



**UiT** Norges arktiske universitet

Fakultet for naturvitenskap og teknologi  
Institutt for fysikk og teknologi

## **Engasjement, motivasjon og læringsutbytte i et digitalt Peer Instruction-basert introduksjonsemne i fysikk**

Sigurd Eriksen Lien

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



“I know words.  
I have the best words.”  
–Donald J. Trump

“I have an idea so smart that my head would explode if I even began to know  
what I was talking about.”  
–Peter Griffin

“My attention span is very short.”  
–will.i.am

“People who think they know everything are a great annoyance to those of us  
who do.”  
–Isaac Asimov

“It is finished.”  
–Jesus Christ



# Sammendrag

Året 2020 har vært et år med drastiske endringer i undervisningsform med en hurtig overgang til digital undervisning grunnet koronapandemien. Denne studien undersøker studentenes motivasjon, læringsutbytte, deltagelse og tidsbruk ved et introduksjonsemne i fysikk, der undervisningen ble lagt om til digital undervisning med Peer Instruction.

Studentens motivasjon ble karakterisert med motivasjonskonstruktene fra Eccles expectancy-value modell for motivasjon, målt med *Interactive Engagement and Motivation in Physics Learning* (IMPEL). Det fremkommer at *interesse og glede, nytteverdi og selvrealisering og identitet* er betydelig bedre enn *mestringsforventning og kostnad*. Sammenlignet med motivasjonen til studentene som tok emnet i 2019 – et år med tradisjonell undervisning – viste en *t*-test ingen statistisk signifikant endring ved et signifikansnivå på 0.05. Læringsutbytte målt med Force Concept Inventory (FCI) viste et normalisert gain på 0.46, som er sammenlignbart med Peer Instruction i fysisk undervisning. Sammenhengen mellom studentenes læringsutbytte og faktorene motivasjon, deltagelse og tidsbruk ble undersøkt og det ble funnet noen interessante, men ikke statistisk signifikante, sammenhenger med lineær regresjon og et 0.05 signifikansnivå.

Basert på resultatene konkluderer vi med at digital undervisning med Peer Instruction tilsynelatende har fungert så godt som vi kunne forventet med fysisk undervisning, både med hensyn på læringsutbytte og motivasjon.



# Takksigelser

Først vil jeg rette en takk til de to veilederne mine: David Andre Coucheron og Børge Irgens. Takk for alle gode råd, innspill og tilbakemeldinger. Uten deres faglige dybde hadde ikke denne oppgaven blitt så bra som den ble. Det rettes også en takk til studentene som frivillig har deltatt i studien.

Jeg vil også rette en takk til alle mine medstudenter på studiet. En spesiell takk rettes til min gode venn og medstudent Øystein Grøntveit Winnberg. Takk for alle fine samtaler om faglige og utenomfaglige ting, både i løpet av innspurten på masterskriving, og gjennom et femårig studieløp.

Til slutt vil jeg takke min forlovede Victoria. Takk for ferdig middag hver kveld, og at jeg slapp å ta buss alle de sene kveldene i innspurten av masterskrivingen.





# Innhold

<b>Sammendrag</b>	<b>iii</b>
<b>Takksigelser</b>	<b>v</b>
<b>Figurer</b>	<b>xi</b>
<b>Tabeller</b>	<b>xiii</b>
<b>1 Introduksjon</b>	<b>1</b>
1.1 Bakgrunn for oppgaven . . . . .	1
1.2 Forskningsspørsmål . . . . .	2
1.3 Forutsetninger for studien . . . . .	3
<b>2 Teori</b>	<b>5</b>
2.1 Læringsteoretisk bakteppe . . . . .	5
2.1.1 Knowledge in Pieces . . . . .	6
2.1.2 Nyansering . . . . .	8
2.2 Peer Instruction . . . . .	10
2.2.1 Problemer med dagens fysikkundervisning . . . . .	10
2.2.2 Aktiv læring og Peer Instruction . . . . .	11
2.2.3 Hvordan undervisning drives i Peer Instruction . . . . .	12
2.2.4 Forutsetninger for at Peer Instruction skal fungere . . . . .	14
2.2.5 Peer Instruction og konstruktivisme . . . . .	15
2.3 Hverdagsforestillinger . . . . .	15
2.3.1 Hva er egentlig hverdagsforestillinger? . . . . .	16
2.3.2 Hverdagsforestillingerers opphav . . . . .	17
2.3.3 Hvordan avdekke elevenes hverdagsforestillinger . . . . .	18
2.3.4 Hvordan utfordre elevenes hverdagsforestillinger . . . . .	19
2.3.5 Diagnostisk test . . . . .	21
2.4 Force Concept Inventory (FCI) . . . . .	21
2.4.1 Hva er Force Concept Inventory? . . . . .	21
2.4.2 Validitet i FCI . . . . .	22
2.4.3 Reliabilitet i FCI . . . . .	23
2.4.4 Bruksområder for FCI . . . . .	24

2.5	Motivasjon og mål på motivasjon . . . . .	24
2.5.1	Motivasjonsteori . . . . .	25
2.5.2	Interactive Engagement and Motivation in Physics Learning (IMPEL) . . . . .	29
2.5.3	Reliabilitet i IMPEL . . . . .	30
2.5.4	Validitet i IMPEL . . . . .	33
2.5.5	Peer Instruction og motivasjon . . . . .	34
2.6	Relevant statistikkteori . . . . .	34
2.6.1	Effektstørrelser . . . . .	35
2.6.2	Lineær regresjon . . . . .	37
2.6.3	t-test . . . . .	39
<b>3</b>	<b>Metode</b>	<b>41</b>
3.1	Utvalg . . . . .	41
3.2	Gjennomføring av emnet . . . . .	45
3.2.1	Undervisningsform . . . . .	45
3.2.2	Fellesundervisning . . . . .	45
3.2.3	Eksamen . . . . .	46
3.3	Instrumenter . . . . .	46
3.3.1	Force Concept Inventory . . . . .	46
3.3.2	IMPEL . . . . .	47
3.4	Prosedyre . . . . .	48
3.5	Analysar . . . . .	49
3.5.1	FCI . . . . .	50
3.5.2	IMPEL . . . . .	50
3.5.3	FCI og motivasjonskonstrukter . . . . .	51
3.5.4	FCI og deltagelse . . . . .	51
3.5.5	FCI og tidsbruk . . . . .	52
3.6	Etiske betraktninger . . . . .	52
<b>4</b>	<b>Resultater</b>	<b>53</b>
4.1	Fysikkstudentenes motivasjon . . . . .	53
4.1.1	Resultater fra reliabilitetsanalyse . . . . .	53
4.1.2	Fysikkstudentenes motivasjon i 2020 . . . . .	54
4.2	Sammenligning av motivasjon i 2019 og 2020 . . . . .	56
4.2.1	Fysikkstudentenes motivasjon i 2019 . . . . .	56
4.2.2	Motivasjon i 2019 og 2020 . . . . .	58
4.3	Sammenligning av læringsutbytte med motivasjon, tidsbruk og deltagelse . . . . .	59
4.3.1	Læringsutbytte . . . . .	60
4.3.2	Motivasjon og læringsutbytte . . . . .	61
4.3.3	Tidsbruk og læringsutbytte . . . . .	64
4.3.4	Deltagelse og læringsutbytte . . . . .	66

<b>5</b>	<b>Diskusjon</b>	<b>69</b>
5.1	Kjennetegn på fysikkstudentenes motivasjon i 2020 . . . . .	70
5.1.1	Reliabiliteten i IMPEL . . . . .	70
5.1.2	Studentenes motivasjon . . . . .	71
5.2	Motivasjon etter ett år med koronapandemi . . . . .	72
5.2.1	Motivasjon i 2020 sammenlignet med motivasjon i 2019	73
5.3	Sammenheng mellom motivasjon, tidsbruk og deltagelse, og konseptuell forståelse etter endt undervisning . . . . .	74
5.3.1	FCI som evaluering av undervisningen . . . . .	74
5.3.2	Motivasjon og konseptuelt læringsutbytte . . . . .	76
5.3.3	Deltagelse og konseptuelt læringsutbytte . . . . .	78
5.3.4	Tidsbruk og konseptuelt læringsutbytte . . . . .	79
5.4	Studiens validitet . . . . .	80
5.4.1	Begrepsvaliditet . . . . .	80
5.4.2	Indre validitet . . . . .	81
5.4.3	Ytre validitet . . . . .	82
<b>6</b>	<b>Avslutning</b>	<b>85</b>
6.1	Konklusjon . . . . .	85
6.2	Implikasjoner for undervisning . . . . .	86
6.3	Videre forskning . . . . .	87
	<b>Vedlegg</b>	<b>97</b>
A.1	Statistiske verdier fra IMPEL – utvalgsgruppen er alle som tok IMPEL . . . . .	97
A.2	Statistiske verdier fra IMPEL – utvalgsgruppen er de som både tok IMPEL og FCI . . . . .	103
A.3	Verdier for skjevhet og kurtose for fordelingen av studentnes gjennomsnitt på hvert konstrukt . . . . .	109



# Figurer

3.1	Venn diagram for utvalget. . . . .	42
3.2	Histogramplot av karakterene i til utvalgsgruppene i fysikk 1 og 2. . . . .	44
3.3	Eksempeloppgave fra FCI. . . . .	47
4.1	Snittsvar IMPEL 2020. . . . .	55
4.2	Snittsvar IMPEL 2019. . . . .	57
4.3	Spredningsplot av motivasjonskonstrukter og deltagelse. . .	59
4.4	Histogramplot som viser fordelingen av andel korrekte svar i FCI pre- og post-test, og spredningsplot for FCI pre- og post-test scores for hver student. . . . .	60
4.5	Spredningsplot for FCI-gain og motivasjonskonstrukter, samt resultater fra lineær regresjon. . . . .	62
4.6	Spredningsplot av FCI post-test score og motivasjonskonstrukter. . . . .	63
4.7	Spredningsplot av FCI-gain og tidabruk. . . . .	65
4.8	Spredningsplot av FCI post-test score og tidsbruk. . . . .	66
4.9	Spredningsplot av FCI-gain og deltagelse, og spredningsplot av FCI post-test score og deltagelse. . . . .	67



# Tabeller

2.1	Nivåer av effekt for Cohen's d. . . . .	37
3.1	Beskrivelse av utvalget. . . . .	43
3.2	Konverterte svaralternativ for IMPEL. . . . .	48
4.1	Resultat fra t-test mellom motivasjonskonstrukt i 2019 og 2020. . . . .	58
4.2	Effektstørrelser for FCI. . . . .	60
4.3	Multippel regresjon på motivasjonskonstrukt og FCI-gain. . . . .	62
4.4	Resultatene fra multippel lineær regresjon på alle motivasjonskonstruktene og post-score fra FCI. . . . .	64
A.1	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «interesse og glede» for alle som tok IMPEL. . . . .	98
A.2	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «nytteverdi» for alle som tok IMPEL. . . . .	99
A.3	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «mestringsforventning» for alle som tok IMPEL. . . . .	100
A.4	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «selvrealisering og identitet» for alle som tok IMPEL. . . . .	101
A.5	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «kostnad» for alle som tok IMPEL. . . . .	102
A.6	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «interesse og glede» for gruppen som tok både FCI og IMPEL. . . . .	104
A.7	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «nytteverdi» for gruppen som tok både FCI og IMPEL. . . . .	105
A.8	Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «mestringsforventning» for gruppen som tok både FCI og IMPEL. . . . .	106

A.9 Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «selvrealisering og identitet» for gruppen som tok både FCI og IMPEL. . . . .	107
A.10 Fullstendige statistiske verdier for indikatorene som gjør opp indeksen som måler «kostnad» for gruppen som tok både FCI og IMPEL. . . . .	108
A.11 Skjevhet og kurtose for fordelingen av gjennomsnittssvar i 2020. . . . .	109
A.12 Skjevhet og kurtose for fordelingen av gjennomsnittssvar i 2019. . . . .	109





# Introduksjon

## 1.1 Bakgrunn for oppgaven

Alle mennesker har en iboende motivasjon. Det er den som bestemmer hva en velger å engasjere seg i (Holt mfl., 2015). En læringsinstitusjon må ta motivasjonen til de lærende i betraktning for å lykkes med å skape innsatsvilje og engasjement, og legge til rette for at de lærende skal få et godt læringsutbytte (Lillemyr, 2007). Motivasjonsteorier har som mål å forklare motivasjon, og så-dan legge til rette for at læringsinstitusjonen skal kunne forstå seg på hva som motiverer den lærende. Denne oppgaven i hovedsak vil søke å ta utgangspunkt i Eccles og Wigfield (2002) sin modell for motivasjon. Teorien er en såkalt «Expectancy-Value theory» som er en todelt teori for motivasjon (Ramberg, 2006). Den ene siden av teorien skal forklare hvor godt de lærende forventer å prestere i oppgaven de engasjerer seg i, mens den andre siden søker å forklare hvorfor de velger å engasjere seg i denne oppgaven (Eccles & Wigfield, 2002).

Kunnskapsdepartementet (2021) skriver at studentene som følge av koronapandemien har hatt nedsatt motivasjon for studiene sine. Dette ble målt ved Studiebarometeret, en studie fremsatt av Kunnskapsdepartementet, og gjennomført av NOKUT høsten 2020, samtidig som emnet denne studien fokuserer på ble gjennomført (Kunnskapsdepartementet, 2021; NOKUT, 2021). I tillegg rapporteres det at studentene føler de har hatt redusert læringsutbytte, selv om de har brukt like mye tid på studiene som studenter i tidligere år har brukt.

Gjennom studier har man over lengre tid observert at konvensjonell under-

visning i fysikk ikke legger til rette for at elever eller studenter skal forstå de underleggende konseptene, og tilegne seg forståelse av fysikk – tross adekvate resultater på konvensjonelle, algoritmiske problemløsningsoppgaver (Hake, 1998; Hestenes, Wells & Swackhamer, 1992; Mazur, 1997). Som en respons til dette har en søkt etter undervisningsmetoder som gir bedring i konseptuell forståelse, samtidig som en ikke undertrykker evnen til klassisk problemløsning (Hake, 1998; Mazur, 1997). En slik undervisningsmetode er Peer Instruction, en studentaktiv undervisningsmetode utviklet av Eric Mazur for undervisning av introduksjonsemne i fysikk ved høyere utdanning (Mazur, 1997). Det har blitt utført en rekke studier på effekten av Peer Instruction, og resultatene er lovende (Crouch & Mazur, 2001; Fagen, Crouch & Mazur, 2002; Hake, 1998; Mazur, 1997).

Universitetet i Tromsø – Norges arktiske universitet har i sin strategi mot 2022<sup>1</sup> skrevet at «UiT skal øke bruken av læringsfremmende teknologi og implementere digitale eksamensformer ved alle studier der det er hensiktsmessig» og «UiT skal legge til rette for fleksible og nettstøttede undervisningsformer». I tillegg skriver de at «UiT skal utvikle og ta i bruk nye pedagogiske virkemidler og studentaktive læringsformer». Dette er en del av et skifte mot mer fleksible og digitaliserte utdanninger og vurderinger. Samtidig som det digitale skiftet, legges det også opp til at en skal bruke mere studentaktive undervisningsformer som Peer Instruction. De fleste studiene gjort på Peer Instruction og andre studentaktive undervisningsformer vektlegger, naturlig nok, fysisk klasseromsundervisning. Da koronapandemien traff Norge for fullt i 2020 ble undervisningsinstitusjonene tvunget til en hurtig omlegging til digital undervisning. Omleggingen har nok blitt gjort på mange måter, men i denne oppgaven har det vært et ønske om å se på hvilke implikasjoner digital undervisning har hatt for en studentaktiv undervisningsform.

## 1.2 Forskningsspørsmål

Ifølge Knight (2004) er Peer Instruction en studentaktiv undervisningsform som er gunstig for store studentgrupper i introduksjonemne i fysikk. I lys av resultatene fra Studiebarometeres som presentert av Kunnskapsdepartementet (2021) har jeg gjennom denne oppgaven ønsket å utforske noen aspekter ved digital undervisning i Peer Instruction. Dette inkluderer en kartlegging av studentenes motivasjon i 2020, og hvordan denne motivasjonen er sammenlignet med motivasjonen til studentene fra 2019. Jeg har også ønsket å se hvilke implikasjoner den digitale undervisningen har hatt for det konseptuelle læringsutbyttet til studentene. I tillegg til dette har det vært ønskelig å utfors-

1. [https://uit.no/om/art?dim=179033&p\\_document\\_id=355830](https://uit.no/om/art?dim=179033&p_document_id=355830)

ke sammenhenger mellom konseptuell forståelse og motivasjon, konseptuell forståelse og deltagelse i undervisningen, og konseptuell forståelse og tidsbruk i emnet. For å kunne svare på dette har følgende forskningsspørsmål blitt formulert.

1. Hva kjennetegner fysikkstudentenes motivasjon?
2. Hvordan er motivasjonen til studentene som tok FYS-0100 i 2020, sammenlignet med studentene som tok tilsvarende emne i 2019?
3. Hva er sammenhengen mellom konseptuelt læringsutbytte, motivasjon, tidsbruk, og deltagelse etter endt digital undervisning i et introduksjons-emne i fysikk?

### 1.3 Forutsetninger for studien

Studien, med tilhørende datainnsamling, har blitt gjennomført på emnet FYS-0100 generell fysikk ved UiT – Norges arktiske universitet, høsten 2020. Fellesundervisningen har blitt gjennomført ved bruk av Peer Instruction, modifisert for digital undervisning. Hvilke modifikasjoner som har blitt gjort vil vi komme tilbake til i seksjon 3.2. I løpet av høsten 2020 ble det samlet inn datamateriale på gruppesamtaler i undervisningen i form av deltagelse i undervisningen. I tillegg ble det gjennomført kartlegging av forståelse av grunnleggende Newtonsk mekanikk gjennom en test kalt *Force Concept Inventory* (FCI). FCI ble gjennomført på to forskjellige tidspunkt: Før formell undervisning, og etter endt formell undervisning. I tillegg har studentgruppen blitt kartlagt gjennom en undersøkelse som heter *Interactive engagement and motivation in physics learning* (IMPEL). Spesielt nyttig for denne studien er kartleggingen av hvilken motivasjon studentene hadde etter endt undervisning, og selvrapportert tidsbruk i FYS-0100.



# /2

## Teori

I dette kapitlet vil vi ta for oss teoribakgrunnen for oppgaven. Først vil vi ta for oss et læringsteoretisk bakteppe i seksjon 2.1. Videre vil vi ta for oss undervisningsformen Peer Instruction i seksjon 2.2, før vi tar for oss teori på hverdagsforestillinger og Force Concept Inventory i henholdsvis seksjon 2.3 og 2.4. Deretter vil vi ta for oss motivasjon, og mål på dette i seksjon 2.5, før vi avslutter med relevant statistikkteori som ikke har fått plass tidligere i teorien.

### 2.1 Læringsteoretisk bakteppe

Det konstruktivistiske synet på vitenskapen er generelt akseptert og mye brukt (Sjøberg, 2014). Et slikt syn fastholder at vitenskap er konstruert av forskere i et vitenskapelig samfunn, som søker å måle og konstruere modeller av den naturlige verden (Sjøberg, 2014). I følge konstruktivisten er verden uavhengig av menneskelig forståelse, men kunnskap om verden er en menneskelig og sosial konstruksjon. Konstruktivismen forkaster synet på vitenskapen som en leverandør av objektiv sannhet om verden, men fremmer et dynamisk syn på vitenskap (Sjøberg, 2014). Den teorien som best forklarer et fenomen i den naturlige verden er kun den best tilgjengelige, og skulle en bedre forklaring bli utarbeidet, vil den gamle bli forkastet. Konstruktivisme som læringsteori tar utgangspunkt i at individer selv konstruerer ny forståelse og kunnskap, basert på det de allerede kan (Angell mfl., 2019; Sjøberg, 2014). Konstruksjonen av

kunnskap starter fra fødselen av. Fra man er født blir man utsatt for en rekke sanseinntrykk og fenomener, og med utgangspunkt i at alle individer har et behov for å mestre sin egen virkelighet, må disse forklares. Vi bygger oss derfor et sett med forestillinger om årsakssammenheng og fenomenenes natur (Sjøberg, 2014). Dette er også sentralt i skapelsen av hverdagsforestillinger, som vi vil komme tilbake til.

I en konstruktivistisk syn på læring er det slik at for at individet aktivt skal søke viten, så må det være en viss tvil om det en opplever – en indre konflikt, eller en kognitiv konflikt (Sjøberg, 2014). I så måte er premissene for denne studien i stor grad konstruktivistiske, og det er naturlig å vie det konstruktivistiske perspektiv ekstra plass. I det følgende vil noen konstruktivistiske perspektiver på kunnskap bli presentert. De er enige om at kunnskap konstrueres av individet, men uenige i strukturen denne kunnskapen har.

### 2.1.1 Knowledge in Pieces

Knowledge in Pieces (KiP) er et epistemologisk perspektiv, og et bredt teoretisk og empirisk rammeverk som tar sikte på å forstå kunnskap og læring i lys av det større teoretiske bakteppet, som er konstruktivisme (diSessa, 2018; Harlow & Bianchini, 2020). Det har hatt god suksess med å forklare læringsfenomen i naturfagsundervisning – mest fremtredende hvordan det forklarer fremvoksende kunnskap i lys av preeksisterende forestillinger (diSessa, 2018). KiP er en alternativ teori innen feltet «konseptuell endring», og begynte i fysikkundervisning for å gi en dypere forståelse av preeksisterende forestillinger. diSessa (2018) presiserer at preeksisterende forestillinger ofte stemples som misoppfatninger, men at dette er noe som bør unngås da «misoppfatning» er et negativt ladet ord som gir inntrykk av at forestillingen må korrigeres så kjapt som mulig. Preeksisterende kunnskap kan videre også inkludere hverdagsforestillinger – forestillinger som har blitt til uten formell undervisning – som er et begrep vi senere vil se nærmere på.

Som navnet tilsier, tar KiP sikte på å være en teori om kunnskap og læring som er kapabel til å forklare både kortsiktige fenomen – læring i fragmenterte deler – og langsiktige fenomen, som blant annet innebærer konseptuell endring. KiP er dermed en reaksjon på en del andre læringsteorier som er forutsigende og høytflyvende, og dermed vanskelig å anvende i kaosen som er læring i den virkelige verden (diSessa, 2018).

KiP har mye til felles med teoriene til Piaget (diSessa, 2018). Hovedsakelig ligger likhetene i begge konstruktivistiske trekk – fokuset på hvordan langsiktige endringer begynner i preeksisterende mentale strukturer. En nøkkelforskjell mellom disse teoriene er, i følge (diSessa, 2018), at KiP forsøker å

kombinere både langsiktige og kortsiktige læringsperspektiver. diSessa (2018) argumenterer for at Piagetistisk psykologi aldri var veldig god på å beskrive hva elevenes kortsiktige tankegang hadde å si for langsiktige endringer. KiP er på mange måter i seg selv iboende konstruktivistisk, i og med at det er et læringsperspektiv som er dynamisk og i stadig utvikling (Harlow & Bianchini, 2020).

Det kom tidligere fram at KiP eksisterer i domenet til konseptuell endring, og samtidig bryter med en del av de tidligere læringsteoriene. Spesielt gjelder dette synet på hverdagsforestillinger og misoppfatninger (Harlow & Bianchini, 2020). Det har lenge vært kjent at elever har dannet seg forestillinger om hvordan verden henger sammen basert på dagligdagse erfaringer – som gir mening i de situasjonene de har kommet ovenfor (Sjøberg, 2014). Det er også kjent at disse forestillingene kan være uforenelige med de vitenskapelige teoriene som utdanningsinstitusjoner den dag i dag søker å formidle (Angell mfl., 2019; Sjøberg, 2014). Vi vil komme tilbake til disse forestillingene, kjent som hverdagsforestillinger og/eller misoppfatninger i seksjon 2.3. Det tidligere synet på disse forestillingene har vært, og er nok i stor grad fortsatt, at de bygger opp et paradigme – et velstrukturert og sammenhengende syn på hvordan verden henger sammen (Harlow & Bianchini, 2020; Sjøberg, 2014). diSessa (1993) tok et litt annet perspektiv på disse forestillingene. Han så ikke på en oppfatning eller en forståelse som en enkel helhet, men heller som et sett med mindre, integrert ideer som aktiveres koordinert (Harlow & Bianchini, 2020). Han kalte disse enkelthetene for *phenomenological primitives*, eller p-prims (Harlow & Bianchini, 2020). For å forklare hva disse koordinerte ideene er, tyr Harlow og Bianchini (2020) til en analogi. De drar paralleller mellom interaksjonene mellom disse ideene, og individene i en stim med fisk, eller en flokk med fugler, som svømmer eller flyr i formasjon som en helhet. Vi har alle sett fascinerende klipp av hvordan stimer eller flokker sammen beveger seg koordinert – og det kan være vanskelig å skille ut enkeltindivider fra helheten. En kan se på fuglene eller fiskene som enkeltindivider, eller som en del av en enhet (flokken).

Harlow og Bianchini (2020) sin analogi for fiskestim eller fugleflokk dras videre av undertegnede. Når en fugleflokk eller en fiskestim blir angrepet av en predator – her en vitenskapelig teori – vil flokken sammen, i koordinerte bevegelser, unngå predatoren. Predatoren søker de svake individene, de som ikke passer inn og kanskje beveger seg litt annerledes. I paradigmet, som er flokken, vil predatoren søke de individene som har svake trekk – hverdagsforestillinger eller misoppfatninger. For at predatoren skal lykkes er det avgjørende at det svake individet tvinges ut i åpent lende – isolert fra flokken. På samme måte må en hverdagsforestilling fram i lyset for at den skal kunne bli tatt opp i et vitenskapsteoretisk synspunkt. Dersom en skal dra analogien for langt, kan prosessen sees i lys av den evolusjonære prosessen naturlig seleksjon. Naturlig predasjon, der de svake individene lukes ut, fører over tid til en sterkere popu-

lasjon ved at det er de sterke individene som når fruktbar alder. På den måten føres de fordelaktige egenskapene videre, mens de ufordelaktige ikke overlever.

I følge Harlow og Bianchini (2020) ble diSessa (1993) møtt med noe motstand fra en del andre lærde. Spesielt gjaldt dette ideen om at det var lite underliggende struktur i intuitive ideer i fysikk. Noen år senere publiserte diSessa og Sherin (1998) artikkelen «What Changes in Conceptual Change?». Det var virkelig her at det rådende synet på langsiktige fenomener i læring ble utfordret. I følge Harlow og Bianchini (2020) besto det gjeldende synet den gang i «konseptuell endring». Nøkkelprinsippet til læringsmodellen konseptuell endring beskriver de som prinsippet om at misoppfatninger som tilhører noviser kan slettes, og erstattes med synspunkt som tilhører eksperter. Piaget hadde nok kalt dette for en akkomodasjonsprosess (Sjøberg, 2014).

Det er også Piagets teori om akkomodasjon som var diSessa (1993) og senere også Hammer (2004) sin huggestabbe. Hammer (2004) fortsatte tenkningen til diSessa (1993) om KiP, men han refererte til det som «manifold ontology» som, noe keitete oversatt, blir «en mangfoldig forståelse av hvilke entiteter som eksisterer i tankesett». Som diSessa (1993) beskrev Hammer (2004) synspunktet sitt som et alternativ til de Piagetistiske ideene på det tidspunktet – at elevene har forestillinger som må erstattes og at kognitiv dissonans og akkomodasjon må eksistere for at læring skal skje Harlow og Bianchini (2020) og Sjøberg (2014). Det samme går igjen i konseptuell endring – en forutsetning om at elevene må bli ukomfortable med de forestillingene de har, oppleve konflikt (kognitiv dissonans), og så finne et nytt (vitenskapelig) konsept som løser den kognitive konflikten (Harlow & Bianchini, 2020). Hammer (2004) argumenterte for at dette siste steget ikke kan finne sted dersom forestillingene ble antatt å være kognitive strukturer fordi studentene, ved å bruke allerede tilstedeværende konseptuelle strukturer for å prosessere den nye informasjonen, ikke ville finne veien til ny konseptuell forståelse. Det diSessa (1993) kalte for p-prims, kalte Hammer (2004) for ressurser – og han forsøkte ikke å beskrive deres natur. Hammer (2004), som diSessa (1993), mente at ressursene verken kunne være korrekte eller feilaktige i seg selv, men at det heller er slik at deres nytteverdi er situasjonsbetinget (Harlow & Bianchini, 2020). Han påpekte også at de er kontekst-avhengige, hvilket betyr at ressurser som aktiveres i en kontekst ikke nødvendigvis aktiveres i en annen.

### 2.1.2 Nyansering

diSessa (1993), Hammer (2004) og Harlow og Bianchini (2020) maler et nokså svart bilde av de perspektivene som lenge har eksistert på hverdagsforestillinger og misoppfatninger. De presenterer konseptuell endring som en endimensjonal sak, der det kun er om å gjøre å slette elevenes forestillinger og erstatte



dem med nye. Selv om denne litteraturen har mange gode poeng hva gjelder startpunktet for læring og kontekstbasert forståelse av konsepter, virker det ikke helt som den evner å se at litteraturen som behandler konseptuell forståelse som et velstrukturert nettverk tar for seg mere enn bare sletting av forestillinger og innprenting av ny kunnskap. Vi skal komme mere tilbake til den klassiske synsvinkelen videre i teorien, men skal bruke Sjøberg (2014) for å illustrere her. Sjøberg (2014) skriver at fordi hverdagsforestillinger ofte er godt innarbeidet gjennom erfaringer fra dagliglivet, er de gjerne klart strukturerte og godt innarbeidet i et system. Han skriver at problemet forekommer i de situasjonene der elevenes hverdagsforestillinger kolliderer med de vitenskapelige ideene som utdanningsinstitusjonen forsøker å formidle. Det som da ofte skjer – i følge Sjøberg (2014) – er at eleven lar vitenskapens teorier og begreper være gyldige kun i en skolekontekst, mens deres hverdagsforestillinger og dagligdagse begreper hersker i den virkelige verden utenfor klasserommet. Dette er ikke så ulikt det diSessa (1993) og Hammer (2004) beskriver, om hvordan forestillinger kan aktiveres i enkelte kontekster, men ikke i andre.

Sjøberg (2014) skriver også at læring må starte der eleven befinner seg – det er allmenn akseptert. Videre skriver han at en må vurdere i hvilke tilfeller en kan bygge videre på på elevenes forestillinger – en assimilasjonsprosess i følge Piaget, og i hvilke tilfeller en ikke kan det, og må bryte ned elevenes forestillinger og erstatte dem med nye og bedre – en akkomodasjonsprosess i følge Piaget. Hvilke, og hvordan, vil vi komme tilbake til i seksjon 2.3.

Det vi til nå har tatt for oss, har beskrevet læring som en individuell prosess. Dette er noe vi på mange måte assosierer med Piaget (Sjøberg, 2014). For Piaget lå kunnskapens natur og individet i sentrum. Derfor kan en argumentere for at denne tilnærmingen er kognitiv. I følge Sjøberg (2014) er den kognitive tilnærmingen viktig, men sannsynligvis for smal og kan ikke alene beskrive læring. Han skriver at læring ikke utelukkende er den kognitive evnen til å forstå det en presenteres – at læring også i stor grad handler om interesse og motivasjon, kanskje vel så mye som noe annet. Sjøberg (2014) siterer Guy Claxton, som så fint sier det slik: «Cognition doesn't matter if you are scared, depressed and bored» (Claxton, 1989). I nyere tid har slike tanker blitt innlemmet i et konstruktivistisk læringsperspektiv, som tidligere har vært tilnærmet rent kognitivt (Sjøberg, 2014).

Alle som leser dette har tilhørighet til et sosialt miljø – vi tilhører i en sosial sammenheng. Ved et slikt sosialt miljø er det en rekke karakteristikk – være det kunnskap, handlinger eller uttrykksformer – som oppfattes som anselige. Den sosiale sammenhengen fører også med seg en rekke språklige premisser – ord og uttrykk har en bestemt betydning, og enkelte språklige virkemidler gir mere mening enn andre (Sjøberg, 2014). Dette gjelder også elevene. Elever er ikke tomme beholdere for kunnskap, men individer med

erfaringer og preeksisterende kunnskap – pregede av det sosiale miljøet de tilhører. Dette gjelder også når de møter naturfaget i skolen. I så måte kan ikke læring kun være en personlig og kognitiv prosess (Sjøberg, 2014). Sjøberg (2014) ettertrykker at når eleven ikke vil ta til seg et naturfaglig konsept, så er det ikke nødvendigvis slik at det er fordi det er for abstrakt – det kan rett og slett hende at det ikke passer inn i elevens sosiale sammenheng. Igjen skriver han at det kan lede eleven til å lage seg to verdener: Skoleverden – hvor faglige uttrykk, konsepter og begrep brukes, og den virkelige verden, hvor dagligdagse uttrykk, begrep og hverdagsforestillinger gjelder. På denne måten kan de få det beste fra begge verdener. De presterer greit i skolesammenheng, og trenger ikke å bryte med den sosiale sammenhengen de tilhører. Hverdagsforestillinger er altså ikke bare tett knyttet til elevenes kognisjon – de er også boltet fast i elevenes sosiale miljø. Det handler altså ikke bare om at elevene ikke forstår at de vitenskapelige konseptene kan forklare verden bedre enn deres hverdagsforestillinger, de kan fortsatt holde fast ved sine forestillinger (Sjøberg, 2014). På grunn av dette har det vært naturlig å utvide det konstruktivistiske perspektivet til å inkludere den sosiale sammenhengen der læring ofte skjer – dette har blitt kalt *det sosial-konstruktivistiske perspektivet* (Sjøberg, 2014). Som Sjøberg (2014) skriver eksisterer elevene i en bestemt sosial sammenheng. Denne sosiale sammenhengen inkluderer ofte jevnaldrende (medelever). For studenter er det andre studenter i samme situasjon. Det å ta tak i, og bruke, den sosiale sammenhengen – der Sjøberg (2014) skriver at læring ofte skjer – og bruke den til å bearbeide elevenes hverdagsforestillinger er en interessant idé. Det er nettopp dette Mazur (1997) snakker om i sin bok *Peer Instruction*, som vi i det følgende skal se mere nøye på, når han skriver at medelever gjerne kan instruere hverandre bedre enn det læreren får til.

## 2.2 Peer Instruction

### 2.2.1 Problemer med dagens fysikkundervisning

I happen to believe that those who currently succeed in the sciences do so in spite of the current educational system, not because of it. I don't think we should be satisfied when a student just knows how to plug numbers into an equation in a given situation, how to solve a differential equation, or how to recite a law of physics (Mazur, 1997, s. 29).

Slik motiverer Harvard-professor Eric Mazur, mannen bak *Peer Instruction*, en overgang fra den konvensjonelle tavleundervisningen til en undervisningsform som i større grad utfordrer den forkunnskapen studenter eller elever sitter med.

Han erfarte at studentene han underviste i et introduksjonskurs i fysikk, hadde suboptimale prestasjoner på fysiske problemstillinger som innebar konseptuell forståelse i en undervisningsform med tradisjonell tavleundervisning (Mazur, 1997). Han tallfestet skuffelsen ved å måle konseptuell forståelse på kraftkonseptet i newtonsk mekanikk ved bruk av Force Concept Inventory (FCI) – en test som måler nettopp dette (Hestenes mfl., 1992; Mazur, 1997). Som en reaksjon utarbeidet han en undervisningsform der studenten står i sentrum. Denne tilnærmingen modifiserer den tradisjonelle formen for undervisning, slik at den omfatter spørsmål som skal avdekke vanskeligheter med fagstoffet, samt aktivisere studentene. Denne tilnærmingen kjenner vi som *Peer Instruction*. Som navnet ymter til, vil studentene, eller elevene, ofte bli instruert av sine medelever i denne undervisningsformen. Dette bygger på en tanke om at de som vet best hva elevene sliter med, er elevene selv, og at de derfor kan ha treffende forklaringer. Peer Instruction har vist seg formålstjenlig, gjennom at det med denne undervisningsmetoden har vist seg at elevene eller studentene har oppnådd en høyere grad av konseptuell forståelse enn de gjorde ved konvensjonell undervisning (Crouch & Mazur, 2001; Fagen mfl., 2002; T. Gok, 2012; Mazur, 1997; Porter, Bailey Lee, Simon, Cutts & Zingaro, 2011).

### 2.2.2 Aktiv læring og Peer Instruction

Aktiv læring er enhver læringsaktivitet der eleven eller studenten interagerer med eller er deltagende i læringsprosessen, som motsetning til å være passiv mottager av informasjon (Knight, 2004). Det kan bety så mangt å være deltagende. Renkl, Atkinson, Maier og Staley (2002) skriver at det er viktig at elevene tar del i «higher order thinking activities» som analyse, syntese, og evaluering.

Peer Instruction er en form for student-sentrert omvendt klasseromssituasjon. Det innebærer at informasjonsflyten fra lærer til elev flyttes ut av klasserommet, og at anvendelse og refleksjon rundt tilegnet kunnskap flyttes inn i klasserommet (Bergmann & Sams, 2012; Crouch & Mazur, 2001). Mazur (1997) utviklet Peer Instruction før introduksjonen av omvendt klasseromsundervisning slik vi kjenner den i dag (Bergmann & Sams, 2012). Forskjellen var at Mazur (1997) la opp til at studentene hans før undervisning skulle lese seg opp på et gitt emne, der Bergmann og Sams (2012) legger opp til at forelesningen kan flyttes til utenfor undervisningstimene ved å ta i bruk digitale hjelpemidler og anvende videoforelesninger. Grunnen til dette er selvsagt at Mazur (1997) utviklet Peer Instruction i en tid der slike hjelpemidler ikke kunne forutsettes tilgjengelige for folk flest. Forutenom dette er likhetene slående. Generelt sett ønsker både Mazur (1997) og Bergmann og Sams (2012) å flytte informasjonsoverføringen fra lærer til elev ut av klasserommet, og legge til rette for å kunne bruke undervisningstimene til å øke forståelse og diskutere fagstoff. Det meste tyder på

at en slik tilnærming til undervisning forbedrer blant annet elevenes forståelse (Bergmann & Sams, 2012; Crouch & Mazur, 2001; Fagen mfl., 2002; T. Gok, 2012; Porter mfl., 2011). Vi har til nå med viten og vilje vært noe vag i beskrivelsen av aktivitetene som foregår i undervisningstimene i Peer Instruction, slik som Mazur (1997) har presentert dem. Det er fordi vi nå skal ta for oss disse detaljene.

### 2.2.3 Hvordan undervisning drives i Peer Instruction

Det følgende vil ha hovedutgangspunkt i Eric Mazurs bok *Peer Instruction: A User's Manual* – der han har gjort en fullstendig presentasjon av Peer Instruction (Mazur, 1997). Derfor vil det ikke nødvendigvis videre bli referert til Mazur (1997) der det ikke føles naturlig å gjøre så. I tilfeller der andre referanser har blitt brukt, vil det komme frem i teksten.

Problemet med tradisjonell undervisning – ifølge Mazur – er at måten den drives på fører til et økt inntrykk av at det er problemløsningsstrategier som er hovedfokuset i fysikk. Han mener at gjennom at læreren holder en monolog, fratrar hen elevene eller studentene muligheten til refleksjon rundt det de har blitt undervist. Som tidligere nevnt, ønsker Mazur at elevene skal lese seg opp på det aktuelle emnet før de møter til undervisning. At han mener at elevene skal «lese seg opp», må sees litt med hensyn på den aktuelle tiden. Da Mazur utviklet Peer Instruction var ikke de digitale hjelpemidlene det de er nå, og det kan godt hende Mazur ville ment at videoforelesninger i forkant av undervisning kunne fungert vel så bra. I tillegg mener han man bør benytte seg av er tiltak som heter *Just-in-Time-Teaching* (JiTT). JiTT er et web-basert tiltak som har som mål å forbedre undervisningen (Cashman & Eschenbach, 2003). Hovedessensen er at det skal komme en tilbakemelding til lærer på hva elevene kan om temaet som skal undervises (rett) før undervisningen starter, slik at læreren får lagt opp eller tilpasset undervisningen, i tråd med det resultatene fra tilbakemeldingen avdekker, «just in time» (Novak, Patterson, Gavrin & Christian, 1999; Simkins & Maier, 2010). Denne tilbakemeldingen når gjerne læreren ved at elevene avlegger et svar på en online quiz fram til en time i forkant av undervisning. Da skal læreren bruke den siste timen i forkant av undervisning til å optimalisere undervisningsøkta (Cashman & Eschenbach, 2003). Effektene av JiTT har blitt dokumentert gjennom studier, og resultatene tyder på at JiTT kan føre til redusert utmattelse og økt oppmøte hos elevene (Cashman & Eschenbach, 2003; Novak & Patterson, 1997). Forberedelse før undervisning og JiTT er viktige ressurser som kan forbedre funksjonen til Peer Instruction.

Peer Instruction søker å nyttegjøre seg elevinteraksjoner i studentgruppen i løpet av undervisningstimen, og rette elevenes oppmerksomhet mot de baken-

forliggende konseptene. Undervisningen burde – ifølge Mazur – være bygget opp av kortfattede presentasjoner av sentrale konsepter. Disse presentasjonene skal, for hver presentasjon, etterfølges av en såkalt «ConcepTest», som består av noen konseptuelle spørsmål angående det presenterte temaet. Til disse spørsmålene skal elevene først gi et selvstendig svar, før de diskuterer svaret med medelevene, før de på nytt skal svare. Hensikten er å fremme refleksjon rundt argumentene de utvikler og bruker, og å gi en ressurs, for både lærer og elevene selv, til vurdering av forståelse av konseptet.

Hver ConcepTest kan beskrives slik:

1. Spørsmål blir stilt (1 min)
2. Elevene får tid til å tenke (1 min, lengere tenketid vil føre til tilbakefall på spørsmålet og redusert tankevirksomhet)
3. Elevene avlegger individuelt svar
4. Elevene drøfter svar sammen med medstudent (1-2 min, rom for at lærer kan delta/observere misforståelser.)
5. Elevene avlegger nytt og revidert svar
6. Lærer teller opp svar som en tilbakemelding
7. Forklaring (2+ min)

Akkurat det at lærer har mulighet til å observere, og får umiddelbar tilbakemelding på elevenes forståelse – er et nøkkelpunkt ved Peer Instruction. Det lar læreren vurdere om et tema trenger videre utdypning og oppklaring i etterfølgelse av en ConcepTest. Simon og Cutts (2012) argumenterer for at læreren har større innflytelse med sin forklaring i kjølvannet av en slik diskusjon. Grunnen til dette, mener de, er at elevenes hjerner blir primet til å danne koblinger mellom lærerens forklaring og deres egen forståelse. Tilbakemeldingen kan komme i flere former. Det kan være flashcards eller håndsopprekning, så vel som digitaliserte spørreskjema. Med den digitale undervisningen som har blitt gjennomført under store deler av koronapandemien – kanskje spesielt ved universiteter – vil det være rimelig å gjøre antagelsen om at sistnevnte har vært den rådende tilbakemeldingsformen i løpet av det siste året. Et noe overraskende resultat – ifølge Mazur – er at tross i redusert tid til problemløsning, ble prestasjoner på klassiske problemløsningsoppgaver forbedret ved bruk av Peer Instruction. Mazur er av den oppfatning at dette illustrerer hvordan økt konseptuell forståelse også forbedrer prestasjoner på konvensjonelle vurderinger. Påstanden underbygges av andre som har oppnådd resultater av samme karakter (Hake, 1998; Tobias, 1992).

## 2.2.4 Forutsetninger for at Peer Instruction skal fungere

For at Peer Instruction skal fungere er det fundamentalt at læringsaktørene i klasserommet har et konstruktivt samarbeid. Det gjelder samarbeidet mellom lærer og elever, men også samarbeidet elevene seg i mellom. Mazur (1997) skriver at for at et samarbeidsorientert læringsmiljø skal kunne eksistere, er det avgjørende at elevene ikke føler konkurranse i klasserommet. Johnson, Johnson, Haugaløkken og Aakervik (2006) har skrevet om samarbeid i skolen. Det siteres:

[...] den måten gjensidig sosial avhengighet tilrettelegges på, bestemmer samspillet mellom individene, som i neste omgang bestemmer resultatene. Positiv gjensidig avhengighet (samarbeid) fører til en stimulerende form for samspill hvor individene oppmuntrer hverandre og hjelper hverandre å lære. Negativ gjensidig avhengighet (konkurranse) fører til en form for hindrende (hemmende) samspill hvor individene forsøker å ta motet fra hverandre og hindre at andre anstrenger seg for å lære. Uten gjensidig avhengighet, det vil si blant individer som arbeider uavhengig av hverandre, er det ikke noe samspill mellom individene.

Gjennom dette underbygges viktigheten av at elevene ikke føler konkurranse seg i mellom. Det er også viktig – ifølge Mazur (1997) – å presisere at karaktersetting ikke er kvotebasert – det vil ikke bli vanskeligere å få en god karakter bare fordi flere andre får en god karakter. Målet med dette er å redusere følelsen av at det er en konkurranse om karakterene i klasserommet. En følelse av konkurranse om karakterene kan føre til at elevene blir motvillige til å hjelpe hverandre. Det har blitt gjort forskning som antyder at et samarbeidsorientert læringsmiljø har potensiale til å øke både interesse for faget og motivasjon, samt være en katalysator for dypere forståelse (Cross, 1998; T. Gok, 2012; Simon & Cutts, 2012).

Mazur (1997) påpeker også viktigheten av samarbeidsvillighet hos elevene hva gjelder aksept av bytte til en ny undervisningsform. Han mener elevene sannsynligvis vil vise motstand til innførsel av et nytt regime dersom en ikke gjør grunnen for et slikt skifte til kjenne. Han formoder derfor transparens, og motivering ved et slikt bytte. Det er flere strategiske valg en kan ta for å motivere dette. En bør fra starten av være klar på at det ikke blir tradisjonell undervisning. Videre motiverer man ved å presentere fordelene elevene vil få av å ta til seg Peer Instruction, og man trekke gjerne fram hvilke resultater andre har hatt med det.

Det er essensielt at et undervisningsopplegg ikke kommer isolert fra vurderingsformen. For mange elever er eksamen det endelige målet, og det er den

som til slutt blir avgjørende for hvor elevene legger innsatsen (Mazur, 1997). Dersom elevene har i bakhodet at eksamen vil teste klassiske, algoritmiske problemløsningsferdigheter, er det her de vil legge fokuset sitt. Det er derfor kritisk at eksamensformatet følger undervisningsformen – hvilket har den betydning at eksamen må inneholde en substansiell andel konseptoppgaver (Mazur, 1997).

### 2.2.5 Peer Instruction og konstruktivisme

Yaoyuneyong og Thornton (2011) poengterer at konstruktivistiske læringsmiljø både skal utfordre og støtte elevenes tankeprosesser, og fasilitere for aktiv læring, der elevene selv kan finne frem til konseptene og prinsippene ved et spørsmål, heller enn å kun bli servert fakta. Videre rapporterte de at de konstruktivistiske læringsmiljøene skal støtte elevmedvirkning, la elevene ta ansvar for egen læring, og støtte et samarbeidsorientert klasserom der elevene jobber sammen i grupper for å løse kvalitative og kvantitative oppgaver. I så måte er Peer Instruction en aktiv læringsstrategi som er basert på den konstruktivistiske og sosialkonstruktivistiske læringsteorien (T. Gok & Gok, 2017). Michinov, Morice og Ferrières (2015) har også indikert at Peer Instruction er basert på en sosialkonstruktivistisk tilnærming til læring, der sosiale interaksjoner spiller en uvurderlig rolle – diskusjonen og samarbeidet mellom medelever i et klasserom har en positiv påvirkning på konstruksjonen av kunnskap. King (2002) forklarer at enkle samarbeidsoppgaver, som å la elevene forklare konsepter til hverandre – er bevis på at den sosialkonstruktivistiske tilnærmingen til læring fungerer; den aktive handling det er å undervise et annet individ krever at elevene tydeliggjør, utdyper og rekonseptualiserer læringsmaterialet.

## 2.3 Hverdagsforestillinger

I all sin eksistens har mennesket hatt behov for forståelse av verden som omgir det og være mester av sin egen virkelighet (Sjøberg, 2014). Med vitebegjær som katalysator har vitenskapen blitt drevet fremover – den har til alle tider vært, og er fortsatt i stadig forandring med hensikt å finne forklaring på de fenomenene vi omgir oss med. På mange måter kan dette påstås å være en analogi til et enkelt menneskes vitebegjær (Sjøberg, 2014). (Sjøberg, 2014) skriver at et menneske, fra dagen det blir født til den dag det ikke lenger er blant oss, har et iboende behov for å være sine omgivelses mester. I iver etter å forklare fenomenene vi observerer hender det at vi danner oss personlige forklaringer som ikke nødvendigvis stemmer over ens med de rådende vitenskapelige teoriene – hverdagsforestillinger. Som tidligere fremkom, er disse forestillingene en del av elevenes kognitive strukturer – deres verdensoppfatning – kalt paradigme

(Sjøberg, 2014). Det er i de tilfellene disse hverdagsforestillingene tydelig bryter med de vitenskapelige teoriene som skal formidles i læringsinstitusjoner at det oppstår problemer (Barke, Hazari & Yitbarek, 2009; Sjøberg, 2014).

I sin bok *Fysikkdidaktikk* siterer Angell mfl. (2019) Ausubel: «If I had to reduce all of educational psychology just into one principle, I would say this: the most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly.» (Ausubel, Novak & Hanesian, 1978, før forordet). Naturfagdidaktikken har gjennom en rekke tiår hatt stor interesse for hvordan en kan ta tak i det eleven vet, eller tror hen vet (Sjøberg, 2014). Ausubel nyter stor enighet blant lærere og pedagoger – en må fastslå hvilke forestillinger elevene tar med seg inn i klasserommet, og ta tak i dem på en måte som danner koblinger mellom elevenes forestillinger og vitenskap – slik kan en få de vitenskapelige ideene til å vedvare i elevenes kognitive strukturer (Barke mfl., 2009).

### 2.3.1 Hva er egentlig hverdagsforestillinger?

I ordlyden til begrepet «hverdagsforestilling» ligger også betydningen – at det viser til de forestillingene som fungerer «til hverdags». Det vil si at de ikke er vanskelige å begripe seg på, og er dannet med bakgrunn i elevenes forutværende erfaringer (Angell mfl., 2019). I følge Stojanovska, Petrusevski, Köller og Karlsen (2015) er det også slik at elevene føler at deres etablerte ideer gir mening – og at de derfor må være korrekte. Det kan – i følge Stojanovska mfl. (2015) – være nyttig å dra skilnad på de forestillingene som har blitt til på tross av – eller på grunn av – formell undervisning, og de som har blitt til i fravær av dette. Det finnes et vell av begreper å velge i, men undertegnede finner Hannisdal og Ringnes (2000) sine begreper både lett forståelige og oversiktlige. De skiller mellom to begreper: Hverdagsforestillinger og misoppfatninger. Hverdagsforestillinger er de forestillingene som elevene har konstruert i fravær av formell undervisning – på bakgrunn av egne erfaringer og kunnskapen de hadde til rådighet da forestillingen ble til. Misoppfatninger, derimot, er forestillinger som har oppstått selv om formell undervisning har vært til stede. Det kan være på tross av, men også på grunn av denne undervisningen.

Dersom en elev tar til seg vitenskapelig teori, er det viktig å ha i bakhodet at konseptene som er tilhørende denne teorien ikke nødvendigvis vil opprettholdes – de kan fort erstattes med forhenværende forestillinger (Barke mfl., 2009). Viden akseptert er det videre at det en har lært først tenderer til å være det som er mest resilient. Det vil si at forestillinger som har blitt dannet forutfor formell undervisning er motstandsdyktige mot endring og vil være lette å falle tilbake på (Stojanovska mfl., 2015). Av denne grunn har ofte hverdagsforestillinger stødigere fundament enn de vitenskapelige teoriene skolen



formidler. En viktig strategi for å innprente vitenskapelige teorier i elevenes hukommelse er å repetere nyervervede ideer, slik at de ikke får anledning til å falle tilbake på gamle hverdagsforestillinger (Barke mfl., 2009; Stojanovska mfl., 2015). Sjøberg (2014) skriver også at en av de viktigste forskjellene på misoppfatninger og hverdagsforestillinger er deres motstandsdyktighet mot endring. Hverdagsforestillinger er mye grundigere innarbeidet hos elevene, og stikker dypere enn misoppfatninger. Misoppfatninger vil derfor lettere la seg korrigere. I praksis kan de dog opptre likt – så likt at de ofte ikke skiller på i litteraturen, de refereres bare til som for eksempel «alternative forestillinger», eller på engelsk «misconceptions». De kan altså være vanskelig å skille før man får oppklaring i deres opphav hos elevene, eller alternativt har forsøkt å korrigere dem. Av den grunn kan det være ideelt å notere seg at når det videre vises til hverdagsforestillinger, så er det ikke utelukket at det, på grunn av deres like natur, også vises til misoppfatninger.

### 2.3.2 Hverdagsforestillingers opphav

Som nevnt tidligere har mennesket et immanent behov for å danne seg forklaringer og tolkninger av verden som omgir det. De forklaringene man konstruerer seg bunner i de erfaringene og den kunnskapen man besitter (Barke mfl., 2009). Går en ut i fra at mennesket ikke kan alt eller har erfart alt – hvilket det sjelden gjør – er det rom for feilaktige forklaringer på erfarte fenomener. Disse forklaringene har vi lært å kjenne som hverdagsforestillinger (Hannisdal & Ringnes, 2000).

Det har vist seg vanlig at hverdagsforestillinger en finner i dag en gang har vært den beste vitenskapelige teorien tilgjengelig (Sjøberg, 2014). I følge Sjøberg (2014) kan dette eksempelvis være forestillingen om at det er øynene som står for aktiviteten når vi ser på noe – det vil si at øynene aktivt sender ut stråler for å avlese det vi retter blikket mot. Det er nesten noe poetisk over det hele – og tankene dras tilbake til den tidligere brukte analogien om at menneskets kunnskapskonstruksjon er analog til vitenskapens utvikling.

Språket kan også fostre flust av hverdagsforestillinger – og kan være støttedrikke for en rekke andre. Sjøberg (2014) slår fast at bakgrunnen for at så mange av de historiske forestillingene får feste i dagens hverdagsforestillinger er at språket speiler fortidens vitenskapelige teorier. Dersom vi skal dra videre på det overnevnte eksempelet kan vi si at vi «kaster blikket», eller alternativt «blikk kan drepe». I løpet av sin skolegang er det sannsynlig at de fleste elever vil finne at begreper fra det hverdagslige språket og det vitenskapelige er uoverenstemmelige. En rekke begreper som brukes i vitenskapen har blitt plukket rett ut av det dagligdagse språket, men har fått en annen vitenskapelig betydning. Dette har skjedd fordi begrepene brukes som presise vitenskapelige

redskaper (Sjøberg, 2014). I tillegg er det slik at en del begreper i vitenskapen oppleves som både abstrakte og lite familiære for elevene. I begge tilfeller er utfallet gjerne at elevene mistolker begrepene og får gale ideer (Stojanovska mfl., 2015). Som vi har vært inne på tidligere kan det også hende at elevene i slike tilfeller danner seg to parallelle verdener – skoleverden og den virkelige verden utenfor skolen (Sjøberg, 2014). Elevene lærer seg at i skoleverden gjelder de vitenskapelige begrepene og teoriene – mens i den virkelige verden er det hverdagslige begreper og forestillinger som spiller hovedrollen (Sjøberg, 2014).

### 2.3.3 Hvordan avdekke elevenes hverdagsforestillinger

Ausubel mfl. (1978) hevder at det mest essensielle i en undervisningssituasjon er å finne ut hva elevene kan, og designe undervisningen deretter. Barke mfl. (2009) understreker at læreren helst skal være klar over de hverdagsforestillingene elevene tar med seg inn i klasserommet, med hensikt å legge opp undervisningen på en hensiktssvarende måte. Han legger til at en ikke kan anta at elevene har et gitt sett med hverdagsforestillinger som er de mest vanlige i befolkningen – hvilke hverdagsforestillinger som råder er høyst variabelt, og kulturell og sosial bakgrunn spiller en stor rolle. Derfor vil det være en nødvendighet å utføre en diagnostisering av hvilke hverdagsforestillinger som er tilstede. De hverdagsforestillingene som blir avdekket skal så inkorporeres i undervisningen. Diagnostiske tester er et solid alternativ når det kommer til diagnostisering av hverdagsforestillinger (Barke mfl., 2009). Barke mfl. (2009) refererer til Taber (2002). Han har bragt til verden begrepet «Learning doctor». I denne metaforiske fremstillingen har læreren, som en lege, i oppgave å diagnostisere og behandle elevene, som i metaforen er pasienter. Taber (2002) trekker også frem diagnostiske tester som et nyttig og viktig redskap i diagnosen. Pasientene skal behandles etter diagnosene som stilles. Det er et formålstjenlig undervisningsopplegg som er behandlingen. Både Stojanovska mfl. (2015) og Barke mfl. (2009) retter en pekefinger mot tanken om at en kan avdekke hele omfanget av hverdagsforestillinger på denne måten. Dette kommer av at elevenes varierende bakgrunn gjør at de har atskillige erfaringer som deres paradigme bygges på. Dette fører til at det blir viktig å la elevene selv finne ut hvilke hverdagsforestillinger de har. Et eksempel på hvordan dette kan gjennomføres, er at en lar elevene få uttrykke sine forkunnskaper på det aktuelle feltet. Dette kan føre til at elevene selv blir ukomfortable med de hverdagsforestillingene de besitter, når det blir klart for dem at det de trodde de visste ikke stemmer over ens med vitenskapelig teori (Sjøberg, 2014).

### 2.3.4 Hvordan utfordre elevenes hverdagsforestillinger

I tillegg til at Sjøberg (2014) skriver det, står det i de nye læreplanenes overordnede del at dersom læring skal forekomme, må det skje med utgangspunkt i de ideene elevene allerede har (Utdanningsdirektoratet, 2017). Det Sjøberg (2014) skriver er egentlig en utvidelse av dette til Piagets konstruktivistiske teorier om assimilasjon og akkomodasjon – han skriver at hver ervervelse av kunnskap vil skje ved endring eller forkastning av elevenes forhenværende kunnskap – deres skjema. Sjøberg (2014) stiller også spørsmål ved hvilke hverdagsforestillinger en kan bygge videre på, og hvilke som må forkastes. Den første utfordringen er, som skrevet i underseksjon 2.3.3, å skaffe seg en oversikt over hvilke hverdagsforestillinger som er utbredte. Dette problemet er ofte (delvis) løsbart ved en diagnostisk test. Dersom de avdekkede hverdagsforestillingene bryter så grunnleggende med de vitenskapelige teoriene skolen søker å formidle, at de ikke har potensiale til å være gjenstand for videreføring – må man utfordre elevene på deres hverdagsforestillinger på en sånn måte at de tvinges til å kvitte seg med de forestillingene de entret klasserommet med – og erstatte dem med vitenskapelige konsepter (Sjøberg, 2014).

Som allerede ymtet til, kan prosessen videre belyses av Piagets teori om den kognitive funksjon (Sjøberg, 2014). Ifølge Piaget er det slik at den kognitive struktur er satt sammen av skjema som hjernen har konstruert (Holt mfl., 2015). Skjemaene er tanker og handlinger som er organisert i symbiose med hverandre, og de er i konstant endring ettersom individet eksponeres for nye inntrykk. I følge Piaget skjer denne prosessen, der nye skjema blir konstruert eller etablerte skjema blir mere intrikate, ved to prosesser – assimilasjon og akkomodasjon (Holt mfl., 2015). Når nye sanseinntrykk blir innforlivet i etablerte skjema, kalles det for assimilasjon. Dersom en bruker en hverdagsforestilling, utvider den, slik at den passer med dagens vitenskapelige teorier – er det rom for å tolke dette som en assimilasjonsprosess. Akkomodasjon må skje i tilfeller der eksisterende skjema ikke kan videreføres – men må endres på eller forkastes og erstattes med nye – og strukturene endres (Holt mfl., 2015; Sjøberg, 2014). Det kan for eksempel sees på som en akkomodasjonsprosess når elevene må kvitte seg med de hverdagsforestillingene de har, for å få på plass vitenskapelig teori, fordi hverdagsforestillingene de besitter er på kollisjonskurs med vitenskapelige konsepter.

I det overnevnte fremkommer det at utidsmessige vitenskapelige teorier dukker opp den dag i dag som hverdagsforestillinger – gjerne har de vært de rådende vitenskapelige teoriene på ett eller annet tidspunkt (Sjøberg, 2014). Denne typen hverdagsforestillinger har potensiale til å være en avsats for videre læring (Barke mfl., 2009). Ifølge Barke mfl. (2009) er det i slike tilfeller viktig å fortelle elevene at deres hverdagsforestillinger har blitt holdt av mange innflytelsesrike vitenskapsmenn, og at de slett ikke er uvanlige. Med dette startpunkt kan en

formidle teoriens utvikling i et historisk perspektiv (Barke mfl., 2009). På denne måten kan elevene få inntrykket av å ha nådd dagens vitenskapelige teorier i takt med vitenskapen.

I arbeid med hverdagsforestillinger vil mange, spesielt støttespillere av den tradisjonelle konstruktivismen, hevde at det er effektivt å skape en kognitiv konflikt hos elevene (Barke mfl., 2009; Sjøberg, 2014; Stojanovska mfl., 2015). Med forbehold om at ikke alle, blant andre diSessa (1993) og Hammer (2004), ville tatt utgangspunkt i en slik synsvinkel – vil vi her ta utgangspunkt i et tradisjonelt konstruktivistisk perspektiv. Stojanovska mfl. (2015) skriver at i de øyeblikkene elever med veletablerte hverdagsforestillinger møter situasjoner som later til å bryte med deres verdensoppfatning står elevene ovenfor et veiskille: De kan enten velge å akkommodere skjemaene sine til å passe de nye sanseinntrykkene, eller de kan avvise de nye opplevelsene fordi de ikke gir mening – vi sier at det skapes en kognitiv konflikt Sjøberg (2014). Barke mfl. (2009) skriver at elevene, i slike øyeblikk, er aller mest mottagelige for påvirkning. Stojanovska mfl. (2015) kommer med forslag til metodikk – egnet for å skape en kognitiv konflikt. Å gi elevene førstehånds erfaring med fenomener, og få aktivt bruk for sansene gjennom praktiske eksperimenter er en utbredt metode som også er velegnet (Angell mfl., 2019; Stojanovska mfl., 2015). Ifølge Holt mfl. (2015) er det også et kjent fenomen fra psykologien at førstehånds erfaring gjør at sannsynligheten for at man husker noe økes. Sjøberg (2014) skriver at dette ikke kan sees på som revolusjonerende – det var nemlig dette John Dewey snakket om for et århundre siden. En av de viktigste tingene John Dewey påpekte var at gjennom refleksjon over erfaringer og sanseinntrykk kan kunnskap og tenkning utvikles (Angell mfl., 2019). I tillegg skriver Stojanovska mfl. (2015) at analogier gjennom sammenligninger med kjente sammenhenger kan gjøre det ukjente kjent. I tillegg nevnes blant annet tegninger, visuelle bilder, og konseptkart.

I en digital undervisningsform vil det være begrensede muligheter for å gi oppgaver som baserer seg på for eksempel praktiske eksperimenter. Brekke (2002) skriver at en gjennom problemstillinger kan skape en kognitiv konflikt, gitt at en er klar over hvilke hverdagsforestillinger elevene har. Disse kan igjen løses gjennom diskusjoner i sosial omgang med andre. I ånden av «just-in-time teaching» og «Peer Instruction» kan det også være anledning til å skape en kognitiv konflikt gjennom at elevene selv ser en demonstrasjonsvideo eller lignende i forkant av undervisning. På denne måten kan de komme til undervisning med en kognitiv konflikt som kan løses gjennom diskusjon med andre.

### 2.3.5 Diagnostisk test

Diagnostiske tester har i det foregående blitt dratt frem som et nyttig verktøy for å avdekke elevenes hverdagsforestillinger (Barke mfl., 2009; Taber, 2002). Målet med en diagnostisk test er, ifølge Wallace (2015), å få kartlagt elevenes potensiale og utviklingsområder, evaluere i hvilken grad de vil behøve hjelp, og finne ut hvilke utfordringer som skal brukes. I tillegg påpekes det at en diagnostisk test gjerne ikke kommer alene – den er et naturlig startsted for utviklingen av en plan for undervisningen. Det er også viktig at en slik test har en formålstjenlig form. Ifølge Treagust (1988) innebærer dette at elevene ikke skal ha vanskeligheter for å forstå hva de svarer på. Som første steg i utarbeidelsen av en undervisningsplan er det naturlig at elevene ikke har hatt noen formell undervisning om temaet den diagnostiske testen skal ta for seg. Derfor er det lite sannsynlig at de har en forståelse av de vitenskapelige begrep som inngår for å forklare de vitenskapelige konseptene som angår de gitte tema som skal undervises. Derfor vil det være naturlig å formulere spørsmålene i en diagnostisk test med et slikt språk at de er lettfattelige for alle – et hverdagslig språk (Treagust, 1988). Treagust (1988) påpeker også viktigheten av at en slik test ikke favoriserer de skrivesterke elevene. På bakgrunn av dette er retningslinjene til Queloz, Klymkowsky, Stern, Hafen og Köhler (2017) for diagnostiske tester relevante. Deres retningslinjer går på at spørsmålene burde være utformet som flervalgsspørsmål, og at svaralternativene må reflektere sunn fornuft. I vårt tilfelle betyr dette at svaralternativene skal reflektere kjente hverdagsforestillinger.

## 2.4 Force Concept Inventory (FCI)

I mange tilfeller kan en som instruktør ønske innsikt i elevenes forståelse av konsepter i fysikk. I Newtonsk mekanikk er kraftkonseptet grunnleggende for forståelse (Mazur, 1997). Vi skal nå ta for oss en test som kartlegger testdeltakerens forståelse av kraftkonseptet. Denne testen har fått navn *Force Concept Inventory*.

### 2.4.1 Hva er Force Concept Inventory?

Force Concept Inventory (FCI) er en flervalgstest hvis formål er å evaluere elevenes forståelse av det grunnleggende konseptet *kraft* i Newtonsk mekanikk (Mazur, 1997). Argumentet for nyttheten til denne typen test er at uten dette konseptet er resten av den Newtonske mekanikken ubrukkelig, om ikke meningsløs (Hestenes mfl., 1992). Testen ble utformet i lys av et tilbakevendende problem i fysikkundervisning; hverdagsforestillinger (Hestenes mfl., 1992).

Hver elev starter fysikkundervisningen med et system av forestillinger som virker fornuftig ut i fra elevens kunnskap og erfaringer. Det har blitt viden akseptert at disse forestillingene spiller en stor rolle i introduksjonskurs i fysikk (Hestenes mfl., 1992). I følge Hestenes mfl. (1992) har det blitt spesifikt vist at hverdagsforestillinger som omhandler bevegelse og krefter er spesielt inkompatible med Newtonske konsepter, at konvensjonell fysikkundervisning gjør lite for å endre på disse hverdagsforestillinger, og at dette er uavhengig av underviseren. Dette vil si at fysikklæreren kan være så kompetent og dedikert hen bare vil, men at effektiv undervisning krever mer enn bare dedikasjon og fagkunnskap. Det krever en forståelse av hvordan elever eller studenter tenker og lærer. Med denne bakgrunn utviklet Hestenes mfl. (1992) FCI som et diagnostisk verktøy for å hjelpe lærere å undersøke hvilke hverdagsforestillinger deres elever eller studenter besitter.

Flervalgsspørsmålene som FCI er satt sammen av skal tvinge den som tar testen til å gjøre et valg mellom de hverdagsforestillingene en besitter og vitenskapelig teori. I så måte følger FCI retningslinjene til Queloz mfl. (2017) for hvordan en diagnostisk test burde se ut. Som nevnt i underkapitlet om diagnostisk test i teorien er det viktig at studentene forstår hva det spørres om. Hestenes mfl. (1992) underbygger dette, og sier at spørsmålene i FCI kan oppfattes trivielle for instruktøren, men at den historiske empirien forteller at disse spørsmålene avdekker store kunnskapshull hva gjelder konseptuell forståelse av Newtonsk mekanikk – som konvensjonelle problemløsningsoppgaver ikke avdekker. Strukturen i FCI kan best tolkes som en nedbrytning av kraftkonseptet inn i seks konseptuelle dimensjoner: Kinematikk, Newtons første lov, Newtons andre lov, Newtons tredje lov, superposisjonsprinsippet og typer krefter. For videre fordypning i strukturen i FCI vises det her til tabell 1 i Hestenes mfl. (1992). Det er viktig å påpeke at alle dimensjonene må til for å skape et helhetlig bilde av en students eller elevs forståelse av kraftkonseptet i Newtonsk mekanikk – det vil si at hele FCI må brukes for å skape et helhetlig bilde (Hestenes mfl., 1992).

### **2.4.2 Validitet i FCI**

Huffman og Heller (1995) publiserte en faktoranalyse av FCI der de sådde tvil rundt validiteten til og tolkningen av FCI. De hadde i sine analyser ikke funnet grupperinger i dataen rundt de forskjellige dimensjonene som FCI består av. Som det tydelig fremgår i Hestenes og Halloun (1995) hadde Huffman og Heller (1995) ikke fått med seg at det ikke var slik FCI skulle brukes, en skal ikke se på enkeltdimensjoner, fordi det ikke danner et helhetlig bilde av forståelsen av kraftkonseptet. Dermed endte Huffman og Heller (1995) med å publisere en faktoranalyse som styrker dette argumentet, heller enn å svekke FCI som diagnoseverktøy. Hestenes og Halloun (1995) skriver at FCIs

åpenbare validitet er utover rimelig tvil. Dette begrunnes i at FCI har blitt brukt og undersøkt av mange fysikkprofessorer, at forslag til forbedringer har kommet og blitt fulgt opp – men at disse i stor grad gikk på mindre ting – som formuleringer eller diagrammer. I hvilken grad FCI måler ens forståelse av det Newtonske kraftkonseptet er et spørsmål om innholdsvaliditeten i FCI. For å kunne si noe om denne må en estimere sannsynligheten for et falskt negativt svar og et falskt positivt svar. Et falskt negativt svar er i denne sammenheng et feil svar som har blitt gitt av noen som egentlig hadde den konseptuelle forståelsen til å svare rett. Det er flere ting som kan føre til dette. En grunn til at dette kan oppstå er at testsubjektet ikke har forstått hva det spørres om. Hestenes mfl. (1992) trekker fram språkbarrierer som en utfordring i denne sammenheng. En annen grunn til falske negativt, som trekkes fram som mest sannsynlig av Hestenes og Halloun (1995) er uforsiktighet og uoppmerksomhet. De mener dette fordi spørsmålene er så klart formulert og det korrekte svaret skiller seg sånn ut i fra resten av svaralternativene at en som har forståelse av konseptet lett vil skille ut det korrekte alternativet. Gjennom deres analyse har Hestenes og Halloun (1995) kommet fram til at sannsynligheten for en falsk negativ er maksimalt 10%. Det er vanskeligere å minimere falske positive. Et spørsmål med fem svaralternativer, der ett er korrekt vil ha en iboende 20% sjans for falskt positivt svar selv ved tilfeldig valg. Hestenes og Halloun (1995) skriver dog at studentenes valg er ikke tilfeldige. Det har gjennom intervjuer kommet fram at de valgene de har gjort er velbegrunnede. Dette betyr at det er tiltak en kan iverksette for å minimere støyen fra falske positive. I FCI er det to tiltak i effekt nettopp av denne grunn (Hestenes & Halloun, 1995). For det første så undersøker FCI de samme dimensjonene med flere spørsmål som undersøker flere kontekster og vinklinger. En falsk positiv på ett spørsmål kan bli delvis kompensert for av et feil svar på et annet spørsmål innen samme dimensjon. For det andre er de svaralternativene gode distraksjoner fra det korrekte svaralternativet for en student som mangler Newtons forståelse. Dette har blitt oppnådd gjennom å bruke hverdagsforestillinger som de andre svaralternativene, som vi har nevnt er i tråd med retningslinjene til Queloz mfl. (2017) for den diagnostiske testen. Disse har blant annet blitt samlet inn gjennom ekstensive intervjuer med elever (Hestenes & Halloun, 1995).

### 2.4.3 Reliabilitet i FCI

Før 2010 hadde det ikke blitt gjort noen kjente studier på reliabiliteten i FCI – hverken intern konsistens eller test-retest reliabilitet (Lasry, Rosenfield, Dedic, Dahan & Reshef, 2011). Lasry mfl. (2011) har gjort reliabilitetsanalyser på FCI. De har funnet at FCI måler ett enkelt konstrukt. Det ble målt høy intern konsistens, samt at høy test-retest reliabilitet viser at totalscoren i FCI er en nøyaktig måling.

De fant først den høye test-retest reliabiliteten noe forunderlig fordi testsubjektene endret svaret sitt i en tredjedel av tilfellene. Det viste seg at grunnen til at test-retest reliabiliteten likevel ble høy var at de individene som hadde svart rett på testen stort sett også hadde svart rett på retesten, mens de som hadde svart feil på testen i mange tilfeller byttet til et annet galt svar på retesten. Som vi har nevnt er det slik at de feilaktige svarene på et FCI-spørsmål skal reflektere hverdagsforestillinger og/eller forestille sunn fornuft. Det er dermed ikke utenkelig at det er flere feilaktige alternativer som virker innbydende for en som ikke har adekvat forståelse av det Newtonske kraftkonseptet – og at dette kan føre til bytte av svaralternativ mellom test og retest.

#### 2.4.4 Bruksområder for FCI

Det er flere bruksområder for FCI (Hestenes mfl., 1992). For det første kan den brukes som et diagnostisk verktøy for å avdekke og klassifisere hverdagsforestillinger. Intervjuer basert på FCI blir trukket fram som en nyttig metode for å få videre innsikt i elevenes hverdagsforestillinger. Hestenes mfl. (1992) påpeker videre at når instruktøren har fått tilstrekkelig innsikt i elevenes hverdagsforestillinger bør hen inkorporere denne innsikten i klassediskusjoner for å undersøke hverdagsforestillingene, og stimulere interaksjon mellom elevene – som kan føre til konseptuell endring. Allerede her har det dannet seg konturene av en elevaktiv undervisningsform, og det er jo nettopp dette Mazur (1997) har tatt utgangspunkt i når han har utviklet Peer Instruction for å inkorporere elevdiskusjoner rundt hverdagsforestillinger i undervisningen.

Dens andre, og kanskje viktigste, bruk er som verktøy for å evaluere egen praksis (Hestenes & Halloun, 1995). I så tilfelle brukes FCI som en førtest før undervisningsperioden starter, og så som en ettertest etter endt undervisningsperiode. Basert på resultatene vil man kunne se om forståelsen hos elev- eller studentgruppen har forbedret seg.

### 2.5 Motivasjon og mål på motivasjon

Motivasjonsteorier søker å anskueliggjøre, predikere og begripe seg på menneskelig atferd (Wæge, 2007). Ifølge Lillemyr (2007) må skolen, eller enhver annen læringsinstitusjon, ta i betraktning elevenes motivasjon for at den skal være suksessfull. Dette begrunnes i at økt motivasjon fører med seg økt tilfredshet, innsatsvilje, engasjement, og læringsutbytte. Hestenes mfl. (1992, s. 9) rapporterer mye av det samme – at to klasser med ganske like elevgrupper, både i form av intelligens og matematisk kompetanse har oppnådd helt forskjellige gains (36% vs. 5%). Instruktørene i de to gruppene mente at for-



skjellen på klassene var holdningen deres – den ene klassen var motiverte og hadde stor sannsynlighet for å initiere aktivitet på egenhånd, der den andre klassen krevde stadig tilsyn fra instruktøren. Dette er interessant, og all den tid det ikke er tallfestede data er det innbydende å undersøke om tallfestet motivasjon og gain i FCI korrelerer slik som Hestenes mfl. (1992) antyder. For å for eksempel kunne sjekke om en undervisningsform har medført økning i motivasjon, kan det, som anskueliggjort ved Hestenes mfl. (1992) sitt eksempel, være interessant å kunne tallfeste motivasjon. Motivasjon er et latent konstrukt, og måles derfor gjerne gjennom kjennetegn på motivasjon (Eccles & Wigfield, 2002; Garcia & Pintrich, 1996; Pintrich, Marx & Boyle, 1993). Disse målene kan inkludere mestringsforventning, nytteverdien en ser i faget, selvrealisering og identitet, interesse og glede for arbeid med faget, og hva det koster en å ta faget (Eccles & Wigfield, 2002). Slike konstrukter måles gjerne heller ikke direkte, men heller gjennom et sett med spørsmål der en ikke nevner for eksempel «mestringsforventning» med ord (Hilsdorf, 2020). Spørsmålene som stilles for å måle det latente konstruktet kalles gjerne en skala. En skala er i bunn og grunn en summert indeks, som består av et sett indikatorer, for eksempel spørsmål, som til sammen kan måle et latent konstrukt (Ringdal, 2018).

### 2.5.1 Motivasjonsteori

Testen brukt i denne studien kalles for *Interactive engagement and motivation in physics learning* (IMPEL) (UiO, udatert). Et dypdykk i motivasjonsteori er mer enn denne seksjonen vil ta for seg. I stedet vil vi fokusere på en spesifikk artikkel som utvikleren av IMPEL gjennom samtaler har kommunisert at hen har lagt hovedfokuset på under utarbeidingen av testen. Kartleggingen tar for seg fem motivasjonskonstrukter, nemlig *interesse og glede for å jobbe med fysikk*, *nytteverdi*, *mestringsforventning*, *selvrealisering og identitet*, og *kostnad*. Deres betydning vil – i det følgende – i all hovedsak være basert på modellen til Eccles og Wigfield (2002), ispedd noe annen relevant teori. I og med at det herved er nevnt at teorien fra Eccles og Wigfield (2002) er gjennomgående i denne seksjonen, vil det videre ikke bli referert til nevnte artikkel – der det ikke føles nødvendig.

Det er flust av teorier som fokuserer på individets tro på egen kompetanse og mestringsevne, forventninger om suksess eller fiasko (Eccles & Wigfield, 2002). Denne troen på egne ferdigheter relaterer direkte til spørsmålet, «Kan jeg gjennomføre denne oppgaven?». Generelt sett er det slik at et individ som besvarer dette spørsmålet bekræftende presterer bedre, og er mer motiverte til å velge mer utfordrende oppgaver.

Eccles og Wigfield (2002) viser til Bandura (1997) sin teori på følelse av mestringssevne (self-efficacy). Bandura definerte et individs følelse av mestringssevne

som den tilliten det har til egen evne til å organisere og utføre et gitt handlingsmønster for å løse et problem eller utføre en oppgave. Han mente også at det er et konstrukt som har flere dimensjoner. Det vil si at det kan variere i generalitet og styrke. Dermed mente han at det er slik at enkelte har en sterk følelse av mestringsevne, og andre ikke: noen har en god følelse av mestringsevne i mange situasjoner, mens andre har smalere følelse av mestringsevne som omfatter et mindre sett av situasjoner. I tillegg mente han at det er slik at noen har følelse av mestringsevne på selv de vanskeligste oppgavene, mens noen kun har det på lettere oppgaver. Banduras teori på følelse av mestringsevne fokuserer på mestringsforventning – dog mente Bandura at det er to typer tro en kan ha til forventning: utfallsforventninger (outcome expectations) – tro på at bestemte handlinger fører til bestemte utfall – og mestringsforventning (efficacy expectations) – troen på at en selv kan utføre bestemte handlinger som er nødvendige for å nå et bestemt utfall. Forskjellen mellom disse to er i all hovedsak at selv om et individ tror at en gitt atferd fører til et utfall, er det ikke sikkert at dette individet har troen på at hen kan agere på en slik måte at hen produserer dette utfallet. Bandura betydret at utfallsforventningen må være til stede, men at det er individets mestringsforventning som i stor grad bestemmer et individs målsetninger, valg av oppgaver, vilje til innsats, og utholdenhet. Mestringsforventningskonstruktet til Bandura har blitt tatt i bruk på en rekke felt, deriblant skole, helse og sport. Det har vist seg at resultater fra studier støtter Banduras teoretiske spådommer. For eksempel har det vist seg slik at høye personlige akademiske forventninger har forutsett påfølgende gode prestasjoner, valg av emner, og yrkesrettede ambisjoner (se Bandura (1997)). Kort fortalt støtter altså Banduras teori at mestringsforventning forteller noe om hvor godt studentene forventer å gjøre det i en oppgave eller et fag, for eksempel fysikk (Wigfield, 1994).

Teoriene om mestringsforventning er et nyttig verktøy for å forklare individets ytelse på forskjellige prestasjonsbaserte oppgaver. De tar derimot ikke for seg grunnen til at noen velger å ta del i forskjellige slike oppgaver. Det er ikke slik at en nødvendigvis gjør en oppgave bare fordi en er sikker på at en kan. Vi skal nå ta for oss hvorfor en velger å investere seg i en oppgave.

Det er flere teorier som fokuserer på å skille mellom indre og ytre motivasjon. Fellestrekk er at individer som er indre motivert, tar del i en aktivitet fordi de er interessert i og har glede av aktiviteten. Når individer er ytre motivert, engasjerer de seg i en aktivitet på grunn av ytre stimuli, som for eksempel å motta ros eller en belønning.

Deci og Ryan (1985) publiserte selvbestemmelsesteorien som en respons til økende bevis på at ytre motivasjon kunne være direkte hindrende for at individet skulle gjennomføre aktiviteter som det fant iboende interessante. Teorien integrerer to perspektiver på motivasjon. Det første perspektivet sier at mennes-

ker er motiverte til å opprettholde et optimalt nivå av stimulering. Det andre sier at mennesker har et grunnleggende behov for selvbestemmelse (eller autonomi), kompetanse og tilhørighet. De argumenterte for det første perspektivet med at et individ vil søke optimal stimulering i aktiviteter det finner utfordrende. Individet vil finne at disse aktivitetene er indre motiverende fordi de har et grunnleggende behov for kompetanse. Videre argumenterte de for at den indre motivasjonen til å holde på med en oppgave kun kan opprettholdes dersom individet føler seg kompetent og føler at det har en stor grad av selvbestemmelse. I følge Eccles og Wigfield (2002) er det gjort funn på at ytre kontroll og negative tilbakemeldinger på kompetanse svekker indre motivasjon – som støtter teorien til Deci og Ryan (1985). På toppen av dette argumenterte Deci og Ryan (1985) for at de grunnleggende behovene for kompetanse og autonomi spiller en rolle også i mere ytre motivert oppførsel. Her tar vi for oss et eksempel, som presentert i Eccles og Wigfield (2002): Se for eksempel for deg en student som bevisst og uten ytre press velger en bestemt studieretning fordi den vil hjelpe ham med å tjene mye penger. Denne studenten er styrt av hans grunnleggende behov for kompetanse og autonomi, men basisen for hans valg av studieretning ligger ikke i studieretningen i seg selv – den er en ytre faktor. Til slutt foreslo Deci og Ryan (1985) at et grunnleggende behov for tilhørighet forklarer hvorfor individer gjør ytre mål til indre mål gjennom internalisering.

Csikszentmihalyi og Csikszentmihalyi (1992) beskrev indre motivert atferd på bakgrunn av de umiddelbare subjektive opplevelsene som oppstår når noen er engasjert i en aktivitet. Csikszentmihalyi og Csikszentmihalyi (1992) beskrev en tilstand han kalte for «flow», eller flyt på norsk. Denne tilstanden kjennetegnes kort fortalt av at et individ er fullstendig fordypet i en aktivitet – så fordypet at individet glemmer tid og sted. Det argumenteres også for at følelsen av flyt er en belønning i seg selv som driver individer til å engasjere seg i en aktivitet. I tillegg virker det å være slik at gjentatte opplevelser av denne tilstanden kun er mulig dersom individet søker mer og mer utfordrende aktiviteter, og dermed øker sin underliggende kompetanse for å møte disse utfordringene. I så måte driver opplevelsen av flyt individet til videre utvikling, gjennom å forsterke den adferden som fører til forbedring.

*Interesse* er også et mye studert konsept. En skiller gjerne mellom individuell og situasjonsrettet interesse. Individuell interesse er en relativt stabil orientering der individet er trukket mot visse tema, mens situasjonsrettet interesse er en emosjonell tilstand som oppstår på grunn av spesifikke trekk ved en aktivitet eller oppgave. Her kan en fort tenke at situasjonsrettet interesse er det samme som vi har diskutert tidligere – at det er utfordringen ved oppgave i seg selv som skaper interesse, men det stemmer ikke helt. Situasjonsbasert interesse i denne sammenheng viser til karakteristikkene ved en (akademisk) oppgave som skaper interesse. Det kan være for eksempel personlig relevans, nyheter

eller forståelighet. Individuell interesse kan både vise til følelsene eller verdiene en person assosierer med en oppgave eller aktivitet – og de er direkte relaterte til den aktiviteten, ikke til ytre faktorer. I følge Eccles og Wigfield (2002) indikerer forskning som er gjort på interesse at individuell interesse er tett linket med dyp læring (e.g. gjenkjenne hovedkonsepter og relasjonene mellom disse, kunne svare på spørsmål om dypere forståelse), heller enn overflatelæring (e.g. kunne svare på enkle spørsmål, enkel gjengivelse av tekst). En skulle derfor kunne tro at interesse skulle være en fordel i en undervisningsform som fokuserer på konseptuell forståelse, slik som Peer Instruction gjør.

Det har vært stor interesse rundt hvordan prestasjonsrelaterte målsetninger påvirker prestasjonsrelatert adferd. Blant andre Bandura (1997) har vist at spesifikke og proksimale, men samtidig utfordrende mål fremmer både mestringsforventning og forbedrede prestasjoner. Disse målene har gjerne vært relativt smale, men det har også blitt definert bredere mål og disse har blitt utforsket. Nicholls, Cobb, Yackel, Wood og Wheatley (1990) definerte to hovedtyper av motivasjonsrelaterte målsetningstyper: egorelaterte mål og oppgaverelaterte mål. Individuer med egorelaterte mål ønsker å maksimere evalueringer knyttet kompetanse som er positive, mens de ønsker å minimere negativ evaluering. De fokuserer gjerne på om det de gjør vil sette dem i et positivt lys, og om de vil prestere bedre enn andre i de oppgavene de påtar seg. Derfor velger de oftere oppgaver de vet de kan gjennomføre, og velger bort utfordringer. Individuer med oppgaverelaterte mål fokuserer på å mestre oppgaver og øke egen kompetanse. Dette kommer til uttrykk gjennom at de tenderer til å fokusere mere på hva de kan lære av en oppgave og hvordan de kan gjennomføre oppgaven. Disse individene velger gjerne oppgaver som er utfordrende, fordi disse oppgavene har større sannsynlighet for å utvide deres kompetanse. Wentzel (1991) har også vist at mål relatert til akademiske prestasjoner inkluderer å se på en selv som suksessfull, pålitelig, ønske å lære nye ting, og ønske å få ting gjort. Ford og C.V. (1987) skrev om de målene et individ har for sitt forhold til miljøet rundt seg. De kom fram til at disse målene inkluderte selvbestemmelse og identitet, mål for sosiale forhold som tilhørighet og sosialt ansvar, og oppgaverelaterte mål som mestring, materiell gevinst, og trygghet. Ford og C.V. (1987) definerte også mål som var sentrale innad i individet selv. Disse inkluderer blant annet affektive mål (som lykke og velvære) og kognitive mål (som utforskning og intellektuell kreativitet).

## 2.5.2 Interactive Engagement and Motivation in Physics Learning (IMPEL)

De som først introduserte de fem motivasjonskonstruktene målt i Interactive Engagement and Motivation in Physics Learning (IMPEL); interesse og glede, nytteverdi, mestringsforventning, selvrealisering og identitet, og kostnad; var Eccles, Adler, Futterman, Goff og Kaczala (1983). Det er disse vi kjenner som underkategoriene i Eccles-modellen i motivasjonsteorien. Eccles og Wigfield (2002) gjengir disse, og vi skal nå ta for oss disse og hvordan de relaterer til teorien vi nettopp har presentert.

Spørsmålene i IMPEL er stilt på en sånn måte at studentene skal svare ett av følgende alternativ:

1. Helt uenig
2. Delvis uenig
3. Hverken enig eller uenig
4. Delvis enig
5. Helt enig

Et eksempel på et slikt spørsmål er «Jeg følte at fysikk passet for meg.»

Vi ser at konstruktet «interesse og glede» fra IMPEL, hvis spørsmål er presentert i tabell A.1, kobles direkte på indre motivasjon av Eccles og Wigfield (2002). Reeve (1989) skriver også at dette konstruktet bidrar til indre motivasjon – interesse bidrar med å vekke og rette oppmerksomheten mot gitte tema – mens glede bidrar med å opprettholde viljen til å arbeide og investere tid. Eccles og Wigfield (2002) never også at et individ engasjerer seg i en aktivitet på bakgrunn av interesse og glede, og argumenterer deretter for at et individ kan finne en aktivitet interessant på basis av at den gir en passende utfordring. Gleden individet føler ved å gjøre en aktivitet kan komme fra en følelse av å være kompetent og selvbestemt. I spørsmålene på dette konstruktet fra IMPEL reflekteres dette ved at det blir spurt om studentene er interesserte i å lære nye ting i fysikk, om de synes det er gøy å lære fysikk og om de synes fysikk er iboende spennende. Det siste, spesielt, gjenspeiler konseptet interesse – som vi har sett på tidligere. Det er også ett spørsmål i konstruktet interesse og glede fra IMPEL som tar for seg adferd som karakteriserer indre motivasjon – flow. Vi beskrev tidligere flow som en tilstand der et individ oppsluker fullt og helt av en aktivitet og mister følelsen av tid og sted. Det er akkurat denne karakteristikken som blir gjengitt i et av spørsmålene i konstruktet interesse og glede – som jo måler indre motivasjon. Ut i fra linken mellom personlig interesse og dyp læring er det mulig å danne seg en forventning om at en vil se en positiv korrelasjon mellom høy score på konstruktet interesse og glede målt ved IMPEL og konseptuell forståelse målt ved FCI (som vi kommer tilbake

til senere).

Eccles mfl. (1983) og Wigfield (1994) skriver at «nytteverdien» studenten ser i et fag eller en arbeidsoppgave er hvor godt det relaterer til fremtidige arbeidsoppgaver eller mål. Deci og Ryan (1985) argumenterer, som vi har sett, for at de grunnleggende behovene for autonomi og kompetanse spiller en rolle i ytre motivert adferd. Spørsmålene som stilles i IMPEL på konstruktet nytteverdi måler i stor grad det samme som det eksempelet fra Eccles og Wigfield (2002) tar for seg. Det blir spurt om i hvilken grad å studere fysikk vil gi studentene de jobbmulighetene de ønsker seg og også direkte om de tror de vil ha nytte av det de har lært i fysikk.

Konstruktet «selvrealisering og identitet» er direkte knyttet til de målrelaterte teoriene vi så på tidligere. Dersom vi ser på spørsmålene som er presentert i tabell A.4 finner vi utsagn som: «Det er viktig for meg å lære fysikk.» Dette er utsagn som går direkte på teorien til Wentzel (1991) om akademiske mål. Andre utsagn, som «Jeg liker å fortelle folk at jeg tar emner i fysikk.», kan sees på som egorelaterte mål ved at individ som setter dette høyt ser på å ta fysikkemner som en prestasjon som setter dem i et positivt lys – og får dem til å se smart ut. Det måles også identitet, gjennom spørsmål om individet føler seg som en fysikkperson. Spørsmål som er rettet mot at individet føler at fysikk passer for det virker å sikte etter å måle en form for tilhørighet. (Wigfield, 1994) skriver at *selvrealisering og identitet* er viktigheten studentene tilskriver arbeidsoppgavene de skal gjøre eller det de skal lære basert på deres oppfatninger av egen identitet og idealer, eller deres kompetanse på et gitt felt.

Wigfield og Eccles (2000) beskriver *kostnaden* et individs avgjørelse av å delta i en aktivitet som «Cost refers to how the decision to engage in one activity (e.g., doing schoolwork) limits access to other activities (e.g., calling friends), assessments of how much effort will be taken to accomplish the activity, and its emotional cost».

### 2.5.3 Reliabilitet i IMPEL

Når vi forsøker å måle et latent konstrukt, slik vi gjør når vi måler motivasjonskonstruktene i IMPEL, så ønsker vi selvfølgelig at alle spørsmålene som utgjør skalaen skal måle det samme – konstruktet vi forsøker å måle. Generelt sett så vil det si at vi ønsker å teste skalaen vår for validitet og reliabilitet. En skala er valid hvis den måler det den skal måle, og reliabel hvis den måler det samme hver gang (Hilsdorf, 2020).

## Sanne verdier

Det finnes to forklaringer på hva sanne verdier er (Ringdal, 2018). Vi skal ta for oss en empirisk forståelse. Den sanne verdien til en entitet er den verdien entiteten *faktisk* har. I nærmere detalj er det gjennomsnittsverdien av  $n$  mål, gjort med samme måleinstrument, der  $n$  går mot uendelig (Ringdal, 2018). Det trengs ikke veldig mye praktisk erfaring for å konkludere med at det er en umulighet å gjøre et uendelig antall mål, men de sanne verdiene er allikevel viktige når en skal ta for seg reliabilitet.

## Definisjon av reliabilitet

I klassisk test-teori er den grunnleggende modellen for et mål gitt som

$$X = t + e, \quad (2.1)$$

der  $X$  er den variabelen en måler,  $t$  er den sanne verdien og  $e$  er (tilfeldige) målefeil (Ringdal, 2018). Under antagelsen om at  $t$  og  $e$  er uavhengige variable vet en, fra loven om at hvis variable er uavhengige så er variansen lineært uavhengig, at variansen til en sum av variable er summen av variablenes individuelle varianser:

$$Var(X) = Var(t) + Var(e). \quad (2.2)$$

Nå skal vi definere reliabiliteten til et mål. Den er den andelen av variansen til  $X$  som blir beskrevet av variansen til den sanne verdien  $t$  (Ringdal, 2018). Den forteller dermed også hvor mye av variansen som må tilskrives målefeil. Formelen blir i følge Ringdal (2018):

$$r_{XX} = \frac{Var(t)}{Var(X)}. \quad (2.3)$$

På bakgrunn av de sanne verdienes flyktige natur er ikke formelen over direkte anvendbar. En kjenner ikke de sanne verdiene, og dermed heller ikke hva målefeilene er. Derfor har en måtte komme fram til andre måter å måle reliabiliteten i en variabel på. Minimumskravet for å beregne reliabiliteten til en variabel er at man enten har en indikator som er målt på to forskjellige tidspunkter, eller at man har to indikatorer som måler samme underliggende variabel på samme tidspunkt (Ringdal, 2018). Dersom en har det førstnevnte, bruker man en metode som kalles for *test-retest*. For denne metoden vises det til Ringdal (2018) for videre lesning. I vår data har vi det sistnevnte og da bruker vi *intern konsistens* for å beregne reliabilitet. Intern konsistens har den fordel at kun ett måletidspunkt er nødvendig. Det uovertruffent mest anvendte målet på intern konsistens er *Cronbach's  $\alpha$* .

### Cronbach's $\alpha$

Cronbach's  $\alpha$  er et mål på reliabilitet i en skala (indeks). Den forteller og hvor stor den indre konsistensen er (Hilsdorf, 2020). Med den indre konsistensen menes det i hvor stor grad hver av spørsmålene (indikatorerne) som utgjør skalaen måler det samme konstruktet. Måler de det samme vil Cronbach's  $\alpha$  være høy, gjør de ikke det vil den være lavere. Formelen for Cronbach's  $\alpha$  er gitt som

$$\alpha = \frac{N\bar{r}}{1 + (N - 1)\bar{r}}, \quad (2.4)$$

der  $N$  er antallet spørsmål, og  $\bar{r}$  er gjennomsnittet av korrelasjonene mellom spørsmålene (gjennomsnittet av entitetene i øvre triangel av korrelasjonsmatrisen). Verdiene av Cronbach's  $\alpha$  vil spenne mellom 0 og 1. Jo høyere verdien er, jo bedre er den indre konsistensen, og dermed reliabiliteten. Generelt sett er en  $\alpha$ -verdi på over 0.7 akseptabelt (Hilsdorf, 2020; Ringdal, 2018). Kvadratroten til Cronbach's  $\alpha$  gir et estimat på korrelasjonen mellom skalaen og den underliggende sanne verdien (Ringdal, 2018).

### Lokalisere spørsmål som bidrar til lavere reliabilitet

Innad i hver enkelt skala som skal måle konstruktene kan det være spørsmål som drar ned reliabiliteten – som ikke måler det samme som resten av spørsmålene. I statistikkprogrammet SPSS som kan brukes til å regne ut Cronbach's  $\alpha$  til en skala finnes det en funksjon som lar deg se hva Cronbach's  $\alpha$  ville vært dersom en enkelt indikator (spørsmål) ble fjernet. Dette verktøyet viser seg nyttig. Det lar en oppdage hvilke spørsmål som må sees nøyere på. For å finne ut om spørsmål som stakk seg ut (senket reliabiliteten) målte det samme som resten av spørsmålene i skalaen, ble det brukt andre statistiske størrelser. Et spørsmål som gjør opp en skala, skal måle det samme konstruktet som de andre spørsmålene i skalaen. Det er dermed naturlig at et slikt spørsmål korrelerer med flere av de andre spørsmålene i skalaen (Ringdal, 2018). Ved bruk av en funksjon i SPSS som lar deg sjekke inter-skala korrelasjon mellom spørsmål, kan man sjekke hvordan enkelte spørsmål som stikker seg ut i reliabilitetsanalysen korrelerer med de andre spørsmålene – og man kan sammenligne med interkorrelasjonen mellom resten av konstruktene. I følge (Ringdal, 2018) er minstekravet for et spørsmåls korrelasjon mot resten av spørsmålene som gjør opp skalaen på 0.4. I tillegg har det vært nyttig å se på gjennomsnitt og standardavvik, samt kurtose og skjevhet for å se om fordelingene til svarene på indikatorerne som har en korrelasjon i nedre sjikte samsvarer i noen grad med resten av spørsmålene i skalaen. Kurtose og skjevhet har blitt anvendt for å evaluere hvordan fordelingen til et enkelt spørsmål avviker fra normalfordelingen. Det finnes Flere tommelfingerregler for hva som er akseptable verdier for skjevhet og kurtose, men i denne oppgaven har vi valgt å bruke  $|skjevhet| < 3$  og



$|kurtose| < 10$  i tråd med (Brown, 2006; Kline, 2011). De argumenterer for at verdier som ligger utenfor  $|skjevhet| < 3$  og  $|kurtose| < 10$  er problematiske avvik fra normalfordelingen. Det har blitt kommunisert fra utvikleren av IMPEL at prosessen beskrevet ovenfor samsvarer med prosessen hen selv har brukt for å evaluere reliabiliteten i spørsmålene. Resultatene fra de statistiske analysene som har blitt gjort er presentert i tabell A.1, A.5, A.3, A.2 og A.4.

#### 2.5.4 Validitet i IMPEL

Som overnevnt har motivasjonskonstruktene i IMPEL blitt basert på Eccles mfl. (1983) sin modell for motivasjonskonstrukt. Alle indikatorene (spørsmålene) er utviklet av fysikkdidaktikkgruppen ved UiO som driver IMPEL-prosjektet (UiO, udatert). Dette uttrykker en representant for denne forskningsgruppen om validiteten i IMPEL<sup>1</sup>:

Alle indikatorer som er knyttet til Eccleskonstruktene i første del av undersøkelsen er utviklet av fysikkdidaktikkgruppen i Oslo. Gruppen har lang erfaring med Eccles-modellen, to av våre medarbeidere har utviklet flere spørreskjemaer med dette som teoretisk rammeverk. I tillegg til at to eksperter på Eccles-modellen har vært med på å utvikle alle items så har vi pilotert spørreskjemaet i to runder i en piloteringsfase. I piloteringsfasen gjennomførte vi (i runde 1) et fokusgruppeintervju med førsteårsstudenter i fysikk på UiO, deretter (i runde 2) samlet vi inn data fra N=97 studenter i sitt første og andre studieår på UiO. Under fokusgruppeintervjuet fikk vi høre hva deltakerne tenkte om hvert spørsmål i spørreskjemaet. Da fikk vi sjekket at respondentene faktisk forstår spørsmålene, og ikke svarer på noe annet enn det vi ønsket å spørre om. Etter fokusgruppeintervjuene, og justeringene som fulgte, gjorde vi en større pilot blant fysikkstudenter ved unitveristet i Oslo. Etter runde 2 i piloteringsfasen justerte vi noen indikatorer som viste høy skjevhet, derfor er noen av spørsmålene litt sterkere formulert enn andre (Eks.: "jeg er veldig interessert i ...", i stedet for "jeg er interessert", og lignende).

Alle spørsmålene i IMPEL, med unntak av ett som ble lagt til i ettertid på bakgrunn av innspill fra en ekstern forsker, ble pilotert på denne måten. Derfor er validitetsargumentene for de fleste spørsmålene ganske sterke.

1. Dette har kommet frem i skriftlig samtale med en representant for forskningsgruppen ved UiO som utviklet IMPEL.

### 2.5.5 Peer Instruction og motivasjon

Ifølge Abeysekera og Dawson (2015) er behovene for kompetanse, relasjoner og autonomi tett knyttet til individets indre motivasjon. I en omvendt undervisningssammenheng blir elevene oppfordret til aktiv deltagelse. Dette kan igjen tilrettelegge for oppfyllelse av behovene for kompetanse og autonomi, bedre enn i tradisjonell undervisning. Som aktivt deltagende vil det også være sannsynlig at elevene knytter tettere relasjoner til både medelever og læreren. Med bakgrunn i disse argumentene går Abeysekera og Dawson (2015) god for at omvendt undervisning kan skape læringsmiljø som imøtekommer elevenes immanente behov for autonomi, kompetanse og relasjoner – og dermed stimulerer til sterkere indre motivasjon. Det kan være problematisk å overbevise elevene om å endre sin atferd uten at de får noe igjen for det (Ryan & Deci, 2000). Mazur (1997) peker på at en kan påføre Peer Instruction en grad av ytre motivasjon ved å belønne eksempelvis deltagelse på førtester. Belønningen kan komme i form av gratispoeng på neste prøve, eller lignende.

T. Gok (2012) har i en studie målt utvikling av motivasjon over en periode med Peer Instruction. Det fremkom ingen signifikant endring i mestringsforventning eller andre motivasjonskonstrukter. T. Gok (2012) stiller seg dog noe kritisk til resultatet fordi Peer Instruction var i bruk over et relativt kort tidsrom – og ikke nødvendigvis hadde rukket å gjøre en forskjell. Allison (2012) har studert åttendeklassingers motivasjon som følge av Peer Instruction. Studien fant noe økt motivasjon som følge av Peer Instructions inntog. Undersøkelsen ble gjennomført i matematikk.

Det har blitt gjennomført studier – både innenfor matematikk og naturfag – som hentyder en samvariasjon mellom konseptuell forståelse og motivasjon hos elevene (Mansukhani, 2010; Yen, Tuan & Liao, 2010). Tross i nevnte studier later det ikke til å være noe rigid teoretisk grunnlag for å hevde at økt konseptuell forståelse korrelerer med motivasjon hos elevene, ei heller et sterkt forskningsrelatert grunnlag. I fall det skulle vise seg å eksistere en slik korrelasjon, vil det kunne ha viktige implikasjoner for synet på hvordan Peer Instruction inndirekte kan føre til økt motivasjon gjennom økt konseptuell forståelse. Det synes å være behov for videre forskning på dette feltet.

## 2.6 Relevant statistikkteori

I denne seksjonen vil relevant statistikkteori, som det ikke har vært naturlig å presentere i følge med annen teori, bli beskrevet.

### 2.6.1 Effektstørrelser

Når en skal evaluere egen praksis ønsker en gjerne å måle hvor stor effekt undervisningen har hatt på elevenes forståelse – man ønsker å måle en effektstørrelse. Det har vært en rekke forskjellige definisjoner av effektstørrelse, men Kelley (2012) har samlet en rekke definisjoner til «a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest». Hvor enn vagt dette fremstår, forteller det at en effektstørrelse skal kvantifisere størrelsen på et fenomen (for eksempel en økning i konseptuell forståelse) for å besvare et spørsmål (for eksempel et spørsmål om hvor stort læringsutbytte undervisningen har gitt).

Det er vanlig å dele effektstørrelser inn i standardiserte effektstørrelser og ustandardiserte effektstørrelser (Kelley, 2012; Lydersen, 2020). Ustandardiserte effektstørrelser er enkle å forstå. De oppgir effekt i samme enhet som det de forsøker å oppgi en effekt på (Fritz, 2011). Det enkleste eksemplet er differansen mellom to entiteter – for eksempel kan en se på differansen i FCI score før og etter instruksjon. Dette kaller vi for standard (FCI) gain. Standardiserte effektstørrelser har blitt skalert med hensyn på variabiliteten i datamaterialet som studeres (Baguley, 2009). De er ikke direkte anvendbare i populasjonen de er beregnet med hensyn på, men gir et godt, universalt grunnlag for å måle statistisk styrke (Kelley, 2012). Cohen's d, som vi vil komme tilbake til, er en mye brukt standardisert effektstørrelse.

#### Gain

Hvor mye den lærende har forbedret seg kan regnes ut ved å kalkulere et gain (Hake, 1998). Gain kan kalkuleres ved  $G = S_f - S_i$ , der  $G$  er gain,  $S_i$  er scoren i prosent før undervisningsperiodens start og  $S_f$  er scoren i prosent etter endt undervisningsperiode. Dette tar dog ikke i betraktning forbedringspotensiale, og favoriserer dårlige resultater på førtesten. En bedre løsning er et normalisert gain, som defineres til å være forholdet mellom  $S_f - S_i$  og den maksimalt mulige gain (Hake, 1998):

$$G = \frac{S_f - S_i}{100 - S_i}. \quad (2.5)$$

Normalisert gain måler altså hvor mange prosent av forbedringspotensialet som har blitt utnyttet. En vil så typisk ta gjennomsnittet for å finne det gjennomsnittlige normaliserte gain for sin undervisning dersom en ønsker å få en generell evaluering av egen praksis. Det skal presiseres at gain, hverken vanlig gain eller normalisert gain er standardiserte effektstørrelser. Selv om normalisert gain er normalisert og du ikke direkte kan lese av en prosentvis forbedring, har det fortsatt ikke tatt i betraktning variabiliteten i populasjonen. Mazur

(1997) skriver at ved konvensjonell undervisning vil man vanligvis oppleve et normalisert gain omkring 0.25, mens man i en mere interaktiv undervisningsform – som Peer Instruction – kan forvente å se normalisert gain i intervallet 0.36 til 0.68. Mazur (1997) oppgir et standard gain på 0.08 ved konvensjonell undervisning, og mellom 0.14 og 0.21 ved Peer Instruction.

Hestenes mfl. (1992) er kritisk til å ta utgangspunkt i et gain når man skal evaluere undervisningen med FCI. De har samlet inn både førtest- og ettertestdata, og har funnet at førtestresultatene er uniformt lave. Av denne grunn mener de at førtesten ikke behøver gjennomføres, og at det kun er resultatene på ettertest som teller. De skriver at gains vil bli store med dårlige førtest resultater, og små med gode førtest resultater. Videre skriver de at resultater på ettertesten blir nært det samme i begge tilfeller – gitt at undervisningen har vært effektiv. Dette fører til et prinsipp om at det ikke lenger er akseptabelt å legge skylden for dårlige resultater på ettertest på elevenes manglende forkunnskaper. Det er sannsynlig at den største mangelen ligger i undervisningen. Hestenes og Halloun (1995) definerer videre to terskelverdier for FCI-score. Det ene kaller de for «Newtonian entry threshold», og det har en verdi på 0.60, eller 60 prosent. Under denne terskelverdien har en, ifølge Hestenes og Halloun (1995), ikke konseptuell forståelse nok til å ta inn over seg viderekomne tema i fysikken. Det andre kaller han for «Newtonian mastery threshold», og det ligger på verdien 0.85 i FCI, eller 85 prosent andel korrekte svar. En score over dette indikerer at en har forutsetning for å forstå de fleste Newtonske konsepter, og at man er en Newtonsk tenker.

### Cohen's d

Cohen's d er en vanlig standardisert effektstørrelse. Vi starter med to gjennomsnitt,  $\mu_1$  og  $\mu_2$ . Disse kan for eksempel være målinger av FCI i en populasjon på to forskjellige tidspunkter. Differansen  $\mu_2 - \mu_1$  vektet så over et standardavvik – ofte sammenslått standardavvik (pooled std.)  $\sigma_{ss} = \sqrt{\frac{\mu_1^2 + \mu_2^2}{2}}$  (Baguley, 2009; Cohen, 1988). Sammensatt blir dette altså:

$$d = \frac{\mu_2 - \mu_1}{\sigma_{ss}}. \quad (2.6)$$

Cohen selv definerte tre nivåer av effekt; lavt, middels og høyt. Disse tilsvarte henholdsvis en  $d$  på 0.2, 0.5 og 0.8 (Cohen, 1988). I senere tid har Sawilowsky (2009) foreslått å legge til veldig lavt, veldig høyt og svært, med verdier på henholdsvis 0.01, 1.20 og 2.0, til skalaen. Vi presenterer dem i tabell her for å være ryddig.

**Tabell 2.1:** Nivåer av effekt for Cohen's d.

Verdi	Effekt
Veldig lavt	0.01
Lavt	0.2
Middels	0.5
Høyt	0.8
Veldig høyt	1.2
Svært	2.0

## 2.6.2 Lineær regresjon

Lineær regresjon brukes for å finne sammenheng mellom avhengig variabel og en eller flere uavhengige variabler. Det er to typer lineær regresjon – enkel og multipel.

### Enkel lineær regresjon

Enkel lineær regresjon er nyttig for å finne en sammenheng mellom to variabler. Den ene er en uavhengig variabel og den andre er en avhengig variabel. Enkel lineær regresjon ser etter et statistisk forhold, men ikke et deterministisk et (Walpole, Myers, Myers & Ye, 2007). Forskjellen er at i et deterministisk forhold går en ut i fra at en variabel nøyaktig kan forklare den andre variabelen. Ideen med enkel lineær regresjon er å finne den linja som best passer dataen. Den best tilpassede linja er den som minimerer den totale prediksjonsfeilen. Prediksjonsfeilen er feilen mellom datapunktet og regresjonslinja. Denne blir som oftest kvadrert for å ta hensyn til at datapunkt både kan ligge over og under linja – slik at de ikke utligner hverandre.

Modellen har formen

$$\hat{y} = a + bx, \quad (2.7)$$

der verdiene  $a$  og  $b$  må være valgt slik at de minimerer feilen (sum of squared errors)

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2, \quad (2.8)$$

der  $x_i$  er det faktiske datapunktet og  $\hat{y}_i$  er prediksjonen fra modellen (Walpole mfl., 2007). Vi skal ikke videre ta for oss hvordan man finner koeffisientene, men det er beskrevet i Walpole mfl. (2007).

Hvor stor andel av variabiliteten i dataen som beskrives av den tilpassede

modellen beskrives av  $R^2$ -verdien (Walpole mfl., 2007). For data  $y_i$  med gjennomsnitt  $\bar{y}$  er  $R^2$ -verdien

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} = 1 - \frac{SSE}{SST}, \quad (2.9)$$

der vi allerede er kjent med SSE, og SST er den totale summen av kvadrater (total corrected sum of squares)

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2.10)$$

som beskriver variansen som ideelt sett hadde vært beskrevet av modellen (Walpole mfl., 2007).

Fra regresjonsanalyse kan man også få ut p-verdier. P-verdien for hver uavhengige variabel (én i tilfellet enkel lineær regresjon) tester nullhypotesen, som sier at variabelen ikke har noen korrelasjon med den avhengige variabelen (Walpole mfl., 2007). Dersom det ikke er noen korrelasjon, er det ingen assosiasjon mellom forandringene i den uavhengige variabelen og den avhengige variabelen. Den andre ord er det utilstrekkelige bevis for å konkludere med at det er en signifikant effekt på populasjonsnivå.

Dersom p-verdien for en variabel er mindre enn det valgt signifikansnivået (typisk 0.05), gir datamaterialet nok grunnlag for å forkaste nullhypotesen. Det vil si at datamaterialet ditt støtter hypotesen om at det er en korrelasjon som er signifikant forskjellig fra null. Variabelen er statistisk signifikant og kan brukes i modellen (Walpole mfl., 2007). Dersom p-verdien er større en signifikansnivået, indikerer det at det er utilstrekkelig bevis for å forkaste nullhypotesen. Det vil si at en ikke kan konkludere med at det er en korrelasjon mellom den uavhengige variabelen og den avhengige variabelen som er signifikant forskjellig fra null (Walpole mfl., 2007).

## Multipel lineær regresjon

På samme måte som enkel lineær regresjon forsøker multipel lineær regresjon å finne en sammenheng mellom en avhengig variabel og uavhengige variabler. Forskjellen her er at det er flere uavhengige variabler. Det gjør at man ikke ser etter en linje som den beste tilnærmingen til dataen, men et plan eller et hyperplan (Walpole mfl., 2007). Det gir en modell på formen

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n. \quad (2.11)$$

Videre har  $R^2$ -verdien og p-verdien samme posisjon i multipel lineær regresjon som i enkel lineær regresjon, men  $R^2$ -verdien erstattes av en justert  $R^2$ -verdi,

som korrigerer for at  $R^2$  øker med antall uavhengige variabler (Walpole mfl., 2007).

### 2.6.3 t-test

En  $t$ -test er en statistisk test som brukes for å sammenligne gjennomsnittene i to grupper. Den er ofte brukt i hypotesetesting for å finne ut om en prosess eller en behandling har en effekt på den interessante populasjonen, eller for å finne ut om to grupper er forskjellige fra hverandre (Walpole mfl., 2007). En  $t$ -test antar at dataen en har er uavhengig, og ikke avviker problematisk fra normalfordelingen (Walpole mfl., 2007). At datamaterialet ikke avviker problematisk fra normalfordelingen kan vi se til ved å sjekke at skjevhet og kurtose er innenfor  $|skjevhet| < 3$  og  $|kurtose| < 10$  (Brown, 2006; Kline, 2011). I hypotesetesting forsøker man å teste om nullhypotesen stemmer, eller om det er en alternativ hypotese som best beskriver relasjonen en forsøker å utforske. I en  $t$ -test er det om gjennomsnittene i to populasjoner er signifikant forskjellige. Nullhypotesen er at gjennomsnittene er like. Den alternative hypotesen er at de er signifikant ulike. Man beregner en testobservator, som vi videre vil referere til som  $t$ -verdien, ved

$$t = \frac{\bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad (2.12)$$

der  $\bar{x}_1$  og  $\bar{x}_2$  er de beregnede gjennomsnittene i de to populasjonene,  $n_1$  og  $n_2$  er antallet entiteter i hver populasjon,  $s_1$  og  $s_2$  er standardavvikene i de to populasjonene, og  $d_0 = \mu_1 - \mu_2$ .  $d_0$  faller til null under den gjeldende nullhypotesen. Den nevnte formelen for  $t$ -verdier gjelder dersom en ikke kan fastslå at populasjonene har like varianser. I dette tilfellet kan heller ikke antall frihetsgrader regnes ut ved  $v = n_1 + n_2 - 2$ , men skal i følge Walpole mfl. (2007) estimeres ved

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}. \quad (2.13)$$

$t$ -verdien måler størrelsen på differansen mellom gjennomsnittene, relativ til variansen i datamaterialet du bruker den på. Jo større  $t$ -verdien er, jo større er bevisene mot nullhypotesen. Det betyr at det er større bevis for at det er en signifikant forskjell jo lengre bort fra 0  $t$ -verdien beveger seg.  $t$ -verdien kan tolkes ved å sammenligne den med kritiske  $t$ -verdier fra  $t$ -fordelingen. En kritisk  $t$ -verdi kan finnes ved å bruke antall frihetsgrader og et signifikansnivå med en «percent point function» (PPF) (Walpole mfl., 2007). Denne funksjonen er noe intrikat, men `scipy`-biblioteket i Python regner den ut for oss. En kan også

bruke en tabell for t-fordelingen for å finne kritisk t-verdi. Den kritiske t-verdien tolkes som ytterpunktene i et konfidensintervall, hvis størrelse bestemmes av signifikansnivået. Har en for eksempel et signifikansnivå på  $\alpha = 0.05$  vil en ha et 95% konfidensintervall. Som et resultat er prosedyren å ikke forkaste nullhypotesen dersom  $-t_{\alpha/2,v} < t < t_{\alpha/2,v}$ , der  $t_{\alpha/2,v}$  er den kritiske t-verdien (Walpole mfl., 2007).

Gitt at nullhypotesen stemmer kan vi regne ut sannsynligheten for å oppnå observasjonen som har blitt observert (t-verdien) eller en enda mere ekstrem verdi (lenger bort fra 0). Det er denne sannsynligheten vi kjenner som p-verdien. Det vil si at dersom p-verdien er liten er det liten sannsynlighet for at vi ville hadde denne observasjonen dersom nullhypotesen stemmer. En liten p-verdi kan få oss til å forkaste nullhypotesen. Dersom p-verdien er mindre enn signifikansnivået vil det få oss til å forkaste nullhypotesen. Videre fordypning i temaet kan for eksempel gjøres i Walpole mfl. (2007).



# / 3

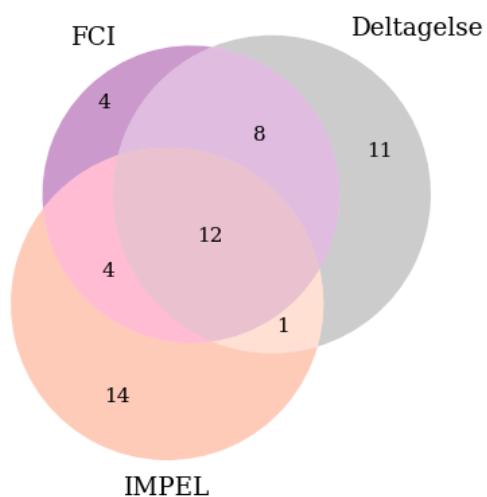
## Metode

I dette kapitlet fremkommer studiens metode. I seksjon 3.1 blir utvalget til studien presentert. Gjennomføringen av emnet som studien har fokusert på blir gjort rede for i seksjon 3.2. Instrumentene som har blitt brukt for å skaffe til veie datamateriale blir presentert i seksjon 3.3. Prosedyren i studien blir lagt frem i seksjon 3.4, og analysene som blir gjort fremkommer i seksjon 3.5. Til slutt vil etiske betraktningen bli redegjort for i seksjon 3.6.

### 3.1 Utvalg

I denne oppgaven har studentene i FYS-0100 generell Fysikk, et introduksjons-emne i fysikk ved UiT Norges arktiske universitet, deltatt. Det var i alt 67 studenter som tok eksamen og var aktuelle for studien. Studenter ble rekruttert i starten av emnet gjennom undervisningen, og deltagelse var frivillig. Samtykke kunne når som helst trekkes tilbake. Studentene samtykket til de forskjellige delene av studien på samme tidspunkt gjennom et samtykkeskjema, men kunne velge å bare samtykke til deler av studien. Alt i alt var det tre forskjellige deler de kunne samtykke til at ble brukt i studien: IMPEL (som inkluderer både motivasjon og tidsbruk, samt personlig data), FCI (som måler konseptuell forståelse), og datamateriale fra undervisning (det relevante for denne studien er deltagelse). Det var 28 studenter som deltok på både før- og ettertest og samtykket på FCI, 31 som deltok og samtykket på IMPEL, og 36 som samtykket på at datamateriale fra undervisning ble innsamlet (hvorav en

har kunnet samle datamateriale på deltagelse for 32 av disse). Hvordan disse overlappet er illustrert i figur 3.1. I tillegg til disse var det 39 studenter som samtykket og deltok i IMPEL i 2019.

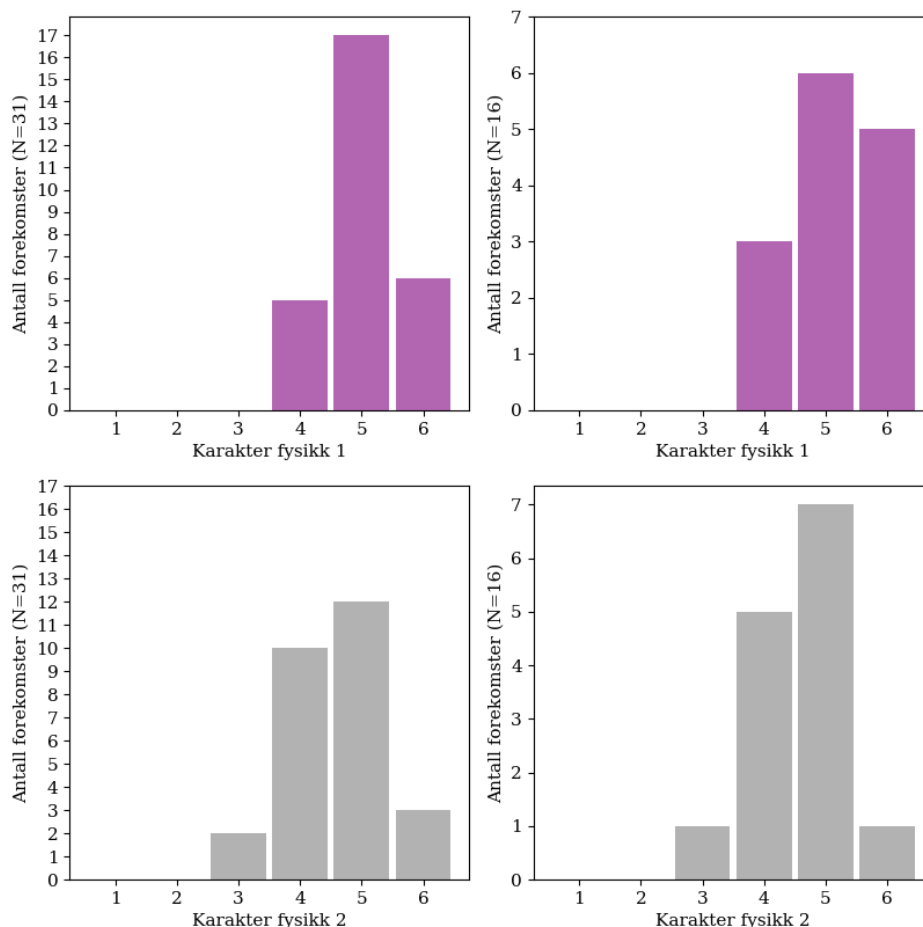


**Figur 3.1:** Venn diagram som viser hvor mange studenter som deltok i hvilke deler av studien.

Gjennom en rekke spørsmål fra IMPEL ble studentene bedt om å oppgi en del informasjon om seg selv. Med utgangspunkt i denne informasjonen kan vi si noe om studentene som deltok i studien. Informasjonen studentene har oppgitt er presentert i tabell 3.1. Karakterene fra fysikk 1 og 2 blir presentert som histogrammer i figur 3.2.

**Tabell 3.1:** Informasjon om deltagerene i studien. I kolonnen til høyre (N=31) er informasjonen til hele gruppen som til IMPEL. I kolonnen i midten (N=16) er informasjonen til de som også har tatt både før- og ettertest i FCI.

<b>Studieprogram</b>	<b>Studie N = 16</b>	<b>Totalt N = 31</b>
Sivilingeniør, AFM	5	6
Sivilingeniør, EKM	5	9
Sivilingeniør, romfysikk	5	9
Fysikk, bachelor	1	2
Geologi	0	2
Kjemi, bachelor	0	1
Matematikk, bachelor	0	1
Matematikk og statistikk	0	1
<b>Kjønn</b>		
Mann	10	17
Kvinne	6	14
<b>Årskull</b>		
1987	1	1
1993	0	1
1995	0	2
1996	1	1
1997	0	4
1998	1	2
1999	2	4
2000	3	7
2001	8	9
Gjennomsnitt	1999	1998
<b>Mors utdanningsnivå</b>		
Tilsvarende doktorgrad	3	3
Tilsvarende mastergrad	5	14
Tilsvarende bachelorgrad	6	8
Fagskole	1	1
Videregående skole	0	2
Ungdomsskole	1	3
<b>Fars utdanningsnivå</b>		
Tilsvarende doktorgrad	1	2
Tilsvarende mastergrad	8	14
Tilsvarende bachelorgrad	3	6
Fagskole	3	4
Videregående skole	1	3
Ungdomsskole	0	2



**Figur 3.2:** Til høyre viser figuren histogrammer av karakterene fra videregående for utvalgsgruppen som tok både FCI før- og ettertest, og IMPEL ( $N = 16$ ). For disse var gjennomsnittskarakteren i fysikk 1 var 5.14, og gjennomsnittskarakteren i fysikk 2 var 4.57. Til venstre sees karakterene for alle som tok IMPEL ( $N = 31$ ). Gjennomsnittskarakteren i fysikk 1 var 5.04, og gjennomsnittskarakteren i fysikk 2 var 4.59 for disse.

Studentene som deltok i undersøkelsen kom fra flere studieretninger, men de fleste var tidlig i sitt studieløp. Emnet FYS-0100 generell fysikk er et introduksjonsemne i fysikk som tar for seg flere temaer som emner senere i utdanningsløpet vil utdype. Det virker rimelig å anta at studentene som tar dette emnet ikke har møtt på Peer Instruction i fysikkfag ved institusjoner for høyere utdanning tidligere. Det har gjennom samtaler med emneansvarlig<sup>1</sup>

1. Emneansvarlig har tilgang til fullt datamateriale på FCI, også de som ikke har samtykket til at det brukes direkte i denne studien. Derfor har fagansvarlig anledning til å si noe om de overordnede trendene som fremkommer her.

kommet frem at at studentene som er med i studien har en gjennomsnittlig poengsum på før- og ettertestene med FCI som ligger veldig nært gjennomsnittet for den totale studentgruppa som tar emnet. Det er et godt argument som taler for at studentgruppa som studeres er forholdsvis representativ for den totale studentgruppa. Dersom vi antar karakter i fysikk 1 og 2 som gode mål, er også gjennomsnittskarakterene for studentene som har tatt både FCI og IMPEL, og alle som har tatt IMPEL ganske like – hvilket taler for at forkunnskapene inn i undervisningen er like gjennom studien. IMPEL fra det foregående året (2019) har også blitt brukt. Det er ingen andre rammebetingelser enn digital undervisning som har blitt endret og som dermed tilsier at utvalget skulle være forskjellig.

## **3.2 Gjennomføring av emnet**

### **3.2.1 Undervisningsform**

Undervisningen i emnet hadde tre forskjellige aktiviteter: Fellesundervisning med Peer Instruction (2x2 timer i uka), regneverksted (2x2) som var ustrukturert regnehjelp og seminar (1x2) som var konseptuelle oppgaver og gjennomgang av oppgaver. Under koronapandemien som har herjet verden det siste året har universitetene gjort sitt for å forhindre smittespredning. Dette innebærer at undervisningen har blitt utført i form av digital hjemmeundervisning. Det ble lagt om til Peer Instruction dette skoleåret. Det ville nok blitt gjort uansett, men digital hjemmeundervisning var en sterk pådriver for å få til dette. Blant andre Mazur (1997) skriver om at gruppene er spontane i Peer Instruction, men situasjonen tatt i betraktning ble det lagt opp til faste grupper i håp om å hjelpe det sosiale. Gruppene ble delt opp ved å bruke FCI førtest. Det var ønskelig at alle gruppene skulle reflektere studentgruppen så godt som mulig, så de ble satt sammen av opp til 5-6 deltagere i hver gruppe, hvor seleksjon ble gjort ut i fra at gruppene skulle ha likt snitt på FCI. Dette ble gjort ved tilfeldig trekking av grupper uten tilbakelegging. Gruppene ble beholdt dersom de hadde et gjennomsnitt på FCI som var like stort som snittet innad i studentgruppen innenfor et gitt avvik.

### **3.2.2 Fellesundervisning**

Før fellesundervisning måtte studentene gjøre en rekke forberedelser, i tråd med teorien til Mazur (1997) – hovedsakelig i form av filmer av pensum, men også litt lesing. Disse ble vanligvis gjort tilgjengelig fredag uken før, og fellesundervisningen var på mandag 10:15. Studentene måtte så avlegge en førtest før undervisning, og de fleste gjorde dette bare minutter før. Førtestene kunne

telle som ett godkjent arbeidskrav dersom de deltok på 70% av dem. De var ikke poenggitte og deltagelse var nok. De ble brukt til å se hva som måtte diskuteres ekstra i fellesundervisning, samt i håp om at det skulle gi bedre læringseffekt – da studentene allerede var primet på fagstoffer før de entret undervisning. Dette ligner på et konsept som kalles for just-in-time teaching, som vi har nevnt i teoridelen, men det er gitt noe lengere tid til forberedelser fordi ren JiTT ikke ville la seg gjøre for en del studenter.

Hvert tema i pensum (bevegelsesmengde, fluidmekanikk, etc.) strakk seg over en undervisningsperiode på 1-2 uker. Undervisningen foregikk etter Peer Instruction, men hadde generelt en liten del mini-forelesninger. Alt ble gjennomført digitalt, og det ble brukt tegnebrett til illustrasjon der det var nødvendig, samt noen enkle demonstrasjoner.

### 3.2.3 Eksamen

Blant andre Mazur (1997) uttrykker, som vi har sett på i teoridelen, at det er viktig at eksamen reflekterer undervisningsformen fordi det er eksamen som er det viktigste målet for de fleste studentene. Eksamen i dette tilfellet var en 4-timers digital hjemmeeksamen, hvor 50% var konseptuelle kortsvarsoppgaver (som de i undervisningen), og 50% var tradisjonelle regneoppgaver.

## 3.3 Instrumenter

Det ble brukt to instrumenter i denne studien. Force Concept Inventory (FCI) ble brukt for å måle studentenes konseptuelle forståelse før og etter undervisning, og IMPEL ble brukt for å måle studentenes motivasjon etter endre undervisning, samt selvrapportert tidsbruk.

### 3.3.1 Force Concept Inventory

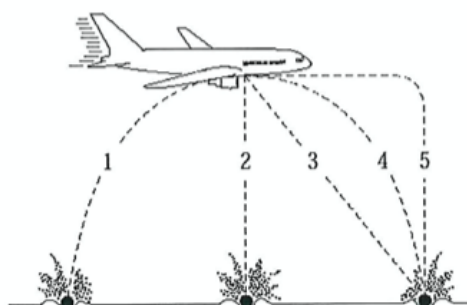
Force Concept Inventory (FCI) er en test som måler konseptuell forståelse av kraftkonseptet i newtonsk mekanikk. Den ble utviklet av Hestenes mfl. (1992) på engelsk men har blitt oversatt til norsk bokmål av forskere ved Universitetet i Oslo (UiO). Testen består av 30 konseptuelle flervalgsoppgaver. Hver oppgave skal teste en spesifikk dimensjon ved kraftbegrepet, og de aktuelle svaralternativene skal undersøke kjente misoppfatninger tilhørende denne dimensjonen. Det mest vanlige bruken av FCI er til evaluering av undervisning, men det er også vanlig å bruke den som en diagnostisk test.

I figur 3.3 er det avbildet et eksempel på en oppgave som ble gitt i FCI-testen høsten 2020. Det er fem svaralternativer, hvor kun ett er riktig.

14. En bowlingkule faller ved et uhell ut av lasterommet til et fly mens det flyr i lufta horisontalt med bakken.

Sett fra bakken, hvilken bane beskriver best bevegelsen til kula etter at den falt ut av flyet?

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



Figur 3.3: Eksempeloppgave fra FCI.

Hestenes og Halloun (1995) argumenterer for at FCI har en validitet som er utenfor rimelig tvil. Dette begrunnes blant annet i at falske negative – hovedsaklig på grunn av at de feilaktige svaralternativene virker usannsynlige for en som har konseptuell forståelse for dimensjonen som testes – beregnes til å ha en sannsynlighet på godt under ti prosent. At testen undersøker forskjellige dimensjoner fra flere vinkler fører til at falske positive også minimeres (Hestenes & Halloun, 1995). Hestenes og Halloun (1995) skriver også for at reliabiliteten har vist seg å være god. Dette har blitt vist med både indre konsistens og test-retest. For mere utdypende teori om FCI, vises det til seksjon 2.4 i teorien.

### 3.3.2 IMPEL

IMPEL-undersøkelsen har blitt utviklet av en IMPEL-gruppen ved UiO, som emnansvarlig samarbeider og gjør datainnsamling med (UiO, udatert). Av denne grunn ble IMPEL foretrukket som instrument for å måle motivasjon. Motivasjon måles i fem forskjellige konstrukter: Nytteverdi, kostnad, interesse og glede, selvrealisering og identitet, og mestringsforventning. Dette er i tråd med og Eccles mfl. (1983) og Eccles og Wigfield (2002) sin modell for motivasjon.

Motivasjonsdelen av IMPEL består av seks til åtte spørsmål per konstrukt, 35 totalt. Hvert av spørsmålene har svaralternativer i form av en 5-delt Likertskala (Likert, 1932). Svaralternativene ble konvertert fra tekststreng til numeriske data som beskrevet i tabell 3.2

**Tabell 3.2:** De konverterte svaralternativene for IMPEL, for at de skal være mulige å arbeide med.

Svaralternativ	Verdi
Helt uenig	0.0
Delvis uenig	0.25
Verken enig eller uenig	0.50
Delvis enig	0.75
Helt enig	1.0

IMPEL besto av én del som måler motivasjonskonstrukter, men IMPEL målte også selvrapportert tidsbruk i emnet og en del personlig informasjon om deltagerne, samt holdning mot emnet.

Validiteten til IMPEL har blitt testet, og en av utviklerene av IMPEL har uttalt seg om reliabiliteten i IMPEL. Dette finnes i seksjon 2.5.4 i teorien. Reliabiliteten har vi gjort analyser på, og det kommer i seksjon 3.5.2.

### 3.4 Prosedyre

Her vil en oversikt over når de forskjellige delene av datainnsamlingen ble gjennomført. I tillegg vil måten datamaterialet ble samlet inn og preprossesert på bli presentert, med unntak av FCI og IMPEL – hvis prosedyre for gjennomføring ble presentert i seksjon 3.3 ovenfor.

1. FCI førtest gjennomført 17. august 2020
2. Skjema for samtykke fylt ut 10. september 2020
3. FCI ettertest gjennomført 19. november 2020
4. IMPEL-undersøkelsen ble gjennomført 20. november 2020
5. Data på deltagelse ble samlet inn 4. februar 2020

Fellesundervisningens periode strakk seg fra 14. august 2020 til og med 9. november 2020. I perioden etter data på deltagelse ble samlet inn, har tiden gått til analyse av datamateriale.

Førtesten til FCI ble gjennomført i en av de første forelesningene, og ettertesten ble gjennomført i en av de siste forelesningene. Den ble implementert i Canvas og ble gjort på norsk bokmål. Studentene fikk 60 minutter til å svare på testen. I tillegg til å bli brukt til å dele inn gruppene, ble FCI brukt for å



kartlegge læringsutbyttet fra emnet. Den ble ikke brukt spesielt med tanke på diagnostisk testing. FCI var frivillig, og samtykke til bruk i forskning ble gjort utenom undersøkelsen.

IMPEL-undersøkelsen ble gjennomført på norsk og implementert i et nettskjema som driftes av UiO. Studentene hadde noen uker på å svare på nettskjemaet. All data ble tatt vare på til senere.

Målet med å inkludere deltagelse i studien var å se i hvilken grad studentene hadde økt læringsutbytte jo mere de deltok i undervisning. Tanken med dette var å gi en pekepinn på om undervisningsformen Peer Instruction i seg selv ga studentene økt læringsbytte i digital undervisning. Deltagelse ble registrert som tilstedeværelse i undervisningen. Dette ble gjort gjennom møtefunksjonen i Microsoft Teams, hvor gruppesamtalene i undervisningen foregikk. Datamaterialet ble registrert manuelt ved å skrive ned hvilke studenter som hadde deltatt i hver enkelt undervisningsøkt.

Det var en del av IMPEL at studentene skulle rapportere eget tidsbruk i emnet. De ble spurt om hvor mye tid de brukte per uke i forelesninger i FYS-0100, i regneverksted i FYS-0100, i seminarer i FYS-0100, og hvor mye tid de brukte på FYS-0100 utenfor dette. De første 3 spørsmålene har blitt samlet til tidsbruk i undervisning, mens det siste står for seg selv som tidsbruk i selvstudie. Vi har også lagt alle 4 sammen til total tidsbruk. Siden vi har tilgang på IMPEL-undersøkelsen fra 2019, har vi også tilgang på datamateriale på selvrapportert tidsbruk fra 2019. I denne undersøkelsen hadde det blitt spurt om hvor mye tid de brukte per uke i forelesninger i FYS-0100, i seminarer i FYS-0100, og hvor mye tid de brukte på FYS-0100 utenfor dette. Det vil si at regneverksted hadde blitt utelatt fra undersøkelsen. Vi ønsket å gjøre sammenligninger på de to årene, slik som det ble gjort med motivasjonskonstruktene fra IMPEL, fordi det er interessant å se om tidsbruk har endret seg i en digital hverdag. Det har dermed blitt besluttet å utelukke regneverksted fra 2020-dataen i akkurat denne analysen. Den vil dog bli brukt i videre analyse, der det gjøres sammensatt analyse med resultater fra FCI.

## 3.5 Analyser

I denne seksjonen vil det bli redegjort for hvilke analyser som har blitt gjort på datamaterialet tilgjengelig. Det som omhandler analyser gjort direkte på FCI vil komme frem i underseksjon 3.5.1, og det som er gjort på IMPEL kommer frem i underseksjon 3.5.2. Sammenligningen av IMPEL og FCI fremkommer i underseksjon 3.5.3, og videre følger sammenligningene av FCI deltagelse og tidsbruk i henholdsvis underseksjon 3.5.4 og 3.5.5.

### 3.5.1 FCI

I det mottatte datamaterialet fra FCI var navnene til studentene erstattet av en tallkode for å sikre anonymitet. Ut i fra denne data ble det regnet ut gjennomsnittlig andel korrekte svar på førtest og ettertest – både for hver enkelt student, men også for hele populasjonen. Det ble også beregnet normalisert gain for gjennomsnittene, samt normalisert gain for hver enkelt student – også her ble det tatt gjennomsnitt av de individuelle gainene. Standard gain ble også beregnet. I tillegg til dette ble effektstørrelsen ved Cohens  $d$  regnet ut. Teorien på disse effektstørrelsene finnes i underseksjon 2.6.1 i teorikapitlet. Andelen riktig svar på før- og ettertest for hver student ble plottet mot hverandre i et spredningsplot. Det ble også plottet som histogrammer for å tydeligere få fram den underliggende fordelingen i datamaterialet. Utrekninger og plotting ble gjort i Python.

### 3.5.2 IMPEL

Det ble gjort reliabilitetsanalyse på indeksene (motivasjonskonstruktene) i IMPEL for å se om de målte det de skulle. Cronbachs  $\alpha$  ble foretrukket til dette formålet, og det ble sjekket om alle indeksene var innenfor akseptable verdier for  $\alpha$  ( $>0.7$  ifølge Hilsdorf (2020) og Ringdal (2018)) som beskrevet i underseksjon 2.5.3 i teoridelen. Videre ble indikatorer sjekket for kurtose og skjevhet, og indikatorer med verdier utenfor tommelfingerregelen til Brown (2006) og Kline (2011) på  $|skjevhet| < 3$  og  $|kurtose| < 10$ , ble forkastet. Dette gjaldt dog ingen indikatorer. Det ble også målt korrelasjon mot resten av indeksen for hver indikator, og det ble beregnet om hver indikator tilførte validitet til indeksen, eller tok bort fra validiteten til indeksen ved å sjekke hvordan hver enkelt indikator påvirket Cronbachs  $\alpha$ . I følge Ringdal (2018) er en korrigeret korrelasjon større enn 0.4 mot resten av indeksen akseptabel. All analysen på IMPELs validitet ble gjennomført i statistikkprogramvaren *Statistical Package for the Social Sciences* (SPSS)<sup>2</sup>. Analysen som har blitt beskrevet ble utført på datamaterialet fra alle som tok IMPEL, men også på datamateriale som hadde blitt filtrert på at de som hadde tatt IMPEL også skulle ha tatt FCI – fordi dette datamaterialet vil bli brukt i senere analyser. Reliabilitetsanalysen finnes i sin helhet i appendiks A.1 og A.2, og førte til at to indikatorer ble fjernet.

IMPEL ble kjørt også høsten 2019 i FYS-0100. For å undersøke endringen fra året før ble gjennomsnittet av hvert konstrukt fra de to årene testet med en tosidig  $t$ -test, beskrevet i underseksjon 2.6.3 i teoridelen, med et signifikansnivå på 0.05. Det ble også gjort et spredningsplot mellom prosentvis deltagelse, og hvert motivasjonskonstrukt, samt total motivasjonsscore fra IMPEL 2020.

2. <https://www.ibm.com/analytics/spss-statistics-software>

### 3.5.3 FCI og motivasjonskonstrukter

I underseksjonene 2.5.1 og 2.5.5 i teoridelen presenterte vi at det har vært gjort studier, og finnes teori som indikerer at det kan være en sammenheng mellom motivasjon og konseptuell og dypere forståelse (Eccles & Wigfield, 2002; Mansukhani, 2010; Yen mfl., 2010). Det har også kommet frem at en dypere forståelse i fysikk involverer en forståelse av konseptene – og sammenhengen mellom dem. I så måte er det interessant å finne ut om motivasjonen, målt ved IMPEL, hos studentene har ført til at de har engasjert seg i fagstoffet på en måte som har gjort at de har oppnådd dypere forståelse i form av konseptuell forståelse, målt ved FCI. IMPEL måler konstruktene interesse og glede, nytteverdi, selvrealisering og identitet, mestringsforventning og kostnad. Det har blitt tatt den frihet å i tillegg samle disse til ett enkelt konstrukt, total motivasjonsscore. Dette ble gjort ved å regne gjennomsnitt av alle svar for hver student, og ble gjort på eget incentiv, ikke med bakgrunn i teori. For hver av de 5 konstruktene, og den samlede motivasjonsscoren, har det blitt gjort spredningsplot mot FCI-gain, og også mot post-score fra FCI. På det samme datamaterialet har det blitt gjort lineær regresjon for å tallfeste en eventuell sammenheng. I tillegg har det blitt gjort multippel lineær regresjon med motivasjonskonstruktene som uavhengige variabler, for å se om en samlet motivasjon henger sammen med endring i forståelse målt ved FCI-gain, eller oppnådd nivå av forståelse målt ved post-score fra FCI. Vi har sjekket p-verdiene som kom ut av lineær regresjon for å kunne se om det er en signifikant korrelasjon. Signifikansnivået har vært 0,05.

### 3.5.4 FCI og deltagelse

Peer Instruction er, som det fremkommer i teoridelen, en undervisningsform som tar sikte på å øke deltagerenes konseptuelle forståelse. En skulle da kunne forvente at – dersom undervisningen har fungert – de deltagerene som har deltatt på en større andel av undervisningen opplevd en større utvidelse av sin konseptuelle forståelse. Når vi vet at FCI måler konseptuell forståelse, er det nærliggende å sjekke hvordan FCI-gain og deltagelse i undervisningen samvarierer. Vi har plottet FCI-gain og FCI post-score mot deltagelse for hver enkelt student i hvert sitt spredningsplot og gjort lineær regresjon på det samme datamaterialet for å kunne tallfeste en eventuell sammenheng. Vi har sjekket p-verdien som kom ut av lineær regresjon for å kunne se om det er en signifikant korrelasjon. Signifikansnivået har vært 0.05.

### 3.5.5 FCI og tidsbruk

Tidsbruk i undervisning, tidsbruk i selvstudie og totalt tidsbruk har alle blitt plottet mot FCI-gain og post-score fra FCI. Som for deltagelse vil det være naturlig å tro at hvis undervisningen har fungert så vil de som rapporterer stort tidsbruk i undervisning per uke ha tilegnet seg mye konseptuell forståelse. For å undersøke om det er en slik relasjon har lineær regresjon blitt benyttet på de samme dataene som er plottet mot hverandre. Vi har sjekket p-verdien som kom ut av lineær regresjon for å kunne se om det er en signifikant sammenheng. Signifikansnivået har vært 0.05.

## 3.6 Ethiske betraktninger

Studentene har samtykket til at datamaterialet som blir brukt i studien blir brukt på slik måte. De ga sitt samtykke i et samtykkeskjema som ble gitt eksternt alle andre datainnsamlinger. Samtykket har vært fritt og informert. Det betyr at de er informert om hva de samtykker til, det er valgfritt, og er ikke bindende – de kan trekke seg når som helst. Dersom de samtykket til å la deres IMPEL data bli studert var de med i trekningen av ett gavekort på 1000kr og to gavekort på 500kr. Det er emneansvarlig som har spurt om studentene kunne tenke seg å delta i studien, i så henseende har det blitt gitt informasjon om at eksamen er anonym, slik at studentene ikke tenker at de blir vurdert annerledes om de ikke deltok i studien. Studien er godkjent av NSD.

I innsamling av datamateriale har jeg ikke vært deltagende i undervisningssituasjonen – jeg har kun brukt datamateriale som jeg har hentet utenom undervisning, og har ikke vært i direkte interaksjon med studentene. Informasjonen jeg har mottatt er andrehånds fra emneansvarlig. I så måte er resultatene upåvirket av at jeg har gjennomført denne studien. Undervisningen er også akkurat den samme for de som deltar i studien og de som ikke gjør det.

# /4

## Resultater

I dette kapitlet vil resultatene fra analysene presenteres. Seksjon 4.1 tar for seg fysikkstudentenes motivasjon. Videre sammenlignes motivasjonen til studentene i 2020 med studentene i 2019 i seksjon 4.2.. Til slutt vil jeg i 4.3 presentere resultater på sammenhengen mellom læringsutbytte og motivasjon, deltagelse og tidsbruk for å undersøke om det er en sammenheng.

### 4.1 Fysikkstudentenes motivasjon

Denne seksjonen skal ta for seg motivasjonen til studentene som tok IMPEL i 2020. I forkant presenterer vi interessante deler av reliabilitetsanalysen som har blitt gjort på indeksene fra IMPEL som måler motivasjonskonstrukter – de delene av den som har ført til forkasting av indikatorer.

#### 4.1.1 Resultater fra reliabilitetsanalyse

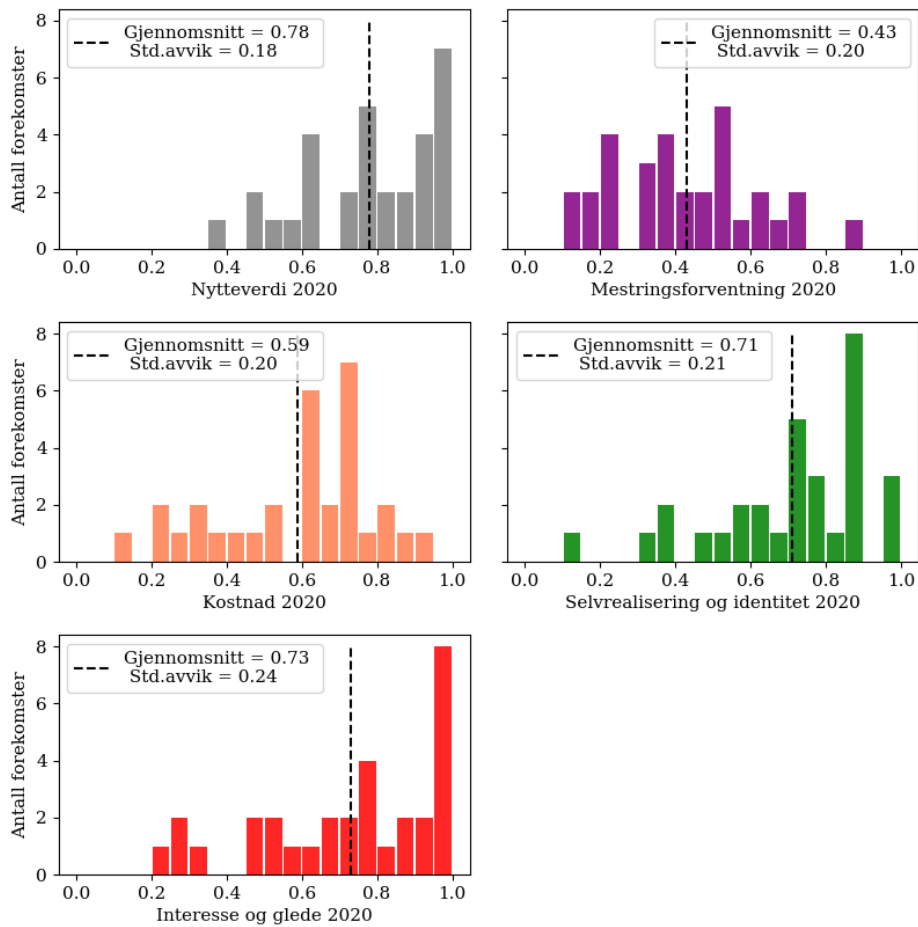
Før reliabilitetsanalysen ble alle indikatorenes skjevhet og kurtose sjekket. Alle indikatorer hadde  $|\text{skjevhet}| < 3$  og  $|\text{kurtose}| < 10$ , som er uproblematisk for videre analyser Brown (2006) og Kline (2011).

I tabell A.4 og A.5 i appendiks A.1 finner vi verdiene for konstruktene selvrealisering og identitet, og kostnad for den fullstendige IMPEL undersøkelsen.

Indikatoren «Jeg følte at fysikk krevde veldig mye tid og arbeid.» fra kostnads-konstruktet i tabell A.5 har en korrelasjon mot indeksen på 0.165, og senker dermed Cronbach's  $\alpha$  betydelig. For konstruktet selvrealisering og identitet er det den siste indikatoren, «Det er viktigere for meg å få gode karakterer i fysikkemner enn i andre emner.» som skiller seg ut med en korrelasjon på 0.273 mot resten av indeksen og senker Cronbach's  $\alpha$  noe. I tabell A.9 og A.10 i appendiks A.2 finner vi verdiene for konstruktene selvrealisering og identitet, og kostnad fra IMPEL for de som også hadde tatt FCI. De samme indikatorene som skiller seg ut også her. For selvrealisering og identitet korrelerer den aktuelle indikatoren negativt med indeksen med en korrelasjon på -0.119, og senker også Cronbach's  $\alpha$  betraktelig. Av denne grunn har disse to indikatorene blitt besluttet fjernet før videre analyse som involverer motivasjonskonstruktene fra IMPEL.

#### 4.1.2 Fysikkstudentenes motivasjon i 2020

Her tar vi for oss resultatene som beskriver fysikkstudentenes motivasjon ved fullført undervisning høsten 2020. Figur 4.1 presenterer et histogram av deltagenes gjennomsnittsverdi for hvert konstrukt. Skjevhet og kurtose for fordelingene som finnes i figur 4.1, er presentert i tabell A.11 i appendiks A.3. Skjevhet og kurtose trengs for å kunne si om fordelingene har problematiske avvik fra normalfordelingen eller ikke i en kommende statistisk test.



**Figur 4.1:** Figuren viser snittsvarene for hver student på hvert konstrukt i IMPEL 2020, samt gjennomsnitt av disse gjennomsnittene med tilhørende standardavvik.

Knytter en hvert gjennomsnitt til den kategorien den er nærmest fra tabell 3.2, finner en at den gjennomsnittlige student er delvis enig i at emnet har god nytteverdi, at emnet fører til god selvrealisering og identitet, og at hen har god interesse og glede av å jobbe med faget. Vi ser at gjennomsnittsstudenten er verken enig eller uenig i at at han har god mestringsforventning, og at det er høy kostnad ved å jobbe med faget. Kostnad er et invers konstrukt, og at det er «bra» om en student føler lite kostnad assosiert med emnet. Hvis en ser på den inverse til kostnaden, som vil være en skala i samme retning som de andre indeksene, er denne verdien på 0.41. Det vil si at både kostnad og mestringsforventning har ganske lave verdier, i nedre sjiktet av å være hverken enig eller uenig i at mestringsforventningen er god og at kostnaden er liten. I tillegg er det interessant at alle verdiene for skjevhet og kurtose, presentert i tabell A.11, er så små som de her, og vi med bakgrunn i disse

beslutter at fordelingene for hver av konstruktene ikke har problematiske avvik fra normalfordelingen.

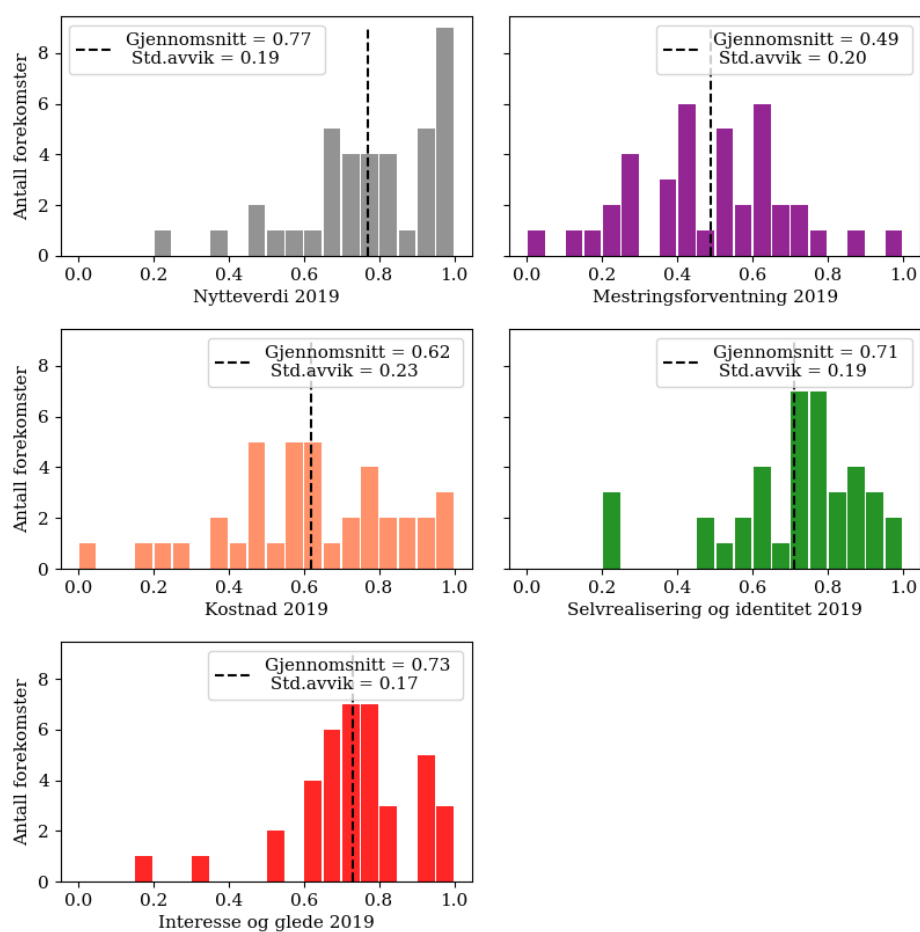
## **4.2 Sammenligning av motivasjon i 2019 og 2020**

I denne seksjonen presenteres resultater på motivasjonskonstrukter fra IMPEL 2019, før disse sammenlignes med tilsvarende datamateriale fra 2020 som presentert i underseksjon 4.1.2.

### **4.2.1 Fysikkstudentenes motivasjon i 2019**

I figur 4.2 finner vi fordelingen av snittsvarene studentene har oppgitt på hvert konstrukt. Gjennomsnittsvarene på hvert konstrukt finnes også i samme figur, med tilhørende standardavvik. Skjevhet og kurtose for fordelingene som finnes i figur 4.2, er presentert i tabell A.12 i appendiks A.3.





**Figur 4.2:** Figuren viser snittsvarene for hver student på hvert konstrukt i IMPEL 2019, samt gjennomsnittet av disse gjennomsnittene med tilhørende standardavvik.

Den gjennomsnittlige studenten er – etter sammenligning med den kategorien den er nærmest fra tabell 3.2 – delvis enig i at emnet har god nytteverdi, at emnet fører til god selvrealisering og identitet, og at hen har god interesse og glede av å jobbe med faget. Den gjennomsnittlige student er verken enig eller uenig i at det hen har god mestringsforventning og at kostnaden med å jobbe med faget er høy. Det er også interessant å notere seg at verdiene for skjevhet og kurtose, presentert i tabell A.12, er lave og at datamaterialet for hvert konstrukt antas å ikke ha problematiske avvik fra normalfordelingen.

## 4.2.2 Motivasjon i 2019 og 2020

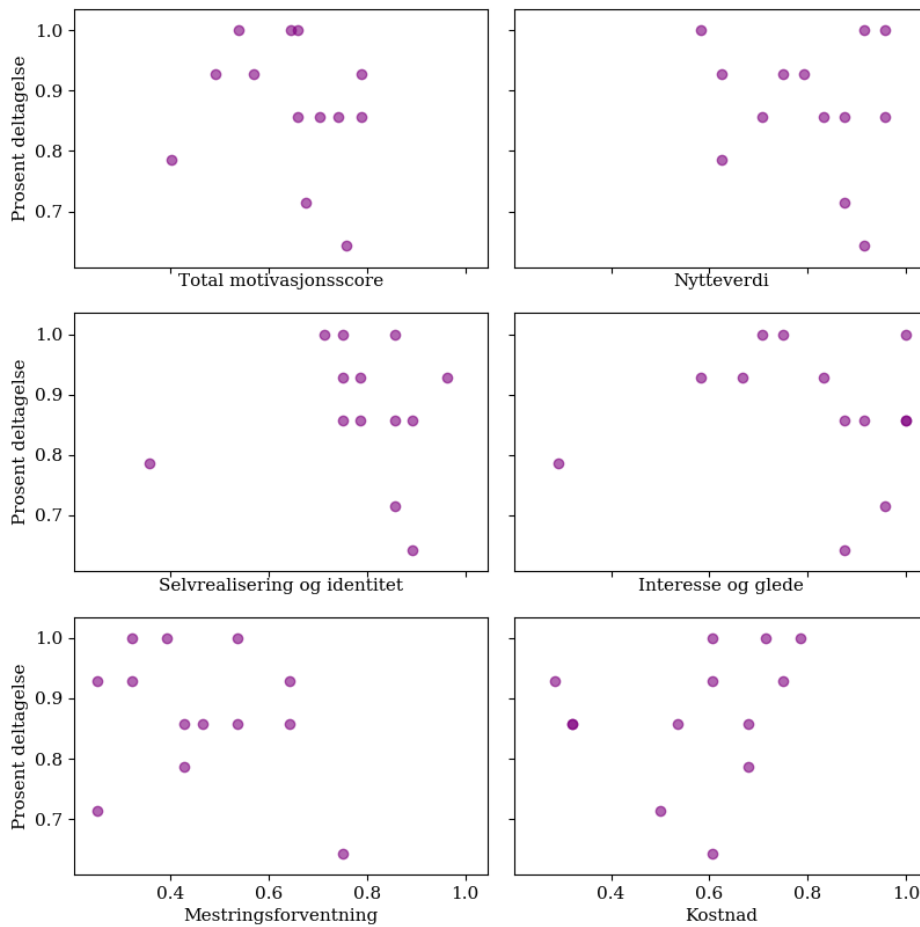
Resultatene på motivasjonskonstrukter fra IMPEL fra henholdsvis 2019 og 2020 har blitt sammenlignet, med mål om å finne ut om det hadde vært en endring i motivasjon mellom disse to årene. For å finne ut om det er en signifikant forskjell på noen av gjennomsnittene har en t-test blitt brukt. Resultater fra denne finnes i tabell 4.1.

**Tabell 4.1:** Resultatene fra t-testen som ble gjort på IMPEL. Vi ser at det ikke er noen signifikante endringer. Signifikansnivået er 0.05.  $t_B$  er beregnet t-verdi, DoF er frihetsgrader,  $t_K$  er kritisk t-verdi, og  $p$  er p-verdien.

Konstrukt	$t_B$	DoF	$t_K$	$p$
Nytteverdi	-0.08	66	2.00	0.93
Mestringsforventing	1.11	66	2.00	0.27
Kostnad	0.61	67	2.00	0.61
Selvrealisering og identitet	0.02	61	2.00	0.98
Interesse og glede	0.06	52	2.01	0.95

Med et signifikansnivå på 0.05 gir ikke resultatene fra tabell 4.1 grunn til å forkaste nullhypotesen om at det ikke er noen signifikant forskjell på gjennomsnittene, for noen av motivasjonskonstruktene.

Jeg ønsket også å undersøke om det var en sammenheng mellom undervisningen med Peer Instruction og fraværet av nedgang i motivasjon. Motivasjonskonstruktene og en total motivasjonsscore er sammenlignet med deltagelse i spredningsplot i figur 4.3.



**Figur 4.3:** Figuren viser spredningsplot av deltagelse mot hver av motivasjonskonstruktene, samt en total motivasjonsscore

Figur 4.3 viser stor spredning i begge retninger, og det fremkommer ingen klar sammenheng. Derfor det ble besluttet å ikke gå videre med en regresjonsanalyse.

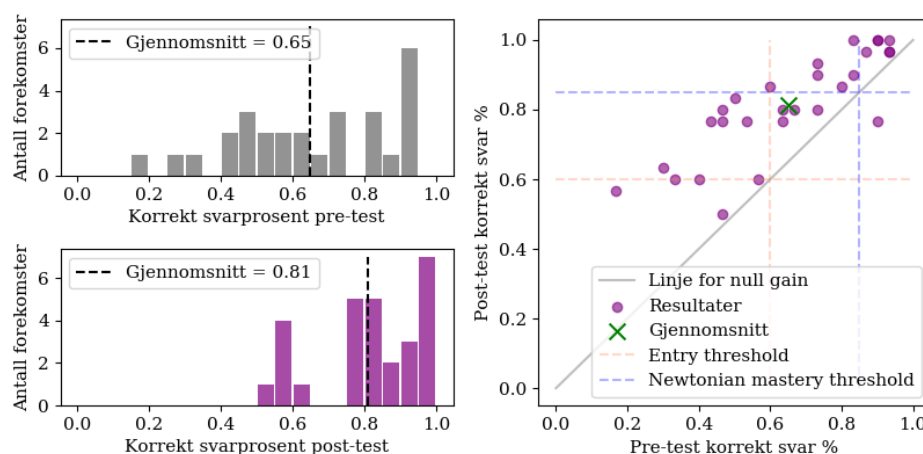
### 4.3 Sammenligning av læringsutbytte med motivasjon, tidsbruk og deltagelse

I denne seksjonen vil resultatene fra sammensatt analyse bli presentert. I de følgende underseksjonene vil vi først ta for oss hvilket læringsutbytte studentene har hatt i digital undervisning med Peer Instruction, deretter vil vi ta for oss

analyser gjort på læringsutbytte og motivasjon, deltagelse, og tidsbruk.

### 4.3.1 Læringsutbytte

I denne seksjonen presenteres resultatene fra analyse som har blitt gjort for å finne læringsutbyttet fra FCI. Figur 4.4 viser i høyre kolonne fordelingen i pre- og post-test fra FCI med histogrammer, og utviklingen til hver student i høyre kolonne ved hjelp av et spredningsplot. Nullgainlinje, samt entry og mastery threshold fra Hestenes og Halloun (1995) er markert i spredningplottet. I histogramplottene er gjennomsnittene markert. Tabell 4.2 presenterer relevante effektstørrelser for FCI.



**Figur 4.4:** I venstre kolonne viser figuren fordelingen av andel korrekte svar i FCI pre- og post-test, presentert ved bruk av histogrammer. Her er også gjennomsnitt markert i hvert av histogramplottene. I høyre kolonne viser figuren FCI pre- og post-test scores for hver enkelt student, samt et gjennomsnitt. I plottet er det også tegnet inn en linje for null gain, der en instans som befinner seg nordvest for linja vil ha positivt gain. I tillegg er Newtonian mastery threshold samt entry threshold markert i figuren.

**Tabell 4.2:** Relevante effektstørrelser basert på FCI og deres verdier for hele populasjonen samlet.

Mål	Verdi
Normalisert gain (gain av gjennomsnittscore)	0.46
Standard gain (post-test score - pre-test score)	0.16
Cohen's d	0.88

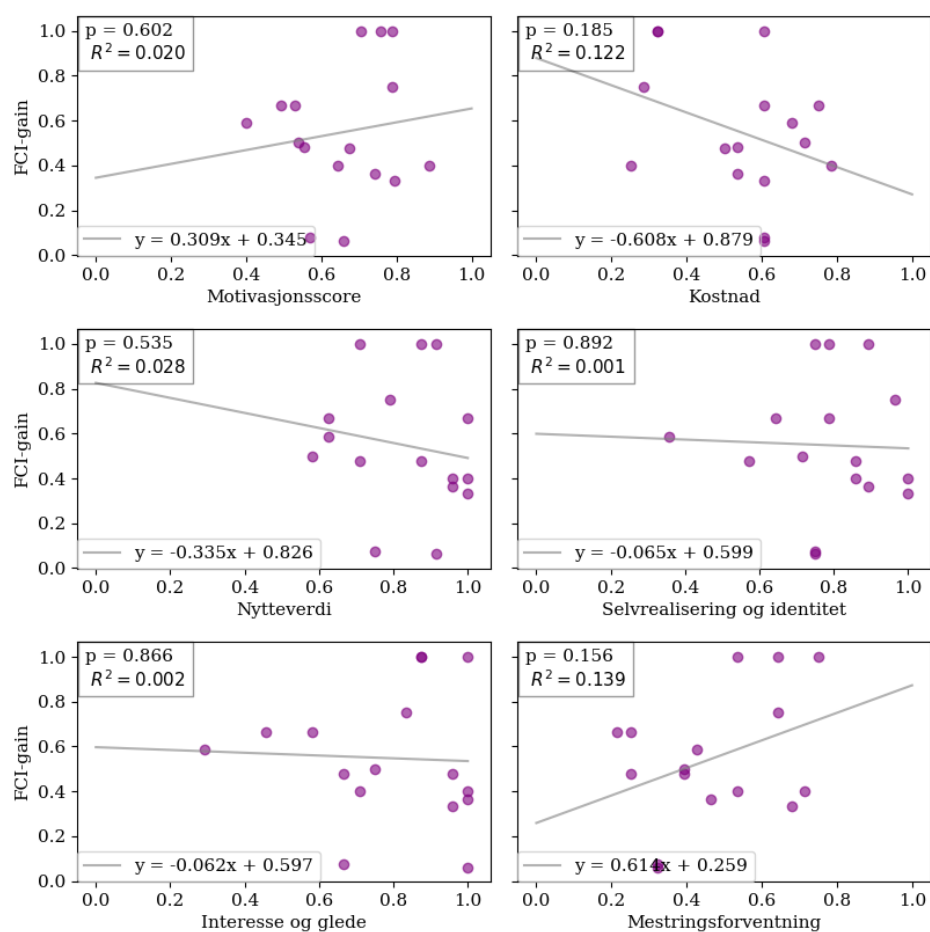
Hestenes og Halloun (1995) og Hestenes mfl. (1992) skriver at det er resultater

på ettertesten som er det mest interessante målet for å evaluere undervisning, og at en snittscore på FCI etter undervisning er et godt mål på undervisningens suksess. Fra histogramplottet nederst til høyre i figur 4.4 ser vi at denne verdien er på 0.81. Sett opp mot terskelverdiene fra Hestenes og Halloun (1995), som presentert i teorien, ser vi at vi er godt over «Newtonian entry level» på 0.60, og rett under «Newtonian mastery level» på 0.85. I tabell 4.2 finner vi også både standard gain, med en verdi på 0.16, samt normalisert gain beregnet av gjennomsnittet fra førtest og gjennomsnittet fra ettertest, i tillegg til gjennomsnittet av normalisert gain beregnet på hvert enkelt individ. Det vanlige gainet ligger i det samme sjiktet som det Mazur (1997) har opplevd ved bruk av Peer Instruction, som er mellom 0.14 og 0.21. Vi ser også at de verdien som er basert på normalisert gain faller innenfor det Mazur (1997) definerer som akseptable verdier for normalisert gain ved en interaktiv undervisningsform som Peer Instruction som er mellom 0.36 og 0.68, og definitivt bedre enn ved konvensjonell undervisning der det ligger rundt 0.25 (Mazur, 1997). Til slutt ser vi av tabell 4.2 at Cohen's  $d$  er på 0.88, som ifølge Cohen (1988) er en høy verdi.

Fra høyre kolonne i figur 4.4 har vi at det før undervisning var 11 studenter som var under «Newtonian entry level», og én som var helt på grensa. Etter undervisning var det kun to som var under, i tillegg til tre helt på grensa. I tillegg til dette har antallet som er over «Newtonian mastery level», altså er Newtonske tenkere, nært fordoblet seg fra 7 til 12. Alle studentene, med unntak av én har også hatt positiv utvikling. Resultatene fra FCI er også presentert som to separate histogrammer for førtest og ettertest i venstre kolonne i figur 4.4.

### 4.3.2 Motivasjon og læringsutbytte

Her vil vi se på resultatene analysene som ble gjort på FCI og IMPEL. Det blir presentert to sett av plots. Begge settene består av spredningsplot, der FCI-gain blir plottet mot motivasjonskonstruktene fra IMPEL i figur 4.5, og FCI post score blir plottet mot motivasjonskonstruktene fra IMPEL i figur 4.6. For å fastslå en eventuell korrelasjon, eller mangel derav, mellom motivasjonskonstruktene og resultater fra FCI har det også blitt gjort lineær regresjon – både enkel og multippel. Resultater fra enkel lineær regresjon mellom resultater fra FCI og motivasjonskonstruktene finnes også i figur 4.5 og 4.6 for regresjonen mellom motivasjonskonstruktene og henholdsvis FCI-gain og FCI post-score. Tilsvarende resultater for multippel regresjon finnes i tabell 4.3 og 4.4.



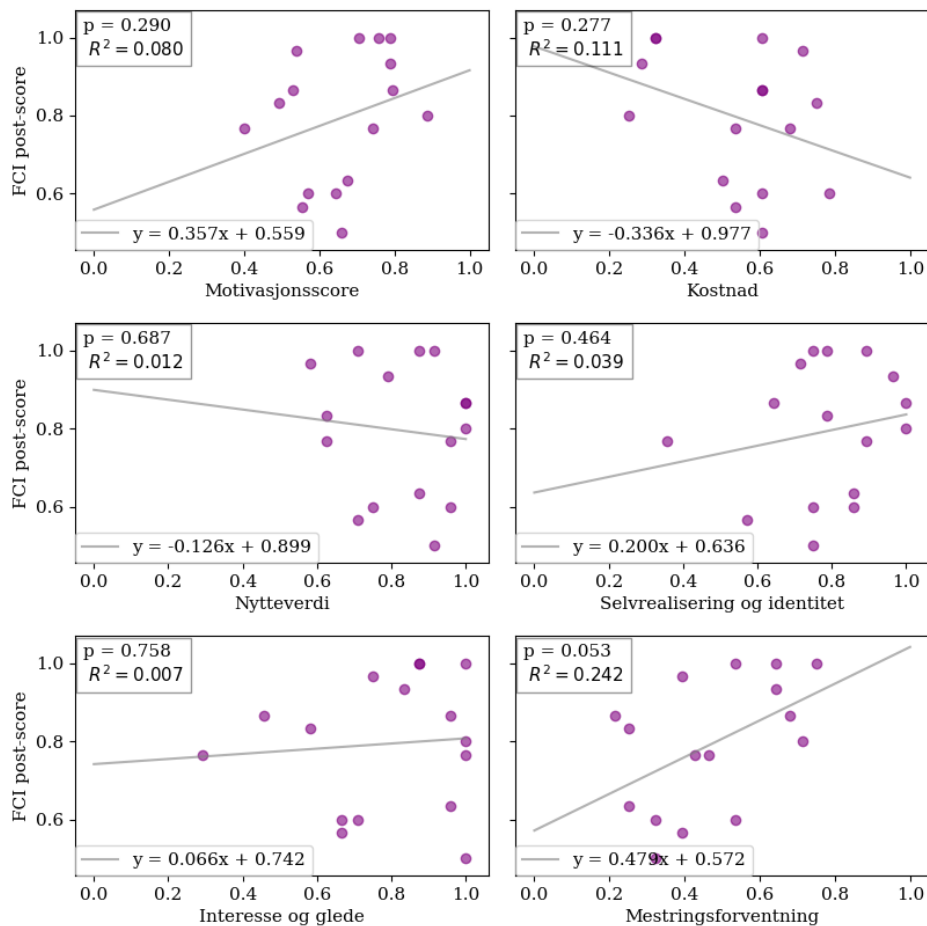
**Figur 4.5:** Figuren viser FCI-gain og motivasjonskonstrukter for hver enkelt student – for hvert enkelt motivasjonskonstrukt, samt en samlet motivasjonsscore basert på scorene fra konstruktene. I tillegg viser figuren resultater fra lineær regresjon. Den oppgitte p-verdien er p-verdien til stigningstallet.

**Tabell 4.3:** Resultatene fra multiplert lineær regresjon på alle motivasjonskonstruktene og FCI-gain. Vi søker en modell på formen  $y = ax_1 + bx_2 + cx_3 + dx_4 + ex_5 + f$ , der FCI-gain =  $y$ , Selvrealisering og identitet =  $x_1$ , Interesse og glede =  $x_2$ , mestringsforventning =  $x_3$ , nytteverdi =  $x_4$  og kostnad =  $x_5$ .

	a	b	c	d	e	f	R <sup>2</sup> (just.)
Koeff. verdi	-0.049	-0.417	0.713	-0.435	0.629	0.659	0.009
p-verdi	0.951	0.506	0.211	0.504	0.288	0.185	

For å undersøke om det er en signifikant korrelasjon kan vi se på resultater fra lineær regresjon som er plottet sammen med datapunktene i spredningsplot-

tet. For at en slik relasjon skal være signifikant burde p-verdien være under signifikansnivået på 0.05. Vi ser at p-verdien til stigningstallet for mestringsforventning er nærmest, på 0.156, som sier at det ikke er noen bakgrunn for å hevde at vi har stadfestet at dersom mestringsforventning eller FCI-gain øker, vil den andre også gjøre det. Den multiple regresjonen som har blitt gjort med motivasjonskonstruktene i hver dimensjon, og FCI-gain som målvariabel, produserer heller ikke noen signifikant korrelasjon i noen av dimensjonene.



**Figur 4.6:** Figuren viser FCI post-test score og motivasjonskonstrukter for hver enkelt student – for hvert enkelt motivasjonskonstrukt, samt en samlet motivasjonsscore basert på scorene fra konstruktene. I tillegg viser den resultater fra regresjon mellom hvert konstrukt og FCI post-score. Den oppgitte p-verdien tilhører stigningstallet.

**Tabell 4.4:** Resultatene fra multippel lineær regresjon på alle motivasjonskonstruktene og post score fra FCI. Vi søker en modell på formen  $y = ax_1 + bx_2 + cx_3 + dx_4 + ex_5 + f$ , der post score fra FCI =  $y$ , Selvrealisering og identitet =  $x_1$ , Interesse og glede =  $x_2$ , mestringsforventning =  $x_3$ , nytteverdi =  $x_4$  og kostnad =  $x_5$ .

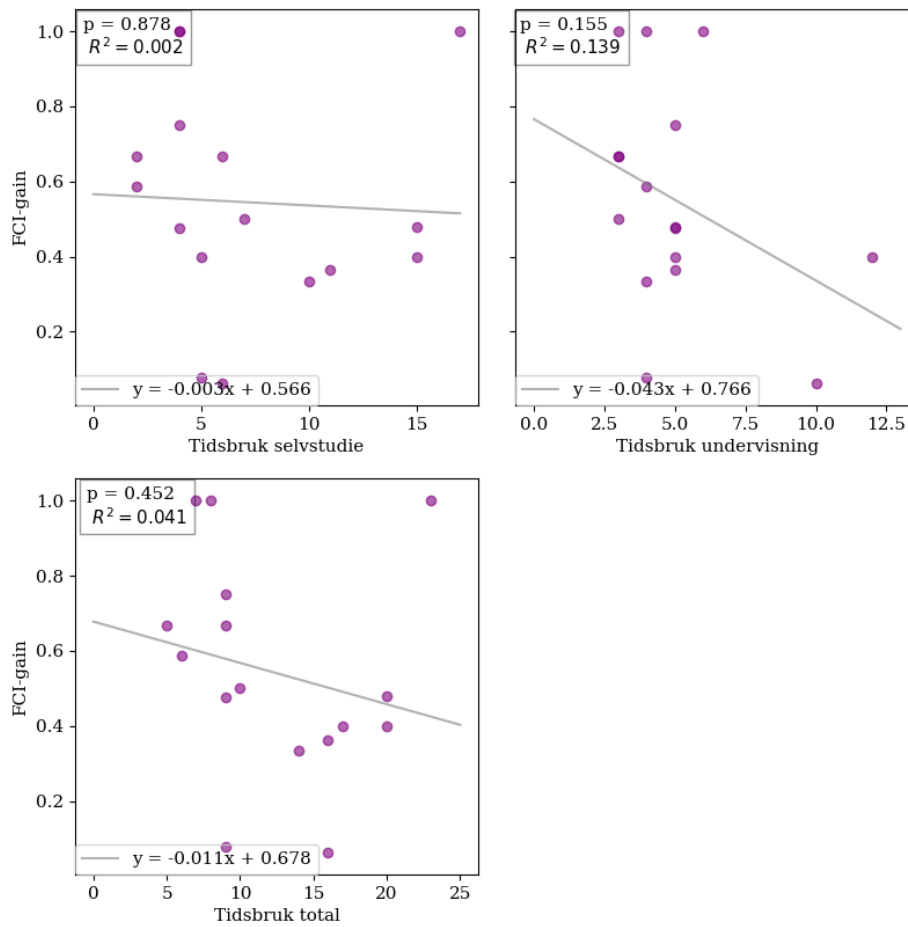
	a	b	c	d	e	f	$R^2$ (just.)
Koeff. verdi	0.327	-0.277	0.423	-0.388	0.246	0.748	0.085
p-verdi	0.474	0.430	0.151	0.298	0.449	0.016	

I figur 4.6 ser vi at det i de fleste tilfeller ikke ser ut til å være noen direkte observerbar relasjon mellom FCI post-score og motivasjonskonstruktene, men også her ser det ut som mestringsforventning og resultatet fra FCI kan korrelere til en viss grad. Dersom vi undersøker regresjonsanalysen som også har blitt plottet i figur 4.6 sammen med datapunktene ser vi at det ikke er noen signifikante resultater med et signifikansnivå på 0.05, men at mestringsforventning er veldig nært med en p-verdi på 0.053. De andre konstruktene produserer ikke resultater nær signifikans. Fra resultatene av multippel regresjon der motivasjonskonstruktene er uavhengige variable, og FCI post-score er målvariabelen, kommer det ikke ut noen signifikante resultater, men også her er dimensjonen med mestringsforventning nærmest.

### 4.3.3 Tidsbruk og læringsutbytte

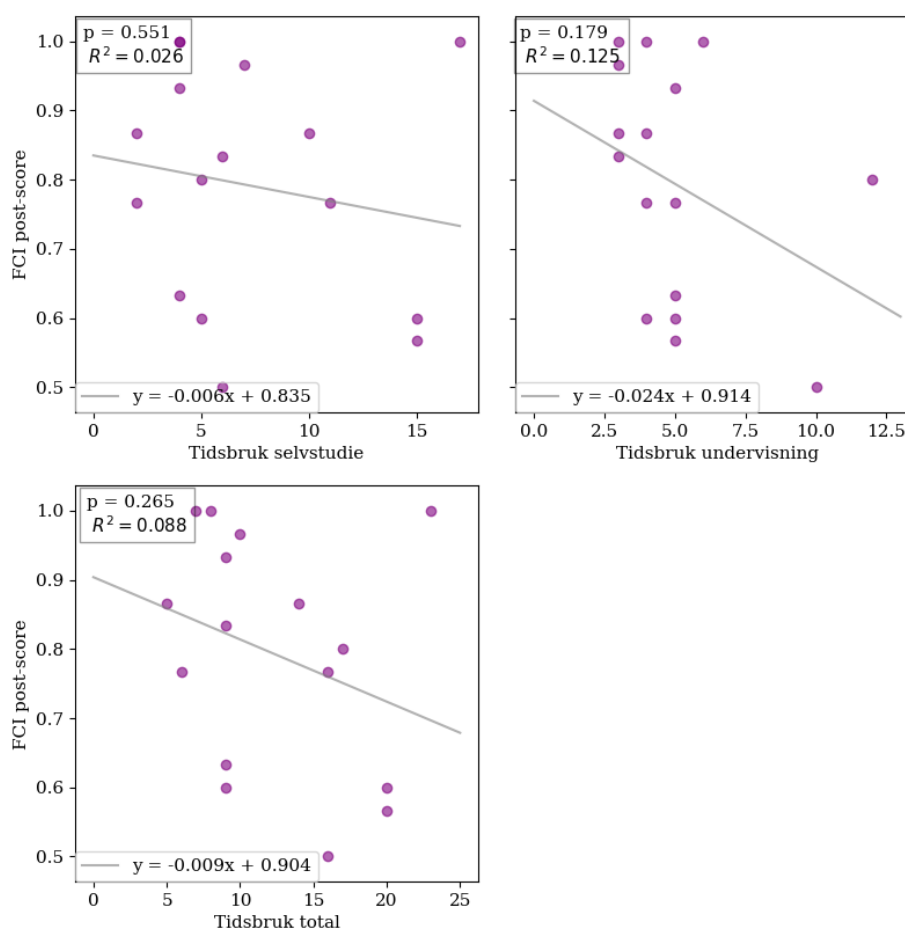
I denne underseksjonen vil vi presentere resultatene som har kommet ut av analysen som har blitt gjort på konseptuell forståelse fra FCI og tidsbruk. Det blir presentert plots på FCI-gain og tidsbruk, samt FCI post-score og tidsbruk. Disse finnes i henholdsvis figur 4.7 og 4.8. Resultater fra lineær regresjon på det samme datamaterialet finnes også i figur 4.7 for tidsbruk og FCI-gain, og i figur 4.8 for tidsbruk og FCI post-score.





**Figur 4.7:** Figuren viser FCI-gain og tidsbruk for hver enkelt student. Tidsbruk er delt inn i tidsbruk i undervisning, tidsbruk på selvstudie, og disse er også lagt sammen til et totalt tidsbruk. I tillegg viser figuren resultater fra regresjonsanalyse mellom disse. p-verdien som oppgis er p-verdien til stigningstallet.

Fra regresjonsresultatene i figur 4.7 ser vi at det ikke er noen signifikante resultater for lineær regresjon mellom tidsbruk og FCI-gain ved et signifikansnivå på 0.05.



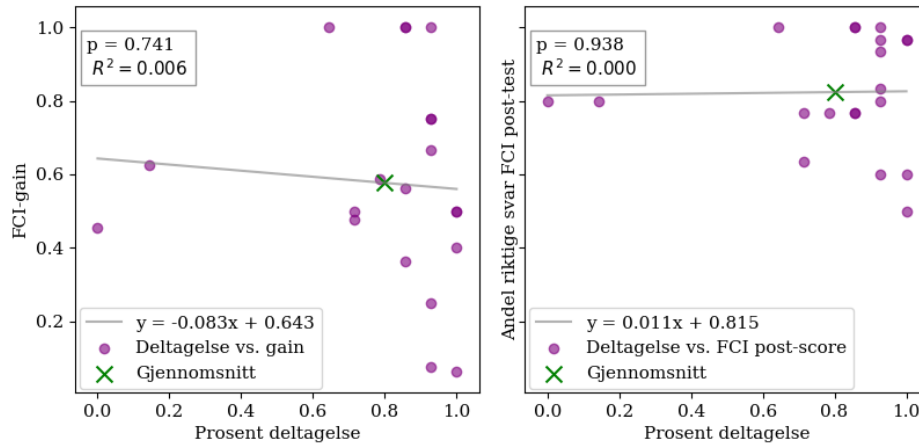
**Figur 4.8:** Figuren viser FCI post-test score og tidsbruk for hver enkelt student, samt resultater fra lineær regresjon på disse. Den oppgitte p-verdien er p-verdien til stigningstallet.

Resultatene fra lineær regresjon i figur 4.8 forteller at det ikke er noen signifikante resultater for lineær regresjon mellom tidsbruk og FCI post-score for p-verdiene gitt et signifikansnivå på 0.05.

#### 4.3.4 Deltagelse og læringsutbytte

Her vil det bli presentert resultater som har kommet ut av analysen som har blitt gjort på konseptuell forståelse målt ved FCI og deltagelse. Et spredningsplot av deltagelse mot økning i konseptuell forståelse målt ved FCI-gain finnes til venstre i figur 4.9. Lineær regresjon gjort på samme datamateriale finnes det resultater fra i samme figur. Tilsvarende plot for deltagelse mot slutt-score i FCI

finnes til høyre i figur 4.9, sammen med tilhørende lineær regresjon.



**Figur 4.9:** Til venstre viser figuren FCI-gain og deltagelse for hver enkelt student, samt lineær regresjon. Til høyre viser figuren FCI post-test score og deltagelse for hver enkelt student, samt lineær regresjon. Det er p-verdien til stigningstallet som er oppgitt.

Ut i fra plottene i figur 4.9 ser det ikke ut til å være noen korrelasjon mellom deltagelse og hverken FCI-gain eller FCI post-score. Dette bekreftes av regresjonsanalysen.



# /5

## Diskusjon

I introduksjonen formulerte jeg tre forskningsspørsmål jeg ønsket å svare på:

1. Hva kjennetegner fysikkstudentenes motivasjon?
2. Hvordan er motivasjonen til studentene som tok FYS-0100 i 2020, sammenlignet med studentene som tok tilsvarende emne i 2019?
3. Hva er sammenhengen mellom konseptuelt læringsutbytte, motivasjon, tidsbruk, og deltagelse etter endt digital undervisning i et introduksjons-emne i fysikk?

I samme rekkefølge vil disse spørsmål blir forsøkt besvart. Resultatene fra IMPEL vil kunne si noe om studentenes generelle motivasjon for faget etter endt undervisningsperiode. Ved å sammenligne disse resultatene med resultater fra IMPEL året før, har jeg sjekket om den generelle trenden i motivasjonsutviklingen for studenter er den samme som for studentene i dette emnet. Før jeg gjør dette vil jeg dog se på reliabilitetsanalysen som har blitt utført på IMPEL. Videre har jeg ønsket å se om motivasjonen de har endt opp med har en sammenheng med den konseptuelle forståelsen de har tilegnet seg. Resultatene fra FCI vil kunne fortelle oss noe om undervisningen har fungert for de aktuelle studentene med hensyn på læringsutbytte i form av konseptuell forståelse. I forsøk på å tallfeste eventuelle sammenhenger har det blitt gjort lineær regresjon. Resultatene fra denne vil bli diskutert her. Med tanken om å bestemme om det er undervisningen som har ført til en økning i konseptuell forståelse har også FCI-resultatene blitt sett i sammenheng med deltagelse

i fellesundervisning. Resultatene fra dette vil bli diskutert her. Det har også vært ønskelig å sjekke om selvrapportert tidsbruk i de andre delene av undervisningen kan ha vært viktig, samt selvrapportert tidsbruk på faget utenom undervisning. Disse resultatene vil også bli diskutert.

Etter en slik kronologisk gjennomgang skal analysen belyses av relevant teori fra teoridelen. Til slutt vil jeg drøfte denne oppgavens begrensninger og svakheter, og så løfte blikket mot videre arbeid og mulige utvidelser.

## 5.1 Kjennetegn på fysikkstudentenes motivasjon i 2020

I denne seksjonen av diskusjonen vil jeg ta for oss fysikkstudentenes motivasjon etter endt undervisning i 2020. I underseksjon 5.1.1 vil jeg diskutere reliabilitetsanalysen som har blitt gjennomført på IMPEL, og i underseksjon 5.1.2 vil jeg diskutere verdiene for studentenes motivasjon som ble målt ved bruk av IMPEL høsten 2020.

### 5.1.1 Reliabiliteten i IMPEL

Siden IMPEL-undersøkelsen er forholdsvis ny har jeg gjort en analyse av reliabiliteten i hvert konstrukt i Eccles-modellen. Ved å bruke Cronbach's  $\alpha$ , som beskrevet i teoridelen, har jeg kunnet sjekke intern konsistens som et reliabilitetsmål for hver enkelt indeks som måler et konstrukt. Som beskrevet i teoridelen, må denne verdien være over 0.7 for at reliabiliteten skal være akseptabel. Som en ser av tabellene i appendiks A.1 og A.2 har den interne konsistensen vært god, både for det fulle datasettet på IMPEL og det reduserte, som var filtrert på at deltagerene også hadde deltatt på FCI før- og ettertest. Med hensikt å gjøre indeksene til bedre mål på konstruktene ble også hver indikator (spørsmål) sjekket. For at det skal bidra positivt til indeksen må det, som kjent fra tidligere i oppgaven, ha en korrelasjon på 0.4 mot resten av indeksen. Den enkleste måten å plukke ut slike spørsmål på var å sjekke hvilke som gjorde at Cronbach's  $\alpha$  sank. Det var flere indikatorer som førte til dette, men de fleste av dem hadde akseptabel korrelasjon mot resten av indeksen. Det var spesielt to indikatorer som stakk seg negativt ut.

Fra indeksen som måler konstruktet *Selvrealisering og identitet* ble indikatoren «*Det er viktigere for meg å få gode karakterer i fysikk enn i andre emner.*» fjernet. Dette var den eneste indikatoren som, for denne indeksen, satte fysikk opp mot andre emner. I så måte skilte den seg ut, de andre indikatorene måler hvordan

deltagerene identifiserer seg med fysikkfaget og hvordan det å ta fysikk hjelper dem å realisere den personen de føler de er, mens denne indikatoren måler om fysikkfaget er viktigere for dem enn andre ting. Derfra argumenteres det for at denne indikatoren ikke passer inn i indeksen fordi den ikke måler det samme som resten av indikatorene. Det gir derfor mening at studentene skal avgi et svar på denne som er forskjellig ifra svarene på resten av indikatorene.

Den andre indikatoren tilhører indeksen som måler konstruktet *Kostnad*. Indikatoren er «*Jeg følte at fysikk krevde veldig mye tid og arbeid.*» Forskjellen på denne indikatoren og de andre indikatorene i indeksen er at denne indikatoren ber studentene tenke tilbake på tiden før undervisning, og evaluere hvordan de følte kostnaden ved å ta fysikk var den gang. I så måte måler den ikke det samme som de andre indikatorene som måler kostnaden de føler de har ved å ta fysikk på testtidspunktet.

Fjerning av disse indikatorene ga en mere reliabel undersøkelse for den gitte utvalgsgruppen målt ved Cronbach's  $\alpha$ . Etter fjerning hadde indeksene Cronbach's  $\alpha$  på mellom 0.766 og 0.911.

### 5.1.2 Studentenes motivasjon

I denne underseksjonen vil vi diskutere kjennetegn ved studentenes motivasjon i 2020.

Fra figur 4.1 ser vi at gjennomsnittsstudenten uttrykker delvis enighet med at de ser en god *nytteverdi* i faget, med en score på 0.78. Som presentert i tabell 3.2 tilsvarer en score på 0.75 delvis enighet. Når det gjelder *mestringsforventning* er den ganske lav, med en score på 0.43. En score på 0.5 tilsvarer å hverken være enig eller uenig, og den gjennomsnittlige student i 2020 trekker fortsatt mot å hverken være enig eller uenig. Den gjennomsnittlige scoren på *kostnad* er 0.59. Det betyr at den gjennomsnittlige student er hverken enig eller uenig i at kostnaden ved å studere fysikk er høy. Videre er gjennomsnittsstudenten delvis enig i at hen identifiserer seg med fysikkfaget og at det bidrar til hens selvrealisering gjennom en verdi på 0.71 på *selvrealisering og identitet*. Verdien på *interesse og glede* er på 0.73.

Eccles og Wigfield (2002) sier at mestringsforventning, til forskjell fra de andre konstruktene, fokuserer på studentenes tro på at de klarer å gjennomføre en oppgave. De andre konstruktene sier noe om hvorfor studentene velger (eller ikke velger) å engasjere seg i oppgaven. Det vil si at mestringsforventning er den eneste blant dem som er prestasjonsrettede. I vårt tilfelle betyr dette at studentene ser større grunn til å engasjere seg i læringsmaterialet, enn de har tro på at de klarer det. Til slutt kan det sies at *nytteverdi*, *selvrealisering* og

*identitet og interesse og glede* har veldig like verdier, men

Ramberg (2006) skriver at mange elever i videregående skole opplever realfagene som vanskelige fag der du ikke får noe gratis. Tanken om at realfagene er vanskeligere enn andre fag forsterkes av at elevene får realfagspoeng: ekstra poeng som forbedrer karaktersnittet ved endt skolegang. Elevene ser på disse poengene som belønning for å ta ekstra krevende fag – fag det koster ekstra mye å gjennomføre (Ramberg, 2006). Fag som engelsk trekkes fram som fag der du får noe gratis ved å for eksempel se på TV eller film. Fremmedspråklige fag og realfagene trekkes fram som fag der du ikke gjør det (Ramberg, 2006). Har fysikkfaget og andre realfag i videregående skole i så måte feilet? De søker å forklare fenomenene som er over alt rundt oss, og vi omgir oss sådan konstant med deres fagområde. Er læringsmaterialet i realfagene så abstrakt at elevene distanserer det fra den virkelige verden? Sjøberg (2014) er inne på mye av det samme – at elevene danner seg to verdener: den virkelige verden der hverdagslige fenomener forklares av hverdagsforestillinger, og skoleverden der vitenskapelige ideer og teorier regjerer. Kanskje kan et fokusskifte mot konseptuell kunnskap være behjelpelig i å løse denne dissonansen mellom det realfagene søker å være, og det de oppleves som for elevene. Det kan i allfall være del i en forklaring på hvorfor studentene som begynner på et fysikkstudie synes at kostnaden er større enn om de skulle studere noe annet, og at grunnen til at de er motivert for studiet heller handler om at de identifiserer seg med fagområdet, ser stor nytteverdi i det, og interesserer seg for å arbeide med det.

Ved å innføre en førtest for motivasjon kunne en ha identifisert studenter som føler en høy kostnad vedrørende studiet og gjennomført intervjuer med disse studentene for å bedre fastslå hva som gjør at de assosierer studiet med høy kostnad. Resultatene fra IMPEL 2019 indikerer også at høy kostnad er en vedvarende trend på tvers av årskull. I en forlengelse av denne studien kunne en altså ha gjennomført en førtest på IMPEL og identifisert om høy kostnad virket å være en trend også der. Dersom dette skulle fremkomme ville det vært interessant å gjennomføre intervjuer for å fastslå hvorfor studenter assosierer fysikkstudiet med høy kostnad.

## 5.2 Motivasjon etter ett år med koronapandemi

Her vil vi diskutere resultatene fra seksjon 4.2 i resultatkapitlet. Det betyr at vi først kort tar for oss studentenes motivasjon i 2019, før vi ser på hvordan denne motivasjonen er sammenlignet med studentenes motivasjon i 2020 i underseksjon 5.2.1.



### 5.2.1 Motivasjon i 2020 sammenlignet med motivasjon i 2019

Vi ser at studentene i 2019 har verdier for motivasjonskonstruktene som ligger innenfor samme kategori som studentene i 2020, hvilket indikerer at tallverdiene er i samme sjikte. Det er også interessant å notere seg at verdiene for skjevhet og kurtose ikke avviker problematisk fra normalfordelingen. Det konstruktet som har endret seg mest er mestringsforventning, med en endring fra 0.49 i 2019 til 0.43 i 2020. Formen på fordelingene i figur 4.2 og 4.1 ser ikke ut til å ha endret seg mye. For å finne ut om fordelingene er forskjellige kunne kanskje for eksempel en Mann-Whitney U test vært brukt.

Fra Kunnskapsdepartementet (2021), som rapporterer resultatene fra studiebarometeret som publisert av NOKUT (2021), hvis undersøkelser ble gjennomført høsten 2020, heter det at studentenes motivasjon på landsbasis har sunket noe fra 2019 til 2020. For å sjekke om dette også er tilfelle i denne studien har det blitt gjort en *t*-test som tester om gjennomsnitt er signifikant forskjellig fra hverandre. En *t*-test forutsetter normalfordelte data, ikke minst for små datasett (Walpole mfl., 2007). I så henseende er det relevant at konstruktene i 2019 og 2020 ikke fraviker problematisk fra normalfordelingen ved skjevhet og kurtose. Resultatet fra *t*-testen indikerer ikke en signifikant endring i noen tilfeller ved et signifikansnivå på 0.05.

Det betyr at man ikke har evidens for å som tilsier at studentene i 2020 har endret motivasjon, sammenlignet med studentene året før. Dette på tross at de har levd med digital undervisning i nærmere ett år. Siden Kunnskapsdepartementet (2021) skriver at motivasjonen til studentene generelt sett har sunket under koronapandemien kunne det vært interessant å finne ut om det er en sammenheng mellom undervisningen og fraværet av nedgang i motivasjon. Dette ble forsøkt gjort gjennom å sammenligne deltagelse og motivasjon, som presentert i slutten av underseksjon 4.2.2. Fra spredningsplottene var det ikke en observerbar sammenheng, og det ble besluttet å ikke gå videre med lineær regresjon. Dersom en hadde visst hvilken motivasjon studentene entret undervisningen med kunne en muligens ha sagt noe om hvordan studentenes motivasjon utviklet seg ved digital undervisning med Peer Instruction. I så måte ville dette vært mere beskrivende for om Peer Instruction kan ha hjulpet motivasjonen i digital undervisning, eller om studentene kom til undervisning med høyere motivasjon enn de hadde ved endt semester.

For underviseren er det interessant at både kostnad og mestringsforventning er lave på tvers av årskull. Det kan være interessant å iverksette tiltak for å bedre disse sidene ved studentenes motivasjon. Det er også interessant at motivasjonen med fysisk undervisning i 2019 ikke er signifikant forskjellig fra motivasjonen med digital undervisning med Peer Instruction i 2020. Det tyder

på at det ikke er noen evidens for at Peer Instruction i en digital undervisningsform gjør skade på studentenes motivasjon.

En lignende studie, enten både før og ettertest med IMPEL eller bare en enkelt kartleggingstest av motivasjon for fysikk eller naturfag, kan bli gjennomført også ved Peer Instruction i videregående, kanskje også i ungdomsskole. Det kan også være interessant å gjøre kartleggingsstudier, lignende det Anwer, Iqbal og Harrison (2012) har gjort i Pakistan, for å kartlegge om det er noen forskjell i motivasjonen til elever i urbane og landlige områder. En kan ta dette videre og forsøke å se på studenters eller elvers motivasjon i forskjellige landsdeler. Det kan enten være gjennom å gjøre kartlegging ved forskjellige skoler rundt om i landet eller ved forskjellige universiteter. En veldig dyp undersøkelse kunne ha tatt for seg både urbane og landlige områder over hele landet, og slik gjort en gjennomgående studie av motivasjon på landsbasis. Dersom en skulle benytte seg av en undersøkelse som IMPEL i grunnskolen, måtte den sannsynligvis ha vært modifisert slik at den undersøkte motivasjonen i naturfag, ikke ren fysikk – da fysikk ikke er et eget fag i grunnskolen. Dette er interessant videre arbeid som kan nå andre grupper enn de IMPEL dekker.

### **5.3 Sammenheng mellom motivasjon, tidsbruk og deltagelse, og konseptuell forståelse etter endt undervisning**

I denne delen av diskusjonen vil resultatene fra seksjon 4.3 bli diskutert. Jeg vil først ta for oss læringsutbyttet studentene har hatt fra digital undervisning i underseksjon 5.3.1. Deretter vil jeg ta for oss sammenligninger mellom motivasjon og konseptuelt læringsutbytte i underseksjon 5.3.2, før jeg ser på en sammenligning mellom deltagelse i undervisning og konseptuelt læringsutbytte i underseksjon 5.3.3. Til slutt vil jeg se på sammenligningen av tidsbruk og konseptuelt læringsutbytte i underseksjon 5.3.4.

#### **5.3.1 FCI som evaluering av undervisningen**

I underseksjon 4.3.1 fremkom det at effektstørrelsene regnet ut på pre- og post-test fra FCI indikerer at effekten av undervisningen har vært høy. Vi observerer for eksempel en Cohen's  $d$  på 0.88. I tillegg observerer vi at fordelingen av gjennomsnitt har blitt flyttet til høyre i figur 4.4, hvilket bekrefter at utviklingen har vært positiv. Dersom en legger til at alle, men unntak av én student, har hatt positiv utvikling er det rimelig å tro at den konseptuelle forståelsen til majoriteten har økt. Dette er på ingen måte uforventet, men det er allikevel et

positivt resultat.

Det kan være interessant for underviseren i fysikk å ta med seg at studentene oppnådde et læringsutbytte som generelt forventes av denne formen for undervisning. I lys av at Felder og Brent (1996) sier at det ikke er noen automatikk i at læringsutbyttet blir større bare fordi en endrer undervisningsformen og at underviseren gjerne må bygge seg opp erfaring for å skape økning i læringsutbytte er det positivt for underviser å ta med seg at studentene har oppnådd et tilfredsstillende læringsutbytte på første gjennomføring.

Det har blitt nevnt at FCI har en verdi som diagnostisk test, og at den undersøker hver dimensjon av kraftbegrepet i forskjellige kontekster. I lys av diSessa (1993) og Sjøberg (2014) kan det, både i forskningsøyemed og i forbindelse med diagnostisering av hverdagsforestillinger, være interessant å undersøke hvilke kontekster som aktiverer kunnskap, og hvilke som ikke gjør det. I undervisningssammenheng gjør det at en kan legge opp undervisningen slik at en får brukt riktig kontekst for å skape en kognitiv konflikt. I Peer Instruction kan en gjerne stille spørsmålet i en ConcepTest på en slik måte at problemstillingen skaper en kognitiv konflikt, slik det ble diskutert i seksjon 2.3 av teoridelen (Brekke, 2002). I forskningsøyemed kan dette være interessant for å sjekke om det er noen kontekster som skaper større utfordringer enn andre.

I en videreutvikling av denne studien kunne det ha vært interessant å implementere FCI med selvtillit – det vil si at for hvert svar så skal testdeltageren vurdere hvor sikker hen er på at det valgte svaralternativet er korrekt. Dette kan være interessant av flere grunner. Som kjent fra seksjon 2.4 i teoridelen består svaralternativene i FCI av ett korrekt svar og flere ukorrekte svar. De ukorrekte svarene representerer utbredte hverdagsforestillinger. Det ville vært interessant å se om de som svarer feilaktig på spørsmål i FCI er like selvsikker på svaret sitt som de som svarer korrekt. Dersom FCI brukes som en diagnostisk test er dette definitivt verdifull informasjon fordi det kan gi en pekepinn på om det feilaktige svaret skyldes en hverdagsforestilling, eller om det kommer av ren gjetning. Selvtillit kan også figurere som tilleggsinformasjon ved bruk av FCI til evaluering av undervisning. En kan få informasjon om de som har valgt korrekt svar er mere sikker på det korrekte svaret etter undervisningsperioden var ferdig – at det i mindre grad er gjetning som har ført til rett svar, men reell forståelse. På samme måte kan en vurdere de feilaktige svarene. En kan sjekke om det etter undervisning er hverdagsforestillinger som fører til ukorrekt svar, eller om det er gjetning.

### 5.3.2 Motivasjon og konseptuelt læringsutbytte

I denne underseksjonen av diskusjonen vil jeg diskutere resultatene som har kommet fram av sammenligning av *konseptuelt læringsutbytte* og motivasjonskonstruktene *kostnad*, *nytteverdi*, *selvrealisering og identitet*, *interesse og glede*, og *mestringsforventning*. Disse har blitt plottet mot hverandre, og det har blitt gjort regresjonsanalyser – både enkel lineær regresjon og multippel lineær regresjon – med motivasjonskonstruktene som uavhengige variable. Målvariabelen (den avhengige variabelen) har vært normalisert FCI-gain som et mål på økning i konseptuell forståelse, og post-score fra FCI, som et mål på den konseptuelle forståelsen studentene har oppnådd etter endt undervisning. Resultatene fra multippel regresjon med alle konstruktene som uavhengige variable ga i begge tilfellene ingen signifikant korrelasjon for noen av konstruktene, med laveste p-verdi for en av deres koeffisienter på 0.151, og et signifikansnivå på 0.05. Resultatene fra enkel lineær regresjon er i de fleste tilfeller heller ikke statistisk signifikant. Det produseres i de fleste tilfeller p-verdier for stigningstallene som er nokså store. Unntaket er p-verdien til stigningstallet i den lineære regresjonen med mestringsforventning som uavhengig variabel og FCI post-score som målvariabel. Denne p-verdien er på 0.053, og er heller ikke signifikant med et signifikansnivå på 0.05.

Jo nærmere en er signifikansnivået, jo større er sjansen for å gjøre feil av type to. Det vil si at en feilaktig beholder nullhypotesen om at det ikke er en signifikant effekt selv om det er det. Det vil også si at testens statistiske power blir lavere – som vil si at sjansen for å finne en effekt dersom den er til stede synker. Lav utvalgsstørrelse gjør også at statistisk power synker. Det skal en større effekt til for å få stor statistisk power dersom en har et lite utvalg. I dette tilfellet har vi en forskjell på 0.06, eller alternativt 6% som tilsvarer en effektstørrelse gitt ved Cohen's  $d$  på 0.31 – altså en liten effektstørrelse. Signifikansnivået er på en måte arbitrært, det bestemmes av testens premisser. I en del tilfeller kan det være alvorlig å forkaste en korrekt nullhypotese, og da setter gjerne signifikansnivået så lavt som 0.01. Signifikansnivå på 0.1 har også blitt brukt innenfor andre forskningsfelt (Walpole mfl., 2007). Når effektstørrelsen dog er så lav som den er her er det ikke veldig stor konsekvens av type 2 feil.

I seksjon 2.5 i teoridelen skrev jeg at Hestenes mfl. (1992) rapporterer at to klasser med nokså like elevgrupper, både i form av intelligens og matematisk kompetanse, oppnådde helt forskjellige gains (35% vs 5%). Forskjellen på de to klassene, mente Hestenes mfl. (1992), samt instruktørene av de to klassene – var motivasjon. Den ene klassen var motiverte og satte igang med aktiviteter på egenhånd, der den andre trengte kontinuerlig tilsyn. Uten at dette er tallfestet så antyder det en korrelasjon mellom motivasjon og gain i FCI. Lillemyr (2007) indikerer mye av det samme, og skriver at dersom en læringsinstitusjon skal lykkes, blant annet med hensyn på læringsutbytte, må

den ta motivasjon på alvor. Derfor skulle en kanskje forvente noe mere signifikante resultater enn det som fremkommer her. Eccles og Wigfield (2002) skriver at de som svarer positivt på spørsmål om mestringsforventning presterer bedre, og er mere motiverte til å velge utfordrende oppgaver. Bandura (1997) skriver også at høy mestringsforventning har forutsett påfølgende gode akademiske prestasjoner. Vi så i underseksjon 5.1.2 at mestringsforventning skilte seg fra de andre konstruktene ved at det hadde lavere score. Eccles og Wigfield (2002) skriver at mestringsforventning skiller seg fra de andre motivasjonskonstruktene ved at det er prestasjonsrettet – det handler om hvor stor tro en har på at en kan prestere noe. Resten av motivasjonskonstruktene sier noe om hvor stor grunn en har til å faktisk engasjere seg i læringsmaterialet. I så henseende er det kanskje ikke så rart at mestringsforventning har vært nærmere å forutse påfølgende prestasjoner enn de andre motivasjonskonstruktene.

Både Bandura (1997) og Eccles og Wigfield (2002) skriver at mestringsforventning har *forutsett* gode resultater. I så måte er ikke premisset for denne studien de samme som for det Bandura (1997) og Eccles og Wigfield (2002) beskriver. Mestringsforventning i denne studiens kontekst måler ikke mestringsforventningen de hadde før de tok emnet som førte til resultatene på konseptuelt læringsutbytte fra FCI. Det måler heller mestringsforventningen de har dannet seg på bakgrunn av disse og andre erfaringer – og som sier noe om hvordan de forventer å prestere i fremtidige emner. I så henseende hadde en førtest på motivasjon, på samme måte som for FCI, vært et verdifullt verktøy for å knytte mestringsforventningen studentene entret undervisningen med opp mot prestasjon eller utvikling av konseptuell forståelse. Dersom en hadde både en førtest og en ettertest på motivasjon, kunne en beregnet en effektstørrelse. Dette ville banet vei for en studie som sammenlignet utviklingen av studentenes motivasjon med for eksempel konseptuell forståelse, men også om en interaktiv undervisningsform som Peer Instruction kan føre til forbedret motivasjon i seg selv. Som omtalt i underseksjon 2.5.5 i teoridelen har det blitt gjennomført studier, både innenfor matematikk og naturfag, som beskriver en sammenheng mellom konseptuell forståelse og motivasjon (Mansukhani, 2010; Yen mfl., 2010). T. Gok (2012) har også gjort en studie på utvikling av motivasjon over en periode med undervisning i Peer Instruction og fant ikke signifikant utvikling i noen motivasjonskonstrukt, men diskuterer at det er mulig Peer Instruction ikke var i bruk lenge nok til å utrette en signifikant forandring. Allison (2012) har gjort en lignende studie på åttendeklassinger etter Peer Instruction ble tatt i bruk, og fant noe økning i motivasjon. Det later dog ikke til å være en sterk forskningsbasert konsensus om at Peer Instruction øker motivasjonen, eller at konseptuell forståelse og motivasjon samvarierer. En studie over et helt semester med Peer Instruction i et introduksjonsemne i fysikk, der motivasjon blir målt både før og etter undervisning kan være et verdifullt bidrag til dette forskningsfeltet som synes å ha behov for fordykning.

### 5.3.3 Deltagelse og konseptuelt læringsutbytte

I denne underseksjonen vil jeg tolke resultatene fra underseksjon 4.3.4 i resultatdelen. Målet med å ta med deltagelse i studien har vært å forsøke å se om det er en sammenheng mellom deltagelse i undervisningen og hvilket læringsutbytte studentene har tilegnet seg.

I regresjonsanalysen med normalisert FCI gain som målvariabel og deltagelse som uavhengig variabel får vi en p-verdi på 0.741. Det forteller oss at det ikke er en signifikant korrelasjon med et signifikansnivå på 0.05. I regresjonsanalysen med FCI post-score som målvariabel og deltagelse som uavhengig variabel får vi en p-verdi på 0.938. Dette forteller oss at det heller ikke her er en signifikant korrelasjon.

Deltagelse som konstrukt er ikke klart definert. Det ble vurdert om en skulle legge vekt på hvor aktiv studentene var i gruppesamtalene, men det ble vanskelig å rettferdiggjøre måling av taletid – det er en forskjell på hvor mye taletid hver enkelt får i en gruppe der 2 som oftest deltar, kontra en gruppe der 4-5 som oftest deltar. Å definere deltagelse som «konstruktiv deltagelse» ble også vurdert – at studentene måtte være til stede og ha sagt noe konstruktivt. Det var flere problemer med dette. Det ene er at «konstruktivt» er ganske høytflyvende. Det kan hende at studenten selv opplevde noe hen sa som konstruktivt, mens jeg ikke gjorde det (eller omvendt). På grunn av datamaterialets natur var det også en utfordring å peke på hvem som snakket og hvem som sa hva. Med datamaterialet menes opptak av samtalene studentene i mellom i Microsoft Teams – som jeg hadde tilgang til, men ikke benyttet meg av i studien. I disse opptakene, spesielt i de gruppene der studentene ikke brukte kamera, viste det seg veldig vanskelig å skille ut hvem som snakket. Til slutt ble det besluttet å kun se på om studentene hadde vært tilstedeværende i gruppesamtalen. Dette var både lett å sjekke gjennom møtefunksjonen i Microsoft Teams, og det følte rettferdig. Datamaterialet ble registrert manuelt ved å skrive ned hvilke studenter som hadde deltatt i hver enkelt undervisningsøkt. Noe av tanken her var, ved å registrere deltagelse som tilstedeværelse, at studentene fortsatt kan få fordelene av å få konseptene forklart av sine medstudenter – som ifølge Mazur (1997) er gunstig, og at det derfor kan være verd å registrere disse også.

Det som har blitt målt er altså deltagelse i form av tilstedeværelse. I en utvidelse av dette kan det være relevant å utvide denne variabelen. Kanskje det skal være et krav om at de har vært aktiv deltagende i diskusjoner, eller kanskje en kan måle taletid. Dersom en gjør begge deler kan det være rom for å sammenligne resultatene fra de to. Et resultat fra en slik analyse kan også potensielt avsløre hva som bidrar til at majoriteten har stort læringsutbytte. Dersom det skulle vise seg at det er de med mest taletid som har størst læringsutbytte, kan det gjøres forsøk på tiltak som skal tilrettelegge for at alle skal få taletid. Det virker

relevant å forsøke å finne ut av hvilke deler av undervisningen som fører til læring, all den tid det er en del av de som har deltatt på alle eller nesten alle forelesningene som har dårligst resultat på FCI, både i form av normalisert gain, men også ligger rundt og under entry threshold fra Hestenes og Halloun (1995) på 0.6, som forteller at de ikke har den konseptuelle forståelsen til å ta til seg viderekomne konsepter i fysikk.

I et sosiokonstruktivistisk læringsperspektiv gir det også mening å utvide deltagelse til å innbefatte interaksjoner med andre. Både det å forklare og få forklart kunnskap til andre som befinner seg i samme situasjon er verdifulle komponenter i det sosiale aspektet ved læring (Sjøberg, 2014).

### 5.3.4 Tidsbruk og konseptuelt læringsutbytte

I denne underseksjonen skal jeg diskutere resultatene fra underseksjon 4.3.3 i resultatdelen på tidsbruk og konseptuelt læringsutbytte. Med et signifikansnivå på 0.05 er det ingen signifikante resultater fra lineær regresjon og det kan ikke sies at det er en sammenheng mellom tidsbruk og læringsutbytte. Dersom en ser på figurene i underseksjon 4.3.3 ser en at det er flere studenter som har høyt tidsbruk i selvstudie, men både lavt normalisert FCI-gain og lav FCI post-test score. Dette tyder på at det er en del studenter som investerer mye tid i faget, men fortsatt ikke får det de ønsker ut av det.

At det er en slik avstand mellom tiden studenter rapporterer at de bruker på faget og læringsutbyttet de får ut er interessant for den som underviser fysikkemner med studentaktive undervisningsformer. Det kan være interessant å finne ut av hvilke studieteknikker som er utbredt blant de som ikke investerer veldig mye tid i faget, men fortsatt får stort læringsutbytte, og iverksette tiltak for å få flere til å prøve slike studieteknikker.

I en utvidelse av studien kunne det være interessant å legge til spørsmål i IMPEL for å kartlegge hva de bruker tid på i selvstudie. Dette ville gjort det enklere å si noe om hvilke arbeidsvaner som for majoriteten fører til lavt læringsutbytte i Peer Instruction. Dersom det for eksempel skulle vise seg at de som investerer mye tid i oppgaveregning gjør det dårlig i Peer Instruction, kan en vurdere om det har oppstått en dissonans med det konseptuelle som foregår i undervisning, og om en skal legge opp til at selvstudiet skal foregå på andre måter. Det samme gjelder i motsatt ende av skalaen. Dersom en observerer at enkelte studieteknikker er utbredt blant de som får mye ut av undervisningen kan en kanskje oppfordre til å prøve ut disse teknikkene.

Det er heller ikke sikkert at en lineær regresjon er det beste å anvende her. Det er tydelig at det er mange som har fått god konseptuell forståelse og legger

lite tid i selvstudie. Det kan godt hende at det er en litt annen type relasjon mellom tidsbruk og konseptuell forståelse. Hvis man ikke legger noe tid i å lære seg fagstoffet, vil man heller ikke lære noe – det er opplagt. Det kan dog hende at dersom en legger litt tid i faget så har en et stort hopp i læringsutbytte fra ingenting, noe resultatene også indikerer. Videre er det ikke nødvendigvis naturlig at læringsutbytte stiger like kraftig med økning i tidsbruk.

## 5.4 Studiens validitet

### 5.4.1 Begrepsvaliditet

Begrepsvaliditet handler om samsvarelsen mellom begrepet slik det er forankret i teori, og begrepet slik det måles. Det handler om en studie faktisk måler det den uttaler å måle – og i så måte operasjonaliserer det teoretiske begrepet på en vellykket måte. Et problem er ofte at operasjonaliseringen ikke kan gjøres ved et direkte mål – men at det teoretiske begrepet må måles gjennom forskjellige indikatorer.

Konseptuell forståelse måles gjennom svar på FCI, som måler forståelse av kraftkonseptet i Newtonsk mekanikk. Ifølge diSessa og Sherin (1998) og Sjøberg (2014) er det slik at konteksten har mye å si for om forståelse blir aktivert. Derfor kan en tro at et spørsmål som stilles i en gitt kontekst ikke nødvendigvis vil aktivere kunnskapen en student har, men heller aktivere en av hverdagsforestillingene i denne konteksten. Hestenes mfl. (1992) skriver at for å delvis kompensere for en slik falsk negativ, der studenten egentlig besitter den aktuelle konseptuelle kunnskapen, angripes hver dimensjon av kraftbegrepet fra flere vinkler. På grunn av at FCI er godt testet og validert (Hestenes & Halloun, 1995), og brukt til å måle FCI før undervisningsstart og etter undervisningsslutt, argumenteres det for at konseptuell forståelse er operasjonalisert på en god måte.

Motivasjonskonstruktene er målt av en rekke indikatorer som til sammen danner en indeks som skal representere hvert motivasjonskonstrukt. At indikatorene måler det de skal har blitt verifisert av eksperter på modellen til Eccles og Wigfield (2002) for motivasjon, som har dannet grunnlaget for IMPEL-undersøkelsen. I tillegg har det gjennom intervjuer blitt verifisert at testdeltagerne forstår hva de svarer på. En av studiens begrensninger er at motivasjonskonstruktene har blitt målt etter endt undervisning. Dette er kanskje aller mest aktuelt for konstruktet mestringsforventning, som forsøker å beskrive hvordan studentene forventer å gjøre det. Mestringsforventning operasjonalisert slik som i denne studien forteller noe om hvordan studentene forventer å gjøre det i fremtidige fysikkemner, ikke i emnet som har blitt studert.



I selvrapportert datamateriale er det også en usikkerhet knyttet til målingens nøyaktighet. En testdeltager vil gjerne overdrive når det gjelder svar der et utslag kan sees som positivt for en selv (Northrup, 1997). Dette kan sees i sammenheng med et bias, kjent fra sosialpsykologien som «the social-desirability bias». Dette biaset beskriver hvordan en person som deltar i en undersøkelse responderer på spørsmål på en måte som kan bli oppfattet gunstig av andre (Krumpal, 2013). I denne studien er nok dette mest aktuelt for tidsbruk, der studentene selv blir bedt om å anslå hvor mange timer i uka de bruker på de forskjellige delene av læringsprosessen.

Som diskutert kan det være flere måter å måle deltagelse på. I denne studien har det vært nok at de har vært til stede i gruppesamtalen i undervisningen. Et problem med dette er at en ikke får innsikt i noen grad av deltagelse. Hverken når det kommer til om de har deltatt samtaler med medstudenter om de aktuelle spørsmålene, men vi vet heller ikke hvor lenge de har vært til stede. En student som i løpet av undervisningsøkta har hoppet inn i gruppesamtalen, for så å hoppe rett ut igjen vil få registrert like mye tilstedeværelse som en som har vært til stede hele undervisningsøkta. Det er allikevel slik at tilstedeværelse, slik som det blir målt, blir målt på en nøyaktig måte siden det blir registrert i møteloggen til Microsoft Teams om en bruker har vært inne i et møte eller ikke.

### 5.4.2 Indre validitet

Indre validitet viser til graden av sikkerhet på at et kausalt forhold som blir testet er troverdig og ikke påvirket av andre faktorer eller variabler (Behi & Nolan, 1996).

Dersom rammebetingelsene endres i løpet av studien kan dette forringe studiens indre validitet (Behi & Nolan, 1996). Det ble dog tidlig kommunisert fra UiT at undervisningen høsten 2020 kom til å bli heldigital. Undervisningen ble også gjennomført med Peer Instruction fra første forelesning, og undervisningsopplegget var klart da studentene startet. Det ble ikke gjort endringer på dette i løpet av semesteret. Samtykkeskjema ble samlet inn tidlig, slik at de som hadde samtykket ble plassert i samme gruppe hele semesteret slik at en kunne studere dem, uten å inndirekte studere noen som ikke hadde samtykket. I denne studiens tilfelle gjelder det hovedsakelig deltagelse. De forble i disse gruppene hele semesteret.

FCI besto av en pre-test og en post-test. Ved å bruke den samme testen to ganger på denne måten risikerer en at resultatet på pre-test påvirker resultatet på post-test. Grunnen til dette er at testdeltagere husker hva det har svart på pre-test og føler de må være konsistent i svarene sine, og dermed avlegger

samme svar på post-test (Behi & Nolan, 1996). Som jeg har vært inne på tidligere ble dette forsøkt unngått ved at pre-test kom tidlig i semesteret og ikke ble vektlagt eller nevnt i undervisningen etter den var avlagt. Det ble i tillegg presisert at det ikke var en vurderingssituasjon, kun en diagnostisk test. Dette kan være en svakhet, men også en styrke. Det betyr at det ikke er en forskjell i hvordan konseptuell forståelse ble målt i løpet av studien.

En har i dette tilfellet ikke en kontrollgruppe der undervisningen har vært konvensjonell tavleundervisning. I stedet er det sammenligninger med andres resultater på FCI ved bruk av Peer Instruction og ved konvensjonell tavleundervisning i fysisk klasseromsundervisning som danner sammenligningsgrunnlaget. En svakhet ved denne oppgaven er altså at en ikke vet hvordan en student som får konvensjonell undervisning i en digital undervisningsform vil gjøre det i forhold til studentene som fikk undervisning med Peer Instruction.

Ekstremt lav score på pre-test fører ofte til større gain enn god score på pre-test, og en kan si at en test-retest studie favoriserer dårlige resultater på pre-test (Hestenes & Halloun, 1995; Hestenes mfl., 1992). Dette har blitt forsøkt unngått i FCI ved at det har blitt brukt normalisert gain, som tar hensyn til studentenes forbedringspotensiale (Hake, 1998).

Med bakgrunn i det overnevnte argumenterer jeg for at studiens indre validitet er god. Unntaket er at en ikke har en kontrollgruppe for digital undervisning, men heller må stole på andres resultater fra fysisk undervisning som sammenligningsgrunnlag.

### 5.4.3 Ytre validitet

Ytre validitet viser til i hvilken grad resultatene fra studien kan bli anvendt, eller generalisert, til andre sammenhenger og grupper.

Det første jeg skal se på er populasjonsvaliditet. Det vil si: hvor godt passer utvalget inn i den populasjonen en ønsker å si noe om? I et introduksjonsemne i fysikk på universitetsnivå er det naturlig at en har en overvekt av studenter som skal gå videre med fysikk, men også enkelte studenter fra andre studieretninger som tar emnet som valgfag. I utvalgsgruppen for alle som tok IMPEL (N=31) er det, som vi ser av tabell 3.1 i metodekapitlet, til stede til sammen fem studenter fra geologi, kjemi, matematikk og statistikk – som ikke er primære fysikklinjer. Av de som i tillegg til IMPEL også har tatt FCI (N=16) er det ingen av disse, kun studenter som går fysikklinjer. Det er denne utvalgsgruppen som studeres når jeg tar for meg sammenhenger mellom konseptuell forståelse og tidsbruk eller motivasjon. Denne gruppen er kanskje representativ for fysikkstudentene i den fulle gruppa, men en kan ikke si at

den er representativ for også de som ikke nødvendigvis skal videre med fysikk. Dette kan være et tilfelle av sampling bias, som blant annet omhandler at når deltagerene plukker seg selv ut til studien, så er det en gitt gruppe som har større sannsynlighet for å delta – for eksempel de som har en større interesse for fagområdet (Svartdal, 2015).

Netland, Sivertsen og Olufsen (2018) har gjort en studie på tilsvarende emne ved samme universitet, og rapporterer en gjennomsnittlig pre-test score på 0.59 på FCI. Studentene i dette emnet hadde en gjennomsnittlig pre-test score på 0.65, og har syv prosent bedre score på FCI enn studentene den gang. Ut over dette er det lite som kan generalisere til en populasjon utenfor den som blir studert. I en utvidelse av studien kunne en kanskje ha undersøkt demografiske variabler for å se mer på representativitet en i utvalgsgruppen.

Viten om at en blir studert kan påvirke hvordan en deltager i en studie oppfører seg – dette kalles for Hawthorne-effekten (Merrett, 2006). I denne studiens tilfelle er dette et problem som i stor grad er eliminert ved at jeg ikke har observert dem som sådan. Datamaterialet på konseptuell forståelse, tidsbruk og motivasjon har blitt samlet inn ved undersøkelser, mens datamateriale på deltagelse har blitt samlet inn etter endt undervisning. Jeg har altså ikke selv vært deltagende i undervisningen, og det er lite sannsynlig at min studie av resultater på test og undersøkelse, samt innsamling av deltagelse etter endt undervisning har påvirket noen deler av studien. Den samme argumentasjonen gjelder også for det vi kjenner som experimenter effect, som omhandler at den som gjennomfører studien kan ha adferd som påvirker studien (Kintz, Delprato, Mettee, Persons & Schappe, 1965).

Med utgangspunkt i manglende bakgrunn for å kunne si noe om populasjonsvaliditeten det vanskelig å anslå studiens grad av ytre validitet, men en kan håpe at videre undersøkelser innenfor fagfeltet i Norge vil bekrefte eller avkrefte dens generalitet.



# /6

## Avslutning

I arbeidet med denne oppgaven har jeg studert kjennetegn ved studentenes motivasjon og læringsutbytte. Forskningsspørsmålene jeg har forsøkt å besvare er:

1. Hva kjennetegner fysikkstudentenes motivasjon?
2. Hvordan er motivasjonen til studentene som tok FYS-0100 i 2020, sammenlignet med studentene som tok tilsvarende emne i 2019?
3. Hva er sammenhengen mellom konseptuelt læringsutbytte, motivasjon, tidsbruk, og deltagelse etter endt digital undervisning i et introduksjons-emne i fysikk?

### 6.1 Konklusjon

Gjennom å studere resultater fra IMPEL-undersøkelsen i 2020 er konklusjonen at studentenes mestringsforventning i fysikk er lavere enn nytteverdien de ser i fysikk, hvilken selvrealisering og identitet de har av det og interessen og gleden ved å jobbe med fysikk. Kostnaden de ser ved å ta emnet er også nokså stor - den inverse kostnaden er i samme størrelsesorden som studentenes mestringsforventning. Det vil si at kjennetegnene på deres motivasjon er at de ser større grunn til å engasjere seg i faget enn de har tro på at de mestrer det. Samtidig indikerer verdiene på motivasjonskonstrukter fra IMPEL 2020 at grunnen til å engasjere seg i større grad er basert på nytteverdien

studentene ser i fysikk, at det hjelper deres selvrealisering og identitet, og at det har interesse og glede av å arbeide med fysikk, enn den er basert på at de ser liten kostnad i å arbeide med fysikk. Ved å sammenligne motivasjonen til studentene i 2019 med motivasjonen til studentene i 2020 fant vi at det ikke var signifikant endring i noen av motivasjonskonstruktene i Eccles-modellen.

Studentenes konseptuelle læringsutbytte som operasjonalisert gjennom normalisert FCI-gain, post-test score på FCI, og Cohen's  $d$  blir vurdert til å være godt jamfør det som regnes som tilfredsstillende i fysisk undervisning. I så henseende konkluderes det med at Peer Instruction har bidratt til et godt læringsutbytte i et digitalt introduksjonsemne i fysikk. I sammenligningen av motivasjonskonstruktene og konseptuelt læringsutbytte, ved både normalisert FCI-gain og post-score fra FCI, fremkom det ingen signifikante resultater med lineær regresjon – hverken enkel lineær regresjon mellom motivasjonskonstruktene og konseptuelt læringsutbytte eller ved multippel regresjon mellom motivasjonskonstruktene og konseptuelt læringsutbytte. Mestringsforventning ved undervisningslutt og post-score fra FCI hadde dog en  $p$ -verdi på 0.053 nært signifikansnivået på 0.05, og det blir argumentert med at fordi mestringsforventning er det eneste konstruktet for forsøker å si noe om forventet prestasjon, gir dette resultatet mening. Det konkluderes også med at mestringsforventningen studentene hadde på begynnelsen av undervisningen ville være et bedre mål på hvordan de forventet å gjøre det i emnet.

Gjennom lineær regresjon på deltagelse og konseptuelt læringsutbytte konkluderes det med at det ikke fremkommer en signifikant korrelasjon. Det konkluderes også med at målet på deltagelse muligens kunne vært endret eller utbedret for å mere nøyaktig registrere deltagelse i undervisning. Det har heller ikke blitt funnet noen korrelasjon mellom tidsbruk i hverken selvstudie eller undervisning og konseptuelt læringsutbytte.

Fremover vil en kunne forvente at det kommer studier som vil sette resultatene fra denne studien i kontekst. Det gjelder både studier på studentenes læringsutbytte i digital undervisning under koronapandemien, men også studier på hvordan studentenes motivasjon har vært.

## 6.2 Implikasjoner for undervisning

For underviseren kan det være interessant at mestringsforventning og kostnad scorer gjennomgående dårlig både i 2019 og 2020. Dersom dette er en vedvarende trend kan det være ønskelig å iverksette gode tiltak for å bedre disse dimensjonene ved motivasjon. Det er også interessant at studentene som fikk digital undervisning med Peer Instruction i 2020 ikke scorer signifikant

forskjellig fra studentene som fikk fysisk undervisning i 2019 på noen motivasjonskonstrukter.

Digital undervisning med Peer Instruction virker å kunne gjennomføres på en tilfredsstillende måte, da studentene som ble undervist digitalt med Peer Instruction oppnådde et konseptuelt læringsutbytte som er typisk for denne typen undervisning i en fysisk klasseromssituasjon.

At flere studenter rapporterer høyt tidsbruk i selvstudie, men oppnår lavt konseptuelt læringsutbytte tyder på at aktivitetene i selvstudie ikke er fruktbare. Det er også flere studenter som rapporterer nokså lavt tidsbruk i selvstudie, men oppnår stort konseptuelt læringsutbytte. Underviser kan ønske å kartlegge hvilke studieteknikker som er utbredt blant studenter som oppnår høyt læringsutbytte uten å bruke veldig mye tid i selvstudie, og iverksette tiltak for å få flere til å prøve disse.

## 6.3 Videre forskning

Det som har blitt diskutert av videre arbeid i denne oppgaven har i stor grad dreid seg om utvidelse eller utvidet bruk av instrumentene som har blitt anvendt. For å kunne bedre kunne stadfeste hvordan motivasjon påvirker læringsutbytte og vice versa mener jeg at motivasjon burde kartlegges både før og etter undervisningen. På denne måten vil en kunne forbedre analysene i denne studien, eller utvikle nye. I tillegg vil en pre-test i IMPEL være verdifull for å si noe om mestringsforventningen til det gitte emnet, heller enn til de kommende emnene. For å kunne kartlegge hva studentene bruker tid på i selvstudie slik at en kan forsøke å finne studieteknikker som fungerer og ikke gjøre det for Peer Instruction, mener jeg at IMPEL-undersøkelsen kunne ha blitt utvidet slik at den tar med spørsmål om hva studentene bruker tid på i selvstudie, i tillegg til hvor mye tid de bruker. I tillegg mener jeg at det er rom for kartlegging av motivasjon knyttet til naturfag og realfag i grunnskole og videregående skole med utgangspunkt i Eccles-modellen for motivasjon. For at dette skal kunne gjennomføres må det utvikles en lignende test som IMPEL for disse sammenhengene og fagområdene.

For å kartlegge om det er en forskjell på selvtillit i svaret hos de som velger en hverdagsforestilling og de som velger korrekt svar på et spørsmål på FCI mener jeg det vil være relevant å implementere en FCI-test med selvtillit der studentene etter hvert svar må vurdere hvor sikker de er på svaret sitt. Jeg mener også FCI har potensiale til å brukes som diagnostisk test på en litt annen bakgrunn enn kun å teste hvilke hverdagsforestillinger som er til stede. Den kan også teste i hvilke kontekster en hverdagsforestilling aktiveres, og i hvilke

kontekster vitenskapelig forståelse blir aktivert.

Jeg mener at deltagelse i undervisning i Peer Instruction kunne vært forsøkt implementert som mere enn tilstedeværelse, for eksempel med et krav om at studentene skal være muntlig aktive før det blir registrert som deltagelse i undervisningen.



# Bibliografi

- Abeysekera, L. & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development*, 34(1), 1–14. doi:10.1080/07294360.2014.934336
- Allison, T. M. H. (2012). *The Impact of Classroom Performance System-Based Instruction with Peer Instruction upon Student Achievement and Motivation in Eighth Grade Math Students* (Doktoravhandling, Liberty University). Hentet fra <https://digitalcommons.liberty.edu/doctoral/490>
- Angell, C., Bungum, B., Henriksen, E. K., Kolstø, S. D., Persson, J. & Renstrøm, R. (2019). *Fysikkdidaktikk* (2. utg.). Oslo: Cappelen Damm Akademisk.
- Anwer, M., Iqbal, H. M. & Harrison, C. (2012). Students' attitude towards science: A case of Pakistan. *Pakistan Journal of Social and Clinical Psychology*, 9(2), 3–9.
- 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>
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: Freeman.
- Barke, H.-D., Hazari, A. & Yitbarek, S. (2009). Students' misconceptions and how to overcome them. I *Misconceptions in Chemistry*. Berlin: Springer-Verlag.
- 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.
- Brekke, G. (2002). Introduksjon til diagnostisk undervisning i matematikk. Trondheim: Matematikksenteret. Hentet fra [http://realfagsloyper.no/sites/default/files/2018-04/Brekke%2C%20G\\_Introduksjon%20til%20diagnostisk%20undervisning%20i%20matematikk%20%E2%80%93%20Utdrag\\_o.pdf](http://realfagsloyper.no/sites/default/files/2018-04/Brekke%2C%20G_Introduksjon%20til%20diagnostisk%20undervisning%20i%20matematikk%20%E2%80%93%20Utdrag_o.pdf)
- Brown, T. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Cashman, E. & Eschenbach, E. (2003). Active learning with Web technology - just in time. doi:10.1109/FIE.2003.1263352

- Claxton, G. (1989). Cognition doesn't matter if you are scared, depressed and bored. I I. Adey, J. Bliss, J. Head & M. Shayer (Red.), *Adolescent Development and School Science*.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. utg.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Cross, K. P. (1998). Why Learning Communities? Why Now? *About Campus*, 3(3), 4–11. doi:10.1177/108648229800300303
- Crouch, C. & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. doi:10.1119/1.1374249
- Csikszentmihalyi, M. & Csikszentmihalyi, I. S. (1992). *Optimal Experience: Psychological Studies of Flow in Consciousness*. Cambridge University Press.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.
- diSessa, A. A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2/3), 105–225. Hentet 19. mai 2021, fra <https://www.jstor.org/stable/3233725>
- 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
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. & Kaczala, C. (1983). Expectancies, values and academic behaviours. I J. Spence (Red.), *Achievement and Achievement Motivation* (s. 75–146). San Francisco: Freeman.
- Eccles, J. S. & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, 53(1), 109–132. doi:10.1146/annurev.psych.53.100901.135153
- Fagen, A., Crouch, C. & Mazur, E. (2002). Peer Instruction: Results from a Range of Classrooms. *The Physics Teacher*, 40, 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
- Ford, M. E. & C.V., N. (1987). A Taxonomy of Human Goals and Some Possible Applications. I M. E. Ford & D. H. Ford (Red.), *Humans as Self-Constructing Living Systems* (s. 289–312). Routledge. Hentet 28. mars 2021, fra <https://www.taylorfrancis.com/https://www.taylorfrancis.com/chapters/edit/10.4324/9780429025297-10/taxonomy-human-goals-possible-applications-martin-ford-nichols>
- Fritz, C. O. (2011). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2. doi:10.1037/a0024338

- Garcia, T. & Pintrich, P. R. (1996). Assessing Students' Motivation and Learning Strategies in the Classroom Context: The Motivated Strategies for Learning Questionnaire. I M. Birenbaum & F. J. R. C. Dochy (Red.), *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge* (s. 319–339). Dordrecht: Springer Netherlands. Hentet fra [https://doi.org/10.1007/978-94-011-0657-3\\_12](https://doi.org/10.1007/978-94-011-0657-3_12)
- Gok, T. (2012). The Effects of Peer Instruction on Students' Conceptual Learning and Motivation. *Asia-Pacific Forum on Science Learning and Teaching*, 13(1), 1–17.
- Gok, T. & Gok, O. (2017). Peer instruction: An evaluation of its theory, application, and contribution. *Asia-Pacific Forum on Science Learning and Teaching*, 18(2). Hentet 21. mai 2021, fra [https://www.eduhk.hk/apfslt/v18\\_issue2/gok/page2.htm](https://www.eduhk.hk/apfslt/v18_issue2/gok/page2.htm)
- Hake, 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). doi:10.1119/1.18809
- 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. & Ringnes, V. (2000). *Kjemi i skolen: undervisning og læring*. Kristiansand: Høyskoleforlaget.
- 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](https://doi.org/10.1007/978-3-030-43620-9_26)
- 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), 502–502. doi:10.1119/1.2344278
- Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30(3), 141–158. doi:10.1119/1.2343497
- Hilsdorf, M. (2020). Cronbach's Alpha: Theory and Application in Python. Hentet 8. mars 2021, fra <https://towardsdatascience.com/cronbachs-alpha-theory-and-application-in-python-d2915dd63586>
- Holt, N., Bremner, A., Sutherland, E., Vlieg, M., Passer, M. W. & Smith, R. E. (2015). *Psychology: The Science of Mind and Behaviour*. McGraw-Hill Education.
- 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
- Johnson, D. W., Johnson, R. T., Haugaløkken, O. K. & Aakervik, A. O. (2006). *Samarbeid i skolen — pedagogisk utviklingsarbeid, samspill mellom mennesker* (4. utg.). Namsos: Pedagogisk Psykologisk Forlag.
- Kelley, K. (2012). On effect size. *Psychological Methods*, 17(2), 137. doi:10.1037/a0028086

- King, A. (2002). Structuring Peer Interaction to Promote High-Level Cognitive Processing. *Theory Into Practice*, 41(1), 33–39. Hentet 21. mai 2021, fra <https://www.jstor.org/stable/1477535>
- Kintz, B. L., Delprato, D. J., Mettee, D. R., Persons, C. E. & Schappe, R. H. (1965). The experimenter effect. *Psychological Bulletin*, 63(4), 223–232. doi:10.1037/h0021718
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*, 3rd ed. New York, NY, US: Guilford Press.
- Knight, R. D. (2004). *Five Easy Lessons: Strategies for Successful Physics Teaching*. San Fransisco, CA: Addison Wesley.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4), 2025–2047. doi:10.1007/s11135-011-9640-9
- Kunnskapsdepartementet. (2021). Slik har koronapandemien påvirket studentene. Regjeringen. Hentet 24. mai 2021, fra <https://www.regjeringen.no/no/aktuelt/studiebarometeret/id2834124/>
- 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
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 55–55.
- Lillemyr, O. F. (2007). *Motivasjon og selvforståelse*. Oslo: Universitetsforlaget.
- Lydersen, S. (2020). Er effekten liten eller stor? *Tidsskrift for Den norske legeforening*, 140(3). doi:10.4045/tidsskr.19.0665
- Mansukhani, V. P. (2010). *Identifying with Mathematics : the effects of conceptual understanding, motivation, and communication on the creation of a strong mathematical identity* (Doktoravhandling, University of California). Hentet 12. november 2020, fra [https://escholarship.org/uc/item/9rk3s88k#article\\_abstract](https://escholarship.org/uc/item/9rk3s88k#article_abstract)
- Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice Hall.
- Merrett, F. (2006). Reflections on the Hawthorne Effect. *Educational Psychology*, 26(1), 143–146. doi:10.1080/01443410500341080
- Michinov, N., Morice, J. & Ferrières, V. (2015). A step further in Peer Instruction: Using the Stepladder technique to improve learning. *Computers & Education*, 91, 1–13. doi:10.1016/j.compedu.2015.09.007
- 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
- Nicholls, J., Cobb, P., Yackel, E., Wood, T. & Wheatley, G. (1990). Students' theories of mathematics and their mathematical knowledge: Multiple dimensions of assessment. *Journal for Research in Mathematics Education*, 21(2), 109–122.

- NOKUT. (2021). Studiebarometeret 2020 – Hovedtendenser. Hentet 24. mai 2021, fra [https://www.nokut.no/globalassets/studiebarometeret/2021/hoyere-utdanning/studiebarometeret-2020\\_hovedtendenser\\_1-2021.pdf](https://www.nokut.no/globalassets/studiebarometeret/2021/hoyere-utdanning/studiebarometeret-2020_hovedtendenser_1-2021.pdf)
- Northrup, D. A. (1997). *The Problem of the Self-report in Survey Research: Working Paper*. Institute for Social Research, York University.
- Novak, G. M., Patterson, E., Gavrin, A. & Christian, W. (1999). *Just-in-time Teaching: Blending Active Learning with Web Technology*. Prentice Hall.
- 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
- Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond Cold Conceptual Change: The Role of Motivational Beliefs and Classroom Contextual Factors in the Process of Conceptual Change. *Review of Educational Research*, 63(2), 167–199. doi:10.3102/00346543063002167
- 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
- Queloz, A. C., Klymkowsky, M. W., Stern, E., Hafen, E. & Köhler, K. (2017). Diagnostic of students' misconceptions using the Biological Concepts Instrument (BCI): A method for conducting an educational needs assessment. *PloS one*, 10(5).
- Ramberg, I. (2006). *Realfag eller ikke? Elevers motivasjon for valg og bortvalg av realfag i videregående opplæring*. Hentet 9. juni 2021, fra <https://nifu.brage.unit.no/nifu-xmlui/handle/11250/2357331>
- Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion*, 13(2), 83–103. doi:10.1007/BF00992956
- 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
- Ringdal, K. (2018). *Enhet og mangfold: Samfunnsvitenskapelig forskning og kvantitativ metode* (4. utg.). Fagbokforlaget.
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:10.1006/ceps.1999.1020
- Sawilowsky, S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, 8(2). doi:10.22237/jmasm/1257035100
- Simkins, S. & Maier, M. (2010). *Just-in-time Teaching: Across the Disciplines, Across the Academy*. Stylus Publishing, LLC.
- Simon, B. & Cutts, Q. (2012). Peer instruction: a teaching method to foster deep understanding. *Communications of the ACM*, 55(2), 27–29. doi:10.1145/2076450.2076459
- Sjøberg, S. (2014). *Naturfag som allmenndannelse: en kritisk fagdidaktikk* (3. utg.). Oslo: Gyldendal Akademisk.

- Stojanovska, M., Petrusovski, V., Köller, H.-G. & Karlsen, S. (2015). Students' alternative conceptions and ways to overcome them. I I. Maciejowska & B. Byers (Red.), *A guidebook of good practice for the pre-service training of chemistry*. Krakow: Jagiellonian University.
- Svartdal, F. (2015). *Psykologiens forskningsmetoder: En introduksjon* (4. utg.). Fagbokforlaget.
- Taber, K. (2002). *Chemical misconceptions: prevention, diagnosis and cure*. Royal Society of Chemistry.
- Tobias, S. (1992). *Revitalizing Undergraduate Science: Why Some Things Work and Most Don't. An Occasional Paper on Neglected Problems in Science Education*. Research Corporation, Book Dept. Hentet 12. november 2020, fra <https://eric.ed.gov/?id=ED357975>
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10, 159–169.
- UiO. (udatert). IMPEL - Interactive engagement and motivation in physics learning - Fysisk institutt. Hentet 23. mai 2021, fra <https://www.mn.uio.no/fysikk/forskning/prosjekter/impel/index.html>
- Utdanningsdirektoratet. (2017). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Hentet fra <https://www.udir.no/lk20/overordnet-del/>
- Wallace, S. (2015). *A dictionary of education*. Oxford, UK: Oxford University Press.
- Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K. (2007). *Probability & statistics for engineers & scientists* (8. utg.). Pearson Education, Inc.
- Wentzel, K. R. (1991). Relations between Social Competence and Academic Achievement in Early Adolescence. *Child Development*, 62(5), 1066–1078. doi:10.2307/1131152
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6(1), 49–78. doi:10.1007/BF02209024
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68–81. doi:10.1006/ceps.1999.1015
- Wæge, K. (2007). *Elevenes motivasjon for å lære matematikk og undersøkende matematikkundervisning* (Doktoravhandling, Norges teknisk-naturvitenskapelige universitet). Hentet 11. november 2020, fra <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/258129>
- Yaoyuneyong, G. & Thornton, A. (2011). Combining peer instruction and audience response systems to enhance academic performance, facilitate active learning and promote peer-assisted learning communities. *International Journal of Fashion Design, Technology and Education*, 4(2), 127–139. doi:10.1080/17543266.2011.564210
- Yen, H.-C., Tuan, H.-L. & Liao, C.-H. (2010). Investigating the Influence of Motivation on Students' Conceptual Learning Outcomes in Web-based

vs. Classroom-based Science Teaching Contexts. *Research in Science Education*, 41, 211–224. doi:10.1007/s11165-009-9161-x





# Vedlegg

## A.1 Statistiske verdier fra IMPEL - utvalgsgruppen er alle som tok IMPEL

**Tabell A.1:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Interesse og glede», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Interesse og glede,  
Cronbach's  $\alpha = 0.901$**

Spørsmål:	Gjennomsnitt:	Standardavvik:	Cronbach's $\alpha$ u/ spørsmål:	Korrelasjon mot resten av skala:	Skewness:	Kurtosis:
Jeg var veldig interessert i å lære nye ting i fysikk.	0.7823	0.2868	0.882	0.733	-1.259	0.709
Jeg synes fysikk var veldig gøy.	0.7903	0.2669	0.888	0.694	-1.045	-0.170
Jeg synes det er veldig gøy å lære fysikk.	0.7742	0.2912	0.877	0.766	-1.146	0.355
Jeg kan bli så engasjert at jeg glemmer tiden når jeg holder på med fysikk.	0.5323	0.3275	0.895	0.667	-0.254	-0.904
Fysikk handler om temaer jeg synes er veldig spennende.	0.7097	0.2508	0.890	0.684	-0.927	0.909
Jeg er veldig interessert i å lære nye ting i fysikk	0.7661	0.3091	0.863	0.853	-1.261	0.683

**Tabell A.2:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Nytteverdi», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Nytteverdi,  
Cronbach's  $\alpha = 0.850$**

Spørsmål:	Gjennomsnitt:	Standardavvik:	Cronbach's $\alpha$ u/ spørsmål:	Korrelasjon mot resten av skala:	Skewness:	Kurtosis:
Jeg følte at det jeg lærte i fysikk ville bli nyttig for meg.	0.8226	0.2753	0.860	0.477	-1.586	1.807
Det jeg lærer i fysikk vil gi meg en fordel på arbeidsmarkedet.	0.7419	0.2188	0.806	0.748	-0.254	-0.946
Det jeg lærer i fysikk er veldig nyttig for meg i fremtiden.	0.7742	0.2611	0.807	0.726	-0.768	0.699
Å studere fysikk vil gi meg akkurat det yrket jeg ønsker meg.	0.6694	0.2692	0.832	0.610	-0.494	-0.245
Å studere fysikk vel gi jobbmuligheter som er attraktive for meg.	0.7742	0.2359	0.808	0.726	-1.221	2.256
Det jeg lærer i fysikk er veldig nyttig for meg videre i utdanningen.	0.8871	0.1560	0.836	0.623	-1.075	0.220

**Tabell A.3:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Mestringsforventning», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Mestringsforventning,  
Cronbach's  $\alpha = 0.876$**

Spørsmål:	Gjennomsnitt:	Standardavvik:	Cronbach's $\alpha$ u/ spørsmål:	Korrelasjon mot resten av skala:	Skewness:	Kurtosis:
Jeg synes det var lett å lære meg fysikk.	0.5161	0.3023	0.880	0.528	-0.252	-0.847
Jeg forventet at jeg kom til å gjøre det bedre enn gjennomsnittet i fysikkemner	0.4758	0.2766	0.846	0.749	0.202	-0.336
Jeg forventet at jeg kom til å lære stoffet lett i fysikkemner.	0.3871	0.2726	0.853	0.699	0.692	-0.027
Jeg kommer til å gjøre det bedre enn de fleste andre studentene i fysikkemnene jeg tar dette semesteret.	0.3790	0.2404	0.857	0.670	-0.289	-0.832
Jeg lærer fagstoffet lett i fysikkemner jeg tar dette semesteret.	0.3790	0.2128	0.852	0.736	0.120	-0.474
Jeg kommer til å gjøre det bedre enn gjennomsnittet i fremtidige fysikkemner.	0.4435	0.2559	0.865	0.605	0.088	0.302
Jeg kommer til å lære fagstoffet lett i fremtidige fysikkemner.	0.4435	0.2390	0.857	0.679	-0.003	0.013

**Tabell A.4:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Selvrealisering og identitet», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Selvrealisering og identitet,  
Cronbach's  $\alpha = 0.874$**

Spørsmål:	Gjennomsnitt:	Standardavvik:	Cronbach's $\alpha$ u/ spørsmål:	Korrelasjon mot resten av skala:	Skewness:	Kurtosis:
Jeg følte at fysikk passet for meg.	0.7984	0.2616	0.848	0.737	-1.160	0.191
Det var viktig for meg å bli flink i fysikk.	0.7661	0.2495	0.870	0.514	-1.212	1.685
Jeg synes at fysikkstudier passer for meg.	0.7500	0.2236	0.853	0.717	-1.195	2.783
Jeg liker å fortelle folk at jeg tar emner i fysikk.	0.6774	0.3240	0.855	0.676	-0.792	-0.277
Folk som kjenner med mener at jeg er en fysikkperson.	0.6210	0.2949	0.855	0.663	-0.675	-0.026
Jeg mener at jeg er en fysikkperson.	0.6371	0.2875	0.842	0.775	-0.970	-0.432
Det er viktig for meg å lære fysikk.	0.7419	0.2188	0.841	0.848	-0.892	0.614
Det er viktigere for meg å få gode karakterer i fysikkemner enn i andre emner.	0.5645	0.2956	0.898	0.273	-0.022	-0.738

**Tabell A.5:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Kostnad», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

<b>Konstrukt: Kostnad, Cronbach's <math>\alpha = 0.821</math></b>								
<i>Spørsmål:</i>	<i>Gjennomsnitt:</i>	<i>Standardavvik:</i>	<i>Cronbach's <math>\alpha</math> u/ spørsmål:</i>	<i>Korrelasjon mot resten av skala:</i>	<i>Skewness:</i>	<i>Kurtosis:</i>		
Jeg følte at fysikk krevde veldig mye tid og arbeid.	0.7984	0.2535	0.845	0.165	-1.4440	2.133		
Å ta emner i fysikk koster meg mer tid og krefter enn om jeg hadde valgt å studere noe annet.	0.7339	0.2135	0.803	0.533	-0.559	-0.035		
Jeg synes fysikkemner er vanskeligere enn andre emner.	0.5403	0.2746	0.797	0.565	-0.340	-0.877		
Jeg ville hatt mer fritid dersom jeg studerte noe annet.	0.7258	0.2446	0.786	0.663	-0.481	-0.724		
Jeg synes det er slitsomt å studere fysikk.	0.5000	0.2739	0.778	0.698	-0.325	-0.950		
Å studere fysikk krever for mye arbeid.	0.5081	0.3194	0.818	0.430	0.141	-1.130		
Jeg får gjort mindre av andre ting jeg liker, fordi jeg tar emne(t) i fysikk.	0.5645	0.3225	0.772	0.719	-0.219	-1.069		
Jeg synes det er stressende å skulle prestere godt i fysikk.	0.5645	0.3160	0.790	0.610	-0.735	-0.618		

## **A.2 Statistiske verdier fra IMPEL – utvalgsgruppen er de som både tok IMPEL og FCI**

**Tabell A.6:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Interesse og glede», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Interesse og glede,  
Cronbach's  $\alpha = 0.911$**

Spørsmål:	Gjennomsnitt:	Standardavvik:	Cronbach's $\alpha$ u/ spørsmål:	Korrelasjon mot resten av skala:	Skewness:	Kurtosis:
Jeg var veldig interessert i å lære nye ting i fysikk.	0.8125	0.3096	0.883	0.838	-1.746	2.343
Jeg synes fysikk var veldig gøy.	0.7969	0.2772	0.886	0.812	-1.089	-0.168
Jeg synes det er veldig gøy å lære fysikk.	0.8750	0.1826	0.918	0.574	-1.174	0.144
Jeg kan bli så engasjert at jeg glemmer tiden når jeg holder på med fysikk.	0.5938	0.3146	0.899	0.752	-0.136	-0.887
Fysikk handler om temaer jeg synes er veldig spennende.	0.7969	0.2086	0.892	0.804	-1.178	1.952
Jeg er veldig interessert i å lære nye ting i fysikk	0.8594	0.2230	0.887	0.827	-1.718	2.697



**Tabell A.7:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Nytteverdi», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Nytteverdi,  
Cronbach's  $\alpha = 0.747$**

<i>Spørsmål:</i>	<i>Gjennomsnitt:</i>	<i>Standardavvik:</i>	<i>Cronbach's <math>\alpha</math> u/ spørsmål:</i>	<i>Korrelasjon mot resten av skala:</i>	<i>Skewness:</i>	<i>Kurtosis:</i>
Jeg følte at det jeg lærte i fysikk ville bli nyttig for meg.	0.8438	0.3010	0.773	0.356	-2.152	4.030
Det jeg lærer i fysikk vil gi meg en fordel på arbeidmarkedet.	0.7656	0.2135	0.646	0.716	-0.129	-1.646
Det jeg lærer i fysikk er veldig nyttig for meg i fremtiden.	0.8281	0.2366	0.671	0.615	-1.266	0.833
Å studere fysikk vil gi meg akkurat det yrket jeg ønsker meg.	0.7656	0.1929	0.749	0.329	-0.113	-1.194
Å studere fysikk vel gi jobbmuligheter som er attraktive for meg.	0.8438	0.2016	0.681	0.603	-0.845	-0.838
Det jeg lærer i fysikk er veldig nyttig for meg videre i utdanningen.	0.9375	0.1443	0.732	0.420	-2.375	5.314

**Tabell A.8:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Mestringsforventning», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Mestringsforventning,**

**Cronbach's  $\alpha = 0.858$**

<i>Spørsmål:</i>	<i>Gjennomsnitt:</i>	<i>Standardavvik:</i>	<i>Cronbach's <math>\alpha</math> u/ spørsmål:</i>	<i>Korrelasjon mot resten av skala:</i>	<i>Skewness:</i>	<i>Kurtosis:</i>
Jeg synes det var lett å lære meg fysikk.	0.5313	0.28687	0.834	0.692	-0.531	-0.855
Jeg forventet at jeg kom til å gjøre det bedre enn gjennomsnittet i fysikkemner	0.5625	0.3228	0.821	0.736	-0.083	-0.267
Jeg forventet at jeg kom til å lære stoffet lett i fysikkemner.	0.5469	0.2772	0.829	0.685	0.522	0.128
Jeg kommer til å gjøre det bedre enn de fleste andre studentene i fysikkemnene jeg tar dette semesteret.	0.4036	0.2720	0.840	0.617	-0.910	0.796
Jeg lærer fagstoffet lett i fysikkemner jeg tar dette semesteret.	0.4375	0.2141	0.838	0.630	0.227	-0.646
Jeg kommer til å gjøre det bedre enn gjennomsnittet i fremtidige fysikkemner.	0.4688	0.1548	0.853	0.530	0.060	0.055
Jeg kommer til å lære fagstoffet lett i fremtidige fysikkemner.	0.4688	0.1797	0.848	0.568	0.192	-0.821

**Tabell A.9:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Selvrealisering og identitet», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

**Konstrukt: Selvrealisering og identitet,  
Cronbach's  $\alpha = 0.806$**

<i>Spørsmål:</i>	<i>Gjennomsnitt:</i>	<i>Standardavvik:</i>	<i>Cronbach's <math>\alpha</math> u/ spørsmål:</i>	<i>Korrelasjon mot resten av skala:</i>	<i>Skewness:</i>	<i>Kurtosis:</i>
Jeg følte at fysikk passet for meg.	0.8594	0.2230	0.749	0.742	-1.718	2.697
Det var viktig for meg å bli flink i fysikk.	0.7969	0.2085	0.777	0.566	-1.178	1.952
Jeg synes at fysikkstudier passer for meg.	0.8594	0.1573	0.803	0.368	-0.653	-0.321
Jeg liker å fortelle folk at jeg tar emner i fysikk.	0.6875	0.2814	0.797	0.465	-0.080	-1.516
Folk som kjenner med mener at jeg er en fysikkperson.	0.7500	0.2415	0.753	0.705	-0.507	-0.735
Jeg mener at jeg er en fysikkperson.	0.7500	0.2236	0.727	0.872	-0.639	0.027
Det er viktig for meg å lære fysikk.	0.7969	0.1875	0.749	0.799	-1.418	4.192
Det er viktigere for meg å få gode karakterer i fysikkemner enn i andre emner.	0.5313	0.2394	0.874	-0.119	0.765	0.121

**Tabell A.10:** Tabellen viser statistiske verdier for spørsmålene som gjør opp konstruktet «Kostnad», samt Cronbach's  $\alpha$  for dette konstruktet. Det har blitt funnet gjennomsnitt, standardavvik, skewness, kurtosis, ethvert spørsmåls korrelasjon med resten av skala, samt hva Cronbach's  $\alpha$  ville vært for konstruktet uten hvert enkelt spørsmål.

<b>Konstrukt: Kostnad, Cronbach's <math>\alpha = 0.739</math></b>								
<i>Spørsmål:</i>	<i>Gjennomsnitt:</i>	<i>Standardavvik:</i>	<i>Cronbach's <math>\alpha</math> u/ spørsmål:</i>	<i>Korrelasjon mot resten av skala:</i>	<i>Skewness:</i>	<i>Kurtosis:</i>		
Jeg følte at fysikk krevde veldig mye tid og arbeid.	0.7656	0.2809	0.766	0.161	-1.426	2.367		
Å ta emner i fysikk koster meg mer tid og krefter enn om jeg hadde valgt å studere noe annet.	0.7344	0.1700	0.730	0.339	0.074	-0.489		
Jeg synes fysikkemner er vanskeligere enn andre emner.	0.4844	0.2809	0.735	0.319	0.137	-1.009		
Jeg ville hatt mer fritid dersom jeg studerte noe annet.	0.6875	0.2500	0.669	0.657	-0.343	-0.738		
Jeg synes det er slitsomt å studere fysikk.	0.5156	0.2494	0.644	0.778	-0.598	-0.835		
Å studere fysikk krever for mye arbeid.	0.4688	0.2720	0.748	0.250	0.277	-0.672		
Jeg får gjort mindre av andre ting jeg liker, fordi jeg tar emne(t) i fysikk.	0.4531	0.2918	0.671	0.619	-0.125	-0.777		
Jeg synes det er stressende å skulle prestere godt i fysikk.	0.4688	0.2869	0.709	0.449	-0.634	-0.958		

### A.3 Verdier for skjevhet og kurtose for fordelingen av studentnes gjennomsnitt på hvert konstrukt

I dette appendikset presenteres tabeller som viser verdier for skjevhet og kurtose i 2019 og 2020, i henholdsvis tabell A.12 og A.11.

**Tabell A.11:** Tabellen viser verdiene for skjevhet og kurtose for fordelingen av studentenes gjennomsnittsvar på hvert konstrukt i 2020. Fordelingene finnes i figur 4.1.

Konstrukt	År	Kurtose	Skjevhet
Nytteverdi	2020	-0.71	-0.58
Mestringsforventning	2020	-0.57	0.30
Kostnad	2020	-0.61	-0.53
Selvrealisering og identitet	2020	0.27	-0.92
Interesse og glede	2020	-0.81	-0.57

**Tabell A.12:** Tabellen viser verdiene for skjevhet og kurtose for fordelingen av studentenes gjennomsnittsvar på hvert konstrukt i 2019. Fordelingene finnes i figur 4.2.

Konstrukt	År	Kurtose	Skjevhet
Nytteverdi	2019	0.57	-0.90
Mestringsforventning	2019	-0.21	0.01
Kostnad	2019	-0.22	-0.30
Selvrealisering og identitet	2019	1.01	-1.11
Interesse og glede	2019	2.10	-1.00





