



UiT Norges arktiske universitet

Fakultet for naturvitenskap og teknologi
Institutt for fysikk og teknologi

Samhandling og læring i et introduksjonsemne i fysikk basert på digital Peer Instruction

Øystein Grøntveit Winnberg

FYS-3907 Mastergradsoppgave i fysikk ved lektorutdanningen, trinn 8-13, Juni 2021

“Du lider ingen nød i furuskog”
–Sigurd Eriksen Lien

Sammendrag

I de senere år har det vært en skifte mot økt bruk av digitale former for undervisning, og i 2020 ble dette spesielt framskyndet fordi brorparten av undervisningen på universitetsnivå måtte foregå digitalt på grunn av koronapandemien. Det er også voksende interesse for mer studentaktive undervisningsformer for å oppnå bedre læringsutbytte. Dette åpner for nye spørsmål rundt digitale studentaktive undervisningsmetoder. Denne studien undersøker: 1) læringsutbyttet i et introduksjonsemne i fysikk gjennomført ved digital Peer Instruction med faste grupper; 2) kjennetegnene til produktive diskusjoner i denne settingen. Forskningsspørsmålene er undersøkt kvantitativt ved før- og ettertester, og kvalitativt ved å analysere opptak av digitale læringssituasjoner.

Studentenes konseptuelle forståelse har blitt kartlagt ved Force Concept Inventory (FCI) ved begynnelsen og slutten av semesteret. Det normaliserte gainet på 0,46 og verdi for Cohens d på 0,88 er sammenlignbare med typiske verdier fra aktive undervisningsformer i en fysisk setting. I Peer Instruction svarer studentene på oppgaver før og etter en gruppediskusjon. Antallet studenter på hver gruppe som svarer riktig før og etter diskusjonen viser at i grupper på tre bør minst én ha svart riktig før diskusjonen for at gruppen skal komme til rett svar. I større grupper bør det være minst to som har svart riktig før diskusjonen. Det er klar sammenheng mellom enkeltstudentenes prestasjon på FCI førtesten og oppgavene før diskusjon. Analysen av opptakene viser at studentene med høy poengsum på FCI førtesten ofte dominerer diskusjonene, mens de andre ikke bidrar mye.

Resultatene viser at digital Peer Instruction med faste grupper gir sammenlignbart læringsutbytte med andre former for aktiv undervisning, men analysen av gruppediskusjonene og diskusjonsoppgavene viser at undervisere bør passe på gruppesammensetning og oppmuntre alle studentene til å delta i diskusjonene.

Takksigelser

Aller først vil jeg rette en stor takk til mine dyktige veiledere: Børge Irgens og David Andre Coucheron. Tusen takk for alle verdifulle innspill, råd og tilbakemeldinger. Deres konstruktive og innsiktsfulle betraktninger har gjort det interessant og givende for meg å jobbe med denne oppgaven. En stor takk rettes også til studentene som deltok i denne studien.

Jeg vil også takke min gode venn, medstudent og samtalepartner Sigurd Eriksen Lien. Jeg har fått mye god innsikt fra våre diskusjoner gjennom hele studieløpet. Det være seg samtaler om fysikk, fotball eller røyefiske, så vel som megetsigende sitater om furuskogens fortrefelighet.

Til slutt en takk til mamma, pappa, Ingunn og Bjørnar for all støtte gjennom studieløpet og i livet ellers.

Innhold

Sammendrag	iii
Takksigelser	v
Figurer	xi
Tabeller	xiii
1 Introduksjon	1
2 Teori	5
2.1 Læringsteoretisk bakteppe	5
2.1.1 Knowledge in Pieces	6
2.1.2 Fysikkfaget	9
2.1.3 Sosio-konstruktivistisk nyansering	10
2.2 Alternative forestillinger i fysikk	11
2.2.1 Å kartlegge alternative forestillinger	12
2.2.2 Å utfordre alternative forestillinger	13
2.3 Konseptuell forståelse	13
2.3.1 Force Concept Inventory	14
2.3.2 Validitet og reliabilitet av FCI	15
2.3.3 Bruk og fortolkning av FCI	16
2.4 Aktiv læring og Peer Instruction	17
2.4.1 Peer instruction	18
2.4.2 Gjennomføring av Peer Instruction	19
2.4.3 Forutsetninger for at Peer Instruction skal fungere . .	21
2.5 Diskusjoner i undervisningen	22
2.5.1 Språk og kommunikasjon i fysikk	22
2.5.2 Argumentasjon og resonnering	22
2.5.3 Kvalitet på diskusjoner	23
2.5.4 Gruppesammensetning og gruppedynamikk	24
2.5.5 Uenighet	26
2.6 Relevant statistikk	27
2.6.1 Effektstørrelser	27

2.6.2	Annet	29
3	Metode	31
3.1	Forskningsdesign	31
3.2	Utvalg	32
3.3	Gjennomføring av emnet	36
3.4	Instrumenter	37
3.5	Opptak	38
3.6	Prosedyre	39
3.6.1	Studiens forløp	39
3.6.2	Analyser	40
3.7	Etiske betraktninger	42
4	Resultater	45
4.1	Læringsutbytte fra Peer Instruction	45
4.1.1	Overordnet læringsutbytte fra undervisning	45
4.1.2	Utbytte fra diskusjon av oppgaver	47
4.1.3	Sammenheng mellom faglig inngangsnivå og overordnet læringsutbytte	51
4.1.4	Sammenheng mellom prestasjon på diskusjonsoppgaver og overordnet læringsutbytte	52
4.2	Produktive diskusjoner	54
4.2.1	Sammenheng mellom faglig inngangsnivå og prestasjon på diskusjonsoppgaver	54
4.2.2	Korrekthet og uenighet	55
4.2.3	Videre analyse av produktive diskusjoner	60
5	Diskusjon	69
5.1	Læringsutbytte	70
5.1.1	Utbytte fra overordnet undervisning	70
5.1.2	Utbytte fra diskusjoner	72
5.2	Produktive diskusjoner	77
5.2.1	Inngangsnivå mot prestasjon på diskusjonsoppgaver	77
5.2.2	Korrekthet og uenighet	78
5.2.3	Hva skjer i diskusjonene?	80
5.3	Studiens validitet	84
5.3.1	Indre validitet	84
5.3.2	Ytre validitet	86
6	Avslutning	87
6.1	Konklusjon	87
6.2	Implikasjoner	88
6.3	Videre arbeid	89

INNHold

ix

Vedlegg

97

Figurer

3.1	Informasjon om deltakere i studien	33
3.2	Informasjon om deltakere med forskjellig samtykke	35
3.3	Alle deltagere i studien	35
3.4	Eksempeloppgave fra FCI	37
3.5	Eksempel på diskusjonsoppgave	38
3.6	Tidslinje	40
4.1	Fordeling av FCI-resultater	46
4.2	Spredningsplott av konseptoppgaver	48
4.3	Histogram med fordeling av konseptoppgaverresultater	48
4.4	Spredningsplott av studenters snittscore på konseptoppgaver	50
4.5	Histogram med snittscore på konseptoppgaver blant studenter	50
4.6	Spredningsplott av individuelle FCI Hake gains	52
4.7	Sammenheng mellom normalisert gain fra diskusjoner og FCI ettertest	53
4.8	Sammenheng mellom resultater på diskusjonsoppgaver og FCI ettertest	53
4.9	Sammenheng mellom FCI førtest og KO-score	54
4.10	Matriser med alle analyseenheter	56
4.11	Uenighet før og etter diskusjon	57
4.12	Antall korrekte ut av diskusjon for ulikt antall riktige inn i diskusjon (treergrupper)	58
4.13	Antall korrekte ut av diskusjon for ulikt antall riktige inn i diskusjon (firergrupper)	59
4.14	Antall korrekte ut av diskusjon for ulikt antall riktige inn i diskusjon (femmergrupper)	59
4.15	FCI-fordeling, eksempelgruppe	62
4.16	Oppgave, analyseenhet 3	62
4.17	Oppgave analyseenhet 11	64
A.1	Oppgave tilknyttet analyseenhet 3	97
A.2	Oppgave tilknyttet analyseenhet 11	99

Tabeller

3.1	Informasjon om studentene,	33
3.2	Informasjon om subutvalg av studenter	34
4.1	Klassifisering av analyseenheter	61



Introduksjon

Gjennomføringen av introduksjonsemner i fysikk både på universitetsnivå og i videregående skole har tradisjonelt vært konvensjonell i den forstand at hovedvekten av undervisningen har blitt lagt på forelesninger hvor fagansvarlig formidler fagstoffet til studentene. Etter forelesningen arbeider studentene med å løse standardiserte regneoppgaver. Studentene kan ofte svare godt på slike oppgaver etter endt undervisning, men studier har vist at en slik tilnærming til undervisning ikke fører til at de oppnår konseptuell forståelse (Hake, 1998; Hestenes, Wells & Swackhamer, 1992; Knight, 2004; Mazur, 1997). Fra dette har behovet for undervisningsstrategier som i større grad fremmer konseptuell forståelse reist seg, og aktive undervisningsformer med samtale og dialog i fokus har i så henseende vist seg å gi gode resultater. (Freeman mfl., 2014).

Hva som inngår i aktiv læring er ikke universelt definert, men kjennetegnes av at studentene aktiv tar del i å bearbeide fagstoff, heller enn å passivt motta det. En slik tilnærming til undervisning er *Peer Instruction*. Denne undervisningsformen er etterprøvd og kvalitetssikret ved forskjellige institusjoner og fagfelt (Crouch & Mazur, 2001; Fagen, Crouch & Mazur, 2002; Hake, 1998; Porter, Bailey Lee, Simon, Cutts & Zingaro, 2011), og legger stor vekt på læring gjennom diskusjon av oppgaver i grupper. Forløpet i en undervisningsøkt i Peer Instruction er i hovedsak lagt opp slik at studentene svarer individuelt på en konseptuell oppgave, hvorpå de diskuterer oppgaven i grupper, for så å avlegge et nytt svar i etterkant av diskusjonen. Dette gjentas så med nye konseptoppgaver. En nærmere gjennomgang av Peer Instruction som undervisningstilnærming vil bli presentert i seksjon 2.4.

I en Peer Instruction-setting vil en høy andel korrekte svar etter diskusjon være en indikator på at studentene har forstått det aktuelle konseptet, og at man kan bevege seg videre til nytt stoff (Mazur, 1997). Underviseren har i liten grad mulighet til å observere og kvalitetssikre det som skjer i gruppediskusjonene, og hvilke faktorer som spiller inn for at en gruppe blir enige om korrekt svar kommer ikke direkte til syne. Her tegner det seg et skille mellom hvorvidt en diskusjon er fører til at mange svarer rett, eller om den i seg selv er utslagsgivende for forståelse. Fra dette framkommer et behov for å få større innsikt i hva som foregår i, og legger premissene for gruppediskusjoner i en Peer Instruction-setting.

Digital undervisning blir vanligere å vanligere, og ved høyere utdanning ser man et skifte mot flere fleksible digitale plattformer for undervisning. Dette har spesielt blitt aktuelt i 2020 som følge av koronapandemien. Det er av interesse å se nærmere på hvordan studentaktiv undervisning fungerer i en digital setting, og denne oppgaven vil derfor undersøke nærmere hvordan undervisning av et introduksjonsemne i fysikk ved bruk av Peer Instruction fungerer. I dette ligger både hvordan det overordnede læringsutbyttet er, men gitt gruppediskusjonens sentrale posisjon i Peer Instruction, er det også interessant å si noe om utbyttet fra disse. Til å gjøre dette har følgende to forskningsspørsmål blitt formulert:

1. Hvordan har læringsutbyttet vært ved digital Peer Instruction?
2. Hva kjennetegner en produktiv gruppediskusjon i denne settingen?

Det aktuelle emnet har tidligere blitt gjennomført på tradisjonelt vis med en hovedvekt på forelesninger, men med innslag av elementer fra Peer Instruction og andre aktive undervisningsformer. Emnet var oppdelt i fellesundervisning, seminarer, regneverksted, obligatoriske innleveringsoppgaver og en avsluttende digital eksamen. Fellesundervisningen i kurset ble høsten 2020 omstrukturert fra å bestå i forelesning til å dreie seg rundt diskusjoner i faste grupper og konseptuelle diskusjonsoppgaver. Gjennomføringen av fellesundervisningen i emnet ble på grunn av den pågående Covid-19-pandemien gjennomført digitalt – med alle de konsekvenser, fordeler og ulemper som naturlig fulgte.

For å evaluere det overordnede læringsutbyttet fra undervisningen har resultatet fra Force Concept Inventory (FCI) blitt brukt. Dette er et kartleggingsverktøy for forståelse av grunnleggende newtonsk mekanikk, i form av en test som studentene gjennomførte før og etter formell undervisning. For å kunne undersøke gruppediskusjonene i seg selv, er studentbesvarelser på alle diskusjonsoppgaver i løpet av undervisningen undersøkt. Den digitale gjennomføringen av fellesundervisningen har åpnet et mulighetsrom for å kunne gå inn og undersøke nærmere hva som skjer i gruppediskusjoner, som ikke ville vært like tilgjengelig ved stedlig undervisning. Ved skjermopptak har gruppesamtaler fra store deler

av semesteret blitt tilgjengelig for analyse, med mål om å kunne finne mer ut av hva som gjør gruppediskusjoner produktive. Gjennom en variert analyse av de tilgjengelige data søker denne oppgaven dypere innsikt i diskusjonene i Peer Instruction.

/2

Teori

2.1 Læringsteoretisk bakteppe

De fleste syn på læring og kunnskap i naturvitenskapen springer ut fra et *konstruktivistisk* fundament. Hovedtanken i konstruktivismen forklarer Sjøberg (2014) som at mennesker i møte med verden rundt seg finner sammenhenger ved først og fremst å tolke sanseinntrykk og lage seg mentale modeller av virkeligheten. Et konstruktivistisk vitenskaps- og kunnskapssyn gir dermed slipp på søken etter en objektiv sannhet, men prøver heller å beskrive verden på best mulig måte ved å utvikle og bruke egnede hjelpemidler – kognitive så vel som mer håndfaste (Angell mfl., 2019). Konstruktivismen avfeier ikke eksistensen av en objektiv virkelighet, men den problematiserer tilgjengeligheten til denne. I grener av konstruktivismen blir en objektiv virkelighet sett på som noe som ikke kan læres. I en diskusjon vedrørende radikal konstruktivisme anskueliggjør Quale (2007) dette på følgende vis:

Spesielt betraktes naturvitenskapen som en slik modell – definert ved et sett av teorier, metoder, prosedyrer og verktøy – som kan brukes til å strukturere en bestemt undermengde av våre opplevelser: nemlig de fenomener som til enhver tid blir ansett å være legitim gjenstand for naturvitenskaplig forskning. (Quale, 2007, p. 181).

Et konstruktivistisk syn på naturvitenskapen kan også ses i parallell med naturvitenskapens historie (Sjøberg, 2014). Den teorien som til en hver tid best

forklarer et fenomen er ved det tidspunkt gjeldende, og skulle en forbedret teori komme til syne, vil denne etter hvert erstatte den forrige. Sjøberg viser også til at man fra fødselen av hele tiden er utsatt for sanseintrykk og fenomener som man ønsker å forklare. Menneskers trang til å forklare sin virkelighet og sine omgivelser er i følge Sjøberg iboende. Fra erfaringer og opplevelser gjennom våre livsløp danner vi oss et syn på virkeligheten, og bygger oss opp noen oppfatninger og forestillinger vedrørende forskjellige fenomeners natur og årsakssammenhenger. Slike forestillinger står sentralt i mange perspektiver på læring i naturvitenskapen, og vil bli et gjennomgangstema videre i denne delen av oppgaven. I det følgende vil noen perspektiver på læring med basis i konstruktivismen bli presentert mer utdypende.

2.1.1 Knowledge in Pieces

Fra et epistemologisk perspektiv presenteres *Knowledge in Pieces* (heretter KiP) fra diSessa (2018) som et stort teoretisk og empirisk rammeverk for å forstå hva kunnskap og læring innebærer. Dette synet på læring ble presentert av diSessa (1993) som et alternativ til de, på det tidspunkt, rådende perspektivene på læring i naturfag (Harlow & Bianchini, 2020). KiP etablerer et grunnlag for å forstå læring i naturvitenskaplige fag, som er konsistent med den større bakenforliggende teorien på læring – i hovedsak konstruktivismen – og diSessa selv påpeker at KiP har hatt signifikant suksess hva gjelder å forklare fenomener knyttet til undervisning i naturfaglige disipliner (diSessa, 2018).

Der KiP på linje med mange andre syn på læring tar utgangspunkt i at ny kunnskap bygger på eksisterende kunnskap, skiller det seg ut ved å ta sikte på å kunne forklare både kortsiktige fenomener – fragmentert kunnskap og læring i øyeblikket (derav navnet), og langsiktige fenomener som blant annet konseptuell endring (diSessa, 2018). Harlow og Bianchini (2020) viser blant annet til Davis, Horn og Sherin (2013) som konstaterer at alle perspektiver på læring som er av formen: «slett gammel kunnskap, erstatt den med korrekt kunnskap», ikke er et *lærings*perspektiv i det hele tatt. KiP søker videre å skape en toveis sammenheng mellom teori og empiri, i den forstand at teorien er strengt avgrenset og bygd på observasjoner samtidig som teorien kan projisere den lærendes gjøren og laden når læring finner sted og dermed gi en allmengyldig betydning til den lærendes handlinger (diSessa, 2018). Både denne toveis sammenhengen mellom empiri og teori og inkluderingen av både kortsiktig og langsiktig læring, gjør at KiP kan ses på som en reaksjon på de teorier på læring som er a priori, men altomfattende, og derfor vanskelig å bruke til å beskrive det som faktisk skjer i læring (diSessa, 2018). For eksempel deler KiP mange aspekter med den konstruktivistiske basisen i Piaget sine teorier, men som diSessa påpeker er ikke piagetistisk psykologi spesielt god på å forklare hvordan kortsiktige endringer hos den lærende og den tenkningen

som skjer i sanntid, henger sammen med langsiktige endringer.

diSessa (1993) peker spesielt på fysikkstudenters intuitive forståelse av verden rundt seg og på forståelsen av kausaliteten i hendelsesforløp. Forut for diSessa (1993) hadde det blitt gjort klart at fysikkstudenter hadde forestillinger om verden som var konsistente med det de opplevde i hverdagen, men ute av takt med fysiske lover og prinsipper. Slike forestillinger blir ofte referert til som misoppfatninger og/eller hverdagsforestillinger, og vi skal se nærmere på disse i seksjon 2.2. Synet på disse forestillingene (blant annet fra McCloskey (1983)), var at de var robuste og sammenhengende selv når de var inkompatible med naturvitenskaplige ideer og prinsipper (Harlow & Bianchini, 2020). diSessa presenterte et litt annet syn på dette, og i motsats til en helhetlig kognitiv struktur, beskrev han strukturer satt sammen av små enheter som han kalte for *fenomenologiske primitiver* (p-prims) (Harlow & Bianchini, 2020). Kognitive strukturer knyttet til preeksiterende oppfatninger ble her ikke sett på som koherente og rigide, men som satt sammen av små enheter, som i forskjellig grad kunne linkes opp mot hverandre eller aktiveres. diSessa (1993) påpekte at slike p-prims ikke nødvendigvis er korrekte eller ukorrekte i seg selv, men at de kan være passende eller upassende i ulike kontekster. En treffende analogi for å forklare synet på forståelse som et sett av små koordinerte enheter, brukes av Harlow og Bianchini (2020), ved å vise til interaksjonene i en fugleflokk eller fiskestim. Hvert individ må agere på en bestemt måte i sammenheng med de andre for å skape en helhetlig bevegelse.

Tidligere i denne seksjonen ble *konseptuell endring* nevnt som et eksempel på langsiktige læringsfenomener. Konseptuell endring blir presentert av Harlow og Bianchini (2020) som en rådende teori for læring, før den ble utfordret av blant andre diSessa ut over nittitallet. Essensen og hovedpremisset som ligger i konseptuell endring kan forklares som at misoppfatninger eller preeksisterende oppfatninger som innhas av noviser innen et felt, kan *erstattes* av synspunkt og tankebaner likt det hos eksperter innen det samme feltet. Denne erstatningen skjer enten ved at nye sammenhenger mellom konsepter forstås (svak konseptuell endring), eller ved at forståelsen av selve konseptet endres (sterk konseptuell endring) (Harlow & Bianchini, 2020). Med KiP og ideen om p-prims søker diSessa og Sherin (1998) å utfordre dette synet. I stedet for ideen om *konsepter* som sådan, introduseres ideen om *koordinasjonsklasser* beskrevet som kunnskapssystemer og bestående av forskjellige typer kognitive bestanddeler som sammen skaper et system av strategier for å kunne forstå verden. I hovedsak presenteres her en mye mer flytende forståelse av konsepter enn det som var rådende. For en mer inngående gjennomgang av koordinasjonsklasser vises det til diSessa og Sherin (1998).

I forlengelsen av diSessa (1993) og diSessa og Sherin (1998) sitt arbeid, trekker Harlow og Bianchini (2020) fram Hammer (2004) og hans beskrivelse av en

såkalt *manifold ontology*, altså en mangfoldig forståelse av kognitive fenomener. Hammer presenterer i likhet med diSessa dette synspunktet som et alternativ til de da rådende synspunkt på læring som noe som skjer gjennom piagetistiske akkomodasjonsprosesser. For eksempel beskriver Harlow og Bianchini (2020) hvordan det å løse kognitive konflikter spiller en stor rolle i synet på læring gjennom konseptuell endring. Løsning av kognitive konflikter ved å akkomodere nye og mer korrekte oppfatninger, kan ikke skje – i følge Hammer (2004) – dersom konsepter blir sett på som helhetlige kognitive strukturer. Bruk av eksisterende strukturer til å prosessere ny informasjon vil ikke i seg selv føre til at oppfatninger endres (Harlow & Bianchini, 2020). Hammer (2004) foreslår dermed også å se på kognitive strukturer i form av små enheter. Hammer kaller dog disse for *ressurser* heller enn *p-prims*.

Sentralt både hos Hammer (2004) og diSessa (1993) er at hverken ressurser eller *p-prims* er korrekte eller ukorrekte i seg selv, men at deres brukbarhet avhenger av kontekst og situasjon. Ressurser som aktiveres i en kontekst vil ikke nødvendigvis aktiveres i en annen kontekst (Harlow & Bianchini, 2020). diSessa (2018) eksemplifiserer dette med en case hvor måten en instruktør stiller spørsmål på er avgjørende for forklaringsmodellene til studenten. Her er studenten bedt om å forklare hva som skjer når en ball kastes rett opp i luften og faller ned igjen. Studenten har en adekvat forklaring i tråd med fysiske prinsipper (kun tyngdekraften virker på ballen), inntil instruktøren spør spesifikt om hva som skjer idet ballen endrer retning. Her endrer studenten forklaringsstrategier og det kommer frem oppfatninger som ikke stemmer overens med de faktiske fysiske prinsippene (Studenten forklarer blant annet at kraften fra den som kastet ballen «er værende» i ballen helt til den snur). Harlow og Bianchini (2020) viser på sin side til en case fra Hammer, Elby, Scherr og Redish (2005), hvor en student bruker forskjellige ressurser i hjemmet og i klasserommet når hen er konfrontert med det samme konseptet. Slike eksempler tydeliggjør i følge Harlow og Bianchini (2020) den kontekstbaserte naturen til kunnskap og forståelse, og at det ikke holder å vite hva en student *kan* for å spå om studenten vil være i stand til å nyttegjøre seg passende ideer og forklaringsmodeller i enhver kontekst.

I Hammer (2004) sitt syn på kognitive strukturer som satt sammen av ressurser, inkluderer han forskjellige klasser av ressurser. En av disse er de han kaller *epistemologiske ressurser*. Disse forklarer hvordan både studenter og lærere tilnærmer seg læring (Harlow & Bianchini, 2020). Epistemologiske ressurser handler altså om hvordan kunnskap blir sett på av et lærende individ. Hvilke epistemologiske ressurser som aktiveres er – som med andre kognitive enheter – kontekstavhengig. For eksempel kan en student i en kontekst tenke på kunnskap som noe en får fra en autoritet, men i andre kontekster se på kunnskap som noe hen selv kan konstruere (Harlow & Bianchini, 2020). Dette tilkjennegjør et metakognitivt perspektiv som også må være til stede for å forstå læring. På

samme måte som studenter kan dra veksler på ulike deler av det de besitter av kunnskap i møte med nytt fagstoff, kan de også dra veksler på ulike syn på kunnskap.

Som poengtert i starten av denne seksjonen, er KiP – like mye som et perspektiv på læring – også et rammeverk for å modellere kognitive prosesser. Som sådan kan rammeverkets oppbygning sies å være et nettverk i tråd med det forenklede bildet man ofte møtes med i forklaringen av hvordan nevroner i hjernen er koblet sammen. Aktivering av ett nevron eller grupperinger av nevroner, leder til aktivering av andre deler av nettverket (Bing & Redish, 2009). I forlengelsen av dette kan aktivering av ressurser like godt være at de slås på som at de slås av. Bing og Redish (2009) bruker strukturene i rammeverket fra KiP til å peke på hvordan man kan modellere hvordan studenter bruker og endrer ulike epistemologiske ressurser i undervisningssettinger. Hvordan en students aktivering av epistemologiske ressurser kommer til syne i argumentasjon og problemløsning peker Bing og Redish på som *epistemological framing*. Innbakt i dette ligger også hvilke hjelpemidler og øvrige ferdigheter studenten nyttegjør seg av i den aktuelle settingen.

2.1.2 Fysikkfaget

De perspektiver på læring som er presentert til nå er i høyeste grad relevante for – for ikke å si basert på – læring i fysikkfaget. Det synes allikevel betimelig å videre rette et grovt, overordnet fokus på hva som kjennetegner fysikk som fagområde. Angell mfl. (2019) legger frem to definisjoner av hva fysikk er, som tenkt presentert for en novise innen faget. Den første er som følger: «Fysikk er vitenskapen om universets og alt stoffs fundamentale bestanddeler og om kreftene som virker mellom dem». En slik definisjon baserer seg på det Angell mfl. beskriver som fysikkens *produkter*, og i produkter ligger begreper, lover og teorier. Den alternative definisjonen er i motsetning til den første fundert ut i fra fysikkens *metoder* og lyder «Fysikk dreier seg om å lage (matematiske) modeller av virkeligheten». I en slik definisjon karakteriseres fysikken ved bruken av hensiktsmessige verktøy og praksiser, og gjerne i form av matematikk. I følge Angell mfl. må elever i møte med fysikkfaget få møte begge disse sidene. Domenet av metoder vil i fysikkundervisning, gjerne forbindes med kvantitativ problemløsning – for eksempel regneoppgaver med utgangspunkt i kjente formler, mens domenet av produkter kan knyttes til den mer kvalitative forståelsen av fysiske fenomener og konsepter.

Ikke ulikt de perspektivene som over ble presentert fra Angell mfl. (2019), postulerer Knight (2004) at et av flere mål med god fysikkundervisning er «å balansere kvantitativ resonnering og problemløsning opp mot kvalitativ resonnering og konseptuell forståelse». Fra dette følger at det å lære seg fysikk

innebærer å beherske fysikkens forskjellige domener. Fysikkens domener kan her henspille på de innholdsmessige ulike grenene av fysikken (mekanikk, elektromagnetisme etc.), så vel som produkter og metoder som beskrevet over. Knight presenterer også tre ulike kategorier for kunnskap i fysikk. Den første av disse er faktakunnskap, og forklares som den kunnskap vi bærer med oss fra erfaringer i den fysiske verden, eller som vi aksepterer på bakgrunn av veletablert konsensus i vitenskapelige miljøer. Faktakunnskap kan på litt sleivete vis ses på som selvstendige fragmenter av informasjon uten noen direkte kobling til hverandre. Kobles imidlertid disse fragmentene sammen til større konstrukt, beredes grunnen for det Knight kaller konseptuell kunnskap, beskrevet som en sammenfattet forståelse av nevnte faktakunnskapkonstrukt. Det er lett å ane at et slikt syn på kunnskap har base i et KiP-perspektiv, og innbakt i det Knight kaller faktakunnskap kan det også ligge ulike epistemologiske synspunkter (jamfør Hammer (2004) i seksjon 2.1.1). Knight (2004) beskriver også en siste form for kunnskap, nemlig prosedyreorientert kunnskap. Denne knytter de to første kategoriene opp mot den kvantitative dimensjonen ved fysikkfaget, og er kunnskap om fysikkens metoder. Det er også et metakognitivt aspekt ved denne typen kunnskap, for i følge Knight handler prosedyreorientert kunnskap om «[...] knowing to use what you know».

Siden læring og fysikklæring i stor grad blir beskrevet som å endre og tilpasse kognitive strukturer, synes det fornuftig å trekke frem hvordan man – i følge Knight (2004)– kan skille mellom hvordan noviser og eksperter innen samme fagfelt organiserer kunnskap. Det pekes på at eksperters kunnskapsstrukturer kan beskrives som hierarkiske, noe som medfører at det å resonnerer på bakgrunn iboende kunnskapstrukturer i møte med ny kunnskap kommer mer naturlig enn hos en novise. Novisers kunnskapsstrukturer beskrives derimot som flytende og kaotiske, og både ny og eksisterende kunnskap vil for en novise vekttes likt. I fysikkfaget tilkjennegjør dette seg i møte med oppgaver og problemløsning. Der en ekspert kan koble oppgaven som skal løses til underliggende fysiske prinsipper, vil novisen kun koble oppgaven til overflateaspekter – det være seg for eksempel en formel fra en lærebok. Dermed må fysikknovisers kognitive strukturer omformes til å bli mer hierarkiske om forståelse i faget skal oppnås, og kanskje er ikke det mulig uten sosial interaksjon. I det følgende skal vi nyansere synet på læring som en individuell prosess, og fremheve viktigheten av støtte fra andre i læringsprosessen.

2.1.3 Sosio-konstruktivistisk nyansering

Til nå i dette kapittelet har læring og tilegning av kunnskap i stor grad blitt sett på som en individuell øvelse. I forlengelsen av de synspunkt på læring med base i konstruktivismen som er blitt presentert, er det derfor relevant å også bringe til bords det Angell mfl. (2019) betegner som *sosialkonstruktivisme* og et

sosiokulturelt syn på læring. Angell mfl. viser til Solomon (1994) som erklærer at «personlig kunnskap er notorisk usikker hvis den ikke blir sosialt forsterket». Solomon (1994) løfter også frem språk som en tydelig katalysator for utvikling av kunnskap. Angell mfl. (2019) påpeker også hvordan sammenhengen mellom språk og læring i realfagene er tett koblet opp mot Vygotskys tanker om *den nærmeste utviklingssonen*, og skillet mellom det en person kan ta til seg av kunnskap på egenhånd og med støtte fra et eller flere individ(er) med høyere grad av ekspertise. En lærer, medstudent eller medelev med høyere innsikt kan virke som et stillas i utviklingen og tilpasningen av kunnskapsstrukturer – fra flytende til hierarkiske (jamfør Knight (2004)) – hos en gitt student eller elev.

Å fremheve det sosiale aspektet ved læring i utvidelsen av et konstruktivistisk læringssyn, innebærer å inkludere den sosiale settingen der læring forgår. Sjøberg (2014) poengterer at alle sosiale miljøer har en rekke karakteristika – blant annet hva gjelder verdensanskuelse og uttryksformer – som fremkommer som naturlige innad, men som kan skille seg fra andre sosiale miljøer. Ord, begreper, og konsepter kan oppfattes forskjellig og ha forskjellig betydning fra miljø til miljø (Sjøberg, 2014). Skoleelever eller studenter ved universitetet vil preges av sine miljø, og med disse rammebetingelsene er det i følge Sjøberg (2014) urealistisk å betrakte læring som en personlig, kognitiv prosess. Som vi har sett, bærer mange konstruktivistiske perspektiver på læring preg av å fokusere på endring eller tilpasning av kognitive strukturer, og alternative forestillinger blir ofte sett på som det som i så måte behøver endres eller tilpasses. Sjøberg (2014) skriver at slike forestillinger ofte er godt innarbeidet fra hverdagen. Derfor må det – når de konfronteres i undervisning – tas høyde for en eventuell endring kun vil skje i konteksten av undervisning, hvor det sosiale miljøet er ens medelever og medstudenter. De samme forestillingene vil ikke nødvendigvis bli korrigert i andre kontekster. Dette kan ses på som å være i tråd med det kontekstorienterte aspektet ved KiP-tilnærmingen til læring. Alt i alt er det uomtvistelig at læring er en kompleks prosess. Å tilegne seg kunnskap innebærer korrigering av kognitive strukturer, men for at det skal skje, spiller både kontekst og sosialt samspill en rolle.

2.2 Alternative forestillinger i fysikk

Det har allerede kommet tydelig fram at læring og forståelse – ikke minst i fysikkfaget – på nærmest uunngåelig vis er knyttet til elevers og studenters preeksisterende kognitive strukturer og oppfatning av verden. I denne seksjonen skal vi prøve å gi et overblikk over denne tematikken, men aller først er det på sin plass å gjøre noen begrepsmessige presiseringer. I denne oppgaven blir begreper som «misoppfatninger», «hverdagsforestillinger», «alternative fore-

stillinger» og «preeksisterende forestillinger» brukt om hverandre. I hovedsak så kan man skille på de oppfatninger og forestillinger elever eller studenter har med seg fra «hverdagen» og tar med seg inn i undervisning, og de oppfatninger og forestillinger som blir formet på grunn- og/eller tross av formell undervisning (Hannisdal, Haugan & Munkvik, 2007; Stojanovska, Petrusovski, Köller & Karlsen, 2015). Som det ligger i begrepene, vil hverdagsforestillinger og preeksisterende forestillinger henviser til første kategori, mens misoppfatninger viser til sistnevnte. Fordi begge typene oppfatninger i natur er like, kan de være vanskelig å skille og blir ofte omtalt om hverandre i litteraturen. I denne oppgaven skal vi prøve å vise til alternative forestillinger som et sekkebegrep, dernest vil misoppfatninger og preeksisterende forestillinger peke på de to underkategoriene

I fysikken tilkjenner seg alternative forestillinger seg ofte i form av de erfaringer man har fått bekreftet å gi mening gjennom personlige erfaringer i den fysiske verden, og slike erfaringer basert på «sunn fornuft» har vist seg å spille en substansiell rolle i introduksjonsemner i fysikk (Hestenes mfl., 1992). Spesielt gjelder dette newtonsk mekanikk – Hestenes mfl. hevder sågar at slike preeksisterende forestillinger er inkompatible med newtonske konsepter i de fleste henseender – og det er noen typiske oppfatninger som ofte går igjen blant noviser innen fysikkfaget. For eksempel er det vanskelig, basert på hverdagslig erfaring, å innse at et legeme kan være i bevegelse uten at det virker krefter på det. En vanlig alternativ forestilling er dermed den om en kraft immanent i objekter som holder dem i bevegelse (ofte referert til som en *drivkraft* eller *impetus*). For en oversiktlig taksonomi over typiske hverdagsforestillinger i newtonsk mekanikk henvises det til Hestenes mfl. (1992).

2.2.1 Å kartlegge alternative forestillinger

Angell mfl. (2019) viser til Ausubel, Novak og Hanesian (1978) som sier at den viktigste faktoren som påvirker læring er hva den lærende allerede kan. Dette peker på et behov for å kartlegge hva elever og studenter sitter på av alternative forestillinger. Disse må så adresseres på en måte som på sikt gjør de forenelige med de forestillinger som er rådende i det aktuelle fagfeltet (Barke, Hazari & Yitbarek, 2009). Barke mfl. påpeker at selv om det er en del alternative forestillinger som er vanlige, er en lærer er nødt til å diagnostisere hvilke forestillinger som råder i en klasse eller studentgruppe. Diagnostiske tester kan være en måte å gjøre dette på. Barke mfl. henviser til Taber (2002) som maler et bilde av læreren som en doktor som skal diagnostisere og behandle elever for de «sykdommer» som opptrer i form av alternative forestillinger. Tester som avslører slike forestillinger er et viktig verktøy her. Basert på resultater fra målrettede designede tester kan tiltak iverksettes for å behandle (les: korrigere) de alternative forestillingene (Barke mfl., 2009). Det er imidlertid ingen garanti

for at en test vil kunne avdekke alt, og Sjøberg (2014) peker på viktigheten av at elevene også selv må bli kjent med de oppfatninger de har som er i konflikt med vitenskapens prinsipper. Undervisningen i seg må altså utfordre misoppfatninger og preeksisterende forestillinger.

2.2.2 Å utfordre alternative forestillinger

Alternative forestillinger, spesielt de som er preeksisterende er ofte godt plantet i den lærendes kognitive strukturer, og kan være vanskelig å utfordre (Sjøberg, 2014). Når studenter eller elever ikke får utfordret sine alternative forestillinger, tvinges de til å ty til strategier som ikke er forenelige med varig konseptuell forståelse – det være seg for eksempel ren pugging av problemløsningsstrategier, eller memorisering av fagstoff. Når det gjelder alternative forestillinger spiller ofte språket en rolle. I lys av newtonsk mekanikk er det for eksempel naturlig å peke på hvordan begreper som «kraft» og «energi» brukes forskjellig i hverdagen og i fysikken. Sjøberg påpeker at det er viktig å skille arenaer her. En feilaktig bruk av begrepet kraft i fysikken er ikke nødvendigvis feil hjemme, så lenge man gjør seg forstått. Angell mfl. (2019) viser til Vygotsky som skiller mellom *spontane* og *vitenskapelige* begreper, hvor spontane begreper er knyttet til preeksisterende forestillinger. Et viktig aspekt ved læring fra dette perspektivet – påpeker Angell mfl. – er at elever må få mulighet til å bruke språk og symboler for å teste og justere dem opp mot språket og symbolene til andre, blant annet en lærer, medelever, eller læreboka. God undervisning innebærer her å ta utgangspunkt i elevens språk, og la eleven få utvikle det mot et vitenskapelig språk gjennom å bruke det. Dette skjer ikke gjennom ren pugging, eller å få det forklart. Dette perspektivet oppsummeres fint av Vygotsky sitert og oversatt av Angell mfl. (2019, p. 214):

Språkets struktur representerer slett ikke en enkel avspeiling av tankens struktur. Derfor kan man heller ikke bare henge språket over tanken som en ferdigsydd kjole. Språket tjener ikke som uttrykk for den ferdige tanken. En tanke som blir omsatt i språk, blir omstrukturert og forandres. Tanken blir ikke uttrykt i ordet, den blir til gjennom ordet.

2.3 Konseptuell forståelse

I litteraturen blir det ofte vist til Blooms taksonomi (Bloom, 1956) for å kartlegge alle aspekter som inngår i læring, kunnskap og forståelse, og da kanskje spesielt det kognitive domenet i taksonomien. Her blir de målene man har vedrørende læring klassifisert i seks nivåer av økende kompleksitet. Disse er kunnskap,

forståelse, anvendelse, analyse, syntese og evaluering. I denne oppgaven skal ikke *konseptuell forståelse* i fysikk defineres presist, all den tid det i seg er en øvelse like kompleks og uoversiktilig som den gjengse didaktiske modell. Allikevel er det verdt å presisere at Blooms taksonomi kan være et fint bakteppe for å evaluere graden av forståelse. Selv om forståelse er nivå to her, er det grunnlag for å si at det å *forstå* et konsept i fysikk, innebærer blant annet å kunne anvende det og koble det opp mot andre konsepter. Her vil altså forståelse i stor grad inkludere syntese, analyse og anvendelse blant annet.

Det synes åpenbart at det i fysikkundervisning er ønskelig å legge til rette for at fysikkstudenten eller fysikkeleven står igjen med tilstrekkelig konseptuell forståelse ved endt undervisningsløp. For å kunne evaluere og kvalitetssikre undervisning er det derfor behov for å kunne kvantifisere og måle konseptuell forståelse. I introduksjonsemner i fysikk er kraftkonseptet instrumentelt i så henseende. I det følgende vil testen *Force Concept Inventory* presenteres og beskrives – en test som tar utgangspunkt i kraftkonseptet og misoppfatninger som berører det, og derfra søker å kartlegge en grunnleggende konseptuell forståelse blant studenter.

2.3.1 Force Concept Inventory

Force Concept Inventory (heretter FCI) består av flervalgsoppgaver som tar for seg kraftkonseptet i Newtons mekanikk fra ulike perspektiver og innfallsvinkler. Utviklingen av denne testen var i sin tid motivert av det faktum at fysikkstudenters misoppfatninger og hverdagsforestillinger spiller en stor rolle for læring i introduksjonsemner i fysikk, og at disse er i konflikt med Newtonske konsepter (Hestenes mfl., 1992). Hestenes mfl. (1992) påpeker også at tradisjonell fysikkundervisning gjør lite for å konfrontere og endre misoppfatningene, og at dette er uavhengig av underviseren. Tilfredsstillende undervisning i fysikk forstås dermed som å kreve at underviseren har kunnskap om hvordan studentene tenker og lærer, i tillegg til øvrig faglig kompetanse og pedagogiske evner. FCI tar sikte på å teste studentenes forståelse av kraftkonseptet i særdeleshet. Uten en «Newtonske» forståelse av krefter, vil resten av Newtons mekanikk for en student framkomme som meningsløs og ubetydelig (Hestenes mfl., 1992). Videre er forståelse av Newtons mekanikk grunnleggende for å beherske andre og viderekommende områder innen fysikken (Hestenes mfl., 1992), og FCI fremstår dermed som et passende verktøy for å måle en slags nødvendig basis av forståelse hos fysikkstudenter.

Det er i hovedsak to aspekter som er karakteriserende for FCI. For det første har testen som formål å dekke alle fasetter ved kraftbegrepet i Newtons mekanikk, og som det framstår fra Hestenes mfl. (1992) og Hestenes og Halloun (1995) kan seks dimensjoner ved kraftbegrepet identifiseres som essensielle for en

fullstendig forståelse. Disse er *kinematikk*, *Newtons første lov*, *Newtons andre lov*, *Newtons tredje lov*, *superposisjonsprinsippet* og *ulike typer krefter*. Flervalgsoppgavene i FCI er designet slik at alle disse dimensjonene dekkes, og slik at flere oppgaver dekker samme dimensjon, men med forskjellige innfallsvinkler. Det andre hovedaspektet ved FCI har å gjøre med svaralternativene til hver oppgave. Testen er designet slik at studenten på hver oppgave blir tvunget til å gjøre et valg mellom et Newtonsk konsept, og tre alternative forestillinger¹ (Hestenes & Halloun, 1995). De gale alternativene representerer misoppfatninger og forestillinger som studenten kan ha med seg inn i undervisningen og som er til hinder for faktisk konseptuell forståelse. Hestenes mfl. (1992) påpeker at oppgavene i FCI gjerne oppfattes som trivielle for det de kaller trente newtonske tenkere, men at de avdekker store konseptuelle hull som tradisjonelle problemløsningsorienterte oppgaver ikke avslører. På denne måten er FCI et verktøy som kan brukes både i diagnostisk øyemed, men også for å evaluere undervisning.

2.3.2 Validitet og reliabilitet av FCI

Det påpekes i Hestenes mfl. (1992) og Hestenes og Halloun (1995) at resultater fra FCI må ses på som et samlet konstrukt og at det testen faktisk måler er misforholdet mellom studentens forestillinger og det newtonske kraftkonseptet. Som en respons til en faktoranalyse publisert av Huffman og Heller (1995) som sår tvil om validiteten og tolkningen av FCI, peker Hestenes og Halloun (1995) på nettopp dette aspektet ved testen, og bruker det til å argumentere for at Huffman og Heller (1995) sine resultater tvert om bidrar til å styrke validiteten til FCI. Huffman og Heller (1995) tok for seg de dimensjonene av kraftbegrepet som FCI baserer seg på, og fremhevet det faktum at faktoranalysen ikke fant grupperinger av disse dimensjonene. Denne erkjennelsen er imidlertid helt naturlig, i følge Hestenes og Halloun (1995), all den tid disse dimensjonene er motivert som identifiserbare for trente newtonske tenkere, og ikke utrente fysikkstudenter som bærer med seg alternative forestillinger.

Hestenes og Halloun (1995) hevder at validiteten til FCI er ut over all rimelig tvil. Blant det som påpekes og diskuteres er forekomster av falske positive og falske negative. Sannsynligheten for at disse opptrer sier noe om validiteten til innholdet i FCI. I denne sammenhengen så vil falske negative være tilfeller hvor en student med kontroll på det newtonske kraftkonseptet har valgt et ikke-newtonsk (altså galt) svar, og falske positive er tilfeller hvor korrekt svar avlegges av ikke-newtonske årsaker. Sannsynligheten for en falsk negativ blir av Hestenes og Halloun konservativt anslått til å være under ti prosent. Dette følger

1. Dette gir navnet til FCI en fin tvetydig tøtsj, som gjør det svært så beskrivende for hva testen består i.

av at det korrekte – altså det newtonske svaret til de fleste spørsmål, fremstår åpenbare og uproblematisk for en som besitter tilstrekkelig forståelse. Derfor vil forekomsten av falske negative i hovedsak komme av uoppmerksomhet og uforsiktighet hos respondenten. Når det gjelder falske positive vises det til to konkrete attributter ved testens design som reduserer disse. Det ene vedrører de ulike dimensjonene ved kraftkonseptet som testen berører. Designet er slik at en dimensjon belyses av ulike spørsmål, og slik kan en falsk positiv på et spørsmål i noen grad kompenseres for av en ukorrekt respons på et annet spørsmål som belyser samme dimensjon. Dessuten vil det faktum at de gale svaralternativene representerer misoppfatninger og hverdagsforestillinger gjøre at disse er mer forlokkende enn det korrekte svaret for ikke-newtonske tenkere. Fordi andelen falske positive er vanskelig å begrense i stor nok grad, må FCI-resultater – muligens untatt respondenter med særdeles høy score – ses på som en øvre grense for respondentens newtonske forståelse (Hestenes mfl., 1992).

Basisen i FCI tar utgangspunkt i en annen, tidligere mye brukt diagnostisk test i fysikk, nemlig *Mechanics Diagnostic* (I. A. Halloun & Hestenes, 1985b). Omtrent halvparten av spørsmålene i FCI er de samme som i denne, og designet av de to testene er smidd over samme lest (Hestenes mfl., 1992). Hestenes mfl. fremhever dog at FCI tilbyr enda mer systematisk og gjennomført oversikt over forskjellige misoppfatninger enn forgjengeren. På bakgrunn av det overstående brukes det faktum at reliabiliteten til *Mechanics Diagnostic* er høy – ettersom resultater er blitt replisert mange ganger, som et utgangspunkt for å stadfeste reliabiliteten til FCI. Det vises videre til ulike intervjuer gjort med studenter vedrørende sine svar på FCI, og Hestenes mfl. (1992) påpeker en påfallende forutsigbarhet hos studentenes responser. Utover dette er det analysene til Lasry, Rosenfield, Dedic, Dahan og Reshef (2011) som er verdt å ta for seg når det gjelder FCIs reliabilitet. Her ble det funnet at FCI – i tråd med testens intensjon – måler ett enkelt konstrukt, og at den indre konsistensen i testen er høy. Lasry mfl. (2011) sine analyser fant også en høy test-retest-reliabilitet for FCI.

2.3.3 Bruk og fortolkning av FCI

Hestenes mfl. (1992) påpeker at FCI kan bli brukt i både diagnostisk øyemed, såvel som til evaluering av undervisning og forskning. Testen kan brukes til å identifisere og klassifisere misoppfatninger hos studenter, og Hestenes mfl. påpeker at intervjuer med studenter basert på FCI kan gi undervisere solid innsikt i hvilke misoppfatninger som råder hos studentgruppen. På grunn av testens relativt høye reliabilitet og validitet kan det også trygt påstås at resultater fra ettertest er en følge av selve undervisningen og ikke av andre omstendigheter, med mindre det undervises for selve FCI (som i memorisering) (Hestenes mfl.,

1992). Det kommer også fram at FCI-score på ettertest er uavhengig førtesten, og at underviser ikke kan skylde på andre omstendigheter enn sin undervisning, dersom FCI ettertest skulle vise presumptivt svake resultater. En annen nyttig anvendelse av FCI kan i følge Hestenes og Halloun (1995) være å bruke førtesten til å sette sammen grupper av studenter, og det yntes frempå at heterogene grupper – hva gjelder fordeling av score på FCI førtest – kan være fordelaktig. Selv om Hestenes mfl. (1992) slår fast at det er ettertesten som skal legge føringer for evaluering av undervisningen, vises det allikevel til at det å se på individuelle forbedringer fra førtest til ettertest og sammenligne disse med potensiell forbedring, kan være til nytte for å undersøke hvorvidt studenters utgangspunkt har noe å si for effekten av undervisningen. Dette er en måte å måle effekt av undervisning på. Vi skal komme tilbake til effektstørrelser i seksjon 2.6.1.

Hestenes og Halloun (1995) identifiserer to grenser som sier noe om studentees konseptuelle forståelse av newtonsk mekanikk. Den første er et *entry threshold*, og er en nedre grense for forståelse. Denne grensen finner vi ved en score på omtrent 60 prosent. I følge Hestenes og Halloun, vil studenter med score under dette kunne beskrives som å bare så vidt ha begynt å bruke newtonske konsepter på en sammenhengende og helhetlig måte i argumentasjon og resonnering. Studenter som scorer under 60 prosent kan også sies å ha en forståelse som er utilstrekkelig for effektiv problemløsning, og vil følgelig antas å slite i videregående emner innen fysikk. Den andre grensen som Hestenes og Halloun (1995) identifiserer er et såkalt *newtonian mastery threshold*. Hvor denne grensen ligger er ikke entydig fra litteraturen, men fra Hestenes og Halloun (1995) kommer det fram at denne kan sies å være ved en score på omtrent 85 prosent. Studenter med en score høyere enn dette kan sies å være newtonske tenkere, og har en god forståelse av den newtonske mekanikken, og et solid fundament for videre forståelse. Denne tolkningen av FCI-resultater som her er beskrevet over er i følge Hestenes og Halloun (1995) konsistent med det de kaller «a *three-stage model of conceptual evaluation in learning Newtonian mechanics*». Her deles forståelse av newtonsk mekanikk inn i tre stadier som grovt gjengitt er henholdsvis forståelse av kraftkonseptet, forståelse av dynamiske konsepter (herunder fart og akselerasjon), og en helhetlig forståelse av interaksjon i mekanikken (inkludert Newtons tredje lov).

2.4 Aktiv læring og Peer Instruction

Fra det konstruktivistiske bakteppet som er etablert, fremkommer det sentralt at elever og studenter i fysikk må konstruere sin kunnskap, heller enn å motta den. Dette innebærer direkte interaksjon med ideer og konsepter. Renkl, Atkinson, Maier og Staley (2002) påpeker viktigheten av at elever og studenter tar del i

såkalte «higher order thinking activities»: analyse, syntese og evaluering. Dette kjenner vi også igjen som de tre siste nivåene i Blooms taksonomi, og det er klart at disse aspektene ved læring krever en aktiv deltakelse fra den lærende. I aktiv læringssetting brukes den formelle undervisningen til å diskutere ideer, til å besvare spørsmål, til å oppklare forvirring – kort og godt utøve alle aktive aspekter ved faget, mens lærer er til stedet. Det er ikke her læreren sin jobb å overføre informasjon som studentene og elevene mottar passivt (Knight, 2004). Det er veldokumentert at den klassiske forelesningstilnærmingen til undervisning ikke er effektiv for å løfte forståelsen til den gjengse student eller elev, og at aktive undervisningstilnæringer i så måte gir bedre resultater (Hake, 1998; Knight, 2004; Van Heuvelen, 1991).

For å fasilitere for aktive undervisningssettinger er det ofte hensiktsmessig å ty til såkalt *omvendt undervisning* (Bergmann & Sams, 2012). Dette innebærer at elevene må tilegne seg informasjon på egenhånd før undervisning, i form av å lese, eller å se videoer. Informasjonen som er tilegnet blir videre behandlet aktivt i formell undervisning. En slik omstrukturering av undervisningen gir også den som underviser mulighet til å i større grad aktivt delta og bistå i læringsprosessen. Denne oppgaven fokuserer på Peer Instruction. Dette er en tilnærming til undervisning som både kan sies å være aktiv og basert på omvendt undervisning. I det følgende vil vi legge frem hovedessensen i denne undervisningsformen.

2.4.1 Peer instruction

Peer Instruction (PI) har etterhvert blitt velkjent og anerkjent som undervisningstilnærming, og det spesielt i fysikkfaget. For å kunne introdusere de sentrale aspektene ved PI er det ryddig å redegjøre for Eric Mazur (1997) sin motivasjon for å utvikle denne tilnærmingen til undervisning. Mazur har lang erfaring fra undervisning av introduksjonsemner i fysikk på universitetsnivå, og i mange år underviste han disse kursene på tradisjonelt og konvensjonelt vis. Det vil si at undervisningen i stor grad besto i enveis overføring av informasjon i form av klassiske forelesninger. Undervisningen var også preget av utledningner av sammenhenger og regning av eksempeloppgaver på tavla (Mazur, 1997). Mye tydet på at undervisningen var effektiv, all den tid studentene oppnådde gode resultater på eksamen, men Mazurs syn på egen praksis endret seg i møte med en serie artikler skrevet av Ibrahim Halloun og David Hestenes (I. A. Halloun & Hestenes, 1985a, 1985b, 1987; Hestenes, 1987). I disse artiklene ble det synliggjort hvor stor betydning studenters preeksisterende forestillinger og oppfatninger hadde for deres læring. Det var nytt for Mazur at disse ideene, intuisjonene og oppfatningene om fysiske fenomener som studentene hadde opparbeidet seg og videre tok med seg inn i fysikkundervisninga var svært resistente, og at konvensjonell fysikkundervisning gjorde lite for å endre på dette.

Dette ble grobunn for Mazurs utvikling av PI som undervisningsform.

Skillet mellom problemløsning og konseptuell forståelse står sentralt i meto-
dikken i PI. Tradisjonell undervisning gjør det mulig – i følge Mazur (1997) –
at studenter gjør det godt på konvensjonelle problemløsningsoppgaver uten å
ha tilegnet seg den grunnleggende konseptuelle forståelsen som oppgavene
bygger på. Dette gjør det også vanskelig for lærere og instruktører å vurdere
om studenter faktisk forstår det de gjør, og selv erfarne lærere kan være lykkel-
lige uvitende om denne diskrepansen hos studentene. En uheldig konsekvens
av konvensjonell undervisning er derfor at studenter blir primet til å tro at
nøkkelen til suksess er å terpe på løsningsstrategier, og å for eksempel lære seg
å bruke riktig formel til riktig problem – en oppfatning som kommer til kort
i møte med nye, mer komplekse problemer. Studentens fokus blir fjernet fra
det å faktisk forstå fysikken, og den kunnskapen studenten sitter igjen med
vil i stor grad være prosedyreorientert, og i svært liten grad konseptorientert
(jamfør seksjon 2.1.2).

Mazurs bekymring for hva hans studenter faktisk satt igjen med av konseptuell
forståelse etter konvensjonell undervisning ble bekreftet som målt ved FCI
– som presentert i seksjon 2.3.1 (Hestenes mfl., 1992). Løsningen til Mazur
ble å direkte konfrontere studentene med sine feilaktige oppfatninger og
forestillinger, og for å få til dette måtte strukturen for undervisningen endres.
PI baserer seg på en omvendt klasserom-tilnærming til undervisning, hvor
tilegning av informasjon skjer utenfor klasserommet eller forelesningssalen,
hvorpå den prosesseres i selve undervisningen. I det følgende vil strukturen i
en typisk PI-økt presenteres, som beskrevet av Mazur (1997).

2.4.2 Gjennomføring av Peer Instruction

Som nevnt må studentene gjennom PI i all hovedsak tilegne seg fagstoff
på forhånd. I Mazurs beskrivelse av et typisk PI-opplegg innebærer dette at
studentene blir bedt om å lese hans notater om tema til den kommende
undervisningssesjonen, i tillegg til relevant fagstoff fra den aktuelle læreboka.
Studentene må også avlegge en kort test som tar for seg det de har lest. Dette er
først og fremst for å sikre at alle er forberedt til undervisning, og testen er ikke
ment for å sjekke forståelsen til studentene som sådan. En mer konseptuell test
rett i forkant av formell undervisning kan allikevel ha fordelaktige effekter for
studentene, og er i tråd med det Cashman og Eschenbach (2003) kaller *Just-in-
Time-Teaching* (JiTT). Ved at læreren får kunnskap om hvilke deler av fagstoffet
som er problematiske for studentene «just in time», kan dette ble behandlet
mens det fortsatt er friskt i minne hos studentene. JiTT er dokumentert å ha
god effekt blant annet på studenters motivasjon (Cashman & Eschenbach, 2003;
Novak & Patterson, 1997). Lærer vil i PI bruke den informasjonen hen får ut

fra en slik førtest til å best mulig planlegge selve undervisningsøkta. I selve undervisningen behandles den informasjonen som studentene har tilegnet seg. Mazur (1997) beskriver at både demonstrasjoner og fragmenter av det som kan ses på som tradisjonelle forelesninger kan inngå i undervisningen, men at forelesningssekvensene nødvendigvis er kortere, mindre rigide, og mer situasjonsavhengige enn i konvensjonell undervisning.

Innholdet i en PI-økt i fysikkemner består gjerne av noen bestemte tema som skal gås gjennom. Underviseren presenterer gjerne ett tema om gangen, ikke som i en standard forelesning, men som en grov overordnet gjennomgang uten å bruke formler eller ligninger. Dette følges opp med en konseptuell oppgave som studentene får begrenset med tid til å svare på. Mazur (1997) kaller disse oppgavene for *ConceptTests*², og vi vil ta for oss disse nærmere senere i denne seksjonen. Studentene skal i første omgang avgi et individuelt svar på oppgaven. Når dette er gjort skal studentene i grupper diskutere og argumentere for sine respektive svar. Studentene får noen få minutter til dette. Gitt for mye tid, vil studentene ty til problemløsningsstrategier i både argumentasjonen og i jakten på riktig svar. Etter endt diskusjon blir den samme oppgaven gitt på nytt og studentene har mulighet til å endre sine individuelle svar. Ut i fra studentenes svar før og etter diskusjonene kan lærer få en umiddelbar respons på hvorvidt studentene forstår det temaet som gjennomgås. Om mange svarer galt også etter diskusjon, er det et tydelig tegn på at temaet krever å behandles grundigere før man går videre til neste tema og dermed neste konseptuelle oppgave.

Det fremkommer tydelig at de konseptuelle diskusjonsoppgavene står sentralt i gjennomføringen av PI, og Mazur (1997) peker på at å lage, plukke ut og komponere utvalget av disse før en undervisningsøkt er en avgjørende faktor for utfallet av økta. Videre pekes det på noen kriterier som et god konseptoppgave bør oppfylle. For det første skal en enkelt oppgave fokusere på kun ett enkelt konsept, og oppgaven skal ikke enkelt kunne løses ved å lene seg på ligninger og formler. Videre skal oppgaven være utvetydig formulert og ha en vanskelighetsgrad som hverken er for høy eller lav. Mazur (1997) peker på at forbedringen blant studentene oftest er størst når omtrent femti prosent svarer korrekt før diskusjon. Det siste kriteriet Mazur peker på er at oppgavene skal ha adekvate flervalgssvar. I dette legger han at ukorrekte svar ideelt sett bør speile studenters mest vanlige misoppfatninger og preeksisterende forestillinger. Det poengteres også at det er studentene selv som bør legge grunnlaget for å identifisere hva som er «gode» gale svar. Ved å lage oppgaver som krever korte tekstvar, kan underviseren telle opp de vanligste gale svarene, og dermed basere fremtidige oppgaver på disse.

2. Mazur kaller faktisk hele prosessen og konteksten som oppgavene inngår i for *ConceptTests*, men vi peker her kun på selve oppgavene i eventuelle videre henvisninger til *ConceptTests*.

2.4.3 Forutsetninger for at Peer Instruction skal fungere

Ved å diskutere oppgavene med hverandre i grupper får studentene muligheten til å artikulere sin fysikkforståelse. De må resonnerer, forklare til hverandre og argumentere for sitt syn. Dette påpeker Mazur (1997) som fordelaktig siden studenter ofte kan forklare ting til hverandre på en måte som er mer hensiktsmessig og effektiv enn det en lærer eller foreleser vil være i stand til. Studenter som nylig har forstått et konsept i fysikken vil være i kontakt med de aspekter av konseptet som er utfordrende og vanskelige, og dette gjør dem i stand til å vektlegge de mest prekære aspektene når de forklarer. Mazur (1997) fremhever det faktum at det er rimelig å anta at nettopp de utfordringer ved fysikkforståelse som er identifiserbare i en grunnleggende fase av læring, i mindre grad er tilgjengelige og synlige for en lærer, foreleser eller professor som har behandlet emnet over mange år. En lærer vil i sin forklaring av et fysikkonsept kanskje vektlegge de presumptivt mest sentrale elementene, som forstått av en ekspert, heller enn de aspektene som er prekære for en novises forståelsen av konseptet.

Den sentrale posisjonen som gruppediskusjoner har i PI gjør at det er høyst prekært for læreren å fasilitere og legge til rette for at studentene eller elevene trygge i denne settingen. Mazur (1997) påpeker hvor viktig det er at læreren er transparent, ærlig og tydelig til studentene i forklaringen av hvordan undervisningen skal forgå. Mazur argumenterer videre for at læreren selv må ta på seg en aktiv rolle i undervisningen og interagere med studentene i deres diskusjoner. Dette gir studentene trygghet, og gir også lærer mer innsikt. Ved å merke seg hvordan studentene diskuterer og forklarer til hverandre kan læreren også bearbeide sine egne forklaringsstrategier. For at PI skal fungere må studentene oppeve et fokus på samarbeid og fravær av konkurranse i undervisningen. Et siste viktig aspekt er at undervisningen også gjenspeiler seg i eksamen. En avsluttende eksamen i et fysikkemne undervist ved Peer Instruction, må i følge Mazur ha et substansielt antall konseptuelle oppgaver.

Det kan argumenteres for at det å argumentere, resonnerer og forklare konsepter til medelever vil føre til læring i tråd med det sosialkonstruktivistiske perspektivet på læring. Å få innspill og synspunkter fra andre, kan føre til at kognitive strukturer modifiseres hos ett individ. Samtidig vil det å forklare og instruere andre gjøre at man blir bevisst hvilke virkemidler, begreper, og øvrige forklaringsstrategier man drar veksler på, og man vil veldig ofte måtte hoppe mellom ulike måter å forklare på. Slik blir man også bevisst sine epistemologiske ressurser (jamfør Hammer (2004) i seksjon 2.1.1). Diskusjonene i en Peer Instruction-setting har en sentral posisjon i denne oppgaven og i det følgende skal vi prøve å finne noen holdepunkter for hva som egentlig inngår i en produktiv diskusjon i denne sammenhengen.

2.5 Diskusjoner i undervisningen

2.5.1 Språk og kommunikasjon i fysikk

Det er allerede blitt gjort klart at språk spiller en sentral rolle i utvikling av kunnskap og i læring i fysikk. Fysikkfaget i seg selv har et språk som preges av tydelige kjennetegn, og som må tilegnes og beherskes dersom man skal lære seg faget. Vi har allerede vært innom begreper som har konkrete betydninger i fysikkfaget, men språket har også tydelige karakteristika i sin fremstilling og syntaks (Angell mfl., 2019). For det første er fysikkspråket multimodalt, og et konsept eller fenomen kan representeres på mange ulike måter og fra mange ulike synsvinkler. Nomaliseringer – å transformere verb til substantiv – er også et typisk kjennetegn. Foreksempel er begrepet «stråling» avledet av verbet «å stråle» (Angell mfl., 2019). Språket i fysikk er også – som i andre vitenskapelige sjangere – formelt, og preges av bruken av passiv form. Alle disse trekkene er ting som en fysikkelev eller fysikkstudent må lære seg å beherske, for å også kunne beherske selve fagstoffet. Dette kan vanskelig gjøres uten å nettopp bruke språket aktivt.

Det er mye som kan greies ut om vedrørende fysikken og naturvitenskapen som en kommunikasjonssjanger, men i denne oppgaven er det mer hensiktsmessig å fokusere på ulike kommunikasjonsformer i undervisningen. Angell mfl. (2019) viser til Mortimer og Scott (2003) sitt rammeverk for kommunikasjonsformer i klasserommet som kartlegger hvorvidt diskusjoner er interaktive eller ikke og hvorvidt de er dialogiske eller autoritative (med autoritativ menes at kunnskap formidles fra lærer til elever, og hvor elevenes synspunkt ikke kommer i spill). I lys av det fokuset vi har på aktiv læring og Peer Instruction er det i disse settingene ønskelig at brorparten av kommunikasjonen er dialogisk og interaktiv. Angell mfl. (2019) viser til Bakhtin og Holquist (1981) som sier at det er vår personlige respons og tolkning av andres utsagn som legger premissene for hvordan vi selv tenker og fortår. Bakhtin og Holquist påpeker at korrigeringer fra læreren er nødvendig for at dialoger skal utvikle faglig forståelse, men at lytting alene aldri vil kunne resultere i korrekt tenkning. Innspill fra andre må bearbejdes og responderes på. Bakhtin og Holquist beskriver på treffende vis den personlige formuleringen man enten utøver høylytt eller som en ren mental øvelse som respons på andres argumenter, som «det indre overbevisende ordet».

2.5.2 Argumentasjon og resonnering

Venville og Dawson (2010) fant i en studie som omhandlet argumentasjon i klasserommet, at det å fokusere og trene på argumentasjon i realfagsundervisning har en positiv innvirkning på elevers argumentasjonskompetanse, men

også på deres forståelse av det faglige innholdet. For å motivere sine funn refererer Venville og Dawson blant annet til Von Aufschnaiter, Erduran, Osborne og Simon (2008). Her hevdes det at elever som aktivt deltar i argumenterende diskusjoner har større sannsynlighet for å oppnå interesse for innholdet i undervisningen enn andre elever. Derfor er det fornuftig å også forvente økt forståelse av fagets innhold hos disse elevene, slik at sammenhengen mellom argumentasjonskompetanse og fagforståelse kan sies å være toveis.

Å dekomponere et argument og antyde hva som inngår i et argument av høy kvalitet er ikke nødvendigvis en enkel øvelse. Det er vanlig å ty til en modell av Toulmin (2003) – noe blant andre Venville og Dawson (2010) gjør – for å rangere argumenter etter deres innhold. Her er et argument av høy kvalitet kjennetegnet ved at det i tillegg til en påstand, inneholder data som støtter opp under påstanden, en klar forbindelse mellom påstand og data, og dessuten betraktninger rundt gyldighetsområdet til påstanden.

2.5.3 Kvalitet på diskusjoner

Knight (2004) beskriver at det er kjennsgjerning at mange studenter og elever ikke vet hva det innebærer å kunne forklare noe. I stedet for faktisk å forklare fenomener og konsepter, tyns det ofte til unnvikende strategier som for eksempel å resitere en formel med ord, eller gjenta et spørsmål med en annen formulering. Med utgangspunkt i tanken om at interaktiv undervisning og elevdiskusjoner i klasserommet er fordelaktige, er det dermed av interesse å ha noen pekepinner på hva som er kjennetegn på en diskusjon av høy kvalitet. Det er nemlig ikke nødvendigvis alltid sånn at elevdiskusjoner er konstruktive hva gjelder elevens læring. Angell mfl. (2019) poengterer at produktive diskusjoner fordrer at elever bygger på hverandres utspill og utfordrer hverandre til å utdype og forklare ulike synspunkt. Videre kategoriserer Angell mfl. gjennom Mercer og Littleton (2007) produktive diskusjoner i to underkategorier: *kumulative* diskusjoner og *eksplorerende* diskusjoner. Førstnevnte innebærer at elever argumenterer ved å bygge på hverandres utsagn uten å nødvendigvis utfordre disse. Eksplorerende diskusjoner vil derimot innebære at elevene også utfordrer og argumenterer mot hverandre.

Det er som sagt ikke alle elevdiskusjoner som er formålstjenlige, og mange elever evner hverken å bygge på, eller utfordre, andres resonneringer og påstander (Mercer & Littleton, 2007). På samme måte som undervisning i argumentasjonskompetanse kan legge grunnlag for økt forståelse gjennom argumentasjon, viser det seg at kvaliteten på elevers dialoger vil øke dersom de får bevisst opplæring i samtale som er utforskende i den forstand at de oppfordres til å utfordre hverandre og etterspørre hverandres synspunkter (Angell mfl., 2019). Videre peker Angell mfl. på at læreren spiller en viktig

rolle for å fasilitere diskusjoner av god kvalitet. Det er en nøkkel at elevenes dialoger blir satt inn i en sammenheng som strekker seg ut over deres eget kunnskapsområde, slik at de også etter dialogen kan se at de kommer videre i utviklingen av faglig forståelse. Derfor er det i følge Angell mfl. prekært at læreren vet hva som foregår i elevdiskusjonene, slik at hen kan kontekstualisere det de har diskutert i lys av fagstoffet for elevene.

Det kan pekes på både teoretisk og empiriske argumenter for at det å tilegne seg kunnskap som en epistemologisk prosess, krever at studentene eller elevene både beskriver, forklarer, spår, argumenterer, vurderer og kritiserer (Osborne mfl., 2016). Osborne mfl. viser til Chi (2009) som definerer tre typer utfall av læringsaktiviteter. De kan være *aktive*, *konstruktive* eller *interaktive*. I denne definisjonen forsås aktiv som at studentene «gjør noe». Et utfall som er konstruktivt innebærer at studentenes undervisningsaktivitet skal kuliminere i et bestemt produkt, mens et interaktivt utfall krever at elevene både deltar i dialogisk aktivitet og produserer noe. Ikke overraskende kom Chi fram til at tilnærminger som ga interaktive utfall var mer effektive enn de som ga konstruktive utfall, som igjen var mer effektive enn tilnærminger med aktive utfall. Osborne mfl. (2016, s. 4) går videre med dette og sier at det han kaller diskursive interaksjoner må kreve at studentene ikke bare konstruerer argumenter, men også evaluerer og bedømmer andres:

While constructing explanations or argument is an activity that undoubtedly has value, deeper understanding requires students to identify flaws in both their own reasoning and that of others. Without such opportunities, human reasoning suffers from confirmation bias (Mercier & Sperber, 2011) and is commonly flawed.

2.5.4 Gruppesammensetning og gruppedynamikk

Denne oppgaven tar spesielt for seg gruppediskusjoner i en Peer Instruction-setting, og derfor er det på sin plass å ta opp noen elementer vedrørende gruppedynamikk og gruppesammensetning som må være til stede for at læring og forståelse skal kunne skje gjennom disse diskusjonene. For at en gruppe skal kunne oppnå gode resultater, så må det legges opp til et godt samarbeid innad i gruppa. Det ble i seksjon 2.4.1 pekt på at en opplevelse av samarbeid blant elevene er en forutsetning for at PI skal fungere, og dette gjelder også innad i diskusjonsgruppene. Samarbeidslæring er et stort felt innen pedagogikken, og her vil vi peke på noen konkrete elementer som kjennetegner såkalte *høytpresterende samarbeidsgrupper* som forklart fra Johnson, Johnson, Haugaløkken og Aakervik (2014).

Høytpresterende samarbeidsgrupper er i følge Johnson mfl. (2014) sjeldne, og

ikke noe man umiddelbart kan forvente å oppnå ved arbeid i grupper. Johnson mfl. definerer ulike typer grupper ut i fra blant annet gruppesammensetning, gruppemedlemmers motivasjon, evne og vilje til å samarbeide, og prestasjonsnivå, og det påpekes at for at en gruppe skal fungere, så må den være mer produktiv enn om gruppemedlemmene jobbet individuelt. Det er ikke ønskelig å lage såkalte *pseudogrupper*, hvor gruppearbeidet kun blir en byrde på produktiviteten. Slike grupper kjennetegnes av at medlemmene ikke kommuniserer og samarbeider, og at noen lurer seg unna eller er «gratispassasjerer». En *samarbeidgruppe* er derimot en gruppe som i sin sammenheng og dynamikk legger til rette for et prestasjonsnivå langt over det individuelt arbeid ville gjort. Slike grupper har ifølge Johnson mfl. (2014) noen typiske kjennetegn som nå vil bli presentert.

For det første har alle medlemmer en opplevelse av å ha et felles mål, noe som vil føre til økt motivasjon til å gjøre en innsats utover den rent individuelle presentasjonen. Dessuten pekes det på følelsen av ansvar. I en samarbeidsgruppe kjenner medlemmene både på ansvar for egen innsats, så vel som for gruppen som helhet. De holder hverandre ansvarlig for å legg ned et tilfredsstillende arbeid, men krever like mye av seg selv. De to neste kjennetegnene går på det sosiale aspektet. For det første er det viktig at gruppen jobber sammen ansikt til ansikt, og gir hverandre nødvendig faglig og sosial støtte. Videre vil en velfungerende gruppe også føre til at medlemmene lærer og utvikler sine sosiale ferdigheter. Det siste kjennetegnet for samarbeidsgrupper er at evaluering og forbedring av læringsprosessen vektlegges. I en velfungerende gruppe viser både enkeltpersonene, og gruppa som helhet evne til selvrefleksjon og vurdering av eget arbeid. Hvis en gruppe oppfyller alle disse aspektene, og utvikler et godt samhold hvor alle har omsorg for hverandre og vilje til å gjøre det beste for at gruppa skal lykkes, kan prestasjonen til gruppa overstige alle forventninger, og gruppa omtales da som en *høytpresterende samarbeidsgruppe*. Det er viktig å påpeke at læreren spiller en viktig rolle i å fasilitere for å kunne oppnå gode grupper, og hen er nødt til å observere gruppearbeidet, samt gripe inn og korrigere ved behov, for å best mulig sørge for at gruppene blir velfungerende.

Det er også som oftest lærerens jobb å sette sammen gruppene i første instans, og det er en del ting som bør tas med i betrakningen når elever deles inn i grupper. For det første er gruppestørrelse en faktor. Johnson mfl. (2014) viser til noen tommelfingerregler hva gjelder antall gruppemedlemmer. En generell tanke er «jo færre, jo bedre», og Johnson mfl. peker på at i miljøer hvor samarbeidslæring ikke er innarbeidet kan det være lurt å dele inn i grupper på to til tre personer. Jo mindre gruppen er, jo lettere er det for læreren å identifisere elevenes vansker i gruppearbeidet. Dessuten vil det være vanskeligere for enkeltelever å sluntre unna og ikke bidra. Til tross for dette er det også noen fordeler i å også ha større grupper. Blant annet vil flere

gruppemedlemmer føre til større spennvidde i kunnskapsnivå, ferdigheter og ekspertise, noe som kan gagne det endelige resultatet. Johnson mfl. fremhever allikevel at læreren må være klar over at større grupper trenger mer oppfølging, og at store grupper krever mer av gruppemedlemmene – både faglig og sosialt – og blir gruppa for stor, vil det føre til minimalt med samspill.

En viktig del av samarbeidslæring er det at elevene på gruppa opplever gjensidig avhengighet. Hvis ikke det er til stede, vil gruppa i hverste fall fungere dårligere enn det individuelt arbeid ville gjort (Johnson mfl., 2014). Johnson mfl. påpeker blant annet at det kan være en god idé å gi hvert gruppemedlem spesifikke arbeidsoppgaver. Gjensidig avhengighet kommer også til spille når det gjelder det faglige nivået til medlemmene. Grupper hvor alle deltagere er på samme faglige nivå kan være lite optimale. I slike grupper vil medlemmene kunne få en følelse, enten av at det ikke er noe å hente fra de andre utover det man allerede vet og kan, eller av at man ikke selv har noe å bidra med til fellesskapet. Johnson mfl. mener at man generelt bør strebe etter å lage heterogene grupper i henhold til faglig nivå og prestasjoner. I heterogene grupper vil mer utførlig tenkning finne sted, medlemmene vil være mer engasjerte, og vil ta hverandres perspektiv når de forklarer og diskuterer. I henhold til blant andre Osborne mfl. (2016) vil dermed heterogene grupper gi diskusjoner av høyere kvalitet. Ved at elevene i større grad vil forklare og rettlede hverandre i heterogene grupper, oppnår man også det som er ønskelig i aktiv læring og Peer Instruction. Elevene tar aktivt del i fagstoffet og underiser hverandre. Alt dette vil, også i følge Johnson mfl. (2014), bidra til å øke forståelse og læring.

2.5.5 Uenighet

Læring i form av kognitive konflikter kan skje basert på førstehånds erfaringer i den virkelige verden, eller gjennom en god konseptuell oppgave i fysikktimen. Ofte må det andre personer til for at en kognitiv konflikt skal kunne oppstå. I diskusjoner med medelever eller medstudenter hvor forskjellige oppfatninger og syn deles og argumenteres for, kan en elev eller student få korrigert sine alternative forestillinger gjennom argumentene og refleksjonene til andre. For at dette skal skje må det være en uenighet mellom synspunkter i utgangspunktet. Johnson mfl. (2014) peker på at det er ønskelig å legge til rette for faglig uenighet (konflikt) for å oppnå gode diskusjoner og gode prestasjoner fra en gruppe, og at det å være uenige om fagstoffet kan virke positivt på elevens evne til logiske resonnementer og ha en langsiktig læringseffekt.

2.6 Relevant statistikk

2.6.1 Effektstørrelser

I undersøkelser av data ønsker man å bruke mål som sier noe om resultatene som produseres. Ofte er man interessert i å undersøke statistisk signifikans – for eksempel om det er grunnlag for å si at det er korrelasjon mellom to datasett. Et ofte like interessant type mål baserer seg på å rett og slett kvantifisere differansen mellom to grupper av data. Slike mål kalles ofte for effektstørrelser (Coe, 2002). Det verserer mange forskjellige definisjoner av, og eksempler på effektstørrelser i litteraturen. I et forsøk på å sammenfatte ulike aspekter ved det som defineres som effektstørrelser, kommer Kelley (2012) med følgende definisjon:

Effect size is defined as a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest. Our definition of effect size is more than the combination of effect and size, as our definition depends explicitly on addressing a research question of interest (Kelley, 2012).

Effektstørrelser er altså i sin natur knyttet til det som undersøkes, og det betyr at ulike effektstørrelser kan være hensiktsmessige å bruke i ulike sammenhenger.

Ofte er det høyst interessant å for eksempel se på differanse mellom sentralverdier i to datasett. Dette kan – relevant for denne oppgaven – være differansen mellom gjennomsnittlig testscore på FCI førtest og ettertest, eller gjennomsnittlig antall korrekte før og etter diskusjon blant studentene på en bestemt konseptoppgave. En slik effektstørrelse refereres ofte til som ustandardisert (Kelley, 2012). I det ligger at den ikke tar hensyn til variabiliteten i datamaterialet. Ustandardiserte effektstørrelser vil av naturlige årsaker oppgis i samme enhet som den data de forsøker å måle, og de er ofte hensiktsmessige å bruke når datamaterialet som undersøkes har en iboende mening eller representeres ved en allment kjent skala (Fritz, 2011). På den andre siden har vi standardiserte effektstørrelser som er skalerte med hensyn til variabiliteten til dataene som undersøkes (Baguley, 2009). Dette gjør at de gir et mer generelt grunnlag for å måle statistisk styrke (Kelley, 2012). For eksempel vil man ofte, når man sammenligner to mengder, ta hensyn til variansen til de to mengdene. I det følgende vil de effektstørrelsene som er blitt brukt i denne oppgaven presenteres.

Gain

I denne oppgaven undersøkes datamateriale som er av typen «før, etter». Det vil si at effektstørrelser er relevante å undersøke. Blant annet tar vi for oss resultater på FCI før og etter undervisning, og besvarelser av konseptoppgaver før og etter diskusjon. Den mest naturlige størrelsen å se på her er differansen mellom to gjennomsnitt. Denne størrelsen vil i vi i denne oppgaven referere til som *gain* der ting ikke presiseres i nærmere detalj, og vi kalkulerer gain på følgende vis:

$$\langle G \rangle = \langle S_f \rangle - \langle S_i \rangle .$$

Her er $\langle G \rangle$, $\langle S_f \rangle$ og $\langle S_i \rangle$ henholdsvis gain, snittscore før en læringsprosess og snittscore etter endt læringsprosess. En læringsprosess refererer i denne oppgaven til gruppediskusjoner (kortsiktige læringsprosesser) og hele undervisningsløpet (langsiktig læringsprosess). Et slikt gain i forbindelse med for eksempel FCI kan si noe om hvordan en student eller en klasse har forbedret sin konseptuelle forståelse (Hake, 1998). Hvis man er kjent med datamaterialet kan slike gain være veldig megetsigende for resultatet av læringsprosessene. allikevel så vil ikke et slikt gain ta forbedringspotensiale i betraktning. Eksempelvis vil en student som har levert et godt resultat på FCI førtest, score lavt på læring målt ved et slikt gain – svake resultater på førtest favoriseres. Dermed foreslår Hake (1998) et mål som tar hensyn til dette, og vi definerer her *normalisert gain*, $\langle g \rangle$, som følger:

$$\langle g \rangle = \frac{\langle G \rangle}{\langle G \rangle_{\max}} = \frac{\langle S_f \rangle - \langle S_i \rangle}{100 - \langle S_i \rangle} .$$

Dette er altså forholdet mellom faktisk gain og det potensielt høyeste gain (hundre prosent), $\langle G \rangle_{\max}$, slik at normalisert gain er et mål på hvor mye av forbedringspotensialet som har blitt utnyttet. Det kan være av interesse å finne gjennomsnittet av individuelle normaliserte gains, så vel som å finne normaliserte gain av gjennomsnittresultater. Selv om normalisert gain er en normalisert effektstørrelse, tar den ikke i betraktning variabiliteten i data, og er dermed ikke standardisert.

Cohens *d*

En ofte brukt standardisert effektstørrelse er Cohens *d* (Cohen, 1988). Utgangspunktet for denne er differansen mellom gjennomsnittene til de to populasjonene man undersøker, $\mu_1 - \mu_2$. Effektstørrelsen blir standardisert ved at man vekter over et standardavvik – vanligvis er dette et såkalt «pooled» standardavvik for de to gjennomsnittene (Baguley, 2009). Dette defineres som

$$\sigma_p = \sqrt{\frac{(\mu_1^2 + \mu_2^2)}{2}} ,$$

og Cohens d er dermed gitt som

$$d = \frac{\mu_2 - \mu_1}{\sigma_p}.$$

For å anskueliggjøre hva som er god effekt som målt ved denne effektstørrelsen definerer Cohen (1988) tre nivåer av effekt. En verdi på 0,2 svarer til et lavt nivå, 0,5 til middels og 0,8 til høyt. Altså vil en differanse i data som målt ved Cohens d være minimal ved en verdi under 0,2 og betydelig ved verdier over 0,8.

2.6.2 Annet

I denne oppgaven brukes lineær regresjon og hypotesetesting – nærmere bestemt en parret t-test – i de kvalitative analysene. For nærmere presentasjon av dette henvises leseren til Walpole, Myers, Myers og Ye (2007) eller annen passende litteratur innen grunnleggende statistikk.

/3

Metode

3.1 Forskningsdesign

Målet med denne oppgaven var blant annet å undersøke hva som kjennetegner konstruktive gruppediskusjoner i en Peer Instruction-setting. Motivasjonen for oppgaven lener seg på premisset om at diskusjoner av oppgaver er den dominerende kilden til læring i Peer Instruction. Diskusjoner av oppgaver kan være suksessfulle ved at mange på gruppa endrer svar til korrekt svar gjennom disse, men de er ikke garantert å være konstruktive i seg selv. Med bakgrunn i dette var det naturlig å velge en flermetodetilnærming i forskningsdesignet, nærmere bestemt et forklarende design (Creswell & Plano Clark, 2017). Et sikt forskningsdesign består av to distinkte faser, en kvantitativ og en kvalitativ. I den første fasen samles og analyseres kvantitative data, hvorpå kvalitativ dataanalyse gjøres i den andre fasen for å forklare og utvide funnene fra den kvantitative fasen. Med en slik tilnærming vil analysen av kvantitative data gi et overblikk og overordnet forståelse relatert til forskningsspørsmålene, mens den kvalitative biten vil kunne gi en mer presis og dypere forståelse av resultatene. I denne studien vil den kvalitative fasen ta for seg og foredle resultatene i forbindelse med forskningsspørsmålet som dreier seg om produktive diskusjoner. Et slikt design er vurdert som hensiktsmessig i denne studien fordi en kvalitativ del vil gi dypere innsikt i hva som skjer i selve gruppediskusjonene utover hvor mange som svarer korrekt på oppgavene.

Den kvantitative delen skal se hvorvidt Peer Instruction har fungert, men også undersøke sammenhengen mellom prestasjoner på konseptoppgavene og

utbytte av undervisningen. Vi ser også på om henholdsvis uenighet og andel korrekte svar før diskusjon har noe å si for resultater på diskusjonsoppgaver. Ut i fra dette blir produktive (suksessfulle) diskusjoner identifisert, og det blir gjort et utvalg av disse til den kvalitative delen. Denne kal gi mer innsikt i hva som kjennetegner produktive diskusjoner, og undersøke hvilke faktorer som spiller en rolle for at en diskusjon i seg selv skal være konstruktiv. Til å gjøre dette er utvalgte gruppediskusjoner blitt observert, transkribert og analysert. Analysen er inspirert av prinsippene i en tematisk analyse (Braun & Clarke, 2006), hvor diskusjonene er blitt gått gjennom flere ganger for å plukke opp gjennomgående mønstre eller tema.

3.2 Utvalg

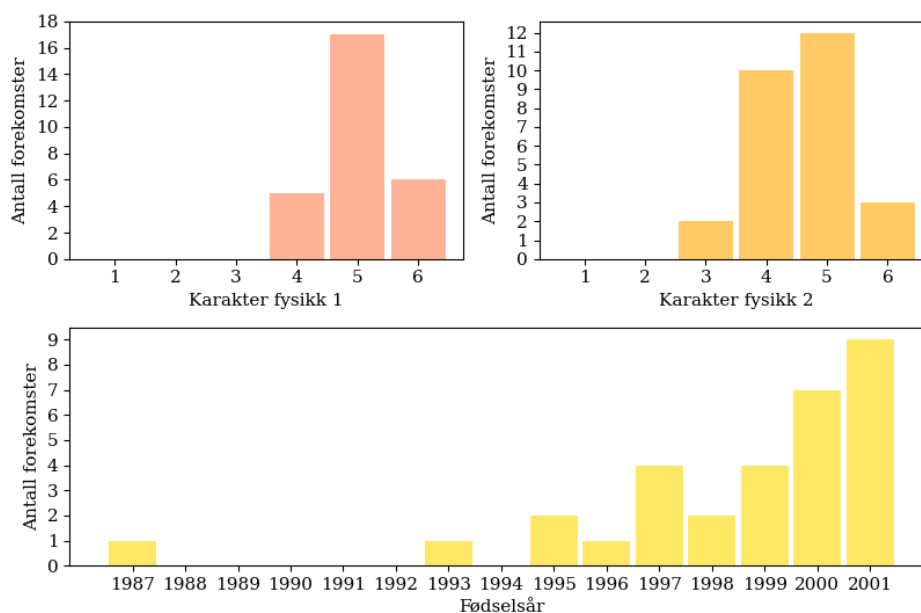
Deltagere i studien er studenter som tok emnet FYS-0100 generell fysikk ved UiT – Norges arktiske universitet høsten 2020. Totalt 38 studenter av 67 som avla eksamen, ble rekruttert gjennom frivillig samtykke uten noen form for insentivordning. Dette er et introduksjonsemne i fysikk og for de fleste studentene er det et førstesemesteremne. Det vil si at majoriteten av deltagerne er førsteårsstudenter. Studentene tilhører studieprogrammene *sivilingeniør i anvendt fysikk og matematikk*, *sivilingeniør i energi, klima og miljø*, *sivilingeniør i romfysikk* og *bachelor i romfysikk*. Dette er alle studier hvor dette emnet gjennomføres første semester, og som har flere og mer avanserte fysikkfag videre i studieløpet.

Gjennom en spørreundersøkelse laget i forbindelse med IMPEL-prosjektet ved universitetet i Oslo, som ser på fysikkstudenters læring og motivasjon¹, ble alle studenter som tok emnet bedt om å oppgi en del informasjon om seg selv i tillegg til å svare på motivasjonsrelaterte spørsmål. Deltagere som deltok i denne undersøkelsen var med i trekningen av tre gavekort, ett på 1000 kroner og to på 500 kroner. Personlig informasjon om studentene blir her brukt til å gi en grov oversikt over hele studentgruppa (utover de som samtykket til å delta i denne studien). Det var 31 studenter som tok IMPEL. I tabell 3.1 er fordelingen av studieprogram, kjønn, mors utdanningsnivå og farsutdanningsnivå presentert for alle studenter som oppga denne informasjonen. Studentenes fordeling av fødselsår fysikkarakterer fra videregående vises i tillegg i histogrammene i figur 3.1. Fra dette framkommer at en gjennomsnittsstudent (som svarte på IMPEL) er født i 1998 og har henholdsvis karakter 5,05 og 4,59 i fysikk 1 og fysikk 2 fra videregående.

1. <https://www.mn.uio.no/fysikk/forskning/prosjekter/impel/>

Tabell 3.1: Informasjon om studentene,

Kjønn		
Mann	17	
Kvinne	14	
Studieprogram		
Antall (N = 31)		
Sivilingeniør, AFM	6	
Sivilingeniør, EKM	9	
Sivilingeniør, romfysikk	9	
Bachelor fysikk	2	
Geologi	2	
Bachelor kjemi	1	
Bachelor matematikk	1	
Matematikk og Statistikk	1	
Foresattes utdanningsnivå		
Mor	Far	
Ungdomsskole	3	2
Videregående skole	2	3
Fagskole	1	4
Bachelorgrad eller tilsvarende	8	6
Mastergrad eller tilsvarende	14	14
Doktorgrad eller tilsvarende	3	2
Totalt antall		
31		



Figur 3.1: Fordelingen av karakterer i fysikk 1 og fysikk 2, samt fødselsår for studentene.

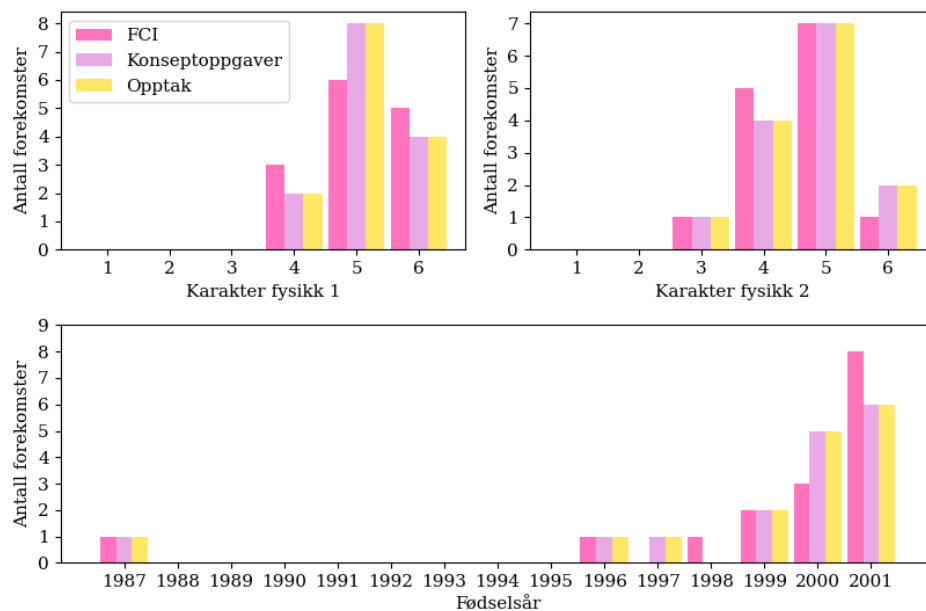
Uavhengig av IMPEL oppga studentene samtykke til at ulike deler av undervisningen kunne brukes som datamateriale. Relevant for denne oppgaven er samtykke til å bruke FCI-data, samtykke til å bruke svarene på konseptoppgaver i undervisningen og samtykke til opptak av diskusjoner. Data ble filtrert på samtykke etterhvert som de ble behandlet. I tabell 3.2 samt i figur 3.2 finner vi informasjon om studentgruppen (som også deltok på IMPEL) som ga samtykke til bruk av FCI-data, bruk av oppgavedata fra undervisningen, og opptak. Siden ikke alle av deltagerne i denne studien også deltok på IMPEL, vil ikke alle være representert her.

Tabell 3.2: Informasjon om studenter filtrert på FCI, konseptoppgaver (KO) og opptak.

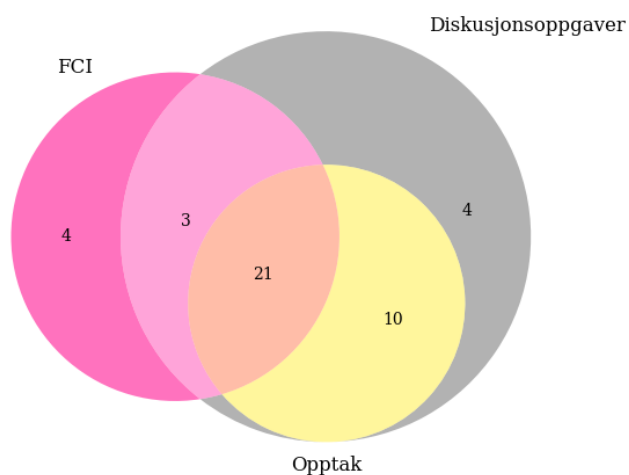
Studieprogram	FCI	KO	Opptak
Sivilingeniør, AFM	5	5	5
Sivilingeniør, EKM	5	6	6
Sivilingeniør, romfysikk	5	5	5
Bachelor fysikk	1	1	1
Kjønn			
Mann	10	10	10
Kvinne	6	7	7
Mors utdanningsnivå			
Ungdomsskole	1	1	1
Videregående skole	0	1	1
Fagskole	1	1	1
Bachelorgrad eller tilsvarende	6	7	7
Mastergrad eller tilsvarende	5	5	5
Doktorgrad eller tilsvarende	3	2	2
Fars utdanningsnivå			
Ungdomsskole	0	0	0
Videregående skole	1	3	3
Fagskole	3	3	3
Bachelorgrad eller tilsvarende	3	2	2
Mastergrad eller tilsvarende	8	9	9
Doktorgrad eller tilsvarende	1	0	0
Totalt antall	16	17	17

Av alle deltakere var det 28 stykker som gav samtykke til innsyn i, samt deltok på, FCI. Tilsammen 38 studenter svarte på diskusjonsoppgaver i undervisningen, og ga samtykke til innsyn i disse. Det var også 31 stykker som deltok i gruppesamtaler og som ga samtykke til innsyn i disse. Av disse var det 24 som både tok FCI og svarte på diskusjonsoppgaver, 21 både tok FCI og deltok i gruppediskusjoner, mens 36 stykker deltok både i diskusjonsoppgaver og svarte på diskusjonsoppgavene. En oversikt over deltagerne i studien finnes som et

Venn-diagram i figur 3.3



Figur 3.2: Fordelingen av karakterer i fysikk1 og fysikk 2, samt fødselsår for studenter med henholdsvis FCI-samtykke, konseptoppgavesamtykke og samtykke til opptak.



Figur 3.3: Oversikt over antall deltagere i de forskjellige aspektene i studien.

3.3 Gjennomføring av emnet

Undervisningen i FYS-0100 generell fysikk, høsten 2020 besto av tre forskjellige hovedaktiviteter, i tillegg til ukesoppgaver og fire skriftlige innleveringer. Fellesundervisning ble gjennomført to dager i uka i totimersøkter og ble gjennomført ved Peer Instruction. Ellers ble det arrangert regneverksteder (to ganger to timer i uka) og seminarer (en gang to timer i uka). Førstnevnte var ustrukturert regnehjelp, mens sistnevnte bestod i diskusjon av konseptuelle oppgaver i tillegg til gjennomgang av regneoppgaver. Siden denne oppgaven tar for seg det som foregikk i fellesundervisningen vil ikke regneverksteder og seminarer beskrives nærmere. På slutten av semesteret ble det avholdt en firetimers digital hjemmeeksamen, hvor halvparten av denne bestod av konseptuelle kortvarsoppgaver lignende diskusjonsoppgavene fra fellesundervisning², og den andre halvparten bestod av mer tradisjonelle regneoppgaver.

Pensum i kurset kan deles inn i tre deler: klassisk mekanikk, fluidmekanikk og termofysikk – hvor mekanikkdelen er den klart største. Fellesundervisningen var grovt sett lagt opp slik at et spesifikt tema ble tatt opp i løpet av en uke. Dette kunne for eksempel være tema som bevegelsesmengde eller gassloven. De mest omfangsrrike temaene ble behandlet over to uker. Deler av fellesøkterne ble satt av til små forelesningsseksjoner, samt enkle demonstrasjoner og simuleringer fra faglærer, mens størsteparten gikk med til at studentene svarte på konseptuelle oppgaver og diskuterte disse i grupper. Diskusjonsgruppene i fellesundervisningen var faste og bestod som regel av fem til seks deltagere. Gruppesammensetningene var ikke helt tilfeldig, men baserte seg på resultater fra FCI førtest, i den forstand at det ble lagt opp til at gruppene skulle ha lik gjennomsnittscore her.

Studentene skulle før hver fellesøkt forberede seg ved å lese tekster og se videoer som tok for seg det aktuelle fagstoffet. Forberedelsene som studentene skulle gjøre før undervisning ble vanligvis gjort tilgjengelige ved slutten av en uke med undervisning, slik at studentene fikk tid til å gjøre dem før første fellesøkt neste uke. Studentene skulle også avlegge en førtest rett før fellesundervisning. Dette gjorde de fleste kun få minutter før selve økta. Et visst antall gjennomførte førtester telte som en obligatorisk aktivitet for å kunne ta eksamen (uavhengig av resultat), og ble brukt av fagansvarlig til å kartlegge fokusområder før den aktuelle økta.

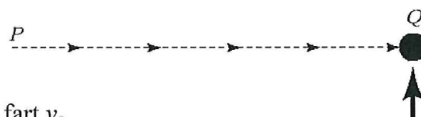
2. Noen av oppgavene fra undervisning ble sågar gitt uendret på eksamen.

3.4 Instrumenter

Denne undersøkelsen har tatt i bruk data fra Force Concept Inventory (FCI) for å måle studentenes konseptuelle forståelse og dermed utbyttet fra overordnet undervisning. I tillegg har svar og resultater på konseptuelle diskusjonsoppgaver i undervisningen blitt brukt for å kunne undersøke studentenes prestasjoner på disse, samt hva som kjennetegner diskusjoner som fører til at mange svarer riktig. Nå vil FCI og konseptoppgavene presenteres nærmere.

I denne oppgaven har FCI blitt brukt til å evaluere i hvor stor grad undervisningen har ført til bedret konseptuell forståelse for studentene, og til å undersøke om det er sammenheng mellom resultater på FCI og hvordan studenter gjør det på konseptoppgavene. FCI er en test, bestående av 30 flervalgsoppgaver, som måler konseptuell forståelse i grunnleggende newtonsk mekanikk³. Testen er utviklet av Hestenes mfl. (1992), og tar utgangspunkt i kraftkonseptet i mekanikken og spesielt alternative forestillinger rundt dette. Hvert spørsmål søker å dekke en bestemt dimensjon ved kraftbegrepet og svaralternativene dekker bestemte kjente misoppfatninger vedrørende den aktuelle dimensjonen. FCI brukes i diagnostisk øyemed, men også til evaluering av undervisning og forskning. I figur 3.4 finner vi en av oppgavene som ble gitt som en del av FCI. Som alle andre oppgavene i FCI har denne fem svaralternativer, hvor kun ett av disse er korrekt.

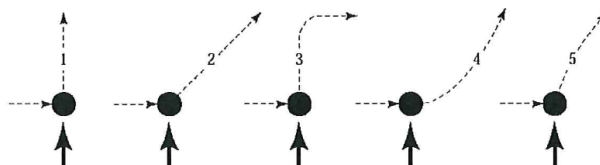
Figuren viser en ishockeypuck som glir på isen fra P til Q. Du ser ovenfra og ned på pucken. Se bort fra friksjonen. Pucken har retlinjet bevegelse med konstant fart v_0 .



I Q får pucken et kort horisontalt støt i retning av den tykke pila. Hvis pucken hadde ligget i ro ved Q, ville støtet gitt pucken horisontal bevegelse med fart v_k i støtretningen.

8. Hvilken bane beskriver best bevegelsen til pucken etter støtet?

- A. Bane 1
- B. Bane 2
- C. Bane 3
- D. Bane 4
- E. Bane 5



Figur 3.4: Eksempeloppgave hentet fra FCI.

Validiteten til FCI er utover all rimelig tvil (Hestenes & Halloun, 1995), noe som kan synliggjøres ved sannsynligheten for forekomster av falske positive

3. Fra Hestenes mfl. (1992) framkommer det at testen måler misforholdet mellom det newtonske kraftkonseptet og testtakerens forestillinger om dette. Det henvises til seksjon 2.3.1.

og falske negative. Sannsynligheten for falske negative anslås konservativt til å være under ti prosent, mens testens design sørger for at også forkomsten av falske positive er forholdsmessig lav (Hestenes & Halloun, 1995). Det er også blitt vist at FCI sin reliabilitet er høy. Blant annet viser analysene til Lasry mfl. (2011) høy indre konsistens og test-retest-reliabilitet. For en grundigere gjennomgang av FCI henvises det til seksjon 2.3.1.

I undervisningen skulle studentene diskutere konseptuelle oppgaver i grupper. Gode diskusjonsoppgaver i Peer Instruction bør i følge Mazur (1997) oppfylle en rekke krav. Blant annet skal et bestemt konsept være i fokus, og det skal hverken være hendig eller nødvendig å løse oppgaven ved regning eller bruk av formler. Hensikten med diskusjonsoppgavene i undervisningen er å konfrontere og utvikle studentenes konseptuelle forståelse. For nærmere beskrivelse av Peer instruction og diskusjonsoppgaver henvises leseren til seksjon 2.4.2. Siden denne studien ser på studentenes utvikling fra diskusjonene ble altså data fra disse oppgavene tatt i bruk. I figur 3.5 presenteres et eksempel på en konseptuell oppgave som ble brukt i undervisning. Denne oppgaven hadde ett korrekt svaralternativ.

1.

Person sitter i en robåt og har med seg en stor stein. Midt utpå innsjøen kaster personen steinen overbord. Vil da vannstanden i innsjøen

A. øke

B. synke

C. forbli uendret

Figur 3.5: Eksempel på konseptuell diskusjonsoppgave brukt i undervisningen.

3.5 Opptak

Fellesundervisningen foregikk digitalt i *Microsoft Teams*⁴, og gruppediskusjonene ble også gjennomført her (i private kanaler). Det var seks grupper som samtykket til at opptak kunne bli brukt. For å kunne gjøre opptak av samtalene krevdes tilgang til gruppekanalene. Opptakene ble gjort på til sammen seks forskjellige enheter – en til hver gruppe. Til dette ble programvaren *OBS Studio*⁵ brukt. Uoverensstemmelse mellom lyd og bilde var problematisk ved flere anledninger på en del av enhetene, trass i alle preventive tiltak. Flere opptak endte opp helt uten lyd eller helt uten bilde. I sum førte dette til at utvalget av opptak ble begrenset. Opptak uten lyd ble forkastet, mens alle opptak hvor det var mulig å høre hva som ble sagt i samtalene ble beholdt.

4. <https://www.microsoft.com/nb-no/microsoft-teams/group-chat-software>

5. <https://obsproject.com/>

3.6 Prosedyre

I det følgende presenteres denne studiens kronologiske forløp. Deretter de analyser som har blitt gjort på datamaterialet.

3.6.1 Studiens forløp

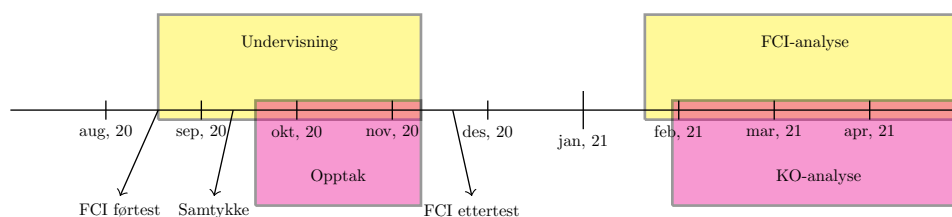
I fysikkemnet denne studien tar for seg ble FCI førtest gjennomført i første undervisningsuke (17. august 2020) og ettertest ble gjennomført i siste undervisningsuke (19. november 2020). På grunn av den digitale settingen hadde studentene 60 minutter til rådighet både på førtest og ettertest, mot 30 minutter som er konvensjonen. Testen ble gjennomført på *Canvas*⁶, og versjonen som ble brukt var en norsk oversettelse fra universitetet i Oslo. Samme versjon ble brukt både ved før- og ettertest. Resultatene fra førtesten ble brukt til å dele studentene inn i faste diskusjonsgrupper. Ellers ble ikke FCI brukt eller nevnt i undervisning, og studentene fikk ikke vite svaret på oppgavene.

Skjema for samtykke til deltagelse i ulike deler av studien ble fylt ut av studentene innen 10. september 2020. Det ble gjort opptak fra og med den 17. september 2020. Mellom 17. august 2020 og 9. november 2020 ble formell fellesundervisning gjennomført, og opptak ble gjort til og med den siste undervisningsøkta den 9. november. All fellesundervisning foregikk som Peer Instruction og store deler av hver økt gikk med til gruppevise diskusjoner av konseptoppgaver. Oppgavene ble fasilitert og besvart ved hjelp av *Learning Catalytics*⁷. De fleste av oppgavene i det aktuelle emnet var hentet fra Mazur (1997). En del supplementerende oppgaver ble laget av fagansvarlig, hvor noen av disse ikke var konseptuelle, men testet spesifikke ferdigheter. De aller fleste oppgavene var flervalgsoppgaver hvor det var ett korrekt svar, men noen hadde også flere riktige alternativer. Det ble også gitt kortsvarsoppgaver, tegneoppgaver og rangeringsoppgaver.

Den 20. januar 2021 ble data fra FCI før- og ettertest gjort tilgjengelig for analyse, mens data fra diskusjonsoppgavene ble tilgjengelig fra 28. januar 2021. En oversikt over tidslinjen for studien er illustrert i figur 3.6.

6. <https://www.instructure.com/en-gb/canvas>

7. <https://www.pearson.com/us/higher-education/products-services-teaching/learning-engagement-tools/learning-catalytics.html>



Figur 3.6: Tidslinje for studien.

3.6.2 Analyser

For å hente ut og prosessere de delene av datamaterialet som var nyttige for denne oppgaven ble datamaterialet fra FCI og diskusjonsoppgaver behandlet ved hjelp av Python⁸. Ingen oppgaver eller responser ble fjernet fra FCI-datamaterialet. For å evaluere læringsutbyttet fra undervisningen ble en rekke størrelser regnet ut. Gjennomsnittlig andel korrekte svar på henholdsvis førtest og ettertest ble funnet for studentgruppa som helhet, men også for hver enkelt student. Ut i fra dette ble gain og normalisert gain (det henvises til seksjon 2.6.1) regnet ut. For å kunne si noe om effekten til undervisningen, utover gain og normalisert gain, ble også Cohens d regnet ut. Verdiene av effektstørrelsene ble sammenlignet med de verdier som anses som gode i henhold til Mazur (1997) og Cohen (1988) (Det henvises til seksjon 2.6.1). Andelen korrekte svar på før- og ettertest ble plottet mot hverandre i et spredningsplott for å visualisere hvordan studentenes utvikling var. Fordelingen av andelen korrekte før og etter undervisning ble også visualisert ved å lage histogrammer. Det skal også nevnes at effektstørrelsene ble regnet ut som såkalte mathced gains. Det vil si at kun resultatene fra studentene representert i både førtest og ettertest ble tatt med.

Fellesundervisningen i det aktuelle fysikkemnet besto av 23 økter spredt utover høstsemesteret. Fra 22 av disse ble diskusjonsoppgaver med svar og resultater gjort tilgjengelige for denne undersøkelsen. Til sammen innebar det 159 oppgaver. Konseptoppgavene ble hentet inn i til sammen 22 .csv-filer – én svarende til hver undervisningsøkt. Her ble ufullstendig data fjernet, det være seg både konseptoppgaver om ikke var poenggivende og responser som av forskjellige årsaker ikke var mulig å behandle. I tillegg var det mange oppgaver som ikke var tolkbare gjennom fordi de var tegneoppgaver eller tekstvarsoppgaver. Disse ble nødvendigvis også fjernet fra datamaterialet. Til sammen ble 64 oppgaver fjernet, og 95 oppgaver ble dermed brukt til analyse.

8. <https://www.python.org/>

Etter preprosessering ble det gjort en rekke undersøkelser. For å kartlegge studentenes besvarelser, og oppgavenes vanskelighetsgrad ble gjennomsnittlig andel korrekte svar før diskusjon og etter diskusjon regnet ut per student og per oppgave. For å nærmere se på utbyttet studentene hadde av diskusjonene ble også verdier for normalisert gain og Cohens d funnet. Det ble laget et spredningsplott med fordelingen av oppgavene, ut i fra andel korrekte svar inn og ut av diskusjon. Histogrammer og spredningsplott ble også laget for studentenes gjennomsnittscore.

Videre ble sammenhengen mellom studentenes FCI-resultater og deres resultater på konseptoppgavene undersøkt. Det ble laget seks spredningsplott. Individuelle normaliserte gain fra konseptoppgavene ble plottet både mot individuelle normaliserte gain fra FCI og score på FCI ettertest, for å sjekke om det var en sammenheng mellom studentens grad av utvikling gjennom gruppediskusjonene og konseptuell forståelse som målt ved FCI. Et datapunkt svarende til en student med unormalt stort negativt gain ⁹ ble fjernet fra datamaterialet. Fordi diskusjonsoppgavene ikke er designet for å måle et konstrukt på samme måte som FCI, var det interessant å også bruke et mindre komplisert mål på prestasjon på diskusjonsoppgavene. Derfor ble snittscore på diskusjonsoppgavene henholdsvis før og etter diskusjon plottet mot FCI-score fra ettertest. I tillegg ble FCI-score på førtest plottet mot snittscore på konseptoppgavene henholdsvis før diskusjon og etter diskusjon. Dette ble gjort for å undersøke hvilken rolle studentenes nivå hva gjelder konseptuell forståelse har å si for hvordan de presenterer på konseptoppgavene i undervisningen. Lineær regresjon ble gjennomført for alle spredningsplottene for å sjekke for korrelasjon.

De som hadde gitt samtykke til innsyn i både diskusjonsoppgaver og opptak var fordelt på seks grupper. En *analyseenhet* ble definert som én oppgave diskutert av én gruppe. En del av disse ble fjernet på grunn av manglende svar innad i gruppene, og dette ga et gjenstående utvalg av 405 forskjellige analyseenheter. Gruppemedlemmer som enten ikke hadde svart før diskusjon eller etter diskusjon ble fjernet fra dataene, og dette medførte at analyseenhetene kunne ha *effektive* gruppestørrelser på en, to, tre, fire eller fem personer. Det fremkom videre at 63 av analysenhetene hadde effektive gruppestørrelser på færre enn tre personer. Fordi grupper på to ble ansett som for små til å gi representative resultater, ble det besluttet å fjerne disse. I søken etter å identifisere produktive diskusjoner, ble fordelingen av alle de resterende analyseenhetene undersøkt videre.

Uenighet innad i gruppene ble undersøkt for å se om dette spilte en rolle for utviklingen i diskusjonene. Det ble laget matriser som viste fordelingen av

9. Fra samtaler med fagansvarlig kom det fram at dette sannsynligvis var en student som ikke har prøvd å svare korrekt på ettertest

antall unike svar i hver diskusjon ut i fra effektiv gruppestørrelse. Fra dette ble tre «grader av uenighet» definert: *helt enig*, *helt uenig* og *hverken/eller*. Grunnet variasjonen i effektive gruppestørrelser ble det besluttet å ikke dele uenighet inn i flere kategorier. En rekke histogrammer med fordelingen av korrekthet og uenighet både inn og ut av diskusjon ble produsert for de ulike effektive gruppestørrelsene. På bakgrunn av de analyser som ble gjort ble det besluttet å definere produktive diskusjoner som diskusjoner hvor færre enn to gruppe-medlemmer hadde korrekt svar inn, og færre enn to gruppe-medlemmer hadde galt svar ut. Disse – til sammen 48 analyseenheter – ble valgt for videre analyse. De gjenværende 48 enhetene ble filtrert videre basert på hvilke diskusjoner det var blitt gjort opptak av, og til slutt sto 14 analysenheter igjen til kvalitativ analyse.

Opptakene av de 14 utsilede diskusjonene ble først sett gjennom en gang for overblikk. Deretter ble hver diskusjon gått gjennom på nytt, og potensielt interessante observasjoner ble notert ned. Ved tredje gjennomsyn ble alle diskusjonene transkribert. Analysen avslørte to hovedkategorier av diskusjoner. Det som tydeligst skilte diskusjoner fra hverandre, var hvorvidt det var kun én person, eller flere i gruppa som ledet an diskusjonen. Det var tydelig at endringen fra galt svar til riktig svar enten skjedde ved at en person overbeviste de andre, eller ved at to eller flere kom med sine argumenter og syn på oppgaven. På bakgrunn av dette ble alle diskusjonene gjennomgått på nytt og forsøkt klassifisert på denne måten. To av disse har blitt brukt som eksempel i denne oppgaven, en fra hver av de to observerte hovedkategoriene. For å undersøke om tendensene beskrevet ovenfor var spesielle for de undersøkte diskusjonene, ble også alle opptak av diskusjoner hvor flere enn en person hadde galt svar ut, gått gjennom (men ikke transkribert), og grovt klassifisert i henhold til det samme systemet. Dette gjaldt til sammen atten diskusjoner. Grunnet denne oppgavens omfang og tidsrammet ble ikke disse studert nærmere.

3.7 Etske betraktninger

Alle deltagere i studien har samtykket til bruk av datamateriale i form av henholdsvis resultater fra FCI, resultater på konseptoppgaver fra undervisning og opptak av gruppediskusjoner. Samtykket har vært fritt og informert, i den forstand at studentene har fått informasjon om hva det er de samtykker til og at samtykket er valgfritt og på ingen måte bindende. Ingen insentiver for å delta i studien ble gitt, og studien er godkjent av NSD¹⁰. Deltagelse og samtykkeerklæring ble fasilitert og etterspurt av fagansvarlig, og derfor ble det blitt gitt grundig informasjon om at studentenes anonymitet ble ivaretatt og at

10. <https://www.nsd.no/>

også eksamen var anonym. Dette for at studentene ikke skulle få en følelse av at det lønnte seg akademisk å delta i studien. Samtykkeskjema ble distribuert og besvart i starten av semesteret. I denne studien har forskjellige data blitt analysert i fellesskap. Studenter som ikke har gitt samtykke til innsyn i minst et aspekt ved datamaterialet som analyseres har dermed blitt fjernet fra den aktuelle delen av studien.

Med unntak av opptakene fra gruppediskusjonen, er datamaterialet som er mottatt andrehånds, og det at jeg har gjennomført studien har i særdeles liten grad påvirket datamaterialet. Undervisningen i sin helhet ble heller ikke påvirket av denne studien. Undervisningen var lik for alle uavhengig av deltagelse, og deltagelse spilte ingen rolle for studentenes vurdering i emnet.

Det synes nødvendig å klargjøre min rolle i undervisningen utover denne studien. Det semesteret undervisningen pågikk var jeg ansatt som seminarlærer i det aktuelle emnet. Dette innebar at jeg rettet innleveringsoppgaver, og ledet ukentlige seminarer med to ulike studentgrupper. Fordi jeg har kjennskap til deler av studentmassen, må det tas høyde for ulike bias det kan ha medført i forbindelse med mine observasjoner av videosamtaleopptak. I tillegg har jeg gjennom min rolle som seminarlærer vært direkte involvert i undervisningen, og dermed potensielt hatt mulighet til å påvirke studentene til å i fellesundervisningen agert og respondert på en måte som var ønskelig fra min side vedrørende denne studien. Jeg har selvfølgelig vært klar over dette, og har utført mine arbeidsoppgaver som seminarlærer uten tanke på studien. Dessuten har jeg på ingen måte vært involvert i fellesundervisningen, som er den delen av undervisningen denne oppgaven tar for seg.

I denne studien er opptak av gruppediskusjoner i Microsoft Teams benyttet. Dette gjør det mulig å identifisere deltagerne, noe som medfører mange etiske hensyn som best mulig må ivaretas. Blant det viktigste er ansvaret for å verne om deltagerens personvern. For å ivareta personvern hensyn, har videomateriell fra denne studien ikke blitt delt eller diskutert med andre enn de som deltar i prosjektet. I praksis har kun jeg, veileder og en masterstudent som tar i bruk deler av det samme datamaterialet sett på og diskutert opptakene. I eksempler hentet fra gruppediskusjoner i denne oppgaven, vil studentene være representert ved ikke-sporbare pseudonymer generert av en *random name generator*¹¹.

I min behandling av videoopptakene er jeg utsatt for en rekke bias og andre faktorer som påvirker hvordan jeg tolker og oppfatter resultatene. Mitt utgangspunkt påvirker det jeg ser. Derfor har jeg vært påpasselig med å gå gjennom opptak flere ganger, også etter transkripsjoner, for å forsøksvis bekrefte at mine

11. <https://www.behindthename.com/random/>

oppfatninger er representative. Utvalgte opptak er også blitt sett av og diskutert med veileder for å påse at mine observasjoner kan støttes opp av andre. I en studie hvor opptak brukes, bør også studiens subjekter få mulighet til å gå gjennom det som produseres fra denne dataen for å kunne se om deres ord og handlinger har blitt oppfattet og tolket på en representativ og objektiv måte. Dette har det ikke vært mulighet til å gjøre i denne studien. Til slutt er det viktig å være klar over de bias jeg også kan ha i den delen av analyse hvor *utvalg* av videoopptak og transkripsjon inngår. I denne studien har videoopptak blitt valgt ut på bakgrunn av hvor mange som har svart korrekt inn og ut av gruppediskusjonene. Det er en fare for at man – når man gjør utvalg i en studie som denne – kan bli ivrig etter å finne interessante ting (gjerne også uavhengig av studiens fokusområde), og at man dermed påvirkes i utvalgsprosessen. Det samme gjelder i analysen av transkripsjonene.

/4

Resultater

Studien inneholder både kvantitative og kvalitative analyser. I seksjon 4.1 tar vi for oss læringsutbyttet fra undervisningen, både overordnet og fra diskusjonsoppgavene spesielt. Dette innebærer resultater fra kvantitative analyser. I seksjon 4.2 tar for seg aspekter som spiller inn på hvorvidt gruppediskusjoner resulterer i mange korrekte svar, og i forlengelsen av dette er produktive, og i denne seksjonen presenteres både resultater fra de kvantitative og de kvalitative analysene.

4.1 Læringsutbytte fra Peer Instruction

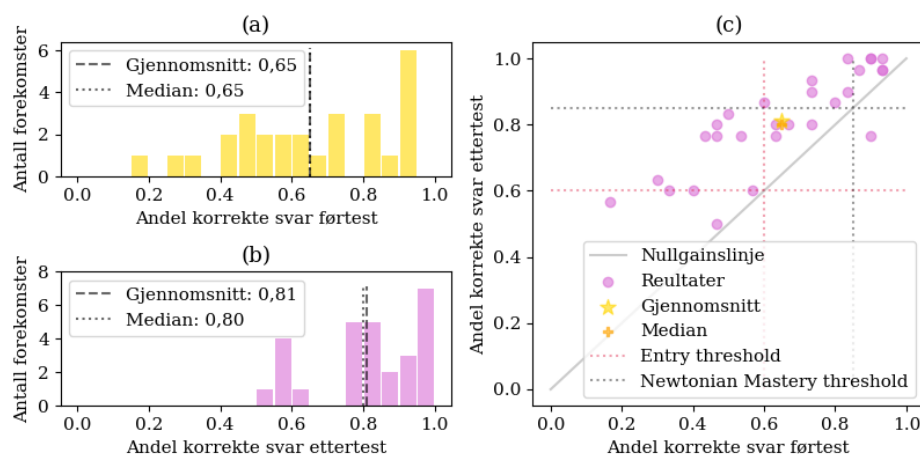
For å evaluere læringsutbyttet har vi sett på utfallet av den overordnede undervisninga i form av konseptuell forståelse som målt ved FCI, samt hvordan utbyttet har vært fra gruppediskusjonene i seg selv ved å kartlegge besvarelsene på diskusjonsoppgavene i fellesundervisningen. I det følgende vil disse resultatene presenteres.

4.1.1 Overordnet læringsutbytte fra undervisning

For å finne ut om undervisningen har fungert har det vært ønskelig å si noe om læringsutbyttet etter undervisning. Til dette har vi brukt FCI. Som det påpekes i seksjon 2.3.1 er dette et mye brukt og anerkjent verktøy for å måle konseptuell

forståelse i fysikk, og dermed godt egnet for å si noe om undervisningens overordnede læringsutbytte. I det følgende presenteres resultatene fra analysen gjort på FCI-data.

I figur 4.1 presenteres et spredningsplott sammen med histogrammer som viser fordelingen av andel korrekte svar på førtest og ettertest.



Figur 4.1: (a) Fordeling av andel korrekte svar på FCI førtest. (b) Fordeling av andel korrekte svar på FCI ettertest. (c) Andel korrekte svar på FCI førtest plottet mot andel korrekte svar på FCI ettertest. Den diagonale linjen representerer null forbedring mellom før- og ettertest. Dottedede linjer er også tegnet inn for Entry threshold og Newtonian Mastery threshold (som presentert i kapittel 2).

Fra analysen framkom det at gjennomsnittlig andel korrekte svar på førtest var 0,65, og at gjennomsnittlig andel korrekte svar på ettertest var 0,81. Forholder vi oss til Hestenes mfl. (1992) (det henvises til seksjon 2.3.1), så er snittscore på FCI ettertest det mest naturlige målet på om konseptuell forståelse har blitt oppnådd etter undervisningen. En score på omtrent 85 prosent vil herfra bli sett på som å være et såkalt «Newtonian mastery threshold», noe som vil si at en individuell score over 0,85 er et tydelig tegn på at respondenten kan sies å være en newtonsk tenker. Hestenes og Halloun (1995) definerte også et «entry threshold» for konseptuell forståelse på rundt 60 prosent, og scoren på omtrent 81 prosent må i lys av dette sies å være et positivt resultat. Fra spredningsplottet i figur 4.1 ser vi at av alle deltagende og samtykkende studenter så var 11 stykker under entry threshold før undervisningsløpet, og at disse er redusert til to ved endt undervisning. Dessuten har antallet newtonske tenkere fordoblet seg. Av alle studentene har bare én endt opp med å ha negativ utvikling, og den konseptuelle forståelsen til majoriteten av studentene har dermed økt, noe som både er forventet og ønskelig. Den overordnede positive utviklingen bekreftes

ved det tydelig skiftet mot høyre i histogrammene i samme figur.

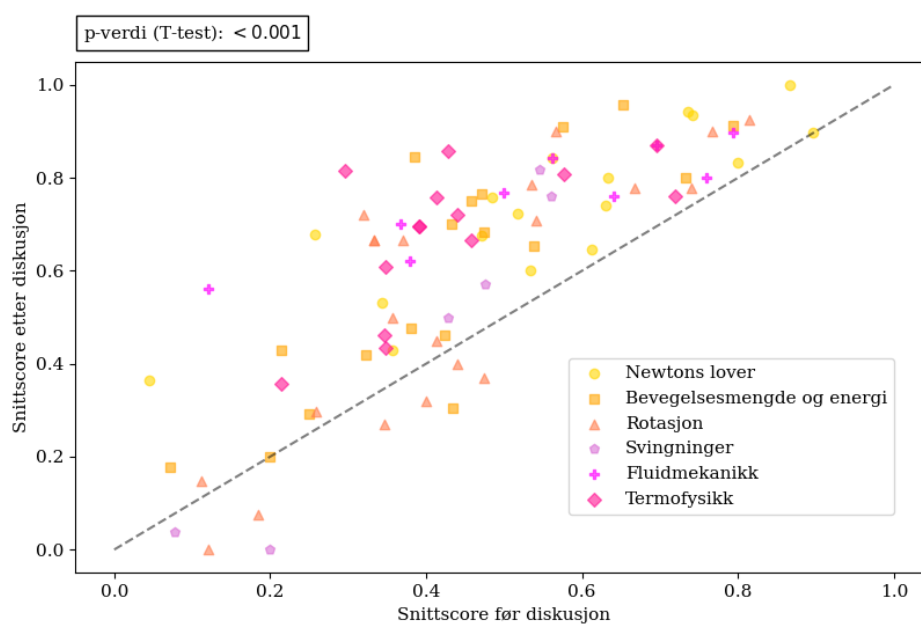
Gjennomsnittscorene fra førtest og ettertest ga et gain på 0,16. Med Hestenes mfl. (1992) i bakhodet skal vi ikke legge alt for stor vekt på dette resultatet, men vi noterer oss at det er en økning. Allikevel kan vi konstatere at et gain på 16 prosent er fint i samsvar med de resultater Mazur (1997) påpeker at man burde forventet ved Peer Instruction. Fordi gain – som motivert i seksjon 2.6.1 – vil favorisere studenter med lav score på førtest, er det mer interessant å se på normalisert gain. Dette ble funnet til å være på 0,46. Mazur (1997) peker på at man ved konvensjonell undervisning vanligvis vil se et normalisert gain på omkring 0,25, mens man ved mer interaktiv undervisning kan forvente verdier mellom 0,36 og 0,68. Vi ser at undervisningen i så måte har gitt forventede resultater. Til slutt ble det funnet en verdi for Cohens d på 0,88. Som påpekt i seksjon 2.6.1 vil en verdi større enn 0,8 være sett på som god, og verdien vi har fått er dermed en tydelig indikator på at differansen i snittscore på FCI førtest og FCI ettertest som oppnådd ved gjennomføring av dette fysikkemnet, er betydelig.

4.1.2 Utbytte fra diskusjon av oppgaver

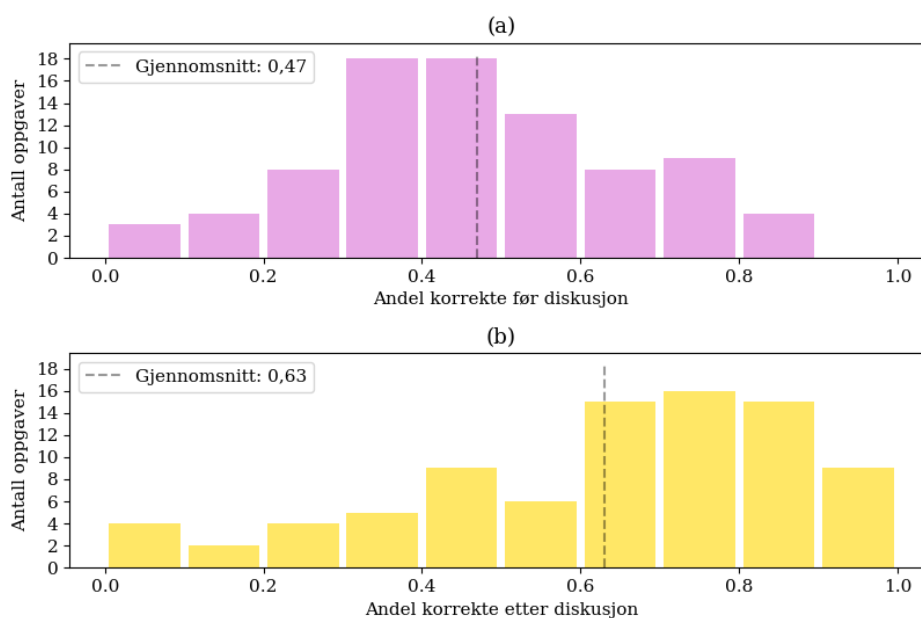
Data med studentenes besvarelser på diskusjonsoppgavene i undervisning har blitt analysert for å finne ut noe om utbyttet fra diskusjonen av disse. Oppgavene i seg selv er blitt kartlagt ut i fra gjennomsnittlig score før og etter diskusjon. På samme måte er studentenes utvikling blitt kartlagt.

Kartlegging av oppgaver

Selve oppgaven kan spille inn på utbyttet av en gruppediskusjon. Hvis oppgaven for eksempel er for vanskelig, eller utydelig formulert, så er det tenkelig at utviklingen fra diskusjonen ikke er ønskelig. For å beskrive og kartlegge diskusjonsoppgavene har spredningsplott (figur 4.2) og histogrammer (figur 4.3) som viser fordelingen av gjennomsnittlig andel korrekte svar for hver oppgave blitt produsert. En parret t-test har blitt utført for å sjekke om det er en signifikant forskjell på gjennomsnittene i før- og etterdata. Dessuten har effektstørrelsen Cohens d blitt regnet ut for å si noe om differansen mellom korrekthet før og etter diskusjon på en gjennomsnittsoppgave.



Figur 4.2: Spredningsplott av konseptoppgaver. Snittscore før diskusjon på førsteaksen. Snittscore etter diskusjon på andreaksen. Oppgaver er fargekodet etter oppgavens tema. p-verdi fra parret T-test er inkludert.



Figur 4.3: Fordelingen av gjennomsnittscoren på hver konseptoppgave henholdsvis før (a) og etter (b) diskusjon.

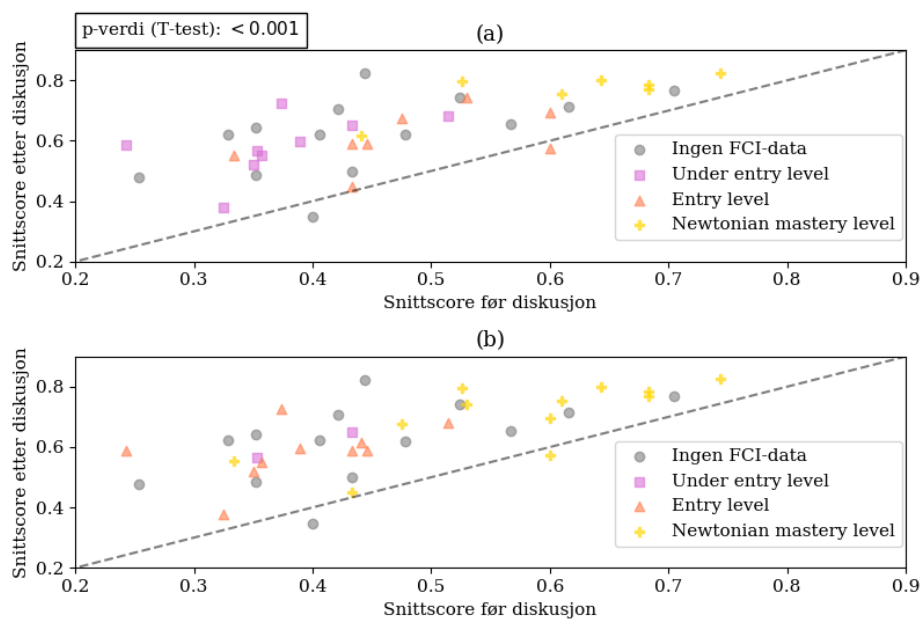
Fra figur 4.2 ser vi at på de aller fleste oppgavene har flere studenter svart korrekt etter diskusjon, enn før diskusjon. Disse oppgavene har derfor gitt flere produktive diskusjoner enn uproduktive. På 11 av oppgavene har imidlertid flere studenter svart galt etter diskusjon enn før diskusjon.

Vi merker oss at hele seks av elleve oppgaver med negativ utvikling er oppgaver som omhandler rotasjon. Videre ser vi at alle oppgaver (med unntak av en) hvor mer enn halvparten svarer korrekt før diskusjon, har positiv utvikling. I følge Mazur (1997) er ønsket vanskelighetsgrad på en konseptoppgave i Peer Instruction slik at omtrent femti prosent svarer riktig før diskusjon, og vi ser fra figur 4.2 at det er mange oppgaver med god utvikling i dette sjiktet, samtidig som det er fare for negativ utvikling dersom færre enn femti prosent svarer korrekt.

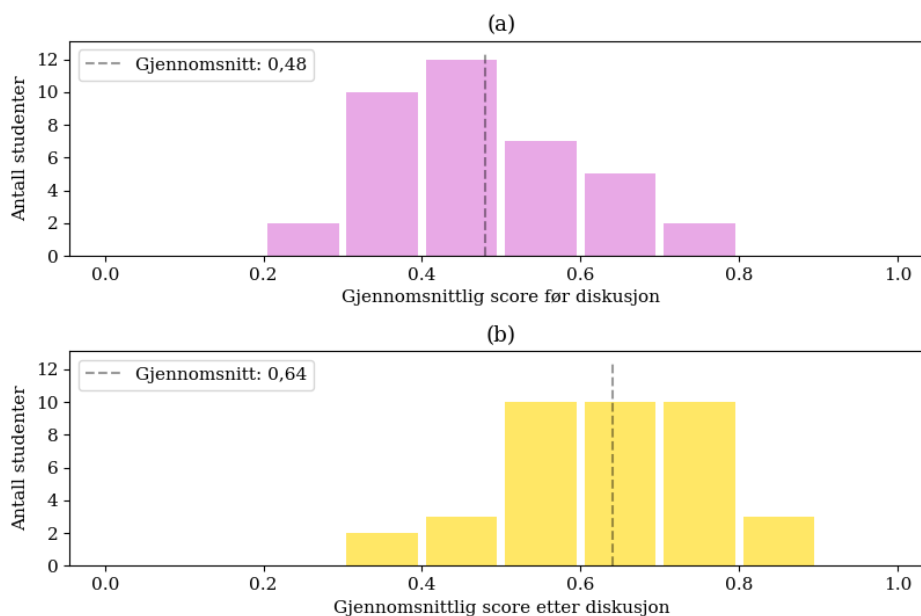
Fra figur 4.3 ser vi igjen at majoriteten av oppgaver gir positiv utvikling. Vi ser også at gjennomsnittsandelen korrekte svar var på 0,47 før diskusjon, og på 0,63 etter diskusjon. Dette gir en differanse på 0,16. Det gir dermed en verdi for Cohens d på 0,74. Som det kom fram i seksjon 2.6.1, vil en verdi på 0,8 være å anses som god, mens en på 0,6 er middels. Vi kan derfor si at for en gjennomsnittsuppgave, har diskusjonen hatt middels til god effekt. Fra t-testen som ble utført fikk vi en svært lav p-verdi, noe som tilsier at det er en signifikant forskjell mellom før- og etterdata.

Kartlegging av studentenes utvikling

For at en student i skal oppleve en diskusjon i Peer Instruction som produktiv, så vil det innebære at studenten svarer korrekt etter diskusjon. For å undersøke hvordan forbedringen til studentene var gjennom diskusjonene ble derfor konseptoppgavedata også brukt til å se på fordelingen av studentenes gjennomsnittlige score før diskusjon og etter diskusjon. Disse er illustrert i figur 4.4 og figur 4.5. Det har også blitt utført en parret t-test for å sjekke om det er en signifikant forskjell på gjennomsnittene i før- og etterdata. Videre er Cohens d regnet ut også her for å undersøke hvor stor effekt diskusjonene har hatt for gjennomsnittsstudenten.



Figur 4.4: Spredningsplott som viser studentenes snittscore på konseptoppgavene før og etter diskusjon. Studentdata er fargekodet på score på FCI førtest (a) og FCI ettertest (b). p-verdi fra parret T-test er inkludert.



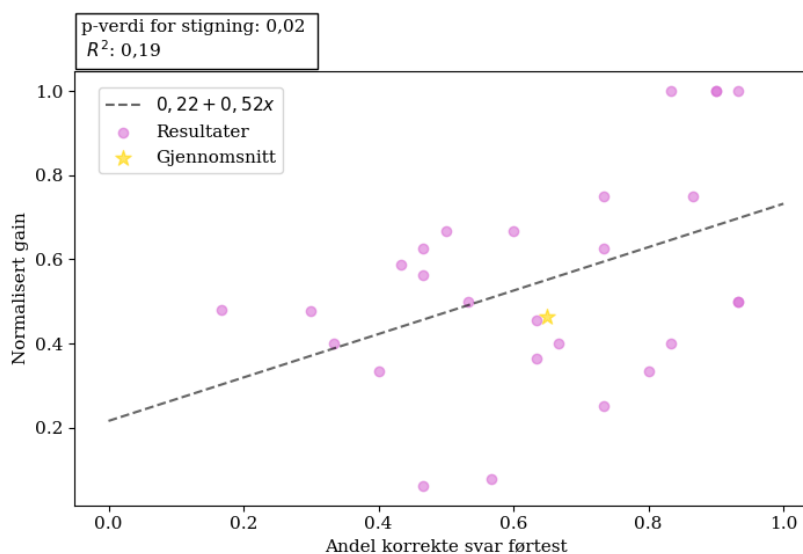
Figur 4.5: Histogram som viser fordelingen av gjennomsnittscoren på konseptoppgavene for hver student henholdsvis før (a) og etter (b) diskusjon.

Fra figur 4.4 ser vi at det kun er to studenter som i snitt har negativ utvikling på diskusjonsoppgavene, og at diskusjonene for gjennomsnittsstudenten jevnt over har vært produktive. Plottene er fargekodet ut i fra henholdsvis FCI førtest og ettertest, og man kan ikke uventet ane en tendens til at de som har en newtonsk forståelse før undervisningsløpet i stor grad svarer korrekt etter diskusjon. I det store og det hele kan vi si at de aller fleste studentene har fått utbytte av diskusjonene i form av å svare korrekt ut av diskusjon, noe som bekreftes av histogrammene i figur 4.5 hvor vi ser en tydelig (og forventet) forskyvning mot høyre. Gjennomsnittstudenten hadde i snitt riktig svar på 48 prosent av spørsmålene før diskusjon og på 64 prosent av spørsmålene etter diskusjon, noe som gir en økning på 16 prosent. Fra dette ble en verdi for Cohens d på 1,37 kalkulert. Denne verdien tyder på at det er en betydelig differanse på studentenes andel korrekte svar før diskusjon og etter diskusjon, og at utbyttet fra diskusjonen for gjennomsnittstudenten i så måte har vært godt. Fra t-testen indikerer en verdi på under 0,001 at det er signifikant forskjell på før- og etterdata.

4.1.3 Sammenheng mellom faglig inngangsnivå og overordnet læringsutbytte

For å si mer om læringsutbyttet, er det interessant å undersøke om den forståelsen studentene går inn i undervisningen med, har innvirkning på utbyttet av undervisningen. For å gjøre dette har individuelle normaliserte gains på FCI blitt plottet mot score på FCI førtest (figur 4.6). For å sjekke for korrelasjon har det også blitt gjennomført en lineær regresjon.

Fra p -verdien på 0,02 for stigningstallet i regresjonen vist i figur 4.6 kan vi bekrefte en sammenheng mellom høy score på FCI førtest og læring som målt ved individuelle normaliserte FCI-gains.

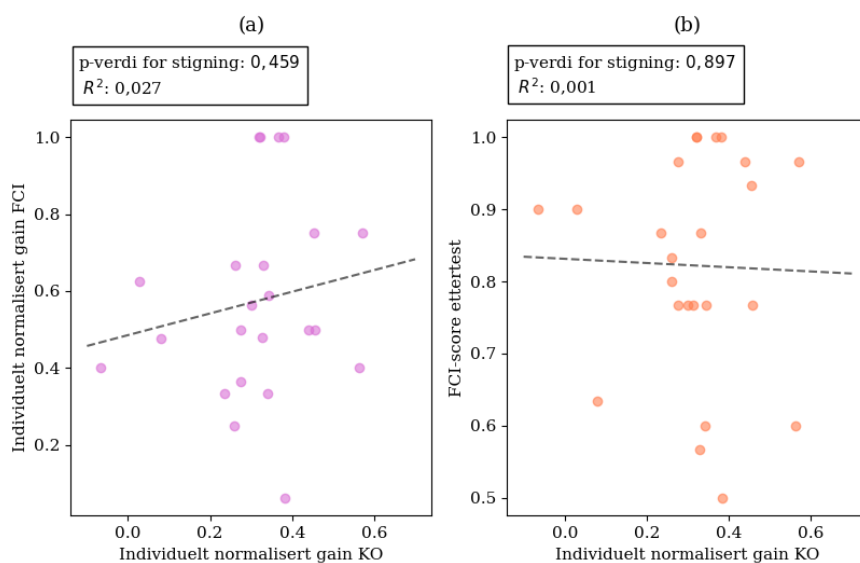


Figur 4.6: Individuelle normaliserte gains plottet mot andel korrekte svar på FCI førtest. Utligger med negativt gain er fjernet. Regresjonslinje fra lineær regresjon gjort på data er tegnet inn, og p-verdi for stigningen til linja samt R^2 -verdi er oppgitt.

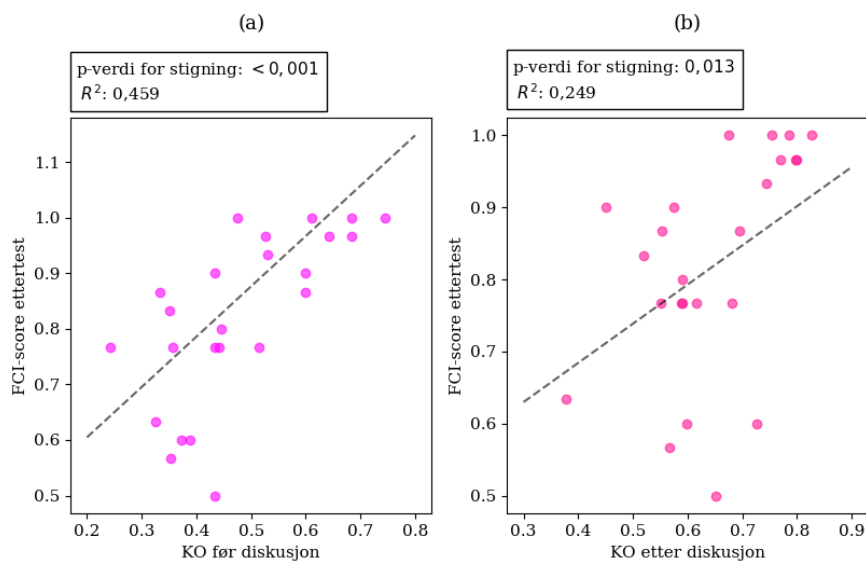
4.1.4 Sammenheng mellom prestasjon på diskusjonsoppgaver og overordnet læringsutbytte

For å nærmere undersøke om studentenes gjennomsnittlige utvikling på diskusjonsoppgavene har noe å si for læringsutbyttet totalt, har utviklingen på diskusjonsoppgavene, som målt ved normalisert gain blitt plottet mot henholdsvis normalisert gain på FCI og score på FCI ettertest (figur 4.7). For å undersøke om det er korrelasjon har det også blitt utført lineær regresjon på begge modellene.

Fra p-verdiene rapportert fra regresjonene i figur 4.7 er det svært vanskelig å påstå at god utvikling på diskusjonsoppgavene fører till bedre læringsutbytte i form av konseptuell forståelse. Fordi konseptoppgavene ikke er ment å måle noe spesifikt, var det også av interesse å undersøke om det var sammenheng mellom henholdsvis gjennomsnittscore på før diskusjon og etter diskusjon, og FCI-score på ettertest. Dette finner vi i figur 4.8. Her ble det også gjort lineær regresjon for å sjekke for korrelasjon



Figur 4.7: Spredningsplott av normalisert gain på konseptoppgavene og henholdsvis normalisert gain fra FCI (a) og FCI score på ettertest (b). Utligger med negativt normalisert FCI-gain er fjernet fra (a). Rapportert p-verdi fra lineær regresjon er inkludert i begge plottene.



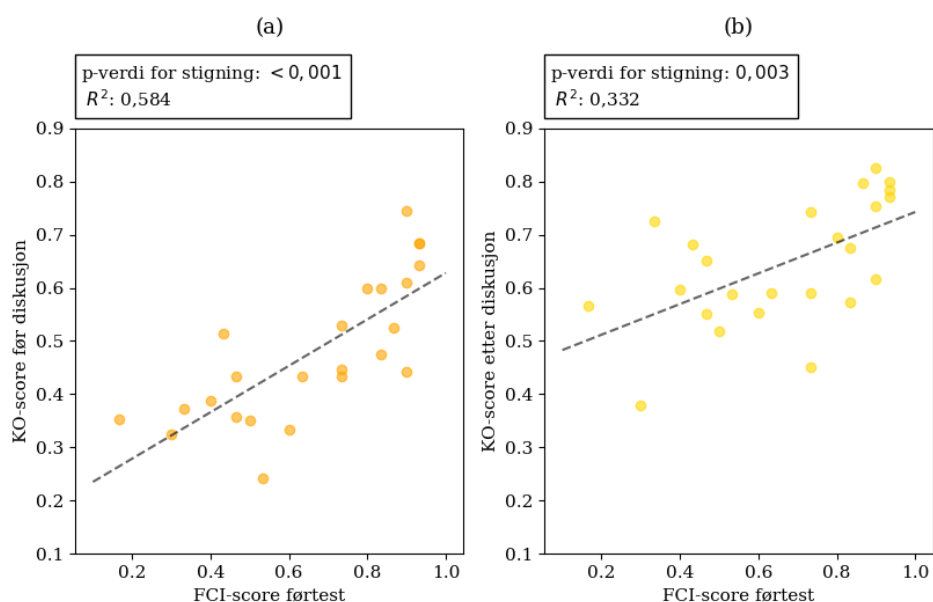
Figur 4.8: Spredningsplott av gjennomsnittsscore på konseptoppgavene henholdsvis før diskusjon (a) og etter diskusjon (b) og score på FCI ettertest (b). Rapportert p-verdi fra lineær regresjon er inkludert i begge plottene.

Fra p-verdiene rapportert i figur 4.8 ser vi at det kan sies å være en signifikant sammenheng mellom score på konseptoppgavene og score på FCI ettertest.

4.2 Produktive diskusjoner

4.2.1 Sammenheng mellom faglig inngangsnivå og prestasjon på diskusjonsoppgaver

En diskusjon i Peer Instruction kan være suksessfull (rent kvantitativt) uten at det foregår noen reell diskusjon, dersom alle gruppemedlemmene vet svaret på det spørsmålet som skal diskuteres. Derfor er det nyttig å undersøke hvorvidt det faglige inngangsnivået til studentene spiller inn på hvorvidt de presterer godt på diskusjonsoppgavene. For å si noe om dette, har sammenhengen mellom FCI-score på førtest og gjennomsnittlig andel korrekte svar på diskusjonsoppgavene henholdsvis før og etter diskusjon blitt undersøkt. I figur 4.9 finner vi spredningsplott av score på FCI førtest mot gjennomsnittlig andel korrekte svar før og etter diskusjon, for hver student. For å se om det er korrelasjon, er det også blitt gjort lineær regresjon på de to modellene.

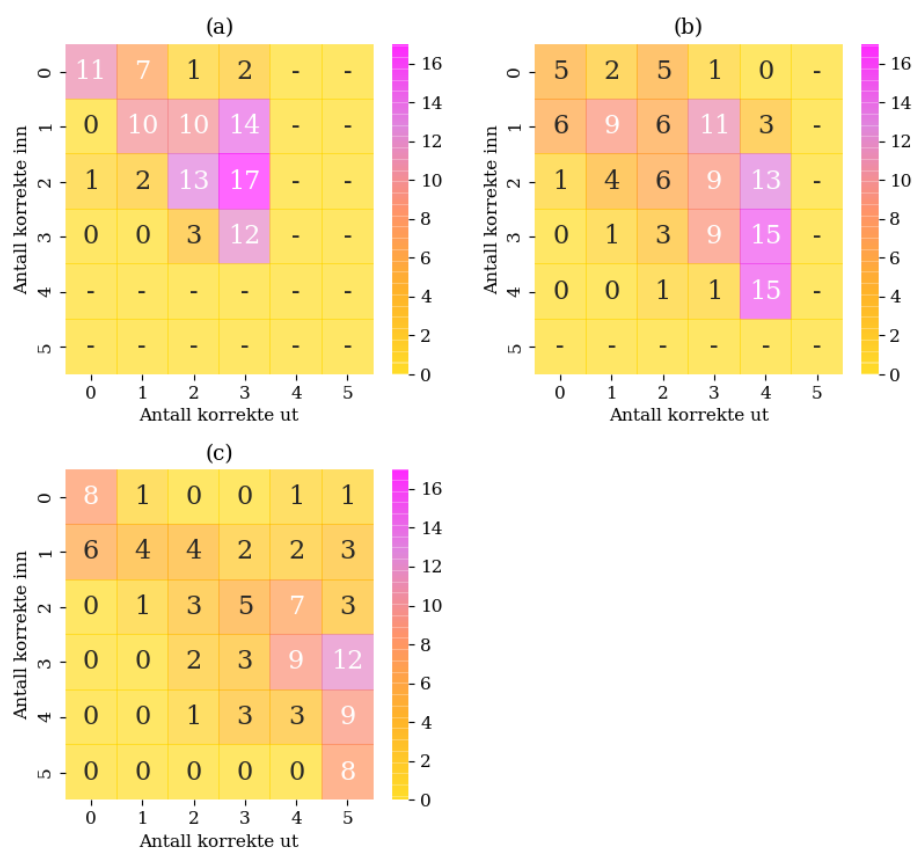


Figur 4.9: Spredningsplott som viser sammenhengen mellom FCI-score på førtest og gjennomsnittlig score på konseptoppgavene henholdsvis før (a) og etter (b) diskusjon, for hver student. Resultater fra lineær regresjon er inkludert.

Fra figur 4.9 ser vi fra p -verdiene at FCI-score på før- og ettertest kan spille en rolle for graden av utvikling på diskusjonsoppgavene. Spredningen i y -retning blir mindre etter diskusjon, noe som både er forventet og ønskelig. Vi ser at det er en ganske tydelig sammenheng mellom score på FCI førtest og score på konseptoppgaver før diskusjon. Dette gir validitet til FCI førtest som et mål på studentenes nivå når de går inn i undervisningen.

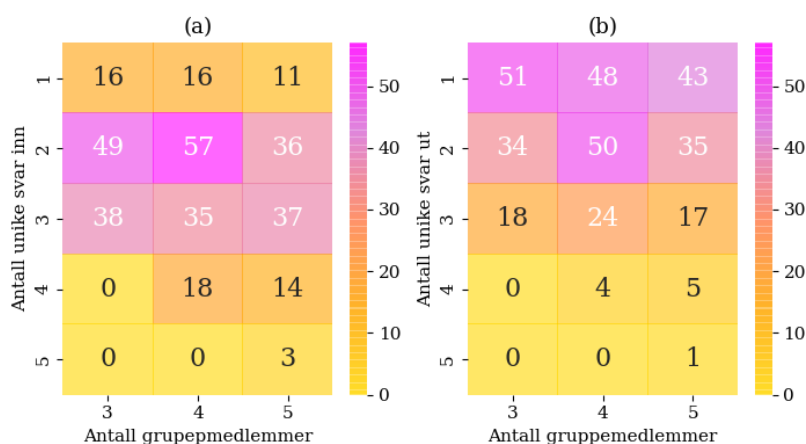
4.2.2 Korrekthet og uenighet

For å undersøke hva som kjennetegner selve diskusjonene i Peer Instruction nærmere, har alle instanser av diskusjoner blitt undersøkt nærmere. En analyseenhet er et tilfelle av en gruppediskusjon (én gruppe diskuterer én oppgave). I det følgende vil *korrekthet* bli pekt på som hvor mange, eller hvor stor andel av en gruppe som har svart korrekt på konseptoppgaven. I figur 4.10 presenteres en oversikt av alle analyseenhetene i form av en matriser hvor enhetene er kategorisert ut i fra antallet korrekte svar inn i diskusjon og antallet korrekte svar ut av diskusjon. Her har effektive ener- og toergrupper blitt fjernet fra datamaterialet. Det er produsert matriser for hver av de gjenværende effektive gruppestørrelser. Videre er uenighet inn og ut av diskusjon undersøkt. Dette ved å se på antallet unike svar opp mot gruppestørrelse. Fordelingen av analyseenhetene i henhold til uenighet inn og ut av diskusjon presenteres i figur 4.11. For å videre se på hvordan uenighet og korrekthet påvirker utfallet av diskusjonene ble det laget histogrammer som viser hvordan korrekthet ut fordelte seg for ulike nivåer av korrekthet inn og uenighet inn. Disse presenteres i figur 4.12, figur 4.13 og figur 4.14. Fordi andel korrekte er et vanskelig mål på korrekthet (De ulike effektive gruppestørrelsene tatt i betraktning), har analyseenhetene blitt kategorisert ut i fra hvor mange som har korrekt inn i diskusjon, og hvor mange som har korrekt ut av diskusjon. Fordelingen av analyseenheter i henhold til korrekthet er illustrert i figur 4.10. Her presenteres fordelingen for de forskjellige effektive gruppestørrelsene.



Figur 4.10: Matriser av alle analyseenheter for forskjellige effektive gruppestørrelser. (a) viser treergrupper, (b) viser firergrupper og (c) viser femmergrupper. Matrisene er visualisert som heatmaps. Analyseenheterne er kategorisert ut i fra antall korrekte gruppe-medlemmer før diskusjon og antall gruppe-medlemmer med korrekt svar etter diskusjon.

Fra figur 4.10 ser vi at hvis ingen har korrekt inn, så er det sjelden at alle blir enige om korrekt svar ut. Motsatt er det – som forventet – ofte at alle svarer korrekt etter diskusjon, dersom alle har korrekt før diskusjon. Kun fem av 40 slike diskusjoner gjør at noen endrer til galt svar. Fra alle tre matrisene ser vi at jo flere som har korrekt inn, jo bedre blir utfallet i form av antall korrekte ut. En interessant tendens er at det på treergruppene ser ut til å holde at en student har korrekt før diskusjon for at diskusjonen ikke skal gi negativ utvikling, mens det ser ut som de større gruppene ofte krever at to stykker har korrekt inn i diskusjonen for at gruppa skal få mange korrekte ut av diskusjonen. Generelt ser vi at jo større grupper, jo flere bør ha korrekt inn i diskusjonen for at utviklingen skal være god, og at det bør være et gruppe-medlem med korrekt svar inn i diskusjonene for at flertallet av gruppa skal bli enig om korrekt



Figur 4.11: Matriser som viser fordelingen av uenighet inn og ut av diskusjon for ulike gruppestørrelser. (a) viser fordelingen før diskusjon og (b) viser fordelingen etter diskusjon.

svar.

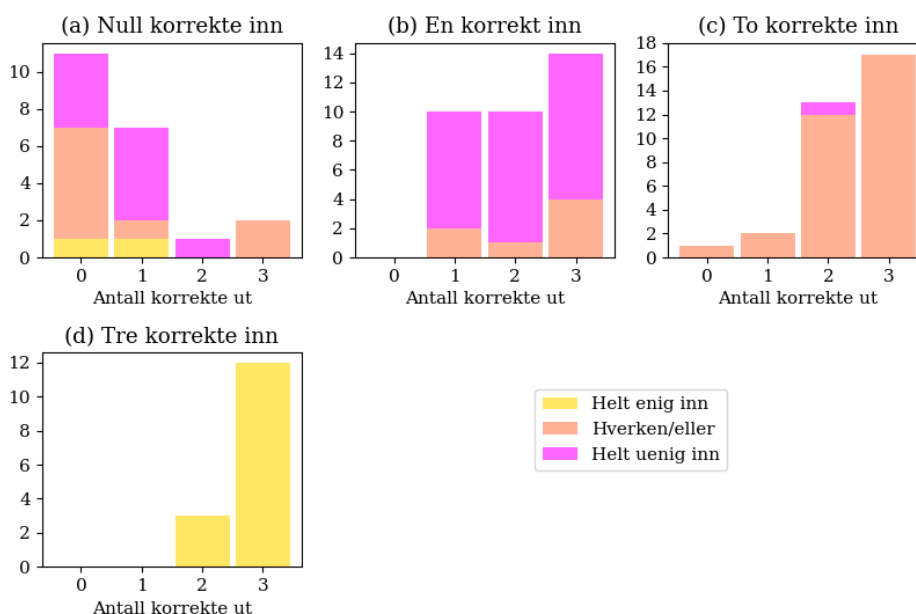
Videre var det interessant å se på hvordan uenighet spilte inn på utfallet av diskusjonene. I figur 4.11 er fordelingen av uenighet inn og ut av diskusjon forsøkt visualisert ved to matriser hvor analyseenhetene fordeles etter antall gruppelemmer og antall *unike* svar.

Som det kommer fram av figur 4.11, er studentene mer uenige før diskusjon enn etter diskusjon (naturlig nok). Majoriteten av diskusjonene ender med at alle på gruppa er enige. Det kommer fram at i majoriteten av diskusjonene er det to eller tre unike svar inn. Hvis ett av disse skulle være korrekt, er det grunn til å tro at det vil føre til positiv utvikling. For å videre undersøke sammenhengen mellom uenighet og korrekthet har det blitt laget histogrammer for de ulike gruppestørrelsene, som viser hvordan gruppelemmenes svar etter diskusjon fordeler seg ut i fra hvor mange som hadde korrekt inn. Histogrammene er fargekodet ut i fra graden av uenighet inn i diskusjonene. *Helt uenig* betyr at alle gruppelemmene har svart forskjellig, *helt enig* betyr at alle har svart det samme, mens *hverken/eller* indikerer alle andre tilfeller. Tre av analyseenhetene var diskusjoner av oppgaver hvor det var ulike svar som kunne gi korrekt score. Disse er justert for i histogrammene.

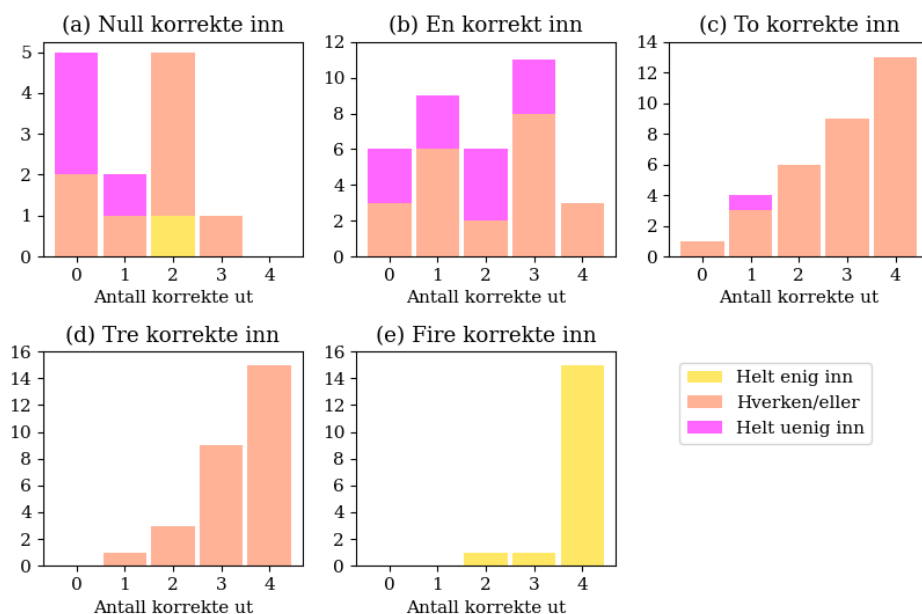
Vi ser fra figur 4.12 at for effektive treergrupper, så bør det minst være ett gruppelem som har korrekt inn i diskusjonen for at det skal være positiv utvikling i gruppa. Uenighet inn i diskusjonen kan bidra til positiv utvikling, så lenge det er minst en på gruppa som har korrekt før diskusjonen. Vi ser at høy grad av enighet jevnt over er relativt fordelaktig for utviklingen i diskusjonen

så fremt ikke alle er enige om galt svar.

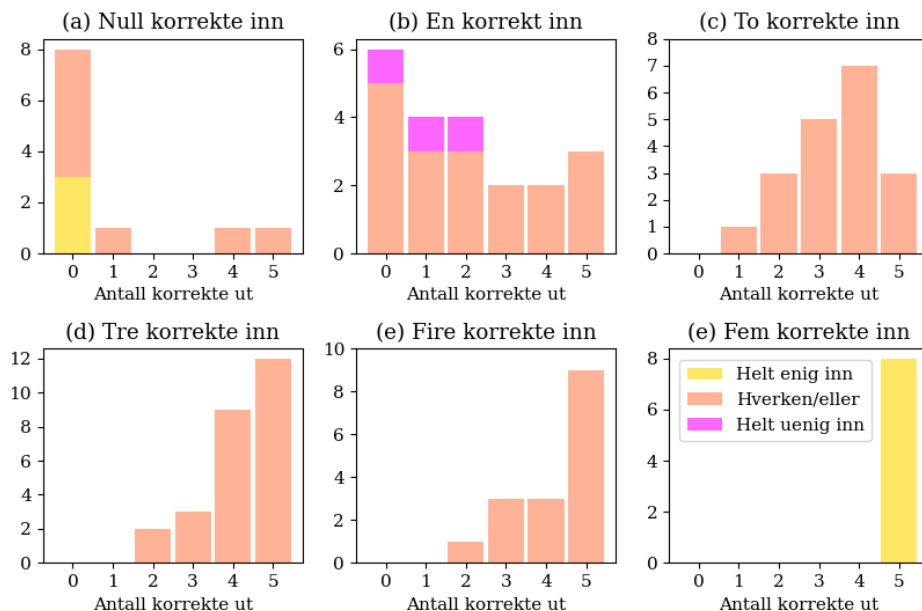
For de effektive firergruppene i figur 4.13 ser vi mye av den samme tendensen, men vi ser (som fra figur 4.10) at det gjerne må to gruppede medlemmer med riktig svar inn, for å i stor grad få en positiv utvikling. Her er det også enda tydeligere enn hos treergruppene, at total uenighet ikke ser ut til å gagne utviklingen



Figur 4.12: Fordeling av antall korrekte etter diskusjon, for ulikt antall korrekte før diskusjon. data for alle effektive treergrupper. Søyler er fargekodet ut i fra uenighet før diskusjon.



Figur 4.13: Fordeling av antall korrekte etter diskusjon, for ulikt antall korrekte før diskusjon. data for alle effektive firergrupper. Søyler er fargekodet ut i fra uenighet før diskusjon.



Figur 4.14: Fordeling av antall korrekte etter diskusjon, for ulikt antall korrekte før diskusjon. data for alle effektive femmergrupper. Søyler er fargekodet ut i fra uenighet før diskusjon.

Til slutt ser vi hos femmergruppene i figur 4.14 en lignende tendens som hos firergruppene. Det må mer enn ett grupped medlem med korrekt svar til for å virkelig få en ønsket utvikling. Igjen ser vi at total uenighet ikke ser ut til å være fordelaktig.

4.2.3 Videre analyse av produktive diskusjoner

For å videre undersøke hva som kjennetegner en produktiv diskusjon ble 14 opptak av diskusjoner gått gjennom og transkribert. Alle disse var kjennetegnet av at det var maks ett grupped medlem som hadde korrekt svar før diskusjon, og maks ett som hadde galt svar etter diskusjon. Analysen av disse fant at diskusjonene generelt skilte seg fra hverandre ut i fra hvor mange som tok ordet, forklarte og argumenterte for sitt svar. Enten var det en person som overbeviste resten av gruppa, eller så var det dialog mellom to eller flere grupped medlemmer som medførte enighet om korrekt svar. Alle de nevnte diskusjonene har derfor blitt klassifisert i henhold til dette.

Fra analysen kom det også fram at det i to¹ av gruppene ofte var samme person som overbeviste de andre, i de tilfellene hvor det skjedde. Nærmere undersøkelser viste at dette for begge gruppene var det grupped medlemmet med høyest score på FCI førtest. I disse tilfellene var det altså det grupped medlemmet som i utgangspunktet hadde størst grad av konseptuell forståelse, som tok ordet og klarte å overbevise de andre på gruppa.

I tabell 4.1 finner vi klassifisering av alle transkriberte analyseenheter i henhold til klassifiseringskriteriene presentert over. Her vil en analyseenhet klassifisert som *D* (dialog) representere en samtale hvor flere enn en person bidro til at gruppeflertallet kom fram til riktig svar. Tilsvarende representerer en klassifisering *M* (monolog) en diskusjon hvor én person overbeviser brorparten av gruppa til å velge riktig svar. Klassifisering 0 viser til diskusjoner hvor det fra opptak og transkripsjoner ikke var mulig å se hva som gjorde at mange svarte riktig i runde to. I tabellen er også andel korrekte inn og ut av diskusjonen tatt med, i tillegg til andel aktive. Med andel aktive menes hvor mange av gruppetotalen som bidro i det hele tatt i samtalen. Derfor kan andel aktive være større enn en, selv om diskusjonen er blitt klassifisert som *M*.

I det følgende vil to av diskusjonene som ble klassifisert bli presentert. Disse er markert med farge og fete typer i tabell 4.1. Begge foregikk i samme gruppe, og der den første ble klassifisert som *M* – en diskusjon der en person overbeviser resten, ble den andre klassifisert som *D* – en diskusjon hvor flere bidro til

1. Til sammen ble fem forskjellige grupper transkribert, men de nevnte to gruppene sto for majoriteten av de 14 analyseenheterne.

å komme fram til riktig svar. I det følgende blir informasjon om gruppa og gruppemedlemmenes utgangspunkt presentert. Videre blir samtalene presentert med eksempler fra transkripsjonene. Fullstendige transkripsjoner for disse to samtalene finnes i seksjon 6.3.

Tabell 4.1: Resultater fra klassifisering av utvalgte analyseenheter

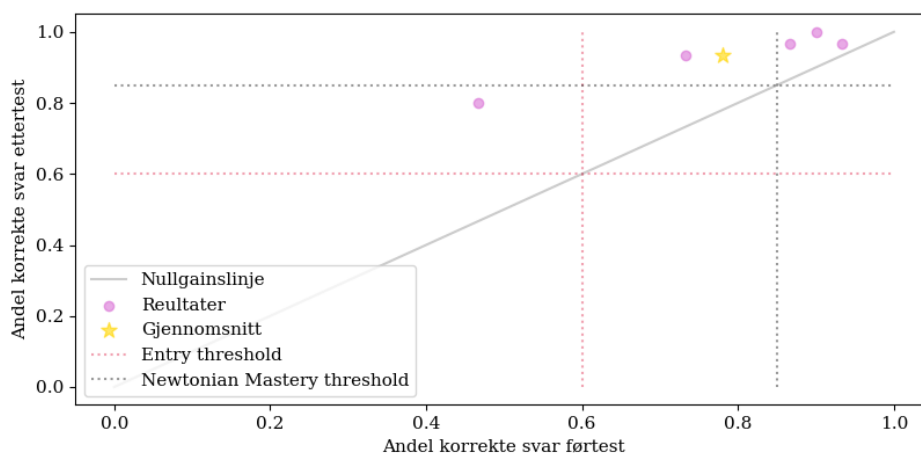
Enhet	Korrekte inn	Korrekte ut	Aktive	Klassifisering
1	1 av 4	3 av 4	2 av 4	<i>D</i>
2	1 av 4	3 av 4	2 av 4	<i>M</i>
3	1 av 4	4 av 4	2 av 4	<i>M</i>
4	1 av 4	3 av 4	3 av 4	<i>M</i>
5	0 av 4	3 av 4	3 av 4	<i>D</i>
6	0 av 4	4 av 4	3 av 4	<i>D</i>
7	1 av 3	3 av 3	0 av 3	<i>0</i>
8	1 av 4	3 av 4	3 av 4	<i>M</i>
9	1 av 4	3 av 4	3 av 4	<i>M</i>
10	1 av 4	2 av 3	2 av 4	<i>M</i>
11	1 av 5	5 av 5	4 av 5	<i>D</i>
12	1 av 5	5 av 5	4 av 5	<i>D</i>
13	0 av 3	3 av 3	3 av 3	<i>D</i>
14	1 av 4	3 av 4	3 av 4	<i>D</i>

Gruppa

Den gruppa som vi tar for oss, bestod av fem samtykkende studenter, alle hannkjønn. Dette var en gruppe som i snitt hadde et resultat på FCI førtest på 0,78, noe som var noe høyere enn snittet til den totale studentmassen. Gruppa hadde god effekt av den overordnede undervisningen, med et normalisert FCI-gain på 0,70 og en Cohens d på 1,18. I figur 4.15 er fordelingen av gruppemedlemmenes FCI-score før og etter diskusjon.

Herfra ser vi at en av studentene ved inngangen til undervisningen var under Hestenes mfl. (1992) sitt entry threshold, mens hele tre stykker allerede før undervisning hadde konseptuell forståelse som tilsa at de kan kalles newtonske tenkere. I henhold til det Johnson mfl. (2014) kaller heterogene grupper, er nok denne gruppa ikke helt optimal. Det at de fleste har et godt faglig utgangspunkt er nok allikevel bedre enn om majoriteten hadde et svakt utgangspunkt. Mazur (1997) skriver at studenter som har god konseptuell forståelse vil kunne argumentere på et høyt nivå ut i fra denne forståelsen, og denne toveis sammenhengen mellom argumentasjonskompetanse og konseptuell forståelse vil nok være til fordel for en gruppe som denne. Det er også verdt å kommentere

at de diskusjonene vi her tar for oss fant sted et godt stykke ut i semesteret, og at gruppemedlemmenes konseptuelle forståelse som målt ved FCI førtest ikke gjenspeiler faktisk status idet diskusjonene fant sted.



Figur 4.15: Fordeling av FCI-score for gruppa som presenteres. (a) Fordeling av FCI score på førtest, (b) fordeling av FCI-score på ettertest, (c) Spredningsplott av FCI-scores.

Analyseenhet 3

I figur 4.16 presenteres oppgaven som ble diskutert i analyseenhet 3.

1.

Person sitter i en robåt og har med seg en stor stein. Midt utpå innsjøen kaster personen steinen overbord. Vil da vannstanden i innsjøen

- A. øke
- B. synke
- C. forbli uendret

Figur 4.16: Oppgaven som diskuteres i analyseenhet 3.

Her skal vi ikke ta for oss selve oppgaven som diskuteres, selv om det ville vært svært interessant å undersøke hvilke oppgaver som ga gode diskusjoner. Vi merker oss allikevel at denne oppgavens gale svaralternativer svarer til oppfatninger mange av studentene nok har, men som ikke stemmer overens med det som er riktig. Korrekt svar er her alternativ B.

Videre tar vi for oss selve samtalen som fulgte av denne oppgaven. Det henvises igjen til seksjon 6.3 for den fulle transkripsjonen. Samtalen ble klassifisert til å være av typen hvor en person gjennom sine argumenter overbeviser de andre

på gruppa. Deltagerne i denne samtalen er *Gawain*, *Mukesh*, *Valdemar* og *Indra*. Inn til samtalen har kun Gawain svart korrekt på oppgaven. Alle de andre på gruppa har svart alternativ C. Diskusjonen forløper seg slik at Gawain som best han kan, prøver å forklare hvordan han tenker for å argumentere for at vannstanden må minke. I linje 3 starter hans første lange resonnement:

Gawain: Altså hvis, hvis steinen er i båten, så vil båten synke ned, og det vil skyve vann ut til siden og oppover.

Valdemar: Mhm

Gawain: Og da vil jo vannstanden være økt tilsvarende så mye som hvis det båten er under vann var fylt med vann, men det er det jo ikke. Det er fylt med en bitteliten – eller ikke bitteliten, en stor – men det er fylt med en stein, så da når du, så da, volumet av båten som er under vann vil være tilsvarende vekta, ehmm, eller være tilsvarende av volum som tilsvarer samme vekten av vann som det er i båten.

Valdemar: Mhm

I linje 7 avslutter Gawain sin forklaring. Fram til dette har han ikke blitt utfordret eller fått bidrag fra de andre, kun kortfattet støtte fra et annet gruppemedlem. Samtalen forløper videre slik:

Gawain: Så hvis du da kaster ut steinen så vil steinen ha veldig mye mindre volum enn det volumet av båten hadde før, for båten vil da flyte høyere opp. Beste måten å se for seg er å se for seg at steinen er veldig tung og båten er veldig lett. Så vi kan si at båten flyter bare oppå vannet, den øker ikke vannivået i det hele tatt når den ikke har en stein i, fordi den er så lett. Også hvis vi legger steinen nedi, så vil da volumet øke fordi båten blir dyttet langt ned. Også hvis vi hiver ut steinen, så vil steinen nesten ikke øke volumet i det hele tatt, mens båten vil bli dyttet opp av vannet som gjør at vannet senker seg.

Indra: Ja, det var en god forklaring.

Valdemar: Så svaret mener du er...

Gawain: Vannet minker.

Valdemar: Så pass ja.

Indra: Jeg tok at den ikke, at den vannstanden var uendret, men jeg er

med på det nå [uhørbart]. Hvis man ser for seg at det er en — som du sier — en liten stein som har uendelig stor masse, så vil jo den fortrenge masse, masse, masse vann, men når du hiver den uti så vil det ikke påføre vannstanden noe økning.

Først når Valdemar spør eksplisitt etter svaret er det tegn til at de andre gruppe-medlemmene aktivt tar stilling til det Gawain forklarer, men det blir ikke gått videre med. Indra kommer til slutt inn og sier hva han svarte, men fordi Gawain har overbevist ham avslutter han med å resitere essensen i Gawains forklaring i stedet for å bygge på den eller utfordre den. Dette er dermed ikke en produktiv diskusjon, som presentert av Mercer og Littleton (2007) via Angell mfl. (2019) i seksjon 2.5.3, og det går an å argumenter for at det er Gawain som får klart mest ut av denne diskusjonen, all den tid han er den eneste som i større grad får anledning til å artikulere sin forståelse. Det må imidlertid ikke utelukkes at Gawains forklaringer behandles hos de andre gruppe-medlemmene uten at det uttrykkes høylydt, og dermed har positiv utvikling gjennom «det indre overbevisende ordet» (jamfør Bakhtin og Holquist (1981) i seksjon 2.5.1).

Analyseenhet 11

I figur 4.17 presenteres oppgaven som diskuteres i denne diskusjonen. Deltagerne i diskusjonen er *Gawain*, *Mukesh*, *Valdemar*, *Nazarius* og *Indra*. Korrekt svar er C, og Indra var den eneste som hadde korrekt inn i diskusjonen. De andre svarte alternativ A før diskusjonen, bortsett fra Gawain som svarte D.

4.

Vi har to kilo kokende kokende vann i en isolert beholder. Vi slipper så en tokilos kobberblokk oppi beholderen. Kobberet tar vi rett fra fryseren så det holder 0 °C. Hvilket utsagn er sann om sluttemperaturen til vannet?

- A. Temperaturen vil ende på 50 °C, altså midt mellom de to temperaturrene, siden begge har like stor masse.
- B. Temperaturen vil ende nært 0 °C siden kobberet leder varmen vekk så godt.
- C. Temperaturen vil ende nær 100 °C siden kobberet har langt lavere spesifikk varmekapasitet enn vannet.
- D. Ingen av utsagnene over er riktige.

Figur 4.17: Oppgaven som diskuteres i analyseenhet 11.

Diskusjonen som forløp fra denne oppgaven ble klassifisert til å være av typen hvor dialog mellom to eller flere gruppe-medlemmer gjør at gruppa blir enige om korrekt svar. Igjen henvises leseren til seksjon 6.3 for den fulle transkripsjonen. Diskusjonen starter med at tre av gruppe-medlemmene sier hva de har svart. Deretter Starter Gawain i linje 6 på et resonnement som skal forklare hvorfor han har svart det han har gjort:

Gawain: Jeg tok ingen av dem, for jeg har gjort noen eksperimenter på dette

her på videregående og hvis jeg husker riktig i alle fall, så er kobber, eh, sin spesifikke varmekapasitet er på ca 600 J/g, mens vann er 4200, så kobber er da ca. en syvendedel, hvis jeg husker riktig, av vann.

Indra: Ja

Valdemar: Hvor vil det ende opp da?

Gawain: Så da...

Mukesh: Så da [uhørbart] temperatur?

Indra: Ja, det var det jeg på en måte tenkte at... [Folk snakker i munnen på hverandre]

Indra: Hva du sa, at?

Nazarius: Hvis de ligger lenge nok, så vil temperaturen bli lik.

Her er både Valdemar og Mukesh tidlig ute med å følge opp ved å stille spørsmål, og Nazarius kommer også med et innspill. Indra – som var den eneste med korrekt svar inn – prøver å komme med sitt syn, men blir avbrutt. I linje 18 følger Gawain opp sitt innledende resonnement:

Gawain: Det jeg tenker er at den er i nærheten av hundre, fordi kobber har ganske lav spesifikk varmekapasitet, men liksom hvor, hvor nærme de definerer som nærme her, det vet vi jo ikke. Så, det er liksom ikke, det er liksom ikke sånn innenfor to tre grader. Det kommer til å være...

Nazarius: Mellom åtti og nitti

Gawain: Sannsynligvis mer enn ti grader forskjell, så... [Uhørbart]

Gawain: ...nitti grader

Indra: Men det er ikke femti da, hvis det er ulik varmekapasitet?

Både Nazarius, og etterhvert Indra, bygger på og utvider forklaringen til Gawain. Dette leder til at gruppa tilsynelatende blir enige om det korrekte svaret:

Gawain: Hvis, ja, hvis jeg husker riktig på kobber så...

Valdemar: Det virker det som du gjør, generelt.

Gawain: Jeg kan jo google det.

Valdemar: Ehm, så da, men da vil jeg tro at nær hundre virker som et logisk svar.

Indra: Mhm.

Gawain: Ja.

Valdemar: Ikke femti i hvert fall.

Nazarius: Nope.

Videre kommer både Gawain, Nazarius og Indra med nye forklaringer som gjør at alle på gruppa til slutt er enige om hva som er riktig svar:

Gawain: Her står det 380, altså nærmere 400. Så, ja da er den nok en del lavere enn det, altså da er det vel mye nærmere hundre, vil jeg tro.

Indra: Ja.

Gawain: Jeg husket feil, men...

Nazarius: Altså, går det an å tenke seg at, kobber vil jo ta opp et x antall ehm mengde med energi, men den energien, men, vil da øke med varme fra da null til et visst antall grader. Men den energien som blir gitt fra vannet til kobberet vil være betydelig mindre enn den energien som, ehh, vil være betydelig mindre enn det det trengs for å senke temperaturen til vann en grad enn det du trenger for å øke temperaturen til kobberet en grad. For kobberet vil jo øke temperaturen mye fortere enn vann synker i temperatur.

Indra: Mhm. [Folk snakker over hverandre]

Indra: [Uhørbart] delta T , så blir jo sluttverdien på T , det blir liksom likevektspunktet, mens den ene starter fra null og den andre starter fra hundre, men c , m -delen vil, eller c vil være ulik på begge sider av ligninga.

Nazarius: Mhm.

Valdemar: Ja.

Gawain: Ja, når jeg googlet det nå så ser jeg at kobber har mindre enn en tiendedel av varmekapasiteten til vann.

Nazarius: Så det vil være nærmere hundre grader.

Hvis vi igjen ser til Mercer og Littleton (2007), så er denne diskusjonen i større grad enn den forrige en produktiv diskusjon, og det kan sågar hevdes at den er en eksplorerende diskusjon. Gruppemedlemmer både bygger videre på, og stiller spørsmål til andres forklaringer. Flere enn en person får her anledning til å forklare og artikulere sitt syn for de andre, der det i den første diskusjonen var bare en person som gjorde det.

/5

Diskusjon

I denne oppgaven har målet vært å si noe om hvordan digital Peer Instruction med faste grupper har fungert som undervisningsform i et introduksjonsemne i fysikk, og da især rette et fokus mot gruppediskusjonene av konseptuelle oppgaver for å finne ut hvilke faktorer som gjør at en gruppe blir enig om korrekt svar på en slik oppgave. For å gjøre dette ble følgende to forskningsspørsmål formulert:

1. Hvordan har læringsutbyttet vært ved digital Peer Instruction?
2. Hva kjennetegner en produktiv gruppediskusjon i denne settingen?

I dette kapitlet vil vi ta tak i resultatene fra studien for å best mulig kunne svare på forskningsspørsmålene. Vi vil peke på hva resultatene indikerer, og diskutere eventuelle implikasjoner. Vi vil i tillegg ta med oss de begrensinger som er i studien, og diskutere mulige utvidelser eller utbedringer som kunne gitt enda klarere svar på forskningsspørsmålet. Det vil også pekes på momenter og resultater som er spesielt interessante å dykke dypere ned i som materie for videre arbeid, så vel som resultater som har implikasjoner for undervisning. Aller sist i dette kapitlet vil studiens validitet diskuteres.

5.1 Læringsutbytte

For å undersøke hvordan læringsutbyttet fra Peer Instruction har vært, ble det overordnede utbyttet - i form av konseptuell forståelse som målt ved FCI – undersøkt, såvel som utbyttet fra selve gruppediskusjonene. Det er fra resultatene god grunn til å hevde at det overordnede læringsutbyttet fra undervisningen har vært godt, og isolert sett har gruppediskusjonene gitt positiv utvikling. I dette delkapittelet vil vi gå gjennom alle relevante resultater og drøfte de nærmere. Det vil bli pekt på begrensninger, men samtidig på hvilke implikasjoner resultatene kan ha for undervisere og hva som vil være interessant videre arbeid.

5.1.1 Utbytte fra overordnet undervisning

Mazur (1997) rapporterer om verdier for normalisert gain fra FCI på mellom 0,36 og 0,68 ved bruk av Peer Instruction i fysikkundervisning, kontra verdier rundt 0,25 ved konvensjonell undervisning. I sin metastudie fant Hake (1998) at undervisning med bruk av interaktive metoder og tilnærminger ga verdier for normalisert gain på rundt 0,48, mot omtrent 0,23 ved konvensjonell undervisning. Resultatene fra undervisningsløpet denne oppgaven tar for seg gav et normalisert gain på 0,46 (se seksjon seksjon 4.1, som er helt på linje med både det som rapporteres fra Mazur og Hake. En Cohens d på 0,88 underbygger at undervisningen overordnet har hatt god effekt på studentenes konseptuelle forståelse. Gjennomsnittstudenten er etter undervisning tett på å kunne kalles en newtonsk tenker om vi bruker Hestenes mfl. (1992) sin terminologi, og gitt validiteten og reliabiliteten til FCI som mål på konseptuell forståelse, kan det med relativt stor grad av trygghet hevdes at studentene etter gjennomføring av dette emnet står bedre rustet til videre fysikkemner, enn om kurset hadde blitt gjennomført på tradisjonelt vis. Mazur (1997) peker på at det å få utbytte av Peer Instruction krever tålmodighet og tiltro til undervisningsformen. Man høster virkelig fruktene av en omstilling til Peer Instruction etterhvert som det blir mer og mer innarbeidet. Fra Felder og Brent (1996) framkommer det at det ikke vil gi umiddelbar og automatisk læringseffekt å legge om undervisningen fra å være lærer- til elevsentrert og at en slik omstilling er krevende både for underviser og for studentene. Vi har her sett at en omstilling til digital Peer Instruction har båret frukter allerede ved første gjennomføring noe som gir grunn til optimisme for at utbyttet skal bli enda større ved videre utvikling og tilpassing av den aktive undervisningstilnærmingen i de kommende årene.

Det har i denne oppgaven gjort rede for hvordan det aktuelle fysikkemnet ble gjennomført (det henvises til seksjon 3.3), og spesielt for høsten 2020 var det at all fellesundervisning foregikk digitalt. Pandemirestriksjoner gjorde at det

i undervisning og planlegging av undervisning var et overordnet fokus på at studentene skulle få en følelse av fellesskap og tilhørighet. Inndeling i faste grupper var blant annet et slikt tiltak. Det er altså faktorer i fasiliteringen av selve undervisningen som kan ha spilt en rolle for hvordan studentenes utbytte har vært. Som presentert i seksjon 2.5.4 kan faste grupper under de riktige forutsetningene gi svært gode resultater. Samtidig er det en fare for at faste grupper i en Peer Instruction-setting over tid vil henfalle til å stole på den på gruppa som har vist seg å ofte ha rett, og kanskje kunne nye impulser i form av ulike grupper på hver nye oppgave vært fordelaktig. Johnson mfl. (2014) sier at uformelle grupper ofte kan være å foretrekke når studenter eller elever i selve undervisningen skal behandle fagstoff over kort tid. I en Peer Instruction-setting kan for eksempel underviseren sørge for at så mange som mulig til enhver tid diskuterer med noen de er uenige med. I en digital setting er også dette enkelt å fasilitere. Med faste grupper går uansett denne muligheten tapt.

Når vi i denne studien har vurdert det overordnede læringsutbyttet, har vi kun sett på FCI-resultatene. Selv om både Hake (1998) og Mazur (1997) peker på en tydelig sammenheng mellom aktiv undervisning og økt evne til å løse problemløsningsoppgaver, har ikke dette blitt undersøkt her all den tid FCI kun ser på konseptuell forståelse. Det aktuelle emnet hadde obligatoriske innleveringer med mange oppgaver som var problemløsningsorienterte, og omtrent halvparten av eksamen besto av denne typen oppgaver. Det å ha fysikkkunnskaper innebærer både kvalitative og kvantitative aspekter. For å undersøke om fellesundervisningen gikk på bekostning av studentenes problemløsningskompetanse kunne et instrument som måler også dette blitt anvendt, eksempelvis *Mechanics Baseline Test* (Hestenes & Wells, 1992) som er designet for å i større grad enn FCI måle kvantitative aspekter rundt studenters fysikkforståelse. Det skal allikevel nevnes at Peer Instruction som undervisningstilnærming ofte kan føre til at studenter presterer bedre også på problemløsningsbaserte oppgaver, enn det konvensjonell undervisning gjør (Mazur, 1997). Det at denne oppgaven i stor grad har likestilt læringsutbytte med konseptuell forståelse utelukker også å knytte på studentenes egen opplevelse av undervisningen. For å si enda mer om hvordan undervisningen har fungert, hadde det derfor vært av interesse å undersøke hvordan den spilte inn på for eksempel studentenes motivasjon.

Selv om det i utdanningsforskning – spesielt innen fysikkfaget – er stor konsensus om at aktive undervisningsformer har store fordeler sammenlignet med konvensjonell undervisning, pekes det ofte på at undervisere ofte er nølende og skeptiske til det å legge om undervisningen (Knight, 2004; Mazur, 1997). Dette kan blant annet bunne i en redsel for at det tar tid å få gode resultater, og at omlegging vil gå utover læringen til studentene på kort sikt. Denne frykten er ikke ubegrunnet: omlegging er utfordrende, og ikke garantert å gi umiddelbar effekt (Felder & Brent, 1996). Vi har allikevel her sett at en omlegging til Peer

Instruction kan gi umiddelbare resultater i form av økt konseptuell forståelse blant studentene, og det er fra dette grunn til å oppfordre undervisere i fysikkemner til å vurdere aktive tilnærminger til fysikkundervisningen.

Det at fellesundervisningen her har blitt gjennomført digitalt er verdt å dvele litt ekstra ved. Peer Instruction som tilnærming til undervisning forklares av Mazur (1997) i konteksten av en tradisjonell fysisk setting for fysikkundervisning, hvor fellesundervisningen foregår stedlig i en forelesningssal. Dette er også de rammer og fasiliteter en typisk underviser i fysikkemner i Norge i utgangspunktet har til disposisjon. Vi har vist en heldigital gjennomføring av Peer Instruction i et introduksjonskurs i fysikk som tilsynelatende ikke står tilbake for resultater i en fysisk setting. Det kan være grunn til å spekulere i om det er aspekter ved den digitale gjennomføringen som er fordelaktige sammenlignet med en konvensjonell stedlig gjennomføring. I en forelesningssal har ingen studenter faste plasser, det er ofte fastmonterte stoler eller seter, og mobiliteten til studentene og underviser er derfor begrenset. I seksjon 2.5.4 viste vi til Johnson mfl. (2014) som blant annet fremhever viktigheten av gjensidig avhengighet i gruppeinteraksjoner i undervisning. Dette innebærer at gruppe-medlemmer har øyekontakt, og rent fysisk er en enhet. Selv om diskusjonsgrupper i en plattform som Microsoft Teams ikke er fysisk samlet, så er de adskilt fra resten av klassen på en måte som ikke er mulig i en forelesningssal. Gruppen kan fokusere på det de skal gjøre uten å bli forstyrret av de andre gruppene. Generelt vil en digital gjennomføring av gruppediskusjoner være mindre kaotisk, og det vil være mindre tid som går med til omorganisering av klasserommet eller forelesningssalen. For underviser vil det også i en digital setting være enklere å interagere individuelt med de forskjellige gruppene slik at hen dermed kan få et dypere innblikk i hva som foregår i samtalene og hva som må prioriteres i videre undervisning. Det må dog påpekes at et generelt overblikk vil gå tapt i en slik setting. I en fysisk setting kan underviser lettere bli klar over hvor hens tilstedeværelse trengs, noe som er vanskelig digitalt. Det fremkommer interessant å undersøke videre hvordan en digital gjennomføring av gruppediskusjoner kan sammenlignes med fysisk oppmøte i deler av undervisning basert på Peer Instruction. Ved å sammenligne resultatene fra både en fysisk og en digital setting, kan man over tid gjøre seg nytte av en digital setting i de tilfeller hvor det viser seg å være en fordel.

5.1.2 Utbytte fra diskusjoner

Gruppediskusjoner av konseptuelle oppgaver utgjør en stor del av Peer Instruction-basert undervisning, så også i det fysikkemnet denne oppgaven tar for seg. For å evaluere studentenes utbytte av undervisningen, var det derfor også nyttig å undersøke hvordan de presterte på diskusjonsoppgavene.

Kartlegging av oppgaver

Resultatene (det henvises til seksjon 4.1.2) viser – som det jo ville vært naturlig å forvente – at antallet korrekte svar på oppgavene var større etter diskusjon enn før diskusjon på et stort flertall av oppgavene. Effektstørrelsene som ble regnet ut viste at på en gjennomsnittsoppgave hadde gruppediskusjonene god effekt. Vi så allikevel en tendens til at oppgaver hvor få svarer korrekt før diskusjon ga en negativ utvikling. Alle oppgaver med negativ utvikling hadde en gjennomsnittlig førscore på under 0,5. Mazur (1997) erfarer at oppgaver hvor omtrent femti prosent svarer korrekt før diskusjon er ønskelig i Peer Instruction, og fra resultatene i denne studien kan det antydes at hvis færre enn femti prosent svarer galt før diskusjon, kan det gi negative resultater. Umiddelbart er det naturlig å tenke seg at de oppgavene hvor det er negative er for vanskelige til å kunne være formålstjenlige. Det kan allikevel være andre ting enn vanskelighetsgrad som spiller inn for disse. Oppgaven kan for eksempel være tvetydig eller utydelig formulert, og dessuten kan formatet på oppgaven spille en rolle. En oppgave hvor svaralternativene skal rangeres vil for eksempel kunne åpne for mange gale svar som er «nært» det korrekte.

Det er verdt å kommentere at flertallet av oppgavene som i snitt ga negativ utvikling var oppgaver som omhandlet rotasjon. Siden dette er et introduksjons-emne i fysikk, er det rimelig å tro at rotasjon er et nytt konsept (i fysikkfaget) for flertallet av studentene¹. Det kan derfor være at de alternative forestillingene (det henvises til seksjon 2.2) studentene har rundt dette temaet i større grad er til stede enn det som går på rettlinjert bevegelse, og dermed krever mye bearbeiding for korrigerings. For eksempel vil man fra erfaring med translatorisk bevegelse kunne være primet til å tenke på hastighet og akselerasjon som størrelser som kun peker parallelt eller antiparallelt med en bevegelse. I motsetning til de andre temaene som i dette kurset kan være nye for studentene – som fluidmekanikk og termofysikk – konfronterer rotasjonskonseptet oppfatninger vedrørende translasjon som de kan tolke som allmenngyldige. I undervisning av translasjon på studieforbereidende nivå gjøres det mange generaliseringer, og disse kan feste seg og kreve mye bearbeiding når de blir utfordret. For eksempel vil man ofte ikke, før rotasjon behandles, ta hensyn til legemers utstrekning og massefordeling. Slike omstillinger kan kreve store omrokninger av kognitive strukturer, og i så fall trengs tid til bearbeiding. I et sosiokonstruktivistisk perspektiv er det i denne sammenhengen kanskje naturlig at utviklingen i diskusjoner blir negativ dersom ingen i diskusjonsgruppa har en forståelse av rotasjon i tråd med korrekt fysisk forståelse. Det kan fra dette argumenteres for at rotasjonsdelen av et slikt introkurs i fysikk krever mye fokus, og kanskje mer fokus enn de delene som ikke omhandler

1. I gjeldende læreplan for fysikk på videregående (Utdanningsdirektoratet, 2006) er enkel uniform sirkelbevegelse det eneste som inngår av rotasjon

mekanikk.

I denne studien har vi ikke gått inn og sett nærmere på selve oppgavene og undersøkt vanskelighetsgrad, utforming, formuleringer og andre faktorer som kan spille inn på hvordan de oppfattes, tolkes og diskuteres. Vi har vært inne på Mazur (1997) sine krav for adekvate oppgaver i Peer Instruction, og en nærmere analyse av oppgavene som har blitt brukt i det aktuelle emnet ut i fra disse kriteria ville vært interessant. Spesielt interessant som en evaluering av emnet, ville det vært å undersøkt de oppgavene som ga negativ utvikling. Et klarere bilde av hva slags oppgaver som gjennom god diskusjon ga forståelse, og hvilke som ikke gjorde det, ville gitt et enda tydeligere svar på om læringsutbyttet fra undervisningen har vært adekvat. Det er ønskelig å utelate oppgaver som gir studentene nye feilaktige konseptuelle oppfatninger heller enn å utfordre og korrigere dem. Også oppgaver som framstår som for lite utfordrende er ønskelige å luke bort. Mazur anslår at et intervall på omtrent 40 til 60 prosent riktige svar før diskusjon, er optimalt, og at oppgaver som havner utenfor dette bør fjernes eller bearbeides. Gjennomsnittandelen korrekte svar på en oppgave før diskusjon var i denne studien på 0,47 (det henvises til seksjon 4.1.2), men spredningen var relativt stor og mange oppgaver hadde for lav andel korrekte til å være optimale ut i fra Mazurs kriterium.

Kartlegging av studentsvar og faglig inngangsnivå

Når det kommer til den gjennomsnittlige students forbedring på konseptoppgavene fant vi at også den var god. Kun to studenter svarte i snitt mer galt etter diskusjon enn før diskusjon. Effektstørrelsene som ble kalkulert (En Cohens d på 1,37) tilsier at studentene hadde en særdeles god effekt av gruppediskusjonene i henhold til det å svare riktig på oppgavene. Fra fordelingen av studentenes svar før og etter diskusjon kan vi se en tendens til at studenter med høy score på FCI førtest i større grad svarte mer korrekt før diskusjon enn studenter med lavere FCI-score. Denne forskjellen ble mindre tydelig etter diskusjon, da spredningen av data ble mindre. Det siste er både en forventet og ønskelig tendens. Et av målene med Peer Instruction er at studentene skal utfordre og løfte hverandre opp til høyere grad av forståelse gjennom gruppediskusjoner, og det vil være å forvente at de med høy grad av forståelse i forbindelse med en oppgave (og som dermed svarer korrekt inn) hjelper de som sliter med den aktuelle oppgaven til å svare riktig etter diskusjonen.

Det at studenter med et høyere faglig inngangsnivå i større grad svarer korrekt på konseptoppgavene før diskusjon er som nevnt i tråd med det som ville vært forventet. Det er allikevel interessant å se at det også er en sammenheng mellom faglig inngangsnivå og overordnet utbytte av undervisningen. Dette kan tyde på at undervisningen totalt sett gagnar de som har god forståelse

fra starten av, mer enn de med lavere grad av konseptuell forståelse. Det kan derfor være av interesse å videre undersøke om dette stemmer, og hvorfor det i så fall gjør det.

Her er det mye som denne oppgaven ikke har tatt i nærmere øyesyn. Det kan meget vel være andre faktorer enn selve undervisningen – for eksempel motivasjon, tidsbruk, interesse med mer – som gjør at studenter med godt faglig utgangspunkt presterer bedre enn de med et dårligere utgangspunkt. En kartlegging av motivasjon gjennom IMPEL er gjort i etterkant av dette kurset, men ikke blitt brukt i denne oppgaven. I en utvidelse ville disse dataene vært interessante å se i sammenheng med studentenes prestasjoner utifra faglig utgangspunkt.

I undersøkelsen av forbedring fra diskusjoner har vi ikke hatt et kvalitetsikret og etterprøvd instrument tilsvarende FCI for den overordnede undervisningen, men i stedet brukt besvarelsene på konseptoppgavene som ble gitt i løpet av undervisningen. Som poengtert i kapittel 3 var omtrent halvparten av disse hentet fra Mazur (1997) og majoriteten var laget for å teste konseptuell forståelse. Kravene Mazur setter for gode oppgaver i Peer Instruction, er ikke ulike malen oppgavene i FCI er designet etter, og det går an å hevde at konseptoppgavene fra undervisning måler en slags konseptuell forståelse oppnådd fra diskusjonene. Dette skal vi allikevel være forsiktig med. Der FCI er et nøye sammensatt verktøy for å måle et bestemt konstrukt, er oppgavene fra undervisning laget til de aktuelle tema som gjennomgås, og er ikke i utgangspunktet ment til å måle et bestemt overordnet konstrukt. Derfor kan vi heller ikke si noe om korrekt svar etter diskusjon av en oppgave faktisk samsvarer med økt konseptuell forståelse. For å få bedre innblikk i studentenes utvikling fra diskusjonene, kunne studentene blitt bedt om å oppgi hvor trygge de var på sine svar. Fra dette kunne man både fått innblikk i hvordan studentenes selvsikkerhet på egen forståelse spilte inn på utviklingen av diskusjonene, og man kunne lettere se hvilke oppgaver studentene *faktisk* oppfattet som enkle eller vanskelige.

For å måle effekten av diskusjonsoppgavene ble normalisert gain brukt som effektstørrelse. En begrensning med diskusjonsoppgavene er at før- og etterdata ikke er adskilt i tid på samme måte som for eksempel FCI. Førdata er svar på oppgaver før diskusjon spret utover et stort tidsrom. Etterdata er svar på oppgaver etter diskusjoner, og spredt utover det samme tidsrommet. Det er altså et tydelig temporalt skille mellom all førdata og all etterdata. I tillegg er studentenes svar etter diskusjon tett knyttet opp mot hva de svarte før diskusjon. Fordi det ikke er et stort tidsrom mellom svarene, og fordi svarene ikke er uavhengige gir det grobunn for ytterligere skepsis vedrørende. Normalisert gain som effektstørrelse her. Kanskje kunne et annet mål på utvikling fra diskusjonene vært bedre egnet.

Sammenhengen mellom forbedring på diskusjonsoppgaver og overordnet læringsutbytte

Fra figur 4.7 i seksjon 4.1.2 ble det funnet å ikke være signifikant korrelasjon mellom individuelle normaliserte gain på konseptoppgavene og det overordnede læringsutbyttet som målt ved både normalisert gain på FCI og score på FCI ettertest. Gitt diskusjonsoppgavens sentrale posisjon i Peer Instruction, er dette et høyst interessant resultat. Tilsynelatende kan det virke som den forbedringen studentene i snitt har fra en diskusjon av en oppgave ikke henger sammen med den overordnede konseptuelle forståelsen de sitter igjen med ved endt undervisning. Dette kan tyde på at utviklingen i diskusjoner kun målt gjennom andel korrekte svar ikke holder for å si noe om diskusjonene har gitt økt læring, og dermed har vært konstruktive. Dette resultatet åpner for å stille spørsmål ved hvor god effekt diskusjonsdelen av undervisningen totalt sett har hatt, men det kan være vel så betimelig å peke på hva som anses som en produktiv diskusjon. Kanskje kan en diskusjon hvor få kommer fram til korrekt svar være vel så produktiv – i form av at den fører til læring – som en diskusjon som fører til mange korrekte svar?

Som det har kommet fram tidligere i denne seksjonen, så er ikke konseptoppgavene laget i den hensikt å direkte måle forståelse. Samtidig er det ikke nødvendigvis korrekt å bruke effektstørrelsen normalisert gain i denne sammenhengen. Til sammen gir dette grunn til å nyansere bildet av forbedring på konseptoppgavene som ikke-korrelert med overordnet forståelse. Fra figur 4.8 i seksjon 4.1.2 fant vi at det var en signifikant korrelasjon mellom studentenes gjennomsnittlige score på konseptoppgavene (både før og etter diskusjon), og overordnet konseptuell forståelse som målt ved score fra FCI ettertest. I Peer Instruction har underviseren liten mulighet til å få oversikt over hva som foregår i diskusjonene. Hen må forholde seg direkte til hvor mange som svarer korrekt på hver enkelt oppgave for å avgjøre om det enten er ønskelig å gå videre til nye tema i undervisningen, eller om det er nødvendig å dvele lengre ved det til enhver tid aktuelle temaet. Dette blir dermed det verktøyet læreren har for å kunne vite om studentene i tilstrekkelig grad har forstått et konsept. Det at selve prestasjonen på konseptoppgavene – i motsetning til normalisert gain – korrelerer med læringsutbyttet, gir grunn til å tro at læreren får verdifull informasjon om studentenes forståelse fra konseptoppgavesvarene i seg selv. Snittscore på diskusjonsoppgavene etter diskusjon er i likhet med et standard gain (score på ettertest minus score på førtest) en et mål som i større grad er anvendbart for den aktuelle konteksten enn det en mer raffinert effektstørrelse som normalisert gain er (det henvises til seksjon 2.6.1). Fordi konseptoppgavene ikke er ment å måle noe spesifikt, kan det hevdes at et slikt «renere» mål på effekt er bedre egnet her.

I analysen av sammenhengen mellom forbedring på konseptoppgaver og det

overordnede læringsutbyttet, er det ikke tatt høyde for at andre deler av undervisningen enn diskusjonene i fellesundervisninga kan spille inn på sistnevnte. Seminarene, regneverksteder, innleveringer og selvstudium kan ha hatt en substansiell innflytelse på læringsutbyttet, men denne studien har ikke tatt sikte på å undersøke dette. I en større undersøkelse av Peer Instruction i fysikkundervisning kunne også disse aspektene blitt vektlagt. Undervisningen i det aktuelle fysikkemnet har tatt for seg flere tema utover newtons mekanikk og kraftbegrepet, som FCI tar utgangspunkt i. I gruppediskusjonene ble oppgaver innen blant annet fluidmekanikk og termofysikk diskutert. Det kan stilles spørsmål ved hvorvidt det er fornuftig å sammenligne prestasjon på disse oppgavene med en overordnet forståelse av krefter, til tross for at en forståelse av det newtonske kraftbegrepet kan sies å være grunnleggende for forståelsen av andre områder i fysikken.

Siden det her kan antydes at utviklingen i antall korrekte inn og ut av diskusjonene ikke nødvendigvis har noe å si for det overordnede læringsutbyttet vil det være interessant å undersøke nærmere hva det som gjør at en diskusjon er produktiv utover det at mange har svart korrekt etter diskusjonen. Igjen vil jeg peke på en praksis hvor studentene oppgir hvor trygge de er på svarene sine som fordelaktig. Det vil gjøre underviseren sikrere på studentenes forståelse, og det vill også gjort en analyse av utbyttet fra diskusjonsoppgavene enklere og mer solid.

5.2 Produktive diskusjoner

For å kunne si noe om hva som gjør gruppediskusjonene i undervisningen produktive, har både kvantitative og kvalitative analyser blitt utført. Vi har tatt for oss sammenhengen mellom studentenes faglige inngangsnivå og hvor stor rolle korrekthet og uenighet inn i diskusjonene spiller for utfallet. En kvalitativ analyse av utvalgte diskusjoner ble også utført. Her vil resultatene fra disse analysene bli drøftet i henhold til det aktuelle forskningsspørsmålet. Det vil bli pekt på begrensninger og aspekter som er interessante for videre arbeid eller har implikasjoner for underviseren.

5.2.1 Inngangsnivå mot prestasjon på diskusjonsoppgaver

Det faglige utgangspunkt ble funnet å spille en rolle for prestasjonen på selve konseptoppgavene. Sammenhengen mellom score på FCI førtest og score på konseptoppgavene henholdsvis før og etter diskusjon ble vist å være signifikant. Dette gir styrket validitet til FCI som et mål på studentenes konseptuelle forståelse inn i undervisning. Resultatene tyder altså på at studenter med et

godt utgangspunkt hva gjelder konseptuell forståelse virker å ha best utbytte av den overordnede undervisningen, og vi har allerede sett at de presterer bedre på konseptoppgavene. Det kan derfor tenkes at for å fasilitere for produktive diskusjoner så bør det være minst en person på gruppa som har gått inn i undervisningen med høy grad av konseptuell forståelse.

Dette er interessant med tanke på gruppesammensetning. Dersom læreren ønsker å sette sammen heterogene grupper, kan FCI førtest være en god indikator for hvilke studenter som oftest vil ha korrekt svar inn i gruppediskusjonene. Dersom det er minst en student med god forståelse på gruppe er det rimelig å tro at denne studenten vil kunne hjelpe gruppemedlemmer med et svakere utgangspunkt til bedre sin forståelse på det konseptet som behandles i den aktuelle oppgaven.

5.2.2 Korrekthet og uenighet

Korrekthet

For å nærmere undersøke hva som kjennetegner gruppediskusjoner hvor mange har korrekt i etterkant, ble alle instanser av diskusjoner undersøkt (det henvises til seksjon 4.2.2). Analysen avslørte at jo flere på gruppa som hadde korrekt svar inn i diskusjon, jo flere hadde korrekt også etter diskusjonen. Det kan også se ut som den effektive gruppestørrelsen har noe å si for utviklingen av diskusjonen. I effektive treergrupper ble det funnet at ett medlem med korrekt svar inn i diskusjonen, holdt til at diskusjonen ga en positiv utvikling, mens det i de større gruppene var en tendens til at flere enn en måtte ha korrekt inn for at dette kunne skje. Det er på sin plass å dvele litt ved det faktum at resultatene viser *effektive* gruppestørrelser. I utgangspunktet var alle grupper bestående av fem til seks personer, men i analysen av diskusjonene har studenter som ikke svarte på oppgaven før og/eller etter diskusjon blitt fjernet fra datamaterialet. Det utelukker ikke at disse studentene kan ha deltatt i diskusjonene. Dessuten kan en gruppe med effektiv størrelse på tre studenter i en instans, bestå av seks medlemmer i en annen. For å gi validitet til de resultatene som tar i betraktning de effektive gruppestørrelsene, kunne alle diskusjoner blitt undersøkt for samsvar mellom deltagelse i diskusjon og hvor mange som svarte på konseptoppgavene.

For å svare på hva som kjennetegner produktive diskusjoner i Peer Instruction ble det i denne delen av studien valgt å ta utgangspunkt i utviklingen i antall korrekte svar fra før diskusjon til etter diskusjon som et mål på produktivitet. Dette er en tilsynelatende fornuftig tilnærming all den tid det er rimelig å hevde at studentene får mest ut av de diskusjonene hvor de sitter igjen med korrekt kunnskap og forståelse. Det må allikevel tas høyde for at diskusjoner

uten en positiv utvikling i henhold til antall korrekte svar også kan være produktive. Dersom dialogen i gruppa er produktiv kan medlemmene sitte igjen med forbedret forståelse av et fysikkonsept uten å nødvendigvis svare riktig på den aktuelle oppgaven.

Det at større grupper helst bør ha flere enn ett gruppemedlem med korrekt svar inn i diskusjonen, er også interessant i henhold til gruppeinndeling. For den som underviser er det ønskelig å sørge for at gruppene består av flest mulig medlemmer som deltar aktivt og som svarer på konseptoppgavene, for at diskusjonene skal føre til at flest mulig har korrekt etterpå. Det er jo dette som på mange måter er målet med Peer Instruction. Det er dermed verdt å ta med seg at større grupper bør tilordnes på et vis som gjør at det så ofte som mulig er to personer som har korrekt svar inn i diskusjonen. Rollen til underviser er også viktig utover det å lage gode grupper. Under premisset om at diskusjoner i seg selv kan være produktive uavhengig av resultat, er det viktig at underviseren klargjør ting for studentene i etterkant av diskusjonen. Da kan brikkene falle på plass hos de studentene som etter gruppediskusjonene har nærmet seg å oppnå forståelse, men som ikke helt kom i mål.

Uenighet

For å se på uenighet inn og ut av diskusjon ble antall unike svar hos gruppene undersøkt. Det ble, ikke uventet, funnet at gruppene ble mer enige etter å ha diskutert oppgavene. I en samtale i Peer Instruction er det ønskelig at studentene skal instruere hverandre, og hvis alle på gruppa er enige om samme svar før diskusjonen, er det rimelig å anta at diskusjonen ikke nødvendigvis blir veldig fruktbar. I seksjon 2.5.5 viste vi til Johnson mfl. (2014) som pekte på at det å være uenige om fagstoffet som behandles kan ha en langsiktig læringseffekt. Gjennom å initielt være uenige er det større rom på gruppa for å kunne argumentere for sine synspunkt. Det er derfor rimelig å ønske en viss grad av uenighet inn i diskusjonene for at diskusjonene skal være produktive. Fra resultatene har vi dog sett at det på ingen måte er noe som tilsier at total uenighet på gruppa har gode effekter på utviklingen gjennom diskusjonen, og at helst bør minst to av gruppemedlemmene være enige om korrekt svar, for at diskusjonen skal gi gode resultater.

Hvis alle på gruppa har svart forskjellig er det et iboende potensiale for uenighet og konflikt i gruppa, og fra det kan det i beste fall komme diskusjoner av høy kvalitet, noe som igjen fører til økt forståelse hos gruppemedlemmene. I verste fall kan det føre til at ingen blir trygge på hva som er riktig svar og dermed mister motivasjon. Resultatene her tyder på at sistnevnte skjer. For grupper med flere enn tre medlemmer er utviklingen negativ i omtrent halvparten av tilfellene hvor alle er uenige inn i diskusjonen. Samtidig kan man postulere at et

utgangspunkt hvor alle er enige potensielt kan være hemmende for diskusjonen. Dette kan dog avhenge av om det er enighet om riktig svar eller ikke. Vi har funnet at i majoriteten av tilfellene hvor alle er enige inn i diskusjon, så er de enige om det korekte svaret, og at dette sjelden fører til at gruppemedlemmer endrer svar.

I denne oppgaven har vi ikke kvantifisert uenighet på annen måte enn at vi har sett på antall unike svar på hver gruppe. Fordi de effektive gruppestørrelsene varierte ble det videre valgt å ikke dele inn enighet i snevrere kategorier enn «helt enig», «helt uenig» og «hverken/eller». Hadde andre mål på uenighet blitt brukt, så kunne man kanskje sette klarere tendenser til hvordan korrekthet ut av diskusjonene avhenger av uenighet inn. Studentsvar på konseptuelle oppgaver som gitt i Peer Instruction kan i stor grad sies å være kategoriske data, og uenighet er et mål på variabiliteten i dataene. Det er mange måter man kan måle dette på, og alle har ulike fordeler og ulemper (Allaj, 2018; Kader & Perry, 2007). I denne oppgaven ble *Shannon-entropi* (Shannon, 1948) vurdert brukt, men de små effektive gruppestørrelsene avslørte svakheter i dette målet.

Til sammen peker resultatene på at antallet gruppemedlemmer som har korrekt inn i diskusjonene spiller en rolle for utviklingen, og at større grupper krever flere med korrekt inn for å få ønsket utvikling. Det er lite som tyder på at total uenighet skaper god utvikling, og noen av gruppemedlemmene bør være enige om korrekt svar for at så mange som mulig skal svare riktig etter diskusjonen. Det kunne vært interessant å undersøkt hva som skjer i de diskusjonene hvor alle er enige om korrekt svar inn og noen endrer til galt svar ut. Det samme gjelder de diskusjoner hvor alle er uenige i utgangspunktet, men i løpet av diskusjonen blir enige om korrekt svar. Igjen hadde analysen blitt enklere og sikrere dersom studentene også oppga hvor trygge de var på sitt svar når de besvarte konseptoppgavene. Det er rimelig å tro at dersom alle har svart riktig inn, men ikke er trygge på svaret sitt, så er det lettere å endre svar etter diskusjonen. På samme måte kan det godt hende at den som til enhver tid er tryggest på sitt svar inn i diskusjon lettest klarer å overbevise de andre, i en diskusjon med stor uenighet inn.

5.2.3 Hva skjer i diskusjonene?

Fra gjennomgangen av de diskusjonene hvor få hadde korrekt svar inn i diskusjon og mange hadde korrekt etterpå ble det funnet at diskusjonene kunne kunne kategoriseres utifra hvor mange som bidro til at gruppa kom fram til riktig svar. Enten var det et gruppemedlem som gjennom sine argumenter og forklaringer overbeviste resten, eller så oppnådde gruppa konsensus gjennom dialog hvor flere av gruppemedlemmene bidro. Ser vi til Mercer og Littleton

(2007) vil ikke den første kategorien av diskusjoner kunne sies å være produktive diskusjoner, så lenge ingen bygger på eller utfordrer argumentene til den som forklarer. De diskusjonene hvor flere bidrar på veien mot korrekt svar er derimot nærmere å kunne defineres som produktive i henhold til Mercer og Littleton og dermed i større grad enn diskusjonene i den andre kategorien føre til læring og forståelse hos deltagerne. I seksjon 2.5.3 ble det vist til Bakhtin og Holquist (1981) som peker på hvordan personlig respons og tolkning av andres resonnementer ligger til grunn for ens egen forståelse. Dersom det er kun en person som aktivt tar del i diskusjonen, vil ikke hen få respons eller motargumenter, og vedkommende har dermed kun sine egne argumenter å tolke og korrigere. Videre vil de tyste tilhørerne ikke få mulighet til å prosessere argumentene til den som forklarer på best mulig vis ved kun å lytte, men det skal ikke utelukkes at informasjonen og argumentene prosesseres i en indre dialog. Det er derfor ikke noen grunn til å totalt avfeie disse diskusjonene som uproduktive eller meningsløse.

I den kvalitative analysen ble det videre funnet at i de diskusjonen hvor en person overbeviste resten, så var dette i mange tilfeller det gruppemedlemmet som hadde størst grad av konseptuell forståelse i starten av undervisningen, som målt ved FCI. Selv om en ikke skal legge for mye i denne observasjonen, er det interessant å knytte dette opp mot det som ble funnet i analysen av FCI-data. Her kom det fram at det var en signifikant sammenheng mellom faglig inngangsnivå – i form av FCI førtest – og det overordnede utbyttet av undervisningen. Vi fant også at det var sammenheng mellom inngangsnivået og prestasjonen på diskusjonsoppgavene i undervisningen. Det er nærliggende å tenke seg at en student som er faglig sterk og som i tillegg er trygg på sin forståelse, lettere vil ta ansvar i en diskusjonssetting. Hvis det i diskusjoner kun er denne personen som bidrar, er det også kun denne som får artikulert og brukt sin forståelse aktivt. De passive tilhørerne gjør ikke det. Til sammen er det rimelig å tenke den som forklarer jevnt over har et større utbytte fra gruppediskusjonene enn de som ikke gjør det. Dette er igjen i tråd med Bakhtin og Holquist (1981) som sier at lytting til andre er utilstrekkelig for å oppnå korrekt forståelse.

I et perspektiv på læring som baserer seg på *Knowledge in Pieces* (KiP) (Det henvises til seksjon 2.1.1) så vil gruppediskusjonene i en Peer Instruction-setting på mange måter dreie seg om å konfrontere og korrigere alternative oppfatninger og kognitive strukturer som avviker fra en koherent, korrekt fysikkforståelse. I gruppediskusjoner utfordres disse gjennom andres påstander, resonnementer og argumenter. I et tilfelle hvor det kun er en person som presenterer sitt synspunkt er det lite som skiller settingen fra en tradisjonell undervisningssetting, hvor en gruppe med studenter passivt mottar kunnskap uten å prosessere den aktivt. Tar vi utgangspunkt i at aktiv læring er å foretrekke foran slik passiv kunnskapsoverføring, er det ukontroversielt å hevde at diskusjoner hvor kun

en student deltar aktivt ikke er ønskelige i Peer Instruction.

Drar vi KiP-perspektivet videre hadde det vært interessant å undersøke hvilke kognitive ressurser studentene tar i bruk i gruppediskusjonene. Et KiP-perspektiv på læring innebærer at de ressursene som aktiveres i læringsprosesser er avhengige av kontekst. Det kan dessuten tenkes at de resursene en student drar veksler på i en gruppediskusjon i formell undervisning ikke korresponderer med de hen bruker i for eksempel en uformell samtale om fagstoff med medstudenter på fritiden. Igjen er det interessant for videre studier å nærmere undersøke konteksten Peer Instruction foregår i. Kanskje vil den digitale konteksten som preger dette fysikkemnet også spille inn på hvordan og i hvor stor grad studentene tar i bruk ulike kognitive ressurser.

Det ble observert at det ofte var studenter med høy konseptuell forståelse som overbeviste de andre i diskusjonene. Tatt i betraktning undervisningens gjennomføring med faste grupper semesteret igjennom, kan det – som tidligere påpekt – tenkes at gruppa underveis lærte seg hvem som ofte svarte riktig på diskusjonsoppgavene. Derfor kan de lettere ha aksepterte, bevisst eller ubevisst, det vedkommende hadde å si sammenlignet med andre gruppe-medlemmers innspill. I seksjon 2.1.1 ble det vist til det Hammer (2004) refererer til som epistemologiske ressurser. I passive settinger hvor studenter mottar kunnskap uten aktivt behandle den, går det for eksempel an å hevde at mottakeren i denne konteksten forfekter et syn på kunnskap som noe som kommer fra en autoritet, heller enn noe man konstruerer selv. Bing og Redish (2009) peker på hvordan studenters bruk av epistemologiske ressurser kommer til syne i argumentasjon og problemløsning, og en analyse i tråd med deres rammeverk for å undersøke studenters bruk av matematikk i problemløsningsoppgaver i fysikk, kunne vært svært interessant å utføre på gruppediskusjonene i dette emnet for å få enda dypere innsikt i hva som gjør diskusjoner produktive.

I denne studien har vi kun skimmet overflaten av hva det er som faktisk foregår i gruppediskusjoner i Peer Instruction. I en utvidelse av studien hadde det vært særdeles interessant å dykke dypere i selve gruppediskusjonen i Peer Instruction og finne ut hva disse kjennetegnes av. Som det flere ganger er blitt poengtert i denne oppgaven er effektene av Peer Instruction som undervisningsform påvist og etterprøvd positive, og det viser seg sågar fra de resultater som er oppnådd fra denne studien. Det er derimot gjort mindre forskning som tar for seg hva som skjer i selve diskusjonene, og en grundig analyse av diskusjonene ville vært av interesse også utover denne spesifikke oppgaven. For fagansvarlig vil en slik dyp innsikt være verdifull informasjon for å best mulig kunne tilpasse undervisningen i fremtiden.

I en studie fra Skottland kartla Wood, Galloway, Hardy og Sinclair (2014) studenters aktivering av ressurser i diskusjoner i Peer Instruction, og fant at en

slik analyse kan gi en dypere forståelse av hvordan dialoger i Peer Instruction kan føre til læring. Wood mfl. fant også at deres analyse kan si noe om hva som skiller suksessfulle økter fra mindre suksessfulle økter. Spesielt peker de på hvordan en slik grundig analyse kan belyse uoverensstemmelser mellom diskusjonsoppgaver intensjon og faktisk effekt på diskusjonene. Et interessant resultat var at det ikke virket å være noen klar sammenheng mellom hvilke ressurser studentene aktiverer, og utviklingen i diskusjonene. Wood mfl. sin studie var av en liten skala ($N = 25$), og forfatterne etterlyser og oppfordrer selv til utvidelse og supplerende undersøkelser til sin studie.

I en annen studie fra 2006 fant James (2006) at kvaliteten på diskusjoner i Peer Instruction henger sammen med studentenes opplevelse av viktigheten ved å svare riktig på oppgavene. I klasser hvor det var lav terskel for gale svar ble det funnet at studentene uavhengig av inngangsnivå bidro i diskusjonene og at det var en større tendens til uenighet etter diskusjon. Om studentene derimot opplevde at det var mindre rom for å svare galt var tendensen at den mest kunnskapsrike studenten dominerte diskusjonene. I lys av dette hadde det vært svært interessant å gå videre med analysen av diskusjoner hvor en person dominerte diskusjonene. Kanskje kan det være at mange diskusjoner forløp slik fordi studentene i noen grad opplevde et press om å svare korrekt? Mazur (1997) poengterer selv at dersom Peer Instruction skal fungere, så må studentene oppleve en følelse av samarbeid i klasserommet, og ikke oppleve at det er en konkurranse om å svare rett på oppgaver. Dersom studentene undervisningen har en opplevelse av at det er selve diskusjonen som er viktig, og ikke resultatet av den, så kan det føre til flere diskusjoner av god kvalitet. Igjen så ser vi at den som underviser spiller en stor rolle for å fasilitere for gode diskusjoner.

I seksjon 2.5.3 ble det vist til at kvaliteten på dialoger i klasserommet øker dersom det blir gitt bevisst og målrettet opplæring i utforskende samtaler. Hvis studentene blir instruert og oppfordret til å hele tiden bygge på, og utfordre hverandres synspunkt, vil kvaliteten på diskusjonene øke, og dermed øker også fagforståelsen. Videre viste vi i seksjon 2.5.2 til Venville og Dawson (2010) som fant at undervisning en sammenheng mellom argumentasjonskompetanse og fagforståelse i realfag. Dersom underviser ønsker at flest mulig skal delta aktivt og konstruktivt i gruppesamtalene i Peer Instruction, kan det være en idé å sette av tid i undervisningen til å instruere og lære studentene opp i å kunne argumentere på en god måte, og å kunne ha utforskende samtaler. Dette vil kunne medføre at det blir færre samtaler hvor kun ett gruppedemedlem bidrar.

5.3 Studiens validitet

I denne studien har vi sjekket for sammenhenger mellom forskjellige størrelser knyttet til både læringsutbytte og til resultater på diskusjonsoppgaver. Vi har også prøvd å si noe kvalitativt om hva som kjennetegner produktive diskusjoner. En studies validitet er stor grad priggitt studiens forskningsdesign. I kapittel 3 beskrev vi hvordan denne studien har et flermetodedesign i tilnærmingen til å besvare forskningsspørsmålene. I det følgende vil vi på bakgrunn av dette evaluere studiens indre og ytre validitet.

5.3.1 Indre validitet

En studies indre validitet peker på hvor trygg man kan være på at de undersøkelser av sammenhenger som gjøres er til å stole på, og ikke påvirkes av andre variabler og faktorer (Behi & Nolan, 1996). At en studie har stor indre validitet innebærer at resultatene fra studien er i tråd med studiens intensjon, og dermed at resultatene fra studien vil være interessante for de studien angår – i dette tilfellet innebærer dette blant annet fagansvarlig for det aktuelle fysikkemnet. Det er flere faktorer som kan spille inn og forringe en studies indre validitet. Faktorer aktuelle for denne studien vil nå gås gjennom.

Ved en bruk av før- og ettertester som instrumenter i en studie er det en fare for at resultater på førtest påvirker resultatene på etter testen, for eksempel ved at deltagere husker hva de svarte på førtest. I denne studien har vi brukt FCI i som et instrument på denne måten. For å unngå denne effekten ble førtest lagt helt i begynnelsen av semesteret og førtest helt mot slutten, slik at tidsrommet mellom de to skulle være størst mulig. I tillegg ble ikke FCI vektlagt i undervisningen og det ble klargjort at den ikke spilte inn på vurderinger eller evalueringer av studentene. Det ble poengtert i seksjon 2.3.1 at FCIs validitet kan forringes dersom det undervise for FCI. Dette er altså ikke blitt gjort her. Samme versjon av FCI ble brukt både til førtest og ettertest, noe som også betyr at studentenes konseptuelle forståelse er blitt målt på samme vis før og etter undervisning. Ofte trues også slike testers validitet av at svake resultater på førtest gir større forbedring til ettertest enn det sterkere resultater gjør. Dette er tatt hensyn til i denne studien ved å beregne normalisert gain, som ser på fraksjonen av potensiell forbedring i stedet for total forbedring.

I denne studien har alle deltagere overordnet deltatt på den samme undervisningen i den samme settingen. Resultatene fra denne studien som tar for seg Peer Instruction har ikke blitt sammenlignet med en kontrollgruppe som for eksempel fulgte tradisjonell undervisning. Dette gjør at man ikke har noe indre sammenligningsgrunnlag for signifikansen til resultatene, noe som kan virke inn på den indre validiteten til studien. Verdier fra FCI-analysen har kun

blitt sammenlignet med resultater fra studier langt unna denne både i tid og lokasjon.

Rammebetingelsene for denne studien har vært like gjennom hele perioden den har pågått. Det var i god tid før semesterstart avgjort at fellesundervisningen skulle bli heldigital, og at den skulle foregå ved Peer Instruction, med et opplegg som var fastsatt på forhånd. Med unntak av opptak ble data samlet inn ved starten og slutten av semesteret, og ingenting ble analysert før etter endt undervisning.

Til slutt må det igjen nevnes at konseptoppgavene fra undervisningen har blitt brukt som et instrument i denne studien, og da til å måle et utbytte av selve diskusjonene. Fordi disse ikke er laget for å måle et slikt konstrukt, synliggjørres her en betydelig svakhet ved denne studien som forringer dens indre validitet. Konseptoppgavene fra undervisningen er i stor grad blitt brukt på samme måte som FCI, og andel korrekte svar før diskusjon er blitt sammenlignet med andel korrekte svar etter diskusjon. Der FCI før- og ettertest er lagt opp til å ikke skulle påvirke hverandre, er dette nærsagt uunngåelig for konseptoppgavene. Premisset for diskusjonene er jo på mange måter hva gruppemedlemmene har svart i forkant. En utvidelse av denne studien som i større omfang kartla hva som gjorde at studenter svarte riktig etter diskusjon ville derfor kunne bedre fastslå validiteten – eller mangelen av denne – til konseptoppgavene som et mål på utvikling i diskusjonene. Kanskje kunne intervjuer med studenter også være til hjelp her. Dersom det kom fram at studentene endret svar fordi de faktisk endret forståelse, ville det vært med på å styrke denne studiens indre validitet.

Spesielt for den kvalitative delen av studien som angår analyse av videopptak, kan det være vanskelig å si noe konkret om validiteten. Derfor er resultatene fra denne delen forsøkt presentert ryddig og oversiktlig, og strategier for fremgangsmåte og analyser er gjort rede for ved transparente beskrivelser av de valg som er gjort. Det som imidlertid må tas i betraktning er de bias jeg som observatør kan ha hatt i arbeide med opptakene. Her er det flere ting som kan forringe studiens indre validitet. Jeg kan underbevisst ha lett etter sammenhenger som jeg ønsket å finne slik at dataanalyse og utvalg kan påvirkes. Fordi analyser og resultater som bunner i observasjon utført av en person aldri kan være fullstendig objektive og påvirket av slike effekter, har deler av gruppediskusjonene som er undersøkt i denne studien, også blitt sett av veileder, og i etterkant har disse blitt diskutert for å se om vi har vært enige om funnene. Transparens og redelighet i presentasjon av metodikk og resultater er vektlagt for å styrke studiens indre validitet. En annen faktor relatert til mine bias som observatør er min rolle som seminarlærer i det aktuelle kurset. I observasjon av studentdiskusjoner kan jeg underbevisst ha tillagt de studentene jeg visste noe om fra seminarer andre holdninger, prestasjoner og intensjoner enn de jeg

ikke kjente til.

5.3.2 Ytre validitet

Den ytre validiteten til en studie viser til i hvor stor grad studien kan bli gjentatt eller reproduisert i andre kontekster, grupper og sammenhenger (Behi & Nolan, 1996). Faktorer som kan være med å forringe en studies ytre validitet handler ofte om studiens deltagere. Populasjonsvaliditeten til en studie sier noe om hvor godt deltagere i studien representerer den gruppen man ønsker å finne ut noe om. I denne studien ønsker man å si noe om hva som foregår i undervisning av introduksjonskurs i fysikk på universitetsnivå. Det er derfor naturlig at det er mange studenter som skal gå videre med fysikk som tar slike kurs, men det kan også være studenter som tar andre studieretninger. Som det kom fram i seksjon 3.2 har de som samtykket til denne studien vært studenter fra studieprogram som tar viderekommende fysikkemner. Kanskje kan rene fysikkstudenters interesse for faget ha ført til at de har hatt lettere for å delta i studien. Det må altså ta høyde for at denne studiens ytre validitet kan være påvirket av et sampling bias, hvor rene fysikkstudenter er overrepresentert. I så måte er utvalget for denne studien representativt for fysikkstudenter heller enn studenter som tar introduksjonemner i fysikk. Tilsvarende kurs ved samme universitet ble undersøkt i en studie av Netland, Sivertsen og Olufsen (2018). Her ble det rapportert resultater fra FCI førtest som var syv prosent lavere enn det som er rapportert i denne studien. Utover dette finnes det lite som gjør at vi kan generalisere til utenforliggende populasjoner.

En annen trussel til studiens ytre validitet er det faktum at deltagerne vet at de blir studert, og at det dermed kan påvirke resultatene. Dette refereres i litteraturen til som *Hawthorne-effekten*, og i denne studien kan den spesielt gjøre seg gjeldende i forbindelse med opptak av gruppesamtaler. Dette ble forsøkt forbygget ved at deltagerne tidlig ble varslet om at det ville bli tatt opptak av gruppediskusjonene, og de ble ikke informert om hvilke diskusjoner som ble tatt opp og hvilke som ikke ble det. Observatør var til stede i alle Teams-samtaler (uten lyd og bilde) etter at det tekniske oppsettet var på plass – også før opptak av diskusjoner ble påbegynt. Dette i håp om at deltakerne i mindre grad agerte på forskjellig vis avhengig om de ble tatt opp eller ikke. Et annet tiltak i denne sammenhengen var å ikke benytte de første opptakene som ble gjort, hvor studentene i større grad enn senere tenker over at de kan bli tatt opp. Igjen spiller også min rolle som seminarlærer inn her. Enkelte studenter hadde gjennom dette kjennskap til at jeg gjorde denne studien, og deres visshet om dette kan i ytterste konsekvens ha gjort dem enda mer selvbevisste i diskusjonene i fellesundervisningen, enn hvis de ikke visste hvem som tok opptak.

/6

Avslutning

Denne oppgaven har tatt for seg et introduksjonsemne i fysikk på universetsnivå hvor fellesundervisningen har blitt gjennomført digitalt med en Peer Instruction-tilnærming. Ønsket har vært å undersøke hvordan læringsutbyttet har vært, og hva som må til for å skape produktive diskusjoner i en slik undervisningssetting. Derfor ble følgende to forskningsspørsmål formulert:

1. Hvordan har læringsutbyttet vært ved digital Peer Instruction?
2. Hva kjennetegner en produktiv gruppediskusjon i denne settingen?

6.1 Konklusjon

Fra gjennomføringen av denne studien kan det konkluderes med at undervisningen i det aktuelle emnet har gitt ønskelige resultater. Det konseptuelle læringsutbyttet – som målt ved Force Concept Inventory (FCI) – er i tråd med det som ville være forventet med en slik aktiv undervisningstilnærming, og indikerer at studentenes læringsutbytte er bedre enn det ville vært ved konvensjonell undervisning. Gruppediskusjonene i undervisningen ble funnet å ha en god effekt, og det ble funnet å være en signifikant sammenheng mellom studentenes score på en gjennomsnittlig konseptoppgaver etter diskusjon og overordnet læringsutbytte som målt ved FCI-score på ettertest.

Fra analyse av gruppediskusjonene ble det funnet at gruppene burde bestå

av minst en person med korrekt svar inn i diskusjonen, for at flertallet skal svare korrekt etter diskusjonen. Mens det for større grupper ble funnet å være fordelaktig at minst to gruppemedlemmer har korrekt svar inn. Videre fant vi at total uenighet innad i en gruppe ikke er ønskelig for god utvikling i diskusjonen.

Fra sammenligning av resultater fra FCI og data fra konseptoppgavene, ble det funnet at studenter med høy score på FCI førtest hadde bedre utbytte av både gruppediskusjonene og den overordnede undervisningen, enn studenter med dårligere score. Vi har altså funnet at den overordnede undervisningen har gagnet de studentene som hadde god konseptuell forståelse, bedre enn de med lavere grad av forståelse. Det ble også funnet at gruppediskusjoner der få gruppemedlemmer hadde korrekt svar inn og mange hadde korrekt svar ut var kjennetegnet enten ved at ett gruppemedlem overbeviste de andre, eller ved at gruppa ble enig om riktig svar gjennom dialog. I de tilfellene hvor én person overbeviste resten var det ofte det gruppemedlemmet med høyest score fra FCI pretest som fikk mulighet til å forklare for de andre, mens de andre ikke deltok aktivt i diskusjonen.

Det konkluderes med at produktive diskusjoner kjennetegnes av at minst en person har korrekt svar inn i diskusjonen. I større grupper kan det være en fordel at minst to stykker er enige om korrekt svar. I diskusjoner med god utvikling er det ønskelig at gruppa gjennom dialog blir enig om korrekt svar, og ikke gjennom at en person overbeviser de andre. Dersom gruppa er ubalansert med hensyn til inngangsnivå fra FCI førtest, kan det være fare for det siste.

6.2 Implikasjoner

Innføring av digital Peer Instruction i det fysikkemnet denne studien tar for seg har vist seg å gi gode resultater ved første forsøk. Den digitale gjennomføringen har tilsynelatende ikke hatt negative effekter for studentenes konseptuelle forståelse ved endt undervisning. Det vil på bakgrunn av dette anbefales å utforske aktive tilnærminger til fysikkundervisning, også ved digital gjennomføring.

Vi har funnet at dersom Peer Instruction gjennomføres ved bruk av faste grupper, så kan gruppesammensetning spille en rolle for utfallet av, og kvaliteten på diskusjonene. Det bør på større grupper, helst være minst to personer som er enige om korrekt svar inn i diskusjonen for at flertallet skal ha korrekt etter diskusjon. For å sørge for at en eller to på gruppa svarer korrekt før diskusjon kan score på FCI førtest være en indikator for sannsynligheten for dette.

6.3 Videre arbeid

Fra denne studien har det kommet fram flere aspekter som ville vært interessante å undersøke i en utvidelse av studien, eller i kommende studier innenfor samme område. For det første ville det vært fordelaktig å benytte flere verktøy enn data fra konseptoppgavene for å si noe kvantitativt om gruppediskusjonene i Peer Instruction. Det hadde vært interessant å undersøke hvordan faktorer som for eksempel motivasjon, tidsbruk og interesse for faget, spilte inn på det overordnede læringsutbyttet såvel som utbyttet fra gruppediskusjonene. En sammenligning av læringsutbytte som målt ved FCI, og eksamensresultater kunne også vært interessant i en evaluering av det aktuelle emnet.

I denne studien har vi tatt for oss svarene til studentene på konseptoppgavene i undervisning. dersom disse hadde vært implementert slik at studentene også oppga hvor trygge de var på sitt svar ville det gitt sikrere analyser og større innsikt i hvorvidt gruppediskusjoner er produktive. Her ble kun et fåtall av diskusjonene fra undervisningen tatt med i den kvalitative analysen. Det framstår som svært interessant å kunne gjøre en større kvalitativ analyse av en større mengde diskusjoner i en slik undervisningssetting, for å få et enda klarere bilde på hva som gjør en diskusjon produktiv.

Det foreligger konsensus om at Peer Instruction fungerer godt og at aktive undervisningsformer er å foretrekke foran konvensjonell undervisning i fysikkfagene. Det er imidlertid gjort lite forskning for å finne ut av hva som skjer i selve gruppediskusjonene i Peer Instruction, og hva som kjennetegner en diskusjon av høy kvalitet. Fra dette fremkommer et behov for å kartlegge og analysere diskusjoner i Peer Instruction i enda større skala enn det som er blitt gjort i denne oppgaven. Dette kan gi svært nyttig innsikt til undervisere som bruker aktive tilnærminger til undervisning i dag og i fremtiden.

Bibliografi

- Allaj, E. (2018). Two simple measures of variability for categorical data. *Journal of Applied Statistics*, 45(8), 1497–1516. doi:10.1080/02664763.2017.1380787
- Angell, C., Bungum, B., Henriksen, E. K., Kolstø, S. D., Persson, J. & Renstrøm, R. (2019). *Fysikkdidaktikk* (2. utg.). Oslo: CAPPELEN DAMM AS.
- Ausubel, D. P., Novak, J. D. & Hanesian, H. (1978). *Educational psychology: a cognitive view*. New York: Holt, Rinehart og Winston.
- Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, 100(3), 603–617. doi:https://doi.org/10.1348/000712608X377117
- Bakhtin, M. M. & Holquist, M. (1981). *The dialogic imagination: four essays*. Austin: University of Texas Press.
- Barke, H.-D., Hazari, A. & Yitbarek, S. (2009). Students' misconceptions and how to overcome them. I *Misconceptions in Chemistry* (s. 21–36). Springer.
- Behi, R. & Nolan, M. (1996). Causality and control: threats to internal validity. *British Journal of Nursing*, 5(6), 374–377. doi:10.12968/bjon.1996.5.6.374
- Bergmann, J. & Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*. International Society for Technology in Education.
- Bing, T. J. & Redish, E. F. (2009). Analyzing problem solving using math in physics: Epistemological framing via warrants. *Physical Review Special Topics - Physics Education Research*, 5(2), 020108. doi:10.1103/PhysRevSTPER.5.020108
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. D. McKay.
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. doi:10.1191/1478088706qpp0630a
- Cashman, E. & Eschenbach, E. (2003). Active learning with Web technology - just in time. doi:10.1109/FIE.2003.1263352
- Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science*, 1(1), 73–105. doi:https://doi.org/10.1111/j.1756-8765.2008.01005.x
- Coe, R. (2002). It's the effect size, stupid: what effect size is and why it is important, University of Exeter, England. Hentet 21. mai 2021, fra <http://www.leeds.ac.uk/educol/documents/00002182.htm>

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. utg.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Creswell, J. W. & Plano Clark, V. L. (2017). Choosing a Mixed Methods Design. I *Designing and conducting mixed methods research* (3. utg., s. 58–89). SAGE Publications.
- Crouch, C. H. & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. doi:10.1119/1.1374249
- Davis, P. R., Horn, M. S. & Sherin, B. L. (2013). The Right Kind of Wrong: A “Knowledge in Pieces” Approach to Science Learning in Museums. *Curator: The Museum Journal*, 56(1), 31–46. doi:https://doi.org/10.1111/cura.12005
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2/3), 105–225. doi:10.1080/07370008.1985.9649008
- diSessa, A. A. (2018). A Friendly Introduction to “Knowledge in Pieces”: Modeling Types of Knowledge and Their Roles in Learning. I G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt & B. Xu (Red.), *Invited Lectures from the 13th International Congress on Mathematical Education* (s. 65–84). doi:10.1007/978-3-319-72170-5_5
- diSessa, A. A. & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20(10), 1155–1191. doi:10.1080/0950069980201002
- Fagen, A. P., Crouch, C. H. & Mazur, E. (2002). Peer Instruction: Results from a Range of Classrooms. *The Physics Teacher*, 40(4), 206–209. doi:10.1119/1.1474140
- Felder, R. M. & Brent, R. (1996). Navigating the Bumpy Road to Student-Centered Instruction. *College Teaching*, 44(2), 43–47. doi:10.1080/87567555.1996.9933425
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. doi:10.1073/pnas.1319030111
- Fritz, C. O. (2011). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2. doi:10.1037/a0024338
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. doi:10.1119/1.18809
- Halloun, I. A. & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, 53(11), 1056–1065. doi:10.1119/1.14031
- Halloun, I. A. & Hestenes, D. (1985b). The initial knowledge state of college physics students. *American Journal of Physics*, 53(11), 1043–1055. doi:10.1119/1.14030
- Halloun, I. A. & Hestenes, D. (1987). Modeling instruction in mechanics. *American Journal of Physics*, 55(5), 455–462. doi:10.1119/1.15130

- Hammer, D., Elby, A., Scherr, R. E. & Redish, E. F. (2005). Resources, framing and transfer. I J. P. Mestre (Red.), *Transfer of Learning from a Modern Multidisciplinary Perspective* (s. 89–119). Greenwich, CT: IAP.
- Hammer, D. (2004). The variability of student reasoning, lecture 3: manifold cognitive resources. I *PROCEEDINGS-INTERNATIONAL SCHOOL OF PHYSICS ENRICO FERMI* (Bd. 156, s. 321–340). IOS Press; Ohmsha; 1999.
- Hannisdal, M., Haugan, J. & Munkvik, M. (2007). *Eureka!: Naturfag for ungdomstrinnet: Grunnbok 9* (1. utg.). Gyldendal Norsk Forlag AS.
- Harlow, D. B. & Bianchini, J. A. (2020). Knowledge-in-Pieces—Andrea A. diSessa, David Hammer. I B. Akpan & T. J. Kennedy (Red.), *Science Education in Theory and Practice: An Introductory Guide to Learning Theory* (s. 389–401). Cham: Springer International Publishing. Hentet 17. mai 2021, fra https://doi.org/10.1007/978-3-030-43620-9_26
- Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, 55(5), 440–454. doi:10.1119/1.15129
- Hestenes, D. & Halloun, I. (1995). Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller. *The Physics Teacher*, 33(8), 502. Publisher: American Association of Physics TeachersAAPT. doi:10.1119/1.2344278
- Hestenes, D. & Wells, M. (1992). A mechanics baseline test. *The Physics Teacher*, 30(3), 159–166. doi:10.1119/1.2343498
- Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30(3), 141–158. doi:10.1119/1.2343497
- Huffman, D. & Heller, P. (1995). What does the force concept inventory actually measure? *The Physics Teacher*, 33(3), 138–143. doi:10.1119/1.2344171
- James, M. C. (2006). The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, 74(8), 689–691. doi:10.1119/1.2198887
- Johnson, D. W., Johnson, R. T., Haugaløkken, O. K. & Aakervik, A. O. (2014). *Samarbeid i skolen: pedagogisk utviklingsarbeid, samspill mellom mennesker* (4. utg.). Oslo: David W. Johnson og Pedagogisk Psykologisk Forlag AS33.
- Kader, G. D. & Perry, M. (2007). Variability for Categorical Variables. *Journal of Statistics Education*, 15(2), null. doi:10.1080/10691898.2007.11889465
- Kelley, K. (2012). On effect size. *Psychological Methods*, 17(2), 137. doi:10.1037/a0028086
- Knight, R. D. (2004). *Five Easy Lessons: Strategies for Successful Physics Teaching*. San Francisco, CA: Addison Wesley.
- Lasry, N., Rosenfield, S., Dedic, H., Dahan, A. & Reshef, O. (2011). The puzzling reliability of the Force Concept Inventory. *American Journal of Physics*, 79(9), 909–912. doi:10.1119/1.3602073
- Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice Hall.
- McCloskey, M. (1983). Naïve theories of motion. I D. Gentner & A. Stevens (Red.), *Mental Models* (s. 299–324).

- Mercer, N. & Littleton, K. (2007). *Dialogue and the Development of Children's Thinking: A Sociocultural Approach*. Routledge.
- Mercier, H. & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(2), 57–74, discussion 74–111. doi:10.1017/S0140525X10000968
- Mortimer, E. & Scott, P. (2003). *Meaning Making In Secondary Science Classrooms*. McGraw-Hill Education (UK).
- Netland, K. Ø., Sivertsen, A. & Olufsen, M. (2018). Innføring av studentaktive arbeidsformer i seminarundervisningen. Hvilken betydning har dette på læringsutbyttet og klasse miljøet? *Nordic Journal of STEM Education*, 2(1), 1–15. doi:10.5324/njsteme.v2i1.2346
- Novak, G. M. & Patterson, E. T. (1997). World Wide Web Technology as a New Teaching and Learning Environment. *International Journal of Modern Physics C*, 08(01), 19–39. doi:10.1142/S0129183197000047
- Osborne, J., Borko, H., Fishman, E., Busch, K., S, M. & A, T. (2016). *Developing and using an instrument to assess the dialectical potential of whole-class discussions in teaching science*.
- Porter, L., Bailey Lee, C., Simon, B., Cutts, Q. & Zingaro, D. (2011). Experience report: a multi-classroom report on the value of peer instruction. I *Proceedings of the 16th annual joint conference on Innovation and technology in computer science education* (s. 138–142). doi:10.1145/1999747.1999788
- Quale, A. (2007). Konstruktivisme i naturvitenskapen: kunnskapssyn og didaktikk. *Nordic Studies in Science Education*, 3(2), 175–188. Number: 2. doi:10.5617/nordina.379
- Renkl, A., Atkinson, R. K., Maier, U. H. & Staley, R. (2002). From Example Study to Problem Solving: Smooth Transitions Help Learning. *The Journal of Experimental Education*, 70(4), 293–315. doi:10.1080/00220970209599510
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
- Sjøberg, S. (2014). *Naturfag som allmenndannelse: en kritisk fagdidaktikk* (3. utg.). Oslo: Gyldendal Norsk Forlag AS.
- Solomon, J. (1994). The Rise and Fall of Constructivism. *Studies in Science Education*, 23(1), 1–19. doi:10.1080/03057269408560027
- Stojanovska, M., Petrusovski, V., Köller, H. G. & Karlsen, S. (2015). Students' alternative conceptions and ways to overcome them. *A guidebook of good practice for the pre-service training of chemistry teachers*. Krakow: Jagiellonian University}.
- Taber, K. (2002). *Chemical misconceptions: prevention, diagnosis and cure*. Royal Society of Chemistry.
- Toulmin, S. E. (2003). *The uses of argument*. Cambridge university press.
- Utdanningsdirektoratet. (2006). Læreplan i fysikk - programfag i utdanningsprogram for studiespesialisering (FYS1-01). Hentet 10.06.2021. Hentet fra <https://www.udir.no/klo6/FYS1-01#>

- Van Heuvelen, A. (1991). Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics*, 59(10), 891–897. doi:10.1119/1.16667
- Venville, G. J. & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, 47(8), 952–977. doi:https://doi.org/10.1002/tea.20358
- Von Aufschnaiter, C., Erduran, S., Osborne, J. & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 45(1), 101–131.
- Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K. (2007). *Probability & statistics for engineers & scientists* (8. utg.). Pearson Education, Inc.
- Wood, A. K., Galloway, R. K., Hardy, J. & Sinclair, C. M. (2014). Analyzing learning during Peer Instruction dialogues: A resource activation framework. *Physical Review Special Topics - Physics Education Research*, 10(2), 020107. doi:10.1103/PhysRevSTPER.10.020107

Vedlegg

Appendiks A: transkripsjoner fra utvalgte analyseenheter

Dette appendikset inneholder transkripsjoner fra de to analyseenhetene som ble undersøkt nærmere i studien. Deltakere i samtalen presenteres først. Deltakerne er anonymisert ved å hente randomiserte navn fra denne tjenesten: <https://www.behindthename.com/random/>. Deretter følger selve oppgaven som diskuteres, og dernest transkripsjonen fra samtale.

Analyseenhet 3

Deltakere: Gawain, Mukesh, Valdemar, Indra,

Oppgave:

1.

Person sitter i en robåt og har med seg en stor stein. Midt utpå innsjøen kaster personen steinen overbord. Vil da vannstanden i innsjøen

A. øke

B. synke

C. forbli uendret

Figur A.1: Oppgave tilknyttet analyseenhet 3.

Transkripsjon:

- 1 Gawain: Det problemet her er egentlig veldig lett, så, bare kanskje ikke så veldig intuitivt. Ehhm...
- 2 Valdemar: [uhørbart]
- 3 Gawain: Altså hvis, hvis steinen er i båten, så vil båten synke ned, og det vil skyve vann ut til siden og oppover.
- 4 Valdemar: Mhm.
- 5 Gawain: Og da vil jo vannstanden være økt tilsvarende så mye som hvis det båten er under vann var fylt med vann, men det er det

jo ikke. Det er fylt med en bitteliten — eller ikke bitteliten, en stor — men det er fylt med en stein, så da når du, så da, volumet av båten som er under vann vil være tilsvarende vekta, ehmm, eller være tilsvarende av volum som tilsvarer samme vekten av vann som det som er i båten.

- 6 Valdemar: Mhm.
- 7 Gawain: Så hvis du da kaster ut steinen så vil steinen ha veldig mye mindre volum enn det volumet av båten hadde før, for båten vil da flyte høyere opp. Beste måten å se for seg er å se for seg at steinen er veldig tung og båten er veldig lett. Så vi kan si at båten flyter bare oppå vannet, den øker ikke vannivået i det hele tatt når den ikke har en stein i, fordi den er så lett. Også hvis vi legger steinen nedi, så vil da volumet øke fordi båten blir dyttet langt ned. Også hvis vi hiver ut steinen, så vil steinen nesten ikke øke volumet i det hele tatt, mens båten vil bli dyttet opp av vannet som gjør at vannet senker seg.
- 8 Indra: Ja, det var en god forklaring.
- 9 Valdemar: Så svaret mener du er...
- 10 Gawain: Vannet minker.
- 11 Valdemar: Så pass ja.
- 12 Indra: Jeg tok at den ikke, at den vannstanden var uendret, men jeg er med på det nå [uhørbart]. Hvis man ser for seg at det er en — som du sier — en liten stein som har uendelig stor masse, så vil jo den fortrenge masse, masse, masse vann, men når du hiver den uti så vil det ikke påføre vannstanden noe økning
- 13 [Stillhet]
- 14 Gawain: Skal vi se. Okei da har vi fire minutter til
- 15 Valdemar: Bare å chille
- 16 Mukesh: Skulle vi ha en så lang diskusjon?
- 17 Gawain: Ja, det er vel sånn i tilfelle man ikke kan svaret fra før da. Jeg [uhørbart].
- 18 Mukesh: Er det ikke vanligvis sånn tre minutter vi får?
- 19 Gawain: Jo, men nå sa han åtte minutter.
- 20 Mukesh: Åja
- 21 [Stillhet]
- 22 [Gruppa snakker om andre ting]

Analyseenhet 11

Deltakere: Gawain, Mukesh, Valdemar, Indra, Nazarius,

Oppgave:

4.

Vi har to kilo kokende kokende vann i en isolert beholder. Vi slipper så en tokilos kobberblokk oppi beholderen. Kobberet tar vi rett fra fryseren så det holder 0 °C. Hvilket utsagn er sann om sluttemperaturen til vannet?

- A. Temperaturen vil ende på 50 °C, altså midt mellom de to temperaturene, siden begge har like stor masse.
- B. Temperaturen vil ende nært 0 °C siden kobberet leder varmen vekk så godt.
- C. Temperaturen vil ende nær 100 °C siden kobberet har langt lavere spesifikk varmekapasitet enn vannet.
- D. Ingen av utsagnene over er riktige.

Figur A.2: Oppgave tilknyttet analyseenhhet 11.

Transkripsjon:

- 1 Valdemar: Jeg bare gønn på med 50 grader.
- 2 Nazarius: Same
- 3 Indra: Hva sa dere? Jeg tok C, hva tok dere?
- 4 Valdemar: 50.
- 5 Indra: 50, ja jeg lurte på om det er det som er riktig.
- 6 Gawain: Jeg tok ingen av dem, for jeg har gjort noen eksperimenter på å dette her på videregående og hvis jeg husker riktig i alle fall, så er kobber, ehh, sin spesifikke varmekapasitet er på ca 600 J/g, mens vann er 4200, så kobber er da ca. en syvendedel, hvis jeg husker riktig, av vann.
- 7 Indra: Ja
- 8 Valdemar: Hvor vil det ende opp da?
- 9 Gawain: Så da...
- 10 Mukesh: Så da [uhørbart] temperatur?
- 11 Indra: Ja, det var det jeg på en måte tenkte at...
- 12 [Folk snakker i munnen på hverandre]
- 13 Indra: Hva du sa, at?
- 14 Nazarius: Hvis de ligger lenge nok, så vil temperaturen bli lik.
- 15 Indra: Det vil den.
- 16 Nazarius: Men hvor, med hvilken temperatur den slutter på, er det som er usikkert.
- 17 Indra: Det vil...
- 18 Gawain: Det jeg tenker er at den er i nærheten av hundre, fordi kobber har ganske lav spesifikk varmekapasitet, men liksom hvor, hvor nærme de definerer som nærme her, det vet vi jo ikke. Så, det er liksom ikke, det er liksom ikke sånn innenfor to tre grader. Det kommer til å være...
- 19 Nazarius: Mellom åtti og nitti
- 20 Gawain: Sannsynligvis mer enn ti grader forskjell, så...
- 21 [Uhørbart]
- 22 Gawain: ...nitti grader
- 23 Indra: Men det er ikke femti da, hvis det er ulik varmekapasitet?
- 24 Gawain: Hvis, ja, hvis jeg husker riktig på kobber så...
- 25 Valdemar: Det virker det som du gjør, generelt.
- 26 Gawain: Jeg kan jo google det.
- 27 Valdemar: Ehm, så da, men da vil jeg tro at nær hundre virker som et logisk svar.
- 28 Indra: Mhm.
- 29 Gawain: Ja.
- 30 Valdemar: Ikke femti i hvert fall.

- 31 Nazarius: Nope.
- 32 Gawain: Her står det 380, altså nærmere 400. Så, ja da er den nok en del lavere enn det, altså da er det vel mye nærmere hundre, vil jeg tro.
- 33 Indra: Ja.
- 34 Gawain: Jeg husket feil, men...
- 35 Nazarius: Altså, går det an å tenke seg at, kobber vil jo ta opp et x antall ehm mengde med energi, men den energien, men, vil da øke med varme fra da null til et visst antall grader. Men den energien som blir gitt fra vannet til kobberet vil være betydelig mindre enn den energien som, ehh, vil være betydelig mindre enn det det trengs for å senke temperaturen til vann en grad enn det du trenger for å øke temperaturen til kobberet en grad. For kobberet vil jo øke temperaturen mye fortere enn vann synker i temperatur.
- 36 Indra: Mhm.
- 37 [Folk snakker over hverandre]
- 38 Indra: [Uhørbart] delta T , så blir jo sluttverdien på T , det blir liksom likevektspunktet, mens den ene starter fra null og den andre starter fra hundre, men c , m -delen vil, eller c vil være ulik på begge sider av ligninga.
- 39 Nazarius: Mhm.
- 40 Valdemar: Ja.
- 41 Gawain: Ja, når jeg googlet det nå så ser jeg at kobber har mindre enn en tiendedel av varmekapasiteten til vann.
- 42 Nazarius: Så det vil være nærmere hundre grader.
- 43 Valdemar: Ja, nå må vi gå tilbake igjen.
- 44 Nazarius: Ja.
- 45 Gawain: Så det er mindre enn ti grader forskjell, eller mindre enn ti grader mindre enn hundre, så over nitti.
- 46 Nazarius: Ja, nærmere nitti.

