt samarbeid!

Takk til alle informanter som har tatt seg tid til å delta i intervju og besvare spørreskjema. Det betyr veldig mye for oss. Også tusen takk til gode venner som har satt av tid til å lese over oppgaven (i vinterferien, påskeferien og pinsen). Det kaller man gode venner!

Den største takken går til vår veileder Jan Dolonen. Takk for at du har stilt høye krav til oss og pushet oss fra start til slutt. Dette har gjort oppgaven mye bedre. Vi setter stor pris på all hjelpen vi har fått fra deg.

Til slutt vil vi takke hverandre for et solid samarbeid. Det har vært flott å ha en god sparringspartner i denne lange og krevende prosessen. Det blir kjipt å ikke ha hverandre i arbeidslivet.

Oslo, mai 2021

Karina Rusås Olsen & Katrin Fallmyr



# Innholdsfortegnelse

<b>Sammendrag .....</b>	<b>iii</b>
<b>Forord.....</b>	<b>v</b>
<b>1 Introduksjon .....</b>	<b>1</b>
1.1 Bakgrunn for valg av tema .....	1
1.2 Presentasjon av tema .....	2
1.3 Formål og forskningsspørsmål.....	4
1.4 Oppgavens struktur og innhold.....	5
<b>2 Teoretisk rammeverk.....</b>	<b>7</b>
2.1 Sosiokulturelt læringssyn .....	7
2.1.1 Individet som del av sosiale praksiser.....	7
2.2 Artefakter og mediering .....	7
2.3 Virksomhetsteori.....	9
2.3.1 Virksomhetssystemets seks komponenter .....	11
2.3.2 Sentrale konsept i virksomhetsteorien.....	12
2.3.3 Virksomhetsteorien overført til denne studien .....	14
2.4 Den proksimale utviklingssonen og scaffolding .....	16
<b>3 Litteraturgjennomgang .....</b>	<b>20</b>
3.1 Formativ vurdering .....	20
3.1.1 Vurderingskultur i høyere utdanning .....	21
3.2 Hverandrevurdering.....	22
3.2.1 Hverandrevurdering og læringseffekt .....	22
3.2.2 Studenters opplevelse av hverandrevurdering.....	24
3.2.3 Faktorer som påvirker hverandrevurdering .....	25
3.3 Rubrikker.....	30
3.4 Læringsdesign med hverandrevurdering .....	35
3.5 Oppsummering av litteraturgjennomgang .....	37



<b>4 Metode</b> .....	<b>38</b>
4.1 <i>Beskrivelse av casene</i> .....	38
4.1.1 Kursdesign og kontekst .....	39
4.1.2 Innsamling av data .....	41
4.1.3 Utforming av forskningsinstrumenter: intervjuguide og spørreskjema .....	42
4.2 <i>Kvantitativ metode og analytisk strategi</i> .....	43
4.2.1 Operasjonalisering av variabler i spørreskjema .....	44
4.2.2 Statistiske analysemetoder .....	45
4.3 <i>Kvalitativ metode og analytisk strategi</i> .....	46
4.3.1 Det kvalitative forskningsintervjuet .....	47
4.3.2 Analyse og fortolkning: tematisk analyse av intervju .....	47
4.4 <i>Validitet og reliabilitet</i> .....	50
4.4.1 Reliabilitet og validitet ved kvantitativ metode .....	50
4.4.2 Reliabilitet og validitet ved kvalitativ metode .....	52
4.5 <i>Etikk og personvern</i> .....	54
4.5.1 Etske problemstillinger .....	55
<b>5 Presentasjon av funn</b> .....	<b>57</b>
5.1 <i>Kvantitative funn</i> .....	57
5.1.1 Deskriptiv statistikk .....	57
5.1.2 Statistiske analyser .....	67
5.2 <i>Kvalitative funn</i> .....	70
5.2.1 Kvalitative funn knyttet til kursdesign .....	70
5.2.2 Kvalitative funn knyttet til bruk av artefakter .....	76
5.2.3 Kvalitative funn knyttet til å gi hverandrevurdering .....	79
5.2.4 Kvalitative funn knyttet til å motta hverandrevurdering .....	81
5.2.5 Kvalitative funn knyttet til formål med hverandrevurdering .....	83
5.3 <i>Oppsummering av funn</i> .....	85
<b>6 Diskusjon</b> .....	<b>87</b>
6.1 <i>Opplevelse av hverandrevurdering i undervisningen</i> .....	87
6.1.1 Rollene som giver og mottaker av vurdering .....	87

6.1.2 Opplevd nyttinghet .....	91
6.2 Faktorer ved hverandrevurdering som påvirker opplevd læring.....	92
6.3 Hvordan kan rubrikker for hverandrevurdering integreres som del av et større læringsdesign på en hensiktsmessig måte? .....	99
6.3.1 Case 1: Hverandrevurdering for førsteårsstudenter høsten 2020 .....	99
6.3.2 Case 2: Hverandrevurdering for masterstudenter høsten 2020 .....	101
6.3.3 Forslag til design av hverandrevurdering.....	103
<b>7 Avslutning .....</b>	<b>107</b>
7.1 Oppsummering av funn .....	107
7.2 Teoretiske implikasjoner - Hvordan samsvarer studien med funn fra tidligere forskning?.....	109
7.3 Svakheter ved studien og videre forskning .....	110
<b>Litteraturliste.....</b>	<b>112</b>
<b>Vedlegg .....</b>	<b>121</b>

## Figurliste

<b>Figur 1.</b> Vygotskys opprinnelige idé om mediering. Illustrert fra Säljö (2006, s. 25). .....	8
<b>Figur 2.</b> Virksomhetssystemets tredjegerasjon med to samhandlende systemer (Engeström, 2010, s. 136) .....	11
<b>Figur 3.</b> Illustrasjon av Leontievs hierarkiske modell av en aktivitet (Kaptelinin & Nardi, 2006, s. 64) .....	14
<b>Figur 4.</b> Modell av virksomhetssystemet i denne studien før gjennomføring av hverandrevurdering .....	15
<b>Figur 5.</b> Korrelasjonsstyrke (Dancey & Reidy, 2007, s. 176) .....	46
<b>Figur 6.</b> Svarfordeling for målinger av variabelen opplevd læring for førsteårsstudenter .....	58
<b>Figur 7.</b> Svarfordeling for målinger av variabelen opplevd læring for masterstudenter .....	59
<b>Figur 8.</b> Svarfordeling for målinger av variabelen innramming for førsteårsstudenter .....	61
<b>Figur 9.</b> Svarfordeling for målinger av variabelen innramming for masterstudenter .....	62
<b>Figur 10.</b> Svarfordeling for målinger av variabelen artefakter for førsteårsstudenter .....	63
<b>Figur 11.</b> Svarfordeling for målinger av variabelen artefakter for masterstudenter .....	64
<b>Figur 12.</b> Svarfordeling for målinger av variabelen studentenes ferdighetsnivå/ innstilling for førsteårsstudenter .....	65
<b>Figur 13.</b> Svarfordeling for målinger av variabelen studentenes ferdighetsnivå/ innstilling for masterstudenter .....	66
<b>Figur 14.</b> Temaer og undertemaer identifisert gjennom tematisk analyse .....	70
<b>Figur 15.</b> Modell av virksomhetssystemet for førsteårsemnet høsten 2020 .....	99
<b>Figur 16.</b> Modell av virksomhetssystemet for masteremnet høsten 2020 .....	101
<b>Figur 17.</b> Modellen viser et forslag for design av hverandrevurdering .....	103

## Tabelliste

<b>Tabell 1.</b> Beskrivelse av kursdesign .....	40
<b>Tabell 2.</b> Beskrivelse av kursdesign .....	42
<b>Tabell 3.</b> Cronbachs alfa for variablene.....	51
<b>Tabell 4.</b> Gjennomsnitt og standardavvik for variabelen opplevd læring for førsteårsstudenter .....	58
<b>Tabell 5.</b> Gjennomsnitt og standardavvik for variabelen opplevd læring for masterstudenter	60
<b>Tabell 6.</b> Gjennomsnitt og standardavvik for variabelen innramming for førsteårsstudenter.	61
<b>Tabell 7.</b> Gjennomsnitt og standardavvik for variabelen innramming for masterstudenter ....	62
<b>Tabell 8.</b> Gjennomsnitt og standardavvik for variabelen artefakter for.....	63
<b>Tabell 9.</b> Gjennomsnitt og standardavvik for variabelen artefakter for masterstudenter .....	64
<b>Tabell 10.</b> Gjennomsnitt og standardavvik for variabelen studentenes ferdighetsnivå/ innstilling for førsteårsstudenter.....	65
<b>Tabell 11.</b> Gjennomsnitt og standardavvik for variabelen studentenes ferdighetsnivå/ innstilling for masterstudenter.....	66
<b>Tabell 12.</b> Resultat fra korrelasjonsanalyse: korrelasjonsstyrke mellom ulike variabler .....	68
<b>Tabell 13.</b> Resultat fra regresjonsanalyse: sammenhenger mellom opplevd læring og forklaringsvariabler .....	69

# 1 Introduksjon

## 1.1 Bakgrunn for valg av tema

Denne oppgaven handler om formativ vurdering og hverandrevurdering i høyere utdanning. Det gjøres en empirisk undersøkelse av nettbasert hverandrevurdering. Formålet er å forstå hvordan hverandrevurdering kan gjennomføres ved bruk av rubrikker som blir utfylt og sendt til mottaker gjennom nettbaserte verktøy. Innenfor høyere utdanning har summativ vurdering, bestående av en bokstav fra A-F eller tallkarakterer fra 1-6, stått sterkt som den viktigste vurderingsformen. Dette til tross for at feedback og vurdering er viktig for at studenter skal få et godt læringsutbytte (Black & William, 2009; Hattie & Timperley, 2007). Studiebarometeret er en årlig undersøkelse gjennomført av Norsk organ for kvalitet i utdanningen - NOKUT som undersøker studenters oppfatning av kvaliteten ved norske studieprogram. Studiebarometeret fra 2020 viser at spørsmålene som gjelder vurdering og veiledning fra faglig ansatte er noen av spørsmålene som får lavest score av norske studenter. På spørsmål om tilfredshet med mengden feedback studenter får fra faglig ansatte svarer studentene i gjennomsnitt 3,1 på skala som går fra 1-10 hvor 1 er dårligst og 10 er best. Studentene er heller ikke fornøyde med de faglig ansattes evner til å gi konstruktiv feedback på studentarbeid som får gjennomsnittlig verdi 3,4 (NOKUT, 2021). Skal en tro studentene har vurderingspraksisen ved norske høyskoler og universiteter stort forbedringspotensiale.

I stortingsmeldingen *Kultur for kvalitet i høyere utdanning* (Kunnskapsdepartementet, 2016, s. 55-59) blir nettopp forbedring av vurderingspraksis trukket frem som et satsningsområde. Dette er begrunnet både i studentenes opplevelse av manglende feedback, samt forskning som viser at riktig vurderingspraksis er læringsfremmende. Undervisere ved norske institusjoner for høyere utdanning rapporterer om høyt tidspress når det gjelder tid tilgjengelig for undervisning og veiledning (NOKUT, 2019). Dette kan trolig være en del av forklaringen for hvorfor feedback og vurdering ser ut til å være et område hvor det er lite tilfredshet blant studentene, til tross for at det finnes mye forskning om hva som kjennetegner en god vurderingspraksis.

## 1.2 Presentasjon av tema

### **Formativ vurdering og hverandrevurdering.**

I utdanningsforskningen skiller en mellom det en kaller *summativ* og *formativ* vurderingspraksis, også kalt vurdering *av* og vurdering *for* læring. Mange forbinder vurdering i høyere utdanning med det å motta bokstavkarakterer eller tallkarakterer i ulike emner, altså såkalt vurdering *av* læring. Ramaprasad (1983, gjengitt i Black & William, 2009) peker på tre sentrale prosesser i formativ vurdering: å stadfeste hvor studenten er i læringen sin, hvor de er på vei og hva som må til for å komme seg dit.

Som nevnt er det ressurskrevende for utdanningsinstitusjonene å skulle gi underveisvurderinger til hver enkelt student. Dette er trolig noe av årsaken til at hverandrevurdering er blitt et område som er blitt utforsket av flere utdanningsforskere. *Hverandrevurdering* er når to likemenn vurderer hverandres arbeid, for eksempel to studenter ved samme studie. I forskningen blir hverandrevurdering omtalt med flere begrep, blant annet *peer assessment* og *peer feedback*. Peer assessment blir brukt om forskning på *prosessen* hvor to likemenn vurderer hverandre, mens begrepet peer feedback blir brukt om samme prosess, men i forbindelse med forskning hvor en undersøker kvaliteten på feedback likemenn gir hverandre (Double et al., 2020). I denne studien oversettes hverandrevurdering til peer assessment da den undersøker *prosessen* som karakteriserer hverandrevurdering.

### **Feedback**

I denne studien blir begrepet *feedback* brukt om tilbakemeldinger som studentene gir og mottar til eller fra sine medstudenter. Konseptene *vurdering* og *feedback* er nært beslektet i pedagogisk praksis, der feedback er en sentral komponent i formativ vurdering (Damsa et al., 2015). Feedback kan defineres slik; "Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way" (Ramaprasad, 1983, referert i Sadler, 1989, s. 4). I pedagogisk praksis blir feedback sett på som en viktig fasilitator for læring og utvikling (Black & Wiliam, 1998; Hattie & Timperley, 2007; Shute, 2008). Feedback skal gi informasjon om gapet mellom nivået den lærende er på, og ønsket måloppnåelse.

## **Nettbasert hverandrevurdering**

Deltakerne i forskningsprosjektet har gjennomført hverandrevurderingen over nett. Å velge nettbasert hverandrevurdering har vist seg å gjøre det logistiske arbeidet enklere å gjennomføre. Double et al. (2020) skriver i sin metastudie om hverandrevurdering at studentresultatene ikke påvirkes av at prosessen blir gjennomført over nett, men at det ikke finnes tidligere studier som sammenligner hverandrevurdering i fysiske og i nettbaserte rom i en og samme studie. Videre hevder de at det er grunn til å tro at de sosiale prosessene i hverandrevurdering som for eksempel diskusjoner og forklaringer blir påvirket av om prosessen blir gjennomført i fysiske eller digitale rom. Denne studien ønsker å belyse hva som karakteriserer hverandrevurdering i digitale læringsomgivelser.

## **Vurdering med rubrikker**

Innenfor vurdering blir *rubrics*, oversatt til *rubrikker*, forstått som et verktøy for vurdering innenfor en bestemt kontekst, for eksempel i forbindelse med vurdering av tekstlige studentoppgaver (Goodrich, 1997). Rubrikker kan forveksles med sjekklister for spesifikke studentoppgaver, da begge disse innehar *evalueringskriterier* som bestemmer innholdet i oppgaven (Andrade, 2005; Goodrich, 1997). Rubrikker skiller seg imidlertid fra sjekklister da de også innehar komponentene *kvalitetsdefinisjoner* og *scoringsstrategi*. Kvalitetsdefinisjoner beskriver hvordan kvaliteten i studentoppgavene blir bedømt. Et eksempel på dette kan være i en oppgave hvor null skrivefeil blir vurdert til høyeste kvalitet innenfor kriteriet rettskriving. Scoringsstrategien er når en skal gjøre en helhetlig vurdering av studentarbeidet, enten innenfor hvert kriterium, eller som en helhetsvurdering av oppgaven, hvor det blir gitt score fra en oppgitt skala (Popham, 1997). Begrepet rubrikker i forbindelse med studentvurdering har ingen fastsatt definisjon, men begrepet har etablert seg innenfor utdanningsforskningen (Goodrich, 1997). Rubrikker kan brukes i flere vurderingssituasjoner, derunder hverandrevurdering.

## **Fagfellevurdering**

I academia blir ny forskning vurdert i såkalt fagfellevurdering for å kvalitetssikre forskningsbidrag. I fagfellevurdering blir nye forskningsbidrag kritisk vurdert ut fra bestemte vurderingskriterier av eksperter innenfor fagfeltet. Kommunikasjon mellom forskerne er avgjørende for at ny forskning kan kvalitetssikres før publisering (Svartdal, 2021). Studenter som ønsker en karriere i academia bør altså øve seg på å gi gode vurderinger av arbeidet til

likemenn. Hverandrevurdering kan dermed være en øvelse i å vise frem uferdig arbeid som skal vurderes av andre. Også i arbeidslivet er det å kunne gi god feedback til kollegaer en egenskap som er viktig å inneha. Ved hverandrevurdering får studentene også øvelse i å *motta* vurderinger, noe det er behov for. Det viser seg at studenter mangler strategier for hvordan de skal ta i bruk feedback de mottar (Esterhazy, 2019).

### **‘Formativ vurderingspraksis i Canvas’**

Denne masteroppgaven er en del av Universitetet i Oslo sitt forskningsprosjekt med navn ‘Formativ vurderingspraksis i Canvas.’ Prosjektet har som formål å forbedre bruken av formativ vurdering i høyere utdanning gjennom å bruke digitale verktøy. Studenter skal bruke Peergrade eller rubrikker integrert i Canvas LMS for å gjennomføre hverandrevurdering. Dette for å fremme studentaktive læringsformer og for å styrke studentenes akademiske dannelse. Begrepet *akademisk dannelse* defineres slik: “... det som gjør en person til akademiker med evne til kritisk refleksjon. Akademisk dannelse er ikke primært å tilegne seg kunnskap, men å kunne kritisere og utvikle den” (Haanes, 2011, s. 7). Gjennom deltakelse i prosjektet er det ønskelig at studentene selv skal forstå verdien av å gjennomføre hverandrevurdering sammen med sine medstudenter. Dette kan være et steg for å bedre vurderingskulturen ved universitetet. Gjennom prosjektet ønsker en å undersøke om nettbasert hverandrevurdering kan føre til mer og bedre læring i aktuelle emner.

## **1.3 Formål og forskningsspørsmål**

Vurdering har blitt beskrevet som “Achilles’ heel of quality” i høyere utdanning (Knight 2002, s.107, referert i Damsa et al., 2015). Det er bred enighet om at det fortsatt er behov for mer fokus på dette området innen utdanningsforskningen for å styrke studentenes læring. Flere studier viser at vurderingskriterier baserer seg på taus kunnskap og erfaring. Dette fører til en “mystifisering” av vurderingsprosessen (Damsa et al., 2015). Med utgangspunkt i dette sammen med at nettopp vurdering fremstår som et forbedringsområde i offentlige kvalitetsrapporter om høyere utdanning, vil denne studien undersøke vurdering i høyere utdanning. I tråd med fokus på studentaktive læringsformer som er et satsningsområde i stortingsmeldingen Kvalitet for høyere utdanning (Kunnskapsdepartementet, 2016, s. 55-59) vil denne studien derfor undersøke fenomenet hverandrevurdering.



I reelle settinger for hverandrevurdering innehar studentene rolle som både *giver* og *mottaker* av vurdering. Denne empiriske studien vil belyse hvordan studentene opplever å være del av denne prosessen både som giver og mottaker av vurdering. Det vil også bli samlet inn empiri fra undervisere som omhandler hvordan de har designet hverandrevurderingen og hvilke erfaringer de sitter igjen med etter gjennomføring. Double et al. (2020) skriver i sin metastudie om hverandrevurdering at forskere må passe seg for å havne i såkalt “blackboxing” og ønsker derfor å undersøke *når* og *hvordan* hverandrevurdering er mest læringsfremmende. Det er lite forskning som undersøker når og hvordan hverandrevurdering er mest læringsfremmende. Denne studien undersøker hverandrevurdering som del av et større kurs- og læringsdesign, og plasserer hverandrevurdering i en kontekst. Formålet er at innsamlet empiri og analyse vil gi kunnskap om hvordan bruk av rubrikker for hverandrevurdering kan integreres i et læringsdesign i høyere utdanning.

Med utgangspunkt i dette presenteres følgende forskningsspørsmål:

- Hvordan opplever studenter og undervisere bruken av hverandrevurdering i undervisningen?
- Hvilke faktorer har påvirkning på en hverandrevurderingsprosess og hvordan påvirker disse faktorene studentenes opplevde læring?
- Hvordan kan rubrikker for hverandrevurdering integreres som en del av et større læringsdesign på en hensiktsmessig måte?

Ved å bruke metodetriangulering vil fenomenet hverandrevurdering undersøkes. Det blir gjennomført kvalitative intervjuer med studenter og undervisere om deres opplevelse av å gjennomføre en hverandrevurderingsprosess. Det blir også samlet inn kvantitative data i form av spørreskjema fra studentene som omhandler deres erfaringer etter gjennomføring av prosessen.

## 1.4 Oppgavens struktur og innhold

Oppgaven er delt inn i syv kapitler.

*Kapittel 1 Introduksjon* presenterer tema og bakgrunn for valg av tema, samt oppgavens forskningsspørsmål.

*Kapittel 2 Teoretisk rammeverk* tar for seg oppgavens teoretiske perspektiv som består av det sosiokulturelle læringssynet og virksomhetsteorien. Sentrale begrep som beskrives er mediering og artefakter, den proksimale utviklingssonen og scaffolding. Videre blir sentrale komponenter og konsepter som inngår i virksomhetsteorien redegjort for, samt hvordan virksomhetsteorien er overført til denne studien.

*Kapittel 3 Litteraturgjennomgang* gjennomgår tidligere forskning om formativ vurdering, herunder hverandrevurdering. Kapitlet består av hverandrevurdering og læringseffekt, studenters opplevelse av hverandrevurdering, samt ulike faktorer som påvirker en hverandrevurdering. Til slutt utforskes bruk av rubrikker og kursdesign i hverandrevurdering.

*Kapittel 4 Metode* fremstiller studiens forskningsmetodiske innramming. Studien anvender en metodetriangulering med både kvantitativ og kvalitativ metode. Videre presenteres en beskrivelse av casestudien, gjennomføring av studien, og metodiske refleksjoner. Deretter presenteres studiens analytiske strategi, samt studiens validitet og reliabilitet. Kapitlet avsluttes med refleksjoner vedrørende studiens etikk og personvern.

*Kapittel 5 Presentasjon av funn* omhandler analyse av både kvantitative og kvalitative data. Først presenteres deskriptiv statistikk, deretter statistiske analyser, herunder korrelasjonsanalyse og regresjonsanalyse. Videre presenteres kvalitative funn knyttet til fem hovedtema, med hver sine undertema.

*Kapittel 6 Diskusjon* forsøker å besvare oppgavens forskningsspørsmål. Her drøftes funn fra analysen i lys av tidligere forskning og studiens teoretiske rammeverk.

*Kapittel 7 Avslutning* består av en oppsummering av oppgaven og svar på forskningsspørsmålene og teoretiske implikasjoner.

## 2 Teoretisk rammeverk

I følgende kapittel presenteres det teoretiske rammeverket. Det *sosiokulturelle læringssynet* er valgt, siden formålet er å undersøke hverandrevurdering i en sosial praksis og de kollektive interaksjonene som inngår. Videre er begrepene *artefakt* og *mediering* sentrale for å forstå hvordan bruken av rubrikk og digitale verktøy fungerer i aktiviteten.

### 2.1 Sosiokulturelt læringssyn

#### 2.1.1 Individet som del av sosiale praksiser

Det sosiokulturelle læringssynet vektlegger at mennesket er sosialt av natur. Læring anses som et produkt av samspill mellom individer, artefakter og kulturell kontekst. Språket er menneskets viktigste redskap fordi menneskene overfører kunnskap mellom hverandre gjennom å bruke språket (Vygotsky, 1978). Språket innebærer i tillegg til muntlig språk, også ulike multimodaliteter som uttrykk for menneskelige kulturer (Säljö, 2006, s. 25-27). Dette innebærer videre at mennesker formes av språket og kulturen de er del av, og at vi derfor også vil forandre oss i takt med samfunnsendringer. Mennesker benytter felles måter å kommunisere og forstå verden på innenfor de *sosiale praksisene* som de er en del av. Mennesker deltar i kommunikative forløp og dette former dem både som *individer* med personlige særegenheter, men også som *samfunnsborgere* som deler fellestrekk med hverandre. Sosiale praksiser oppstår gjennom at mennesker *handler* med omverdenen og skaper kunnskap sammen med andre individer (Säljö, 2006).

### 2.2 Artefakter og mediering

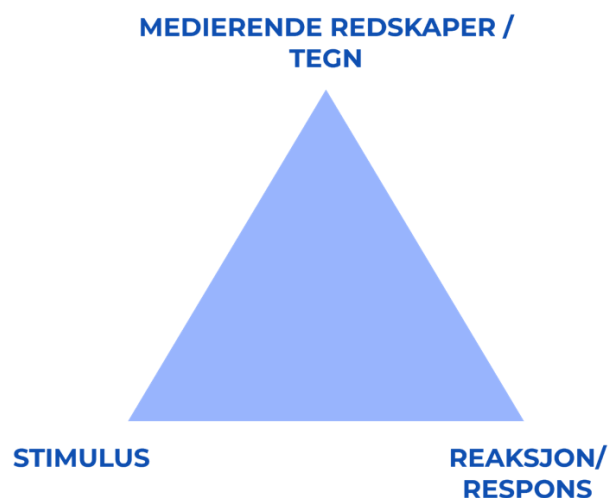
I det sosiokulturelle læringssynet blir mentale funksjoner og handlinger *mediert* gjennom bruk av *artefakter* (Vygotsky, 1978). Mennesker bruker redskaper i sitt daglige virke og er dermed i særstilling fra andre arter. Säljö (2006, s. 83) skriver: "I den sosiokulturelle tradisjonen er den grunnleggende antakelsen (...) at vi mennesker lever i en sosial verden der de fysiske objekter inngår i sosiale praksiser (å leke, å spise, å gå til butikken med med)."

## Artefakter

Vygotsky (1978) skiller mellom to ulike typer redskaper: verktøy (tools) som er fysiske og tegn (signs) som er intellektuelle (psychological tools). Verktøy kan for eksempel være hammer, datamaskine eller programvare. Felles for disse er at de er orientert mot den ytre verden. Tegn er redskaper i språklig form, som for eksempel begreper, symboler og skriftspråk. De språklige redskapene er orientert mot individets mentale prosesser. *Artefakter* er et begrep som brukes om de fysiske redskapene som er menneskeskapt, og som er utviklet med en spesiell hensikt innenfor en sosial praksis. Artefakter kan derfor ikke skilles fra menneskene som lager dem, da de er resultater av menneskelige handlinger. Når mennesket utvikler artefakter, prøver de å gi den aktuelle gjenstanden bestemte egenskaper for å ta den i bruk i bestemte situasjoner. Dette krever kunnskap av produsenten, og kunnskap hos den som skal bruke gjenstanden om hva som nyttiggjør bruken av den. For å lage nye artefakter må vi nyttiggjøre oss av samfunnets kollektive erfaringer (Säljö, 2006, s. 27).

## Mediering

I det sosiokulturelle læringsperspektivet (Vygotsky, 1978) er læring *mediert*. Læring skjer gjennom deltakelse i sosiale praksiser og gjennom å bruke artefakter som har utspring fra kulturen de er en del av. Begrepet *mediering* brukes om hvordan mennesker fortolker verden ved bruk av artefakter. Opprinnelig ble idéen om mediering fremmet som en kritikk av behaviorismens forklaring av læring som et resultat av stimulus og respons. Redskaper og tegn står mellom stimulus og respons (se figur 1), og vi vil derfor aldri forstå menneskelig læring uten å ta tegn og redskaper med i beregningen.



**Figur 1.** Vygotskys opprinnelige idé om mediering. Illustrert fra Säljö (2006, s. 25).

Bruk av tegn og redskaper er så integrert i menneskers leve- og tenkemåte at det omformer hele strukturen i våre psykologiske prosesser, eller arbeidsprosesser hvis det er snakk om redskaper (Säljö, 2006, s. 26-27). Rommetveit skriver: “De medierende redskapene er tegn eller symboler som gjør at vi ikke bare reagerer på signaler; de innebærer at vi også kan tolke omverdenen, ta stilling til den og handle på ulike måter” (1974, sitert i Säljö, 2006, s. 26-27).

I denne studien er artefaktene rubrikker som fylles ut av studentene med feedback til medstudenter. Rubrikkene fylles ut ved å bruke enda to artefakter. Den første er datamaskin med internett og tekstbehandlingsprogram. Datamaskinen brukes for å få tilgang til verktøyene Canvas og Peergrade som blir brukt for å gjennomføre hverandrevurderingen. For at studentene skal nyttiggjøre seg av disse artefaktene må de ha kunnskap om hvordan disse skal brukes på en hensiktsmessig måte. I denne studien antas det at studentene har kunnskap om hvordan de bruker datamaskin, internett og tekstbehandlingsprogram.

## 2.3 Virksomhetsteori

Virksomhetsteorien representerer en tilnærming innenfor det sosiokulturelle perspektivet på læring. Den plasserer menneskelige handlinger i sosiale praksiser ved bruk av artefakter i et system (Engeström, 2010). Den kan ses på som et overordnet teoretisk rammeverk for å studere ulike former for menneskelig praksis, med mål om å forstå enkeltmennesket, og det sosiale miljøet enkeltmennesket tar del i (Engeström, 2005; Kaptelinin & Nardi, 2006). Menneskelige handlinger er et sentralt begrep innenfor virksomhetsteorien, og forklarer hvordan samfunn og verden utvikler seg gjennom menneskelig aktivitet eller virksomhet.

Virksomhetsteorien har sine røtter i russisk psykologi, og ble utviklet av Lev Vygotsky og Aleksei Leontiev. Teorien henter inspirasjon fra tyske filosofer som Kant, Hegel, og Marx I senere tid har Yrjøl Engeström vært sentral i utviklingen av virksomhetsteorien. Med bakgrunn i virksomhetsteoriens historiske utvikling, kan den deles inn i tre generasjoner (Engeström, 2005; Kaptelinin & Nardi, 2006).

### **Førstegenerasjons virksomhetsteori**

*Førstegenerasjons virksomhetsteori* tar utgangspunkt i Vygotskys (1978) triangulære modell av menneskelig handling, se figur 1. Dette er en videreutvikling og nyansering av *stimulus* -

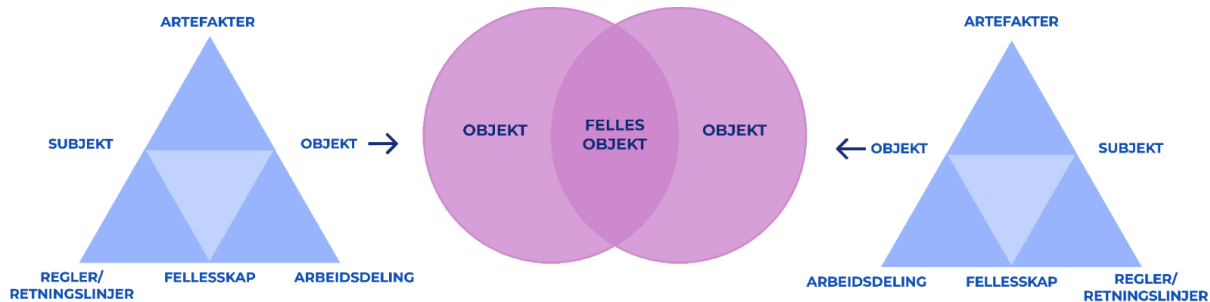
*respons-modellen* kjent innen behaviorismen. Vygotsky introduserte en tredje faktor, *medierende artefakter* siden mennesker sjelden samhandler direkte med omgivelsene (Engeström, 2005; Kaptelinin & Nardi, 2006; Säljö, 2006). De medierende artefaktene, kan sammenlignes med *rubrikkene* som er sentrale verktøy for studentene i forbindelse med hverandrevurderingen. I dette tilfellet kan en si at studentenes handlinger (hverandrevurderingen), medieres ved hjelp av rubrikken.

### **Andregenerasjons virksomhetsteori**

Da Leontiev utviklet *andregenerasjons virksomhetsteori* tok han høyde for forskjellen mellom individuelle handlinger og den kollektive aktiviteten. Den klassiske representasjonen av menneskelig handling hadde en begrensning ved at den ikke tok høyde for det kollektive aspektet som finner sted i sosiale praksiser. Leontiev var opptatt av fellesskapets innvirkning på menneskelig praksis, men inkluderte imidlertid aldri det kollektive aspektet i modellen av et virksomhetssystem (Engeström, 2010). For å illustrere fellesskapets betydning for individets handlinger, bruker Leontiev et jaktlag som eksempel. Et jaktlag kan deles i to grupper, en gruppe som lager lyder for å få byttet til å bevege seg i en bestemt retning, og en annen gruppe som står klar for å drepe dyret. Begge gruppene motiveres av å få dyret drept slik at de kan få mat. Dersom en derimot ser den ene gruppen jegere som jager dyret, isolert for seg selv, virker handlingen svært irrasjonell (Kaptelinin & Nardi, 2006, s. 58). Individenes handlinger kan bare forstås som en del av det større systemet, nemlig jakt-aktiviteten i sin helhet.

### **Tredjegerasjons virksomhetsteori**

Med utgangspunkt i Vygotsky og Leontievs ideer, konstruerte Engeström en modell for det som kalles *tredjegerasjons virksomhetsteori*. For å overkomme begrensningene i den klassiske representasjonen av menneskelig handling, utviklet Engeström en mer kompleks modell bestående av to eller flere virksomhetssystemer. Som illustrert i figur 2, er det inkludert et felles objekt som kobler to systemer sammen. Hvert virksomhetssystem består av seks komponenter, som er i gjensidig påvirkning til hverandre, og som påvirker utfallet av aktiviteten (Engeström, 2010).



**Figur 2.** Virksomhetssystemets tredje generasjon med to samhandlende systemer (Engeström, 2010, s. 136)

### 2.3.1 Virksomhetssystemets seks komponenter

#### Virksomhetssystemets objekt

Et kjerneelement i virksomhetssystemet er *objektet*, siden alle menneskelige aktiviteter er rettet mot et objekt. Dette mente Leontiev er et karakteristisk trekk ved menneskelig aktivitet. “Any activity of an organism is directed at a certain object; an ‘objectless’ activity is impossible” (Leontiev 1981, referert i Kaptelinin & Nardi, 2006, s. 55). Objektet gir motivasjonskraft som former og gir aktiviteten retning, (Engeström, 2010). Det understrekes at objektet ikke må forveksles med mål. Objektet kan være noe materielt, men det kan også være noe mindre håndfast som en plan eller en idé, så lenge den kan deles av deltakerne i aktiviteten. I motsetning til individuelle handlinger, som ofte har et klart definert objekt, så er objektet for en felles aktivitet ikke endelig og helt bestemt, men utvikler seg underveis i prosessen (Virkkunen & Kuutti, 2000). Det enkelte individ som tar del i aktiviteten, kan dermed ha ulik oppfatning av objektet, avhengig av individets rolle i arbeidsfordelingen og erfaring. Objektet vil også ofte være påvirket av andre i virksomhetssystemet som det interagerer med. Dette kan illustreres med et eksempel om en gruppe universitetsstudenter som samarbeider om et prosjektarbeid. Studentene kan ha ulike tanker om hvordan det endelige sluttproduktet skal se ut, og underveis i prosessen påvirker studentene hverandre slik at deres syn på sluttproduktet endrer seg.

#### Subjekt - objekt - artefakt

Toppen av virksomhetsmodellen til Engeström (2010) tilsvarende modellen konstruert av Vygotsky, som forklarer hvordan *handling* blir *mediert* av *artefakter*, se figur 1. *Subjektet* refererer til deltakerne av aktiviteten, som kan være enkeltindivider eller en gruppe

mennesker. Ved et universitet ville dette vært undervisere og studenter. *Verktøyet* som kan være fysisk eller psykologisk, er det som medierer relasjonen mellom subjekt og objekt. Ved et universitet kan dette for eksempel være lærebøker, tavle med kritt eller programvarer som studentene bruker. *Objektet* henviser til retningen på subjektets handling og kan karakteriseres som handlingens formål. Ved et universitet kan objektet for eksempel være å gjøre ferdig en gruppepresentasjon som flere studenter samarbeider om.

### **Regler - fellesskap - arbeidsfordeling**

For å få frem kompleksiteten i menneskelig praksis, inkluderer Engeström (2010) den kollektive og kontekstuelle dimensjonen i modellen. Analyseenheten går dermed fra å være individorientert, til å inkludere *regler, fellesskap og arbeidsfordeling*. Med *regler* menes skrevne og uskrevne regler, som regulerer aktivitetene i systemet. Eksempler på skrevne regler ved et universitet kan være reglement knyttet til plagiat og juks ved eksamen, eller sykefravær ved praksisutplassering. Uskrevne regler kan derimot være at studentene må slå mobilen på lydløs i forelesning, og at en rekker opp hånden før en tar ordet. Fellesskapet bestemmer regler som individet må forholde seg til i en gitt sosial praksis. Samtidig påvirker enkeltindivider fellesskapet og er med å utforme gjeldende regler. *Arbeidsdeling* handler om hvordan fordeling av arbeidsoppgaver er blant deltakerne i systemet. De faglige ansatte har arbeidsoppgaver knyttet til undervisning og forskning, mens studentene har studentarbeid knyttet til egen læring og utdanning. *Fellesskap* refererer til alle aktørene involvert i et enkelt virksomhetssystem.

### 2.3.2 Sentrale konsept i virksomhetsteorien

#### **Virksomhetens kontekst og historiske utvikling**

Et sentralt aspekt i virksomhetsteorien er fenomeners *dynamiske utvikling* over tid, som i eksempelet med prosjektarbeid. Interne motsetninger er drivkraften til endring og utvikling av virksomhetssystemet. Det kan være motsetninger innad i studentgruppen som former og driver utviklingen av aktiviteten i en bestemt retning (Engeström, 2010). Selv om noen av deltakerne i virksomhetssystemet kan være mer sentrale og ha mer makt, understreker Virkkunen og Kuutti (2001) at deres meninger ikke skal ha en høyere prioritet i analysen, men at en derimot skal beskrive den historiske utviklingen av aktiviteten og dens formål. Prinsippet om utvikling kan knyttes til Vygotsky (1978) som mente at en også må inkludere



konteksten og dens historiske utvikling for å få en fullstendig forståelse for fenomenet som undersøkes.

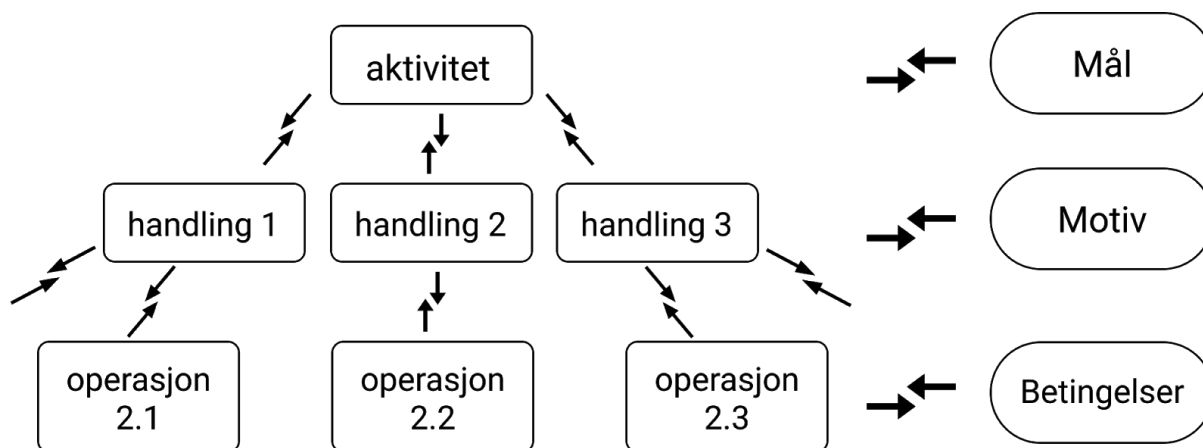
Fra et virksomhetsteoretisk ståsted ser en all praksis som et resultat av bestemte historiske utviklinger under bestemte forhold. Utviklingen foregår kontinuerlig, og er med på å utvikle selve praksisen (Kaptelinin & Nardi, 2006). Utfallet av aktiviteten vil ofte bli en del av eller påvirke et annet virksomhetssystem, noe tredjegenasjon virksomhetsteori tar høyde for. En analyse av objektet, dets kontekst og utvikling, er derfor nødvendig for å forstå menneskelig praksis utført av et enkeltindivid eller en gruppe mennesker.

### **Motsetninger og spenninger i virksomhetssystemet**

Som beskrevet ovenfor er et sentralt prinsipp i virksomhetsteorien at motsetninger og spenninger er en viktig drivkraft for utvikling av aktiviteter. Engeström (2010) bruker begrepet *contradictions* for å forklare hvordan motsetninger og spenninger oppstår i et virksomhetssystem. I denne sammenhengen er ikke en motsetning det samme som en konflikt. Inkludering av nye element i en aktivitet (for eksempel en ny teknologi) kan føre til en motsetning der gamle elementer kolliderer med det nye elementet (Engeström, 2010). En slik motsetning vil være en sentral kilde for utvikling av aktiviteten. Motsetninger kan oppstå innad i et virksomhetssystem og mellom to eller flere virksomhetssystemer.

### **Virksomhetssystemets hierarkiske aktivitetsstruktur**

Innenfor virksomhetsteorien kan en forstå menneskets interaksjon med verden ved å bryte aktiviteter opp og organisere dem i et hierarkisk system. Dette kan organiseres i tre lag, bestående av *aktiviteter*, *handlinger* og *automatiske operasjoner* (se figur 3) (Leontiev, 1974, referert i Kaptelinin & Nardi, 2006, s. 62- 64). Øverst i hierarkiet er selve aktiviteten som er orientert mot aktivitetens motiv. Det midterste laget består av handlinger. Handlinger er bevisste målrettede prosesser som gjennomføres for å oppfylle hovedmotivet til aktiviteten. En aktivitet kan bestå av flere handlinger, som igjen kan brytes ned til enkeltoperasjoner. Ifølge virksomhetsteorien er enkeltoperasjoner automatiske prosesser, som for eksempel handlinger en har utført gjentatte ganger slik at de automatiseres. De ulike nivåene er dynamiske, ved at det er mulig å bevege seg mellom de ulike lagene. For eksempel når en skal lære å kjøre bil, vil det å skifte gir i starten være en bevisst handling, mens den etter hvert automatiseres (Leontiev, 1974, referert i Kaptelinin & Nardi, 2006, s. 62-64).



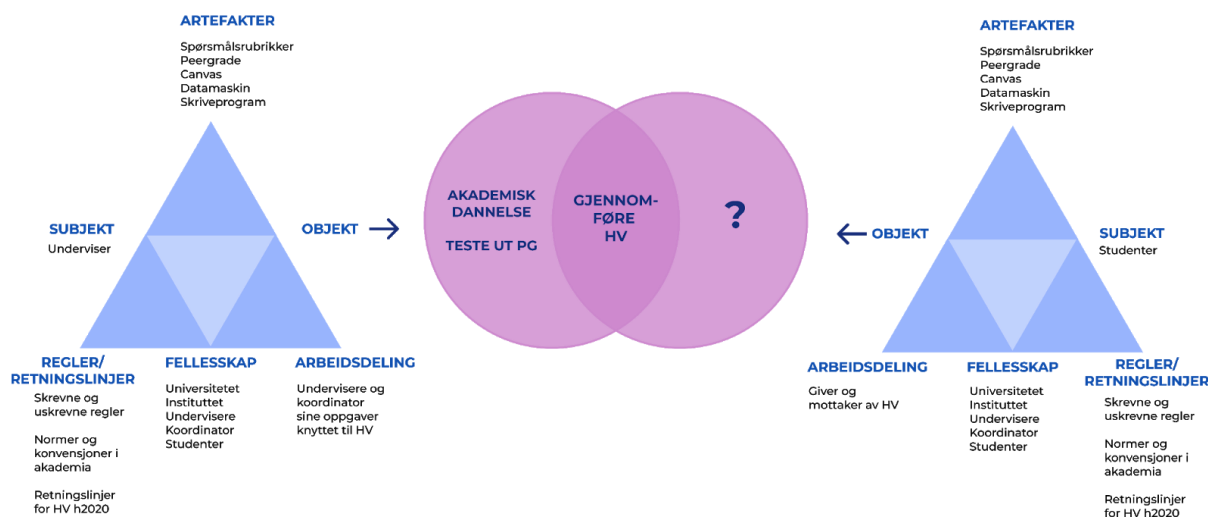
*Figur 3. Illustrasjon av Leontievs hierarkiske modell av en aktivitet (Kaptelinin & Nardi, 2006, s. 64)*

Den hierarkiske aktivitetsstrukturen kan eksemplifiseres ved en student som skal gi og motta feedback på en skriftlig oppgave. Målet med aktiviteten vil være å både gi og motta feedback som skal bidra til at studentene sine tekster blir bedre. Veien mot å gjennomføre hverandrevurdering, vil bestå av flere handlinger. Når det gjelder å gi feedback må studentene lese en medstudent sin tekst, vurdere teksten, for deretter å gi feedback. Handlinger knyttet til å motta feedback vil være å tolke feedbacker, vurdere kvaliteten på feedbackene for deretter å endre egen tekst basert på disse. En del av disse handlingene vil igjen bestå av en rekke automatiske operasjoner. For eksempel vil det å trykke på tastaturet være automatiske operasjoner for mange studenter når de bruker datamaskinen til å skrive en oppgave. For en person som ikke er erfaren med å skrive på datamaskin, vil dette derimot være bevisste handlinger. Med erfaring kan denne prosessen gå fra å være en bevisst handling til en automatisk prosess. Dette illustrerer det dynamiske aspektet ved virksomhetssystemets struktur.

### 2.3.3 Virksomhetsteorien overført til denne studien

Som konseptuelt rammeverk benyttes tredjegerasjons virksomhetsteori. Dette fordi den tar høyde for interaksjonen mellom to eller flere virksomhetssystemer. En analyse av hverandrevurdering utført høsten 2020 ved hjelp av virksomhetsteorien, kan belyse både studentenes og lærernes virksomhetssystem, og deres interaksjon med hverandre. Fokuset i denne oppgaven er hvordan de ulike komponentene i et virksomhetssystem påvirker hverandre, og hvordan de medierende artefaktene, blant annet *rubrikker*, støtter handlingene. Studentene og undervisernes forståelse av den kollektive aktiviteten (hverandrevurderingen)

og dens formål kan hjelpe oss til å forstå hva som kan endres for å forbedre prosessen. Figur 4 viser virksomhetssystemet i denne studien med gjennomføring av hverandrevurdering.



**Figur 4.** Modell av virksomhetssystemet i denne studien før gjennomføring av hverandrevurdering

### Virksomhetssystem 1: Underviserne

Virksomhetssystemet kan forstås både fra undervisere og studenters ståsted. Underviserne har planlagt, organisert og gjennomført hverandrevurderingen i sitt emne. Virksomhetssystemet kan derfor også forstås med underviserne som *subjekter* i aktiviteten, se triangelet til venstre i figur 4. For å legge til rette for hverandrevurdering mellom studentene, har underviserne tatt i bruk *medierende artefakter*, hvor *rubrikkene* har en sentral rolle som artefakt. Underviserne bruker også det digitale verktøyet Canvas for å sende ut rubrikkene studentene skal bruke, og selve hverandrevurderingen blir gjennomført enten i Canvas eller i det digitale verktøyet Peergrade. Disse verktøyene utgjør sammen de medierende artefaktene som står mellom *subjekt* (underviserne) og *objekt*. Objektet motiverer aktiviteten, og for underviserne er objektet at studentene gjennomfører hverandrevurderingen seg imellom. Også utprøving av artefakten Peergrade kan forstås som et objekt. Forskningsprosjektet ‘Formativ vurderingspraksis i Canvas’ har som formål å utvikle normative retningslinjer for bruk av digitale verktøy i akademia, deriblant verktøy for hverandrevurdering. Peergrade er verktøyet som skal prøves ut. Videre er et mål for forskningsprosjektet at studentene skal forbedre sin akademiske dannelse gjennom å delta i hverandrevurdering. *Arbeidsfordeling* viser hvordan oppgaver fordeles mellom deltakerne i aktiviteten. Underviserens rolle er å fasilitere at studentene gjennomfører hverandrevurderingen. Dette innebærer blant annet utforming av rubrikker og organisering av aktiviteten med innleveringsfrister og lignende. *Fellesskapet* i

dette virksomhetssystemet består av studenter, undervisere og koordinatorene som står ansvarlig for det faglige innholdet og organiseringen av undervisningen. Disse jobber sammen for et felles formål. Alle deltakerne i fellesskapet må forholde seg til regler, men i dette virksomhetssystemet er det underviserne som utformer de eksplisitte reglene for aktiviteten, som for eksempel innleveringsfrister og ordkrav.

### **Virksomhetssystem 2: Studentene**

I forbindelse med gjennomføring av hverandrevurdering, har studentene fått i oppgave å vurdere en tekst skrevet av en medstudent. I denne konteksten er studentene *subjektet* i aktiviteten, se høyre side av figur 4. For å gjennomføre oppgaven, benytter studentene et utvalg *medierende artefakter* der hovedverktøyet som benyttes er *rubrikkene*. *Rubrikkene* ligger tilgjengelig i Canvas og Peergrade. I arbeidet med hverandrevurderingen er rubrikken det medierende verktøyet mellom *subjektet* (student) og *objektet* (vurdering av en medstudents tekst). Rubrikkene skal fungere som en støtte og guide for studentene når de vurderer hverandres tekster. *Objektet* kan i denne sammenhengen forstås som selve vurderingen studentene skriver etter å ha lest teksten til en medstudent. Å transformere objektet “hverandrevurderingen” til et ønsket resultat, er det som motiverer aktiviteten. I modellen står det spørsmålsteget ved objektet til studentene fordi dette ikke kan slås fast før de har gjennomført aktiviteten og dataene knyttet til dette er analysert. Dersom en finner at studenter og undervisere har motstridende motiver for handlingen, kan dette være uheldig fordi ulike aktører i samme virksomhetssystem drar i ulike retninger. *Arbeidsfordeling* viser til hvordan ulike oppgaver fordeles mellom deltakerne i aktiviteten. I dette tilfellet vil studentene inneha rollen som både giver og mottaker av en vurdering. *Fellesskapet* i dette virksomhetssystemet består av forskjellige personer som har en rolle i kurset. Det er både studenter som er deltakere i kurset, samt undervisere og koordinatorene. Som student og deltaker i undervisningen, er det ulike *regler* å forholde seg til. Det kan være tidsfrister for levering av hverandrevurderingen eller uformelle regler, som at hverandrevurderingen skal være av en viss kvalitet.

## 2.4 Den proksimale utviklingssonen og scaffolding

Læring må tilpasses den lærendes utviklingsnivå (Vygotsky, 1978). Det går et skille mellom hva den lærende er i stand til å tilegne seg av kunnskap og ferdigheter alene, og hva den er i

stand til å lære seg uten hjelp av andre. Den lærende er i stand til å lære seg atskillig mer med hjelp av en *mer kompetent annen* som kan veilede i læringsprosessen. Den mer kompetente annen er en person som kan mer enn den lærende, innenfor det aktuelle domenet, og dermed kan overføre sin egen kunnskap (Vygotsky, 1978). Den mer kompetente annen finner vi ofte i rollen som lærer eller forelder, men det kan også være en medstudent som er mer kompetent innenfor det aktuelle domenet.

“Zone of Achieved Development” oversatt til *det faktiske utviklingsnivået* er det nivået i utviklingen den lærende allerede har opparbeidet seg gjennom deltakelse i læringsfellesskap. Den allerede etablerte kunnskapen bygges videre på, og ved hjelp av en mer kompetent annen, kan den lærende tilegne seg kunnskaper som ligger i det som kalles for “Zone of Proximal Development,” oversatt til *den proksimale utviklingssonen* (Vygotsky, 1978). Definisjonen for den proksimale utviklingssonen lyder: “The distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (s. 86). Dette skjer ved at den mer kompetente annen modellerer, forklarer eller gir hint om den aktuelle oppgaven som den lærende står overfor. På denne måten kan den lærende lære seg noe som den ikke ville vært i stand til å lære på egenhånd.

Som en konkretisering av Vygotskys idé om den proksimale utviklingssonen introduserte Wood et al. (1976) begrepet *scaffolding*, som på norsk oversettes til stillasbygging, hvor en ser for seg at en mer kompetent annet bygger et stillas rundt den lærende. Ved å bruke ulike undervisningsteknikker kan en legge til rette for at den lærende nærmer seg læringsmålet gradvis innenfor sin proksimale utviklingssone: “To enable a child or novice to solve a problem, carry out a task or achieve a goal that would be beyond his unaided performance” (Wood et al., 1976, s. 90). Denne forståelsen er lik begrepet den proksimale utviklingssonen, men scaffolding bidrar med en utvidet forståelse av hvordan en mer kompetent annen kan støtte den lærende i praksis. Scaffolding er en *metode* for å få den lærende til å utvikle seg innenfor sin utviklingssone. Scaffolding kan integreres i flere læringskontekster og innebærer at den lærende interagerer med omverden for å ta til seg kunnskap, og passer med det inn i det sosiokulturelle læringssynet.

Proessen innebærer at en mer kompetent annen har et bevisst forhold til oppgaven som den lærende skal løse. En mer kompetent annen hjelper den lærende til å kun fokusere på de deloppgavene som det er mulig for den lærende å oppnå på nåværende nivå i læringsprosessen (Wood et al., 1976). Etter hvert som den lærende opparbeider seg videre forståelse for oppgaven, vil den lærende trenge mindre støtte, og en kan begynne med gradvis *fading* hvor stillaset fjernes del for del. Begrepet *fading* blir av Collins et al. (1989) brukt slik: “Once the learner has a grasp of the target skill, the master reduces (or fades) his participation, providing only limited hints, refinements, and feedback to the learner, who practices successively approximating smooth execution of the whole skill “ (s. 456). Dette innebærer at en mer kompetent annen trekker seg vekk og lar den lærende løse stadig mer krevende deler av oppgaven på egenhånd. Til slutt vil ikke den lærende ha behov for mer veiledning og stillaset kan fjernes. Medstudenter kan også scaffolde hverandre, men det viser seg at scaffolding mellom to uerfarne er mindre eksplisitt og gjennomtenkt enn den gitt av lærer. En uerfaren, for eksempel en student, synes det er vanskeligere enn en lærer å identifisere sin medstudent sitt ferdighetsnivå og behov videre (Tabak & Baumgartner, 2004).

Van de Pol et al. (2010) har gjennomgått studier hvor scaffolding er brukt som rammeverk og har identifisert felles karakteristikk. De finner at den første felles karakteristikken er *contingency*, som blir referert til som tilpasning, respons eller kalibrert støtte av den lærende. Den andre karakteristikken er *fading*, hvor stillaset/støtten rundt den lærende gradvis fjernes. Dette fører så til den tredje karakteristikken, *transfer of responsibility*, som innebærer at den lærende selv tar kontroll over aktiviteten som den har lært, og dermed også egen læring.

### **Scaffolding med rubrikker**

I formativ vurdering bør vurderingen som den lærende mottar være tilpasset den enkelte og hva den kan være i stand til å oppnå (Black & William, 2009; Hattie & Timperley, 2007). Vurdering må altså være tilpasset den enkeltes proksimale utviklingssone, som van de Pol et al. (2010) refererer til som *contingency*, og dermed kan en forstå formativ vurdering som en form for scaffolding. Shepard (2005) skriver at scaffolding og formativ vurdering er “essentially the same thing” da både scaffolding og formativ vurdering har som mål å fremme læring innenfor den proksimale utviklingssonen. I vurderingspraksis kan rubrikker være et verktøy for vurdering, for eksempel i forbindelse med vurdering av tekstlige studentoppgaver

(Goodrich, 1997). Bruk av rubrikker til hverandrevurdering kan forstås som en form for scaffolding i høyere utdanning. Rubrikkene blir utformet av underviser, og underviser utformer rubrikkene på en måte som den mener er passende for studentene ved det aktuelle emnet. Førsteårsstudenter vil for eksempel motta rubrikker som er mindre krevende å fylle ut enn rubrikker som masterstudenter benytter. Underviseren utformer rubrikkene med utgangspunkt i hva studentene trenger å forbedre seg i. For eksempel kan underviser lage rubrikker som fokuserer på drøfting, dersom den vet at studentene trenger øvelse i dette. Ved å tilpasse rubrikkene til studentenes behov vil de kunne hjelpe studentene med å utvikle seg innenfor sin proksimale utviklingssone, og rubrikkene er derfor en form for scaffold.

## 3 Litteraturgjennomgang

Litteraturen er valgt ut for å belyse oppgavens forskningsspørsmål. Her presenteres funn fra både kvantitative og kvalitative studier som undersøker ulike aspekter ved formativ vurdering i høyere utdanning. Utvalget av litteratur vil være relatert til følgende:

- Formativ vurdering
- Hverandrevurdering
- Bruk av rubrikker for vurdering
- Kursdesign i hverandrevurdering

### 3.1 Formativ vurdering

Vurdering blir forstått som en komponent i pedagogisk praksis som er avgjørende for den lærendes utvikling og læring (Black & Wiliam, 1998, Black & Wiliam, 2009; Hattie & Timperley, 2007). En skiller mellom *formativ* og *summativ* vurdering, også kalt vurdering *for* og vurdering *av* læring. Formativ vurdering er en vurderingsform som sikter mot å fremme læring, mens summativ vurdering er en oppsummering av hva den lærende har lært mot slutten av et læringsforløp (Sadler, 1989; Black & Wiliam, 2009). Videre er konseptene *vurdering* og *feedback* nært beslektet i pedagogisk praksis (Damsa, et al., 2015). Formativ vurdering kan organiseres på ulike måter, men hovedfunksjonen er å gi feedback på studentens arbeid for å forbedre læringen (Sadler, 1989). Med bakgrunn i dette betraktes feedback som en potensiell viktig faktor for kvaliteten i høyere utdanning (Esterhazy et al., 2020).

#### **Produktive feedback- praksiser i høyere utdanning**

Studier viser at feedback er en av faktorene som har sterkest påvirkning på studentprestasjoner. Ulike typer feedback gir ulik effekt. Feedback som inkluderer informasjon om oppgaven og hvordan forbedre den har høyere effekt enn studier med lavere effekt var feedback i form av ros, belønning og straff (Hattie & Timperley, 2007). Generelt oppfattes feedback som god når den fremmer selvtillit, øker motivasjon og bygger relasjoner (Esterhazy et al., 2020). Videre bør feedback formidles i et spesifikt og tydelig språk, som gjør at studenten forstår feedbacken og kan ta den i bruk (Shute, 2008). Betydningen av hvordan feedback blir gitt og formulert kan ses i lys av enkelte studier som viser at feedback



har begrenset verdi eller at den faktisk kan ha en negativ effekt på læring. Det kan være feedback som er formulert for kritisk eller kontrollerende (Shute, 2008).

I en casestudie med ingeniørstudenter undersøker Esterhazy (2018) relasjonen mellom ulike faktorer i kurset og hvordan disse har betydning for fremveksten av det som hun kaller for “feedback encounters”, oversatt til tilbakemeldingsmøter. Studien viser at tilbakemeldingsmøter var produktive når studentene hadde mulighet til å bli kjent med og utforske verktøy og konvensjoner som er nødvendig for å delta i den pedagogiske praksisen. Det er derfor viktig å ta i betraktning det enkelte fagets praksis og konvensjoner når en designer for læringsaktiviteter som sikter mot at studentene skal være aktive deltakere. Et annet studie som undersøker hvordan studenter tar i bruk feedback de har mottatt, er en litteraturgjennomgang gjennomført av Winstone et al. (2017). Denne studien tar for seg ulike karakteristikk ved *konteksten*, *feedbacken*, *sender* og *mottaker* av feedback, som påvirker hvor produktiv feedback-praksisen er. Studien understreker studentens rolle som aktiv og engasjert bruker av feedback, for at den skal ha en effekt. Ved formativ vurdering ligger ansvaret for læringsprosessen både på lærer og på studenten. Det forutsetter at studenten er aktiv og engasjert i egen læring (Winstone et al., 2017).

Esterhazy et al. (2020) har undersøkt bruken av ulike former for feedback-praksiser. I deres forskning gjennomgår de tre casestudier ved ulike fagdisipliner i Norge. Deres funn er i tråd med andre studier som viser at kriterier for hva som regnes som feedback av god kvalitet, kan best forstås i sammenheng med det aktuelle læringsmiljøet. Effektiviteten til de ulike formene for feedback avhenger av strukturelle, relasjonelle og epistemiske dimensjoner i det konkrete læringsmiljøet. Det argumenteres for at det må utvikles mer kontekstsensitive måter for å evaluere og forbedre kvaliteten på feedback i høyere utdanning (Esterhazy et al., 2020).

### 3.1.1 Vurderingskultur i høyere utdanning

For å unngå at vurderingsprosessen blir oppfattet som uklar av studenter, er det nødvendig med eksplisitte og transparente vurderingskriterier. Dette for at studenter skal få en forståelse for hvilke kriterier de vurderes ut fra. En studentgruppe som drar særlig nytte av dette er førsteårsstudenter (Damsa et al., 2015). Funn fra studier viser også at flere studenter i høyere utdanning opplever at de ikke har gode nok strategier for hvordan de faktisk kan ta i bruk feedback som de har mottatt (Winstone et al., 2017). Den pedagogiske praksisen ved flere

utdanningsinstitusjoner sikter mot mer studentsentrerte tilnærminger, der studenten er mer aktiv i egen læringsprosess (Black & Wiliam, 2009; Damsa et al., 2015; Winstone et al., 2017). Dette kan legges til rette for ved å bruke egen- og hverandrevurdering. Disse vurderingsformene gir gode muligheter for at studentene kan internalisere akademiske standarder og deretter være bedre rustet til å overvåke egen læringsprosess og løse problemer ut fra disse standardene (Gibbs, 2006).

### **Karakteristikker for hvordan studenter tar i bruk vurdering**

Hva som blir vektlagt i vurderinger sender signaler til studenter om hvilke fagområder de bør fokusere på. En etnografisk studie av studenters erfaringer gjennomført ved Massachusetts Institute of Technology (MIT) viser hvordan studenter strategisk jobber seg gjennom pensum på jakt etter det en får på eksamen, og ignorerer resten av pensum (Snyder, 1971). Snyder refererer til dette som "*The Hidden Curriculum*". Studentenes oppfatning av vurderingskravene er sterkt knyttet til deres tilnærminger til læring (Weurlander et al., 2012). Studentenes fokus på karakterer kan gå på bekostning av feedbacken de mottar (Gibbs, 2006). Winstone et al. (2017) finner i sin litteraturgjennomgang at studier viser at studenter ofte har en begrenset forståelse av formålet med feedback. Det vil si at de forstår at feedback skal fasilitere læring, men anerkjenner ikke i like høy grad deres ansvar i prosessen. De fleste studier som har undersøkt hvordan studenter tar i bruk feedback de har mottatt, baserer seg på selvrapporteringsdata fra studenter. Få studier har undersøkt hvordan studenter faktisk bruker, eller eventuelt ikke bruker feedback (Esterhazy et al., 2020).

## **3.2 Hverandrevurdering**

Hverandrevurdering (peer assessment) er en metode hvor likemenn, for eksempel studenter, vurderer hverandres faglige oppnåelse. Feedback i hverandrevurdering kan gis i flere former, blant annet som skriftlige kommentarer, karakterer eller gjennom muntlig dialog (Topping, 1998).

### **3.2.1 Hverandrevurdering og læringseffekt**

Hverandrevurdering har vist seg å ha læringseffekt på tvers av kontekster og utdanningstrinn (Double et al., 2020; Topping, 1998). Når det gjelder kvaliteten på feedback som studentene gir til hverandre har flere studier konkludert med at feedback gitt av studenter holder like høy kvalitet som feedback fra lærer (Topping, 1998; Zheng et al., 2018).

Double et al. (2020) har gjort en metastudie (54 studier) om hverandrevurdering i høyere utdanning og på lavere skoletrinn. I studien blir læringseffekt av hverandrevurdering sammenlignet med læringseffekt av andre vurderingsformer. Denne studien konkluderer med at deltakelse i hverandrevurdering er mer læringsfremmende enn (1) ingen vurdering og (2) vurdering fra lærer. Sammenlignet med (3) egenvurdering, er hverandrevurdering like læringsfremmende. Hverandrevurdering har vist seg å ha positiv effekt på studenter sine skriveferdigheter (Double et al., 2020). I en metastudie fra Huisman et al. (2018) finner en at studentene som har mottatt vurdering fra medstudenter forbedrer sine skriveferdigheter sine like mye som dem som mottar vurdering fra lærere.

### **Hvorfor er hverandrevurdering læringsfremmende?**

En av faktorene som gjør at hverandrevurdering fremmer læring er at studentene blir tvunget til å bruke mer tid på oppgaven gjennom aktiviteten hverandrevurdering og deloppgaver knyttet til denne (Topping, 2005).

Flere studier vektlegger læringsprosessene forbundet med å *gi* hverandrevurdering. Kognitive prosesser som inngår i å *gi* hverandrevurdering innebærer blant annet å sammenligne eget arbeid med andres. Dette gjør at studentene øker sitt eget refleksjonsnivå og kan forstås som en metakognitiv handling (Topping, 2005; van Popta et al., 2017). Videre er det å *gi* hverandrevurdering en god øvelse for at studentene skal lære seg å ta i bruk vurderingene som de *mottar*. Nicol (2012, 2014) forklarer denne prosessen som bestående av to handlinger: å *gjøre en evaluering* og å *gi en forklaring*. Når studentene gir vurderinger, interagerer de med faginnholdet, behandler det, sammenligner det med alternativt innhold, ser det fra forskjellige perspektiver, og skaper ny kunnskap. Det andre elementet i reflekterende kunnskapsbygging som Nicol nevner, er å *gi en forklaring* for vurderingen. Når studentene blir spurt om å *gi en forklaring*, må de aktivt bruke følgende kognitive prosesser: overvåke, evaluere og øve på egen forståelse. Det krever at studentene vurderer arbeidet for å identifisere feil, mangler i forståelsen, misforståelser og alternative synspunkter, og å konstruere skriftlig feedback. På den måten blir studentenes egne ideer om emnet tilgjengelig for egnevaluering og oppdatering.

I en kvalitativ studie av To og Panadero (2019) med 11 førsteårsstudenter som gjennomførte hverandrevurdering viser det seg at hverandrevurdering også har positiv innvirkning på egenvurdering. Dette fordi studentene fikk økt forståelse av hva som er kvalitet, mer selvstendige vurderinger og økt selvrefleksjon. Ved å sammenligne eget arbeid med andres kan en både lære av andre, og reflektere over egne svakheter, noe som forbedrer evnen til egenvurdering (Panadero, 2016).

### 3.2.2 Studenters opplevelse av hverandrevurdering

Vickerman (2009) har undersøkt studenters erfaringer med bruk av hverandrevurdering med 90 studenter ved et idrettsstudie i UK. Studiet ble gjennomført gjennom spørreskjema med Likert-skala og åpne spørsmål. På spørsmål om studentene er enige i at hverandrevurdering forbedrer deres evne til å reflektere kritisk svarte 78% av de spurte at de er enige i dette. To og Panadero (2019) sin studie viser at studentene får nye oppfatninger av eget arbeid gjennom å gjennomføre hverandrevurdering. Videre kom det frem at studentene mente at hverandrevurderingen hjalp dem til å gjøre selvrefleksjon. Det er gjort funn som viser at studenter synes at det å lese andres tekster kan gi dem inspirasjon til problemløsning og dermed hjelpe dem med å løse utfordringer de står overfor i eget arbeid (Cartney, 2010).

Cartney (2010) har utført en casestudie om hverandrevurdering med 45 studenter. I denne studien uttrykker studentene at de er engstelige for å vurdere sine medstudenter, blant annet fordi de er redde for å gi en vurdering som ikke er korrekt. Videre uttrykte studentene bekymring for å måtte vurdere en tekst som var så dårlig at de ikke ville klare å si noe positivt. Dette finner en også i studien til Xiao og Lucking (2008). Det er gjort flere studier som viser at studentene foretrekker å motta vurdering fra lærer fremfor medstudenter (Misiejuk et al. 2021; Panadero, 2016; To & Panadero, 2019; Weaver, 1995). I flere studier forklarer studentene at de ikke stoler like mye på vurderingene fra medstudenter som vurderingene fra lærer (Cartney, 2010; To & Panadero, 2019). Noen forklarer også at de ikke stoler på sin egen akademiske kompetanse, og at vurderingene som de gir derfor er mer rettet mot språklige feil og mangler, da de mener det er denne typen feedback de er mest kompetent til å gi.

Cartney (2010) sin studie finner at studenter blir demotivert av å motta korte og lite utfyllende vurderinger fra sine medstudenter dersom de selv mente at de hadde levert en utfyllende

vurdering. I denne studien var studentene inndelt i grupper, og det viste seg at gruppene hvor alle deltakerne hadde vist høyt engasjement satt igjen med en bedre opplevelse av prosessen.

### **Å ta i bruk vurdering fra medstudenter**

Misiejuk et al. (2021) har brukt *learning analytics* for å få innsikt i hverandrevurderingsprosessen. Fokuset for denne forskningen har vært på ‘backward evaluation’ - altså hvordan studentene reagerer på og tar i bruk vurderingen som de mottar av medstudenter. Denne delen av hverandrevurdering er lite forsket på. I studien har de fått tilgang til et stort datasett (N= 7660) fra det nettbaserte vurderingsverktøyet Peergrade. I denne studien finner forskerne det de kaller for “The feedback gap” som beskriver et sprik mellom vurderingen som studentene mottar, og hvordan de tar i bruk vurderingen for å gjøre endringer i eget arbeid. Videre viser denne studien at de fleste vurderingene som ble gitt ikke var nok spesifiserte, og at dette kan være med å forklare hvorfor studentene ikke tar i bruk vurderingen som de mottar. Hvordan studentene tar i bruk vurderingen de mottar blir påvirket av både ytre og indre forhold (Misiejuk et al., 2021). Disse funnene samsvarer med andre funn som omhandler studenters bruk av vurdering som de mottar fra lærer. Videre viser Misiejuk et al. (2021) sin studie at studentene som klarte å nyttiggjøre seg av vurderingen som de mottok av medstudenter, var mer åpne for å innrømme feil og mangler. Studentene som ikke nyttiggjorde seg av vurderingene var mer defensivt innstilt, og oftere misforstod og/eller var uenige i vurderingen som de mottok.

### **3.2.3 Faktorer som påvirker hverandrevurdering**

Forskere har undersøkt betydningen av ulike faktorer forbundet med gjennomføring av hverandrevurdering. De ulike faktorene som påvirker prosessen og læringseffekten vil nå gjennomgås.

#### **Nettbasert og ikke-nettbasert**

Double et al. (2020) sin metastudie som tar for seg variabler som påvirker hverandrevurdering, viser at det ikke er noen forskjell i læringseffekten om den gjennomføres nettbasert eller ansikt til ansikt. van Popta et al. (2017) har i sin metastudie sammenlignet syv empiriske studier og en litteraturgjennomgang som omhandler hverandrevurdering i nettbaserte løsninger. Resultatene viser at nettbasert gjennomføring også gir læringseffekt.

I studien til van Popta et al. (2017) undersøkte forskerne læringsutbytte studentene får av å gi nettbasert hverandrevurdering og kognitive prosesser forbundet med dette. Det var fire studier som omhandlet det å gi hverandrevurdering, og disse viser at denne prosessen øker refleksjonsnivået hos giver. Dette fordi det involverer kognitive prosesser som selvregulering, overvåkning av egen læringsprosess, og utvikling av kritisk sans til eget og andres arbeid. Dette er de samme kognitive prosessene som er involvert i hverandrevurdering gitt ansikt-til-ansikt (Nicol 2012; Nicol, 2014; Topping, 2005; van Popta et al., 2017).

I Cartney (2010) sin casestudie var hverandrevurderingen designet slik at den skulle foregå gjennom mail. I denne studien forklarte flere av studentene at de hadde valgt å gå vekk fra dette. Disse valgte å gjennomføre hverandrevurderingen ansikt-til-ansikt. Studentene begrunnet dette valget med at det var lett å misforstå hverandre med nettbasert gjennomføring. Gjennomføring ansikt-til-ansikt åpner for å kunne stille spørsmål underveis og dermed oppklare eventuelle misforståelser. Cartney tror dette skyldes at studentene ikke er vant til å kommunisere i nettbaserte læringsmiljøer. Denne studien er dog fra 2010, så det er grunn til å tro at dagens studenter vil være mer komfortable med å bruke nettbaserte verktøy, også til denne typen formål. En av gruppene valgte imidlertid å gjennomføre hverandrevurderingen gjennom mail. Dette begrunnet de med at det var lavere terskel for å være ærlig over nett, fremfor ansikt-til-ansikt.

## **Roller**

De fleste studier som er gjennomført om hverandrevurdering har et design hvor deltakerne innehar rollen som både *giver* og *mottaker* av vurdering. Dette gjør at en ikke kan slå fast om det er rollen som giver eller mottaker som er læringsfremmende, eller eventuelt om det er en kombinasjon av disse som gir læring (Double et al., 2020). Flere studier (Nicol, 2012; Nicol, 2014; Winstone et al. 2017) har dog identifisert kognitive prosesser i hverandrevurdering forbundet med det å gi hverandrevurdering som er læringsfremmende, se punkt 3.2.1.

Det er imidlertid gjort noen få studier som skiller mellom disse rollene. Blant annet gjorde Lundstrom og Baker (2009) en casestudie med 91 andrespråksstudenter i høyere utdanning. I studien skulle studentene gi eller motta skriftlig feedback på tekster som de skrev i kurset. I denne studien ble studentene delt i to grupper, hvor den ene gruppen bare skulle gi feedback,

mens den andre gruppen bare skulle *motta* feedback. Funn fra studien viser at studentene som kun skulle *gi* feedback forbedret egen skriving i større grad enn dem som bare skulle motta feedback.

### **Format**

Double et al. (2020) har i sin metastudie sammenlignet hvordan ulike format for hverandrevurdering påvirker læringseffekten. Når de sammenlignet hverandrevurdering som inneholder *skriftlige komponenter* konkluderer de med at dette ikke utgjør noen signifikant effekt sammenlignet med hverandrevurdering uten skriftlige komponenter.

Hverandrevurdering gjennom *dialog* har en signifikant positiv effekt sammenlignet med hverandrevurdering uten dialog. Bruk av *graderingskriterier* har signifikant positiv effekt på studenter i høyere utdanning, men ikke for elever på lavere skolenivå.

Kombinasjon av både skriftlige kommentarer fra medstudent *og* gradering på skala har vist seg å gi mer forbedring av skriftlige prestasjoner enn bare en av disse vurderingene alene (Huisman et al., 2018). I Cartney (2010) sin casestudie valgte studentene mellom å gjennomføre skriftlig eller muntlig vurdering. Flere av studentene som gjennomførte hverandrevurdering i muntlig format, uttrykte etterpå at de hadde gått glipp av læring knyttet til å kunne uttrykke seg presist i skriftlig form. Studentene som gjennomførte hverandrevurdering muntlig begrunnet derimot dette med at det var fordelaktig, fordi det åpner for å stille spørsmål og oppklare misforståelser. Studentene trekker frem de samme fordelene med muntlig gjennomføring i van den Berg et al. (2006) sin studie som evaluerer syv ulike kursdesign med hverandrevurdering.

Hverandrevurdering kan gjennomføres synkront eller asynkront. Liu et al. (2018) har undersøkt dette i sin casestudie med 18 andreårsstudenter ved et universitet i Kina. I studiet sammenlignes studenter som gjennomfører hverandrevurdering hvor eksperimentgruppen diskuterer vurderingen de har gitt/mottatt fra medstudenter underveis. Studiet viste at det synkrone læringsdesignet hadde en positiv læringseffekt, og trekker frem muligheten til å stille spørsmål til giver av vurdering som et positivt aspekt.

## **Anonymitet**

Double et al. (2020) sin metastudie viser at anonymitet ikke har påvirkning på læringseffekten. I casestudiet til Liu et al. (2018) sammenlignet en grupper som gjennomførte nettbasert hverandrevurdering. Den ene gruppen diskuterte vurderingen i anonyme grupper på nett, mens den andre gruppen hadde samme aktivitet i ikke-anonyme grupper. Også dette studiet konkluderte med at anonymitet ikke utgjorde forskjell.

På den andre siden er det gjort forskning som viser at studenter er redde for å gi kritisk feedback til sine medstudenter (Cartney, 2010; Panadero, 2016; To & Panadero, 2019). Dette blir begrunnet med at studentene ikke ønsker at mottaker av vurderingen skal bli lei seg. Noen studenter er også redde for at en kritisk vurdering skal ha negativ påvirkning på relasjonen de har til mottakeren av vurderingen (Cartney, 2010; To & Panadero, 2019). Det blir også begrunnet med at studentene ikke stoler på sin egen kompetanse, og de frykter å fremstå som inkompetente foran sine medstudenter (Cartney, 2010; To & Panadero, 2019). Et konkurransepreget studiemiljø har negativ innvirkning på læringseffekten av hverandrevurdering, og derfor kan anonymitet være fordelaktige i denne typen studiemiljøer (To & Panadero, 2019).

Rotsaert et al. (2018) gjennomførte en casestudie med 46 studenter hvor de hadde en gradvis overgang fra anonym til ikke-anonym hverandrevurdering. I denne studien ble anonymitet brukt som *scaffolds* før studentene ble mer trygge på å både gi og motta hverandrevurdering. Etter hvert som studentene ble mer trygge i emnet og i å gi hverandrevurdering ble den gjennomført ikke-anonymt (*fading*). Hverandrevurderingene økte i kvalitet etter flere runder. Det viste seg at ikke-anonym feedback fikk like høy kvalitet som de anonyme etter å ha gjort det over tid.

I en casestudie av van den Bos et al. (2019) med 114 universitetsstudenter blir anonym og ikke-anonym hverandrevurdering sammenlignet etter klassifisering i "feedback-typer". Blant annet skilles det mellom det som blir kalt *lower order* og *higher order feedback*. Lower order feedback fokuserer på språklige feil og mangler som for eksempel skrivefeil, struktur og grammatikk. Higher order feedback handler om tekstens innhold derunder blant annet faglig innhold og argumentasjon. Studien viser at de anonyme studentene ga mer higher order feedback enn de ikke-anonyme studentene.



## **Hyppighet**

Double et al. (2020) har undersøkt om hvor mange ganger studenter gjennomfører hverandrevurdering har innvirkning på læringseffekten. I studien sammenlignet de studier hvor hverandrevurdering bare var gjennomført én gang, med studier hvor aktiviteten ble gjennomført flere ganger med samme informanter. Funnene viser at det ikke var signifikant forskjell mellom effekten i studier hvor det bare var gjennomført én gang og studier der det var gjennomført flere ganger. Dette motstrider med funn fra flere casestudier som finner at kvaliteten på feedback som studenter gir blir bedre av øvelse (Rotsaert et al., 2018). Misiejuk et al. (2021) sin forskning viser at studentene ikke tar i bruk vurderingen som de mottar, og at dette skyldes at vurderingen som blir gitt er for lite spesifikk og holder for lav kvalitet. Dette påvirkes av indre forhold hos studenten, som for eksempel ferdigheter i å gi hverandrevurdering. Videre peker Esterhazy (2018) på at studenter er dårlige på å ta ansvar for å bruke feedback som de mottar, og at de trenger trening i nettopp dette.

## **Studentenes ferdighetsnivå**

Det er gjort få studier som undersøker hvorvidt studentenes individuelle ferdighetsnivå påvirker læringseffekten av hverandrevurdering. Li og Gao (2016) undersøkte dette i sin kvasiekperimentelle studie med 150 lærerstudenter i USA. I dette studiet ble studentene klassifisert etter prestasjonene lav, middels og høy. Studentene som ble klassifisert med lavt og middels prestasjonsnivå hadde høyest utbytte av å gjennomføre hverandrevurdering. Studentene som allerede presterte høyt akademisk, fikk ikke utbytte av å gjennomføre hverandrevurdering. Flere andre studier viser det samme (Davies, 2000; Li, 2012). Li og Gao forstår disse funnene i sammenheng med Vygotsky (1978) sin teori den proksimale utviklingssonen, hvor studenter lærer av en signifikant annen, for eksempel studenter som har høyere faglig kompetanse.

I To og Panadero (2019) sin forskning med førsteårsstudenter, uttrykker deres undervisere bekymring for at studentene ikke er modne nok til å gjennomføre hverandrevurdering enda. I denne studien er flere av studentene enige med sine undervisere, og uttrykker bekymring for at de ikke har kompetanse til å gi en god vurdering til sine medstudenter. For å kunne gi en god hverandrevurdering må studentene ha nok kompetanse til å vurdere styrker og svakheter ved andres arbeid (Cartney, 2010; Liu. et al., 2018).

I Alfalay (2004) sin casestudie med 78 språkstudenter i høyere utdanning blir studentenes personlighetstrekk vurdert opp mot kvaliteten til både egen-og hverandrevurderingen som de ga. Studien viser at studentene ofte overvurderer kvaliteten til både eget og medstudenters sitt arbeid, men overvurdering av medstudenters arbeid forekom hyppigst. Studenter med dårlig akademisk selvtillit var mest kritisk til eget arbeid. Når feedback fra lærer og medstudenters ble sammenlignet viser det seg at studentene med lavest selvtillit hadde minst sammenfallende vurdering av medstudenters arbeid, hvor de overvurderte medstudenters sitt arbeid.

### 3.3 Rubrikker

Innenfor vurdering blir *rubrics*, oversatt til *rubrikker*, forstått som et verktøy for vurdering innenfor en bestemt kontekst, for eksempel i forbindelse med vurdering av tekstlige studentoppgaver (Goodrich, 1997). Rubrikker kan forveksles med sjekklister for studentoppgaver, da begge disse innehar *evalueringskriterier* som bestemmer innholdet i oppgaven (Andrade, 2005; Goodrich, 1997). I denne delen gjennomgås studier som har undersøkt bruk av rubrikker innenfor ulike fagfelt i høyere utdanning. Tradisjonelt har rubrikker blitt brukt av lærere i arbeid med vurdering av studentarbeid, men rubrikker er også nyttige som *instruksjonsverktøy* (Andrade, 2005). Eksempelvis kan rubrikker være et nyttig verktøy for studenter som har behov for veiledning i hvordan skrive gode akademiske tekster. Rubrikken kan instruere og støtte studenten i å gjøre pålitelige vurderinger av kvaliteten på eget arbeid. Studier har utforsket om inkludering av rubrikk i undervisningen kan føre til en forbedret vurderingspraksis og en økt prestasjon hos studenter. Flere studier viser at rubrikker *kan støtte både undervisning og læring* (Reddy & Andrade, 2010).

#### **Studenters opplevelse med bruk av rubrikker**

Studier som har undersøkt studenters erfaring med rubrikker, viser at studenter generelt verdsetter rubrikker fordi de beskriver lærerens forventninger, tydeliggjør læringsmål og bidrar til at vurderingskriterier blir mer transparente (Reddy & Andrade, 2010). Andre studier viser at studentene synes rubrikker er nyttige både ved planlegging og underveis i arbeidsprosessen (Reynolds-Keefer, 2010). Andrade og Du (2005) har utført et forskningsprosjekt med 14 lærerstudenter hvor de undersøker studentenes opplevelse med bruk av rubrikker for vurdering og hverandrevurdering. Studien viser at studentene er positive

til å bruke rubrikker. Blant annet finner de at rubrikken bidrar til å avklare lærerens forventninger, identifisere styrker og svakheter i eget arbeid, og at karaktergivningen virker mer rettferdig. Dette fordi studentene vet mer om hva læreren vektlegger i vurderingen når kriteriene blir mer transparente. Studenter rapporterte også at rubrikken påvirker kvaliteten på arbeidet, ved at for eksempel rubrikken gjorde det enklere å skrive gode oppgaver. Den økte kvaliteten begrunnes med at rubrikken klargjorde lærernes forventninger og formatet til oppgaven. Videre fremheves det at flere følte seg mindre nervøs for å levere oppgaver, fordi rubrikken hjalp dem med å overvåke at alle kriterier dekket (Andrade & Du, 2005).

Reynolds-Keefer (2010) har gjort en studie i forlengelse av arbeidet til Andrade og Du (2005), som undersøker rubrikkers påvirkning på lærerstudenters læring. I likhet med Andrade og Du viser denne studien også at rubrikker kan bidra til mer transparente vurderingskriterier. Rubrikken kan på den måten bidra til at noe av den "tause" kunnskapen som eksisterer i akademia, overføres fra fagperson til student. Den tause kunnskapen kan omfatte perspektiver på fagstoff som ikke alltid blir eksplisitt uttalt til studentene. Taus kunnskap kan føre til uklare forventninger og at det oppstår misforståelser mellom lærer og student (Bacchus et al., 2020). Transparente vurderingskriterier er en viktig forutsetning for en vurderingspraksis som sikter mot økt læring (Damsa et al., 2015; Nordrum, et al., 2013).

Reddy og Andrade (2010) har sett nærmere på mulige årsaker til studentenes positive holdning til bruk av rubrikker. En faktor som de trekker frem, er at studentene deltok i å utforme rubrikkene. Dette bidro til å tydeliggjøre vurderingskriteriene overfor studentene. En annen faktor er at rubrikken ble gjort tilgjengelig for studentene *før* de begynner på selve oppgaven. Dette samsvarer med studien til Schneider (2006) som sammenlignet to studentgrupper. Den ene studentgruppen (A) fikk utlevert rubrikker før de begynte med oppgaven, mens den andre studentgruppen (B) fikk rubrikkene når de var ferdig med oppgaven. Undersøkelsen viste at 88% av studentene i gruppe A mente rubrikken var nyttig, mot 10% av studentene i gruppe B (Reddy & Andrade). Disse funnene kan forstås i sammenheng med forskning om formativ vurdering som innebærer at vurdering må gis før det aktuelle arbeidet ferdigstilles (Black & Wiliam, 2009; Hattie & Timperley, 2007).

Empiriske studier har også belyst forholdet mellom studentenes nervøsitet for oppgaver og bruken av rubrikk. Både resultatene til Reynolds-Keefer (2010) og Andrade (2005) viser at

flertallet av studentene opplevde en redusert følelse av nervøsitet når rubrikker var inkludert i oppgaven. For eksempel opplevde studenter en økt kontakt med lærer, og at det var lavere terskel for å stille læreren spørsmål. Noen var mindre nervøs for å levere oppgaven, fordi rubrikken bidro til å tydeliggjøre oppgavens vurderingskriterier. Imidlertid kunne rubrikken for noen føre til et økt nivå av nervøsitet fordi de opplevde at det ble stilt høyere krav for kvaliteten på oppgaven (Reynolds-Keefer, 2010).

### **Bruk av rubrikk og opplevd læringsutbytte**

Forskere har vært interessert i å utforske en eventuell kobling mellom bruk av rubrikk og økt læringsutbytte. Flere forskningsresultater antyder en generelt høyere akademisk prestasjon og dypere læring hos studenter som har brukt rubrikk i deres arbeid, men forskningsresultatene er ikke samstemte (Reddy & Andrade, 2010). En studie som ikke viste noen signifikant forskjell i bruk av rubrikk eller ikke, er studien til Green og Browser (2006). Reddy og Andrade (2010) understreker at det var få deltakere (N = 16) i studiet, og poengterer at å utlevere rubrikker ikke nødvendigvis fører til økt læringsutbytte. De konkluderer med at studenter må engasjeres med rubrikken, som for eksempel å være med å konstruere rubrikken, og å aktivt bruke den i egen- og hverandrevurdering.

### **Bruk av rubrikk som scaffolding i undervisningen**

Studier har undersøkt i hvilken grad bruk av rubrikker kan fungere som et verktøy for scaffolding for studenter i høyere utdanning. I Double et al. (2020) sin studie ble designet for hverandrevurdering klassifisert som *åpent* (freeform) eller som *scaffolding*. Læringsdesign hvor studentene bare fikk grunnleggende instruksjoner ble klassifisert som åpne. Bruk av for eksempel rubrikker eller ulike scoringssystemer, ble klassifisert som scaffolding. I denne studien finner en ingen effekt av å bruke scaffolding sammenlignet med å ikke bruke scaffolding. Dette metastudiet har begrenset betydning når det gjelder dette fenomenet, da det inkluderte få studier med åpne læringsdesign.

Imidlertid tyder studiet til Andrade og Du (2005) på at rubrikker kan støtte opp under en formativ vurderingspraksis. Black og Wiliam (2009) beskriver vurderingsprosessen med utgangspunkt i *tre trinn*. Denne prosessen består av å stadfeste hvor studenten er i læringen sin, hvor de er på vei, og hva som må til for å komme seg til ønsket måloppnåelse. De fleste studentene brukte rubrikker for å muliggjøre de tre trinnene i vurderingsprosessen. Dette viser

rubrikkens potensiale for å fremme selvregulerende atferd, som å sette seg mål, og vurdere og revidere eget arbeid (Andrade & Du). Et eksempel på en slik metakognitiv strategi kan illustreres ved studenter som rapporterte at det de lærte ved å bruke rubrikk i et fag, hadde en overføringsverdi som de benyttet seg av i andre fag. Reynolds-Keefer (2010) støtter påstanden til Andrade og Du (2005) om rubrikkers potensielle betydning som et nyttig verktøy i en formativ vurderingsprosess.

### **Undervisernes opplevelse av bruk av rubrikker**

Mens studiene i artikkelen til Reddy og Andrade (2010) rapporterer om et generelt positivt syn på rubrikk blant studenter, rapporteres det om motstridende holdninger blant underviserne. Av positive funn, rapporteres det om at rubrikken bidro til å gjøre objektive vurderinger og at det var effektivt å bruke. Et interessant funn er forskjellen mellom studentene og undervisernes oppfatning av formålet med rubrikker. Studenter koblet dem med læring og prestasjon, mens underviserne fokuserte nesten utelukkende på rubrikken som et verktøy for å objektivt og nøyaktig tildele karakterer (Reddy & Andrade, 2010). I følge Reddy og Andrade (2010) kan noe av motstanden mot å bruke rubrikker, skyldes undervisernes begrensede forestilling om formålet med rubrikk. Det nevnes også at flere av underviserne i høyere utdanning har begrenset erfaring som lærer, og mindre kunnskap om ulike vurderingsformer.

### **Potensielle begrensninger ved bruk av rubrikk**

Til tross for bred støtte for bruk av kriteriebasert vurdering, eksempelvis rubrikk, finnes det studier som peker på uheldige tendenser som kan oppstå. Blant annet hevdes det i artikkelen til Nordrum et al. (2013) at det mangler en felles forståelse av hvordan kriteriene skal brukes. Dette kan føre til en subjektiv tolkning av den enkelte lærer. I tillegg advares det mot at bruk av kriteriebasert vurdering kan redusere studentens autonomi ved at de får en overdreven tillit til kriteriene. Det advares også mot at for detaljerte kriterier fører til at de brukes mer som en sjekkliste i stedet for et verktøy for læring. Torrence (2007, referert i Nordrum et al., 2013) betegner dette som vurdering *som* læring, i stedet for vurdering *av* eller vurdering *for* læring. Andrade og Du (2005) rapporterte om flere positive funn, men peker også på funn som kan begrense rubrikkens nytteverdi. I deres studie kom det frem at noen studenter bare leste deler av rubrikken og at de så på rubrikken som en oppskrift for å gi lærerne det hen vil ha. Dette er i tråd med studiet til Reynolds-Keefer (2010). De fleste studentene i denne studien rapportere

at de brukte rubrikken som en sjekklister, og noen av studentene sjekket ikke rubrikken før de fikk oppgaven med tilhørende rubrikk i retur.

En mulig fallgrube kan være at lærerne formulerer rubrikken på en måte som ikke er forståelig for studenten, og som dermed fører til at studenten ikke klarer å ta rubrikken i bruk. Bacchus et al. (2020) viser til tidligere forskning de har gjennomført (Colvin et al., 2016). Denne forskningen indikerte at mange studenter ikke synes rubrikken var nyttig fordi det var benyttet et komplekst språk og dermed ble rubrikken for avansert å bruke. I studiet svarte 71 % av studentene at de alltid leste rubrikken, og bare 43 % sa at de vanligvis forstod hva som krevdes av dem etter å ha lest rubrikken (Colvin et al., 2016). Dette understreker betydningen av gode spørsmålsformuleringer i rubrikker.

### **Design av rubrikk og integrering av rubrikk i undervisningen**

Studier har undersøkt hvordan utformingen av rubrikk og integrering av rubrikk i undervisningen kan gjøres best mulig. Ifølge Reddy og Andrade (2010) må en forstå kursdesignet i et helhetlig perspektiv, hvis en skal undersøke om bruk av rubrikk faktisk fører til ønsket effekt eller eventuelt hvorfor ønsket effekt ikke oppnås. Videre har utformingen av rubrikken betydning for læringseffekt.

Studier har også undersøkt effekten av ulike vurderingsverktøy og kombinasjonen av ulike former for kriteriebasert vurdering. Bacchus et al. (2020) har gjennomført en studie som undersøker to strategier for å øke studentenes engasjement i vurderingspraksisen; *bruk av eksempler sammen med rubrikk* og *studentsamarbeid om utforming av rubrikk*. Studiet har gjennomført en nettbasert spørreundersøkelse (N= 105) og fokusgruppeintervju (N=28). Funn fra studien viste at lærerdesignede rubrikker alene ikke gir tilstrekkelig klare og transparente kriterier. Den mest effektive strategien var å bruke eksempler sammen med rubrikk, selv om rubrikker utformet av studenter også hadde noe positiv verdi. Til tross for at studien rapporterer om at bruk av eksempler har positiv effekt, var effekten svakere enn ønsket.

Nordrum et al. (2013) har sammenlignet studenters erfaring med to former for kriteriebasert vurdering; *tekstkommentar* og *feedback i rubrikk*. Studien ble gjennomført med førsteårsstudenter som tok et skrivekurs i engelsk ved et universitet i Sverige. Målet med studien var å belyse hvordan formativ vurderingspraksis kan optimaliseres. Studien viser at

ulike funksjoner ble tilskrevet de to ulike formene for vurdering. Tekstkommentarene fokuserte oftere på språklige feil. Feedback med rubrikk ga derimot et mer helhetlig bilde over presentasjon og innhold. Den var dermed et nyttig verktøy å bruke for fremtidige oppgaver, men ble mindre brukt i revisjonsprosessen. Ved valg mellom de to vurderingsformene, foretrakk majoriteten av studentene å bruke tekstkommentar. Det var derimot flere som mente at det var ønskelig med en kombinasjon av både tekstkommentar og rubrikk. Nordrum et al. (2013) hevder at en kombinasjon av begge kan gi en synergieffekt og optimalisere av vurderingspraksisen.

### **Fremtidig bruk av rubrikk**

Oppsummert rapporterer studier om flere positive funn vedrørende bruk av rubrikk i undervisningen. Samtidig er det fortsatt behov for studier som belyser flere aspekter ved bruk av rubrikker. Flere av studiene som refereres til baserer seg på data hvor studentene selv har rapportert om hvordan de brukte rubrikken, mens det er begrensede empiriske bevis for hvordan studentene faktisk bruker rubrikker til sin fordel (Andrade & Du, 2005; Misiejuk et al., 2021; Nordrum et al., 2013). Dermed er det behov for mer kunnskap om dette for å skape bedre forståelse for hvordan rubrikker kan integreres i et helhetlig læringsdesign på en hensiktsmessig måte.

## **3.4 Læringsdesign med hverandrevurdering**

### **Læringskultur**

For at hverandrevurdering skal være effektivt er det nødvendig med et kulturelt skifte i høyere utdanning hvor feedback og vurdering er en naturlig del av samspillet mellom studentene (Esterhazy & Damsa, 2017). Gjennom utdanningsløpet er mange studenter vant til at vurdering er noe som gis av lærer, og de opplever derfor at det er høy terskel for å delta i hverandrevurdering (Cartney, 2010). Flere skoleforskere (Boud & Molloy, 2013; Esterhazy & Damsa, 2017) tar til orde for at lærernes rolle må endre seg fra å være noen som gir studentene input, til å være “designere” av læringsmiljø og fasilitator av studentaktive læringsformer, se punkt 3.1. Fordi studiemiljøer er forskjellige er Esterhazy (2018) kritisk til å lete etter en ideell løsning som skal passe alle settinger.

## **Organisering**

Studenter er bekymret for å virke inkompetente foran sine medstudenter (Cartney, 2010; To & Panadero, 2019). Videre frykter studenter at de mangler kompetanse for å kunne gi en god hverandrevurdering (Cartney, 2010; To & Panadero, 2019). På bakgrunn av dette finner studier at hverandrevurdering bør foregå i mindre grupper på maksimum fem personer, slik at fallhøyden er mindre for studentene (Cartney, 2010; van den Berg et al., 2006). Van den Berg et al. (2006) finner imidlertid at organisering i grupper der to studenter vurderer hverandre er den organiseringen som er enklest å gjennomføre. Anonymisering inngår også i hvordan hverandrevurderingen organiseres, se punkt 3.2.1. En annen forutsetning for vellykket hverandrevurdering er at deltakerne har kunnskap om det aktuelle fagområdet som de skal vurdere (Nicol, 2012; Nicol et al., 2014).

## **Oppgavedesign**

For å undersøke faktorer som kan påvirke hverandrevurderingen har van Popta et al. (2017) gjennomgått fire relevante studier. Ut fra disse konkluderes det med at oppgaveformulering er en viktig faktor. Oppgavene må være av en slik art at studentene blir oppfordret til å *gi forklaring*, heller enn å gi en vurdering i form av nivåbasert måloppnåelse. Dette henger sammen de kognitive prosessene forbundet med hverandrevurdering (Nicol, 2014; Nicol et al., 2014; Topping, 2005). Videre mener van Popta et al. at oppgavene må være designet på en måte som gir studenter mulighet til å foreslå endringer. Størrelsen på produktet som skal vurderes må ikke være for stor. Dersom det er et skriftlig arbeid som skal vurderes av medstudenter må ikke teksten overskride 5-8 sider (van den Berg et al., 2006).

## **Timing**

I formativ vurdering må vurderingen mottas underveis i arbeidet, slik at mottaker av feedback har muligheten til å revidere arbeidet, se punkt 3.1.1 (Black & William, 2009; Gibbs, 2006; Hattie & Timperley, 2007). Videre er det viktig at vurderingen mottas relativt raskt etter at arbeidet er levert. Jo fortare vurderingen mottas, jo enklere vil studenten oppleve at det er å ta den i bruk (Gibbs, 2006). Studenten må få nok tid på seg til å bruke vurderingen som den mottar fra medstudenten til å kunne gjøre nødvendige endringer før arbeidet skal vurderes av lærer (van den Berg et al., 2006).



### 3.5 Oppsummering av litteraturgjennomgang

*Formativ vurdering* er læringsfremmende på tvers av kontekster og utdanningstrinn, og *feedback* er en viktig forutsetning for læring og utvikling. Innenfor høyere utdanning er det behov for forbedring av vurderingspraksisen som utøves. Dette skyldes delvis at studentene ikke har verktøy for å benytte seg av feedback som de mottar, men også at vurderingskriteriene baserer seg på “taus kunnskap”. Det finnes fortsatt lite forskning om hvordan studenter tar i bruk vurdering som de mottar, og de fleste studier som finnes baserer seg på selvrapporteringsdata. Flere forskere tar til orde for et kulturelt skifte i høyere utdanning hvor feedback får en naturlig plass, og undervisere blir “designere” av læringsmiljø.

Studier viser at feedback gitt av studenter holder like høy kvalitet som feedback gitt av lærer. Likevel finner flere studier at studenter er mer kritiske til å motta feedback fra medstudenter. Læringseffekten av hverandrevurdering knyttes til metakognitive prosesser som gjør at studentene øker sitt eget refleksjonsnivå, og kan dermed forstås som en form for *egenvurdering*. Disse effektene er i hovedsak knyttet til det å gi hverandrevurdering, men det er gjort få studier som skiller mellom disse rollene.

Studier peker på ulike faktorer som påvirker en hverandrevurderingsprosess. Studentene kvier seg for å vurdere medstudenter sitt arbeid, og dette kan skyldes flere faktorer som ikke-anonymitet, bruk av dialog, synkronitet og hyppighet. Studenters *ferdighetsnivå* viser seg å ha betydning for den enkelte student sitt læringsutbytte hvor de lavt- og middels presterende studentene får utbytte av hverandrevurdering, mens de høytpresterende har minst utbytte. Et fåtall av studiene om hverandrevurdering i høyere utdanning, undersøker hverandrevurdering fra både underviser og studenters side. Det ser ut til å være et språk i litteraturen når det gjelder studier som belyser hverandrevurdering fra et institusjonelt perspektiv.

Flere studier har vist at rubrikker *kan* støtte undervisning og læring under riktige forutsetninger. Rubrikker med transparente vurderingskriterier kan være med å bidra til at studentene opplever karakterene de mottar som mer rettferdige. Studier har funnet positive effekter av bruk av rubrikk i undervisningen. Dette forutsetter blant annet at rubrikken er godt utformet og at den brukes aktivt av studentene.

## 4 Metode

I dette kapittelet presenteres valg av metode og forskningsdesign. Videre vil det bli redegjort for prosessen med å utarbeide forskningsinstrumenter, innsamling av data og analyse av datamaterialet. Avslutningsvis reflekteres det over etiske problemstillinger, samt at oppgavens reliabilitet og validitet diskuteres. I denne studien er det gjennomført en *metodetriangulering*, hvor det anvendes både *kvantitative og kvalitative metoder* for å best mulig besvare oppgavens forskningsspørsmål.

### Forskningsdesign - metodetriangulering

I denne studien fant vi det mest hensiktsmessig å gjennomføre det som i litteraturen kalles en metodetriangulering. Det innebærer at en studie baseres på en kombinasjon av forskjellige data og metoder, for å belyse et fenomen fra flere synsvinkler (Grønmo, 2016, s. 67).

Bruk av flere metoder kan bidra til en mer samlet og helhetlig forståelse av fenomenet som undersøkes. “Ved å foreta kvalitative og intensive oppfølginger av kvantitative surveyundersøkelser kan vi kombinere generell oversikt med spesiell innsikt på strategisk viktige områder” (Grønmo, 2016, s. 231). Begrepene kvantitativ og kvalitativ refererer først og fremst til egenskaper ved *dataene* som samles inn og analyseres. Data kan karakteriseres som kvantitative dersom de uttrykkes i form av rene tall eller andre tallmengder. Data som ikke uttrykkes på den måten, er kvalitative (Grønmo, 2016, s. 22). Eksempelvis vil en persons argumentasjon om hvorfor ta i bruk hverandrevurdering i høyere utdanning, være kvalitative data. Metodene som anvendes er *spørreskjema* og *intervju*. Vi startet med å gjennomføre en spørreundersøkelse, og deretter intervjuet et utvalg personer. Det kvalitative intervjuet, fungerte som en oppfølger til spørreskjemaet. På denne måten ga den kvantitative tilnærmingen et oversiktsbilde over generelle tendenser i dataene, mens den kvalitative tilnærmingen åpnet opp for å gå mer i dybden på interessante mønstre i datasettet.

### 4.1 Beskrivelse av casene

Denne studien tar utgangspunkt i to casestudier av to ulike kurs ved et universitet i Norge: (a) førsteårsstudenter ved et bachelorprogram i pedagogikk, og (b) fjerdeårsstudenter ved et masterprogram i pedagogikk. Kursene ble valgt fordi de har arbeidet med hverandrevurdering høsten 2020, som er perioden hvor det ble samlet inn data. Begge kursene var fra før valgt ut av forskningsprosjektet “Formativ vurderingspraksis i Canvas”, se punkt 1.2.

## **Covid-19 og påvirkning på prosjektet**

Vi ble involvert i forskningsprosjektet høsten 2020. Midtveis i semesteret ble all undervisning lagt over til digital undervisning. Førsteårsemnet var planlagt å gjennomføres i digitale flater hele semesteret. Førsteårsstudentene hadde kjent sine medstudenter i svært kort tid da universitetet stengte ned. Slik var dette et godt tidspunkt for å bruke nettopp nettbasert hverandrevurdering. Masteremnet ble overført til digitale flater midtveis i semesteret på grunn av Covid-19. I tillegg til prosjektets forskningsmessige verdi, ble prosjektet også en etterspurt arena for faglig og sosial kontakt mellom studenter som ivaretar smittevern. Videre har universitetet det siste året blitt tvunget til å bruke digitale verktøy for all pedagogisk virksomhet og undervisning, noe som gjør at nettbasert hverandrevurdering fremstår som et svært relevant forskningsområde.

### **4.1.1 Kursdesign og kontekst**

#### *Førsteårsstudenter*

Totalt var 249 studenter påmeldt kurset, og til sammen ni forelesere hadde ansvar for ulike deler av undervisningen. Målet med kurset var å gi en innføring i fagfeltet pedagogikk, samt gi en oversikt over viktige delområder og forskningstradisjoner i pedagogikkfaget.

Kurset varte i 12 uker og integrert i kurset var det forelesninger, seminarer og skriveverksted. Som følge av covid-19 ble alle former for undervisning holdt digitalt. Studentene måtte gjennomføre to skriftlige arbeidskrav for å kunne stille til eksamen. En del av arbeidskravet var å gjøre hverandrevurdering i verktøyet Peergrade. Studentene ble tildelt en tilfeldig valgt medstudents tekst, som de skulle vurdere og gi en skriftlig feedback på. Det ble utdelt rubrikker med bestemte kriterier som studentene skulle bruke til å vurdere teksten. Studentene fikk tilbud om individuell veiledning i forbindelse med arbeidskravet. Kurset ble avsluttet med to hjemmeeksamener.

#### *Masterstudenter*

Masterkurset var et 13 ukers langt kurs, ved det samme universitetet som førsteårskurset. Kurset omfattet 21 fjerdeårsstudenter og ble undervist av én hovedlærer. Kurset handler om pedagogikkfaglige problemstillinger forbundet med danning og demokratisering. Undervisningen ble hovedsakelig gjennomført i form av forelesninger og seminarer. Kurset

hadde én foreleser, mens seminarene hadde 2-3 undervisere samtidig. Av obligatoriske aktiviteter måtte studentene delta i minimum to av tre integreringsseminarer, levere to utkast til skriftlig hjemmeeksamen, og holde et muntlig fremlegg. Kurset ble avsluttet med innlevering av en individuell semesteroppgave (hjemmeeksamen). I starten av semesteret valgte studentene tema for semesteroppgaven. Studentene ble delt inn i “skrivegrupper” hvor de skulle veilede hverandre. I midten av semesteret skrev studentene et sammendrag og en sammenstilling av to pensumtekster som var relevant for den enkelte students oppgave. Studentene ble tildelt en medstudents tekst som de skulle kommentere og gi feedback på. Denne hverandrevurderingen ble gjennomført i den digitale læringsplattformen Canvas.

**Tabell 1.** Beskrivelse av kursdesign

	<b>Førsteårsstudenter</b>	<b>Masterstudenter</b>
Nivå	Bachelor	Master
Studiepoeng	20	10
Antall undervisere	9 forelesere	1 foreleser 3 lærere under seminar
Antall studenter	249 påmeldt i kurs 165 gjennomførte HV	21 påmeldt i kurs 21 gjennomførte HV
Kursperiode	12 uker	13 uker
Feedback	Skriveverksted, individuell veiledning med lærer, hverandrevurdering,	Hverandrevurdering, skrivegrupper, feedback fra lærer på hverandrevurdering
Undervisnings- og læringsaktiviteter	Forelesning Seminar Skriveverksted	Forelesning Seminar Skrivegrupper Muntlig fremlegg

### **Beskrivelse av Peergrade og Canvasrubrikker**

Peergrade er et digitalt verktøy for å fasilitere hverandrevurdering. Det brukes av studenter for å gi og å motta skriftlig feedback på skriftlig arbeid. Integrert i Peergrade finnes

vanlige funksjoner som inngår i konseptet hverandrevurdering. Studenter kan sende inn tekst individuelt eller i gruppe, gi feedback på andres tekster, og motta feedback på egne tekster. Peergrade tilbyr ulike funksjoner som gjør at læreren kan tilpasse hverandrevurderingen til sin klasse. For eksempel kan deltakerne gjøres anonyme, slik at det ikke står hvem studenten har gitt eller mottatt feedback til/fra. Det er rubrikker integrert i verktøyet, som læreren kan bruke til å tilpasse den enkelte oppgave. Se vedlegg 1 for å se spørsmålsrubrikker brukt i førsteårs- og masteremnet. Canvasrubrikker kan brukes for å gjennomføre hverandrevurdering. Canvasrubrikker muliggjør at studenter kan gi vurdering til medstudenter basert på rubrikker som faglærer har utformet. Sammenlignet med Peergrade åpner ikke hverandrevurdering i Canvas for like mange funksjoner når det gjelder individuelle tilpasninger.

#### 4.1.2 Innsamling av data

##### **Utvalg og rekruttering av informanter**

Det er gjennomført en *strategisk utvelging*, der det er rekruttert personer som på ulike måter har tatt del i forskningsprosjektet om hverandrevurdering. Strategisk utvelging er basert på en systematisk utvelgelse av personer som har egenskaper eller kvalifikasjoner som er strategiske for å besvare forskningsspørsmålene (Thagaard, 2018, s. 54). Utvalget i denne studien består av studenter fra to ulike studieretninger, samt ansatte ved instituttet. Det er gjennomført intervjuer med personer som har ulike roller i prosjektet, for å kunne belyse fenomenet *hverandrevurdering* fra ulike perspektiver.

Studentene fikk informasjon om forskningsprosjektet i en forelesning i starten av høstsemesteret. De ble informert om prosjektets formål, og hva det ville innebære for dem å delta i prosjektet. De som ønsket å delta, bekreftet dette i en digital samtykkeerklæring via Nettskjema, som er en selvbetjent skjemaløsning for datainnsamling. I samtykkeerklæringen fikk studentene informasjon om prosjektets formål, hva det ville innebære for dem å delta i prosjektet, samt deres rettigheter og kontaktinformasjon. 55 av 165 førsteårsstudenter som hadde gjennomført hverandrevurdering besvarte spørsmålene i spørreskjemaet, mot 7 av 21 masterstudenter. Studentene som samtykket til å delta i fokusgruppeintervju, ble senere kontaktet per e-post eller telefon for å avtale tidspunkt for intervju.

Rekrutteringen av informanter viste seg å være mer utfordrende enn forespeilet. Studentene ble oppfordret til å besvare spørreskjemaet i forelesningen. Det var imidlertid vanskelig å

følge opp for å hente inn flere svar via digitale kanaler. Informasjonen ble gitt som kunngjøring i Canvas, eller via mail. Vi opplevde at studentene ikke svarte på disse henvendelsene, men kunne ikke vite om det var fordi de ikke ønsket å delta eller fordi de ikke hadde lest våre forespørsler. For å overholde regelen om frivillig deltakelse i forskningsprosjekter, satt dette oss i et etisk dilemma. På den ene siden var det viktig for oss å få bekreftet at studentene hadde mottatt informasjonen fra oss og dermed hadde tatt et aktivt standpunkt om deltakelse i prosjektet. Samtidig ønsket vi ikke å være for pågående, og dermed risikere at studentene følte seg presset til å delta. Førsteårsstudentene ble oppringt av koordinator ved studieprogrammet, mens masterstudentene mottok en tekstmelding med forespørsel om å delta i prosjektet. Ved å kontakte deltakerne via mobiltelefon fikk vi respons og positive svar fra mange nok informanter.

**Tabell 2.** *Beskrivelse av kursdesign*

<b>Datamateriale</b>	<b>Informanter</b>
Fokusgruppeintervju med studenter	Fokusgruppeintervju førsteårsstudenter: (N=6) Fokusgruppeintervju masterstudenter (N=3)
Individuelle intervju med undervisere/ koordinator	Underviser førsteår Koordinator førsteår Underviser master.
Spørreskjema	Førsteår N =55 Master N = 7

#### 4.1.3 Utforming av forskningsinstrumenter: intervjuguide og spørreskjema

I denne studien er forskningsinstrumentene intervjuguider for kvalitative intervju (individuelle og fokusgruppe) og spørreskjema med Likert-skala. Utformingen av forskningsinstrumentene er dels teoridrevet da de er utformet i tråd med Engeströms virksomhetsmodell. Ved hjelp av virksomhetsmodellen (se figur 4) ble det utarbeidet fire dimensjoner som skulle binde forskningens teoretiske rammeverk og forskningsinstrumenter sammen. Dette resulterte i følgende dimensjoner:

- *Forhold subjekter (studenter/ undervisere) - objekt (hverandrevurdering)*
- *Forhold subjekt (undervisere) - subjekt (studenter)*

- *Forhold subjekter (student/undervisere) - artefakter (PG, Canvasrubrikker)*
- *Forhold artefakt (PG, Canvasrubrikker) - objekt (hverandrevurdering)*

### **Utarbeiding av spørreskjema**

Spørreskjemaet er utformet med *Likert-skala*. Hvert spørsmål har en gradert skala med fem svaralternativer, der 1 står for “*i svært liten grad*” og 5 står for “*i svært stor grad*”. Spørsmålene er utformet slik at informantene skal uttrykke sin holdning til et bestemt fenomen (se vedlegg 2). I denne studien er fenomenet hverandrevurdering og det teoretiske rammeverket som er presentert.

### **Utarbeiding av intervjuguide**

I studien ble det benyttet en semistrukturert intervjuguide (se vedlegg 3). Den bestod av forhåndsdefinerte tema og forslag til spørsmål og påstander. Spørsmålene omhandlet informantenes erfaringer, synspunkt og holdninger til bruken av hverandrevurdering i undervisningen. Intervjuguiden var med på å strukturere, og fungerte som en sjekklister under selve intervjuet. Den ble utformet med utgangspunkt i oppgavens tema, med mål om å belyse ulike sider ved forskningsspørsmålene. Intervjuguidene var konstruert slik at flere av de samme spørsmålene ble stilt til ulike informanter. For eksempel ble det brukt tilnærmet identiske intervjuguiden i begge fokusgruppeintervjuene. På denne måten ble intervjuene til en viss grad sammenlignbare i ettertid.

## **4.2 Kvantitativ metode og analytisk strategi**

Kvantitativ metode brukes til innsamling av data i form av tall eller mengder, i motsetning til kvalitative data som uttrykkes tekstlig. Måling og kalkulasjoner er vanlige former for kvantitative data, og disse kan uttrykkes i form av tabeller og grafer. Kvantitativ forskning er basert på strukturerte skjema hvor en finner ut hvilken verdi hver enhet har på ulike variabler (Grønmo, 2016, s. 247-248). Et av måleinstrumentene i kvantitativ metode er spørreskjema som er brukt i denne studien. Gjennom å bruke spørreskjema er det mulig å få frem studentenes holdninger og erfaringer knyttet til hverandrevurdering som de gjennomførte høsten 2020. Svarfordelingen for målingene (spørsmålene) og derunder gjennomsnittsverdi og standardavvik for hvert svar vil bli presentert. Deretter presenteres resultatene av en *bivariat korrelasjonsanalyse* og *multippel regresjonsanalyse* for spørreskjemaet som

førsteårsstudentene har fylt ut (N=55). For å gjennomføre målingene har vi benyttet dataprogrammet *Stata*. Grunnet få svar på spørreskjemaet til masterstudentene (N=7) vil det ikke være hensiktsmessig å gjennomføre statistiske målinger på dette datasettet. For å kvalitetssikre målingene er det også gjennomført reliabilitetsanalyser med Cronbachs alfa.

#### 4.2.1 Operasjonalisering av variabler i spørreskjema

Spørreskjemaet omfatter spørsmål om studentenes opplevelse av hverandrevurdering gjennomført høsten 2020, samt deres syn på fenomenet hverandrevurdering. En variabel er et konsept som kan måles, og som varierer mellom individer og situasjoner (Howitt & Cramer, 2017, s. 23). I analysen er spørsmålene operasjonalisert i fire grupper variabler hvor én av disse er *avhengig variabel*, og de andre tre er *forklaringsvariabler*. Målet er å undersøke hvorvidt og eventuelt hvordan den avhengige variabelen endrer seg som følge av endringer i forklaringsvariablene. En forklaringsvariabel, også omtalt som *uavhengig variabel* er en variabel som benyttes for å forklare virkningen på den avhengige variabelen (Dancey & Reidy, 2007, s. 11- 12). Se vedlegg 4 for oversikt over hvilke spørsmål som inngår i hver variabel.

##### **Avhengig variabel:**

*Opplevd læring*: Denne variabelen tar for seg spørsmål som handler om studentenes læringsutbytte. Derav om de synes det var lærerikt å gi eller få feedback, om hverandrevurdering gjorde besvarelsen deres bedre, og om de synes at hverandrevurdering har en verdi. Videre tar den for seg om hverandrevurdering er nyttig med tanke på eksamen, samt om studentene ønsker at det skal bli obligatorisk med hverandrevurdering våren 2021. Opplevd læring representerer *objektet* i virksomhetsmodellen.

##### **Forklaringsvariabler:**

*Opplevd bruk av artefakter* består av spørsmål som handler om Peergrade/Canvasrubrikker og formulering av spørsmålsrubrikker. Denne gruppen samsvarer med komponenten som omhandler *artefakter* i virksomhetsmodellen.

*Studentenes ferdighetsnivå/ innstilling til hverandrevurdering* består av spørsmål som handler om studentenes opplevelse av egen kompetanse og strategier, feedback som de ga/ mottok,



samt deres innstilling til hverandrevurdering. Denne gruppen representerer virksomhetsmodellens *subjekt*.

*Innramming* består av spørsmål som omhandler informasjonen som ble gitt til studentene og om de fikk opplæring i å bruke det digitale verktøyet. Videre består denne gruppen av spørsmål som omhandler studentenes forståelse av hvorfor en skal gjøre hverandrevurdering, samt om studentene har diskutert hverandrevurderingen med medstudenter eller seminarleder/foreleser. Innrammingen representerer nederste del av virksomhetsmodellen som består av *arbeidsdeling, fellesskap og regler*.

## 4.2.2 Statistiske analysemetoder

### **Deskriptiv statistikk**

De kvantitative dataene vil bli presentert ved gjennomsnitt og standardavvik for hver påstand i spørreskjemaet, og for hver variabel påstandene er operasjonalisert i. Dette innebærer deskriptive data med (1) svarfordelinger for hver påstand og hver variabel, samt (2) gjennomsnittsverdi og standardavvik knyttet til hver påstand og hver variabel. Standardavvik er et mål for spredningen av verdier i et datasett, og sier noe om hvor mye de enkelte verdiene avviker fra gjennomsnittet (Howitt & Cramer, 2017, s. 79).

### **Bivariat korrelasjonsanalyse**

Studien anvender en bivariat korrelasjonsanalyse for å undersøke den statistiske sammenhengen mellom ulike variabler i studien. Korrelasjonsanalysen gir informasjon om det er samsvar mellom to variabler og styrken i denne samvariasjonen (Howitt & Cramer, 2017, s. 105). Korrelasjonskoeffisienten kan være enten positiv eller negativ, avhengig av om begge verdiene øker (positiv korrelasjon), eller om en øker og den andre synker (negativ korrelasjon). At to variabler samvarierer, er likevel ikke tilstrekkelig til at en kan konkludere med at det eksisterer en årsakssammenheng (Johannessen et al., 2010). Korrelasjonsanalyse inkluderes for å undersøke graden av samsvar mellom ulike variabler for fenomenet som studeres. Det anvendes *Spearman's korrelasjonskoeffisient* ( $\rho$ ) for å måle korrelasjonen mellom ulike variabler. Dette er fordi det er brukt Likert-skala, og dermed består datasettet av ordinale data noe som gjør at Spearman's  $\rho$  er best egnet (Dancey & Reidy, 2007). Et positivt tall betyr at det er positiv korrelasjon, mens et negativt tall betyr at det er negativ korrelasjon. I følge Dancey og Reidy (2007, s. 176) indikerer en korrelasjonskoeffisient på 0

ingen korrelasjon,  $\pm 0,1 - 0,3$  svak korrelasjon,  $\pm 0,4 - 0,6$  moderat korrelasjon, og  $\pm 0,7 - 0,9$  sterk korrelasjon. En korrelasjon på  $\pm 1$  indikerer *perfekt* korrelasjon. Denne studien anvender denne skalaen for å måle styrke av korrelasjon.

Perfect	+1	-1
Strong	+0.9	-0.9
	+0.8	-0.8
	+0.7	-0.7
Moderate	+0.6	-0.6
	+0.5	-0.5
	+0.4	-0.4
Weak	+0.3	-0.3
	+0.2	-0.2
	+0.1	-0.1
Zero	0	

*Figur 5. Korrelasjonsstyrke (Dancey & Reidy, 2007, s. 176)*

### Multipel regresjonsanalyse

En multipel regresjonsanalyse gjør det mulig å undersøke om det er en statistisk sammenheng mellom flere variabler og undersøke i hvorvidt og eventuelt i hvilken grad forklaringsvariablene påvirker en avhengig variabel (Dancey & Reidy, 2007).

Regresjonsanalyse kan dermed gi et mer presist og fullstendig bilde av datasettet, fordi den sammenligner de ulike forklaringsvariablene og dermed kan fortelle oss hvilke(n) av disse som påvirker den avhengige variabelen i størst grad. Formålet med å gjennomføre en multipel regresjonsanalyse i denne studien er å sammenligne de ulike forklaringsvariablenes påvirkning på den avhengige variabelen *Opplevd læring*.

### 4.3 Kvalitativ metode og analytisk strategi

Kvalitativ metode er en paraplybetegnelse som rommer flere ulike forskningsmetoder. En av dem er det *kvalitative forskningsintervjuet*, som er benyttet i denne studien. En sentral målsetting med kvalitative tilnærminger er søken etter å oppnå en forståelse av sosiale

fenomener (Thagaard, 2018, s. 11). I denne studien er hensikten med det kvalitative forskningsintervjuet å undersøke undervisernes og studentenes opplevelse av og deres refleksjoner rundt fenomenet hverandrevurdering. Herunder undersøkes studentenes rolle som både giver og mottaker av vurdering. Dette kan bidra til en dypere forståelse av funn fra kvantitative data.

#### 4.3.1 Det kvalitative forskningsintervjuet

Det kvalitative forskningsintervjuet søker å forstå verden sett fra intervjupersonenes side (Kvale & Brinkmann, 2012, s. 21). Det er gjennomført tre *individuelle intervju* med to undervisere og en koordinator, og *fokusgruppeintervju* med to ulike studentgrupper. Totalt ble 13 stykker intervjuet. Alle intervjuene var konstruert som *semistrukturerte intervju* (Kvale & Brinkmann, 2012, s. 47). Som konsekvens av covid-19 ble intervjuene gjennomført via digitale videomøter i Zoom. Under intervjuet hadde alle deltakerne kamera på, og intervjuene ble tatt opp ved bruk Nettskjema-diktafon. I ettertid ble lydopptakene transkribert.

##### *Fokusgruppeintervju med studenter*

Det er gjennomført to fokusgruppeintervju med to studentgrupper. Det ene intervjuet var med førsteårsstudenter (N = 6), mens det andre intervjuet var med masterstudenter (N = 3). Intervjuene ble gjennomført høsten 2020, etter at studentene var ferdig med arbeidskrav tilknyttet hverandrevurderingen. Varigheten av intervjuene var fra 60 til 80 minutter. Det ene intervjuet startet med fire studenter, men på grunn av tekniske problemer forlot den ene informanten intervjuet midtveis. Det ble dermed bare tre studenter fra masteremnet som deltok i fokusgruppeintervjuet.

##### *Intervju med undervisere og koordinator*

Det er gjennomført individuelle intervju med to undervisere, og én koordinator tilknyttet prosjektet. Intervjuene ble gjennomført høsten 2020 eller på nyåret 2021. Varighet av intervjuene var fra 30-70 minutter.

#### 4.3.2 Analyse og fortolkning: tematisk analyse av intervju

Det er anvendt tematisk analyse som metode for å analysere det kvalitative datamaterialet. Tematisk analyse består av forskjellige tilnærminger, men består likevel av noen felles karakteristikk. Det er en fleksibel metode som ofte anvendes i kvalitativ forskning, som

brukes for å identifisere, analysere og rapportere mønstre i et datasett (Braun & Clarke, 2006). Vi har valgt å benytte tematisk analyse beskrevet av Virginia Braun & Victoria Clarke, som karakteriseres som en *refleksiv tilnærming* til tematisk analyse. Den refleksive tilnærmingen fremhever viktigheten av forskerens subjektivitet som analytisk ressurs og forskeren som en aktiv, engasjert aktør i analysearbeidet.

I det følgende gjennomgås de seks fasene i analysemodellen presentert av Braun og Clarke (2006). En detaljert beskrivelse av analysearbeidet kan bidra til å synliggjøre valg og vurderinger som er gjort underveis, og slik bidra til et mer transparent forskningsarbeid. Selv om analysen her presenteres i seks faser, understrekes det at det har vært en dynamisk prosess, der vi har beveget oss frem og tilbake mellom de ulike fasene.

#### *Fase 1: Bli kjent med datamaterialet*

Første del av analyseprosessen handler om å bli godt kjent med datasettet. I denne studien innebar det å gjennomføre detaljerte transkriberinger av intervjuene, lese over transkriberingene flere ganger og notere ned underveis når det dukket opp interessante aspekt, som det var interessant å undersøke nærmere. Dette arbeidet legger grunnlaget for resten av prosessen, så vi var opptatt av å bruke god tid på å bli kjent med datasettet, og ikke forhaste dette arbeidet. Gjennom hele prosessen ble det brukt en refleksjonslogg hvor vi skrev ned våre refleksjoner og valg vi hadde tatt underveis i analysen.

#### *Fase 2: Generere koder*

Neste fase i prosessen innebar å produsere koder fra datamaterialet. I refleksiv tematisk analyse blir en kode konseptualisert som en *analytisk enhet* som fanger en observasjon, og som vanligvis bare belyser en del av datasettet, brukt av forskeren til å utvikle temaer (Braun & Clarke, 2020). En kode kan bestå av ett eller flere ord. I denne analysen består kodene for det meste av korte setninger som er valgt ut for å fremheve meningsinnholdet på en best mulig måte. Vi plasserer oss mellom en *induktiv* og *deduktiv* tilnærming til analysearbeidet. Den tematiske analysen er til dels drevet av studiens teoretiske perspektiv og analytiske interesse. Intervjuguiden er utformet med utgangspunkt i teoriperspektivet, og konseptualisering av koder er gjort med utgangspunkt i prosjektets forskningsspørsmål. Samtidig har vi utført en induktiv tilnærming ved at vi ikke hadde forhåndsvalgte koder, og at vi i analysen var åpne for fenomener som ikke stod beskrevet i teori. En refleksiv tilnærming

til tematisk analyse karakteriseres ved at forskerne har en aktiv rolle i å identifisere mønstre og velge ut det som er av interesse for å belyse forskningsspørsmålene (Braun & Clarke, 2020). Kodene ble generert underveis som vi ble kjent med datamaterialet. De ble skrevet manuelt, og organisert i en tabell sammen med det transkriberte materialet (se vedlegg 5).

### *Fase 3: Søke etter tema*

Utvikling av temaer er en dynamisk prosess, som blir gjort med utgangspunkt i kodingen. Et tema defineres som “mønstre med felles mening, samlet av et sentralt konsept eller idé (Braun & Clarke, 2020, s. 14). I motsetning til koder, er temaer mer mangfoldig, og belyser flere observasjoner, og deler av datasettet. Etter å ha kodet datamaterialet, kom vi frem til temaer som kunne dekke de ulike kodene. For å få en oversikt over relasjonen mellom koder, undertema og tema konstruerte vi et digitalt tankekart. Her brukte vi fargekoder for å skille utsagn fra de ulike intervjuene fra hverandre. Slik kunne en også lettere se og sammenligne utsagn fra ulike intervjuer.

### *Fase 4: Gjennomgang av temaer*

Utforming av tankekartet og utviklingen av tema var en dynamisk prosess der det ble gjort flere revideringer underveis. Noen tema ble ekskludert, mens andre ble slått sammen. For å sikre at dette arbeidet ble gjort riktig, leste vi over det transkriberte materialet på nytt slik at ikke viktige mønstre i datasettet gikk tapt i prosessen.

### *Fase 5: Definere og navngi temaer*

Vi kom frem til fem tema med tilhørende undertema som vi syntes var godt dekkende for datamaterialet: kursdesign, artefakter, å gi hverandrevurdering, å motta hverandrevurdering og opplevd formål med hverandrevurdering. *Kursdesign* ble valgt som tema fordi både studenter og undervisere var opptatt av innrammingen av aktiviteten, og hadde mange meninger vedrørende dette. *Artefakter* ble valgt som tema fordi bruken av spørsmålsrubrikk og valg av teknologi (Peergrade/ Canvasrubrikk) viste seg å være viktige faktorer som påvirket studentenes opplevelse av å gjøre hverandrevurdering. Videre valgte vi å skille det å gjøre hverandrevurdering i to deler, å gi hverandrevurdering og å motta hverandrevurdering. Dette gjorde vi fordi hverandrevurdering består av to ulike komponenter, samt at informantene hadde ulike opplevelser vedrørende det å gi og å få feedback. Det ble dermed naturlig å analysere de to komponentene hver for seg. Det siste temaet, *opplevd formål med*

*hverandrevurdering* beskriver de ulike informantenes opplevde formål med å gjøre hverandrevurdering.

#### *Fase 6: Rapportere*

Siste fasen består av å skrive den endelige analysen og en utfyllende rapport, som i dette tilfellet blir å skrive masteroppgaven. I de to neste kapitlene vil de ulike temaene gjennomgås og vises med utdrag og eksempler fra datasettet. Deretter drøftes denne studiens funn opp mot funn fra tidligere forskning.

## 4.4 Validitet og reliabilitet

For å undersøke studiens generaliserbarhet og gyldighet vil dataenes *reliabilitet* og *validitet* analyseres. Reliabilitet refererer til datamaterialets *pålitelighet*, og kommer til uttrykk ved at en får lignende resultater dersom en bruker samme fremgangsmåte ved ulike datainnsamlinger om det samme fenomenet. Validitet, eller *gyldighet*, brukes om i hvilken grad man ut fra dataene kan trekke gyldige slutninger om fenomenet man undersøker (Grønmo, 2016, s. 240 - 242). I det følgende det bli redegjort for ulike mål for validiteten og reliabiliteten tilknyttet denne studien.

### 4.4.1 Reliabilitet og validitet ved kvantitativ metode

#### **Reliabilitet i kvantitative data**

I denne studien er de kvantitative dataene samlet inn gjennom spørreskjema med *Likert-skala*. For å sikre at disse dataene har høy reliabilitet er det gjennomført en analyse av indre konsistens med *Cronbachs alfa* ( $\alpha$ ). Indre konsistens gir uttrykk for reliabilitet og brukes om evnen til stabilitet og konsistens. Indre konsistens brukes om samsvaret mellom ulike testledd som sammen skal gjenspeile et bestemt fenomen (Grønmo, 2016). Ved å bruke cronbachs alfa kan en måle den statistiske sammenhengen mellom spørsmålene i spørreskjemaet som måler de samme fenomenene. Cronbachs alfa måler ikke reliabiliteten til spørreskjemaet, men kan derimot konstatere om skårene for testen i et spesifikt utvalg har høy reliabilitet (Grønmo, 2016, s. 247 - 248). En Cronbachs alfa-verdi på 1 regnes som perfekt indre konsistens i skårene for det aktuelle utvalget. En alfa-verdi høyere enn 0,6 regnes som akseptabel indre konsistens, mens en verdi over 0,8 regnes som god. Er verdien høyere enn 0,9 kan dette

imidlertid være et tegn på at spørreskjemaet har for mange lignende og dermed overflødige spørsmål (Ursachi et al., 2015). Det påpekes likevel at det er lite sannsynlig at overflødige spørsmål påvirker analysene negativt, og setter en tentativ øvre grense på 0,95. Lengden på spørreskjemaet, altså antall spørsmål per skala, vil påvirke alfa-verdien. Skalaer bestående av få spørsmål gir generelt lavere alfa-verdier enn lengre skalaer (Ursachi et al., 2015).

I analysen er flere spørsmål slått sammen til skalaer, se vedlegg 4. Skalaer er variabler som er konstruert etter spørsmål som måler det samme. Grupperingen som er brukt for å beregne indre konsistens samsvarer med grupperingen av forklaringsvariabler i korrelasjonsanalysen. Variabelen *artefakter* fikk en Cronbachs alfa-verdi som er lavere enn anbefalt, men som vi likevel velger å bruke da skalaen til variabelen bare inneholder tre spørsmål (Ursachi et al., 2015).

**Tabell 3.** Cronbachs alfa for variablene

Variabel	Antall spørsmål	Cronbachs alpha	Indre konsistens
Bruk av artefakter	3	Førsteår: 0,49 Master: 0,47	Tilstrekkelig
Ferdighetsnivå /innstilling til hverandrevurdering	7	Førsteår: 0,79 Master: 0,79	God
Opplevd læring	7 førsteår 6 master	Førsteår: 0,89 Master: 0,85	God
Innramming	7 førsteår 6 master	Førsteår: 0,79 Master: 0,75	God

### Validitet i kvantitative data

Skalaene har fått tilstrekkelige Cronbachs alfa-verdier, men en kan ikke være sikker på at disse resultatene er gyldige uten å undersøke hvorvidt spørsmålene som stilles måler fenomenet som en ønsker å måle. *Innholdsvaliditet* er et mål på om hvert begrep som blir brukt er operasjonalisert og definert på en måte som er treffende for den teoretiske

betydningen av begrepet (Grønmo, 2016, s. 252-253). Utformingen av spørsmål til spørreskjemaer og intervjuguider dekker ulike dimensjoner av fenomenet som undersøkes. For eksempel ble det inkludert spørsmål som belyste ulike element som inngår i en hverandrevurdering. Det var viktig å unngå bruk av ukjente fagbegreper i spørsmålene, slik at informantene ikke skulle være usikker på betydningen av enkelte begrep. For å sikre lik forståelse blant informantene, ble det forklart hva en rubrikk er, samt lagt ved skjerm bilde.

Også undersøkt studiens *interne validitet* og *ytre validitet* er undersøkt. Intern validitet uttrykker om undersøkelsen er gjennomført på en tilfredsstillende måte slik at konklusjonene om årsakssammenhenger er gyldig innenfor innrammingen som undersøkelsen inngår i (Grønmo, 2016, s. 253 - 254). Informantene har fått samme undervisning og deltatt i samme forløp når det gjelder hverandrevurdering i sitt aktuelle emne, hvilket styrker undersøkelsens interne validitet. Med dette skal informantene ha samme forutsetninger. På den andre siden kan det være en svakhet at spørreskjema er utført digitalt, da dette trolig hever terskelen for å stille spørsmål underveis dersom noe er uklart. Ytre validitet er et uttrykk for om resultatene i undersøkelsen kan generaliseres til naturlige situasjoner. Altså at de samme fenomenene vil vise seg også utenfor innrammingen som undersøkelsen inngår i (Grønmo, 2016, s. 254). En svakhet ved denne undersøkelsen er at det var få informanter og spesielt få masterstudenter som fylte ut spørreskjemaet (førsteårsstudenter N = 57, masterstudenter N = 7). Dette gjør at studien har begrenset verdi når det gjelder overføring til andre populasjoner.

#### 4.4.2 Reliabilitet og validitet ved kvalitativ metode

##### **Reliabilitet i kvalitative data**

I kvalitative studier er det som regel ikke mulig å teste og beregne reliabiliteten ved hjelp av standardiserte metoder som brukes ved kvantitative studier. Det er likevel gjennomført en reliabilitetsvurdering ved at det er foretatt en kritisk gjennomgang av spesifikke deler av datamaterialet på ulike tidspunkt. Vi har lest over transkriberingen gjentatte ganger på ulike tidspunkt, samt gjennomgått kodingsarbeidet ved ulike tidspunkt. Hensikten med dette er å avklare *stabiliteten* i våre beskrivelser av fenomenet som undersøkes (Grønmo, 2016, s. 249).

En annen måte å vurdere reliabiliteten i kvalitative studier er i form av *ekvivalens*, som er å sammenligne ulike forskeres beskrivelser av de samme forholdene (Grønmo, 2016, s. 250).



Siden vi har vært to personer involvert i forskningsarbeidet, er dette en egnet metode for å måle studiens reliabilitet. Vi har foretatt koding av det samme datamaterialet, og lest gjennom hverandres transkripsjoner for å sikre lik beskrivelse av de samme forholdene. Ekvivalensen vurderes ut fra samsvaret mellom forskernes beskrivelser. Vi opplever at vår beskrivelse av dataene er tilnærmet identiske og har derfor høy ekvivalens.

### **Validitet i kvalitative data**

*Kompetansevaliditet* referer til forskerens kompetanse for innsamling av data i et bestemt forskningsfelt. Spesielt ved kvalitative studier er forskerens kompetanse viktig, fordi innsamlingen av kvalitative data gjennomføres på en mer fleksibel måte som stiller høyere krav til forskerens erfaring. Vi som masterstudenter har begrenset erfaring og kompetanse med denne typen datainnsamling. Vår begrensede erfaring kan være en faktor som påvirker studiens validitet. Imidlertid fikk vi god støtte og veiledning fra mer erfarne forskere via deltagelse i prosjektet “Formativ vurderingspraksis i Canvas”. Kompetansevaliditet refererer ikke bare til den metodiske gjennomføringen av datainnsamlingen, men også om forskerens teoretiske forståelse av feltet som undersøkes. En grundig gjennomgang av den spesifikke forskningsdiskursen har bidratt til at vi har utviklet en god teoretisk forståelse av fagfeltet som studeres. Dette styrker datainnsamlingen og gyldigheten av våre tolkninger.

*Analytisk generalisering* handler om *ytre validitet* i kvalitativ forskning. Kvantitative studier søker å gjøre det som kalles statistiske generaliseringer, mens casestudier søker å gjøre analytiske generaliseringer. I analytisk generalisering forsøker forskeren å generalisere et bestemt resultat, samt undersøke om resultatet samsvarer med en mer overordnet teori (Yin, 2009). For at en skal kunne si noe om hvorvidt funn fra en enkel casestudie samsvarer med funn fra tidligere forskning, må teorien testes gjentatte ganger. Hensikten med denne studien er først og fremst å undersøke fenomenet hverandrevurdering i en bestemt kontekst. Yin (2009) understreker at forskere innenfor det kvalitative forskningsfeltet bør søke etter å oppnå det som kalles for analytisk generalisering. En forutsetning for analytisk generalisering er at det foreligger rikholdige kontekstuelle beskrivelser. Vi har dermed forsøkt å gi en detaljert beskrivelse av konteksten til de to casene, samt innsamling og analyse av data.

## 4.5 Etikk og personvern

Vitenskapelig forskning stiller strenge krav til etiske retningslinjer. Reliabilitet og validitet avhenger av at disse etiske retningslinjene overholdes. Retningslinjene dreier seg om personvern, samtykke, datahåndtering og sikkerhet (Thagaard, 2018, s. 20-27). Det vil her redegjøres for hvordan denne studien har holdt seg innenfor disse retningslinjene, samt drøfte etiske problemstillinger som oppsto underveis.

For å kvalitetssikre at forskningsetiske retningslinjer overholdes er beskrivelse av prosjektet 'Formativ vurderingspraksis i Canvas' sendt inn for godkjenning av Norsk senter for forskningsdata (NSD). Prosjektet er godkjent av NSD den 09.09.2020, se vedlegg 6. Siden vår masteroppgave er knyttet til dette prosjektet, er den skrevet inn og godkjent i samme søknad.

### **Personvern, informert samtykke, datahåndtering og sikkerhet**

I dette prosjektet er det ikke samlet inn sensitive data, men vi har lagret informantenes fulle navn og studentmail. Dette for å kunne kontakte studentene som ønsket å delta i intervju. Deltakerne i en studie skal informeres om forskningsprosjektet og dets formål før de eventuelt avgir et *informert samtykke* til å delta (Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH), 2018). Dette står beskrevet i samtykkeskjemaet (se vedlegg 7) som studentene mottok. I samtykkeskjemaet står det beskrevet hvordan persondata vil håndteres, hvordan sikre informantenes anonymitet og betydningen av et frivillig samtykke. I dette prosjektet var det ekstra viktig å presisere betydningen av frivillighet, og å klargjøre at dette ikke ville ha betydning for studentenes videre studieløp. Det står også opplyst om at de kan trekke seg fra studien når som helst dersom de ønsker det.

NSD stiller krav til å informere informantene om hvordan forskningsdataene skal håndteres. Informantene har samtykket til at deres persondata håndteres frem til prosjektet avsluttes. Samtykkeskjemaene som informantene har utfyllt inneholder navn og studentmail, og disse er samlet inn og ligger lagret i Nettskjema. Lydopptak av intervjuer er tatt opp med Nettskjema-diktafon. Disse har blitt samlet inn og lagret i Nettskjema, og deretter overført til UiO sin interne Onedrive. Denne OneDrive-en åpner for lagring av gule data som ikke er klassifisert som åpne eller fortrolige. Avgjørelsen er tatt av sikkerhetshensyn og i henhold til norsk

personvernlovgivning. Disse dataene har deltakerne i prosjektet tilgang til - totalt fem personer. Etter prosjektets avslutning vil disse dataene bli slettet fra Nettskjema og OneDrive.

#### 4.5.1 Ethiske problemstillinger

##### **Parallele roller som underviser og forsker**

En av deltakerne i prosjektet hadde rollen som både underviser for førsteårsstudentene, emneansvarlig og forsker i 'Formativ vurderingspraksis i Canvas'. Fokusgruppeintervjuet med førsteårsstudentene ble ledet av den aktuelle underviseren. Når forskeren har parallelle roller følger et ansvar om å fremtre med klarhet overfor forskningsdeltakere (NESH, 2018). Som følge av dette klargjorde forskeren at informantenes uttalelser i intervjuet ikke ville ha noen påvirkning på eksamen eller andre akademiske prestasjoner. Videre ble det presisert at hen ikke ville "ta seg nær" av kritikk som kunne oppstå i intervjuet. En av informantene kommenterte at forskeren hadde parallelle roller, hvor forskeren gjentok sine presiseringer. Ut fra uttalelsene til informantene virket de ikke å være særlig påvirket av de parallelle rollene til den aktuelle forskeren. De pratet tilsynelatende fritt og ofte kritisk om både undervisning og kursdesign. Med dette tror vi at dataene som fremkommer ikke i særlig stor grad er preget av forskerens parallelle roller. Dette kan vi imidlertid ikke være helt sikre på, og i ettertid tenker vi at det hadde vært mer ryddig om en annen deltaker i prosjektet hadde gjennomført intervjuet.

##### **Godtgjørelse til informanter**

Studentene som deltok i fokusgruppeintervjuet mottok godtgjørelse for sin deltakelse i form av gavekort. Selv om det er en omdiskutert praksis, er betaling av informanter noe som har vært praktisert i forskningen i mange år. Betaling av informanter kan være nødvendig for rekruttering til forskningsprosjekter. Det kan oppfattes som et vederlag for informantens tidsbruk, som et lokkemiddel og som en takk for innsatsen. Det er godt dokumentert at betaling øker deltakelsen (Stunkel & Grady, 2011, referert i NESH, 2014). Likevel oppgir også informanter andre grunner til å stille seg til disposisjon, som blant annet et ønske om å bidra til forskning. Det finnes også grupper som støtter formålet med forskningen, men som er usikre på om deltakelsen er verdt bryet. I slike tilfeller trenger ikke betaling å være problematisk (Draper et al., 2009, referert i NESH, 2014). I denne studien ble informantene tilbudt et gavekort for at det skulle være verdt bryet med å stille opp i intervjuet.

Slik vi oppfatter retningslinjene til NESH er ikke betaling av informanter nødvendigvis et forskningsetisk problem. Til tross for at studentene har mottatt et gavekort tror vi ikke at forskningsdataene er påvirket av godtgjørelsen i særlig stor grad. 11 masterstudenter fikk en tekstmelding med invitasjon om å delta i et intervju og opplysning om gavekort. Av disse var det bare fire personer som ønsket å delta. Forskningsintervjuet fant sted bare fire dager før masterstudentene skulle eksamen. Dette rettferdiggjør en kompensasjon for deltakelsen. Ingen av studentene etterspurte gavekortet når intervjuet var avsluttet. Samlet sett forteller dette oss at studentene ikke var motivert av gavekortet i særlig stor grad.

## 5 Presentasjon av funn

I dette kapitlet presenteres, tolkes og analyseres data. Det vil først bli presentert funn fra kvantitative data, og deretter funn fra kvalitative data. I det følgende vil både begrepene *tilbakemelding* og *feedback* benyttes. Dette fordi begrepet tilbakemelding er brukt i spørreskjemaet, samt det begrepet som benyttes av informantene i intervjuene.

### 5.1 Kvantitative funn

Her presenteres deskriptiv statistikk. Deretter vil funn gjort av statistiske analyser bli presentert. Det kvantitative datamaterialet er av mindre størrelse (førsteårsstudenter N=55, masterstudenter N=7). Med det sagt vil de kvantitative dataene bidra til å presentere et helhetlig bilde av fenomenet hverandrevurdering og studentenes opplevde læring og erfaring knyttet til dette i de to aktuelle casene.

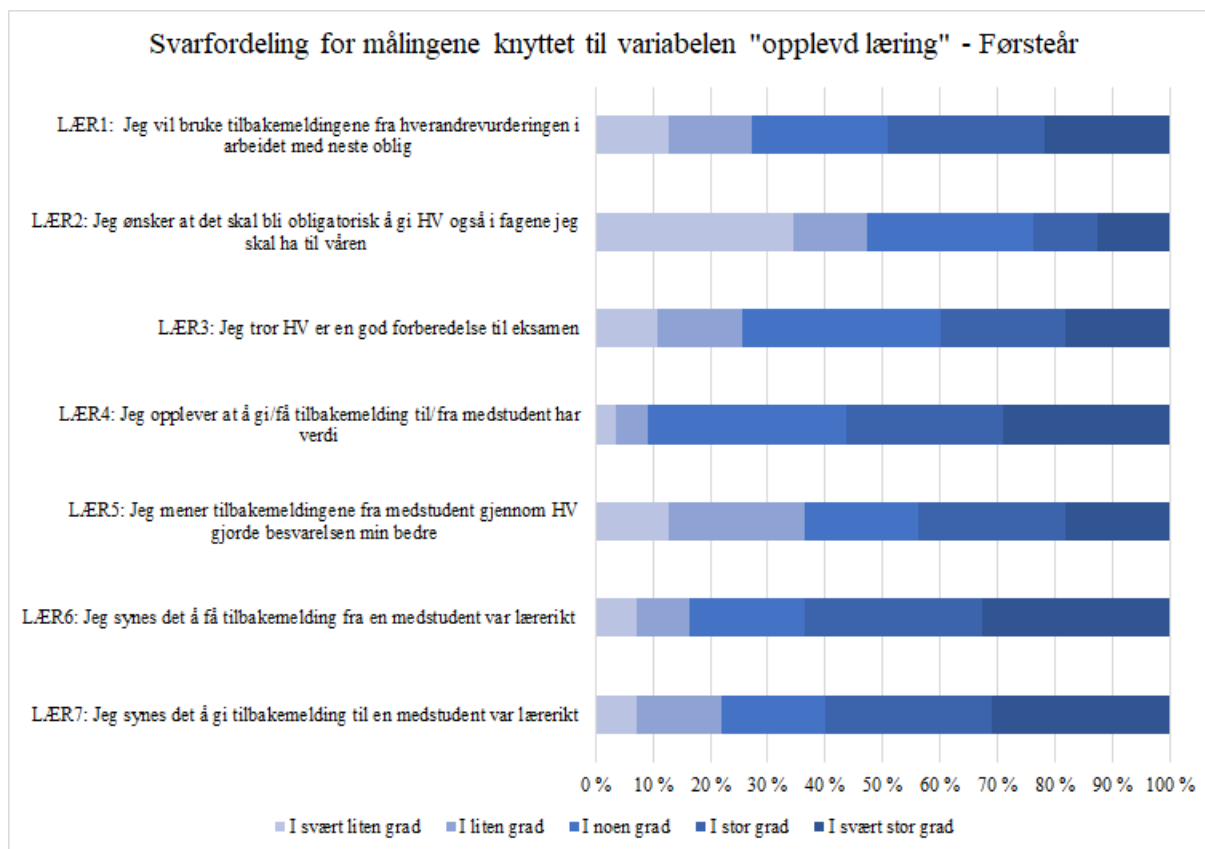
#### 5.1.1 Deskriptiv statistikk

Resultatene er kategorisert etter følgende fire variabler: *opplevd læring*, *innramming*, *artefakter* og *studentenes ferdighetsnivå/innstilling*, se punkt 4.2.1.

##### 5.1.1.1 Kvantitative funn knyttet til variabelen ‘opplevd læring’

Det er totalt syv av påstandene i spørreskjemaet som undersøker studentenes opplevelse av læring i forbindelse med hverandrevurdering.

## Førsteårsstudenter



Figur 6. Svarfordeling for målinger av variabelen opplevd læring for førsteårsstudenter

Tabell 4. Gjennomsnitt og standardavvik for variabelen opplevd læring for førsteårsstudenter

	LÆR1	LÆR2	LÆR3	LÆR4	LÆR5	LÆR6	LÆR7	Gj.snitt
Gj.snitt	3,31	2,55	3,22	3,73	3,13	3,72	3,62	3,33
St.avvik	1,32	1,40	1,22	1,06	1,32	1,22	1,27	1,26

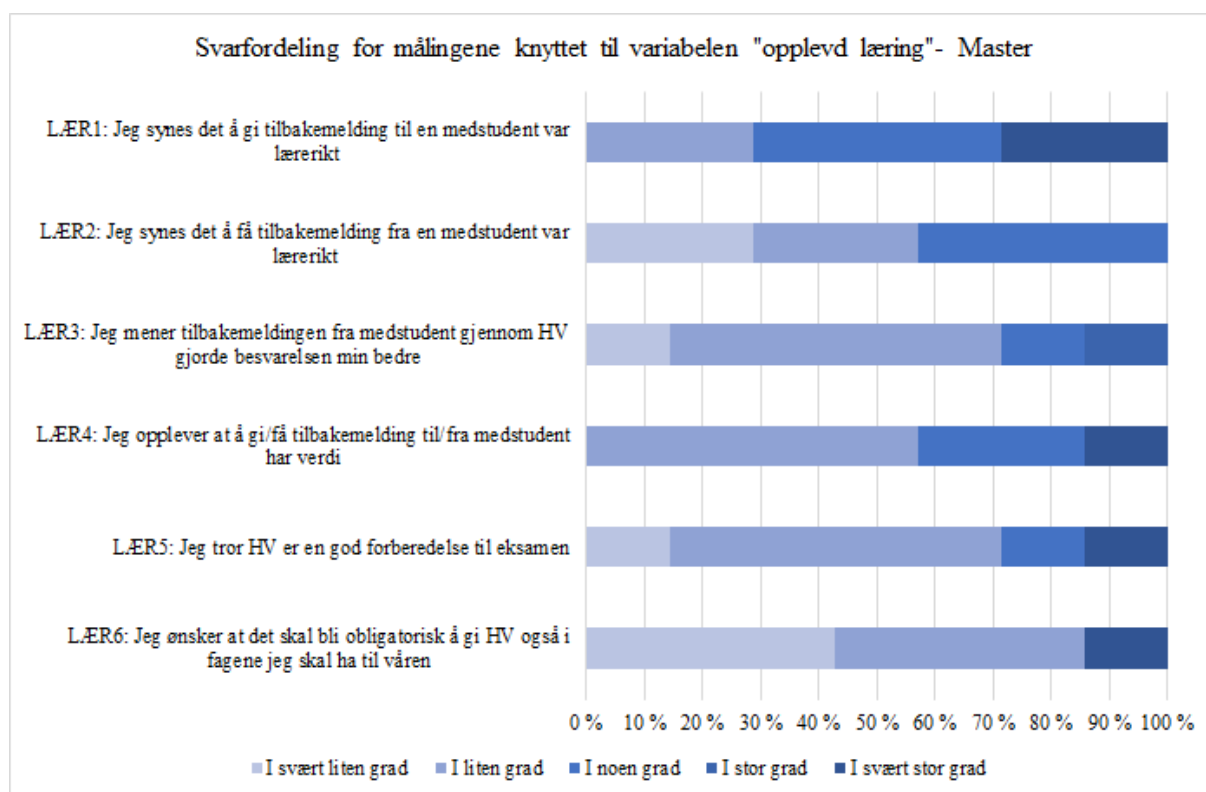
Figur 6 viser at ved påstander som omhandler *opplevd læring* har flertallet av studentene gitt høye verdier, noe som indikerer at studentene er mer enig enn uenig i påstandene. Dette kan en også se i tabell 4 med oversikt over gjennomsnittsverdier og standardavvik.

Gjennomsnittsverdier for alle påstander, utenom påstand 2, har en verdi over 3. Altså er førsteårsstudentene i snitt mer enig enn uenig i påstandene knyttet til opplevd læring. Påstand 4, 6 og 7 har høyest gjennomsnittsverdi. Dette er påstander som måler studentenes opplevde

verdi av å gi eller få tilbakemelding, samt i hvilken grad studentene opplevde at å gi eller få tilbakemelding var lærerikt. Svarfordelingen for påstand 6 og 7 viser ganske like resultat. 60 % av studentene svarte at de *i stor grad* eller *i svært stor grad* opplevde at å gi tilbakemelding var lærerikt. 64 % av studentene svarte at de *i stor grad* eller *i svært stor grad* opplevde at å få tilbakemelding er lærerikt. Dette kan tyde på at studentene synes at det å *gi* og å *få* tilbakemelding opplevdes omtrent like lærerikt.

Påstand 4 skiller seg ut ved at over 90 % av studentene svarer fra *i noen grad* til *i svært stor grad* at det å gi/få tilbakemelding fra medstudent har en verdi. Samtidig ønsker få studenter at hverandrevurdering skal være obligatorisk til våren. Bare omtrent 24 % har svart at de *i stor grad* eller *i svært stor grad* ønsker at det skal bli obligatorisk å gi hverandrevurdering til våren. Selv om studentene syntes at hverandrevurdering var lærerikt, ønsker de ikke nødvendigvis flere obligatoriske arbeidskrav.

Standardavvikene for de ulike påstandene, samt gjennomsnittlig standardavvik for variabelen *opplevd læring* viser gjennomgående relativt høye standardavvik. Påstand 4 som måler om studentene opplever at å gi/få tilbakemelding har en verdi, har lavest standardavvik (1,06).  
*Masterstudenter*



Figur 7. Svarfordeling for målinger av variabelen *opplevd læring* for masterstudenter

**Tabell 5.** Gjennomsnitt og standardavvik for variabelen *opplevd læring* for masterstudenter

	LÆR1	LÆR2	LÆR3	LÆR4	LÆR5	LÆR6	Gj.snitt
Gj.snitt	3,29	2,14	2,29	2,71	2,43	2	2,48
St.avvik	1,25	0,9	0,95	1,11	1,27	1,41	1,49

Svarfordelingen for masterstudentene holder seg mer på venstresiden av skalaen, noe som indikerer lavere grad av enighet i påstandene. Dette kan en også se i tabell 5.

Gjennomsnittsverdien for variabelen *opplevd læring* er lavere for masterstudentene (2,48) sammenlignet med førsteårsstudentene (3,33). I likhet med førsteårsstudentene ønsker få masterstudenter at det skal bli obligatorisk med hverandrevurdering til våren. 86 % av masterstudentene ønsker i *svært liten grad* eller i *liten grad* at hverandrevurdering skal bli obligatorisk til våren. Påstanden som måler hvorvidt studentene syntes at det å gi tilbakemelding til medstudent var lærerikt har 3,29 i gjennomsnittsverdi. Til sammenligning er gjennomsnittsverdien 2,14 for påstanden som måler om studentene syntes at det var lærerikt å få tilbakemelding fra medstudent. I motsetning til førsteårsstudentene, opplevde masterstudentene det som mer lærerikt å *gi* enn å *få* tilbakemelding fra medstudent. Differansen i gjennomsnittsverdien mellom de to påstandene er også betydelig større, sammenlignet med førsteårsstudentene.

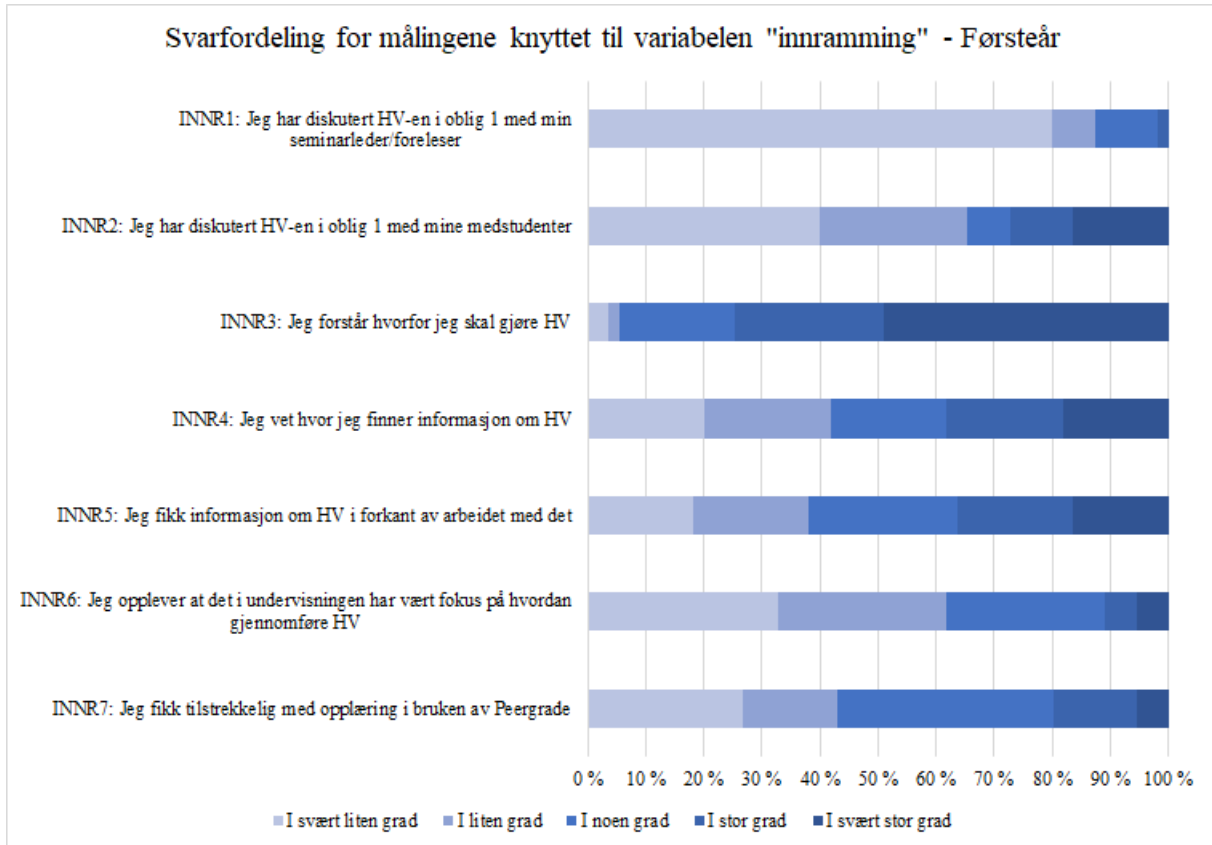
Det gjennomsnittlige standardavviket for variabelen *opplevd læring* er også relativt høy for masterstudentene. Standardavviket tyder på at masterstudentene har en større spredning i sine resultater sammenlignet med førsteårsstudentene. Dette kan trolig skyldes at utvalget for masterstudentene (N=7) er mindre. Samtidig har to av påstandene et standardavvik rett under 1, som indikerer at studentene er noe mer samstemte. Påstanden som måler om det skal bli obligatorisk med hverandrevurdering til våren har størst standardavvik for både førsteårsstudentene (1,40) og masterstudentene (1,41).



### 5.1.1.2 Kvantitative funn knyttet til variabelen 'innramming'

Variabelen *innramming* ble målt av totalt syv påstander.

#### Førsteårsstudenter



Figur 8. Svarfordeling for målinger av variabelen *innramming* for førsteårsstudenter

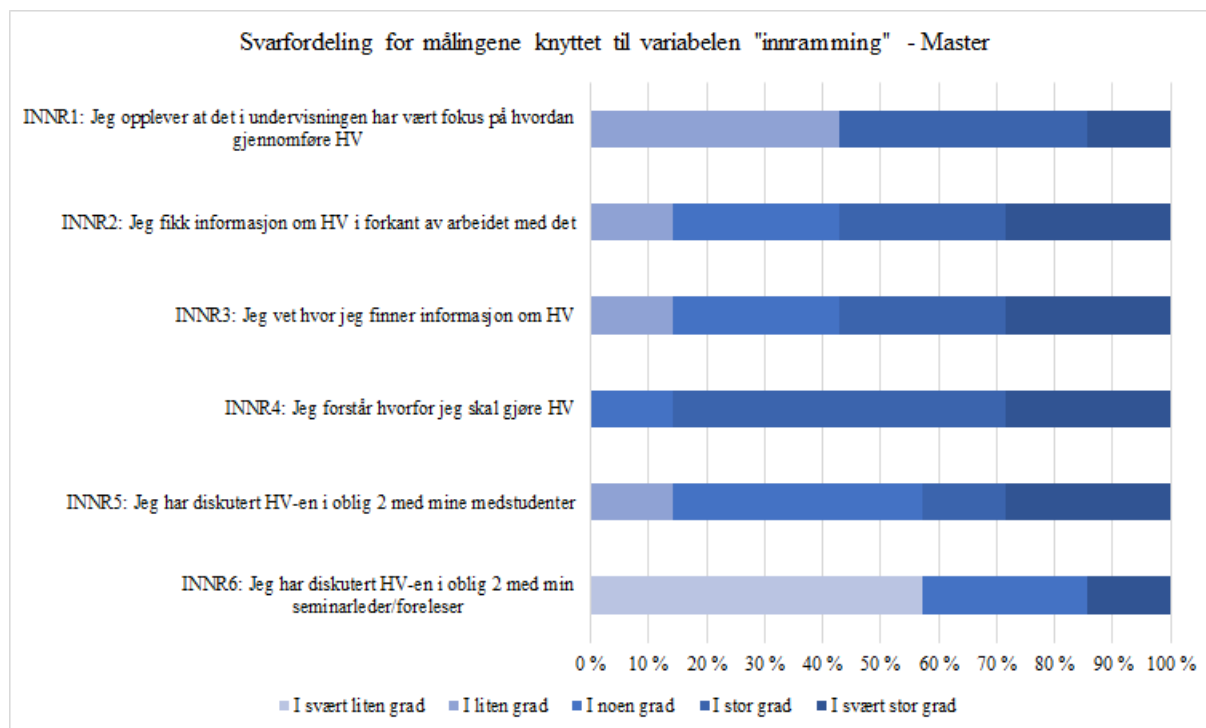
Tabell 6. Gjennomsnitt og standardavvik for variabelen *innramming* for førsteårsstudenter

	INNR1	INNR2	INNR3	INNR4	INNR5	INNR6	INNR7	Gj.snitt
Gj.snitt	1,35	2,38	4,15	2,95	2,96	2,22	2,58	2,66
St.avvik	0,75	1,51	1,04	1,40	1,35	1,13	1,18	1,19

Figur 8 viser at 75 % har svart at de *i stor grad* eller *i svært stor grad* forstår hvorfor de skal gjøre hverandrevurdering. Svært få har diskutert hverandrevurderingen i oblig 1 med sin seminarleder/foreleser. Hele 87 % svarer at de *i svært liten grad* eller *i liten grad* har diskutert dette med seminarleder/foreleser. Standardavviket for påstanden 1 er 0,75, som viser

forholdsvis lite spredning i svarene. Av figur 8 og tabell 6 kan en se at det er flere av studentene som har diskutert hverandrevurdering med medstudenter, sammenlignet med seminarleder/foreleser. Påstand 4 har størst standardavvik (1,40) og dermed størst spredning i svarene. Dette forteller at det er stor variasjon i om studentene vet hvor de finner informasjon om hverandrevurdering, selv om de tar samme emne.

### Masterstudenter



Figur 9. Svarfordeling for målinger av variabelen innramming for masterstudenter

Tabell 7. Gjennomsnitt og standardavvik for variabelen innramming for masterstudenter

	INNR1	INNR2	INNR3	INNR4	INNR5	INNR6	Gj.snitt
Gj.snitt	3,29	3,71	3,86	4,14	3,57	2,14	3,45
St.avvik	1,25	1,11	1,68	0,69	1,13	1,57	1,24

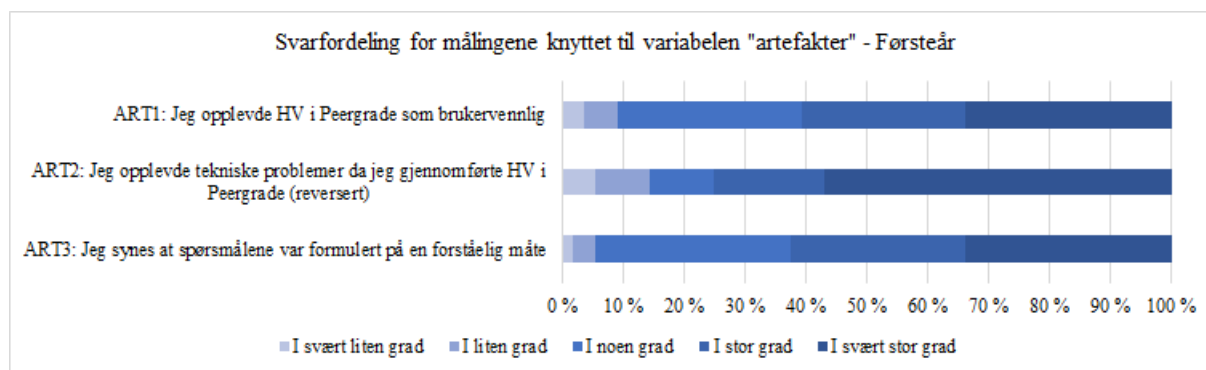
Flertallet av studentene har gitt høye verdier, som indikerer høyere grad av enighet med påstandene. Gjennomsnittsverdien for variabelen *innramming* er høyere for masterstudentene (3,45) sammenlignet med førsteårsstudentene (2,66). I likhet med førsteårsstudentene, ser en at masterstudentene også har en god forståelse for hvorfor de skal gjøre hverandrevurdering.

86 % har svart at de *i stor grad* eller *i svært stor grad* forstår hvorfor de skal gjøre hverandrevurdering. Standardavviket for påstand 4 er 0,69, som også viser lite spredning i svarene. Flertallet av studentene har diskutert hverandrevurdering med sine medstudenter. I likhet med førsteårsstudentene har færre diskutert hverandrevurdering med sin seminarleder/foreleser, selv om gjennomsnittsverdien for denne påstanden er noe høyere for masterstudentene (2,14).

### 5.1.1.3 Kvantitative funn knyttet til variabelen ‘artefakter’

Variabelen *artefakter* ble målt av totalt tre påstander.

#### Førsteårsstudenter



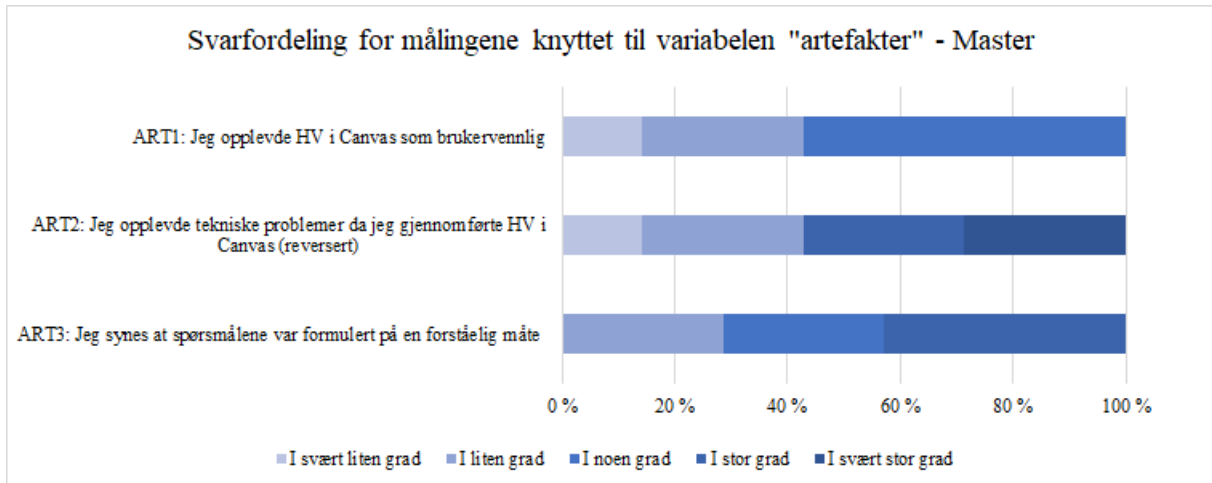
Figur 10. Svarfordeling for målinger av variabelen *artefakter* for førsteårsstudenter

Tabell 8. Gjennomsnitt og standardavvik for variabelen *artefakter* for

	ART1	ART2 (reversert)	ART3	Gj.snitt
Gj.snitt	3,87	4,17	3,95	4
St.avvik	1,02	1,18	0,91	1,04

Svarfordelingen i figur 10 viser at en overveiende del av svarene indikerer stor grad av enighet i påstandene. Dette leses også av gjennomsnittsverdien for variabelen *artefakt* som er 4. Resultat fra figur 10 og tabell 8 viser at studentene virker å være ganske fornøyd med Peergrade og spørsmålsrubrikken. Påstand 2 er reversert og omkodet, da den var formulert negativt i spørreskjemaet.

## Masterstudenter



**Figur 11.** Svarfordeling for målinger av variabelen artefakter for masterstudenter

**Tabell 9.** Gjennomsnitt og standardavvik for variabelen artefakter for masterstudenter

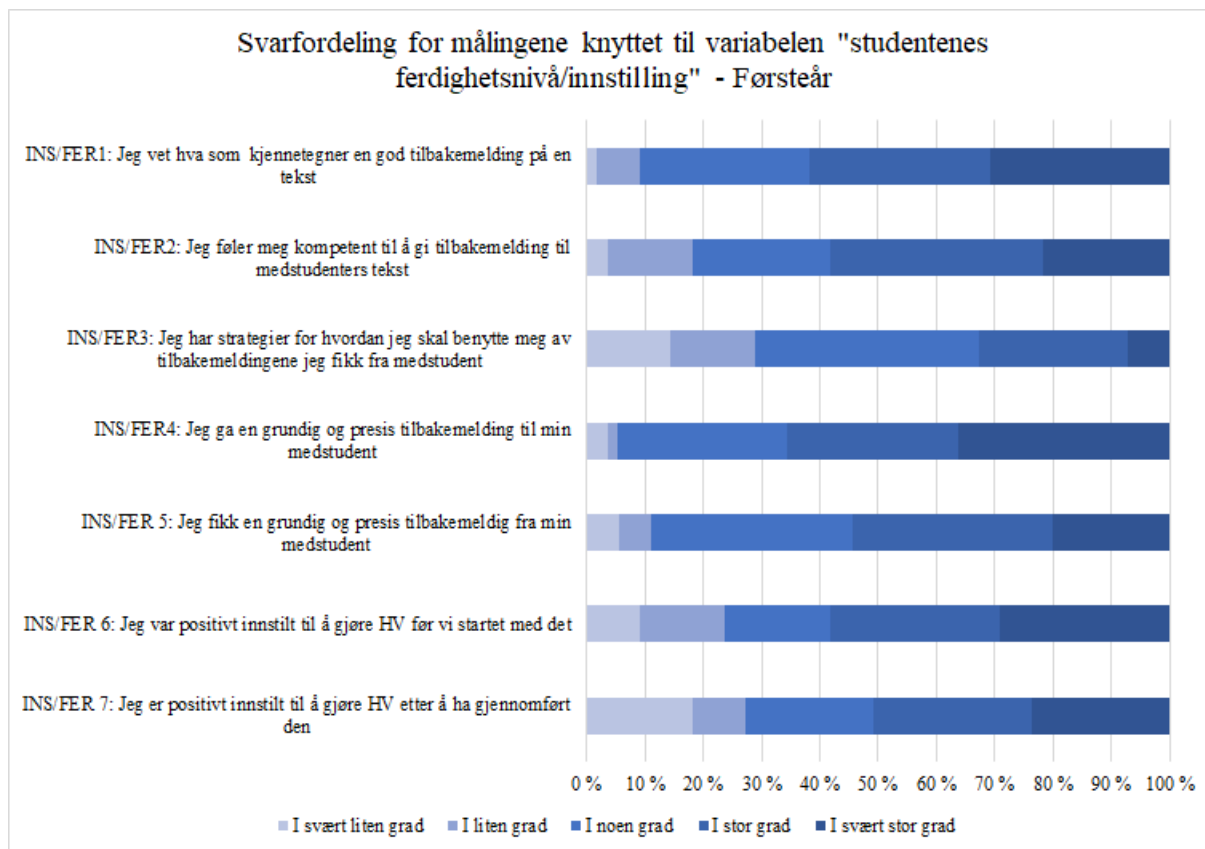
	ART1	ART2 (reversert)	ART3	Gj.snitt
Gj.snitt	2,43	3,29	3,14	2,95
St.avvik	0,79	1,6	0,9	1,1

Den gjennomsnittlige verdien for variabelen *artefakt* er lavere for masterstudentene (2,95), sammenlignet med førsteårsstudentene (4). Det kan dermed se ut til at masterstudentene var mindre fornøyd med Canvasrubrikker og hvordan rubrikkene var formulert. Den ulike svarfordelingen mellom studentgruppene kan også være et resultat av at de har brukt ulike digitale verktøy. Påstanden som måler om studentene opplevde tekniske problemer da de gjennomførte hverandrevurdering i Canvas, har et standardavvik på 1,6. Dette er et ganske høyt standardavvik for en skala som går fra 1-5, og dette indikerer uenighet.

### 5.1.1.4 Kvantitative funn knyttet til variabelen 'studentenes ferdighetsnivå/innstilling'

Variabelen *studentenes ferdighetsnivå/innstilling* ble målt av totalt syv påstander.

## Førsteårsstudenter



**Figur 12.** Svarfordeling for målinger av variabelen studentenes ferdighetsnivå/ innstilling for førsteårsstudenter

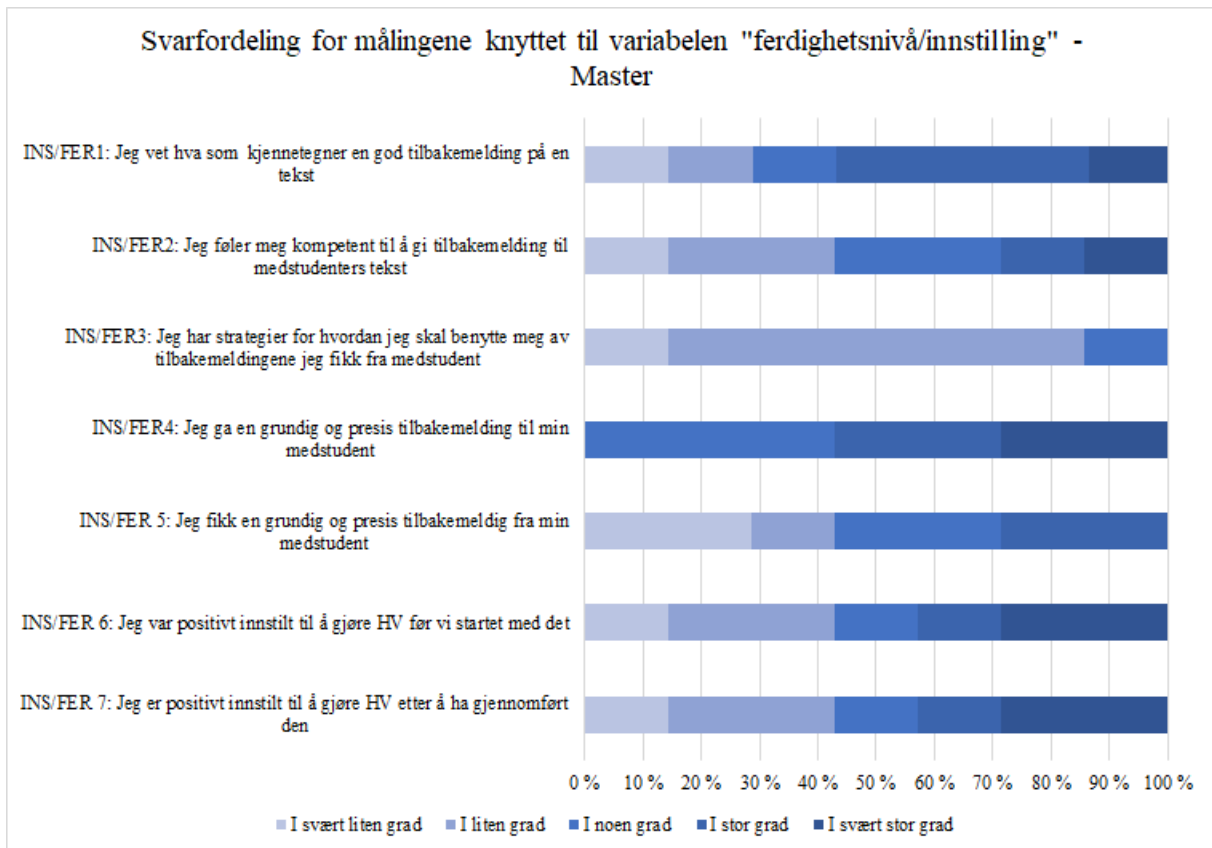
**Tabell 10.** Gjennomsnitt og standardavvik for variabelen studentenes ferdighetsnivå/ innstilling for førsteårsstudenter

	INS/FER1	INS/FER2	INS/FER3	INS/FER4	INS/FER5	INS/FER6	INS/FER7	Gj.snitt
<b>Gj.snitt</b>	3,82	3,58	2,96	3,93	3,58	3,55	3,29	3,53
<b>St.avvik</b>	1,02	1,1	1,13	1,03	1,05	1,3	1,4	1,15

Figur 12 viser at flertallet av studentene gitt høye verdier til *ferdighetsnivå/ innstilling*. Over 90 % av studentene har svart at de *i noen grad* til *i svært stor grad* vet hva som kjennetegner en god tilbakemelding på en tekst. Mange av studentene har også både gitt og fått en grundig og presis tilbakemelding. 66 % av studentene svarte at de *i stor grad* eller *i svært stor grad* ga en grundig og presis tilbakemelding. 55 % av studentene besvarte av de *i stor grad* - *i svært*

stor grad fikk en grundig og presis tilbakemelding. Tabellen viser en liten nedgang i gjennomsnittsverdien for studentenes innstilling til å gjøre hverandrevurdering etter de hadde arbeidet med den (før = 3,55, etter = 3,29).

### Masterstudenter



Figur 13. Svarfordeling for målinger av variabelen studentenes ferdighetsnivå/ innstilling for masterstudenter

Tabell 11. Gjennomsnitt og standardavvik for variabelen studentenes ferdighetsnivå/ innstilling for masterstudenter

	INS/FER1	INS/FER2	INS/FER3	INS/FER4	INS/FER5	INS/FER6	INS/FER7	Gj.snitt
Gj.snitt	3,29	2,86	2	3,86	2,57	3,14	3,14	2,98
St.avvik	1,38	1,35	0,58	0,9	1,27	1,57	1,57	1,23

Masterstudentenes svarfordeling for målingene knyttet til variabelen ferdighetsnivå/innstilling er noe mer fordelt utover hele skalaen. Den totale gjennomsnittsverdien for denne variabelen er 2,98, med et standardavvik på 1,23. Flertallet av studentene opplevde at de ga en mer

grundig og presis tilbakemelding, enn hva de mottok. 57 % av studentene mener at de *i stor grad* eller *i svært stor grad* ga en grundig og presis tilbakemelding til medstudent. 29 % mente at de *i stor grad* mottok en grundig og presis tilbakemelding fra medstudent. Ingen av studentene mente at de *i svært stor grad* mottok en grundig og presis tilbakemelding. Et interessant resultat er at hele 86 % av masterstudentene mener at de *i svært liten grad* eller *i liten grad* har strategier for hvordan de skal benytte seg av tilbakemeldingen mottok. Dette er også påstanden med lavest standardavvik av alle påstandene i hele spørreskjemaet, som betyr at flertallet av studentene mener at de i liten grad har strategier de kan benytte seg av i dette arbeidet. Tabellen viser ingen endring i gjennomsnittsverdi for studentenes innstilling til å gjøre hverandrevurdering etter de har arbeidet med den. Påstanden som måler studentenes innstilling til å gjøre hverandrevurdering før og etter å ha arbeidet med hverandrevurdering viser en gjennomsnittsverdi på 3,14.

### 5.1.2 Statistiske analyser

Kvantitative studier baserer seg på en begrenset mengde data fra en målpopulasjon. Spørreskjemaet for førsteåret er besvart av 55 studenter av en målpopulasjon på totalt 165 studenter som gjennomførte hverandrevurdering høsten 2020. I følge standarder for utvalgsstørrelser burde utvalgsstørrelsen vært 61 respondenter ved konfidensintervall 90 %, som innebærer at målingene med 90% sikkerhet måler verdien for vårt aktuelle utvalg. Med et utvalg på 165 respondenter får en derimot en feilmargen på 11 %, som innebærer større usikkerhet rundt målinger gjort av dataene. En feilmargen over 10 % anbefales i utgangspunktet ikke (Dancey & Reidy, 2007), men vi velger å godta feilmarginen.

#### **Bivariat korrelasjonsanalyse: korrelasjon mellom mulige forklaringsvariabler og studentenes opplevde læring av hverandrevurdering**

For å gjennomføre korrelasjonsanalyse er det brukt Spearmans rangkorrelasjonskoeffisient. Analysen presenterer korrelasjoner mellom avhengig variabel (*opplevd læring*) og forklaringsvariabler (*artefakter*, *studentenes ferdighetsnivå/innstilling*, og *innramming*). Korrelasjonen mellom de ulike variablene presenteres i tabell 12. Denne analysen er kun gjennomført med data fra førsteårsstudentene (N=55), da datautvalget fra masterstudentene er for lite (N = 7).

Korrelasjon mellom avhengig variabel opplevd læring og forklaringsvariablene innramming, artefakt og ferdighetsnivå/ innstilling.

Tabell 12. Resultat fra korrelasjonsanalyse: korrelasjonsstyrke mellom ulike variabler

	Opplevd læring	Innramming	Artefakter	Ferd.niv./innst
Opplevd læring	1,00			
Innramming	0,53**	1,00		
Artefakter	0,44**	0,39**	1,00	
Ferd.niv./innst	0,72**	0,58**	0,58**	1,00

\*P ≤ 0.05, \*\* P ≤ 0.01

Alle forklaringsvariablene har en signifikant positiv korrelasjon med den avhengige variabelen *opplevd læring* ved signifikansnivå 0,01. Dette viser at funnene ikke skyldes tilfeldigheter og er spesielt viktig i studier med få informanter som denne.

Ut fra tabellen ser en at korrelasjonskoeffisienten mellom *opplevd læring* og *innramming* som indikerer en *moderat* positiv korrelasjon mellom de to variablene ( $r_s = 0,53$ ,  $p < 0,01$ ). Dette resultatet *kan* indikere at *opplevd læring* påvirkes av forklaringsvariabelen *innramming*.

Korrelasjonskoeffisienten mellom *opplevd læring* og *artefakt* har lavest verdi ( $r_s = 0,44$ ,  $p < 0,01$ ) av de tre forklaringsvariablene. Den er likevel holdbar, og kan indikere en *moderat* positiv korrelasjon mellom studentenes opplevelse av læring i hverandrevurdering og bruk av artefakter.

Korrelasjonen mellom *opplevd læring* og *ferdigheter/ innstillinger* viser høyest verdi, og har en korrelasjonskoeffisient på 0,72 ( $p < 0,01$ ). Dette indikerer at det er en *sterk* positiv korrelasjon mellom studentenes opplevelse av læring og deres ferdigheter og innstilling til hverandrevurdering. Dette resultatet *kan* tyde på at studentenes opplevelse av egne



ferdigheter, samt deres innstilling til hverandrevurdering er den variabelen som i høyest grad har hatt påvirkning på om studentene opplevde at hverandrevurdering har vært lærerikt.

### Multippel regresjonsanalyse

Det ble utført en multippel regresjonsanalyse for å undersøke om variablene *innramming*, *artefakt* og *ferdighetsnivå/innstilling* påvirker *opplevd læring* av hverandrevurdering.

**Tabell 13.** Resultat fra regresjonsanalyse: sammenhenger mellom opplevd læring og forklaringsvariabler

		Forklaringsvariabler		
Avhengig variabel	Koeffisienter	Innramming	Artefakter	Ferdighetsnivå/ innstilling
Opplevd læring	<b>p</b>	0,04*	0,61	0,00**
	<b>β</b>	0,27	0,06	0,50
	<b>R<sup>2</sup></b>	0,55		

\* $P \leq 0.05$ , \*\*  $P \leq 0.01$

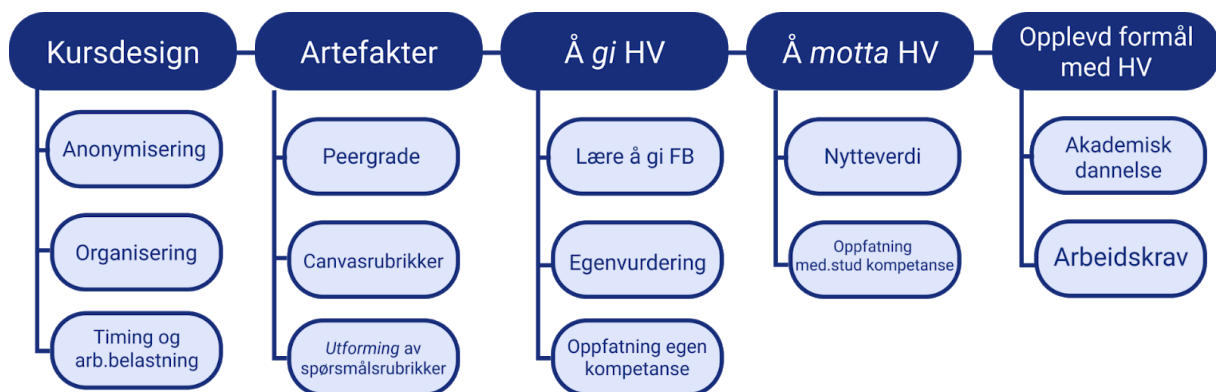
$R^2$  er et mål for hvor stor forklaringskraft forklaringsvariablene har, og hvor mye av variansen i variabelen *opplevd læring* som forklares av disse.  $R^2$  går fra 0 til 1. I denne målingen er  $R^2$  målt til 0,55, som indikerer at over halvparten av variasjonen knyttet til studenters *opplevd læring* kan forklares av forklaringsvariablene.

P-verdiene viser signifikansnivå for forklaringsvariabelen sin effekt på den avhengige variabelen. Beta er et mål for hvor mye *opplevd læring* endres når for eksempel *artefakt* endres med en skalaenhet, mens *innramming* og *ferdighetsnivå/innstilling* holdes konstante. Sammen gjør disse to verdiene at en kan avgjøre hvilke forklaringsvariabler som har påvirkning på den avhengige variabelen *opplevd læring*. Tabell 13 viser at studentenes *ferdighetsnivå/innstilling* er variabelen som har størst betydning på den avhengige variabelen. Denne variabelen er signifikant og indikerer den høyeste positive påvirkningen på studentenes opplevde læring ( $\beta = 0,50$ ,  $p < 0,01$ ). Også variabelen *innramming* har signifikant betydning

for *opplevd læring* og indikerer positiv påvirkning på den avhengige variabelen ( $\beta = 0,27$ ,  $p = 0,04$ ).

## 5.2 Kvalitative funn

Funn fra kvalitative data er delt inn i temaer i henhold til Braun og Clarke (2006) sin metode for tematisk analyse, se punkt 4.3.2. Gjennom tematisk analyse ble det identifisert fem temaer med tilhørende undertemaer, som anses som betydningsfulle for å besvare oppgavens forskningsspørsmål.



**Figur 14.** Temaer og undertemaer identifisert gjennom tematisk analyse

Tema og undertemaer fungerer som paraplybegreper og oppsummer hva kodene uttrykker, se punkt 4.3.2. Sitater fra informantene som deltok i forskningsintervjuer, samt sitater fra fritekstspørsmål i spørreskjema vil bli referert til. Informantene er enten førsteårsstudenter, masterstudenter, koordinator eller undervisere for studentene. For å skille mellom informantenes sitater og våre fortolkninger, blir sitater skrevet i anførselstegn og kursiv.

### 5.2.1 Kvalitative funn knyttet til kursdesign

Dette tema handler om hvordan de to aktuelle kursene, førsteårskurset og masterkurset, er designet. Her diskuterer informantene ulike elementer i kursdesignet som påvirket deres opplevelse av å delta i hverandrevurderingsprosessen. I analysen ble det identifisert tre undertemaer: *anonymisering*, *organisering*, og *timing/arbeidsbelastning*.

## **Anonymisering**

De to studentgruppene gir inntrykk av at de i ulik grad er opptatt av hvorvidt hverandrevurderingen skal gjennomføres anonymt eller ikke. Førsteårsstudentene virket å være mest opptatt av denne problemstillingen. Ordet *anonym* er nevnt 23 ganger i intervjuet, og blir nevnt flere ganger også på spørsmål som ikke direkte omhandler anonymitet. Studentene diskuterer både fordeler og ulemper ved å være anonym og ved å være ikke-anonym. Blant førsteårsstudentene virker det å ha vært noe forvirring om hvorvidt hverandrevurderingen skulle gjennomføres anonymt eller ikke. Det at flere av studentene var usikre på om hverandrevurderingen skulle leveres anonymt eller ikke, kan være en faktor som er med på å forklare hvorfor studentene var opptatt av å diskutere dette temaet under intervjuet. Noen av førsteårsstudentene mente at det var enklere å være ærlig og at det opplevdes mindre angstfylt når en var anonym. En av studentene sier at hen ikke føler seg moden nok til å være ikke-anonym. Samtidig blir det også sagt av en annen førsteårsstudent at ikke-anonymitet er en hindring en må komme over:

*“(...) Ehm.. så veldig stor forståelse for at det er en frykt ved at andre skal lese tekstene sine og føle på det, men jeg tror det er en viktig hindring som man må komme seg over for å begynne å komme videre da”.*

Under intervjuet med både førsteårsstudenter og masterstudenter nevnes det at en fordel ved å være ikke-anonym er at en kan spørre medstudenten dersom noe er uklart, og at det gir press om å gi en god vurdering til sin medstudent.

Masterstudentene uttrykker ikke et ønske om at hverandrevurderingen skal være anonym, og snakker heller ikke særlig om det i intervjuet. Ordet *anonym* er nevnt seks ganger, og kun som svar på spørsmål som handler eksplisitt om anonymitet. Studentene gir inntrykk av at de er inneforstått med at feedback i akademia og i arbeidsliv ikke gjennomføres anonymt, og at dette er noe de er vant med. I motsetning til førsteårsstudentene er masterstudentene en mindre studentgruppe. Angående studentgruppen sier en masterstudent følgende:

*“Jeg tenker at i en så liten gruppe som det vi er hvor alle kjenner hverandre, så synes jeg ikke det er viktig egentlig. Fordi det er ikke alltid vi skjønner helt hva kommentaren går ut på, og da kan det være greit å kunne spørre den medstudenten da”.*

Uttalelsen ovenfor kan indikere at størrelsen på studentgruppen kan ha vært en faktor som gjorde at masterstudentene ikke i like stor grad var opptatt av om hverandrevurderingen skulle skje anonymt eller ikke. En mindre studentgruppe kan gi bedre muligheter for å bli kjent med sine medstudenter. Det sosiale miljøet blant studentene var også noe underviser for masterstudentene var opptatt av, ved at hen planla for gruppearbeid og samarbeid i undervisningen. Av positive sider ved å gjennomføre en ikke-anonym hverandrevurdering, sier en masterstudent:

*“Og så føler jeg også det at det ikke er anonymisert gir meg som skriver teksten litt press på at okei, jeg skal få til en god vurdering fordi noen jeg kjenner skal lese teksten min. Også meg som skal gi en kommentar og tilbakemelding, det gir meg litt mer lyst til å gi en god tilbakemelding fordi jeg kjenner personen jeg skal vurdere”.*

Utsagnet ovenfor tilfører et annet perspektiv i diskusjonen ved at en kan oppleve en form for positivt press ved at en ønsker å skrive en god tekst og å gi en god feedback, fordi en kjenner personen som en både gir og mottar feedback til eller fra. Den siste innleveringen skulle blant annet gjennomføres anonymt, men studentene skrev navnet sitt på innleveringen likevel. Dette indikerer også at det å være anonym eller ikke er en problemstilling som masterstudentene ikke var særlig opptatt av.

## **Organisering**

Her presenteres funn knyttet til *organisering* av hverandrevurderingen. Med organisering menes gruppering av studenter, arbeidsformer og informasjon gitt til studentene.

Under intervjuet forteller masterstudentene at rammene rundt hverandrevurderingen vanskeliggjorde arbeidet. En utfordring som både masterstudentene og underviser for masterstudentene peker på handler om at studenter hadde skrevet tekster om ulike tema. Angående dette sier en masterstudent:

*“Og også det at vi skal gi en sånn hverandrevurdering, at vi blir spleiset sammen med personer som skriver om hvertfall cirka de samme temaene, eller om de samme tekstene. Så de har litt kunnskap om det du skriver om. For sånn som vi har gjort det nå, så har vi faktisk*

*gjort det i akkurat motsatt rekkefølge. Vi har bare først fått en tekst som vi skal kunne noe om, og etterpå har vi sittet i gruppe og diskutert kjempeforskjellige tekster som vi ikke egentlig hadde noe bakgrunn for å kunne noe om”.*

Det var en felles enighet blant masterstudentene under intervjuet at de ønsket en organisering som sikret at en både ga og mottok feedback om samme tema som de selv hadde valgt for eksamen. Studentene mente at det ville føre til at de kunne gi bedre feedback som også omhandlet tekstens innhold, og ikke bare kommentere på struktur og språk. Vedrørende denne problemstillingen sier underviser for masterstudentene:

*“Jeg skjønnte ikke at det gikk an å kommentere på et sammendrag av en tekst der de ikke hadde lest teksten. Men det klarte studentene da... Altså de skulle bare liksom ha gjennomført det. Altså de tenkte ikke at nå skulle de hjelpe medstudenten sin, eller nå måtte de liksom gjøre en skikkelig jobb”.*

Ut fra det som både masterstudenter og underviseren forteller, kan det se ut som at det har vært litt uklare retningslinjer vedrørende arbeidet med hverandrevurderingen. Underviseren forteller at hen ikke hadde tatt høyde for at studentene kunne vurdere hverandres tekster uten å ha lest seg opp på tema for teksten de skulle vurdere. Hen mente at noe av poenget med hverandrevurderingen var nettopp det at studentene måtte lese om et nytt tema på forhånd. Under intervjuet gir likevel alle masterstudentene inntrykk av at de er positive til fenomenet hverandrevurdering. Det blir sagt at de har tro på at det er nyttig innenfor andre rammer hvor studentene gir og mottar feedback som omhandler samme tema. Flere sier også at de har gjort hverandrevurdering før, men med en annen organisering, og at dette har vært mer vellykket. To av masterstudentene beskriver det slik:

*“Jeg tenker sånn at før man eventuelt skal gi den formelle hverandrevurderingen, sånn inne på Canvas med de spørsmålene, så tenker jeg at hvis man først har sittet og pratet sammen, sånn en gruppe eller noen studenter, har sittet og pratet litt sammen om hvordan man ser på hverandres tekster og sånn, så man får litt øving i den måten å kommunisere på, da”.*

*“(...) Tilbakemeldingen man gir fortløpende i små grupper eller hva det måtte være. Uformelle grupper - det er jo hverandretilbakemelding. Så, da drar jeg jo nytte av det”.*

Sitatene ovenfor beskriver hvordan masterstudentene tenker at hverandrevurdering i grupper kan organiseres. Underviseren til masterstudentene mener at organiseringen ble begrenset av koronarestriksjonene, og at dette fikk negative konsekvenser for gruppesammensetningen. Høsten 2020 innførte universitetet såkalte “basisgrupper” hvor de samme studentene, 3-6 personer, skulle samarbeide om alt av kollokvie- og gruppearbeid med hensyn til smittevern. Underviseren sier at hen liker å rullere grupper midtveis i semesteret. I sitatene nedenfor beskriver underviseren for masterstudentene hvordan de ulike gruppene av studenter ble påvirket av å være organisert i faste basisgrupper gjennom hele semesteret, og hvorfor hen pleier å unngå en slik organisering.

*“Så de gruppene som fungerte bra, de liksom bare løfter hverandre opp, og de som var sånn middels, de ble middels hele semesteret, på grunn av at det utviklet seg slike gruppenormer, da. Tidligere så har jeg jo mikset og trikset på hvem som gir kommentar til hvem, men det kunne jeg ikke gjøre nå da”.*

Videre snakker underviseren om ulike utfordringer som kan oppstå i et gruppearbeid. Hen beskriver hvordan studenter i et gruppearbeid også kan overta hverandres misoppfatninger, samt hvordan studentene tilpasser seg hverandre i en gruppe:

*“Det at de kan overta hverandre sine misforståelser. Det som er faren i et gruppearbeid så er det en som har en forståelse av en teori som er litt på siden, og den personen snakker overbevist om at sånn er det bare. Og da er de andre studentene, å ja, dette kan jo stemme. Nei, hvorfor får vi så dårlig karakter til eksamen? (latter). Nei, det er jo fordi at de har på en måte forsterket hverandre i en misforstått tolkning av en tekst (...). En annen ulempe er jo at studentene når de er veldig opptatt av å jobbe med hverandre i gruppe, så innretter de seg ofte etter den svakeste i gruppa”.*

### **Timing og arbeidsbelastning**

Med *timing* menes bestemmelse av tidspunkt for gjennomføring av hverandrevurdering og aktiviteter i forbindelse med dette. Med *arbeidsbelastning* menes summen av arbeid som studentene må gjøre i det aktuelle emnet, men også totalt i løpet av semesteret i sitt studieprogram. Dette kan være obligatoriske oppgaver, eksamener, forelesninger og seminarer.

Timing og arbeidsbelastning er et tema som går igjen i alle intervjuene, og som viser seg å være av betydning når informantene reflekterer over kursdesignet. Angående eksamen sier noen av studentene at oppgaven med å skrive sammendrag var timet bra med tanke på at det fant sted før eksamen. Samtidig sier andre førsteårsstudenter at de opplever at hverandrevurderingen stjeler tid før eksamen. En av studentene mislikte at en mottok hverandrevurdering i Peergrade fredagen, og dermed måtte bruke helgen på å ta i bruk hverandrevurdering fra medstudent. Dette fordi egen tekst hadde innlevering allerede mandagen. Flere av de andre studentene sa seg enig i at dette var en uheldig organisering. Underviser for førsteåret nevner at lite tid kan være en utfordring når en skal planlegge for hverandrevurdering. Underviseren skulle ønske at studentene kunne ha brukt lengre tid på hverandrevurderingene:

*“Vi har nok et problem og det er at semesteret ikke er lengre. For det hadde vært veldig stilig om vi kunne brukt enda mer tid på det”.*

Både masterstudentene og underviser for masterstudentene mener at timingen kunne vært annerledes. På spørsmål om hva som kunne vært designet annerledes i hverandrevurderingen svarer underviser for masterstudentene følgende:

*“Kanskje begynt senere i semesteret. Altså jeg begynte vel andre eller tredje uken eller noe sånt, veldig fort, rett på liksom. Så jeg ville kanskje ha gitt dem litt mer tid. Til å bare svelge inn hva emnet handlet om, at de fikk lest litteraturen godt på forhånd”.*

Også masterstudentene er enige i at hverandrevurderingen kom for tidlig i semesteret. Det førte til at studentene fikk dårlig tid til å lese seg opp på temaet til medstudentens tekst, og dermed valgte å prioritere å undersøke eget tema. Vedrørende timing og arbeidsbelastning sier en masterstudent følgende:

*“For det var så ticht gap på alt vi skulle gjøre. Så da prioriterer du jo de fremføringene og de tingene du selv skal ha”.*

Masterstudentene kommenterte timing og arbeidsbelastning både under intervjuet og i fritekstspørsmål i spørreskjemaet. Flere av dem kommenterte også at de på grunn av at de hadde flere oppgaver samtidig måtte velge hvilke oppgaver de skulle prioritere. Det kan dermed se ut som at flere studenter opplevde en periode med høy arbeidsbelastning, som påvirket deres arbeid med hverandrevurderingen. Et eksempel fra fritekstspørsmål i spørreskjemaet:

*“(..) Også mulig hverandrevurderingen har vært nedprioritert da vi har hatt 4-5 oblig i dette faget pluss hverandrevurdering. Blir litt mye mtp at dette kun er 1 av 3 fag”.*

### 5.2.2 Kvalitative funn knyttet til bruk av artefakter

Hverandrevurderingen ble gjennomført digitalt gjennom å bruke artefaktene Peergrade (førsteårsstudenter) eller Canvasrubrikker (masterstudenter). Dette temaet består av tre undertemaer: *Peergrade* (førsteår) og *Canvasrubrikker* (master) og *utforming av spørsmålsrubikker*.

#### **Peergrade**

Koordinator for førsteårsstudentene sier at verktøyet Peergrade ble valgt i forbindelse med forskningsprosjektet ‘Formativ vurderingspraksis i Canvas’. Det var begrunnet i et ønske fra instituttet om å teste ut nye digitale verktøy i undervisningen, se punkt 4.1. Grunnen til at akkurat Peergrade ble valgt var fordi verktøyet åpner for egenvurdering, og en forventning om at verktøyet ville forenkle det logistiske arbeidet. Verktøyet viste seg imidlertid å være vanskelig å integrere med universitetets systemer som er utformet med hensyn til GDPR. Dette medførte en god del ekstraarbeid for koordinatoren. Videre sier koordinatoren at Peergrade har en øvre maksgrense for antall ord, noe som setter føringer for hverandrevurderingen og som kan være negativt i enkelte sammenhenger. Underviseren til førsteårsstudentene synes at Peergrade fungerte greit til formålet, men understreker også at selve verktøyet ikke er vesentlig i pedagogisk arbeid. Underviseren sier:

*“Men i praksis så kunne vi jo ha brukt et hvilket som helst verktøy som gjorde den jobben som Peergrade gjorde. Så for å være helt ærlig så var jeg ikke så himla interessert i Peergrade*



*annet enn at det var et verktøy for å gjøre den jobben vi skulle ha gjort, så ja... Det var sikkert helt greit, det var det vi måtte bruke”.*

Førsteårsstudentene brukte Peergrade for å gjennomføre hverandrevurderingen. På spørsmål om de opplevde Peergrade som brukervennlig var det bare én student som svarte. Denne studenten sa blant annet at Peergrade var:

*“Veldig lett forståelig og brukervennlig (...) så det er ikke der problemene ligger”.*

Studenten fikk ingen innvendinger mot dette, noe som indikerer at de andre førsteårsstudentene virket å være enige i uttalelsen. Altså merket ikke studentene noe til de tekniske utfordringene som koordinator opplevde.

### **Canvasrubrikker**

Masterstudentene gjennomførte hverandrevurderingen i Canvas. Underviser til masterstudentene ønsket ikke å bruke Peergrade for å gjennomføre prosjektet, og fikk bistand til å sette opp hverandrevurderingen i Canvas. Hen begrunner dette blant annet med at Peergrade var et ukjent verktøy. Videre sier underviser at hen ønsket at hverandrevurderingen skulle ha *“refleksjon i høysetet”* og bestå av åpne spørsmål fremfor *“kunnskapstesting”*. Videre sier informanten at hen ikke har undersøkt verktøyet Peergrade:

*“Nei, jeg har ikke sjekket Peergrade i det hele tatt. Jeg har ikke vurdert det. Men jeg bare hørte når de snakket om det, og så tenkte jeg at ja, men sånn tenker ikke jeg. Sånn vil ikke jeg ha det. (...) Jeg trodde at dette var et verktøy som var veldig begrensende”.*

Masterstudentene fortalte at de opplevde utfordringer med å gjennomføre hverandrevurdering i Canvas. Flere påpekte at det de hadde skrevet i rubrikkene forsvant ved tidsavbrudd. Når dette skjedde måtte studentene starte på nytt med å fylle inn rubrikkene. Dette opplevdes demotiverende, og la tidspress på studentene når de fylte ut rubrikkene. Videre var det forvirring rundt sende-knappen. Flere trodde at denne knappen symboliserte *“å lagre”* og opplevde dermed at vurderingen ble sendt når de derimot ønsket å lagre. Disse måtte trekke tilbake besvarelsen og begynne på nytt igjen. Studentene som brukte Peergrade var altså mer fornøyd med brukervennligheten enn dem som brukte Canvasrubrikker. Dette til tross for at

underviser for masterstudentene valgte å bruke Canvasrubrikker nettopp fordi hen trodde det ville være enklere for studentene. Det er underviser som selv utformer spørsmålene som skal integreres i Peergrade, og det er derfor ingen grunn til å anta at Peergrade ville være mer begrensende enn Canvasrubrikker.

### **Utforming av spørsmålsrubrikker**

Spørsmålsrubrikkene som førsteårsstudentene skulle fylle ut var mer spisset enn dem for masterstudentene, se vedlegg 1. Flere av rubrikkene dreier seg om hvorvidt formelle krav for en akademisk tekst er oppfylt. Underviser for førsteårsstudentene sier at hverandrevurdering åpner for at studentene kan få feedback som omhandler språklige feil og mangler som undervisere ikke har kapasitet til å kommentere. Førsteårsstudentene mente at hver enkelt spørsmål var utformet med for høy ordgrense. Flere sier at de heller kunne ønske at hverandrevurderingen som *helhet* hadde en satt ordgrense. Dette ville gitt rom for å selv vurdere og bruke tid på aktuelle deler av teksten som trengte forbedring. Videre etterspør hele 8 av totalt 21 studenter som fylte ut fritekstsvaret i spørreskjemaet, mer rom for å gi hverandre kritikk. En av studentene skriver i fritekstsvaret i spørreskjemaet at hen skulle gi feedback til en tekst med åpenbare feil og mangler, men spørsmålsrubrikken åpnet ikke for å kommentere disse feilene:

*“Synes det burde vært mer åpent for å gi konstruktiv kritikk i hverandrevurderingen. Spørsmålsboksene var relativt konkrete, og på den utfyllende boksen skulle man i utgangspunktet kun si hva som var positivt med oppgaven. Jeg opplevde på de jeg skulle lese gjennom at det var en del store og åpenbare mangler”.*

Underviser til masterstudentene ønsket derimot at spørsmålsrubrikken skulle åpne opp for refleksjon, og dermed være ganske åpne. Under intervjuet fikk masterstudentene se bilde av noen av spørsmålsrubrikkene. De ble spurt om hva de syntes om spørsmålene, og hvilke spørsmål som eventuelt traff godt eller mindre godt. Masterstudentene sier at spørsmålsrubrikken var vanskelig å fylle ut. Dette ble begrunnet med at det var vanskelig å svare på spørsmål som gjaldt faglig innhold om et tema som de ikke hadde kunnskap om selv. Om dette skriver en av masterstudentene i fritekstsvaret:

*“Så var det egentlig bare spørsmål I jeg kunne svare på. For jeg hadde ikke lest den andre teksten”.*

### 5.2.3 Kvalitative funn knyttet til å gi hverandrevurdering

Her vil det bli presentert funn knyttet til studenter og underviseres erfaring vedrørende det å gi hverandrevurdering. Dette temaet er delt inn i følgende undertemaer: *å lære å gi feedback, egenvurdering og studentenes opplevelse av egen kompetanse.*

#### **Å lære å gi feedback**

Underviser for førsteårsstudentene har vektlagt overfor studentene hvorfor det er viktig å gi hverandre kritikk i et faglig fellesskap. I denne sammenhengen presenterte underviseren en definisjon av kritikkbegrepet for studentene. Undervisere sammenligner hverandrevurdering med fagfelleevaluering i akademia, og mener med dette at studentene bør bli komfortable med å vurdere hverandres arbeid. Underviser for masterstudentene trekker frem lignende begrunnelse for hvorfor det er viktig å gi hverandre feedback. Hen ønsket å vektlegge overfor studentene at:

*“Vi vurderer ikke hverandre, vi hjelper hverandre”.*

Førsteårsstudentene gir uttrykk for at de er redde for at mottaker av vurderingen skal bli lei seg, og de ønsker ikke å oppfattes som *“slem”*. Dette til tross for at deres underviser har presentert dem for definisjonen av kritikkbegrepet, hvor det ble understreket at kritikk er:

*“Å få tilbakemelding på hva som kunne vært annerledes, og ikke en liste på hva som er feil”.*

Masterstudentene er derimot ikke bekymret for å bli oppfattet som *“slemme”* eller bekymret for at mottaker av vurderingen skal bli *“lei seg”*. De gir uttrykk for at både det å gi og motta vurdering bør være en naturlig del av å være deltaker i et faglig fellesskap.

#### **Egenvurdering**

Underviser for masterstudentene sier at det er nettopp det å gi hverandrevurdering som er det mest læringsfremmende i en hverandrevurdering. Videre trekker underviseren kobling mellom det å vurdere andre og det å gjøre egenvurdering:

*“Nei, det er jo at studentene vurderer seg selv. (...) Ja, det gjør de jo også hele tiden. De leser en tekst og så vurderer de teksten, men de vurderer jo også sin egen forståelse av teksten, og så har de samtale med hverandre om teksten da, og så vurderer de. Åja, kanskje den tenker litt annerledes enn meg”.*

Videre sier underviseren at hen har vektlagt læringsutbytte knyttet til det å gi hverandrevurdering når hen har snakket med studentene om hverandrevurderingen. Én av masterstudentene gir uttrykk for at det å gi hverandrevurdering kan forstås som en form for egenvurdering, selv om det ikke blir eksplisitt uttalt. Denne studenten skulle gi vurdering av en tekst som handlet om samme tema som sin egen tekst, og mente derfor at å gi vurdering var til hjelp for å sjekke om hen hadde forstått tema likt som sin medstudent.

Underviser for førsteårsstudentene gjør også en implisitt kobling mellom hverandrevurdering til det å gjøre egenvurdering. Hen mener at det å lese tekster på eget faglig nivå kan hjelpe deg å få et kritisk blikk på egen tekst:

*“Men tanken er selvfølgelig at man gjør noe annet når man tar kritikerrollen, så har du plutselig en del ting å tenke på som du ikke har tenkt på før. Blant annet så får du en anledning til å lese besvarelser som er på ditt eget nivå”.*

Også førsteårsstudentene forstår det å gi hverandrevurdering i sammenheng med egenvurdering. Flere sier at det er nyttig å se andres arbeid da de har en tekst å sammenligne sitt eget arbeid med, og videre at de kan hente inspirasjon fra medstudentens tekst:

*“Når jeg ser på andre sin tekst og hvordan de former det opp og jeg gir både ros og ris til han da, så går jeg tilbake igjen til egen tekst med ganske nye øyne. Og tenker at, har min tekst de samme tendensene?”.*

### **Studentenes oppfatning av egen kompetanse**

Flere av studentene uttrykker usikkerhet når det gjelder egen kompetanse i å gi en god vurdering. Eksempelvis sier en førsteårsstudent:

*“Så jeg følte at det var mer jeg selv som var redd for at det ikke skulle hjelpe noen andre (...) Men at det jeg gjorde rett og slett ikke var godt nok da”.*

Også masterstudentene er usikre på egen kompetanse knyttet til det å gi gode nok vurderinger, men begrunner dette med at de ikke hadde tilstrekkelig kunnskap om temaet for teksten som de skulle vurdere:

*“Det jeg fant mest utfordrende var å vurdere andre sine tekster. Når jeg selv ikke har lest tekstene, eller har selv tid til å gå dypt inn i disse. Føler ikke alltid at min kommentar på andres arbeid kan være brukbart, når jeg ikke har et dypere forståelse av tekstene selv. Og en annen sier: Og da ble det etterhvert litt sånn at man kanskje skrev litt mer over... til en generell tilbakemelding, da. Enn hvis man hadde lest litt mer, kanskje?”*

Altså begrunner masterstudentene sin kompetanse knyttet til det å gi hverandrevurdering med at de har for liten kunnskap om det aktuelle temaet som teksten handler om.

Førsteårsstudentene er derimot mer bekymret for egen kompetanse til å vurdere hva som kjennetegner en god akademisk tekst.

#### 5.2.4 Kvalitative funn knyttet til å motta hverandrevurdering

Her presenteres funn knyttet til det å motta hverandrevurdering. I analysen ble det identifisert to undertemaer: *nytteverdi* og *studentenes oppfatning av andres kompetanse*.

##### **Nytteverdi**

Underviser for førsteårsstudentene mener ikke at vurdering fra medstudenter skal eller kan erstatte vurdering fra lærer, men mener at det er en god øvelse å motta vurdering fra andre på samme faglige nivå. Underviseren forteller at dette ble presisert overfor studentene. Videre sier underviseren at medstudenter har anledning til å vurdere feil og mangler ved en tekst som en underviser ikke har kapasitet til å kommentere:

*“Og så vil nok studentene være flinkere til å si ifra om ting som er, som er kanskje litt mer konkrete som at “her har du referert feil” eller et eller annet sånn, som kanskje en foreleser kanskje ikke gidder, eller som de ser at her er det andre ting som må ordnes opp i først, før vi går på det nivået. For studentene trenger det jo”.*

Underviser for masterstudentene sier selv at hen ikke hadde fokus på det å *motta* hverandrevurdering, da det var nyttigheten knyttet til det å *gi* hverandrevurdering som hen hadde fokusert på:

*“Jeg var veldig opptatt av hvordan de skulle skrive, eller hvordan de skulle tenke når de gir tilbakemeldinger. Jeg var ikke så opptatt av hvordan de skulle bruke tilbakemeldingene for det tenkte jeg sånn at, det kan de bruke eller ikke bruke, det må være opp til dem liksom”.*

Underviseren sier likevel at hen i ettertid ser at det kunne ha vært mer fokus på det å ta i bruk feedback på en god måte. Disse funnene viser forskjeller i hvordan underviserne for de to studentgruppene har presentert og vektlagt mottakersiden ved hverandrevurdering overfor studentene.

Masterstudentene forteller at de opplevde en begrenset nytteverdi av det å *motta* hverandrevurdering. En av studentene fortalte at hen hadde vært heldig og dermed skulle motta feedback fra en medstudent som skrev om samme tema som hen selv. Denne studenten mottok feedback som gikk i dybden på det faglige innholdet, noe som opplevdes som meningsfullt. De som derimot fikk feedback fra en student som ikke hadde kunnskap om temaet mottok feedback som omhandlet språklige momenter ved teksten, som for eksempel struktur og skrivefeil. Denne formen for feedback opplevdes som overfladisk og lite nyttig. Samtidig er det grunn til å tro at feedback som omhandler formelle eller språklige feil og mangler er mer nyttig for førsteårsstudentene da disse har mindre erfaring og kompetanse i det å skrive akademisk.

### **Studentenes oppfatning av medstudenters kompetanse**

Førsteårsstudentene er stort sett enige i at de stoler mer på vurderinger gitt av undervisere enn av medstudenter. Dette inntrykket bekreftes også av underviseren deres som oppfatter at noen av studentene synes at det å bli vurdert av medstudenter er *“under deres verdighet”*.

Førsteårsstudentene forklarer denne usikkerheten ved å fortelle at de er bekymret for å overta misoppfatninger fra sine medstudenter. De synes selv at de er *“forvirra”* og har inntrykk av at andre i kullet er like forvirret som dem selv. Videre blir det sagt at kvaliteten på vurderingen som de mottar er avhengig av om det er en *“flink”* student som gir den. Det blir dermed opp

til tilfeldigheter om de får en vurdering av høy kvalitet. Likevel nevner en av studentene at det er nyttig å motta vurdering fra medstudenter “*fordi de ofte er på det samme nivået som jeg er*”. Masterstudentene opplevde derimot at det å motta vurdering hadde begrenset nytteverdi fordi medstudenten *ikke hadde kompetanse* om det aktuelle temaet.

Funn knyttet til oppfatning av medstudenters kompetanse står i sammenheng med funn som omhandler å *gi* hverandrevurdering og oppfatning av egen kompetanse. Det kan tolkes som at både førsteårs- og masterstudenter opplevde begrenset nytteverdi forbundet med det å motta hverandrevurdering. Begrunnelsen for dette er derimot ikke den samme. Der hvor førsteårsstudentene ikke stoler på sine medstudenter i like stor grad som underviser, opplevde masterstudentene at verdien var begrenset fordi medstudenten manglet kunnskap om det aktuelle temaet som hen skulle vurdere.

### 5.2.5 Kvalitative funn knyttet til formål med hverandrevurdering

Her vil funn knyttet til ulike aktørers opplevde formål med hverandrevurdering presenteres. Disse aktørene er instituttet, undervisere (for master og førsteår) og studenter (master og førsteår). Følgende undertema presenteres: *akademisk dannelse* og *arbeidskrav*.

For instituttet er hverandrevurderingen en aktivitet motivert av formålet for forskningsprosjektet ‘Formativ vurderingspraksis i Canvas’, se punkt 2.3.4 og 4.1 Undervisernes formål med aktiviteten er også forankret i forskningsprosjektet.

#### **Akademisk dannelse**

Underviser for førsteåret forstår aktiviteten hverandrevurdering i sammenheng med begrepet akademisk dannelse, hvor studentene trer inn i et faglig fellesskap. Hen beskriver at studentene ved gjennomføring av hverandrevurdering utsettes for noe som er nytt, men som de bør bli fortrolig med. Underviseren forteller at å gjennomføre hverandrevurdering kan bidra med å ufarliggjøre både det å gi og å få kritikk. Videre mener hen at det handler om å ha et prosessblikk på egen skriving, da “*en tekst aldri er perfekt*”.

Underviser for masterstudentene er enig i at hverandrevurdering er knyttet til deltakelse i et sosialt og faglig fellesskap. Underviseren er kritisk til å bruke ordet *vurdering*, og

understreker at vurdering ikke er et formål i seg selv. I masteremnet ble derfor begrepet *medstudentkommentar* brukt om det som i denne oppgaven omtales som hverandrevurdering:

*“Vurdering er jo et så stygt ord. (Latter) Ja, men det er så slem. (...) Vi vurderer ikke hverandre, vi hjelper hverandre”.*

Både førsteårs- og masterstudenter har forståelse for at det er viktig å ta del i et faglig fellesskap som også innebærer at de gir feedback på hverandres arbeid. Vi opplever likevel førsteårsstudentene som mer kritisk til å gjennomføre hverandrevurdering enn det masterstudentene er. Det blir riktignok uttalt flere ganger i intervjuet med førsteårsstudentene at studentene forstår at de må lære seg å vurdere hverandre for fremtidig studie- og arbeidsliv. Likevel gir flere av førsteårsstudentene uttrykk for at akkurat hverandrevurdering var meningsløst. Disse tendensene ser vi ikke hos masterstudentene, hvor alle gir uttrykk for at de i utgangspunktet er positivt innstilt til fenomenet hverandrevurdering:

*“Men jeg tenker teorien er jo kjempefin. Det er jo god læring for både jeg som vurderer, og for den som jeg skal vurdere”.*

### **Arbeidskrav: obligatoriske oppgaver og eksamen**

Med arbeidskrav menes i denne oppgaven obligatoriske oppgaver og eksamener i emnet. Kvalitative analyser viser at førsteårsstudentene har større fokus på hverandrevurderingens nytthet knyttet til arbeidskrav i emnet, sammenlignet med deres underviser. Førsteårsstudentene nevner ordet *oblig* fire ganger, mens *eksamen* blir nevnt ni ganger. Underviser for førsteåret nevner ordet *oblig* én gang og ordet *eksamen* blir nevnt to ganger. Dette kan tolkes som at der hvor underviseren ser på hverandrevurdering som nyttig i et langtidsperspektiv, er flere av studentene mer opptatt av hvilken nytte hverandrevurderingen vil gi på kort sikt. Med det sagt inneholder ikke intervjuguiden for underviserne spørsmål som eksplisitt nevner ordet eksamen, mens intervjuguiden for studentene inneholdt étt spørsmål knyttet til eksamen, og étt oppfølgingsspørsmål hvor ordet eksamen ble stilt.

Masterstudentene ser også ut til å være opptatt av arbeidskrav, og nevner ordet *eksamen* ti ganger. Underviser for masterstudentene nevner ordet *eksamen* fire ganger, og da for å forklare formålet med hverandrevurderingen:



“(…) Nei altså, overordnet så var det at de skulle skrive en god eksamensoppgave”.

### 5.3 Oppsummering av funn

Av kvantitative funn viser det seg at førsteårsstudentene opplever det å gi og det å motta hverandrevurdering som mer lærerikt enn masterstudentene. Dette står i motsetning til kvalitative funn som viser at førsteårsstudentene er mer skeptiske til å gjennomføre hverandrevurdering enn masterstudentene, herunder både det å gi og det å motta vurdering. I intervjuene kommer det derimot frem at masterstudentene er mer positive til både å gi og å motta vurdering enn førsteårsstudentene. Dette funnet forklares i intervjuene hvor masterstudentene uttrykker at de er mer inneforstått med betydningen av fagfellelvurdering. Masterstudentene opplevde imidlertid at *innrammingen* høsten 2020 fungerte dårlig. Utfordringer ved innrammingen høsten 2020 for masterstudentene skyldes i hovedsak at studentene ikke hadde kunnskap om temaet for teksten som de skulle vurdere, noe som skyldtes en misforståelse mellom underviser og studenter. Førsteårsstudentene oppgir i spørreskjema at det å *gi* og å *motta* hverandrevurdering er omtrent like lærerikt, men masterstudentene oppgir derimot at det å *gi* hverandrevurdering er mer lærerikt enn det å *motta*. Også dette kvantitative funnet forklares av kvalitative data fra intervjuet, hvor masterstudentene også begrunner dette at en skulle *gi/motta* feedback om en tekst med ukjent tema. Førsteårsstudentene uttrykte i intervjuene usikkerhet hvorvidt studentgruppen med seg selv inkludert innehar tilstrekkelig kompetanse for å vurdere andres tekster på grunn av lite akademisk erfaring. Denne bekymringen deles av masterstudentene, men begrunnes derimot med manglende kunnskap om det aktuelle temaet.

Der hvor underviserne ønsker at hverandrevurdering bidra til faglig fellesskap og akademisk dannelse, er studentene mer opptatt av hvordan hverandrevurdering vil være nyttig i nærstående arbeidskrav. Likevel anerkjenner studentene også at hverandrevurdering vil være nyttig i et langtidsperspektiv. Skillet i opplevd formål og tidsperspektiv for hverandrevurdering er mer fremtredende i førsteårsemnet enn i masteremnet. Både førsteårs- og masterstudentener hadde innvendinger angående formulering av spørsmålsrubrikk. Førsteårsstudentene etterspør flere spørsmålsrubrikker som åpnet for konstruktiv kritikk. Masterstudentene synes at spørsmålsrubrikken var vanskelig å fylle ut og også dette forklares av at masterstudentene ikke hadde kjennskap til tema for medstudenters tekst.

Funnene viser at det er mange enkeltfaktorer som spiller inn på gjennomføring av hverandrevurdering. Timing viste seg å ha stor betydning for både undervisere og studenter, hvor de opplevde at hverandrevurderingen kom for tett på andre arbeidskrav, og at det ble høy arbeidsbelastning. Førsteårsstudentene er opptatt av om hverandrevurderingen gjennomføres anonymt og flere uttrykker i intervjuet et ønske om anonymitet. Masterstudentene har derimot ikke ønske om anonymitet. Når det gjelder organisering hadde masterstudentene ønsket mer muntlig og uformeller rammer rundt hverandrevurdering, mens førsteårsstudentene uttrykker ikke samme ønske. Det er utført statistiske analyser med kvantitative data fra førsteårsstudentene. Korrelasjonsanalyse viser signifikant positiv korrelasjon mellom den avhengige variabelen *opplevd læring* og samtlige forklaringsvariabler, hvor *ferdighetsnivå/innstilling* har høyest positiv korrelasjon. Ved multippel regresjonsanalyse hadde variabelen *ferdighetsnivå/innstilling* høyest positiv påvirkning på studentenes *opplevde læring*. Dette forteller oss at studentene som føler seg kompetente til å både gi og motta feedback, brukte tilstrekkelig tid på oppgaven samt var positivt innstilt opplevde hverandrevurdering som lærerikt.

## 6 Diskusjon

I lys av forskningsspørsmålene (se punkt 1.3) diskuteres studiens funn sett i sammenheng med tidligere forskning og teori. Funn fra denne studien samsvarer i stor grad med funn fra tidligere forskning. I diskusjonen sammenlignes funn fra kvantitative og kvalitative analyser. Dette bidrar til en mer samlet og helhetlig forståelse av karakteristikker ved hverandrevurderingen høsten 2020. Der hvor kvantitative funn fremstår som overraskende utnyttes potensialet som ligger i å bruke metodetriangulering. Ved hjelp av kvalitative data fremmes en dypere og mer helhetlig forståelse av kvantitative funn som isolert sett kan fremstå som vanskelige å forklare.

### 6.1 Opplevelse av hverandrevurdering i undervisningen

#### 6.1.1 Rollene som giver og mottaker av vurdering

Denne studien er designet med studenter i rollene som både *giver* og *mottaker* av vurdering i likhet med de fleste andre studier om hverandrevurdering (Double et al., 2020). Funn fra kvantitative data tyder på at flertallet av studentene opplevde at det å *gi* hverandrevurdering var lærerikt. Andre studier har også vektlagt læringsprosessen som er forbundet med å *gi* feedback (Cartney, 2010; Nicol, 2012; Nicol, 2014; Panadero, 2016; Topping, 2005; To & Panadero, 2019; van Popta et al., 2017; Vickerman, 2009; Winstone et al., 2017). Når studentene gir feedback bidrar dette til egenvurdering fordi det å lese andres tekster setter i gang en prosess hvor studenten vurderer og reflekterer over eget arbeid.

Funn fra kvantitative data viser at førsteårsstudentene synes at de både ga og mottok god feedback til og fra sine medstudenter. Masterstudentene skiller seg fra førsteårsstudentene ved at det er et større sprik hvor studentene synes det er mer lærerikt å *gi* vurdering. Dette funnet finner en også i tidligere studier (Nicol, 2014; Nicol, 2012; Winstone et al., 2017).

Førsteårsstudentene skulle gi feedback med fokus på språklige og formelle feil i teksten. Underviseren begrunner dette med at ferske studentene trenger denne formen for feedback (*lower order feedback*). Masterstudentene skulle derimot gi feedback med fokus på tekstlig innhold (*higher order feedback*). Det som derimot skjedde var at masterstudentene også ga lower order feedback og begrunnet dette med at innholdet i teksten de skulle vurdere var

ukjent. Med dette forklares hvorfor førsteårsstudentene opplevde at å *få* feedback var nyttig, i motsetning til masterstudentene.

En annen faktor som påvirket masterstudentenes opplevde læring knyttet til å *motta* vurdering er undervisernes *innramming* av aktiviteten. Variabelen *innramming* har en signifikant positiv korrelasjon med opplevd læring, og i regresjonsanalysen har også *innramming* signifikant positiv påvirkning på opplevd læring. Dette kan forklares av at underviser for masterstudentene forklarer at hen i hovedsak hadde vektlagt læring knyttet til det å *gi* hverandrevurdering. Strategier for hvordan studentene kunne *ta i bruk* feedback de mottok ble ikke diskutert i undervisningen. Dette kan ha påvirket studentene slik at de ikke brukte like mye tid og energi på å ta i bruk feedback, og kan dermed være en mulig forsterkende faktor for hvorfor de opplevde mindre læring knyttet til å motta feedback.

### **Utfordringer med å *gi* hverandrevurdering**

Funn fra kvantitative og kvalitative data viser noe sprikende resultat når det gjelder hvordan studentene opplevde å *gi* vurdering. Kvantitative funn viser at over halvparten av førsteårsstudentene føler seg kompetent til å gi feedback på teksten til en medstudent. I intervjuet uttrykker førsteårsstudentene derimot usikkerhet knyttet til egen kompetanse i å vurdere andres tekster, noe som samsvarer med funn fra tidligere studier (Cartney, 2010; To & Panadero, 2019). En forutsetning for at studenter skal føle seg kompetent til å vurdere andres tekster, er at studentene innehar kunnskap om hva som kjennetegner en god akademisk tekst, samt eksplisitte og transparente vurderingskriterier (Andrade & Du, 2005; Black & William, 2009; Damsa et al., 2015; Reddy & Andrade, 2010), noe som var tydeliggjort i denne studien (se vedlegg 1). At førsteårsstudentene var usikker på egen kompetanse kan imidlertid forklares med funn fra tidligere forskning som viser at studenter trenger øvelse i å gi feedback til medstudenter for å bli trygg i rollen som giver av feedback (Rotsaert et al., 2018).

En annen utfordring som ble diskutert i intervjuet med førsteårsstudentene var at det kunne oppleves ubehagelig å vurdere andres tekster, fordi studentene ikke ønsket å være "*slem*". Lignende funn ble også gjort av Cartney (2010). Underviser for førsteårsstudentene presenterte studentene for kritikkbegrepet, og understreket at hverandrevurdering gjennomføres for å hjelpe hverandre til å bli bedre. Det er likevel mulig at studentene har

negative assosiasjoner til begrepet *vurdering*. Underviser for masterstudentene reflekterte også over vurderingsbegrepet og fortalte at hen ikke likte å bruke det ordet, og i stedet for valgte å bruke “medstudentkommentar”. Denne utfordringen understreker behovet for å gi tydelig informasjon om formålet med læringsaktiviteten, samt at studenter trenger trening i å vurdere både eget og andres arbeid for å bli trygg i rollen som giver av feedback.

Masterstudentene uttrykte derimot ingen bekymringer knyttet til å kritisere medstudenters tekster, og gir uttrykk for at dette er en naturlig del av å delta i et faglig fellesskap. Dette kan trolig skyldes at masterstudentene er mer inneforstått med *konvensjoner* i academia og kan videre forstås i sammenheng med studentenes ferdighetsnivå og kompetanse. Dersom studentene har lite erfaring med å vurdere tekster, vil det naturligvis være mer utfordrende å vurdere styrker og svakheter ved andres arbeid (Cartney, 2010; Liu et al., 2018).

Et interessant funn fra kvantitative data er at masterstudentene selv synes at de er mindre kompetente til å vurdere andre sine tekster sammenlignet med førsteårsstudentene, noe som også begrunnes med at teksten studentene skulle vurdere hadde ukjent tema. Dette står i strid med normative retningslinjer for hverandrevurdering, hvor det anbefales at studentene har kunnskap om det aktuelle fagområdet som de skal vurdere (Nicol, 2012; Nicol et al., 2014). Underviseren for masterstudentene omtalte dette som en misforståelse, og forklarte at hen hadde tatt det som en selvfølge at studentene måtte sette seg inn i tema for å kunne gi feedback på teksten til medstudenten. Denne problemstillingen understreker viktigheten av å formidle tydelige forventninger til studentene, samtidig som den viser hvilke uforutsette konsekvenser som kan oppstå. Med dette tatt i betraktning svarer masterstudentene likevel at de fikk informasjon om hverandrevurdering i forkant av arbeidet med den (3,71), og at de vet hvor de finner informasjon om hverandrevurdering (3,86).

### **Utfordringer med å *motta* hverandrevurdering**

Resultat fra kvantitativ analyse viser at det var en større prosentandel av masterstudentene som manglet strategier for å ta i bruk feedback sammenlignet med førsteårsstudentene. Dette kan ha påvirket hvordan studentene benyttet seg av feedbacken de mottok. Dette samsvarer med tidligere studier som viser lignende funn hvor studenter opplever at de ikke har tilstrekkelige strategier for hvordan de kan bruke feedback (Esterhazy, 2018; Esterhazy et al., 2020; Winstone et al., 2017). Det er imidlertid interessant at førsteårsstudentene rapporterer om at de i større grad synes at de har strategier for å ta i bruk feedback tatt i betraktning at

masterstudentene er kommet lenger i studieløpet og dermed har flere års erfaring fra høyere utdanning. Med bakgrunn i dette ville det være naturlig å anta at flertallet av masterstudentene opplevde å ha strategier for hvordan de skal ta i bruk feedback. En skal samtidig ta i betraktning at resultater fra spørreskjema for masterstudentene baserer seg på et lavt antall informanter, og en må dermed ta høyde for at formuleringer av påstander i spørreskjema kan gi rom for flere tolkninger (se punkt 4.4.1 om definisjonsmessig validitet). Påstanden *“Jeg har strategier for hvordan jeg skal benytte meg av tilbakemeldingen jeg fikk fra medstudent”*, kan tolkes av masterstudentene som at de ikke opplevde å ha strategier for å ta i bruk feedback de fikk til denne konkrete oppgaven. Flere av masterstudentene fortalte at feedbacken hadde begrenset nytteverdi fordi medstudenten ikke hadde kunnskap om temaet hen skriver om. Med bakgrunn i dette kan studentenes opplevelse av å ikke ha strategier for å ta i bruk feedback, være begrenset til denne konkrete situasjonen, fordi påstanden kan tolkes på ulike måter.

Både førsteårs- og masterstudenter diskuterer i intervjuene ulike betenkeligheter ved å motta feedback fra medstudenter. Førsteårsstudentene gir inntrykk av at de stoler mer på feedback gitt av underviser, og at de er bekymret for å overta misoppfatninger. Samtidig blir det også sagt av en førsteårsstudent at dersom medstudenten er på samme faglig nivå som seg selv, er det nyttig å motta feedback fra en medstudent. Flere studier som har undersøkt kvaliteten på feedback som studenter gir, viser at feedback fra student holder like høy kvalitet som feedback gitt av lærer (Topping, 1998; Zheng et al., 2018). Det er også gjort studier som viser at hverandrevurdering faktisk er mer læringsfremmende enn vurdering fra lærer (Double et al., 2020). Underviser for førsteårsstudentene mener at feedback fra medstudent ikke kan erstatte feedback fra lærer, men kan gi en annen type feedback. Hen mener at hverandrevurdering åpner for at studentene kan gi hverandre feedback på tekstlige feil og mangler som for eksempel skrivefeil og referansefeil. Dette er momenter ved tekster som ansatte ved universitetet ikke har kapasitet til å prioritere å rette, men som underviseren mener at studentene trenger feedback på. Masterstudentene uttrykker i intervjuene at de i utgangspunktet er positive til det å bli vurdert av medstudenter, men opplevde hverandrevurderingen høsten 2020 som lite nyttig. Dette kan ha sammenheng med at masteremnet var en mindre studentgruppe sammenlignet med førsteårsemnet. Dette innebærer at de fleste masterstudentene kjenner hverandre, og det er trolig tryggere å stole på vurderingene til sine medstudenter. Videre hadde de en viss forsikring om at medstudentene

hadde akademisk kompetanse da de allerede har fullført en bachelorgrad. Masterstudentene er trolig også mer inneforstått med *normer og konvensjoner* i academia, hvor fagfelleevaluering er en naturlig del av det faglige fellesskapet

Funn fra kvantitative data viser derimot at flere av masterstudentene opplevde lite nytte av å motta feedback fra medstudent høsten 2020. *Ingen* av masterstudentene som besvarte spørreskjemaet mente at det i *svært stor* eller *stor grad* var lærerikt å få feedback fra medstudent. Under intervjuet kommer det tydelig frem at dette har bakgrunn i at studentene skriver tekster om ulike tema. En av studentene fortalte at hen hadde vært heldig fordi hen mottok feedback fra en student som hadde likt tema. Det førte til at denne studenten mottok feedback som gikk mer i dybden på det faglige innholdet. Dette kan karakteriseres som en form for higher order feedback. De som derimot fikk feedback fra en student som ikke hadde kunnskap om temaet mottok derimot feedback omhandlet språk og struktur i teksten. Denne formen for feedback kan karakteriseres for lower order feedback, og opplevdes som overfladisk og lite nyttig for masterstudentene. Det samme fant Cartney (2010) i sin studie. Samtidig er det grunn til å tro at feedback som omhandlet språklige feil og mangler er mer nyttig for førsteårsstudentene da de har mindre erfaring i å skrive akademiske tekster.

### 6.1.2 Opplevd nytthet

Under intervjuene sier både studenter og undervisere at hverandrevurdering kan ha fordeler som de vil dra nytte av senere i studiet eller i arbeidslivet, og det trekkes en parallell mellom hverandrevurdering og fagfelleevaluering i academia. I intervjuet forteller førsteårsstudentene at selv om det kan oppleves ubehagelig å vurdere andres tekster, men studentene uttrykker også forståelse for at det er noe de bør trene på. For studentene er det av betydning at tiden og arbeidet de bruker på hverandrevurdering vil forberede dem på kommende eksamener. Begge studentgruppene gir uttrykk for at de i perioder opplever en høy arbeidsbelastning og at de må prioritere hvor mye tid og arbeid som skal legges ned i hvert enkelt emne. Studentene ønsker at arbeidet de gjør skal være eksamensrelevant. Selv om det ikke ble eksplisitt uttalt at de strategisk velger oppgaver som gir dem best utbytte til eksamen, ble det fortalt at de måtte prioritere hvilke oppgaver som skulle ha mest fokus. Dette er likt det Snyder (1971) beskriver som “The Hidden Curriculum”, som beskriver hvordan studenter prioriterer å bruke tid på det som det er lurt å kunne til eksamen. Det fremstår derfor som at det å engasjere studentene i

oppgaver som ikke gir dem direkte utbytte til eksamen fortsatt er en aktuell problemstilling for undervisere.

## 6.2 Faktorer ved hverandrevurdering som påvirker opplevd læring

### **Peergrade og Canvasrubrikker**

Data fra spørreskjemaet viser at førsteårsstudentene synes at nettbasert hverandrevurdering i Peergrade fungerte godt. Funn fra kvalitative data samsvarer med kvantitative data, eksempelvis med uttalelsen “*Veldig lett forståelig og brukervennlig (...) så det er ikke der problemene ligger*”. Tidligere forskning viser også at nettbasert hverandrevurdering fungerer like godt som ikke-nettbasert (Double et al., 2020; van Popta et al., 2017). I lys av virksomhetsteorien forstår vi det slik at *artefakten* Peergrade er velegnet til å gjennomføre hverandrevurderingen.

Masterstudentene ga artefakten Canvasrubrikker lavere verdi enn førsteårsstudentene ga artefakten Peergrade. I intervjuet gir studentene uttrykk for at Canvas fungerte dårlig til formålet og kommer med eksempler på dårlig brukervennlighet. Det er imidlertid ikke grunn til å tro at Canvasrubrikker fungerte dårlig fordi det var nettbasert. Dette fordi Peergrade fikk høy verdi og masterstudentene gir dessuten uttrykk for at de er vant til å bruke internett og nettbaserte verktøy i studiene. Det som er interessant er at underviser for masterstudentene tok et bevisst valg om å bruke Canvasrubrikker, som var et kjent verktøy for studentene, og begrunnet dette med at hen ønsket at refleksjon og selve hverandrevurderingen skulle være “*i høysetet*”. Det som derimot skjedde var det motsatte, hvor verktøyet som ble brukt fikk økt oppmerksomhet av studentene, da studentene som brukte Peergrade gir verktøyet mindre oppmerksomhet enn hva masterstudentene gjorde.

### **Utforming av spørsmålsrubrikker**

Bruk av spørsmålsrubrikker for hverandrevurdering har i flere studier vist å ha en læringsfremmende effekt (Andrade & Ying, 2005; Reddy & Andrade, 2010; Reynolds-Keefer, 2010). Dette blant annet fordi rubrikkene klargjør lærerens forventninger, og fungerer som scaffolds underveis i skrivearbeid.



Når det gjelder formuleringen i spørsmålsrubrikkene svarer studentene i spørreskjemaet at de synes at spørsmålene var utformet på en *forståelig* måte. Dette betyr likevel ikke at studentene var enige i at utformingen av selve spørsmålene var god. I intervjuet kom det frem at førsteårsstudentene ønsket flere spørsmål som åpnet for å gi hverandre kritikk. En overvekt av spørsmålene i førsteårsstudentene sine rubrikker (se vedlegg 1) hadde detaljerte kriterier. Dette kan ha gjort at studentene opplevde rubrikkene mer som en sjekklister enn som et verktøy for læring, som Torrence (2007, referert i Nordrum et al., 2013,) kaller for vurdering *som* læring, fremfor vurdering *av* eller *for* læring. Lignende funn ble gjort i studien til Nordrum et al. (2013). Det advares mot at dette kan gjøre at studentene får en overdreven tillit til kriteriene, og dermed mister sin egen autonomi. Dette støttes av van Popta et al. (2017) som skriver at spørsmålene må oppfordre studentene til å gi *forklaringer*. Samtidig gir førsteårsstudentene uttrykk for at de tviler på egne ferdigheter i å gi hverandrevurdering, og at de heller ikke stoler på medstudenters vurderinger. Dette kan indikere at studentene ikke føler seg modne for å bare ha åpne spørsmålsformuleringer enda. De er på et stadie hvor de spesifikke spørsmålsformuleringene derimot fungerer som scaffolds beskrevet i flere studier (Andrade & Ying, 2005; Reddy & Andrade, 2010; Reynolds-Keefer, 2010). Spørsmålsrubrikkene bør bli mer åpne og øke i vanskelighetsgrad etter hvert som studentene blir tryggere i emnet (Rotsaert et al., 2018; Vygotsky, 1978). Høsten 2020 hadde derimot underviseren tatt et bevisst valg om å bruke spørsmålsrubrikker som utnytter muligheten som ligger i at studenter kan kommentere feil og mangler som underviser ikke har kapasitet til. Med det sagt burde det være minst ett åpent spørsmål som åpner for å gi kritikk i form av å kommentere åpenbare feil og mangler i teksten.

Underviser til masterstudentene hadde som formål at spørsmålsformuleringene skulle åpne for refleksjon. Studentene synes derimot at dette var vanskelig med den gjeldende spørsmålsutformingen, noe som også forklares av at spørsmålene var fagspesifikke om et tema som de ikke hadde kunnskap om. Tidligere studier har vist at forkunnskaper innenfor det aktuelle domenet påvirker utfallet av hverandrevurderingen (Misiejuk et al., 2021; Nicol, 2012; Nicol et al., 2014), noe som samsvarer med våre funn. Med utgangspunkt i dette mener vi at spørsmålsrubrikkene bør være utformet med utgangspunkt i studentenes forkunnskaper, eller organisert slik at en forsikrer seg om at studentene opparbeider seg forkunnskaper om temaet som de skal vurdere på forhånd.

## **Studentenes ferdighetsnivå og gruppesammensetning**

Statistiske analyser viser at studentenes *ferdighetsnivå/innstilling* er variabelen som har mest positiv påvirkning på studentenes opplevde læring. Dette betyr at studenter som opplever at de selv har høyt ferdighetsnivå/innstilling i høyere grad opplevde at å gjennomføre hverandrevurdering var lærerikt enn dem som selv synes at de har et lavere ferdighetsnivå. De kvantitative funnene står i motsetning til tidligere studier om studenter ferdighetsnivå og hverandrevurdering. Disse studiene finner at det er lavt- og middelspresterende studenter som har høyest utbytte av hverandrevurdering, mens de høyest presterende har lavest utbytte (Davies, 2000; Li, 2012; Li & Gao, 2016). I denne studien er imidlertid læring målt som *opplevd læring* fra selvrapporteringsdata, og en kan dermed ikke være sikre på at studentene som selv synes at de er høytpresterende faktisk er det og omvendt. Forklaringsvariabelen *ferdighetsnivå/innstilling* inneholder også spørsmål som ikke direkte handler om personlig ferdighetsnivå, men som derimot handler om *holdning* til å gjennomføre aktiviteten. Virksomhetsteoretisk gir dette mening, da studenter som har en positiv holdning til og opplever et formål med aktiviteten hverandrevurdering (*objekt*) synes at hverandrevurdering er lærerikt. Altså har de et felles formål og samme motivasjon som sine undervisere for å gjennomføre hverandrevurdering, noe som gjør at aktiviteten oppleves meningsfull.

I intervjuet ga førsteårsstudentene uttrykk for at de kvier seg for å ta imot feedback fra medstudenter. Noen sa derimot at de kunne ha flaks og motta feedback av en mer kompetent student, som for eksempel hadde studert tidligere og var erfaren med akademisk skriving. Dette var imidlertid vanskelig å avgjøre da noen av hverandrevurderingene ble gjennomført anonymt. Masterstudentene ga uttrykk for at de ønsket at studenten som skulle vurdere dem hadde kunnskap om det aktuelle temaet som teksten handlet om. Studentene uttrykte at de var redde for å overta medstudenter sine misforståelser, en bekymring som også deles av underviser for masterstudentene. Disse funnene viser at både førsteårs- og masterstudenter mener at de lærer av en *mer kompetent annen* (Vygotsky, 1978) fremfor en mindre kompetent medstudent.

Underviser for masterstudentene mente at inndeling i såkalte “basisgrupper” høsten 2020 hadde uheldige konsekvenser for gruppedynamikken. Noen av gruppene fungerte dårlig og underviseren sier at det var fare for at studentene innretter seg etter den svakeste i gruppen, og mener at dette er uheldig for de høytpresterende studentene. Dette kan ha hatt påvirkning på

hverandrevurderingen. Studentene hadde gruppediskusjoner om fagstoffet før det ble gjennomført hverandrevurdering. Dette kan ha hatt påvirkning på forkunnskapene til studentene før gjennomføring av hverandrevurdering. Dette forstås også i sammenheng med begrepet en *mer kompetent annen*. Altså kan de lavest presterende studentene ha lært mye av å diskutere med studenter som er flinkere enn dem selv, men ikke motsatt.

### **Anonymitet**

Double et al. (2020) sin metastudie om hverandrevurdering viser at anonymitet ikke utgjør en signifikant effekt på læringsutbytte. Det samme finner Liu et al. (2018) i sin casestudie. I denne studien er læring målt som opplevd læring fra selvrapporteringsdata via spørreskjema og fokusgruppeintervju. En kan derfor ikke måle om anonymitet utgjør en læringseffekt, men en kan derimot diskutere studentenes syn på anonymitet og faktorer som påvirker hvordan de betrakter anonymitet kontra ikke-anonymitet i hverandrevurdering.

Funn fra kvalitative data viser at førsteårsstudentene er veldig opptatt av anonymitet i hverandrevurdering. Noen studenter forklarer at anonym hverandrevurdering er positivt og begrunner dette blant annet med at det gjør det enklere å være ærlig. Videre sier studentene at det oppleves mindre angstfylt at andre leser teksten din dersom du er anonym, noe som samsvarer med funn fra tidligere studier (Cartney, 2010; Panadero, 2016; To & Panadero, 2019). En av studentene knytter dette til eget modenhetsnivå. Hen sier at hen ikke føler seg moden nok enda, men at hen er inneforstått med at ikke-anonymitet er en hindring å komme over. Ingen andre studenter hadde innvendinger mot dette.

Masterstudentene har derimot ikke et ønske om anonymitet. I en av obligene skulle studentene være anonyme, men flere sier at de likevel skrev navn på oppgaven slik at mottaker kunne spørre giver om oppklaring dersom noe var uklart. Studentene uttrykker også et ønske om å diskutere hverandres tekster i mindre grupper, et format som forutsetter ikke-anonymitet. Videre sier studentene at ikke-anonymitet legger et positivt press deltakerne fordi en ønsker at medstudentene skal oppfatte en som kompetent og hardtarbeidende.

Den sprikende oppfatningen om anonymitet mellom førsteårs- og masterstudenter kan trolig knyttes til studentenes ferdighetsnivå som er diskutert tidligere. Videre forklaring kan være at masterstudentene er mer kjent med konvensjoner og normer i academia, og derunder

fagfelleevaluering. En tredje forklaring kan knyttes til antallet studenter i emnet. Førsteårssemnet hadde 165 studenter som gjennomførte hverandrevurdering, mot 21 masterstudenter. Videre var det i masteremnet blitt gjennomført flere økter med klasseromsdiskusjon i mindre grupper hvor studentene skulle bli kjent med hverandre. Masterstudentene gir uttrykk for at studiekulturen i emnet er god og at det er et uformelt miljø. Dette kan til sammen forklare hvorfor førsteårsstudentene ønsker anonymitet, men ikke masterstudentene. Med utgangspunkt i dette tror vi at en god organisering når det gjelder anonymitet for førsteårsstudentene hadde vært lignende som i studiet til Rotsaert et al. (2018). I denne studien gjennomførte studentene hverandrevurdering med en gradvis overgang fra anonym til ikke-anonym feedback, noe som viste seg å være vellykket. Dette forstår vi i sammenheng med *scaffolding* (Wood et al., 1976) hvor læreren bygger stillas rundt den lærende gjennom anonymisering som gradvis blir tatt vekk (*fading*) (Collins et al., 1989; Van de Pol et al., 2010).

### **Format**

Hverandrevurdering kommer i ulike format. Hverandrevurderingen høsten 2020 ble gjort skriftlig og asynkront da studentene skulle gi skriftlig feedback på skriftlig arbeid, og denne ble sendt over nett. Dette ble gjort gjennom å fylle ut rubrikker. I Double et al. (2020) sin metastudie kommer det frem at skriftlige komponenter i hverandrevurdering er effektivt, men utgjør ikke en signifikant forskjell sammenlignet med hverandrevurdering uten skriftlige komponenter. Double et al. (2020) sin metastudie finner imidlertid at bruk av rubrikker i hverandrevurdering ikke utgjør en signifikant forskjell sammenlignet med mer åpne design. En studie som ikke viste noen signifikant forskjell i bruk av rubrikk eller ikke, er studien til Green og Browser (2006). I denne studien er det ikke sammenlignet ulike format med hverandre, og en kan dermed ikke si noe om hvilket format som var mest effektivt. Derimot kan en gjennom funn fra kvantitative og kvalitative data bedømme hvorvidt studentene selv synes at de ulike aspektene knyttet til format ga opplevd læring. Når det gjelder *synkronitet* viser Liu et al. (2018) sin studie at studentene som gjennomførte synkron hverandrevurdering gjorde det bedre enn studentene som gjorde hverandrevurdering ikke-synkront. Dette begrunnes i hovedsak av at synkronitet åpner for å stille hverandre spørsmål om vurdering gitt/mottatt og slik oppklare misforståelser.

Førsteårsstudentene nevner ikke et ønske om å diskutere hverandres tekster i grupper, noe som ville ha forutsatt ikke-anonymitet. Mange av studentene gir uttrykk for at de ikke er modne til å være ikke-anonyme, og det er dermed grunn til å tro at de heller ikke ønsker å diskutere teksten sin i grupper. Når det gjelder masterstudentene blir det sagt at de hadde ønsket en arena for å diskutere tekstene sine synkront med sine medstudenter. Dette begrunner de med at det ville åpnet for dialog, som igjen ville være nyttig for å oppklare misforståelser. Dette er den samme årsaken som Liu et al. (2018) finner i sin studie. Forskjeller mellom førsteårs- og masterstudenter tolker vi i sammenheng med studentenes modenhetsnivå som er drøftet tidligere. Masterstudentene har høyere ferdighets- og modenhetsnivå enn førsteårsstudentene på den måten at de er mer aksepterende og inneforstått med at både det å gi og det å få feedback er en naturlig del av å delta i et faglig fellesskap.

### **Timing og arbeidsbelastning**

Timing og arbeidsbelastning er et tema som går igjen i samtlige intervjuer med både studenter og undervisere. Når det gjelder timingen for vurdering ble den gitt før skrivearbeidet ble gjennomført. Hverandrevurderingen forstås dermed som *formativ* vurdering (Black & William, 2009; Gibbs, 2006; Hattie & Timperley, 2007). Feedbacken ble gitt relativt raskt etter at arbeidet var levert for vurdering, noe som øker sannsynligheten for at studenten tar i bruk feedbacken (Gibbs, 2006). I tillegg fikk studentene tilstrekkelig tid til å ta i bruk feedbacken i henhold til forskning av van den Berg et al. (2006). Feedbacken ble også mottatt raskt etter levering til medstudent. Hverandrevurderingen fulgte altså normative prinsipper for timing i formativ vurdering.

Studenter og undervisere sier at timingen for hverandrevurdering kom for tett opp mot andre arbeidskrav. Dette begrunnes delvis av den *totale arbeidsbelastningen* i emnet, men også totalt i studieprogrammet. Eksempelvis sier en av masterstudentene: “*For det var så tight gap på alt vi skulle gjøre. Så da prioriterer du jo de fremføringene og de tingene du selv skal ha*”. Det kan virke som at det at studentene selv opplever høy arbeidsbelastning i studiene gjør at de nedprioriterer hverandrevurderingen, som igjen påvirker utfallet av aktiviteten negativt. Også underviser for førsteåret ser ut til å anerkjenne at høy arbeidsbelastning er en utfordring for studentene: “*Vi har et problem, og det er jo at semesteret ikke er lengre*”. Med det sagt er nettopp det å gjøre seg kjent med *normer og konvensjoner* i academia en av grunnene til at en

skal gjennomføre hverandrevurdering, eller som underviser for førsteårsstudentene sier: “*lære å bli student*”. Dermed er det kanskje både forventet og positivt at førsteårsstudentene opplever høy arbeidsbelastning det første semesteret som student. Også i overgangen fra bachelor til masterstudier vil studentene oppleve økt arbeidsbelastning. Høyere arbeidsbelastning er altså noe som studentene må bli vant til. Så til tross for at studentene oppgir at den høye arbeidsbelastningen har negativ påvirkning på hverandrevurderingen har den kanskje samtidig positiv påvirkning på deres akademiske dannelse.

I litteraturen står det skrevet om design og tilrettelegging for feedback (Esterhazy & Damsa, 2017) og kursdesign av hverandrevurdering (Berg et al., 2006; van Popta et al., 2017), men ingen av disse trekker frem arbeidsbelastning som en faktor av betydning slik vi finner i denne studien. Vi har heller ikke funnet andre studier hvor faktoren arbeidsbelastning er studert eller diskutert.

### **Studiekultur**

Begge underviserne ser hverandrevurdering i sammenheng med *konvensjoner og fellesskap* i høyere utdanning. Formålet med hverandrevurderingen var blant annet at det skulle styrke det faglige fellesskapet i studentgruppene. Masterstudentene ga imidlertid uttrykk for at de allerede var godt integrert i den *sosiale praksisen* (Säljö, 2006) og hadde etablerte kollokviegrupper. Underviseren la også opp til muntlige gruppesamtaler i undervisningsøktene, såkalte “feedback encounters” i klasserommet beskrevet av Esterhazy (2018). Studiemiljøet i masterklassen karakteriseres som sosialt og åpent med kultur for faglige diskusjoner. Hverandrevurderingen skulle imidlertid gjennomføres skriftlig og asynkront. Slik vi tolker det opplevde disse studentene at det å delta i skriftlig hverandrevurdering faktisk *tok dem vekk* fra sine allerede etablerte sosiale praksiser. Dermed fikk hensikten med hverandrevurderingen fra prosjektgruppen og underviseren sin side motsatt effekt enn hva som var hensikten.

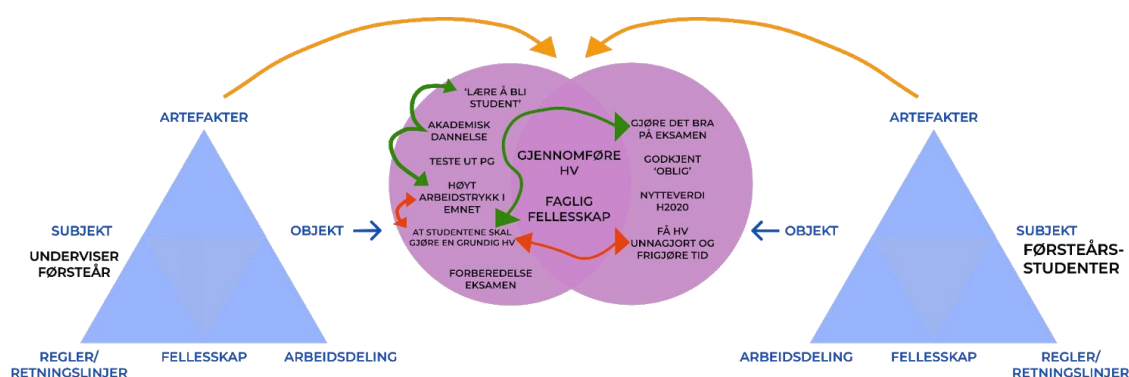
Disse funnene bekrefter at design av vurderingspraksiser må utformes i henhold til gjeldende kontekst for studiemiljøet (Esterhazy et al.; 2020; Esterhazy, 2018). Førsteårsstudentene opplevde i større grad at organiseringen med skriftlig og asynkron hverandrevurdering var velegnet. Dette forstår vi også i lys av studiekulturen. Førsteårsemnet hadde langt flere deltakere enn masteremnet, og studentene var nye på universitetet. Dessuten hadde studentene

startet midt i covid-19-pandemien og hadde dermed færre møteplasser for å bli kjent med sine medstudenter. Disse studentene opplevde den skriftlige hverandrevurderingen som en arena for faglig fellesskap i større grad enn masterstudentene som allerede var godt integrert i studiemiljøet. Altså passet designet bedre for konteksten til førsteårsstudentene.

### 6.3 Hvordan kan rubrikker for hverandrevurdering integreres som del av et større læringsdesign på en hensiktsmessig måte?

Ved hjelp av virksomhetsmodellen vil innrammingen av førsteårs- og masteremnet høsten 2020 diskuteres. Med utgangspunkt i dette fremstilles en modell for hvordan hverandrevurdering med rubrikker kan integreres i et læringsdesign. I modellene nedenfor er komponenter som harmonerer særlig godt markert med grønn pil. Komponenter som står i motsetning til hverandre, *contradictions*, (Engeström, 2010) til hverandre er markert i rødt. Komponenter som både harmonerer og står i motsetning til hverandre samtidig er markert med gul pil.

#### 6.3.1 Case 1: Hverandrevurdering for førsteårsstudenter høsten 2020



Figur 15. Modell av virksomhetssystemet for førsteårsemnet høsten 2020

Figur 15 viser hvordan de ulike komponentene står i gjensidig påvirkning til hverandre og påvirker gjennomføringen av hverandrevurdering utført i førsteårsemnet. Studenter og underviser i førsteårsemnet ser hverandrevurderingen i ulike tidsperspektiv. Underviseren ønsker at hverandrevurdering skal bidra til akademisk dannelse og fellesskap på lengre sikt, mens studentene ønsker at hverandrevurderingen skal hjelpe dem med nærstående obliger og

eksamener. Imidlertid anerkjenner både studenter og underviser at hverandrevurderingen vil være nyttig i et lang- og korttidsperspektiv, men vektlegger disse perspektivene i ulik grad.

*Artefakten* Peergrade var velfungerende for læringsaktiviteten (*objekt*). Både resultat fra spørreskjema og intervju støtter opp under funn som viser at studentene opplevde artefakten Peergrade som et brukervennlig verktøy for formålet. Studentenes opplevelse av Peergrades brukervennlighet kan betraktes i lys av virksomhetssystemets hierarkiske aktivitetsstruktur, se figur 3. Studentenes bruk av Peergrade har tilsynelatende fungert i tråd med det som kalles for *automatiske operasjoner*, ved at studentene ikke brukte energi på verktøyet. Et verktøy som er intuitivt å bruke, frigir tid til å fokusere på læringsaktiviteten hverandrevurdering, og er dermed velegnet som medierende artefakt fra studentenes side. Koordinator opplevde derimot Peergrade som lite brukervennlig, og måtte bruke lang tid for å integrere Peergrade med universitetet sine datasystemer. Formålet med 'Formativ vurderingspraksis i Canvas' var blant annet å teste ut nye verktøy for å forenkle hverandrevurdering og for koordinator fungerte dermed ikke Peergrade godt nok for dette da det førte til merarbeid.

Spørsmålsrubrikken fungerte delvis etter hensikten. Rubrikken var designet for å møte studentenes ferdigheter innenfor *den proksimale utviklingssonen*, og støttet studentene i arbeidet med å gi feedback til medstudent. Studenter opplevde rubrikken som begrensende i form av at den ikke åpnet for å gi konstruktiv kritikk. Det kan ha ført til tilfeller der studenter ikke fikk utnyttet sitt potensiale for å opptre som en *mer kompetent annen* (Vygotsky, 1978). Bruken av spørsmålsrubrikken og formålet med hverandrevurdering kan dermed beskrives som to komponenter som står i *motsetning* til hverandre. Dette ble oppdaget av underviser for førsteåret, og rubrikken ble revidert til neste runde med hverandrevurdering ved at det ble lagt til spørsmål som også åpnet for kritikk. Dette eksempelet illustrerer hvordan to motsetninger, i dette tilfellet relasjon *artefakt - objekt*, kan utvikle og endre eksisterende praksis. Den gule pilen mellom artefakt og objekt i figur 15, illustrerer dermed at artefakten Peergrade fungerte godt som medierende artefakt, mens artefakten spørsmålsrubrikken fungerte mindre godt.

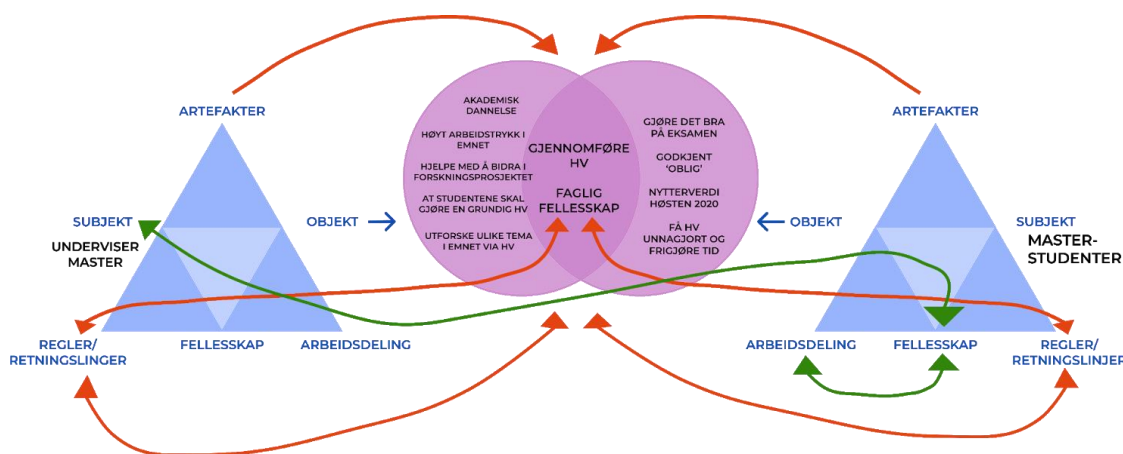
Det å ha høyt arbeidstrykk i emnet har grønn pil og står dermed i harmoni til det å lære å bli student og å opparbeide akademisk danning. I academia er det ønskelig at studentene bruker mye tid på studiet og tilhørende arbeidskrav, og dermed er det ønskelig med et høyt



arbeidstrykk. Dette går inn under det som underviser for førsteåret kaller for å “*lære å bli student*”, og på lengre sikt opparbeide det som blir kalt for akademisk danning.

Det å gjøre en grundig hverandrevurdering er et mål med aktiviteten, sett fra undervisernes side. Som vist i figur 15 er det både motsetninger og harmoni i relasjonen mellom det å gjøre en grundig hverandrevurdering og andre komponenter som inngår i aktiviteten. Å gjøre en grundig hverandrevurdering (*objekt for subjekt underviser*) harmonerer med å gjøre det bra på eksamen (*objekt for subjekt student*). Videre er det en rød pil i relasjon mellom å gjøre en grundig hverandrevurdering og et høyt arbeidstrykk (*objekt for subjekt underviser*) i emnet. Det uttrykker et motsetningsfylt forhold mellom det vi forstår som to av formålene til underviserne. Det å ha et høyt arbeidstrykk i ett fag kan gå på bekostning av å gjøre en grundig hverandrevurdering i et annet fag. Figur 15 viser også en motsetning i relasjonen mellom undervisernes *objekt* (å gjøre en grundig hverandrevurdering) og studentenes *objekt* (å få unnagjort hverandrevurdering og frigjøre tid). Ved at studentene prioriterer eksamen høyere enn andre arbeidskrav, førte dette til at andre arbeidskrav, for eksempel hverandrevurderingen, fikk lavere prioritet.

### 6.3.2 Case 2: Hverandrevurdering for masterstudenter høsten 2020



Figur 16. Modell av virksomhetssystemet for masteremnet høsten 2020

Figur 16 viser ulike komponenter som inngår i virksomhetssystemet for masteremnet. Masterstudentene har flere års erfaring fra academia og vi opplever at de i intervjuet viser faglig modenhet og forståelse for prinsippene bak hverandrevurdering (*objekt*). I

virksomhetssystemet for masteremnet oppstod det, slik vi oppfatter det, en motsetning mellom flere av komponentene.

Også i masteremnet finner vi at studenter og underviser ser formålet med hverandrevurderingen i ulike tidsperspektiv, men spriket mellom langtids- og korttidsperspektivet fremstår litt mindre sammenlignet med førsteårsemnet. Også dette henger sammen med faglig modenhet, hvor det virker som at masterstudentene er mer inneforstått med fenomenet fagfelleevaluering. Underviseren virker også å ha litt mer fokus på kommende eksamen, noe som vi tror skyldes at studentene i større grad har kommet lengre på veien til akademisk dannelse, og derfor ikke trenger like mye fokus på dette.

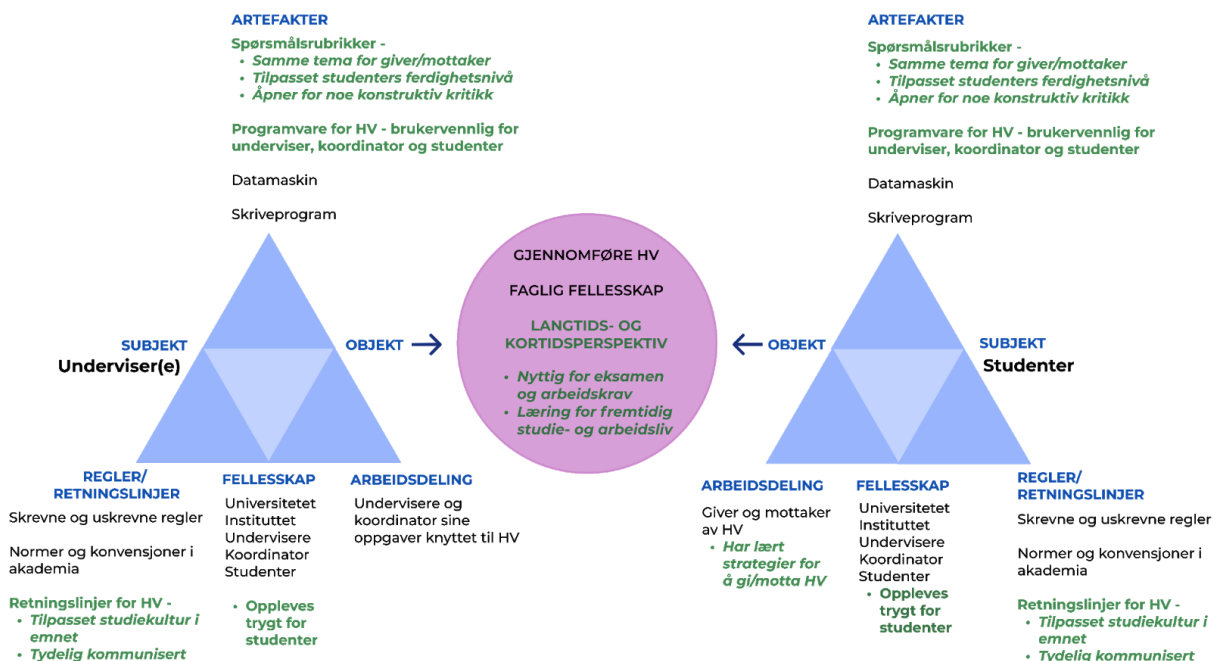
En av motsetningene som vi finner i virksomhetssystemet for masteremnet gjelder i relasjon mellom *regler/retningslinjer* og *objekt* som er å gjennomføre hverandrevurdering. Motsetningen uttrykkes i form av mangelfull eller utydelig informasjon om retningslinjer for hverandrevurderingen som ikke er forstått av studentene. Funn viser at den utydelige kommunikasjonen fikk uheldige konsekvenser for gjennomføringen av hverandrevurdering.

En annen motsetning som vi finner i virksomhetssystemet i figur 16, er i relasjon mellom *artefakt* og *objekt*. Canvas ble valgt fordi det var et kjent verktøy for studentene, og dermed var det ikke nødvendig for verken studenter eller underviser å bruke tid på å lære seg et nytt verktøy. Selv om Canvas i utgangspunktet var et kjent verktøy for studentene, var det første gang de brukte Canvasrubrikker. Funn viser at studentene ikke opplevde Canvas som særlig godt egnet til å mediere aktiviteten hverandrevurdering, da det oppstod utfordringer knyttet til brukervennlighet. Studentene var vant til å bruke digitale verktøy i undervisningen, og dermed burde bruken av verktøyet fungere som en *automatisk operasjon*, men problemer knyttet til brukervennlighet gjorde at det ble en bevisst *handling* i stedet for (se figur 3). Dette funnet står i kontrast til funn fra førsteårsemnet som brukte Peergrade og opplevde det som brukervennlig.

Figur 16 viser en grønn pil i relasjonen mellom underviser for masterstudentene og *felleskap* hos studentene. Denne er plassert fordi underviseren hadde klart å skape et positivt fellesskap mellom masterstudentene. I intervjuet sier underviseren at det i emnet var lagt til rette for faglig diskusjon i mindre grupper i undervisningsøktene der det var lav terskel for å diskutere

faglige problemstillinger. Studentene gir i intervjuet uttrykk for at de er enige i at emnet hadde godt faglig fellesskap og videre at de hadde ønsket at også hverandrevurderingen skulle finne sted innenfor rammene i det uformelle studiemiljøet som allerede var etablert. Figur 16 viser også grønn pil mellom *arbeidsdeling* og *fellesskap* hos studentene. Det gode faglige fellesskapet hadde positiv påvirkning på *arbeidsdeling* for studentene som består av å både gi og å motta hverandrevurdering. Fellesskapet i masteremnet opplevdes som trygt, og det opplevdes derfor mindre angstfylt å både gi og motta feedback til/fra medstudenter. Figur 16 viser rød pil mellom *regler/retningslinjer* og den delen av *objektet* som er *faglig fellesskap*. Retningslinjene for gjennomføringen av hverandrevurderingen stod i motsetning til etablert faglig fellesskap blant masterstudentene. Masteremnet hadde allerede et godt faglig fellesskap som utspilte seg i mindre formelle rammer enn den skriftlige hverandrevurderingen. Den samme motsetningen finner vi ikke hos førsteårsstudentene fordi disse enda ikke hadde etablert et faglig fellesskap i samme grad som masterstudentene. Dette viser at ulike *regler/retningslinjer* fungerer ulikt i ulike *fellesskap*.

### 6.3.3 Forslag til design av hverandrevurdering



Figur 17. Modellen viser et forslag for design av hverandrevurdering

Med utgangspunkt erfaringer gjort høsten 2020 i førsteårs- og masteremnet er det foreslått et nytt design av hverandrevurdering satt i virksomhetssystemet. Figur 17 viser et designforslag for hverandrevurdering hvor subjektene er underviser(e) og studenter. Det er lagt ved nye punkter under hver komponent som er markert i grønn og disse representerer endringsforslag. Med dette kan interne motsetninger i virksomhetssystemet som vi identifiserte høsten 2020 være drivkraft til endring og utvikling som beskrevet av Engeström (2010).

### **Artefakter**

Spørsmålsrubrikkene som skal besvares bør omhandle tema som er kjent for studentene slik at de har mulighet til å også gi feedback som omhandler innhold. For å sikre dette bør studentene pares med en medstudent som skriver om samme tema, *eller* organiseres slik at underviser(e) forsikrer seg om at giver av feedback har kunnskap om det aktuelle temaet. Videre bør spørsmålsrubrikkene være utformet slik at de *scaffolder* studentene, altså at de er tilpasset studentenes ferdighetsnivå. Minst én av spørsmålsrubrikkene bør åpne for å gi konstruktiv kritikk som omhandler valgfrie deler av teksten slik at studentene ikke blir tvunget til å gi feedback til en tekst uten å kunne påpeke åpenbare feil og mangler.

Når det gjelder bruk av artefakter viser det seg at bruk av datamaskin og skriveprogram fungerer som *automatiserte operasjoner*. Hverandrevurderingen bør gjennomføres med en programvare som er brukervennlig både for studenter, undervisere *og* koordinator.

### **Arbeidsdeling**

Når det gjelder *arbeidsdeling* i underviser(e) sitt system består denne av underviser(e) og koordinator(er) sine arbeidsoppgaver knyttet til å gjennomføre hverandrevurdering. For studentene består denne arbeidsdelingen av å være både giver og mottaker av vurdering. I en hverandrevurdering bør studentene inneha strategier for å både gi *og* motta feedback. Dette fordi utformingen av feedback og studentens engasjement med feedbacken de mottar påvirker studentens læringsutbytte.

### **Fellesskap**

Dersom den sosiale praksisen (Säljö, 2006) oppleves trygt for studentene er det trolig enklere å både gi og motta feedback til/fra medstudenter. Underviserne bør legge til rette for et studentaktivt læringsmiljø hvor vurdering og feedback er en naturlig del av den aktuelle

sosiale praksisen. Med det sagt tar det tid å etablere et faglig fellesskap blant studenter, og det vil dermed være vanskelig å etablere allerede første semester som student på universitet. I førsteårsemner det lurt å trygge studentene i det å gi og å motta feedback på andre måter, for eksempel slik som var gjort høsten 2020 hvor førsteårsstudentene ble presentert for en definisjon av kritikk-begrepet.

### **Regler/retningslinjer**

En forutsetning for å gjennomføre en vellykket hverandrevurdering krever at det formidles tydelige og klare forventninger og retningslinjer for arbeidet studentene skal gjøre. Videre må retningslinjene for arbeidet tilpasses studentenes ferdighetsnivå. Eksempelvis kan det å gjennomføre hverandrevurdering ved å være anonym, være en form for *scaffolding* for studenten. Etter hvert som studenten blir trygg på den sosiale praksisen hen tar del i samt får erfaring og utvikler forståelse for aktiviteten, vil studenten trenge mindre støtte. Etter hvert som studentene får erfaring med å gjennomføre hverandrevurdering vil målet være at studentene ikke har behov for å være anonym når de gir eller mottar feedback fra medstudent. En kan begynne med en gradvis *fading* hvor stillaset fjernes.

Læringsdesignet bør tilpasses en spesifikk kontekst, og unngå å lete etter én løsning som skal passe til alle sosiale praksiser. Dette gjelder også formatet, derunder gruppesammensetning, anonymitet og synkronitet/ asynkronitet.

En annen faktor som må tas høyde for når en planlegger for hverandrevurdering, er *timing*. Med det mener vi at feedback må mottas underveis i arbeidet, slik at studenten har mulighet til å ta i bruk og gjøre revideringer i arbeidet på bakgrunn av mottatt feedback. Dette er i tråd med det som defineres som *formativ vurdering*. Videre ønsker vi å understreke betydningen av studentenes opplevelse av den totale *arbeidsbelastningen* i emnet, men også i studieprogrammet som studenten gjennomfører. Når det gjelder total arbeidsbelastning i studieprogrammet blir ikke det tatt høyde for i denne virksomhetsmodellen, men det er likevel noe som kan påvirke studentenes innsats, tidsbruk og dermed også utfall av hverandrevurderingen.

## **Felles objekt**

I lys av virksomhetsteorien er en forutsetning for vellykket gjennomføring av hverandrevurdering at undervisere og studenter oppfatter et felles overordnet formål med aktiviteten. *Opplevd formål* blir altså en faktor som påvirker prosess og opplevd læring i hverandrevurdering. Det er ønskelig at både studenter og undervisere oppfatter at det å gjennomføre hverandrevurdering har samme formål. Det å gjennomføre hverandrevurdering samt å delta i faglig fellesskap bør være et felles mål for aktiviteten. En hverandrevurdering bør designes slik at både studenter og undervisere opplever at aktiviteten er nyttig *både* på kort *og* lang sikt. Med kort sikt menes at hverandrevurderingen skal oppleves nyttig med tanke på nærstående arbeidskrav og eksamen. Med lang sikt menes at studentene opplever at hverandrevurderingen gir læring som er nyttig for fremtidig studie- og arbeidsliv. Dersom studentene hovedsakelig prioriterer deloppgaver i hverandrevurdering som vil være nyttig på kort sikt, kan de gå glipp av viktig læring i et langtidsperspektiv. Ved at undervisere og studenter i større grad opplever felles *objekt* vil det trolig være lettere å nå det felles formålet.

## 7 Avslutning

Denne oppgaven har undersøkt hvordan studenter og undervisere ved et norsk universitet opplever å bruke hverandrevurdering med rubrikker i undervisningen. Videre er det undersøkt hvilke faktorer som påvirker prosessen og studentenes opplevelse av læringsutbytte i forbindelse med hverandrevurdering. For å undersøke dette er det benyttet metodetriangulering med kvalitative intervjuer og spørreskjema med Likert-skala.

### 7.1 Oppsummering av funn

Resultater fra denne studien underbygger potensialet for læring som finnes i en hverandrevurderingsprosess. Funn viser at førsteårsstudentene synes at å både gi og å motta hverandrevurdering er omtrent like lærerikt, mens masterstudentene oppgir at det å gi hverandrevurdering er mer lærerikt enn det å motta. Førsteårsstudentene er derimot mer skeptiske til gjennomføring av hverandrevurdering enn masterstudentene. Førsteårsstudentene begrunner dette med vanskeligheter med å stole på egen og medstudenters faglige kompetanse. Masterstudentene er mer positive til å gjennomføre hverandrevurdering. At masterstudentene er mer positive til hverandrevurdering forstås i sammenheng med studentenes ferdighetsnivå og deres kjennskap til normer og konvensjoner i akademia. Masterstudentene uttrykket derimot at hverandrevurderingen høsten 2020 var lite lærerik, og begrunnet dette med *innrammingen* rundt aktiviteten. Dette skyldes i hovedsak at studentene ikke hadde tilstrekkelig kunnskap om temaet for teksten de skulle vurdere og at hverandrevurderingen ikke var utformet i henhold til studiemiljøet i masteremnet. Masteremnet hadde allerede et etablert faglig fellesskap og studiemiljø og opplevde derfor at asynkron skriftlig hverandrevurdering tok dem vekk fra sine etablerte møteplasser. Skriftlig og asynkron hverandrevurdering for å fremme faglig fellesskap fungerer dermed bedre i andre studiekulturer med mindre etablert faglig og sosialt fellesskap, som for eksempel i førsteårsemnet.

Denne studien viser at enkeltfaktorer i et kursdesign har stor påvirkning på en hverandrevurderingsprosess. For det første er en viktig faktor utformingen av rubrikk, som hadde stor betydning for studentenes opplevelse av deltakelse i hverandrevurdering. Studien viser viktigheten av at den må tilpasses studentenes ferdigheter, samt åpne for konstruktiv kritikk. Førsteårsstudentenes spørsmålsrubrikk i første runde var utformet med fokus på å

oppfylle formelle krav i akademisk skriving, noe som fungerte etter hensikten. Studentene kunne dog ønske flere spørsmålsrubrikker som åpnet for konstruktiv kritikk.

Masterstudentenes spørsmålsrubrikk var utformet med hensikt om å åpne for kritisk tenking og refleksjon. Masterstudentene opplevde spørsmålsrubrikken som lite nyttig da den var utformet fagspesifikk om et tema som var ukjent for flere av studentene. Dette var en hovedfaktor som gjorde at masterstudentene ikke opplevde feedbacken de mottok som nyttig.

Med kvantitative data fra førsteårsstudentene ble det utført statistiske analyser.

Korrelasjonsanalyse viser statistisk signifikant positiv korrelasjon mellom den avhengige variabelen *opplevd læring* og samtlige forklaringsvariabler (*innramming, artefakter, ferdighetsnivå/innstilling*), hvor *ferdighetsnivå/innstilling* har høyest verdi. Dette forteller oss at alle variablene korrelerer positivt med studentenes opplevde læring høsten 2020. Ved multippel regresjonsanalyse hadde variabelen *ferdighetsnivå/innstilling* høyest positiv påvirkning på studentenes *opplevde læring*, men også *innramming* har signifikant positiv påvirkning. Dette forteller oss at studentene som føler seg kompetente til å både gi og motta feedback, brukte tilstrekkelig tid på oppgaven, samt var positivt innstilt opplevde hverandrevurdering som lærerikt.

De ulike aktørene som planlegger for og deltar i hverandrevurdering opplever både felles og ulike formål med å gjennomføre hverandrevurdering. Studien finner ulik vektlegging mellom korttids-og langtidsperspektiv mellom undervisere og studenter når det gjelder opplevd formål. Undervisere for både førsteårs- og masterstudenter ønsker at hverandrevurdering skal bidra til at studentene skal bli aktive deltagere i et faglig fellesskap. Underviser for førsteåret vektlegger også akademisk dannelse som et viktig mål med å gjennomføre hverandrevurdering. Det fremtrer et skille mellom undervisere og studenter, hvor underviserne vektlegger at hverandrevurdering gir læring og nytte i et langtidsperspektiv, mens studentene vektlegger nytthet knyttet til nærstående arbeidskrav og eksamener. Studentene anerkjenner likevel at hverandrevurdering også er nyttig i et langtidsperspektiv. Dette skillet er mindre fremtredende i masteremnet, hvor studentene har høyere ferdighetsnivå og er bedre kjent med normer og konvensjoner i academia.



## 7.2 Teoretiske implikasjoner - Hvordan samsvarer studien med funn fra tidligere forskning?

Funn fra denne studien samsvarer i stor grad med funn fra tidligere forskning.

Hverandrevurderingen var utformet i henhold til normative retningslinjer som inngår i formativ vurdering og feedbackpraksis (Black & Wiliam, 1998, Black & Wiliam, 2009; Hattie & Timperley, 2007; Sadler, 1989; Shute, 2008). Hverandrevurderingen kunne likevel vært bedre tilpasset den aktuelle konteksten som beskrevet av Winstone et al. (2017) og Esterhazy et al. (2020). Masterstudentene manglet kunnskap om det aktuelle fagområdet som de skulle vurdere, noe som påvirket opplevelsen av å delta i hverandrevurdering negativt. Dette sto i strid med normative retningslinjer for hverandrevurdering av Nicol (2012) og Nicol et al. (2014) som anbefaler at studentene har kunnskap om det aktuelle fagområdet som de skal vurdere. Tidligere forskning peker på at studenter mangler strategier for å ta i bruk feedback som de mottar (Esterhazy et al., 2020; Winstone et al., 2017).

Hverandrevurdering har vist seg å gi læring på tvers av kontekster (Double et al., 2020). Av kvalitative analyser fremkommer det at både undervisere og studenter knytter læringsutbyttet til de samme metakognitive prosessene som beskrevet i tidligere forskning (Nicol, 2012, Nicol, 2014; Panadero, 2019; Topping, 2005; van Popta et al., 2017; Vickerman, 2009). Tidligere forskning peker på at studentene kvir seg for å gi negativ feedback til medstudenter (Cartney, 2010). Dette funnet samsvarer med denne studiens funn når det gjelder førsteårsstudentene, mens masterstudentene ikke gir uttrykk for at de kvier seg for dette. Tidligere forskning viser at studenter stoler mer på vurdering gitt av underviser enn av medstudenter (Misiejuk et al. 2021; Panadero, 2016; To & Panadero, 2019; Weaver, 1995) noe funn fra både førsteårs- og masterstudenter bekrefter, selv om masterstudentene er mindre kritiske enn førsteårsstudentene.

Tidligere forskning (Double et al., 2020; Liu et al., 2018) viser at nettbasert hverandrevurdering er like effektivt som ikke-nettbasert, noe som samsvarer med funn fra denne studien. Heller ikke anonymitet har vist seg å påvirke læringsutbytte (Double et al., 2020), men casestudier viser imidlertid at studentene opplever det som en fordel å være anonym for å kunne gi en higher order feedback (Rotsaert et al., 2018) og dette samsvarer med hva noen av førsteårsstudentene forklarer. Tidligere forskning viser motstridende funn når det gjelder hvorvidt øvelse i å gi hverandrevurdering gir økt kvalitet (Double et al., 2020,

Misiejuk et al., 2021; Rotsaert et al., 2018) noe som ville vært interessant å undersøke våren 2021. Det er gjort studier som viser at høytpresterende studenter får lavest utbytte av hverandrevurdering (Davies, 2000; Li, 2012; Teekell, 1989) og disse funnene ses i sammenheng med funn fra denne studien som viser at studentene foretrekker å bli vurdert av underviser eller studenter som er flinkere enn dem selv.

Rubrikkene høsten 2020 var utformet med formål om å beskrive lærerens forventninger, tydeliggjøre læringsmål og bidra til at vurderingskriterier blir mer transparente som beskrevet av Andrade & Reddy (2010). Masterstudentene opplevde ikke at rubrikkene fungerte etter formålet og begrunnet dette med at de manglet kunnskap om det aktuelle temaet som de skulle vurdere. Førsteårsstudentene sine spørsmålsrubrikker var mer spisset og kunne også brukes som en sjekklister (Torrence, 2007, referert i Nordrum et al., 2013; Reynolds-Keefer, 2010). I denne studien tolkes ikke dette negativt fordi førsteårsstudentene trengte å bli kjent med formelle krav i akademisk skriving.

### 7.3 Svakheter ved studien og videre forskning

I denne studien baseres analysearbeidet på selvrapporderingsdata. Dette innebærer at det er studentenes subjektive oppfatninger som vurderes. En kan dermed ikke fastslå kvaliteten på feedback som studentene ga eller hvorvidt de tok denne feedbacken i bruk i de skriftlige innleveringene. Vi hadde vanskeligheter med å rekruttere nok studenter til å svare på spørreskjema som beskrevet i punkt 4.1.2. I både førsteårs- og masteremnet deltok færre informanter enn det som er ideelt. I førsteårsemnet var det få informanter som besvarte spørreskjemaet, dette førte til en noe for høy feilmargin i de statistiske analysene vi utførte. På grunn av få informanter fra masteremnet valgte vi å ikke gjennomføre statistiske analyser med dette datasettet, da feilmarginen ville blitt for høy, til tross for at de samme analysene ville ha vært interessante å gjennomføre. Også den deskriptive statistikken resulterte i svekket validitet grunnet få informanter. Koordinator for førsteåret uttrykker at hen opplever det krevende å nå frem til studentene med informasjon til tross for at den blir gitt gjentatte ganger i flere kanaler. Dette gir oss grunn til å mistenke at det er de mest engasjerte studentene som har fått med seg informasjonen og valgt å besvare spørreundersøkelsen. Dette kan kanskje ha hatt påvirkning på hvor godt informantene representerer målpopulasjonen. Videre var det uheldig at fokusgruppeintervjuet med førsteårsstudentene ble gjennomført av en underviser i emnet som de kjente til fra før. Dette er beskrevet i punkt 4.5.1.

Selv om det foreligger en betydelig mengde forskningsartikler vedrørende ulike aspekter ved hverandrevurdering, ser vi likevel et behov for videre forskning på enkelte områder hvor det i dag ser ut til å være et sprik i litteraturen. Ut fra denne studien viser det seg at de fleste studiene som undersøker fenomenet hverandrevurdering baserer seg på selvrapporteringsdata. Fremtidig forskning bør undersøke hvordan studenter faktisk engasjerer seg med og hvordan de benytter feedback de har mottatt. Videre bør faktoren arbeidsbelastning inkluderes i studier som undersøker hverandrevurdering, da det ser ut til å være en sentral faktor i denne studien. Det er grunn til å tro at det også kan ha hatt påvirkning på casestudiene referert til i litteraturgjennomgangen om hverandrevurdering med studenter i høyere utdanning. Med bakgrunn i det er det naturlig å tro at arbeidsbelastning har vært en påvirkende faktor i tidligere studier gjort i lignende kontekster med studenter som informanter. En årsak til dette kan være at emnene hvor hverandrevurderingen er gjennomført har studenter fra flere studieprogram, og det derfor er vanskelig å kartlegge arbeidsbelastning da den er individuell for hver student. For oss virker det som at det er et gap i litteraturen når det gjelder den totale arbeidsbelastningens påvirkning på en hverandrevurderingsprosess, da dette viser seg å være en faktor med stor betydning for både undervisere og studenter i denne studien.

Et fåtall av studiene om hverandrevurdering i høyere utdanning undersøker hverandrevurdering fra både underviser og studenters side, og det ser ut til å være et sprik i litteraturen når det gjelder studier som belyser hverandrevurdering fra et institusjonelt perspektiv.

# Litteraturliste

- AlFallay, I. (2004). The Role of Some Selected Psychological and Personality Traits of the Rater in the Accuracy of Self - and Peer Assessment. *System* (32), 407–425.  
<https://doi.org/10.1016/j.system.2004.04.006>
- Andrade, H. (2005). Teaching With Rubrics: The Good, the Bad and the Ugly. *College Teaching*, 53 (1), 27-31. <https://doi.org/10.3200/CTCH.53.1.27-31>
- Andrade, H. & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical assessment, Research and Evaluation*, Vol.10, Article 3.  
<https://doi.org/10.7275/g367-ye94>
- Bacchus, R., Colvin, E., Knight, B.C. & Ritter, L. (2020). When rubrics aren't enough: Exploring exemplars and student rubric co-construction. *Journal of Curriculum and Pedagogy*, 17(1), 48-61. <https://doi.org/10.1080/15505170.2019.1627617>
- Black, P. & William, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, Vol 5 (1). <https://doi.org/10.1080/0969595980050102>
- Black, P. & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*. 21(1), 5–31.  
<https://doi.org/10.1007/s11092-008-9068-5>
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V. & Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*.  
<https://doi.org/10.1080/14780887.2020.1769238>

- Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment and Evaluation in Higher Education*, 35 (5), 551-564. <https://doi.org/10.1080/02602931003632381>
- Collins, A., Brown, J. S. & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. I L. B. Resnick (Red.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (s. 453–494). Lawrence Erlbaum Associates, Inc
- Colvin, E., Bacchus, R., Knight, E. & Ritter, A. (2016). Exploring the way students use rubrics in the context of criterion referenced assessment. *Research and Development in Higher Education: The Shape of Higher Education*, 39, 45-52. HERDSA (Higher Education Research & Development Society of Australia Inc).
- Damsa, C., de Lange, T., Elken, M., Esterhazy, R., Fossland, T., Frølich, N., Hovdhaugen, E., Maassen, P., Nerland, M., Nordkvelle, Y., Stensaker, B., Tømte, C., Vabø, A., Wiers-Jenssen, J. & Aamodt, P.O. (2015). Quality in Norwegian Higher Education: A review of research on aspects affecting student learning. *Nordic Institute for Studies in Innovation, Research and Education (NIFU)*, 24, 3-64.
- Dancey, C. P. & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson/Prentice Hall.
- Davies, P. (2000). “Computerized Peer Assessment.” *Innovations in Education and Training International* 37 (4), 346–355. <https://doi.org/10.1080/135580000750052955>
- De Nasjonale Forskningsetiske komiteene. (2014, 10.oktober). *Betaling av forskningsdeltakere*. <https://www.forskningsetikk.no/ressurser/fbib/personvern/betaling-av-forskningsdeltakere/>
- De nasjonale forskningsetiske komiteene. (2018, 4. desember). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. <https://www.forskningsetikk.no/retningslinjer/hum-sam/forskningsetiske-retningslinjer-for-samfunnsvitenskap-humaniora-juss-og-teologi/>

- Double, K., McGrane, J. & Hopfenbeck, T. N. (2020). The Impact of Peer Assessment on Academic Performance: A Meta-analysis of Control Group Studies. *Educational Psychology Review*, 32 (1), 481- 589. <https://doi.org/10.1007/s10648-019-09510-3>
- Engeström, Y. (2005). Activity theory and individual and social transformation. Engeström, Y., Miettinen, R., & Punamaki, R-L (Red.), *Perspectives on Activity Theory*. 19-38. Cambridge University Press.
- Engeström, Y. (2010) Expansive Learning at Work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133-156. <https://doi.org/10.1080/13639080020028747>
- Esterhazy, R. (2018). What matters for productive feedback? Disciplinary practices and their relational dynamics. *Assessment & Evaluation in Higher Education*, 43 (8), 1302-1314. <https://doi.org/10.1080/02602938.2018.1463353>
- Esterhazy, R. (2019). A practical guide to feedback. How to give feedback that students actually use. *UiO. LINK - senter for læring og utdanning. F-link*, 03, 1-12. <https://www.uio.no/link/ressurser/rapporter/f-link/f-link-vol-03-feedback.pdf>
- Esterhazy, R. & Damşa, C. (2017). Unpacking the feedback process: an analysis of undergraduate students' interactional meaning-making of feedback comments. *Studies in Higher Education*, 44(2), 260-274. <https://doi.org/10.1080/03075079.2017.1359249>
- Esterhazy, R., Fosslund, T. & Stalheim, O. R. (2020). What Counts as Quality Feedback? Disciplinary Differences in Students' and Teachers' Perceptions of Feedback. *Quality work in Higher Education*, 155 - 174. [https://doi.org/10.1007/978-3-030-41757-4\\_9](https://doi.org/10.1007/978-3-030-41757-4_9)
- Johannessen, A., Tufte, P. A. & Christoffersen, L. (2010). Introduksjon til samfunnsvitenskapelig metode. Abstrakt forlag.

- Gibbs, G. (2006). How assessment frames student learning. *Innovative Assessment in Higher Education*. Bryan, C. & Clegg, K. (Red.), *Innovative Assessment in Higher Education*, (s. 23- 36). Routledge.
- Goodrich, H. (1997). Understanding rubrics. *Educational Leadership* 54 (4), 14–17.
- Green, R. & M. Bowser. (2006). Observations from the Field: Sharing a Literature Review Rubric. *Journal of Library Administration* 45 (1–2), 185–202.  
[https://doi.org/10.1300/J111v45n01\\_10](https://doi.org/10.1300/J111v45n01_10)
- Grønmo, S. (2016). *Samfunnsvitenskapelige metoder*. (2. utgave). Fagbokforlaget.
- Haanes, V. L. (2011). Dannelsesaspekter i utdanning. Rapport fra en arbeidsgruppe nedsatt av UHRs utdanningsvalg. *Universitets- og høgskolerådet (UHR)*, ( 4-41).  
<https://www.uhr.no/f/p1/iebabbb879-62ba-41c0-bb7a-2e89717edce2/dannelserapportendelig.pdf>
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*. 77 (1). <https://doi.org/10.3102/003465430298487>
- Howitt, D. & Cramer, D. (2017). *Understanding Statistics in Psychology with SPSS* (7th). Pearson.
- Huisman, B., Saab, N., van den Broek, P. & van driel, J.H. (2018). The impact of formative peer feedback on higher education students' academic writing: a Meta Analysis. *Assessment & Evaluation in Higher Education*, 44 (6)  
<https://doi.org/10.1080/02602938.2018.1545896>
- Kaptelinin, N. & Nardi, B.A. (2006). *Acting with technology. Activity Theory and Interaction Design*. Massachusetts Institute of Technology.
- Knight, P. T. (2002). The Achilles' Heel of Quality: The assessment of student learning. *Quality in Higher Education*, 8(1), 107–115.  
<https://doi.org/10.1080/13538320220127506>
- Kvale, S. & Brinkmann, S. (2012). *Det kvalitative forskningsintervju*. (2.utgave). Gyldendal Akademisk.

- Li, L. (2012). How do Students of Diverse Achievement Levels Benefit from Peer Assessment? *International Journal for the Scholarship of Teaching & Learning* 5 (2). <https://doi.org/10.20429/ijstl.2011.050214>
- Li, L. & Gao, F. (2016). The effect of peer assessment on project performance of students at different learning levels. *Assessment & Evaluation in Higher Education*, 41(6), 885-900. <https://doi.org/10.1080/02602938.2015.1048185>
- Liu, X., Li, L. & Zhang, Z. (2018). Small group discussion as a key component in online assessment training for enhanced student learning in web-based peer assessment. *Assessment & Evaluation in Higher Education*, 43(2), 207-222. <https://doi.org/10.1080/02602938.2017.1324018>
- Lundstrom, K. & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30–43. <https://doi.org/10.1016/j.jslw.2008.06.002>
- Meld. St. 16 (2016-2017). *Kultur for kvalitet i høyere utdanning*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/>
- Misiejuk, K., Wasson, B., & Egelandstad, K. (2021). Using learning analytics to understand student perceptions of peer feedback, *Computers in Human Behavior*, Vol 117. <https://doi.org/10.1016/j.chb.2020.106658>
- Nicol, D. (2012). Resituating feedback from the reactive to the proactive. I D. Boud & E. Molloy (Red.), *Feedback in Higher and Professional Education: Understanding it and doing it well* (s. 34–49). Routledge.
- Nicol, D. (2014). Guiding principles of peer review: Unlocking learners' evaluative skills. I C. Kreber, C. Anderson, N. Entwistle, & J. McArthur (Red.), *Advances and innovations in university assessment and feedback* (s. 195–258). Edinburgh University Press.
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), s. 102–122. <https://doi.org/10.1080/02602938.2013.795518>



- NOKUT. (2019). *Tidspress i undervisning og veiledning av studenter i høyere utdanning*.  
[https://www.nokut.no/globalassets/nokut/rapporter/ua/2019/amundsen\\_oygarden\\_tidspress-i-undervisning-og-veiledning-i-hoyere-utdanning\\_17-2019.pdf](https://www.nokut.no/globalassets/nokut/rapporter/ua/2019/amundsen_oygarden_tidspress-i-undervisning-og-veiledning-i-hoyere-utdanning_17-2019.pdf)
- NOKUT. (2021). *Studiebarometeret 2020 - Hovedtendenser. (Rapport 1/ 2021)*.  
[https://www.nokut.no/globalassets/studiebarometeret/2021/hoyere-utdanning/studiebarometeret-2020\\_hovedtendenser\\_1-2021.pdf](https://www.nokut.no/globalassets/studiebarometeret/2021/hoyere-utdanning/studiebarometeret-2020_hovedtendenser_1-2021.pdf)
- Nordrum, L., Evans, K., & Gustafsson, M. (2013). Comparing student learning experiences of in-text commentary and rubric-articulated feedback: strategies for formative assessment, *Assessment & Evaluation in Higher Education*, 38(8), s. 919-940,  
<https://doi.org/10.1080/02602938.2012.758229>
- Popham, J. W. (1997). What's wrong—and what's right—with rubrics. *Educational Leadership*, 55 (2), 72–75.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28 (1), 4–13. <https://doi.org/10.1002/bs.3830280103>
- Reddy, Y.M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448,  
<https://doi.org/10.1080/02602930902862859>
- Reynolds- Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research, and Evaluation*, Vol 15(8), 1-9. <https://doi.org/10.7275/psk5-mf68>
- Rotsaert, T., Panadero, E. & Schellens, T. (2018). Anonymity as an instructional scaffold in peer assessment: its effects on peer feedback quality and evolution in students' perceptions about peer assessment skills. *European Journal of Psychology of Education*, 33, 75–99. <https://doi.org/10.1007/s10212-017-0339-8>
- Sadler, D, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. <https://doi.org/10.1007/BF00117714>

- Schneider, J.F. (2006). Rubrics for Teacher Education in Community College. *The Community College Enterprise* 12, no. 1, 39–55.
- Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 67 – 70.
- Shute, V. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78, 153-189. <https://doi.org/10.3102/0034654307313795>
- Svartdal, F. (2021, 10.januar). *Fagfelle vurdering*. Store norske leksikon. <https://snl.no/fagfelle vurdering>
- Säljö, R. (2006). *Læring og kulturelle redskaper*. Oslo: Cappelen akademisk forlag
- Snyder, B. R. (1971). *The hidden curriculum*. Knopf.
- Tabak, I. & Baumgartner, E. (2004). The Teacher as Partner: Exploring Participant Structures, Symmetry, and Identity Work in Scaffolding. *Cognition and Instruction*, 22(4), 393-429. [https://doi.org/10.1207/s1532690Xci2204\\_2](https://doi.org/10.1207/s1532690Xci2204_2)
- Thagaard, T. (2018). *Systematikk og innlevelse*. En innføring i kvalitative metoder. (5.utgave). Fagbokforlaget.
- To, J. & Panadero, E. (2019). Peer assessment effects on the selfassessment process of first-year undergraduates. *Assessment & Evaluation in Higher Education*, 44 (6), 920-932. <https://doi.org/10.1080/02602938.2018.1548559>
- Topping, K. (1998). Peer Assessment Between Students in Colleges and Universities. *Review of Educational Research*, 68 (3), 249-276. <https://doi.org/10.3102/00346543068003249>
- Topping, K. (2005). Trends in Peer Learning. *Educational Psychology*, 25 (6), 631- 645. <https://doi.org/10.1080/01443410500345172>

- Ursachi, G., Horodnic, I. & Zait, A. (2015). How Reliable are Measurement Scales? External Factors with Indirect Influence on Reliability Estimators. *Procedia Economics and Finance*, 20, 679-686. [https://doi.org/10.1016/S2212-5671\(15\)00123-9](https://doi.org/10.1016/S2212-5671(15)00123-9)
- van den Berg, I., Admiraal W. & Pilot, A. (2006). Peer assessment in university teaching: evaluating seven course designs. *Assessment & Evaluation in Higher Education*, 31(1), 19-36. <https://doi.org/10.1080/02602930500262346>
- van den Bos, A. H. & Tan, E. (2019). Effects of anonymity on online peer review in second-language writing. *Computers & Education*, Vol 142, 1- 13. <https://doi.org/10.1016/j.compedu.2019.103638>
- Van De Pol J., Volman M. & Beishuizen J. (2010). Scaffolding in Teacher - Student Interaction: A Decade of Research. *Educational Psychology Review*, vol. 22 (3), 271-296. <https://doi.org/10.1007/s10648-010-9127-6>
- van Popta, E., Kral, M., Camp, G., Martens, R.L. & Simons, P.R.-J. (2017). Exploring the Value of Peer Feedback in Online Learning for the Provider. *Educational Research Review*, Vol 20, s.24- 34. *Educational Research Review*, 20, 24-34. <https://doi.org/10.1016/j.edurev.2016.10.003>
- Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34 (2), 221-230. <https://doi.org/10.1080/02602930801955986>
- Virkkunen, J. & Kuutti, K. (2000). Understanding organizational learning by focusing on “activity systems”. *Accounting Management and Information Technologies*, 10 (4), 291-319. [https://doi.org/10.1016/S0959-8022\(00\)00005-9](https://doi.org/10.1016/S0959-8022(00)00005-9)
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Weaver, M. E. (1995). Using peer response in the classroom: Students’ perspectives. *Research and Teaching in Developmental Education*, 12 (1), 31–37.

- Weurlander, M., Söderberg, M., Scheja, M., Hult, H. & Wernerson, A. (2012) Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37 (6), 747-760. <https://doi.org/10.1080/02602938.2011.572153>
- Winstone, N. E., Nash R. A., Parker, M. & Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist*, 52(1), 17-37. <https://doi.org/10.1080/00461520.2016.1207538>
- Xiao, Y. & R. Lucking. (2008). The impact of two types of peer assessment on students' performance and satisfaction with a wiki environment. *Internet and Higher Education*, 11 (3–4), 186–93. <https://doi.org/10.1016/j.iheduc.2008.06.005>
- Yin, R. K. (2009). *Case study research: Design and methods* (4. utgave). Sage.
- Zheng, L., Cui, P., Li, X., & Huang, R. (2018) Synchronous discussion between assessors and assessees in web-based peer assessment: impact on writing performance, feedback quality, meta-cognitive awareness and self-efficacy, *Assessment & Evaluation in Higher Education*, 43 (3), 500-514, <https://doi.org/10.1080/02602938.2017.1370533>

# Vedlegg

**Vedlegg 1:** Spørsmålsrubrikker i førsteårsemnet og masteremnet høsten 2020

**Vedlegg 2:** Spørreskjemaer

**Vedlegg 3:** Intervjuguider for undervisere, koordinator og studenter

**Vedlegg 4:** Operasjonalisering av variabler og skalaer for måling av indre konsistens i spørreskjemaene

**Vedlegg 5:** Eksempel på koder i tematisk analyse

**Vedlegg 6:** Godkjenning av prosjekt NSD

**Vedlegg 7:** Samtykkeskjemaer

## Vedlegg 1: Spørsmålsrubrikker i førsteårsemnet og masteremnet høsten 2020

### Rubrikker førsteårsemnet

Oblig 1	Oblig 2
Oblig1 (utkast) Levert: 177 av 249 Gjennomført hver.vurd 165 av 177	Oblig2 (utkast) Levert: 175 av 249 Gjennomført hver.vurd 164 av 175
Besvarelse og oppgavetekst. Gjør besvarelsen det oppgaveteksten ber om?	Beskriv kort innholdet i teksten (minimum 100 ord)
Inneholder besvarelsen et sammendrag av: a) Kapittel 2 i Kvello om transaksjonsmodellen? (ja/nei)	Nevn noe du synes var bra i besvarelsen (min 100 ord)
Inneholder besvarelsen et sammendrag av: b) Kapittel 11 «Social Cognitive Views of Learning and Motivation» i Educational Psychology (Woolfolk, 2019)? (ja/nei)	Nevn noe du synes kan forbedres i besvarelsen (min 100 ord)
Inneholder besvarelsen et sammendrag av: c) En valgfri tekst fra pensum? (ja/nei)	Hva lærte du av å gjøre egen-/hverandrevurdering denne gangen? (min 100 ord)
Er besvarelsen på 750 ord eller færre? (ja/nei)	
Har besvarelsen en overordnet overskrift? (ja/nei)	
Har hvert sammendrag en overskrift? (ja/nei)	
Har besvarelsen innledning? (ja/nei)	
Har besvarelsen topptekst med navn, emnekode og ordantall?	

(ja/nei)	
Har besvarelsen litteraturliste med tre referanser? (ja/nei)	
Inneholder besvarelsen studentens egne meninger? (ja/nei)	
Består besvarelsen av løpende tekst? (ja/nei)	
Er referansene i besvarelsen i henhold til <u>APA - malen</u> ? (ja/nei)	
Inneholder besvarelsen et rimelig utvalg av elementer fra kildetekstene? (Gjengir sammendraget sentrale poeng fra teksten den baserer seg på?) (ja/nei)	
Er besvarelsen skrevet med skrifttypen Times New Roman 12 pkt? (ja/nei)	
Refleksjon	
Beskriv kort de viktigste poengene i de tre sammendragene (min 150 ord)	
Nevn noe du synes var bra i besvarelsen (min 100 ord)	
Hva lærte du av å gjøre egen-/hverandrevurdering? (min 150 ord)	

## Eksempel på rubrikk i masteremnet

### Hvordan kommentere på din medstudents sammendrag?

Hvordan kommentere på din medstudents sammendrag?	
Kriterier	Vurderinger
1. Er referansen til teksten korrekt angitt ?	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
2. Din kommentar til hvordan sammendraget angir tekstens tema	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
3. Din kommentar til hvordan sammendraget beskriver tekstens sentrale budskap	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
4. Din kommentar til hvordan sammendraget peker på hvilke argumenter forfatteren bruker for å underbygge sitt budskap	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
5. Din kommentar til hvordan sammendraget presiserer tekstens sentrale begreper og hva som er forholdet mellom disse begrepene	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.



## Vedlegg 2: Spørreskjemaer

### Spørreskjema til førsteårsstudenter

Merknad: Førsteårsstudentene gjennomførte også egenvurdering høsten 2020, og spørreskjema inneholder derfor spørsmål som omhandler dette. Disse spørsmålene har vi ikke brukt i våre analyser.

---

I denne spørreundersøkelsen ønsker vi å undersøke hvordan du opplevde bruken av hverandrevurdering i xxxx. Med hverandrevurdering mener vi da du brukte verktøyet Peergrade til å gi tilbakemelding til din medstudent.

Spørreundersøkelsen er helt anonym, og en del av prosjektet "Formativ vurderingspraksis i Canvas". Det er frivillig å delta, men vi håper du kan sette av 3-4 minutter å gjennomføre denne undersøkelsen.

---

- Da du gjennomførte hverandrevurdering - hvor lang tid brukte du på å lese, evaluere og kommentere besvarelsen til din medstudent?

Under 10 minutter

Under 20 minutter

Under 30 minutter

Over 40 minutter

### Peergrade

Her vil vi høre om din erfaring med hverandrevurderingsverktøyet Peergrade.

*Skalaen for hvert spørsmål er fra 1 - 5.*

*1 = I svært liten grad, 5 = I svært stor grad*

- Jeg opplevde Peergrade som brukervennlig

- Jeg opplevde tekniske problemer da jeg gjennomførte hverandrevurdering i Peergrade
- Jeg synes at spørsmålene var formulert på en forståelig måte.
- Jeg fikk tilstrekkelig med opplæring i bruken av Peergrade

## Egenvurdering

Med egenvurdering mener vi da du svarte på spørsmål om og evaluerte din egen tekst.

*Skalaen for hvert spørsmål er fra 1 - 5*

*1 = I svært liten grad, 5 = I svært stor grad*

- Jeg fikk informasjon om egenvurdering før jeg gjennomførte den.
- Jeg synes egenvurdering gjorde besvarelsen min bedre.
- Jeg synes egenvurdering er en god forberedelse til eksamen.
- Å gjennomføre egenvurderingen har gjort at jeg tenker annerledes om å skrive akademiske oppgaver
- Jeg forstår hvorfor jeg skal gjøre egenvurdering.
- Jeg synes det å gjøre egenvurdering var lærerikt.
- Jeg har gjennomført hverandrevurdering i Peergrade

## Hverandrevurdering

Med hverandrevurdering mener vi da du brukte verktøyet Peergrade til å gi tilbakemelding til din medstudent.

*Skalaen for hvert spørsmål er fra 1 - 5*

*1 = I svært liten grad, 5 = I svært stor grad*

- Jeg opplever at det i undervisningen har vært fokus på hvordan gjennomføre hverandrevurdering.
- Jeg fikk informasjon om hverandrevurdering i forkant av arbeidet med den.
- Jeg vet hvor jeg finner informasjon om hverandrevurdering.
- Jeg forstår hvorfor jeg skal gjøre hverandrevurdering.
- Jeg har diskutert hverandrevurdering i oblig 1 med mine medstudenter
- Jeg har diskutert hverandrevurdering i oblig 1 med min seminarleder/ foreleser

- Jeg var positivt innstilt til å gjøre hverandrevurdering før vi startet med det.
- Jeg er positivt innstilt til å gjøre hverandrevurdering etter å ha gjennomført den
- Jeg synes det å gi tilbakemelding til medstudent gjennom hverandrevurdering var lærerikt.
- Jeg synes det å få tilbakemelding fra medstudent gjennom hverandrevurdering var lærerikt.\*
- Jeg vet hva som kjennetegner en god tilbakemelding på en tekst
- Jeg føler meg kompetent til å gi tilbakemelding på medstudenters tekst
- Jeg mener tilbakemeldingen fra medstudent gjennom hverandrevurdering gjorde besvarelsen min bedre
- Jeg vil bruke tilbakemeldingene fra medstudent i arbeidet med neste oblig.
- Jeg opplever at å gi/få tilbakemelding til/fra medstudent har verdi
- Jeg har strategier for hvordan jeg skal benytte meg av tilbakemeldingene jeg fikk fra medstudent
- Jeg tror hverandrevurdering er en god forberedelse til eksamen
- Jeg ønsker at det skal bli obligatorisk å gi hverandrevurdering også i fagene jeg skal til våren
- Jeg ga en grundig og presis tilbakemelding til min medstudents tekst.
- Jeg fikk en grundig og presis tilbakemelding fra min medstudent.
- Har du forslag til hvordan vi kan forbedre egen- og hverandrevurderingen, eller kommentarer til hvordan den blir brukt i xxxx?

## Spørreskjema til masterstudenter

I denne spørreundersøkelsen ønsker vi å undersøke hvordan du opplevde bruken av hverandrevurdering i xxxx. Med hverandrevurdering mener vi da du i Canvas kommenterte på din medstudents sammendrag og sammenstilling av to pensumtekster.

**Spørreundersøkelsen er helt anonym**, og en del av prosjektet "Formativ vurderingspraksis i Canvas" ved IPED. Det er frivillig å delta, men vi håper du kan sette av 3-4 minutter å gjennomføre denne undersøkelsen.

- 
- Da du gjennomførte hverandrevurdering - hvor lang tid brukte du på å lese, evaluere og kommentere besvarelsen til din medstudent?

Under 10 minutter

Under 20 minutter

Under 30 minutter

Over 40 minutter

## Spørsmålsrubrikker i Canvas

*Skalaen for hvert spørsmål er fra 1 - 5*

*1 = I svært liten grad, 5 = I svært stor grad*

- Jeg opplevde hverandrevurdering i Canvas som brukervennlig.
- Jeg opplevde tekniske problemer da jeg gjennomførte hverandrevurdering i Canvas.
- Jeg synes at spørsmålene var formulert på en forståelig måte.

## Hverandrevurdering

Her ønsker vi informasjon om deres opplevelse av hverandrevurdering som prosess.

*Skalaen for hvert spørsmål er fra 1 - 5*

*1 = I svært liten grad, 5 = I svært stor grad*

- Jeg opplever at det i undervisningen har vært fokus på hvordan gjennomføre hverandrevurdering.
- Jeg fikk informasjon om hverandrevurdering i forkant av arbeidet med det.
- Jeg vet hvor jeg finner informasjon om hverandrevurdering.
- Jeg forstår hvorfor jeg skal gjøre hverandrevurdering.
- Jeg har diskutert hverandrevurderingen i oblig 2 med mine medstudenter
- Jeg har diskutert hverandrevurderingen i oblig 2 med min seminarleder/foreleser.
- Jeg var positivt innstilt til å gjøre hverandrevurdering før vi startet med det.
- Jeg synes det å **gi** tilbakemelding **til** medstudent gjennom hverandrevurdering var lærerikt.\*
- Jeg vet hva som kjennetegner en god tilbakemelding på en tekst
- Jeg føler meg kompetent til å gi tilbakemelding på medstudenters tekst
- Jeg mener tilbakemeldingen fra medstudent gjennom hverandrevurdering gjorde besvarelsen min bedre
- Jeg opplever at å gi/ få tilbakemelding til/ fra medstudent har en verdi
- Jeg har strategier for hvordan jeg skal benytte meg av tilbakemeldingene jeg fikk fra medstudent
- Jeg tror hverandrevurdering er en god forberedelse til eksamen.
- Jeg ønsker at det blir obligatorisk å gi hverandrevurdering også i fagene jeg skal ha til våren
- Jeg **ga** en grundig og presis tilbakemelding **til** min medstudent.
- Jeg **fikk** en grundig og presis tilbakemelding **fra** min medstudent
- Har du forslag til hvordan vi kan forbedre hverandrevurderingen, eller kommentarer til hvordan den blir brukt i xxxx?

## **Vedlegg 3: Intervjuguider for undervisere, koordinator og studenter**

### Intervjuguide for undervisere - førsteår og master

#### **Underviserens bakgrunn**

- Kan du fortelle litt om din faglige bakgrunn og erfaring? Utdanning, arbeid o.l.
- Hva er din rolle i kurset xxxx?
  - Hvor lenge har du vært ansvarlig for kurset?
- Hvordan vil du vurdere din egen digitale kompetanse?
- Hvilke holdning har du til bruk av teknologi/internett i undervisning i høyere utdanning?  
Evt: kan du utdype: hva mener du/har du erfart at er positive og negative aspekter ved bruk av læringsteknologi?
- Muligheter og utfordringer
- Bruker du andre digitale verktøy/ applikasjoner bruker du i undervisningen?  
Ev.hvilke?

#### **Pedagogiske refleksjoner**

- Hva mener du at kjennetegner en god vurderingspraksis?
  - Kan du utdype underviser og studentens rolle?
- «Studentene kan få mer nyttig veiledning av medstudenter enn hva de kan få av for eksempel av seminarledere og professorer»
  - Positive, negative sider.

#### **Hverandrevurdering - prosessen**

- Hva var formålet med hverandrevurderingen?
  - Hva var målet at studentene skulle lære av dette arbeidet?
  - I hvilken grad synes du at arbeidet med hverandrevurderingen (evt. også egenvurdering) fungerte som planlagt?
- Hva var din rolle i arbeidet med hverandrevurdering?
  - Hvordan støtte studentene i prosessen?

#### **Kursets design / kontekst**

- Kan du fortelle litt om informasjonen dere ga studentene vedrørende hverandrevurderingen?
  - Hvordan gi hverandrevurdering/egenvurdering (kun underviser førsteår)?
  - Strategier for å ta i bruk tilbakemeldinger.
- Hvordan planlagte du for vurdering/tilbakemelding i kurset?
- Hvilke undervisningsformer har dere brukt i kurset utenom egen/hverandrevurdering, og hvilke pedagogiske overveielser har du tatt når du valgte disse?
  - Har dere noen føringer/ retningslinjer fra instituttets side om hvilke undervisningsmetoder dere bør bruke?
  - Gjelder noen av disse vurdering av studentarbeid?
- “Jeg opplever at det er kultur for at undervisere på instituttet diskuterer faglige og pedagogiske problemstillinger”
- Har dere i kollegiet diskutert bruk av læringsteknologi, og har dere felles forståelse av hva som kjennetegner hensiktsmessig bruk?
  - Hva opplever du skal til for å få det til å fungere optimalt?

## **Bruk av digitale verktøy**

### **Førsteår**

- “Jeg synes det fungerte godt å gjennomføre hverandrevurdering i Peergrade”.
  - Fordeler, ulemper.
  - Added value.
  - Hvorfor ble akkurat verktøyet Peergrade brukt?

### **Master**

- Hvorfor valgte du å gjennomføre opplegget i Canvas fremfor å bruke Peergrade?
- “Jeg synes det fungerte godt å gjennomføre hverandrevurdering i Canvas”.

## **Sluttevaluering**

- Hvilke tilbakemeldinger har du fått fra studentene på arbeidet med egen/hverandrevurdering?
- Intensjon vs. faktisk utfall. Har det fungert etter planen?
- På hvilken måte opplever de at det har vært nyttig/ mindre nyttig for studentene?

- Hvis du skulle ha designet opplegget på nytt, ville du ha gjort noe annerledes, evt. hva?

## Intervjuguide for koordinator i prosjektet

### Ansattes bakgrunn

- Kan du fortelle litt om din faglige bakgrunn og erfaring? Utdanning, arbeid o.l.
  
- Hva er din rolle i kurset xxxx, og hva innebærer denne rollen?
  - Hvor lenge har du hatt denne rollen?

- Implementering (prosjektgjennomføring, verktøy og teknisk gjennomføring)

Hvem var ansvarlig for å sette opp spørsmålsrubrikkene?

### Hverandrevurdering - prosessen

- Kan du fortelle litt om formålet med å innføre hverandrevurdering og om forskningsprosjektet?
  - litt om kontekst og historien bak.
  - Er det et bevisst valgt at hverandrevurdering prøves ut på akkurat førsteårskull? Evt hvorfor?
- Hva har vært din rolle i prosjektet “Formativ vurderingspraksis i Canvas”?

### Kursets design / kontekst

- Kan du fortelle litt om informasjonen dere ga studentene vedrørende hverandrevurderingen?
  - I hvilke kanaler? Flere kanaler/bare én?
  - Fikk de informasjon om hvorfor de skulle ha hverandrevurdering (pedagogisk formål)?
- Finnes det noen føringer/ retningslinjer fra instituttets side om hvilke undervisningsmetoder undervisere ved iPed bør bruke?
  - Gjelder noen av disse vurdering av studentarbeid?

### Digitale verktøy



- Hvorfor ble akkurat verktøyet Peergrade valgt?
  - Hvorfor falt valget på akkurat Peergrade fremfor andre Canvas eller andre verktøy?
  - Hvordan har du/dere lært å bruke verktøyet Peergrade? Har dere funnet ut av det selv eller har det vært eksterne aktører inne for å vise dere?
  - Fordeler/ ulemper.
  - PG vs Canvas.
  - Fremtidig bruk av PG?
- Hvem var ansvarlig for å implementere PG i Canvas? Praktisk/ teknisk gjennomføring.

### **Vår 2020 vs. høst 2020**

Dere har gjennomført hverandrevurdering med to forskjellige kurs.

- Vil du gjenfortelle hva som fungerte godt/mindre godt forrige runde? (V20)
- Hva fungerte godt/mindre godt i siste runde? (H20)
- Hvordan vil du sammenligne de to rundene?
  - Hva var grunnen tror du til at de to rundene fikk forskjellig utfall?

### **Sluttevaluering**

- Hvilke tilbakemeldinger har dere fått fra studentene på arbeidet med egen/hverandrevurdering?
- Intensjon vs. faktisk utfall. Har det fungert etter planen?
  - Positive, negative erfaringer.
- På hvilken måte opplever du at det har vært nyttig/ mindre nyttig for studentene?
- Hvis dere skulle ha gjennomført lignende opplegg på nytt, ville du ha gjort noe annerledes, evt. hva?
- Hvilke erfaringer tar dere med dere videre fra arbeidet med hverandrevurdering i høst?
  - Kunne informasjonsflyten vært bedre?
- Hvilken betydning får forskningsprosjektet for Instituttet .
  - Hvordan skal dataene som er samles inn brukes videre av instituttet?

## Intervjuguide for studenter - førsteår og master

Merknad: førsteårsstudentene gjennomførte også egenvurdering høsten 2020. Dette er ikke fokus for denne oppgaven, men var del av intervjuguiden. Disse spørsmålene er markert.

### **Del 1: Hovedinndeling av spørsmålene for fokusgruppeintervju**

Under intervjuet kan det være greit å se til disse hovedpunktene.

- 1: Læringsprosessen
- 2: Hverandrevurdering- aktiviteten
- 3: Teknisk perspektiv

### **Start med påstander til diskusjon:**

Si: Jeg vil nå komme med en påstand som jeg ønsker at dere skal diskutere:

*Hverandrevurdering (oblig?)*

*a) «Man kan få mer nyttig veiledning av medstudenter enn hva man kan få av for eksempel av seminarledere og professorer»*

Kom med eksempler! Påstand B hvis behov.

*(b) «Medstudenter har for lite faglig innsikt til at de kan gi gode hverandrevurderinger»*

*"Egenvurderingen ga meg nyttige verktøy som jeg kan bruke til eksamensforberedelser"*

(kun førsteårsstudenter)

### **Hvordan opplevde du/dere om å GI hverandrevurdering?**

- I hvilken grad fant dere det motiverende? *Utdyp*
- Kan dere gi eksempler på ting dere syntes var bra med det?
- Kan dere gi eksempler på ting dere syntes var utfordrende med å gi hverandrevurdering?
- I hvilken grad lærte dere noe av å gi hverandrevurdering? *Utdyp*
- I hvilken grad fant dere det eksamensrelevant? *Utdyp*
- Hvordan kan du bruke det du har lært?
- På hvilken måte er å gi hverandrevurdering nyttig for arbeid senere i studiene? Hvordan og på hvilken måte? *Utdyp*

## Hvordan opplevde du/dere om å FÅ hverandrevurdering?

- I hvilken grad fant dere det motiverende? *Utdyp*

Oppfølgingsspørsmål: har du endret mening etter arbeidet med hverandrevurdering?

- Kan dere gi eksempler på ting dere syntes var bra med det?
- Kan dere gi eksempler på ting dere syntes var utfordrende med å få hverandrevurdering?
- I hvilken grad lærte dere noe av å få hverandrevurdering? *Utdyp*
- I hvilken grad fant dere det eksamensrelevant? *Utdyp*
- Hvordan kan du bruke det du har lært?
- På hvilken måte er å få hverandrevurdering nyttig for arbeid senere i studiene? Hvordan og på hvilken måte? *Utdyp*

## Påstand:

- *“Jeg ønsker at det skal bli obligatorisk å gi hverandrevurdering også i fagene jeg skal ha til våren”*
- *Oppfølgingsspørsmål: “På bakgrunn av erfaringene fra denne runden vil jeg gjennomføre neste hverandrevurdering på en annen måte”*

## Spørsmålene

(Ha klart skjermbilde av et eksempel på en rubrikk som en kan vise til studentene via skjermdeling)

Vise til førsteårsstudenter:

The screenshot shows a digital interface for a question and answer session. At the top, it is labeled 'S1' and contains the text 'Beskriv kort hovedpoengene i medstudentens sammendrag'. Below this, it is labeled 'Q1 · Text' and contains the same text: 'Beskriv kort hovedpoengene i medstudentens sammendrag'. Underneath the question, there is a label 'Response' followed by a large, empty rectangular text input area for the student's answer.

Vise til masterstudenter:

### Hvordan kommentere på din medstudents sammendrag?

Hvordan kommentere på din medstudents sammendrag?	
Kriterier	Vurderinger
1. Er referansen til teksten korrekt angitt ?	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
2. Din kommentar til hvordan sammendraget angir tekstens tema	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
3. Din kommentar til hvordan sammendraget beskriver tekstens sentrale budskap	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
4. Din kommentar til hvordan sammendraget peker på hvilke argumenter forfatteren bruker for å underbygge sitt budskap	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.
5. Din kommentar til hvordan sammendraget presiserer tekstens sentrale begreper og hva som er forholdet mellom disse begrepene	Dette området vil benyttes av den som vurderer for å gi kommentarer relatert til dette kriteriet.

### Hvordan opplevde dere å bruke rubrikken? (for eksempel når dere:

- ga hverandrevurdering?
- fikk hverandrevurdering?
- gjorde egenvurdering? (kun førsteårsstudenter)

### Bidro rubrikken til at dere reflekterte over akademisk oppgaveskriving på en annen måte enn tidligere?

- På hvilken måte? Hvorfor? Tenk høyt
- Hva var forskjellen i læringsutbytte av å gjøre egenvurdering vs. hverandrevurdering? (kun førsteårsstudenter)
- Var noen spørsmål mer utfordrende enn andre? Hvorfor? På hvilken måte?
- Hvordan synes dere det var å vurdere hvorvidt en medstudent hadde besvart oppgaven?
- Hva tenker dere om anonymitet?

- Hva var bra/mindre bra med spørsmålene i rubrikken?

**Påstand:**

- *“Jeg ønsker at det skal bli obligatorisk å ha hverandrevurdering også i emnene jeg skal ha til våren”* (kun førsteårsstudenter)

**Peergrade** (kun førsteårsstudenter)

Påstand: *“Jeg opplevde PG som brukervennlig”*

- Hvordan opplevde dere brukervennligheten til PG?
- Hvordan opplevde dere å bruke hverandrevurderingsverktøyet PG?

Gi eks. På noe som fungerte bra, og noe som ikke fungerte så bra.. Utdyp...

- *“Jeg fikk ikke nok opplæring i bruk av PG før hverandrevurderingen”*

**Canvas** (kun masterstudenter)

*“Jeg opplevde hverandrevurdering i Canvas som brukervennlig”*

Var det enkelt å finne frem? Vite hvor dere skulle klikke osv?

- Hvordan opplevde du å bruke rubrikk i Canvas som digitalt verktøy for hverandrevurdering?
- Var det brukervennlig å bruke Canvas slik dere gjorde for å gi hverandrevurdering? Var det enkelt å finne frem? Vite hvor dere skulle klikke osv?
- Opplevde noen tekniske problemer?

**Hva synes dere var mest lærerikt og hvorfor?** (Kun førsteårsstudenter)

Hverandrevurdering

Egenvurdering

Refleksjonsnotat

**Avslutning:** Hvordan ville en ideell hverandrevurdering vært?

## **Vedlegg 4:**

### ***Operasjonalisering av variabler***

#### ***Skalaer for måling av indre konsistens i spørreskjemaene***

I dette vedlegget står spørsmålene som inngår i hver skala for måling av indre konsistens (Cronbachs alfa) listet opp sammen med resultatet fra hver skala. Spørreskjemaet er kategorisert etter skalaene:

- Artefakt
- Studentenes ferdighetsnivå/innstilling
- Opplevd læring
- Innramming

#### **Spørreskjema førsteår**

##### **Bruk av artefakter**

3. Jeg opplevde Peergrade som brukervennlig
4. Jeg opplevde tekniske problemer da jeg gjennomførte hverandrevurdering i Peergrade
5. Jeg synes at spørsmålene var formulert på en forståelig måte.

$\alpha = 0,49$

##### **Studentenes ferdighetsnivå/innstilling**

25. Jeg vet hva som kjennetegner en god tilbakemelding på en tekst
26. Jeg føler meg kompetent til å gi tilbakemelding på medstudenters tekst
30. Jeg har strategier for hvordan jeg skal benytte meg av tilbakemeldingene jeg fikk fra medstudent
33. Jeg ga en grundig og presis tilbakemelding til min medstudents tekst.
34. Jeg fikk en grundig og presis tilbakemelding fra min medstudent.
21. Jeg var positivt innstilt til å gjøre hverandrevurdering før vi startet med det.
22. Jeg er positivt innstilt til å gjøre hverandrevurdering etter å ha gjennomført den

$\alpha = 0,79$

### **Opplevd læring**

23. Jeg synes det å gi tilbakemelding til medstudent gjennom hverandrevurdering var lærerikt.

24. Jeg synes det å få tilbakemelding fra medstudent gjennom hverandrevurdering var lærerikt.

27. Jeg mener tilbakemeldingen fra medstudent gjennom hverandrevurdering gjorde besvarelsen min bedre

28. Jeg vil bruke tilbakemeldingene fra medstudent i arbeidet med neste oblig

29. Jeg opplever at å gi/få tilbakemelding til/fra medstudent har verdi

31. Jeg tror hverandrevurdering er en god forberedelse til eksamen

32. Jeg ønsker at det skal bli obligatorisk å gi hverandrevurdering også i fagene jeg skal ha til våren

$\alpha = 0,89$

### **Innramming**

6. Jeg fikk tilstrekkelig med opplæring i bruken av Peergrade

15. Jeg opplever at det i undervisningen har vært fokus på hvordan gjennomføre hverandrevurdering. \*

16. Jeg fikk informasjon om hverandrevurdering i forkant av arbeidet med den.

17. Jeg vet hvor jeg finner informasjon om hverandrevurdering.

18. Jeg forstår hvorfor jeg skal gjøre hverandrevurdering.

19. Jeg har diskutert hverandrevurdering i oblig 1 med mine medstudenter

20. Jeg har diskutert hverandrevurdering i oblig 1 med min seminarleder/ foreleser

$\alpha = 0,79$

## Spørreskjema master

### Bruk av artefakter

2. Jeg opplevde hverandrevurdering i Canvas som brukervennlig.
3. Jeg opplevde tekniske problemer da jeg gjennomførte hverandrevurdering i Canvas.
4. Jeg synes at spørsmålene var formulert på en forståelig måte

$\alpha = 0,47$

### 2) Studentenes ferdighetsnivå/innstilling

15. Jeg vet hva som kjennetegner en god tilbakemelding på en tekst
16. Jeg føler meg kompetent til å gi tilbakemelding på medstudenters tekst
19. Jeg har strategier for hvordan jeg skal benytte meg av tilbakemeldingene jeg fikk fra medstudent
22. Jeg **ga** en grundig og presis tilbakemelding **til** min medstudent
23. Jeg **fikk** en grundig og presis tilbakemelding **fra** min medstudent
11. Jeg var positivt innstilt til å gjøre hverandrevurdering før vi startet med det
12. Jeg er positivt innstilt til å gjøre hverandrevurdering etter å ha gjennomført den

$\alpha = 0,79$

### Opplevd læring

13. Jeg synes det å gi tilbakemelding til medstudent gjennom hverandrevurdering var lærerikt.
14. Jeg synes det å få tilbakemelding fra medstudent gjennom hverandrevurdering var lærerikt.
17. Jeg mener tilbakemeldingen fra medstudent gjennom hverandrevurdering gjorde besvarelsen min bedre
18. Jeg opplever at å gi/få tilbakemelding til/fra medstudent har verdi
20. Jeg tror hverandrevurdering er en god forberedelse til eksamen
21. Jeg ønsker at det skal bli obligatorisk å gi hverandrevurdering også i fagene jeg skal ha til våren

$\alpha = 0,85$

### Innramming

5. Jeg opplever at det i undervisningen har vært fokus på hvordan gjennomføre hverandrevurdering



6. Jeg fikk informasjon om hverandrevurdering i forkant av arbeidet med det.
  7. Jeg vet hvor jeg finner informasjon om hverandrevurdering
  8. Jeg forstår hvorfor jeg skal gjøre hverandrevurdering
  9. Jeg har diskutert hverandrevurderingen i oblig 2 med mine medstudenter
  10. Jeg har diskutert hverandrevurderingen i oblig 2 med min seminarleder/foreleser
- $\alpha = 0,75$

## Vedlegg 5: Eksempel på koder i tematisk analyse

Utdrag fra fokusgruppeintervju med førsteårsstudenter som diskuterer påstanden: "Man kan få mer nyttig veiledning av medstudenter enn hva man kan få av for eksempel seminarledere og professorer".				
Navn	Transkripsjon	Koder	Undertema	Tema
Line (student)	Jeg synes selve tilbakemeldingen på oppgaven ville vært mer verdifull hvis den kom fra en seminarleder eller professor, enn en medstudent. Fordi at det er mye lettere å stole på den tilbakemeldingen man får, og målet skal jo være at eh oppgaven blir bedre, og man kan jo få da tilbakemelding fra en medstudent som er på dårligere nivå enn deg selv. Hvis man da på en måte skal blindt stole på den tilbakemeldingen, så kan man risikere at teksten blir dårligere. Eh.. Er det en professor eller seminarleder så vet du at det er en som kan mer. Eh.. og har rett da, i større grad enn en medstudent.. Så det er ja, egentlig først og fremst derfor.	Mer verdifullt med FB fra lærer  Lettere å stole på FB fra lærer  Lettere å stole på FB fra lærer	Nytteverdi  Oppfatning av medstudents kompetanse  Oppfatning av medstudents kompetanse	Å motta HV  Å motta HV  Å motta HV
Nora (student)	Ja, nei eh... for det første kommer det an på hva den medstudenten skriver, eh selvfølgelig om man får noe ut av det eller om det bare er sånn positivt; ja dette her er bra, så.. personlig, av de tilbakemeldingene jeg har fått så har det ikke vært veldig nyttig. Men det er også vanskelig fordi man vet jo at det sitter noen andre som har skrevet denne teksten, så man må jo ta hensyn til det ikke sant. Og så hvert fall første gangen så var det jo ikke anonymt. så da kunne de jo se at det var jeg som hadde skrevet så det ut som. Så jeg tror de kunne vært litt mer tydelig på det.. men om det er så mye mer nyttig å få det av medstudent, det tenker jeg det kommer veldig an på...	Medstudents kompetanse  Ikke nytte av mottatt FB  Vanskelig å gi ærlig FB pga ikke anonym  Tydeligere retningslinjer	Oppfatning av medstudents kompetanse  Nytteverdi  Anonymisering  Organisering	Å motta HV  Å motta HV  Kursdesign  Kursdesign

Merknad: informantene er anonymisert med bruk av fiktive navn.

## **Vedlegg 6: Godkjenning av prosjekt NSD**



### **NSD sin vurdering**

#### **Prosjekttittel**

Formativ vurderingspraksis i Canvas

#### **Referansenummer**

805073

#### **Registrert**

04.09.2020 av Jan Arild Dolonen - janad@uio.no

#### **Behandlingsansvarlig institusjon**

Universitetet i Oslo / Det utdanningsvitenskapelige fakultet / Institutt for pedagogikk

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

Kjetil Horn Hogstad, k.h.hogstad@iped.uio.no, tlf: 22844883

#### **Type prosjekt**

Forskerprosjekt

#### **Prosjektperiode**

28.09.2020 - 31.12.2021

#### **Status**

09.09.2020 - Vurdert

## Vurdering (1)

---

### 09.09.2020 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 09.09.2020, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

#### MELD VESENTLIGE ENDRINGER

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

[https://nsd.no/personvernombud/meld\\_prosjekt/meld\\_endringer.html](https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html)

Du må vente på svar fra NSD før endringen gjennomføres.

#### TYPE OPPLYSNINGER OG VARIGHET

Prosjektet vil behandle alminnelige kategorier av personopplysninger frem til 31.12.2021.

#### LOVLIG GRUNNLAG

Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 og 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen vil dermed være den registrertes samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a.

#### PERSONVERNPRINSIPPER

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

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen
- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål
- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet
- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

#### DE REGISTRERTES RETTIGHETER

Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.

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

#### FØLG DIN INSTITUSJONS RETNINGSLINJER

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

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

#### OPPFØLGING AV PROSJEKTET

NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Simon Gogl

Tlf. Personverntjenester: 55 58 21 17

## **Vedlegg 7: Samtykkeskjemaer**

### Samtykkeskjema for undervisere og koordinator

Masteroppgave om 'Formativ vurderingspraksis i Canvas'

---

Dette er et spørsmål til deg om å delta i et intervju i forbindelse med vår masteroppgave som er del av prosjektet "Formativ vurderingspraksis i Canvas". I dette skrevet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

---

#### **Formål**

Vi skal undersøke undervisere og studenter sin opplevelse av arbeid med egen- og hverandrevurdering i fagene xxxx og xxxx som har blitt gjennomført høsten 2020. Vi ønsker å undersøke studenter og undervisere sin erfaring med digitale verktøy for å gjennomføre egen- og hverandrevurdering, i tillegg til at vi vil undersøke egen- og hverandrevurdering som læringsprosesser.

#### **I masteroppgaven ønsker vi å belyse følgende forskningsspørsmål:**

- Hvordan opplever studentene og undervisere bruken av hverandrevurdering i undervisningen?
- Hvilket læringsutbytte opplever studentene at de får av å bruke Peergrade/Canvasrubrikker i henhold til læringsmålene i emnet?

#### **Hvem er ansvarlig for masteroppgaven?**

Katrin Fallmyr og Karina Rusås Olsen

Masterstudenter ved Institutt for pedagogikk, Universitetet i Oslo.

Veileder for oppgaven er Jan Arild Dolonen ved Utdanningsvitenskapelig fakultet, Universitetet i Oslo.

#### **Hvorfor får du spørsmål om å delta?**

Du får spørsmål om å delta fordi du underviser eller har vært med å koordinere emner ved IPED hvor Peergrade/ Canvasrubrikker har vært prøvd ut for hverandrevurdering.

#### **Hva innebærer det for deg å delta?**

Tilknyttet masteroppgaven ønsker vi å gjennomføre intervju med undervisere og andre som har kjennskap til og som har bidratt til prosjektet “Formativ vurderingspraksis i Canvas”.

Intervjuet med deg vil ta ca. 45 minutter.

Intervjuene vil med din godkjenning bli tatt opp på lydopptaker via Diktafon-appen, og lagret på Nettskjema.

### **Det er frivillig å delta**

Det er frivillig å delta i intervjuet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

### **Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger**

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Dataene vi samler inn vil lagres på sikker server adskilt fra øvrige data slik at ingen uvedkommende kan få innsyn i dine opplysninger.

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er 1. juli 2021. Lydopptak og personopplysninger vil bli slettet.

### **Dine rettigheter**

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

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

Hva gir oss rett til å behandle personopplysninger om deg?

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Institutt for pedagogikk ved Universitetet i Oslo har

NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

### **Hvor kan jeg finne ut mer?**

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



Masterstudenter

Katrin Fallmyr (katrifal@uio.no) og Karina Rusås Olsen (karinaol@uio.no)

Veileder: Jan Arild Dolonen (janad@uv.uio.no)

Institutt for pedagogikk ved prosjektleder Kjetil Horn Hogstad (k.h.hogstad@iped.no) Tlf.:  
22844883.

UiO's personvernombud: Roger Markgraf-Bye (personvernombud@uio.no)

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt  
med: NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller  
på telefon: 55 58 21 17

Med vennlig hilsen

Katrin Fallmyr og Karina Rusås Olsen

Masterstudenter

### **Samtykkeerklæring**

Jeg har mottatt og forstått informasjon om prosjektet, og har fått anledning til å stille  
spørsmål. Jeg samtykker til:

å delta i intervju

at mine opplysninger behandles frem til prosjektet er avsluttet

## Samtykkeskjema for studenter

### Vil du delta i forskningsprosjektet 'Formativ vurderingspraksis i Canvas'?

---

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke hvordan egen- og hverandrevurdering i Canvas kan brukes til å styrke studentaktive læringsformer og akademisk dannelse i emnet xxxx. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

---

#### Formål

Institutt for pedagogikk (IPED) har fått midler til å undersøke hvordan egen- og hverandrevurdering i Canvas kan brukes til å styrke studentaktive læringsformer og akademisk dannelse for bachelor- og masterstudenter. I løpet av høsten 2020 og våren 2021 ønsker vi å prøve ut verktøyene "Peergrade" og "rubrikker i Canvas" for å undersøke om de kan bidra til å styrke IPED studentenes akademiske evner i form av lesing, vurdering og skriving. Forskningsspørsmålene handler om studenters og underviseres opplevelse av egen- og hverandrevurdering i undervisningen, og om det gir opplevd læringsutbytte. Prosjektet er et forskningsprosjekt finansiert av interne midler ved UiO i programmet "Prosjektmidler for digital vurdering 2020/21". I prosjektet vil forskere fra IPED delta, samt masterstudenter fra IPED som skriver masteroppgave om temaet. Opplysningene vi samler inn i forhold til egen- og hverandrevurdering er del av den ordinære undervisningen i emnene, men vår forskning vil ikke ha noen negative konsekvenser for studentene.

#### Hvem er ansvarlig for forskningsprosjektet, og hvorfor får du spørsmål om å delta?

IPED ved UiO er ansvarlig for prosjektet, og du får spørsmål om å delta fordi du er student på emner ved IPED hvor disse to verktøyene skal prøves ut.

#### Hva innebærer det for deg å delta?

I prosjektet ønsker vi å gjøre 3 ulike former for datainnsamlinger. Den første er en anonym spørreundersøkelse rundt bruken av egen- og hverandrevurdering, og opplevelse av læringsutbytte. Den anonyme spørreundersøkelsen gjennomføres H-2020, og vil ta ca. 5 minutter. Den andre datainnsamlingen er gruppeintervju av studenter som er villige til å stille,

og handler også om bruken av egen- og hverandrevurdering, og opplevelse av læringsutbytte. Intervjuet vil gjennomføres H-2020, og vil ta ca. 45 minutter. Den tredje formen for datainnsamling er innholdet i egen- og hverandrevurderingene. Det for å undersøke kvaliteten på tilbakemeldinger, og om tilbakemeldingene er saklige og gjør bruk av relevante begreper og referanser fra pensum.

### **Det er frivillig å delta**

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle personopplysninger om deg som prosjektet har samlet inn vil da bli slettet. Som tidligere nevnt vil det ikke ha noen negative konsekvenser for deg som student i emnet hvis du ikke vil delta, eller om du senere velger å trekke deg. Du må gjennomføre egen- og hverandrevurdering som normalt i emnet, men forskningsprosjektet vil ikke benytte seg av opplysninger om deg eller som du har gitt.

### **Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger:**

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Bare prosjektgruppen fra IPED, som også inkluderer masterstudenter som arbeider med sin masteroppgave, vil ha tilgang til dine opplysninger. Navneliste og koblingsnøkkel lagres på sikker server adskilt fra øvrige data slik at ingen uvedkommende kan få innsyn i dine opplysninger. Du vil ikke bli gjenkjent i eventuelle publikasjoner (rapporter, artikler eller masteroppgaver).

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Opplysningene anonymiseres når prosjektet avsluttes, noe som etter planen er 31. desember 2021. Etter dette slettes lydopptak, og prosjektets navneliste og koblingsnøkkel, slik at prosjektets data anonymiseres.

### **Dine rettigheter:**

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

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og

- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

### **Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Institutt for pedagogikk ved universitetet i Oslo har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

### **Hvor kan jeg finne ut mer?**

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

- Institutt for pedagogikk ved prosjektleder Kjetil Horn Hogstad, e-post: k.h.hogstad@iped.uio.no, Tlf.: 22844883.
- Veileder for masterstudentene: Jan Arild Dolonen, e-post: janad@uv.uio.no
- UiO's personvernombud: Roger Markgraf-Bye, epost: personvernombud@uio.no
- Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:
- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Kjetil Horn Hogstad, prosjektleder

### **Samtykkeerklæring**

Jeg har mottatt og forstått informasjon om prosjektet *Formativ vurderingspraksis i Canvas*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

Du må velge minst ett svaralternativ.

- å delta i anonymt spørreskjema
- å delta i intervju
- å delta med tekstlig innhold fra egen- og hverandrevurdering gitt i Peergrade eller Canvasrubrikker
- at mine opplysninger behandles frem til prosjektet er avsluttet