



Uit

NORGES
ARKTISKE
UNIVERSITET

Fakultetet for naturvitenskap og teknologi | Instituttet for kjemi

En helhetlig forskningsprosess

En kvantitativ studie om læring av fagkunnskap i organisk kjemi gjennom praktisk arbeid

—

Janita Fosslund Hansen

Masteroppgave i Lektorutdanningen 8-13.klasse | KJE-3907

Kjemididaktikk, Juni 2019



Sammendrag

Praktisk arbeid har en uvurderlig og uerstattelig plass i naturfagundervisningen. Både læreplanen i kjemi og fagfornyelsen i naturfag som inntreer i 2020 beskriver fagenes praktiske aspekt som svært sentralt. Dessverre rapporter forskning fra flere hold at praktisk arbeid som undervisningsmetode ikke lever opp til sitt fulle potensial slik det blir praktisert i dag. Især viser det seg å være vanskelig å påvise en sammenheng mellom praktisk arbeid og læring av spesifikk fagkunnskap.

Gjennom dette masterprosjektet ble det utviklet et undervisningsopplegg av praktisk karakter med hensikt å fremme elevers læring av fagbegreper i *organisk kjemi 2*. Essensen av opplegget er et labforsøk i form av en *helhetlig forskningsprosess*. For å kartlegge elevenes teoretiske kunnskaper før og etter *forskningsprosessen* ble det henholdsvis benyttet en pretest og en posttest.

Det ble avdekket en gjennomsnittlig signifikant *forbedring* av testskår på posttesten. Basert på elevenes egne evalueringer av arbeidsmetoder, samt anbefalinger fra faglitteratur, tyder det på at undervisningopplegget i sin helhet påvirket testresultatene i større grad enn kun labforsøket alene.

Forord

Denne masteroppgaven er siste ledd i en 5-årig master for *Lektorutdanningen 8-13.trinn* ved Universitet i Tromsø. Studietiden har vært givende og lærerikt på mange måter.

Jeg vil først og fremst rette en takk til mine veiledere Hans-Georg Køller, Institutt for lærerutdanning og pedagogikk, og Jørn Hansen, Institutt for kjemi, for gode tilbakemeldinger og innspill. Jeg vil også takke alle elevene som deltok i prosjektet, samt faglærerne for tillitten til å ta over klassene deres i forskningsperioden.

Takk til gode medstudenter og familie for korrekturlesing, støtte og hverdagsgleder.

Tromsø, juni 2019

Janita Fosslund Hansen

Innholdsfortegnelse

1	INNLEDNING	1
1.1	BAKGRUNN	1
1.2	PROBLEMSTILLING	2
2	TEORI	5
2.1	PRAKTISK ARBEID – EN BEGREPSAVKLARING	5
2.2	ET TEORIFUNDAMENT FOR PRAKTISK ARBEID I NATURFAG	5
2.3	OVERORDNET MÅLSETTING MED NATURFAGSUNDERVISNING I SKOLEN	7
2.4	PRAKTISK ARBEID I NATURFAGSUNDERVISNING	9
2.4.1	<i>Mål med praktisk arbeid i naturfag</i>	9
2.4.2	<i>Praktisk arbeid som undervisnings - og læringsmetode</i>	10
2.4.3	<i>Planlegging og evaluering av praktisk arbeid</i>	11
2.4.4	<i>Ulike typer praktisk arbeid</i>	13
2.5	MISOPPFATNINGER I NATURFAG	14
2.6	SYNTESE OG ANALYSE AV DIBENZALACETON	15
3	METODER OG MATERIALE	19
3.1	BEGRUNNELSE FOR VALG AV METODE	19
3.2	UTVALG	20
3.3	GJENNOMFØRING OG INNSAMLING AV DATA	21
3.3.1	<i>Undervisningsopplegg</i>	21
3.3.2	<i>Pretest</i>	22
3.3.3	<i>Labforsøk</i>	22
3.3.4	<i>Labheftet</i>	23
3.3.5	<i>Observasjon</i>	23
3.3.6	<i>Etterarbeid</i>	24
3.3.7	<i>Posttest</i>	24
3.4	VURDERING AV TESTER OG LABHEFTE	24
3.5	STATISTISKE ANALYSEMETODER	25
3.5.1	<i>Deskriptiv statistikk</i>	25
3.5.2	<i>Statistisk inferens</i>	27
3.6	STUDIENS KVALITET	29
3.6.1	<i>Validitet</i>	29
3.6.2	<i>Reliabilitet</i>	31
3.7	FORSKNINGSETIKK	32
4	RESULTATER OG ANALYSE	33
4.1	DESKRIPTIV STATISTIKK	33
4.1.1	<i>Pretest og posttest</i>	33

4.1.2	<i>Pretest vs. Posttest</i>	37
4.1.3	<i>Korrelasjon</i>	39
4.2	STATISTISK INFERENS	42
4.2.1	<i>Outliers</i>	42
4.2.2	<i>Normalfordelt data</i>	43
4.2.3	<i>Paret t-test og konfidensintervall</i>	45
4.3	TEMABASERT POENGSKÅR.....	46
4.4	ELEVEVALUERING AV UNDERVISNINGSMETODER	47
5	DISKUSJON	49
5.1	FRA UTVALG TIL POPULASJON	49
5.2	STYRKET FAGKUNNSKAP SOM FØLGE AV LABFORSØK.....	50
5.3	STYRKET FAGKUNNSKAP SOM FØLGE AV UNDERVISNINGSSOPPLEGG.....	50
5.4	ELEVENES LÆRING AV FAGBEGREPENE	52
5.4.1	<i>Reaksjonstyper og reaksjonsmekanismer</i>	52
5.4.2	<i>NMR</i>	54
5.4.3	<i>Massespektrometri</i>	55
5.4.4	<i>OmkrySTALLISERING</i>	56
5.5	EVALUERING OG PLANLEGGING AV LABFORSØK: SYNTSE AV DBA.....	57
5.6	FORSØK SOM EN HELHETLIG FORSKNINGSPROSESS	61
5.7	SVAKHETER MED STUDIEN.....	62
6	AVSLUTNING	65
6.1	KONKLUSJON	65
6.2	VEIEN VIDERE	67
	REFERANSELISTE	68
	VEDLEGG 1: SAMTYKKESKJEMA	72
	VEDLEGG 2: GODKJENNING FRA NSD	73
	VEDLEGG 3: VURDERINGSKRITERIER FOR PRETEST OG POSTTEST	74
	VEDLEGG 4: PRETEST	75
	VEDLEGG 5: POSTTEST	77
	VEDLEGG 6: LABHEFTE	80

Tabelliste

TABELL 1. ANTALL FRIHETSGRADER I PRAKTISK ARBEID, SOM ANGITT I HERRON (1971).....	13
TABELL 2. BESKRIVELSE AV SENTRALMÅL OG SPREDNINGSMÅL.....	26
TABELL 3. MULIGE UTFALL AV PARET T-TEST.....	28
TABELL 4. SENTRAL - OG SPREDNINGSMÅL AV PRETEST OG POSTTEST	33
TABELL 5. ANTALL ELEVER MED POSITIV, NEGATIV ELLER INGEN ENDRING I SKÅR.....	37
TABELL 6. SKÅRDIFFERANSEN TIL ELEVER MED $\geq 80\%$ RETT PÅ PRETESTEN	37
TABELL 7. GJENNOMSNIITT, STANDARDAVVIK OG VARIASJONSBREDDE AV UTVALGETS SKÅRDIFFERANSE.....	38
TABELL 8. KORRELASJONSANALYSE MELLOM PRETEST-POSTTEST, OG LABHEFTE-POSTTEST	39
TABELL 9. SKJEVHET, KURTOSE, STANDARDFEIL OG Z-SKÅR FOR δ	44
TABELL 10. RESULTATER FRA PARET T-TEST OG KONFIDENSINTERVALL.....	45
TABELL 11. SKÅRFORDELING PÅ HVER AV OPPGAVENE I PRETESTEN OG POSTTESTEN	46

Figurliste

FIGUR 1: DEN PROKSIMALE UTVIKLINGSSONEN (IMSEN, 2005, S. 259)	7
FIGUR 2: PISAS DEFINISJON AV NATURVITENSKAPLIG ALLMENNDANNELSE (BYBEE MFL., 2009)	7
FIGUR 3: HENSikten MED PRAKTISK ARBEID (MILLAR, LE MARÉCHAL & TIBERGHEN, 1999, S. 40	10
FIGUR 5: PLANLEGGING OG EVALUERINGSPROSESSEN AV EN PRAKTISK AKTIVITET (MILLAR MFL., 1999)	11
FIGUR 6: REAKSJONSLIKNING FOR DANNElse AV DBA	15
FIGUR 7: «KJEMIENS TRE DIMENSJONER» (RINGNES & HANNISDAL, 2014).....	16
FIGUR 8: UNDERVISNINGSPLEGG OG DATAINNSAMLINGSMETODER	21
FIGUR 9: DISTRIBUTUSJON AV ELEVENES SKÅR PÅ PRETESTEN	35
FIGUR 10: DISTRIBUTUSJON AV ELEVENES SKÅR PÅ POSTTESTEN.....	36
FIGUR 11: SPREDNINGSPLOTT AV PRETEST – OG POSTTESTSKÅR.....	40
FIGUR 12: SPREDNINGSPLOTT AV LABHEFTET OG POSTTEST.....	41
FIGUR 13: BOKSPLOTT AV SKÅRDIFFERANSEN (δ) MED OUTLIER	42
FIGUR 14: HISTOGRAM AV SKÅRDIFFERANSEN (δ)	43
FIGUR 15: NORMALFORDELT Q-Q PLOT AV SKÅRDIFFERANSEN (δ).....	44
FIGUR 16: ELEVEVALUERING AV UNDERVISNINGSMETODER/ARBEIDSMETODER.....	48

1 Innledning

Målet med denne masteroppgaven er å undersøke om et undervisningsopplegg som inkluderer praktisk arbeid kan bidra til å øke elevenes læring av spesifikke fagkunnskaper i *kjemi 2*. Faglitteratur skrevet om praktisk arbeid i skolen tar utgangspunkt i flere naturvitenskapelige fag, men denne teorien er i stor grad overførbar mellom de ulike retningene. Ordet *naturfag* brukes derfor i denne oppgaven som en felles betegnelse for alle de naturvitenskapelige fagdisiplinene. Praktisk arbeid er, og har i lang tid vært, en verdifull og ubestridelig del av naturfagundervisningen (P. M. Kind, 1999). Likevel har det opp gjennom årene blitt stilt kritiske spørsmål fra flere hold rundt elevenes faktiske læringsutbytte fra en slik undervisningsmetode (Abrahams & Millar, 2008; Berry, Mulhall, Gunstone & Loughran, 1999; Hofstein & Lunetta, 2004). Gjennom denne masteroppgaven belyses dilemmaet rundt praktisk arbeid som læringsstrategi både alene og som en del av et større undervisningsopplegg.

1.1 Bakgrunn

Alle læreplaner i Kunnskapsløftet skal i løpet av de neste årene revideres gjennom *fagfornyelsen*. Som følge av dette implementeres *kjerneelementer* – det viktigste og mest sentrale elevene skal lære i hvert fag (Kunnskapsdepartementet, 2018b). Arbeidet med kjerneelementer i kjemifaget starter først i april 2019 og vil ikke tre i kraft før om et par år. Det er derimot allerede fastsatt kjerneelementer i naturfag hvor ett av de, «*Naturvitenskaplige praksiser og tenkemåter*», tilskriver praktisk arbeid en sentral plass i faget. Det er rimelig å anta at kjemifaget også vil ha et tilsvarende fokusområde. Av den grunn er det hensiktsmessig at praktisk arbeid utøves på en slik måte at det tjerner elevenes læring i faget. Spørsmålet om hvordan elever kan lære fra praktisk arbeid er grunnsteinen i problemstillingen til denne masteroppgaven.

Egne erfaringer med praktisk arbeid i naturfag tilsier at det klassiske slagordet «I do, and I understand», ikke alltid stemmer. Driver i P. M. Kind (1999, s. 90) skrev en omformulering av slagordet som jeg tror mange kan relatere til: «I do and I am even more confused». En slik situasjon er uønsket, og i den forbindelse utviklet jeg et undervisningsopplegg av praktisk karakter med et fokus på refleksjon over egne handlinger. Utformingen av en praktisk aktivitet burde ta utgangspunkt i aktivitetens læringsmål, og ut fra samsvaret mellom disse to faktorene kan aktivitetens vellykkethet bedømmes (Sjøberg, 2011). I denne studien tar den praktiske aktiviteten form som en *helhetlig forskningsprosess*, noe som vil forklares

ytterligere i neste kapittel. For å få et innblikk i hvordan et undervisningsopplegg med en *helhetlig forskningsprosess* påvirker elevenes faglige prestasjoner, ble det gjennomført en pretest og en posttest, henholdsvis før og etter et praktisk forsøk.

1.2 Problemstilling

Det foreligger lite bevis for at praktisk arbeid fremmer læring av teoretisk fagkunnskap (Abrahams & Millar, 2008; Abrahams & Reiss, 2012; Hofstein & Lunetta, 2004; Séré, 2002). I denne masteroppgaven er jeg derfor ute etter å undersøke hvordan praktisk arbeid kan være med på å fremme læring av bestemte fagbegreper. Dette gjøres ved å gjennomføre et undervisningsopplegg som inkluderer en spesifikk type praktisk forsøk, og uttrykkes gjennom følgende forskningshypotese:

Et undervisningsopplegg som innebærer et praktisk forsøk i form av en helhetlig forskningsprosess, vil styrke elevenes fagkunnskaper om sentrale begrep i organisk kjemi

Med utgangspunkt i det praktiske forsøket *syntese av dibenzalaceton (DBA)*, valgte jeg fem fagbegreper som skulle undersøkes nærmere: *addisjonsreaksjon, eliminasjonsreaksjon, kjernemagnetisk resonansspektroskopi (NMR), massespektrometri (MS) og omkrystallisering*. Fagbegrepene gjenspeiles i kompetansemålene i læreplanen for kjemi (KJE1-01; Utdanningsdirektoratet, 2006, s. 6) og er uthevet med fet skrift:

- **gjøre rede for reaksjonstypene oksidasjon, substitusjon, addisjon, eliminasjon, kondensasjon og hydrolyse, og gjøre forsøk med minst to av dem**
- **forklare reaksjonsmekanismen ved addisjon og eliminasjon**
- utføre analyser med kolorimetri **og tolke enkle massespektre og ¹H-NMR-spektre**
- **gjøre rede for og utføre kromatografi, destillasjon og omkrystallisering**

Hva en *helhetlig forskningsprosess* innebærer er imidlertid ikke selvforklarende. I denne masteroppgaven brukes begrepet for å beskrive at elevene er involvert i et forløp hvor de skaper, bearbeider og analyserer et produkt med tilhørende etterarbeid. Det ble vektlagt at elevene skulle lage et eget produkt, da eierskapsfølelse til oppgaven kan føre til mer motiverte og engasjerte elever (Berry mfl., 1999).

Den helhetlige forskningsprosessen beskrives gjennom fem steg som innebærer at elevene skal:

- 1) skape sitt eget produkt gjennom syntese
- 2) gjennomføre isolering og opprensingsarbeid
- 3) beregne prosentutbytte og drøfte feilkilder
- 4) analysere produktet
- 5) tolke analysespekter

En myte i naturfagsundervisning er at det finnes én måte å utøve naturvitenskap på (McComas, 1998). Denne metoden navngis ofte «*the scientific method*», som i hovedtrekk går ut på å formulere en hypotese, samle inn data for å bekrefte/avkrefte hypotesen og til slutt trekke en konklusjon. Dette representerer ikke alle aspektene ved naturvitenskapen og blir feilaktig fremstilt i lærebøker som en «fasit» på hvordan forsøk og eksperimenter skal utføres (Lederman, Antink & Bartos, 2014). Den *helhetlige forskningsprosessen* bygger derfor ikke på denne «modellen». Det er viktig å understreke at en forskningsprosess i skolesammenheng er noe ganske annet enn i forsker-verden, og jeg ønsker ikke å trekke noen paralleller.

Jeg tok et bevisst valg om å bruke begrepet *fagkunnskap* fremfor *kompetanse* i forskningshypotesen. Grunnen til dette er at kompetansebegrepet, ifølge *overordnet del* av læreplanen, omfatter flere komponenter som blant annet evne til kreativitet og kritisk tenkning. Disse elementene kan ikke bedømmes gjennom studiens datamateriale.

Fagkunnskap i denne masteroppgaven defineres analogt med *overordnet del* av læreplanens kunnskapsdefinisjon: «*Kunnskap innebærer å kjenne til og forstå fakta, begreper, teorier, ideer og sammenhenger innenfor ulike fagområder og temaer*» (Kunnskapsdepartementet, 2018a, s. 11).

2 Teori

I dette kapitlet defineres begrepet *praktisk arbeid*, og dets sentrale plass i naturfagundervisningen drøftes i lys av konstruktivistiske læringsteorier, overordnede mål med naturfagsundervisning i skolen, samt tidligere forskning innenfor naturfagsutdanning. Videre vil kapitlet ta for seg utfordringer knyttet til bruk av praktisk arbeid som undervisnings – og læringsmetode, og se på hvordan ulike typer praktisk arbeid kan klassifiseres. Elevene har ofte ikke-vitenskapelige forestillinger som de tar med seg i klasserommet, og da dette kan påvirke deres læring, vies et delkapittel til faglitteratur om misoppfatninger. Til slutt gis det en enkel, teoretisk forklaring av studiens fem fagbegreper og hvordan de henger sammen med labforsøket.

2.1 Praktisk arbeid – en begrepsavklaring

I denne masteroppgaven defineres begrepet *praktisk arbeid* etter internasjonal litteratur innen naturfaglig utdanningsforskning:

Any type of science teaching and learning activity in which students, working either individually or in small groups, are involved in manipulating and/or observing real objects and materials (...) as opposed to virtual objects and materials such as those obtained from a DVD, a computer simulation, or even from a text-based account.

(Millar (2011) i Abrahams & Reiss, 2012, s. 1036)

Denne definisjonen er temmelig vid og omfatter flere typer praktisk arbeid utført av *elever*. Definisjonen legger ingen føringer for *hvor* den praktiske aktiviteten foregår, derimot karakteriseres praktisk arbeid av *hva* som gjøres (Abrahams & Reiss, 2012)

2.2 Et teorifundament for praktisk arbeid i naturfag

John Deweys velkjente prinsipp «learning by doing» danner grunnlaget for en pedagogisk ide som store deler av naturfagets praktiske aspekt tar utgangspunkt i (P. M. Kind, 2003).

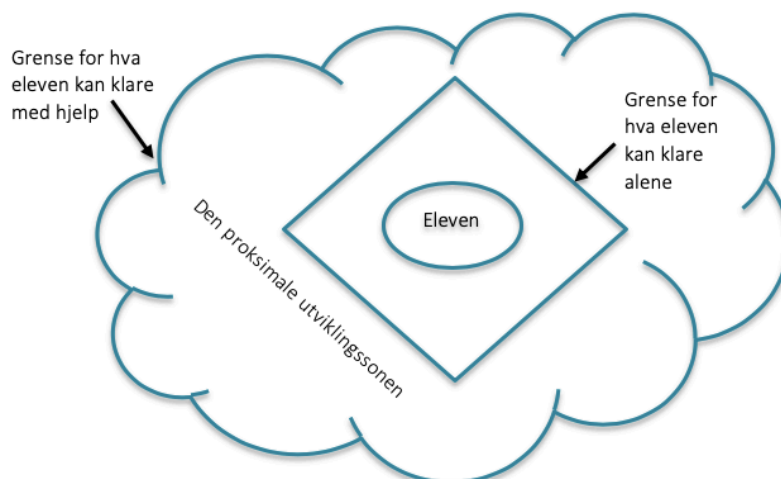
Prinsippet innebærer at læring ikke kommer av ytre stimuleringer, men ved aktiv samhandling med den fysiske verden. Læring oppstår når et individ klarer å se en sammenheng mellom handling og konsekvens (Lyngsnes & Rismark, 2014).

Et læringssyn som beskrevet over, samsvarer med en *konstruktivistisk* tankegang om hvordan kunnskap tilegnes. Dewey blir ofte omtalt som konstruktivismens far da han var en av de første i sin tid til å vektlegge elevens aktive rolle i en læringssituasjon (Imsen, 2005). Fra

midten av 1900-tallet har konstruktivismen vært et innflytelsesrikt, felles internasjonalt konstrukt både innenfor naturvitenskap og naturfagdidaktikk (Cooper & Stowe, 2018; Ringnes & Hannisdal, 2000). Konstruktivismen har mange ansikt og forgreiner seg i ulike teorier, men likevel er det én påstand som danner en felles basis: ny kunnskap dannes fra tidligere erfaringer der individet konturerer kunnskap gjennom en aktiv læringsprosess (Quale, 2003). Millar (2004) understreker viktigheten av å skille mellom konstruktivismen i vitenskapelig og didaktisk sammenheng. Der forskere gjennom tiden konstruerer kunnskap som driver naturvitenskapen fremover, konstruerer skoleelever kunnskap som allerede er anerkjent. Elevenes misjon er å gjøre den allerede kjente kunnskapen til sin egen gjennom observasjon og erfaringer.

En annen kjent skikkelse innenfor konstruktivismen er Jean Piaget og hans *kognitive konstruktivisme*. Denne konstruktivistiske retningen fokuserer på individets mentale prosesser i en læringssituasjon (Imsen, 2005). Piaget mente at mennesket organiserte tankene sine i kognitive strukturer kalt *skjema* som inneholdt individets tidligere erfaringer. Skjemaene kan ifølge Piaget endres gjennom *assimilasjon* og *akkomodasjon*, og det er kun gjennom disse prosessene at læring oppstår. Assimilasjon forekommer i en situasjon der individet bruker et foreliggende skjema til å tolke nye observasjoner og fenomener. Akkomodasjon på den andre siden, oppstår når et skjema må erstattes eller modifiseres som en følge av ubalanse mellom en ny observasjon og eksisterende erfaringer (Lyngsnes & Rismark, 2014).

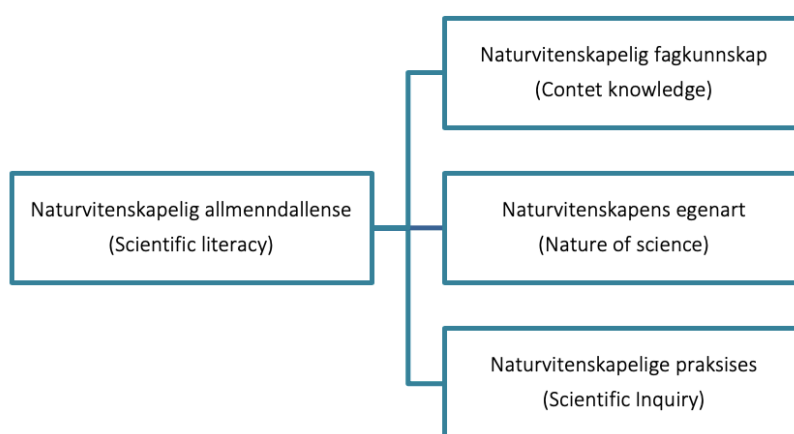
I Piagets verden er læring en personlig sak mellom individet og omverden, og dette synet har møtt en del utfordringer da det ikke tar høyde for læring i sosiale sammenhenger. Et svar på kritikken av Piagets individualistiske syn er Lev Vygotskys *sosiokulturelle konstruktivisme*, der språk og sosial samhandling betegnes som nøkkelkomponenter i læringsprosessen. Vygotsky utviklet en modell kalt den *proksimale utviklingssonen*, presentert i Figur 1, som illustrerer grenser for hva en elev kan klare på egenhånd, og ved hjelp fra andre som for eksempel en lærer. Læring oppstår når eleven befinner seg innenfor den proksimale utviklingssonen, og grensene rundt utvides etterhvert som eleven oppnår ny kunnskap som følge av hjelp og støtte (Imsen, 2005).



Figur 1: Den proksimale utviklingssonen (Imsen, 2005, s. 259)

2.3 Overordnet målsetting med naturfagsundervisning i skolen

Naturfaget verdsettes ofte som et produkt, en prosess og en sosial institusjon (Millar, 2004). Disse tre dimensjonene inngår i en *naturvitenskapelig allmenndannelse*, som på engelsk ofte betegnes *scientific literacy*. Essensen av konstruktet er at folk flest skal ha tilstrekkelig med naturfaglige kunnskaper og ferdigheter for å kunne ta selvstendige og reflekterte beslutninger om eget liv i møte med samfunnet (Sjøberg, 2011). I PISA 2006 defineres konstruktet som en sammensetting av tre komponenter: *naturvitenskapelig fagkunnskap*, *naturvitenskapens egenart* og *naturvitenskapelige praksiser* (Bybee, McCrae & Laurie, 2009). Disse komponentene er tett knyttet opp mot de tre dimensjonene nevnt ovenfor, og er alle tre forutsetning for naturvitenskapelig allmenndannelse. Denne sammenhengen er illustrert i Figur 2 med samsvarende engelsk oversettelse.



Figur 2: PISAS definisjon av naturvitenskapelig allmenndannelse (Bybee mfl., 2009)

Millar (2004) hevder at skolens læreplan har *to* eksplisitte mål med naturfagsundervisning, der allmenndannelse i naturfag er *ett* av dem. Under formålsparagrafen i læreplanen i naturfag står det følgende: «*En viktig del av allmennkunnskapen er å kjenne til at naturvitenskapen er i utvikling, og at forskning og ny kunnskap i naturvitenskap og teknologi har stor betydning for samfunnsutviklingen og for livsmiljøet*» (Utdanningsdirektoratet, 2013; NAT1-03, s. 2). Av de tre komponentene fra Figur 2, har skolen tradisjonelt sett vært orientert mot å formidle *naturvitenskapelig fagkunnskap* med fokus på naturfagets produkt-dimensjon i form av lover, begreper og teori (Abrahams & Millar, 2008). De siste årene har imidlertid undervisning *om* naturvitenskapelig kunnskap i større grad blitt vektlagt. Dette innebærer at elevene skal få en forståelse for hvordan naturfaglig kunnskap dannes og hva som karakteriserer den (P. M. Kind, 2003). Dette betegnes ofte som kunnskap om *naturvitenskapens egenart* og innebærer alle de tre dimensjonene ved naturvitenskapen. Den siste komponenten i en naturvitenskapelig allmenndannelse, *naturvitenskapelige praksiser*, er nært knyttet til prosess-dimensjonen. Lederman mfl. (2014) understreker at dette konstruktet innebærer mer enn bare det rent praktiske, og at kognitive ferdigheter også må trekkes inn.

Læreplanens *andre* overordnede mål sikter mot naturfag som yrkesforberedende (Millar, 2004). I norsk skole kan det tenkes at dette målet i større grad er rettet mot elever som tar programfag (eksempelvis kjemi, biologi, fysikk) på videregående skole, fremfor de som har naturfag som fellesfag. I læreplanen for kjemi kommer målet om rekruttering til naturvitenskapelig jobber tydelig frem: «*Kjemikere er viktige bidragsyttere i utviklingen av bioteknologi, nanoteknologi, medisin, farmasi, miljøfag, nye materialer og nye energikilder. Gjennom programfaget skal den enkelte få innsikt i hvilken betydning kjemisk forskning har for teknologisk og økonomisk utvikling*» (Utdanningsdirektoratet, 2006; KJE1-01, s. 2). PISA-rapporten fra 2015 viser at 29 prosent av norske 15-åringer ser for seg et realfaglig yrke i fremtiden. Omtrent 60 prosent av elevene sa seg enig, eller svært enig i at deres motivasjon for faget var knyttet til en fremtidig karriere (Kjærnsli & Jensen, 2016).

2.4 Praktisk arbeid i naturfagsundervisning

Praktisk arbeid i naturfagsundervisning kan spoles helt tilbake til slutten av 1800-tallet. Da ble læringsformen benyttet til å fremme praktiske ferdigheter som gagnet samfunnet i form av industri og jordbruk (Cooper & Stowe, 2018). Mye har skjedd på 200 år, og praktisk arbeid i naturfag har i dag flere ulike målsettinger og former.

2.4.1 Mål med praktisk arbeid i naturfag

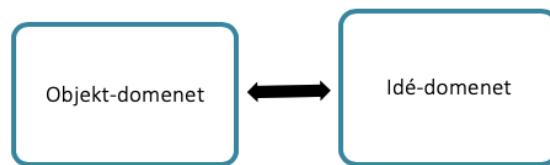
Blant forskere hersker en viss enighet om hva som er målet med praktisk arbeid i naturfagsundervisning (eks. Hofstein & Lunetta, 2004; P. M. Kind, 2003; Millar, 2004; Séré, 2002). Målene gjenspeiler de overordnede målene med naturfagsundervisning i skolen fra forrige kapittel, samt de tre komponentene i en naturvitenskapelig allmenndannelse i Figur 2. I tilfeldig rekkefølge kan målene oppsummeres slik:

1. Elevene skal få en forståelse for naturvitenskapelige begreper, teorier og modeller (kunnskap om *naturvitenskapelig fagkunnskap*)
2. Elevene skal lære seg praktiske ferdigheter i form av prosedyrer og bruk av instrumenter, samt forbedre evnen til problemløsning (kunnskap om *naturvitenskapelige praksiser*)
3. Elevene skal lære hvordan naturvitenskapelig kunnskap etableres og hva som karakteriserer slik type kunnskap (kunnskap om *naturvitenskapens egenart*)
4. Elevene skal oppleve interesse og motivasjon for faget

I læreplanens *Formål* og under hovedområdet *Forskerspiren* kommer de fire målene frem: «*Å arbeide både praktisk og teoretisk i laboratorier og i naturen med ulike problemstillinger er nødvendig for å få erfaring med og utvikle kunnskap om metoder og tenkemåter i naturvitenskapen*» og «*Varierte læringsmiljøer [...] vil berike opplæringen i naturfag og gi rom for undring, nysgjerrighet og fascinasjon*» (Utdanningsdirektoratet, 2013; NAT1-03, s. 2). Mer spesifikt for kjemi, står det i læreplanen at «*utviklingen av kjemisk viten skjer i en vekselvirkning mellom eksperimenter og teori*» og at «*opplæringen i kjemi skal knytte teori til praktisk laboratoriearbeid*» (Utdanningsdirektoratet, 2006; KJE1-01, s. 2).

I 1980-årene ble det fjerde målet som omhandler affeksjonsverdien av praktisk arbeid fremhevet som en av de viktigste grunnene til å integrere praktiske aktiviteter i undervisningen. Flere studier på den tiden viste at praktisk arbeid, spesielt i kjemi, var en av grunnene til at elever ønsket å spesialisere seg innen faget på videregående (Lunetta, Hofstein

& Clough, 2007). På begynnelsen av 1990-tallet ble det et fokusskifte mot det første målet, som også er denne masteroppgavens utgangspunkt, om å styrke elevenes fagkunnskaper. P. M. Kind (2003) hevder at det er dette målet med praktisk arbeid som verdsettes høyest i skolen, selv om det siden tusenårsskiftet har blitt mer bevissthet rundt praktisk arbeid for å fremme kunnskap om naturvitenskapens egenart (mål 3). Millar, Tiberghien og Le Maréchal (2002) mener at den fundamentale hensikten med praktisk arbeid er å hjelpe elevene å koble sammen to domener: domenet av objekter og ting som kan observeres (objekt-domenet), og domenet av teoretisk kunnskap (idé-domenet). Figur 3 illustrerer denne linken mellom praksis og teori.



Figur 3: Hensikten med praktisk arbeid (Millar, Le Maréchal & Tiberghien, 1999, s. 40)

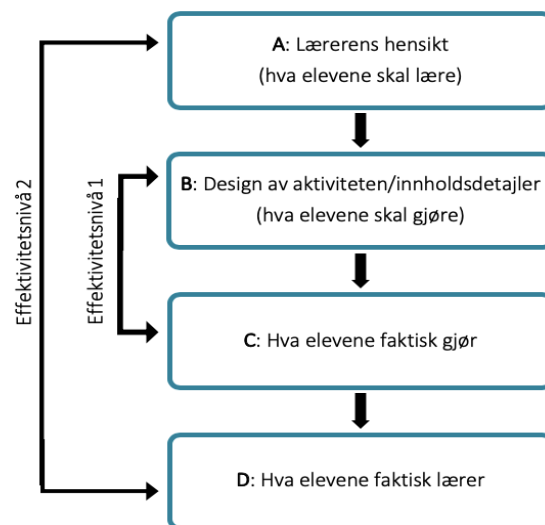
2.4.2 Praktisk arbeid som undervisnings - og læringsmetode

Gjennom *Forskerspiren* i den nåværende læreplanen og *kjerneelementene* i den kommende læreplanen, er praktisk arbeid i naturfag fremstilt som en måte elevene kan oppnå kunnskap på (Kunnskapsdepartementet, 2018b; Utdanningsdirektoratet, 2013). Frem til 1970-tallet ble praktisk arbeid sett på som en bauta som med en selvfølge skulle integreres i naturfagundervisningen. Etter som tiden gikk ble det derimot stilt spørsmål ved praktisk arbeid som en undervisnings – og læringsmetode (Hofstein & Lunetta, 1982). Lunetta og Tamir i Lunetta mfl. (2007) etterlyste i 1979 et større samsvar mellom aktivitetens mål, teorien den baserte seg på, og selve gjennomføringen. I 1990 sa Hodson (s. 175) at praktisk arbeid var både overbrukt og underbrukt i undervisningen. Overbrukt i den forstand at lærere benyttet praktisk arbeid *utelukkende* til å nå læringsmål, og underbrukt i den forstand da dets potensiale sjeldent ble utnyttet. Abrahams og Millar (2008) fant ut at praktisk arbeid var en gunstig undervisningsmetode til å få elevene til å gjøre det de skulle med utstyr og materiale slik læreren ønsket. Derimot fant de lite bevis for at praktisk arbeid hjalp elevene å knytte sammen sine fysiske handlinger med den teoretiske fagkunnskapen som lå til grunn. Hodson (1996) understreker at det ikke er godt nok å bare «gjøre noe praktisk», men at elevene i tillegg må engasjeres kognitivt når de utfører handlingene. Dette viser seg imidlertid å være mer komplisert enn først antatt. Berry mfl. (1999) rapporterer at de fleste elever fokuserer på å fullføre oppgaven fremfor å lære av den, og at de færreste i ettertid husker labforsøk de gjør

på skolen. Det er ingen tydelig sammenheng mellom hvor ofte elever gjør praktisk arbeid i naturfag og deres skår på naturfagsprøver ifølge PISA rapporten fra 2006. Likevel avdekker rapporten at det er en svak *positiv* korrelasjon mellom elever som gjør moderate mengder praktisk arbeid og deres naturfagsskår. Derimot skårer elever som gjør svært *mye* eller svært *lite* praktisk arbeid dårligere på prøver (Kjærnsli, Lie, Olsen & Roe, 2007).

2.4.3 Planlegging og evaluering av praktisk arbeid

Det gir ingen mening å bedømme effektiviteten av praktisk arbeid *generelt*. Praktisk arbeid er et paraplybegrep som innebærer et mangfold av arbeidsformer, og effektiviteten må ses i sammenheng med aktivitetens formål (Sjøberg, 2011). Millar, Le Maréchal og Tiberghien (1999) utviklet et rammeverk for å vurdere effektiviteten av en spesifikk praktisk aktivitet. De foreslo en modell som innebar planlegging – og evalueringsprosessen av en praktisk aktivitet gjennom fire steg (bokser). Modellen som er presentert i Figur 4, kan inndeles i to *effektivitetsnivå* og er oversatt til norsk. Effektivitetsnivå 1 oppfylles dersom om elevene *gjør* det de skal, mens en aktivitet er på effektivitetsnivå 2 når elevene *lærer* det de skal.



Figur 4: Planlegging og evalueringsprosessen av en praktisk aktivitet (Millar mfl., 1999)

Boks A er startpunktet av planleggingen. Her avklares lærerens hensikt med timen – hva hun eller han ønsker at elevene skal lære. Lærerens hensikt kan være ulik fra aktivitetens formelle mål. Eksempelvis kan det formelle målet være å lage en blyakkumulator, mens læreres grunntanke kan være at elevene skal få en bedre forståelse for stoffenes plassering i spenningsrekka, eller bruke et voltmeter for å måle spenning. Hart, Mulhall, Berry, Loughran og Gunstone (2000) hevder at det er hovedsakelig to grunner til praktisk arbeid mislykkes i å fremme læring blant elever. Det første problemet, som Séré (2002) også identifiseres som det

mest kritiske, er at praktiske aktiviteter ofte har alt for mange målsettinger. For mange fokuspunkter kan gjøre at elevene opplever et kognitivt kaos som undergraver læringsprosessen. Det andre problemet er at elevene ofte ikke er klar over aktivitetens formelle mål, og at deres tanker om hva som er hensikten med aktiviteten ikke stemmer overens med lærerens (Lunetta mfl., 2007). Det er lærerens oppgave å ta stilling til disse utfordringene i planleggingsfasen. Læreren må være klar over hva som forventes at han/hun skal lære bort, samt ta et standpunkt til hvilke oppfatninger han/hun ønsker at elevene skal sitte igjen med etter undervisningen (Lederman mfl., 2014).

I **boks B** velger, eller utformer læreren en praktisk aktivitet som skal hjelpe elevene å nå målene i boks A. Her skal læreren ikke bare planlegge hva elevene *fysisk* skal gjøre med objekter og material (praktisk aktivitet), men også ha en klar tanke om hva elevene skal *tenke* under aktiviteten (kognitiv aktivitet). For eksempel kan den praktiske delen være å bruke et lakmuspapir for å måle pH av ulike løsninger og observere farge på papiret, mens den kognitive delen handler om at elevene skal gjenkjenne et mønster for fargeforandringen som følge av endret pH-verdi.

Boks C er det første steget i evalueringen, og stiller spørsmål ved hva elevene *faktisk gjør* under aktiviteten både på det mentale og fysiske planet. Om elevene gjør det de skal med fysiske ting er relativt lett å observere og det kan være flere grunner til at elever mislykkes på dette punktet. For eksempel kan elevene misforstå veiledningen som blir gitt, eller utstyret kan være defekt. Dersom elevene heller ikke benytter den teoretiske kunnskapen læreren hadde forestilt seg, oppstår det også en sprik mellom boks B og C. Det er imidlertid vanskeligere å oppdage mentalt aktivitet og krever at læreren kommuniserer med elevene under aktiviteten (Abrahams, 2009). Effektivitetsnivå 1 oppfylles når elevene gjør det læreren i utgangspunktet hadde tenkt og det er et samsvar mellom boks B og C. At elevene når nivå 1 er som regel en forutsetning for å nå nivå 2. Dette kan avspeiles i den utbredte bruken av «kokebok»-forsøk der lærere gir ut detaljerte steg-for-steg instruksjoner for å sikre at elevene i det minste når nivå 1 (Millar, 2004).

I det siste steget, **boks D**, skal læreren evaluere om elevene har oppnådd de satte læringsmålene fra boks A. Dette vil i så oppfylle effektivitetsnivå 2. Millar mfl. (1999) påpeker at læring utenom de satte målene ikke tas i betraktning. I likhet med boks C, skiller også boks D mellom elevenes praktiske og mentale læringsutbytte. Eksempelvis kan det være et tilfelle der elevene kan prosedyren for tynnsljikt-kromatografi, men ikke forstår

betydningen av stoffenes polaritet. Ringnes og Hannisdal (2014) beskriver læring som en kompleks prosess som forutsetter at kunnskap lagres i langtidsmindet, og påpeker at dette kan være utfordrende å kartlegge. Samtidig vitner det om at elevenes minner fra tidligere utført praktisk arbeid begrenses til fysiske handlinger, og da er det gjerne snakk om et forsøk av ekstraordinær karakter (Abrahams & Reiss, 2012; Berry mfl., 1999).

Enkelte lærere har en idé om at elevene skal lære teoretisk fagkunnskap direkte fra praktisk arbeid (Séré, 2002), noe som P. M. Kind (2003, s. 234) omtaler som en «foreldet empiristisk og positivistisk tankegang». Gjennom et PISA+ prosjekt ble det oppdaget at også flere norske naturfagslærere delte dette synet (Arnesen & Ødegaard, 2010). Dette kan komme av at lærere ikke er klar over de mentale utfordringene elevene står ovenfor når de skal koble sammen domene av observasjoner og ideer (se Figur 3). En slik overseelse kan igjen medføre at lærere ikke er klar over sin rolle i å hjelpe elevene å lage denne koblingen (Berry mfl., 1999; Hofstein & Lunetta, 2004; Séré, 2002).

2.4.4 Ulike typer praktisk arbeid

Det finnes ingen universell måte å kategorisere ulike typer praktisk arbeid på. Praktiske aktiviteter blir ofte tilegnet navn som for eksempel labforsøk, elevforsøk eller eksperiment. Disse ordene kan imidlertid være tvetydige, noe som gjør det vanskelig å skille mellom dem. En annen måte å klassifisere praktisk arbeid er gjennom *frihetsgrader* (Kolstø & Knain, 2011). Herron (1971) beskriver frihetsgrader som et mål på aktivitetens «åpenhet» som kan deles inn i fire nivåer. Tabell 1 viser en oversikt over hans klassifiseringer som tar utgangspunkt i hvorvidt problemstilling, fremgangsmåte og resultat er kjent for elevene i forkant. Hvor mange frihetsgrader som skal tilskrives en aktivitet avhenger utelukkende av dens mål og hensikt (Millar mfl., 2002).

Tabell 1. Antall frihetsgrader i praktisk arbeid, som angitt i Herron (1971)

Frihetsgrader	Problemstilling/ hypotese	Fremgangsmåte	Resultat/ produkt
0	Gitt	Gitt	Gitt
1	Gitt	Gitt	Elev
2	Gitt	Elev	Elev
3	Elev	Elev	Elev

PISA 2006 viser at en sterk vektlegging av forsøk der elevene skal utforske egne ideer korrelerer med lave prestasjoner i faget (Kjærnsli mfl., 2007). Begrepet «discovery-learning»

brukes ofte om forsøk med 3 frihetsgrader der elevene gjennom en induktiv løsningsstrategi skal «oppdage» et fenomen. Denne tilnærmingen har møtt kritikk både for å gi et feilaktig bilde av hvordan naturvitenskapelig forskning praktiseres, samt for å være krevende å gjennomføre for mange elever (P. M. Kind, 1999). Hodson i Kolstø og Knain (2011) påpeker at det ikke kan forventes at elevene skal oppdage noe de ikke har begreper om fra før av.

En studie av Overton, Potter og Leng (2013) undersøkte hvordan elever gikk frem for å løse det de kaller «*open-ended problems*», som på norsk ofte kalles åpne eller utforskende forsøk. Elevene fikk en gitt problemstilling de skulle løse med ukjent fremgangsmåte og resultat, noe som samsvarer med et forsøk av frihetsgrad 2. Ingen kom frem til det riktige svaret. Forskerne konkluderer med at mange elever vil slite med å løse slike problemer og vil trenge støtte fra læreren. Åpne forsøk har vist seg å øke interesse og motivasjon for faget, og slike typer forsøk begrunnes ofte i at elevene får en mer positiv holdning (Lunetta mfl., 2007). Samtidig er det mulighet for at elevene lettere gir opp dersom de støter på utfordringer i en slik situasjon, og risikoen for at de utarbeider en ukorrekt fremgangsmåte som leder til feilaktige data er større (Berry mfl., 1999).

Praktisk arbeid med 0 eller 1 frihetsgrad knyttes ofte til tradisjonelle «kokebok»-forsøk der elevene får utdelt et arbeidshefte med fremgangsmåten beskrevet steg for steg. Slike forsøk kan med fordel brukes for å illustrere et spesifikt fenomen som gir elevene mulighet til observasjon (Herron, 1971). TIMSS-undersøkelsen fra 1995 viser at 89 % av naturfagslærerne brukte øvelser der prosedyren var gitt (P. M. Kind, 1999). En årsak til den utbredte praktiseringen av «kokebok»-forsøk virker å være en kombinasjon av tidsrammen til en undervisningsøkt, samt at desto mer åpen en aktivitet er, jo større pedagogiske krav stilles det til læreren (Hofstein & Lunetta, 2004). «Kokebok»-aktiviteter har blitt kritisert for å være lite mentalt givende da elevene ofte ikke trenger å «tenke selv» for å fullføre en aktivitet. Hovedfokuset deres blir gjerne på å få gjennomført opplegget fremfor å skjønne hva de gjør (Berry mfl., 1999; Séré, 2002).

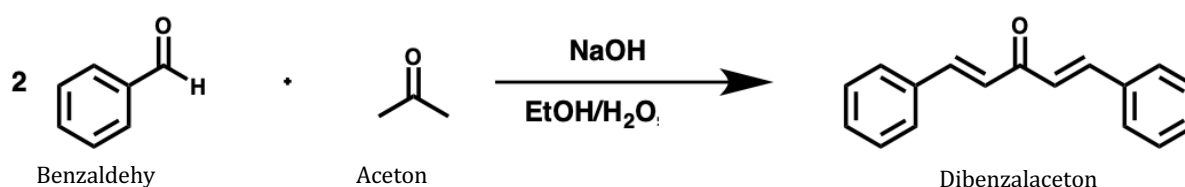
2.5 Misoppfatninger i naturfag

Mennesker konstruerer kunnskap gjennom hele livet, og i møte med den materielle verden dannes forestillinger om fenomener og sammenhenger i naturen. Det hender at disse forestillingene ikke er forenelig med akseptert vitenskap, og Ringnes og Hannisdal (2014) deler forestillingene inn i to kategorier som jeg vil bruke i denne masteroppgaven: *hverdagsforestillinger* og *misoppfatninger*. Hverdagsforestillinger utvikles *før* formell

undervisning på skolen og oppstår gjerne når et individ tolker en observasjon eller hendelse ut fra egne erfaringer. Misoppfatninger derimot, dannes *gjennom undervisning* og kan være et resultat av misvisende lærebøker eller utydelig undervisning. Misoppfatninger kan være lettere å håndtere da de er mer påvirkelige enn hverdagsforestillinger som kanskje har slått rot allerede i barndommen (Barke, Hazari & Yitbarek, 2009). På høyere faglig nivå i naturfag, slik som i *kjemi 2* på videregående, er misoppfatninger gjerne utbredt i større grad enn hverdagsforestillinger. Dette kommer av at elevene som regel ikke har noen tidligere assosiasjoner til de teoretiske fagbegrepene som kan være både komplekse og abstrakte. Dersom elevene opplever at et nytt inntrykk ikke passer med tidligere erfaringer, eller et foreliggende skjema som Piaget ville ha sagt, oppstår det en kognitiv konflikt. Det er ved en slik ubalanse at elevene er mest mottakelig for ny lærdom (Barke mfl., 2009). Etter elevenes misoppfatninger har blitt kartlagt burde de utfordres, og da kan blant annet praktisk arbeid være en hensiktsmessig metode til det formålet.

2.6 Syntese og analyse av dibenzalacetone

Dannelsen av den organiske forbindelsen dibenzalacetone (DBA) skjer gjennom en kryssaldolkondensasjon med basen natriumhydroksid (NaOH) som katalysator. Figur 5 viser den overordnede reaksjonen der aceton og benzaldehyd reagerer i en løsning av etanol og vann. DBA brukes blant annet i solkrem da det absorberer UV-stråling og dermed beskytter huden.



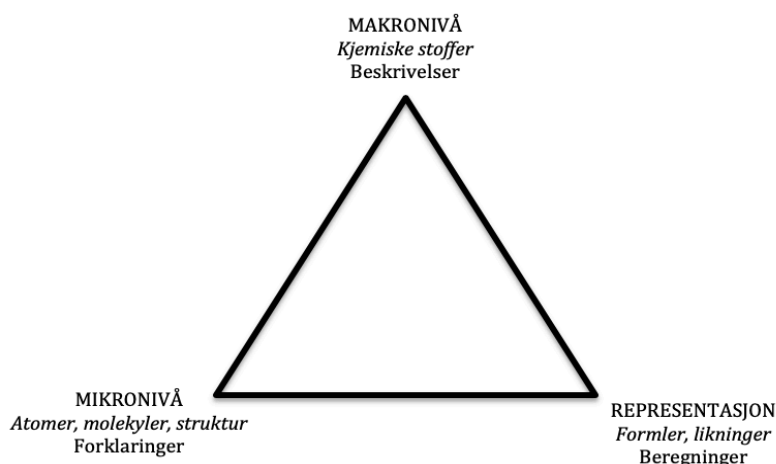
Figur 5: Reaksjonslikning for dannelse av DBA

Reaksjonsmekanismer

Kompetansemålene tilsier at elevene skal kunne redegjøre og forklare reaksjonsmekanismen for addisjonsreaksjon og eliminasjonsreaksjon. Disse reaksjonene kan igjen deles i flere underkategorier, men da dette ikke er et fokus på videregående presenteres reaksjonene i tråd med elevenes lærebok «*Kjemi 2 – Studiespesialiserende utdanningsprogram*» av Brandt og Hushovd (2011, s. 171). I boken beskrives addisjonsreaksjon som en reaksjon der et nytt stoff dannes som følge av en reaksjon mellom en umettet forbindelse (minst én dobbelt – eller trippelbinding i molekylet) og et annet molekyl som binder seg til dobbel/trippelbindingen. Boken definerer en eliminasjonsreaksjon som en reaksjon der et lite molekyl spaltes av slik at

det blir dannet en dobbeltbinding i det nye produktet. Addisjon og eliminasjon er altså to motsatte reaksjoner.

Disse reaksjonstypene forekommer mellom reaktantene i syntesen, og etterhvert som DBA dannes kan det observeres fargeforandring i løsningen fra blank til gul. Reaksjonen betraktes da på *makronivå*. Gulfargen i løsningen skyldes de konjugerte dobbeltbindingene i DBA, en forklaring av fargeendringen gjennom molekylenes oppbygning og struktur på *mikronivå*. Figur 5 illustrerer den kjemiske reaksjonslikningen i form av symboler og er en beskrivelse av reaksjonen på *representasjonsnivå*. Disse tre nivåene utgjør en modell i kjemifaget kjent som «Johnstone's triangle» og representerer tre måter å tenke kjemi på (Cooper & Stowe, 2018). Ringnes og Hannisdal (2014, s. 23) oversetter disse til «Kjemiens tre dimensjoner» som illustreres i Figur 6. Elevenes evner til å koble sammen de tre nivåene er avgjørende for deres læring i faget generelt, men spesifikt i denne masteroppgaven brukes triangelen i forbindelse med læring av reaksjonsmekanismer.



Figur 6: «Kjemiens tre dimensjoner» (Ringnes & Hannisdal, 2014)

OmkrySTALLISERING

OmkrySTALLISERING er en teknikk som brukes for å isolere og rense stoffer. Metoden utnytter at løseligheten av et stoff påvirkes av temperaturendringer. Råproduktet som skal renses blir tilsatt et passende løsningsmiddel som igjen varmes opp mot kokepunktet slik at produktet løses opp. I omkrySTALLISERING av DBA er etanol et godt løsemiddel, og for å unngå at for mye av produktet løses opp, selv i avkjølt løsemiddel, brukes en fortynnet 70%- etanol blanding. Etterhvert som temperaturen stiger øker produktets løselighet, og når alt fast stoff er løst opp avkjøles løsningen sakte for å oppnå mest mulig rene krystaller. Ved en forhastet avkjøling risikeres det at urenheter «fanges» i det nydannede krystallgitteret (Mohrig, Alberg,

Hofmeister, Schatz & Hammond, 2014). Løseligheten til produktet avtar gradvis med avkjølingen og omdannes til fast stoff igjen gjennom krystallisering. Samtidig forblir eventuelle urenheter oppløst i løsningsmidlet og kan filtreres bort (Mohrig mfl., 2014).

NMR og MS

Elevenes produkt, DBA, analyseres ved de instrumentelle teknikkene *kjernemagnetisk resonansspektroskopi* (NMR) *massespektroskopi* (MS). Analysemetodene brukes primært for å bestemme molekylstrukturer, men kan også brukes for å oppdage eventuelle urenheter i produktet. NMR baseres på atomkjerners absorpsjon av lavenergetisk elektromagnetisk stråling i radiofrekvensområdet. I NMR brukes atomkjerner med oddetall protoner og/eller nøytroner, da disse kjernene har *spinn* (Mohrig mfl., 2014). En mye brukt atomkjerne er kjernen til hydrogenisotopen protium (^1H), og i *kjemi 2* introduseres elevene kun for

^1H -NMR. Det er fire typer informasjon som må tas i betraktning ved analyse av et ^1H -NMR-spekter (Mohrig mfl., 2014):

- Antall signaler: indikerer antall ulike «hydrogenmiljø»
- Arealet under signalene: indikerer det relative antallet hydrogenatomer i gruppen som gir opphav til signalet
- Signalenes posisjon på den horisontale akse: viser kjemisk skift og kan indikere funksjonelle grupper. Slik informasjon finnes i tabeller.
- Hvor mange topper signalet er splittet i: kalles finstruktur og indikerer hvor mange *nabo*hydrogenatomer gruppen som gir opphav til signalet har.

MS tar utgangspunkt i atommasse. Helt grunnleggende forklart ioniserer et massespektrometer molekyler i en forbindelse, og splitter de i mindre fragment-ioner som skilles basert på atommasse. Den horisontale akse på et massespekter representerer ratioen mellom ionets masse og ladning, men siden ladningen på ioner ofte er +1, samsvarer ratioen med atommassen i dalton (u). Den vertikale akse indikerer relativ forekomst av hvert fragment-ion (Mohrig mfl., 2014). I en analyse av et massespekterspektrum forsøkes det å identifisere hvilke topper i spekteret som representerer de ulike fragment-ionene molekylet ble splittet i.

3 Metoder og materiale

Dette masterprosjektet er en kvantitativ studie med et eksperimentelt forskningsdesign. For å samle inn data ble det benyttet en pretest, posttest, et labhefte og observasjonsnotater. Testene og labheftet ble summativt vurdert basert på poengskår og analysert ved statistiske metoder i form av deskriptiv statistikk og statistisk inferens. Dette kapitlet vil også presentere studiens utvalg, samt beskrivelsen av undervisningsopplegget i sin helhet. Til slutt diskuteres studiens validitet, reliabilitet og etiske overveielser knyttet til gjennomføringen av masterprosjektet.

3.1 Begrunnelse for valg av metode

Tradisjonelt sett skiller det mellom *kvalitative* og *kvantitative* tilnærminger til forskning. Førstnevnte er ute etter en dypere forståelse av det som studeres, gjerne med få deltakere, mens kvantitative metoder er opptatt av å tallfeste data fra et større utvalg (Bjørndal, 2017). Hvilke av de to tilnærmingene som burde benyttes avhenger av forskningsspørsmålets natur og hva som skal undersøkes (Creswell, 2014). Med utgangspunkt i denne avgjørelsen bestemmes forskningens design og metode.

På bakgrunn av antall informanter (78 elever), samt at datamaterialet hovedsakelig består av testskår i form av tall, har studien et *kvantitativt* preg. Jeg var ute etter å sammenligne testresultater med et fokus på systematikk og kontroll i gjennomføringsfasen, noe som karakteriserer en kvantitativ forskningsstrategi (Bjørndal, 2017). Innenfor kvantitativ forskning skiller det ofte mellom to typer forskningsdesign: *ikke-eksperimentell* og *eksperimentell*. Ikke-eksperimentelle design er ute etter å kvantifisere og beskrive egenskapene til en populasjon, mens eksperimentelle design søker svar på om en spesifikk behandling eller intervensjon påvirker et utfall (Creswell, 2014). Studiens forskningshypotese går ut på å undersøke hvordan elevenes fagkunnskaper blir påvirket av et praktisk undervisningsopplegg, og derfor kan tilskrives et *eksperimentelt* forskningsdesign.

Det finnes ingen eksplisitt betegnelse av begrepet *metode* innenfor forskning. Creswell (2014) definerer begrepet som måten data samles inn og analyseres på, og det er denne beskrivelsen av forskningsmetode jeg har valgt å benytte i masteroppgaven. Det ble i hovedsak samlet inn data gjennom en *pretest* før forsøket, og en *posttest* etter forsøket. Disse var utformet som skriftlige tester der oppgavene tok utgangspunkt i læreplanen i kjemi. For å kunne besvare forskningshypotesen går jeg i denne masteroppgaven går jeg ut fra at elevenes testskår til en

viss grad reflekterer deres forståelse. I tillegg gjorde jeg observasjoner under selve labforsøket og samlet inn elevenes labhefter som supplerende datakilder.

Craig i Cooper og Stowe (2018) uttrykker at det kan være utfordrende å gjøre en fullstendig vurdering av elevenes kunnskap basert på tradisjonelle tester. Samtidig har skriftlige tester og prøver lenge vært, og er enda, den mest brukte måten å kartlegge elevenes nåværende kunnskaper på (Engh, 2014). Eksempler på dette er nasjonale prøver, eksamen, og internasjonale undersøkelser som PISA og TIMSS. Ettersom kvantitativ forskning ofte jobber med et relativt stort antall informanter, kreves det at data samles inn på effektiv og oversiktlig måte. Av den grunn ligger det i kvantitativ forskningstradisjon å anvende tester som måleinstrument (Engh, 2014). Da testene og labhefte ble vurdert i form av numeriske poengskår var det mest hensiktsmessig å bruke *statistiske metoder* for å analysere datamaterialet. Teorien bak metodene som beskrives i kapittel 3.5 er omfattende, men nødvendig for å få en forståelse for resultatenes betydning. Hvordan testene og labheftet ble *vurdert* presenteres i kapittel 3.4.

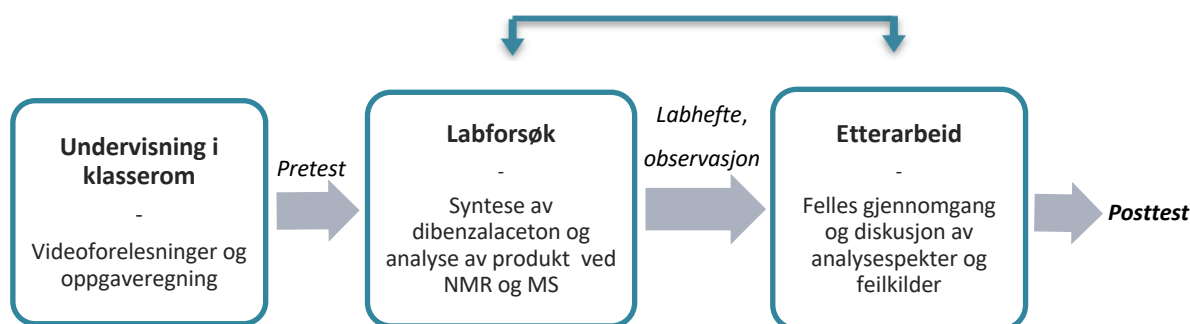
3.2 Utvalg

Problemstillingen som undersøker elevers læring av fagbegreper i organisk kjemi, kan sies å omfatte alle elever som tar programfaget *kjemi 2* ved en norsk videregående skole. Disse elevene utgjør enheter i *populasjonen* av interesse, men ettersom det ikke er mulig å undersøke alle enhetene, velges en mindre gruppe fra populasjonen kalt *utvalg* (Lund, Fønnebo & Haugen, 2006). I 2018 var det 5356 elever som fikk standpunktkarakter i *kjemi 2* (Utdanningsdirektoratet, 2018) og studiens utvalg på 78 elever utgjør dermed 1.5 % av populasjonen.

I utvelgelsen av informanter i denne masteroppgaven var det viktig at de skulle ha visse egenskaper som var representative for populasjonen, men også at de gjenspeilet problemstillingen direkte. Informantene ble innhentet med utgangspunkt i at de studerte *kjemi 2* og betegnes av Thagaard (2018) som en *strategisk utvalgsmetode*. Da jeg ikke hadde mulighet til å reise utenbys ble kun informanter innenfor et visst område forespurt om å delta i masterprosjektet. Dette kan beskrives som et *tilgjengelighetsutvalg*, og står i kontrast til et sannsynlighetsutvalg der hver enhet i populasjonen har like stor sannsynlighet for å bli valgt (Thagaard, 2018). Av pragmatiske, naturlige årsaker bestod utvalget av faste skoleklasser fremfor enkeltelever fra ulike skoler.

3.3 Gjennomføring og innsamling av data

Prosjektet ble initiert ved å opprette kontakt med alle kjemilærere innenfor et visst område. Etter å ha avgjort hvilke klasser som skulle delta satt jeg opp, i samråd med de respektive faglærerne, en tidsplan for utføring av prosjektet i hver klasse. For å oppnå høyest mulig kontroll over prosjektet inntok jeg rollen som både lærer og forsker, og faglærerne hadde dermed ingen faglig innvirkning på elevene i forskningsperioden. Masteroppgavens problemstilling innebærer at elevene skal gjennomgå et *undervisningsopplegg* som består av tre trinn: *undervisning*, et *praktisk labforsøk*, og *etterarbeid*. Labforsøket og etterarbeidet er en del av den *helhetlige forskningsprosessen* som er beskrevet i kapittel 1.2. I løpet av undervisningsopplegget ble det samlet inn datamateriale i form av *pretest*, *posttest*, *labhefte* og *observasjonsnotater*. Læreplanen i kjemi var grunnsteinen i utformingen av alle elementene i gjennomføringsprosessen illustrert i Figur 7.



Figur 7: Undervisningsopplegg og datainnsamlingsmetoder

Jeg ønsker å presisere at på bakgrunn av studiens omgang og utfordringer knyttet til å få nok informanter i utgangspunktet, ble det i samråd med mine veiledere besluttet å ikke ta i bruk en kontrollgruppe. I stedet ble det prioritert å gjennomføre undervisningsopplegget på flest mulig elever slik at jeg skulle få nok data til en kvantitativ undersøkelse. Det finnes imidlertid mye forskningslitteratur på elevers læringsutbytte av praktisk arbeid, og problemstillingen vil derfor blir drøftet i lys av dette.

3.3.1 Undervisningsopplegg

Læringsutbytte fra praktisk arbeid kan styrkes dersom elevene i forkant har kjennskap til de teoretiske ideene som skal brukes under aktiviteten (Abrahams & Millar, 2008; Abrahams & Reiss, 2012; Berry mfl., 1999; Hart mfl., 2000; Séré, 2002). Det ble derfor gjennomført klasseromsundervisning forut av labforsøket i form av videoforelesninger. Videoene ble spilt inn og redigert ved UiTs eget filmstudio *Ressurssenter for undervisning, læring og teknologi*

(Result). Videoenes innhold ble planlagt og strukturert i samarbeid med mine veiledere og resulterte i fem små videosnutter der jeg presenterer fagbegrepene ved bruk av en digital tavle. Videoundervisningen sikret at elevenes møte med fagstoffet i klasserommet var likt på tvers av klassene. I tillegg hadde alle elevene samme lærebok fra Brandt og Hushovd (2011) og ble oppfordret til å gjøre spesifikke oppgaver jeg hadde plukket ut på forhånd i tilknytning til videoforelesningene. Klasseromsundervisningen gikk over totalt seks undervisningsøkter på 45 minutter.

3.3.2 Pretest

Etter at klasseromsundervisning var avsluttet, gjennomførte elevene en pretest samme morgen som labforsøket skulle utføres. Hensikten med pretesten var å kartlegge elevenes fagkunnskap om fagbegrepene etter *kun* undervisning. Pretesten er å finne i Vedlegg 4 og bestod av 6 oppgaver som totalt ga 30 poeng. I tillegg til læreplanen, ble gamle eksamensoppgaver i kjemi fra Utdanningsdirektoratets nettsider og elevenes lærebok brukt som inspirasjon. Testens innhold ble rådført med veiledere mine og pilotert av to medstudenter hvor den ene hadde kjemibakgrunn fra høyere utdanning, mens den andre kun hadde *kjemi 2* fra videregående. Piloteringen var meget nyttig med tanke på testens omfang i form av tidsbruk samt oppgavenes «vanskelighetsgrad», og testen ble modifisert i henhold til tilbakemeldingene. Elevene fikk 45 minutter til disposisjon for å gjøre testen som ble samlet inn fortløpende.

3.3.3 Labforsøk

Fundamentet i masteroppgaven er at det praktiske forsøket, syntese av DBA, skulle gjennomføres som en del av den *helhetlige forskningsprosessen*. Gjennom forøket skulle elevene være med på å skape noe eget som de skulle både bearbeide og analysere. Etter Herrons (1971) klassifisering av praktisk arbeid har studiens labforsøk null frihetsgrader der både problemstilling, fremgangsmåte og produkt var kjent for elevene. Laboratoriefasilitetene på de respektive skolene var til dels gammeldagse med mangel på både utstyr, avtrekksskap og kjemikalier. Dårlig standard på skolelaboratorium er ifølge Roberts (2002) et gjennomgående problem som kan hemme elevens læringsutbytte av praktiske aktiviteter. Av den grunn fant labforsøkene sted på et laboratorium i Forskningsparken (tilknyttet UiT) med tilstrekkelig materiell og plass. Labforsøket beskrevet i kapittel 2.6 ble valgt på anbefaling fra veilederne mine da det lar seg gjennomføre innenfor en rimelig tidsramme, og som regel gir et akseptabelt utbytte av produkt, så fremt prosedyren følges.

Hensikten med labforsøket var aldri at elevene skulle lære fagbegrepene *direkte* fra gjennomføringen. Det er åpenbart at elevene ikke kan lære teorien bak en reaksjonsmekanismene ved å observere en løsning på makronivå. Dette gjelder også tolkning av NMR og MS spekter. Labforsøket ble derimot brukt for å skape assosiasjoner til teorien fra videoforelesningene, og jeg ønsket å skape «knagger» elevene kunne henge teoretisk fagkunnskap på.

3.3.4 Labheftet

Elevene jobbet i par under labforsøket slik at de kunne samarbeide, men fikk likevel utdelt hvert sitt labhefte. Labheftet er å finne i Vedlegg 6 og inneholder følgende: en beskrivelse av forsøkets prosedyre, 10 refleksjonsspørsmål, en «resultatdel» der elevene skulle beregne utbytteprosent, samt en diskusjonsdel der feilkilder skulle drøftes. Labheftets trinnvise forklaring kan bære preg av den beryktede «kokebok»-stilen som ofte får kritikk for å være «mekanisk», lite mentalt utfordrende og umotiverende (Berry, Gunstone, Loughran & Mulhall, 2001; Hart mfl., 2000) Jeg ønsket derfor, i den grad det var mulig, å unngå en slik tilnærming ved å i tillegg legge til spørsmål etter hvert steg som ga grobunn for refleksjon.

3.3.5 Observasjon

Bjørndal (2017, s. 33) beskriver observasjon som en «oppmerksom iakttakelse» og skiller mellom to typer observasjon avhengig av forskerens deltakelse i situasjonen. Under labforsøket tok jeg i bruk *observasjon av andre orden*. Det er en form for observasjon der forsker selv er en del av situasjonen, og står i motsetning til observasjon av første orden der observasjon er forskerens primære oppgave. Observasjon kan foregå under strenge, strukturerte forhold der observasjoner noteres ned i forutbestemte kategorier, eller den kan være åpen med et vidt fokus (Bjørndal, 2017). Jeg ønsket i hovedsak å observere hva elevene gjorde under forsøket i form av fysiske handlinger, samt hvilke spørsmål som ble stilt i tilknytning til forsøket eller labheftet. Etter hver gjennomføring av forsøket skrev jeg ned observasjoner jeg betraktet som relevante i det Bjørndal (2017) betegner som en ustrukturert logg uten forhåndsbestemte kategorier. Det er mange åpenbare svakheter med en slik åpen og ustrukturert tilnærming. Blant annet fikk jeg ikke med meg alt som foregikk da hver klasse bestod av minst 20 elever fordelt på to rom. Jeg kan også ha glemt observasjoner og ha skrevet feilaktige notater i loggen som en følge av dette. I tillegg var observasjonene basert utelukkende på min subjektive oppfatning av situasjonen. Data fra observasjon i denne masteroppgaven blir brukt som en supplerende metode for å få et bilde av hva elevene fysisk gjorde på laben og vil derfor kun bli brukt til å drøfte nettopp dette.

3.3.6 Etterarbeid

Vygotskys erkjenner at læring og utvikling er noe som oppstår i sosiale kontekster (Imsen, 2005). Da analysespektrene var klare dro jeg derfor tilbake til klassene og gjennomgikk de sammen med elevene. Særlig kunne spekteret fra NMR være komplisert å tolke for elevene alene da det inneholdt kompleks-splitting, et tema utenfor læreplanens kompetanser. Klassene diskuterte også eventuelle feilkilder i forsøket og hvordan dette kunne gjenspeiles i analysespektrene og utbytte av produktet.

3.3.7 Posttest

Posttesten ble gjennomført i påfølgende kjemitime, omtrent én uke etter labforsøket. Testen som er å finne i Vedlegg 5 ble også pilotert på medstudenter og inneholdt samme antall spørsmål og poeng som pretesten. Intensjonen med posttesten var å teste de samme fagkunnskapene som pretesten for å kunne sammenligne testskår. Posttesten inneholdt også et evalueringsskjema der elevene ble bedt om å krysse av for hvilken arbeidsmetode de foretrakk for å lære om de ulike fagbegrepene (addisjonsreaksjon og eliminasjonsreaksjon ble slått sammen til én kategori, *reaksjonsmekanismer*). Alternativene var følgende: *lese selv*, *praktisk forsøk*, *undervisning*, *undervisning + forsøk* og *annet*.

3.4 Vurdering av tester og labhefte

Testene og labheftet ble vurdert *summativt* med hensikt å undersøke elevenes fagkunnskaper på et visst tidspunkt. Dysthe (2008) beskriver summativ vurdering som vurdering *av* læring der formålet er å fastsette en karakter eller poengskår. For å vurdere kvaliteten av kompetanse i et fag har UDIR utviklet nasjonale, veiledende *kjennetegn på måloppnåelse* på bakgrunn av kompetansemålene i læreplanen. Disse kjennetegnene er ment å gi felles retningslinjer for vurderingsarbeid (Utdanningsdirktoratet, 2016, s. 1). For *kjemi 2* er kjennetegn på måloppnåelsene ganske generell (Utdanningsdirektoratet, 2019), og av den grunn utarbeidet jeg egne retningslinjer for å vurdere testene. Retningslinjene som jeg har valgt å kalle *Vurderingskriterier* finnes i Vedlegg 3, og er spesifikt rettet mot de fem fagbegrepene.

Poengene på hver oppgave ble vektlagt basert på hvor mange, og i hvilken grad elementer fra *Vurderingskriteriene* måtte anvendes for å besvare spørsmålet. Eksempelvis krevde oppgaven om NMR at elevene kunne skille mellom ulike hydrogenmiljø, si noe om arealet under toppene, slå opp kjemisk skift i tabell og bruke n+1-regelen. Dersom eleven klarer å sette sammen denne informasjonen for å besvare spørsmålet betegnes det som *høy måloppnåelse*. Hvis eleven bruker et par av begrepene kategoriseres det som *middels måloppnåelse*, mens

lav måloppnåelse bedømmes dersom svaret er både ukorrekt og mangelfullt. En slik oppgave som NMR-oppgaven vil naturligvis «veie» mer enn en avkryssingsoppgave

Poengfordelingen av refleksjonsspørsmålene i labheftet ble vektlagt på lik vis som i testene, og endte opp på totalt 26 poeng. Spørsmålene måtte derimot vurderes på en litt annen måte ettersom enkelte av spørsmålene innebar elementer utenfor *Vurderingskriteriene*. Det ble ikke laget egne retningslinjer for vurdering av labheftet da dette kun ble brukt som en supplerende datakilde. Spørsmålene ble vurdert i grad av korrekthet og *Vurderingskriteriene* ble trukket inn der det var hensiktsmessig.

3.5 Statistiske analysemetoder

Innsamlet data som består av pretest, posttest og labheftet er av kvantitativ art, og av den grunn var det naturlig å anvende statistiske analyser på datamaterialet. Analysene ble utført i IBMs programvare *SPSS Statistic Versjon 25* og alle resultater og analyser ble vurdert på et $\alpha = 0.05$ signifikansnivå.

Statistikk handler om å tallfeste observasjoner og analysere data, og kan deles inn i to hovedgreiner: *deskriptiv statistikk* og *statistisk inferens* (Lund mfl., 2006). I denne masteroppgaven brukes deskriptiv statistikk i form av *gjennomsnitt*, *median*, *standardavvik* og *korrelasjonsanalyse* for å beskrive og kvantifisere innsamlet data. Den andre retningen, statistisk inferens, handler om å utlede *statistiske slutninger* fra utvalg til populasjon (Martin & Bridgmon, 2012) og brukes i denne studien i form av *statistisk signifikanstesting* og *konfidensintervall*. Da problemstillingen omhandler elevenes fagkunnskaper før og etter undervisningsopplegget, er det i hovedsak **skår differansen δ** mellom testene som er av interesse. Statistisk inferens vil derfor gjennomføres på utvalgets *gjennomsnittlige skår differanse* for å kunne vurdere i hvilken grad resultatene kan videreføres til hele populasjonen.

3.5.1 Deskriptiv statistikk

Deskriptiv statistikk brukes for å beskrive egenskaper eller karakteristikk ved et datasett og gir for øvrig ingen indikasjoner på om data er «bra» eller «dårlig». Gjennomsnitt og median er to typer *sentralmål* som gir en typisk verdi for datasettets midtpunkt, mens standardavvik er et *spredningsmål* som beskriver spredning av data rundt midtpunktet. Hva disse målene innebærer kan leses i Tabell 2, og i denne masteroppgaven brukes de til å beskrive elevenes testskår på tvers av klassene, samt samlet skår for utvalget.

Tabell 2. Beskrivelse av sentralmål og spredningsmål

Gjennomsnitt	Beregnes ved å summere alle verdiene i et datasett og dividerer på antall elementer. En fordel med gjennomsnitt er at det inkluderer alle verdiene i datasettet. Samtidig kan dette være en ulempe da ekstreme verdier kan bidra til å gi et feilaktig bilde av dataen. I denne studien brukes gjennomsnitt til å beregne gjennomsnittskår for pretest og posttest, samt gjennomsnittlig differanse i skår for utvalget.
Median	Median er den midterste observasjonen av data sortert i stigende rekkefølge. Dersom antall observasjoner er partall tas gjennomsnittet av de to midterste observasjonene. En ulempe med median er at den ikke bruker all informasjon fra datamaterialet og at den kan «skjule» ekstreme verdier i datasettet. I denne studien sammenlignes median med gjennomsnittet for å vurdere om data er normalfordelt.
Standardavvik	Beregner dataverdiene gjennomsnittlige avstand fra gjennomsnittet, og indikerer spredning og variasjon i datasettet. Standardavviket kan altså gi en indikasjon på fordelingen av skår.

Korrelasjon

I motsetning til sentral – og spredningsmål som opererer med *en* variabel, måler en korrelasjonsanalyse styrken og retningen mellom *to* variabler, og sier noe om hvordan disse variablene forholder seg til hverandre (Ross, 2014). Korrelasjon måles ofte i en korrelasjonskoeffisient og i denne studien brukes *Pearson korrelasjonskoeffisient* (ρ). ρ måler *lineær* korrelasjon og har en verdi mellom -1 og +1. -1 tilsvarer en perfekt negativ korrelasjon, en verdi på 0 indikerer ingen korrelasjon, og +1 er en perfekt positiv korrelasjon. Martin og Bridgmon (2012) beskriver korrelasjonsstyrken som lav ved verdi $\rho \leq \pm 0.39$, moderat for verdier mellom ± 0.40 og ± 0.69 og høy ved verdier for $\rho \geq \pm 0.70$, og det er disse verdiene som brukes i masteroppgaven. En negativ ρ korresponderer med en *negativ* korrelasjon der to variabler beveger seg i motsatt retning av hverandre. Det vil si at når én variabel øker, minker den andre. En *positiv* korrelasjon er tilfellet der to variabler beveger seg i samme retning. En lav ρ -verdi betyr ikke nødvendigvis at det ikke er noen sammenheng mellom variablene, men bare at forholdet ikke er *lineært*. Derfor kan det i tillegg være nyttig å fremstille observasjonene grafisk et spredningsplott. Det er viktig å påpeke at korrelasjon ikke sier noe om årsakssammenheng, og at en variabel nødvendigvis ikke er en konsekvens av en annen (Ross, 2014). I dette studiet ble det utført korrelasjonsanalyse for å undersøke tallmessige sammenhenger mellom skår på labheftet og posttesten, og mellom pretesten og posttesten. En slik analyse vil eksempelvis gi en indikasjon på om elever som skårer høyt på pretesten også skårer høyt på posttesten eller motsatt.

3.5.2 Statistisk inferens

I statistisk inferens utføres det målinger på ett eller flere utvalg for å kunne trekke slutninger om de *samme* målingene på vegne av populasjonen (Martin & Bridgmon, 2012). For å finne ut hvor sikker disse målingene er, brukes statistiske tester som beregner *sannsynligheten* for at målingene oppstod *tilfeldig* (Creswell, 2014). Dersom denne sannsynligheten er liten, er resultatet *statistisk signifikant*. I denne masteroppgaven ble statistisk inferens benyttet til å beregne utvalgets *gjennomsnittlige* skår differanse (δ) som heretter vil bli betegnet **utvalgsdifferansen** \bar{x}_δ . Hvis signifikanstesten, *paret t-test*, kalkulerer sannsynligheten for at \bar{x}_δ er et sporadisk resultat til under 5 % (bestemt av $\alpha = 0.05$), kan \bar{x}_δ sies å være statistisk signifikant. I tillegg ble det generert et *95% - konfidensintervall* som estimerer et verdiområde der populasjonens gjennomsnittlige δ , betegnet **populasjonsdifferansen** μ_δ , *sannsynligvis* er å finne.

Paret t-test og hypoteser

Den statistiske signifikanstesten *paret t-test* brukes i situasjoner der data som kommer fra samme utvalg skal sammenlignes (Martin & Bridgmon, 2012). Denne testen egner seg godt i denne studien der gjennomsnittsskåren mellom pretest og posttest utført av samme individ skal overveies. Statistiske signifikanstester opererer ofte med *hypoteser*, som er påstander eller antakelser om egenskapene til en populasjon (Kristensen & Wikan, 2016). Paret t-test har to konkurrerende hypoteser: *nullhypotesen* H_0 og den *alternative hypotesen* H_1 . H_0 antar *ingen* differanse mellom gjennomsnittlig pretest og -posttestskår, mens den alternative hypotesen H_1 antar det motsatte. I stedet for å bekrefte en av hypotesene, tar paret t-test utgangspunkt i at H_0 er sann og avgjør sannsynligheten for at dette stemmer. Denne sannsynligheten kvantifiseres gjennom en p-verdi, og dersom p-verdien er lavere enn signifikansnivået forkastes H_0 . I tillegg genererer paret t-test en *t-verdi* som indikerer standardfeilen, eller feilmarginen, fra *forventet* verdi gitt at H_0 er sann. Denne t-verdien sammenlignes igjen opp mot en *teoretisk* t-verdi fra t-fordelingen $t_{\alpha/2}$ (Ross, 2014). Tabell 3 gir en oversikt over de to hypotesene til paret-test, og under hvilke forutsetninger H_0 forkastes. Det oppgis forkastningsområde for generelle tilfeller, men også for hvilke spesifikke t-verdier og p-verdier som gjelder med et signifikansnivå på $\alpha = 0.05$.

Tabell 3. Mulige utfall av paret t-test

Null hypotese H_0	Alternativ hypotese H_1	Forkaster H_0 når
$\mu_\delta = \mu_{\text{posttest}} - \mu_{\text{pretest}} = 0$	$\mu_\delta = \mu_{\text{posttest}} - \mu_{\text{pretest}} \neq 0$	1. $t_{\frac{\alpha}{2}} < t < -t_{\frac{\alpha}{2}}$ 2. $p < \alpha$
Ingen forskjell i skår mellom pretest og posttest	Forskjell i skår mellom pretest og posttest	1. $1.99 < t < -1.99$ 2. $p < 0.05$

Dersom p-verdien er lavere enn 0.05 er sannsynligheten for å observere utvalgsdifferansen \bar{x}_δ under nullhypotesens antakelser (ingen endring i skår), så liten at H_0 forkastes (Lund mfl., 2006). I tillegg vil en slik p-verdi kunne indikere at det er en *statistisk signifikant differanse* i skår mellom pretest og posttest, samt at dette resultatet mest sannsynlig ikke er sporadisk og oppstod ved en tilfeldighet. I forbindelse med hypotesetesting vil det imidlertid alltid være en viss sannsynlighet for å trekke feil konklusjon om hvilken hypotese som forkastes.

Sannsynligheten for å feilaktig forkaste H_0 er lik signifikansnivået (Ross, 2014), i dette tilfellet 5 %. Sannsynligheten for å feilaktig beholde H_0 er derimot mer komplisert å beregne og vil ikke utdypes på bakgrunn av at denne situasjonen ikke oppstod i studiens tilfelle.

Konfidensintervall

Konfidensintervall er tett knyttet opp til hypotesetesting og kan også avgjøre om nullhypotesen H_0 skal forkastes. Et konfidensintervall angir et intervall, eller verdiområde, som med en viss sannsynlighet inneholder populasjonens sanne verdi. Dersom verdien av H_0 ligger innenfor et $(1-\alpha)\%$ konfidensintervall, kan ikke H_0 forkastes på α -nivå (Ross, 2014). Da studien opererer med et α -nivå = 0.05 ble det kalkulert et 95%- konfidensintervall for å angi et intervallområde der populasjonsdifferansen μ_δ med 95% sannsynlighet kan forventes å bli funnet. Med andre ord gir konfidensintervallet en indikasjon på hvilken skår differanse (δ) elevene *sannsynligvis* oppnår fra pretest til posttest. Dersom verdien av H_0 (null) befinner seg innenfor konfidensintervallet impliserer det at elevene ikke hadde noen endring i skår fra pretest til posttest.

Forutsetninger for paret t-test

Paret t-test bygger på antakelser om kontinuerlige variabler, at skår differansen (δ) er tilnærmet normalfordelt og at det ikke er noen outliers i δ . Dersom disse kravene ikke er oppfylt i tilstrekkelig grad kan det true den statistiske validiteten (Lund mfl., 2006). Elevenes skår differanser er kontinuerlige da δ kan ha alle verdier innenfor et visst intervall. Eventuelle

outliers kan oppdages gjennom analyse med SPSS gjennom et boksplokk. Det er imidlertid mer omfattende å påvise en normalfordeling. *Sentralgrenseteoremet*, et svært fundamentalt teorem i statistikk, sier at fordelingen av utvalgsgjennomsnittet vil gå mot en normalfordeling dersom utvalgsstørrelsen er *stor nok*. Dette vil skje uavhengig av sannsynlighetsfordelingen til populasjonen (Martin & Bridgmon, 2012). Da det er usikkerhet knyttet til om utvalgsstørrelsen i denne studien er tilfredsstillende for teoremet, var det nødvendig å kjøre en normalitetstest av datasettet i SPSS. SPSS sin normalitetstest genererer *histogram*, *normalfordelingsplokk* (Q-Q plot) og verdier for *skjevhet* og *kurtose*.

Dersom normalfordelt data fremstilles i et *histogram* vil dataen ta form som en symmetrisk «bjelle». Et histogram vil gi en generell, visuell beskrivelse av fordelingen av data, og vil ikke i seg selv være tilstrekkelig for å avgjøre normalitet. (Kristensen & Wikan, 2016). I et *Q-Q plokk* plottes hver observasjonsverdi opp mot med dens forventede verdi i en normalfordeling. SPSS kalkulerer lineær regresjon der punktenes avvik fra linja avgjør normalitet (Martin & Bridgmon, 2012). *Skjevhet* og *kurtose* er mål på henholdsvis asymmetri og «spisshet» i en sannsynlighetsfordeling. For å avgjøre om fordelingen har en signifikant skjevhet og kurtose i forhold til normalfordelt data, beregnes en standardiskår som testes på α -nivå (Eikemo & Clausen, 2012). I en normalfordeling vil 68% av datasettet være ± 1 standardavvik (s) fra gjennomsnittet. Halvparten av dette, altså 34 %, havner $+1s$ over gjennomsnittet, mens andre halvparten er $-1s$ under. 27% havner $\pm 2s$ unna gjennomsnittet, og med et signifikansnivå på $\alpha = 0.05$ vil kun data til og med $2s$ bli tatt med i betraktning ($68+27 = 95\%$).

3.6 Studiens kvalitet

Det finnes bestemte krav for at noe skal kunne betraktes som kunnskap. *Validitet* og *reliabilitet* er sentrale begreper i kvalitetssikring av forskning. Jeg ønsker i dette kapitlet å diskutere validitet og reliabilitet både masteroppgaven generelt, samt testene som ble brukt til datainnsamling.

3.6.1 Validitet

Generelt handler validitet om hvor relevant eller godt data representerer fenomenet som forskes på. Validitet bør ses på som et spørsmål om *grad* i stedet for en absolutt tilstand (Christoffersen & Johannessen, 2012). I denne kvantitative studien er det hovedsakelig tre typer validitet som bør tas hensyn til: indre validitet, ytre validitet og statistisk validitet. I tillegg er det aktuelt å ta stilling til validiteten av tester som måleinstrument da studien benytter en pre – og posttest.

Indre validitet sier noe om med hvilken sikkerhet forskningsfunnene kan forklares gjennom den antatte forskningshypotesen, og ikke andre eksterne faktorer (Cohen, Manion & Morrison, 2018). Masteroppgavens indre validitet vil være et spørsmål om hvorvidt elevenes testresultater skyldes undervisningsopplegget og/eller labforsøket som en *helhetlig forskningsprosess*. Det ble ikke foretatt noe form for kjemiundervisning i tidsrommet mellom labforsøket og posttesten, så elevene fikk dermed ingen ny «input» fra skolens side. Hva elevene derimot gjorde på fritiden i form av egenstudier er åpenbart vanskelig å si noe om, og dette kan selvsagt ha vært en faktor som påvirket elevenes testskår. Samtidig kan det argumenteres for at den korte tidsperioden mellom labforsøket og posttesten i kombinasjon med forpliktelser i andre fag, ikke gav elevene mye tid til å lese på egenhånd.

I hvilken grad forskningsresultater kan generaliseres er et spørsmål om *ytre validitet*. (Cohen mfl., 2018). I denne masteroppgaven handler ytre validitet om hvilke slutninger som kan trekkes om fagkunnskapen til populasjonen, som følge av undervisningsopplegget. Generaliserbarhet må evalueres med hensyn utvalgets representativitet, og selekteringsprosessen av utvalget er avgjørende (Lund mfl., 2006). Som nevnt i kapittel 3.2 ble informantene plukket ut basert på tilgjengelighet, noe som kan svekke utvalgets representativitet. Selv om utvalget oppfylte det mest kritiske vilkåret om å studere *kjemi 2* ved en videregående skole, kan det likevel være kritiske forskjeller mellom ulike deler av landet, eller mellom urbane og landlige strøk.

Ytre validitet henger igjen sammen med *statistisk inferens* (se kapittel 3.5.2) som er ute etter å danne et *matematisk* grunnlag for å trekke slutninger fra utvalg til populasjon. Påliteligheten og nøyaktigheten av disse slutningene vedrører *statistisk validitet*. En slutning har en høy statistisk validitet dersom det kan bevises at funnene er forekommet systematisk og ikke tilfeldig (Creswell, 2014).

Validitet av en test går ut på testens målekyndighet. Testens innhold må være presist, korrekt og representativt i forhold til hva som skal undersøkes, og språkbruken må være utvetydig (Cohen mfl., 2018). Oppgavene på pretesten og posttesten ble utformet med inspirasjon fra tidligere eksamensoppgaver fra Utdanningsdirektoratet (UDIR), samt oppgaver fra elevenes egen lærebok. Da UDIRs oppgaveformuleringer er fagvurdert til å måle kompetansemål fra læreplanen styrkes testenes *innholds validitet*. Cohen mfl. (2018) bruker begrepet *concurrent validity* for å betegne i hvilken grad forskningens testresultater vil korrelere med andre

testresultat innen samme fagfelt. I likhet med studiens tester vurderes også eksamen og øvrige skoleprøver i hovedsak summativt, noe som til en viss grad ivaretar denne validitetstypen.

3.6.2 Reliabilitet

Cohen mfl. (2018) beskriver reliabilitet som et paraplybegrep for blant annet pålitelighet, konsistens og reproduserbarhet. Reliabilitet handler også om dataens målenøyaktighet, hvordan den samles inn på og hvordan den bearbeides (Christoffersen & Johannessen, 2012). Dersom en undersøkelse gjentas flere ganger over tid under like forhold hvor tilsvarende resultater oppnås alle gangene, kan forskningen sies å være reliabel. Reliabilitet påvirkes av tilfeldige målefeil, men kan styrkes ved å minimere eksterne faktorer, samt fjerne ekstreme verdier fra datasettet (Creswell, 2014). Jeg brukte videoforelesninger i klasseromsundervisningen for å forsterke studiens reproduserbarhet. Dersom studien skulle repeteres vil forskeren(e) kunne ta i bruk de samme videoforelesninger da de ligger tilgjengelig på Result UiTs egne nettsider i ubestemt tid. Oppgaveregningen knyttet til undervisningen vil også kunne gjentas da det ble benyttet samme læreverk for hele utvalget. Dette impliserer at undervisningsopplegget kan gjentas utover utvalget på klasser som skal introduseres for organisk kjemi. Reliabiliteten kan også kartlegges gjennom signifikans testing der et statistisk signifikant resultat tyder på at andre deler av populasjonen vil oppnå tilsvarende resultater.

Reliabilitet i forhold til pretesten og posttesten handler om at elevenes testskår er konsistent i den betydning at de ville oppnådd samme resultat dersom de utførte testen under andre omstendigheter (Cohen mfl., 2018). Elevenes fysikalske, sosiale og emosjonelle tilstand er faktorer som kan påvirke testresultatene, men som åpenbart ikke er mulig å kontrollere. Et annet spørsmål knyttet til testenes reliabilitet er hvordan oppgavene på testene både vektlegges og vurderes. Det faktum at testene ble utviklet i samarbeid med mine veiledere i tillegg til å bli pilotert av medstudenter, styrker reliabiliteten til datamaterialet. Alt av vurderingsarbeid ble utført av meg alene, og det er sannsynlig at et annet vedkommende ville ha vurdert testene annerledes. Av den grunn utarbeidet jeg retningslinjer for vurdering i form av *Vurderingskriterier* i hensikt om at alle testene skulle bli rettet så likt som mulig. Disse retningslinjene vil bære preg av mine personlige meninger angående hvilke kunnskaper som kvalifiseres som høy/middels/lav grad av måloppnåelse.

3.7 Forskningsetikk

Det råder en allmenn enighet om at forskning ikke skal gå på bekostning av grunnleggende menneskerettigheter og *Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora* har nedfelt etiske retningslinjer for å ivareta dette (NESH, 2016). Retningslinjer om ivaretagelse av personopplysninger, informasjonsplikt, samtykke og konfidensialitet er noen av de viktigste etiske overveielser i forbindelse med dette studiet. En vurdering fra Norsk senter for forskningsdata (NSD) bekrefter at prosjektet vil være i samsvar med personvernlovgivningen (se Vedlegg 2). Testene og labheftet krever ingen personlige opplysninger fra informantene utover navn og klasse, og svarene på testene er ikke av personlig art da det omhandler fagkunnskap. Kun faglærer og meg selv har hatt innsyn i elevenes besvarelser som for øvrig blir oppbevart med begrenset tilgjengelighet for faglærer. Ved bearbeidelse av datamaterialet ble navnene omgjort til koder i form av «T1», der bokstaven indikerer hvilken klasse eleven tilhører. All registrering av data og omkodingsdokumentet blir oppbevart på en ekstern harddisk der ingen andre har tilgang og vil bli slettet ved prosjektets slutt. Disse kodene vil dessuten ikke bli benyttet i selve mastergradsavhandlingen. I forkant av prosjektet ble det delt ut et informasjonsskriv/samtykkeskjema (se Vedlegg 1) til alle elevene i de aktuelle klassene. Informantene ble fullstendig informert om alle stegene i prosjektet og hva som var målet med forskningen da det ikke ble ansett for å påvirke deres prestasjoner på en faglig prøve. Alle informantene var over 18 år og derfor samtykkekompetent. Det ble informert om at deltakelse var frivillig og at det var mulig å trekke seg fra prosjektet til enhver tid uten noen konsekvenser.

4 Resultater og analyse

Statistiske analyser av pretesten, labheftet og posttesten presenteres i dette kapitlet. Først fremstilles deskriptiv statistikk i form av tabeller og søylediagram. Den deskriptive statistikken tar for seg testskårene for hver enkelt klasse, alle klassene samlet og skårddifferansen (δ) mellom testene. Det ble kjørt tester i SPSS for å sjekke om datamaterialet innfridde kravene til å kunne gjennomføre paret t-test, og disse resultatene vil også legges frem selv om de ikke har en direkte tilknytning til problemstillingen. Videre presenteres analyseresultat fra paret t-test og konfidensintervall utført på utvalgsdifferansen (\bar{x}_δ). Elevenes prestasjoner innenfor de ulike fagbegrepene oppsummeres i en poengtabell før elevevalueringen av arbeidsmetoder fremstilles grafisk til slutt.

4.1 Deskriptiv statistikk

Samtlige elever som deltok i dette masterprosjektet betraktes hovedsakelig som *ett* utvalg fra en større populasjon, selv om de tilhører ulike kjemiklasser. Det kan likevel være interessant å undersøke sentral -og spredningsmål på tvers av de fire klassene da dette kan gi et bilde av variasjon i populasjonen. Klassene refereres til som K1, K2, K3 og K4 og bestod av henholdsvis 18, 18, 26 og 16 elever.

4.1.1 Pretest og posttest

Gjennomsnitt (\bar{x}), median og standardavvik (s) av poengskår på pretest og posttest for de fire klassene, samt hele utvalget er presentert i Tabell 4. Gjennomsnittet er også oppgitt i prosent. For å skille mellom målinger fra pretest og posttest brukes henholdsvis tallene 1 og 2 i senket skrift. Eksempelvis vil notasjonen \bar{x}_1 betegne gjennomsnittet for *pretest*. Total skår på testene var 30 poeng (p).

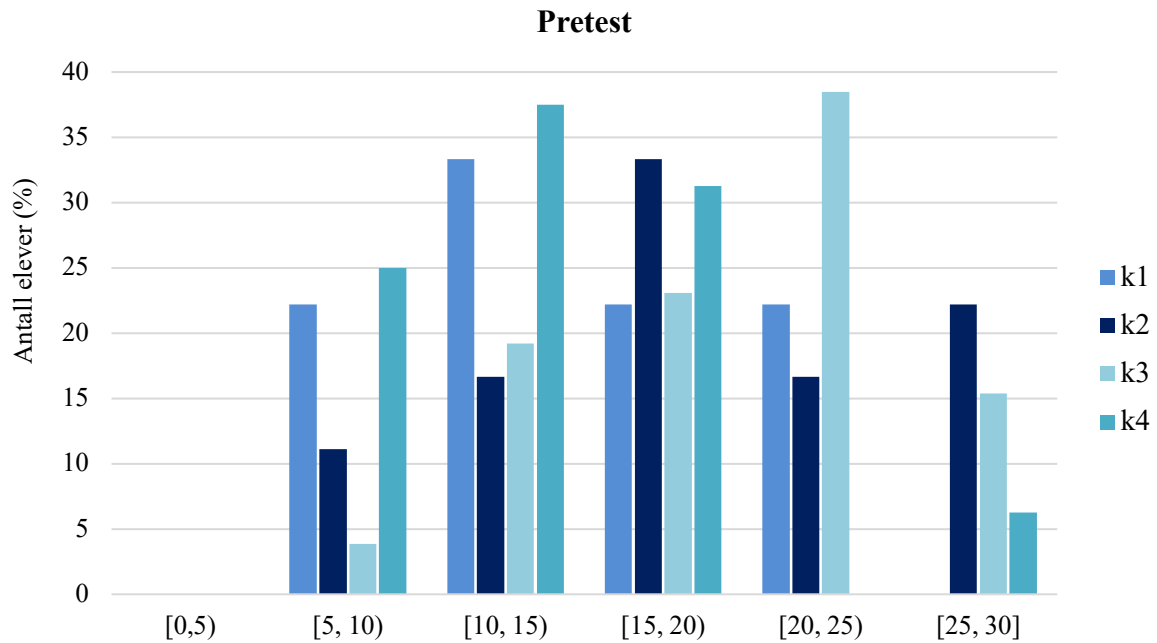
Tabell 4. Sentral - og spredningsmål av pretest og posttest

	K1		K2		K3		K4		Alle	
	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>
\bar{x}	13.6	14.3	18.4	16.3	19.1	20.4	13	23	16.4	18.5
	45%	48%	61%	54%	64%	68%	43%	77%	55%	62%
Median	13	14.5	18	18	20	19.5	14	26.5	16	18
s	6.1	6.4	6.1	7.4	5.8	6.1	5	7.4	6.3	7.3

Med unntak av K2, hadde hele utvalget og de resterende klassene en positiv endring fra \bar{x}_1 til \bar{x}_2 . Et fremtredende resultat var at K4 hadde lavest gjennomsnittsskår på pretesten, men høyest på posttesten med en skårifferanse (δ) på hele 10 p. Utvalgsdifferansen (\bar{x}_δ) var på 2 p (avrundet) og ettersom K2 hadde en nedgang på tilsvarende poeng er det trolig at K4s relativt høye δ hadde en innvirkning på \bar{x}_δ . I kontrast til resultatene fra K4, var det minimal forskjell mellom \bar{x}_1 og \bar{x}_2 for K3, som for øvrig var klassen med høyest gjennomsnittsskår på pretesten. De resterende klassene hadde en gjennomsnittlig forbedring på mellom 1-2 p. Generelt ligger medianen nær gjennomsnittet (± 1 p), noe som tyder på en tilnærmet normalfordeling av δ . K4 skiller seg igjen ut ved at medianen for posttesten er forholdsvis høyere enn gjennomsnittet. Det er et tilfelle som indikerer at datasettet inneholder *flere* lave verdier enn høye i forhold til gjennomsnittet. Fordelingen vil være venstreforskjøvet. K1 hadde tilnærmet lik spredning av skår på begge testene, mens de andre klassene og utvalget hadde *større* variasjon på posttesten enn på pretesten. Klassene og utvalget hadde en gjennomsnittlig spredning på mellom 5-7 p fra gjennomsnittet (avrundet), et resultat som ikke er overraskende da testene ble utført av ulike individer. Dersom prosentene konverteres til tallkarakterer 1-6 kan gjennomsnittsskårene for begge testene sies å ligge på karakteren 3-4 (40%-75%).

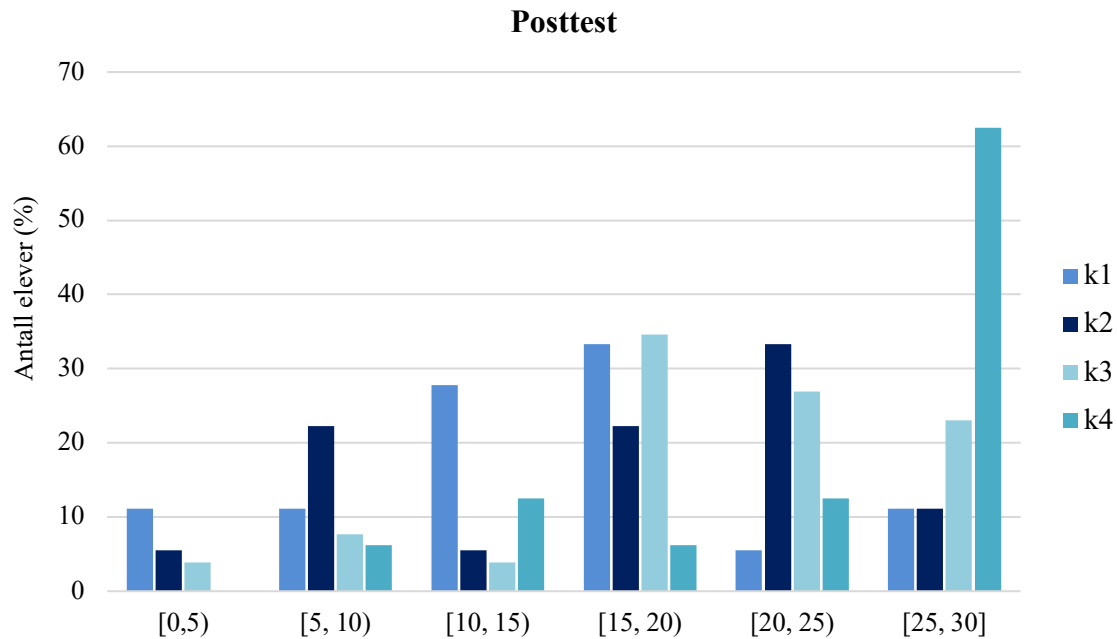
Søylediagram

Poengfordelingen av pretest og posttest for klassene visualiseres i separate søylediagram i henholdsvis Figur 8 og Figur 9. For å oppnå et markant bilde av fordelingene er grensene på y-aksen tilpasset etter høyeste frekvens. Eksempelvis går enhetene på pretestens y-akse opp til 40%, mens y-aksen for posttesten går til 70%. Av den grunn er også skalaen for enhetene ulik på y-aksen. Pretesten har en skala på 5%, mens posttesten har en skal på 10%. X-aksen representerer testskår i intervaller på 5 p. Eksakte verdier for hver av elevenes skår vil derfor ikke kunne leses direkte av diagrammene, men intervallene kan likevel gi en god indikasjon på klassefordelingen av skår.



Figur 8: Distribusjon av elevenes skår på pretesten

Figur 8 viser at ingen av klassene fikk under 4 p på pretesten og at samlet sett skårer flest elever mellom 15 og 19 p. Både gjennomsnittsskåren for utvalget (16.4) og medianen (16) ligger innenfor sistnevnte intervall. Flestparten av elevene fra K1 skåret mellom 10 til 14 p. Ingen elever fra K1 skåret i det høyeste intervallet, men ellers var klassens poeng fordelt likt utover intervallene. I K2 skåret de fleste elevene mellom 15 og 19 p, og den største andelen av elever i det høyeste intervallet kom fra denne klassen. Med en frekvens på 38 % skåret flestparten av elevene fra K3 mellom 20 og 24 p. Dette resultatet sett sammen med det faktum at kun 4% av K3 skåret i det laveste intervallet, forklarer klassens forholdsvis høye gjennomsnittsskår på 19.1 p. Skårene for K4 ligger fortrinnsvis mellom 5 og 19 p med en hovedvekt av skår på 10 til 15 p. 25 % av elevene fra K4 skåret i intervallet (5, 10], samtidig som kun 4% skåret i det høyeste. Sistnevnte resultat samsvarer med at K4 hadde den laveste gjennomsnittsskåren på 12.9 p.



Figur 9: Distribusjon av elevenes skår på posttesten

I likhet med pretesten skårer majoriteten av utvalget i intervallet [15, 20). Samtidig er det flere elever som skårer i de to øverste intervallene på posttesten (20-30 p), og i motsetning til pretesten er alle klassene representert i disse intervallene. På pretesten var det ingen som skåret under 4 p, men fra Figur 9 ser vi at noen få elever havner inn under dette poengsjiktet på posttesten.

K1 har flest antall elever som skårer mellom 15 og 19 p, ett intervallnivå høyere enn på pretesten. Det er også elever fra K1 som skårer i det høyeste intervallet, noe de gjorde på pretesten. Samtidig er de fleste elevene i det laveste intervallet fra K1. Også majoriteten av elevene i K2 skårer ett intervallnivå høyere på posttesten enn på pretesten og har dermed flest elever i intervallet [20, 25). De har derimot en nedgang på 11% i det høyeste intervallet fra pretest til posttest noe som merkes på klassens nedgang i gjennomsnitt. Flesteparten av elevene i K3 skårer mellom 15 og 19 p, hvorav den største andelen på pretesten skåret i intervallnivået over. Når det skal sies skårer færre elever fra K3 under 15 p på posttesten og 50 % av K3 er fordelt på de to høyeste intervallene. Hele 63 % av elevene fra K4 skårer mellom 25 og 30 p på posttesten, samtidig som ingen skårer i det laveste intervallet. Dette er en kontrast til deres resultat på pretesten der de i gjennomsnitt gjorde det dårligst, og resultatet på posttesten avviker til dels fra fordelingen til de tre andre klassene. Poengfordelingen til K4 er tydelig forskjøvet mot høyre og høye skårverdier noe som gjenspeiles i klassens høye gjennomsnittsskår på posttesten.

4.1.2 Pretest vs. Posttest

I seksjonene fremover, denne inkludert, betraktes ikke klassene hver for seg, men samlet som ett utvalg. Det er i hovedsak *skårdifferansen* δ (posttestskår – pretestskår) som er av interesse for problemstillingen. δ for utvalget angis i Tabell 5 og deles inn i positiv ($\delta > 0$), negativ ($\delta < 0$) eller ingen endring ($\delta = 0$).

Tabell 5. Antall elever med positiv, negativ eller ingen endring i skår

	Negativ endring	Positiv endring	Ingen endring
Antall elever	29	45	4
Prosentandel	37%	58%	5 %

Over halvparten (58%) av elevene oppnådde en høyere skår på posttesten, noe som korresponderer med at utvalgsgjennomsnittet for posttesten var høyere enn for pretesten. På den andre siden hadde 37 % en nedgang i skår. En problematikk knyttet til denne analysen er at elever med en høy poengskår på pretesten har en øvre begrensning til å kunne forbedre skåren sin sammenlignet med en elev med lavere poengskår. Eksempelvis vil elever med en pretestskår på 25 p maksimalt kunne forbedre skåren med 5 p, i kontrast til en elev som fikk 10 p på pretesten. Å forbedre skåren med et utgangspunkt på 25 p vil i tillegg kreve tilnærmet full uttelling på posttesten. Av den grunn presenteres skårendringen til elever som fikk $\geq 80\%$ (minst 24 p) riktig på pretesten i Tabell 6. Det var totalt 13 elever som fikk 24 p eller mer på pretesten.

Tabell 6. Skårdifferansen til elever med $\geq 80\%$ rett på pretesten

	Negativ endring	Positiv endring	Ingen endring
Antall elever	7	5	1
Prosentandel	54%	38%	8%

Det viste seg at blant elevene med ≥ 24 p på pretesten, hadde over halvparten (54%) en negativ endring til posttesten. Én elev holdt seg på samme poengsum. Det skal sies at de 7 elevene med negativ endring utgjorde knapt 9 % i forhold til hele utvalget.

Utvalgsdifferanse

Utvalgsdifferansen $\bar{x}_\delta = \bar{x}_2 - \bar{x}_1$ beregnes fra Tabell 4, og viser utvalgtes gjennomsnittlige skårifferanse (δ). Spredningen av δ analyseres ved standardavvik og variasjonsbredde, og er sammen med \bar{x}_δ fremstilt i Tabell 7. Variasjonsbredden tar for seg differansen mellom høyeste og laveste skår og viser den mest «ekstreme» verdien av δ .

Tabell 7. Gjennomsnitt, standardavvik og variasjonsbredde av utvalgets skårifferanse

\bar{x}_δ	s	Variasjonsbredde
2.1	7.1	-41

I gjennomsnitt økte elevene skåren sin med $\bar{x}_\delta = 2.1$ p fra pretest til posttest, noe som tilsvarer en forbedring på 6.7 %. Om denne differansen er statistisk signifikant testes ved statistisk inferens i kapittel 4.2.3. Det er viktig å huske at gjennomsnitt er et statistisk mål som påvirkes av «ekstreme» verdier, og et resultat om en forbedring i skår betyr ikke nødvendigvis at *flere* elever hadde $\delta > 0$ fremfor $\delta < 0$. Skjønt Tabell 5 viser at elevene faktisk hadde en forbedring.

Utvalget hadde et standardavvik på 7.1 p. Dersom vi antar en normalfordeling kan det tenkes at flest elever hadde en skårifferanse på mellom -5 og +9 p (2.1 ± 7.1). Et viktig poeng er at standardavvik aldri kan bli null eller mindre, men dette impliserer ikke at δ ikke er negativ. Variasjonsbredden tar ikke hensyn til gjennomsnitt, men tydeliggjør at datasettet inneholder en relativt «ekstrem» verdi $\delta = -41$. Dette kommer av en elev som fikk 2 p på pretesten og 23 p på posttesten.

4.1.3 Korrelasjon

Pearsons korrelasjonskoeffisient ρ gir indikasjon på om det foreligger lineær sammenheng mellom pre- og posttestskår. ρ kan være sensitiv for ekstreme verdier, så korrelasjonsanalysen ble utført uten outliers (se kapittel 4.2.1). Gjennom korrelasjonsanalysen beregnes også en p-verdi som indikerer analysens signifikansnivå. Dersom p-verdien $< \alpha = 0.05$ kan korrelasjonen mellom de to variablene sies å være signifikant. Det ble kjørt korrelasjonsanalyse mellom pretest og posttest, og labhefte og posttest hvor resultatene fremsettes i Tabell 8.

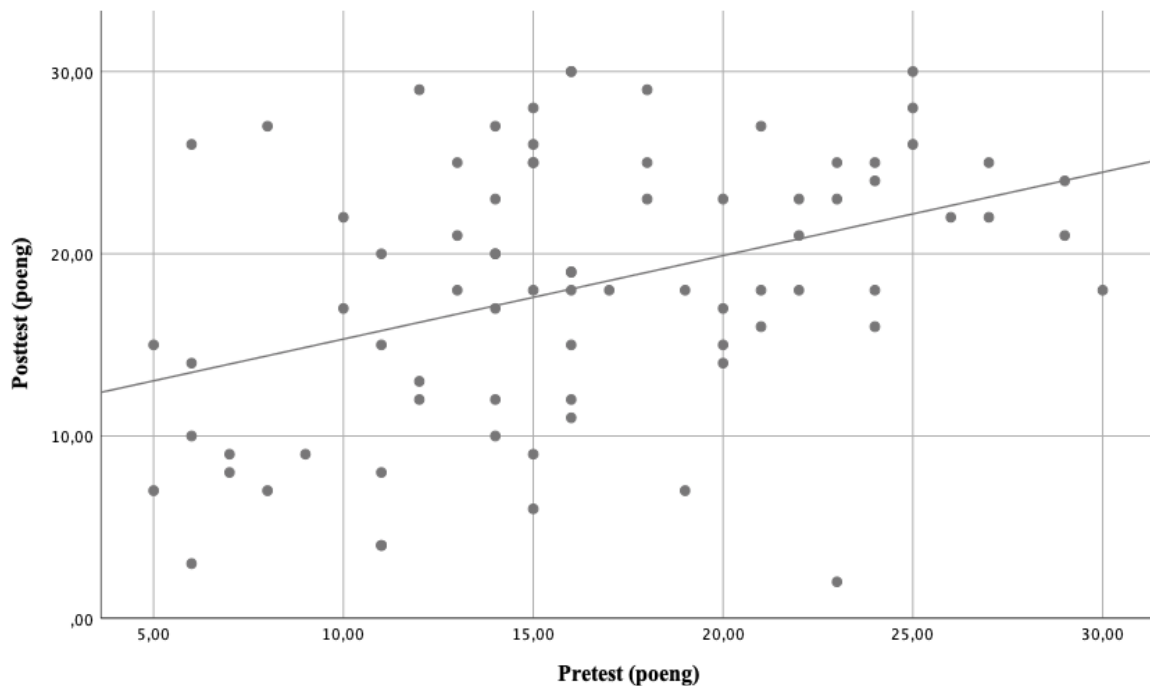
Tabell 8. Korrelasjonsanalyse mellom pretest-posttest, og labhefte-posttest

	n	ρ	Signifikans
Pretest – posttest	78	0.44	0.00354
Labhefte – posttest	58	0.46	0.000241

Begge korrelasjonene har et signifikansnivå $\ll \alpha = 0.05$, noe som betyr at det er tilstrekkelig med bevis for at variablene, altså poengskårene, har et lineært forhold der ρ er statistisk signifikant ulik null. $\rho > 0$ indikerer en positiv korrelasjon mellom både pretest-posttest og labhefte-posttest, og etter Martin og Bridgmon (2012) sin definisjon er styrken på korrelasjonene *moderat* (ρ mellom ± 0.40 og ± 0.69). Dette impliserer at vi til en viss grad kan forvente å finne *like* verdier av pretesten/labheftet og posttesten koblet sammen. Siden korrelasjonen er moderat må vi regne med at det er situasjoner med motsatt utfall også. Dersom vi hadde fått en perfekt positiv korrelasjon ($\rho = 1$) ville dette ha indikert at det verken var nedgang eller forbedring fra den ene testen til den andre. Ettersom utvalgsgjennomsnittet $\bar{x}_\delta > 0$ var det ikke forventet å observere $\rho = \pm 1$. Et moderat avvik fra en perfekt korrelasjon sett sammen med et positivt utvalgsgjennomsnitt, kan brukes som evidens for å styrke forskningshypotesen.

Spredningsplott

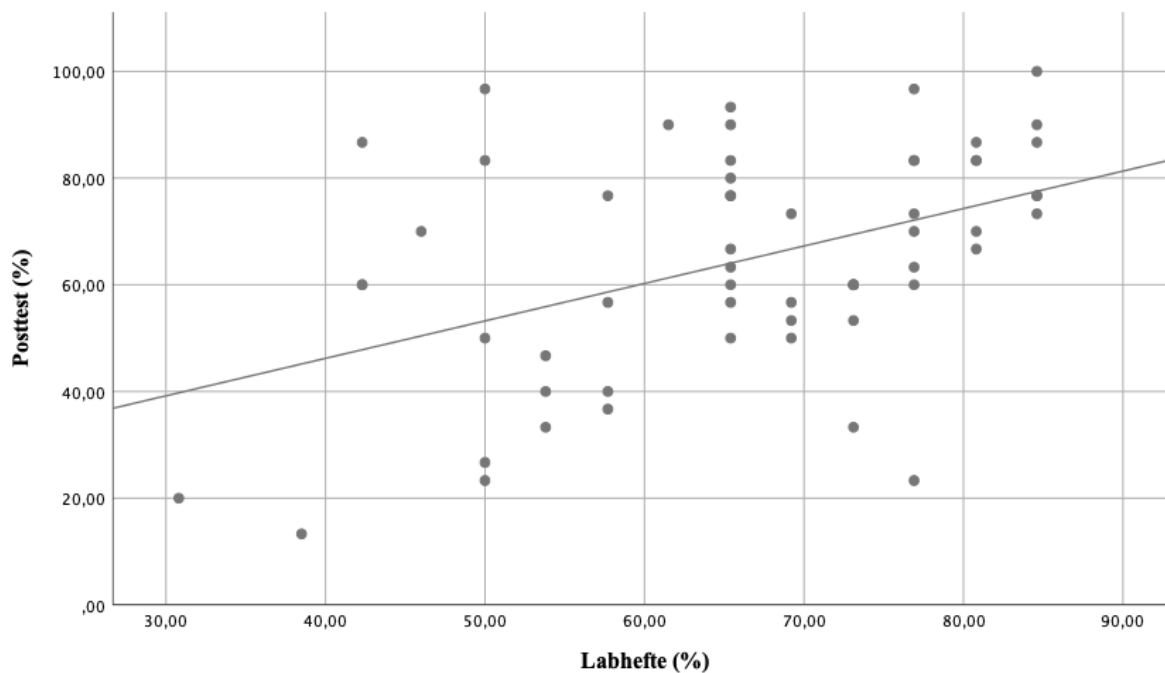
Korrelasjon mellom skårene kan visualiseres gjennom et spredningsplott der hver av skårene plottes mot hverandre i et koordinatsystem. Det kalkuleres en lineær regresjonslinje, og avstanden fra punktene til linja indikerer styrken av korrelasjonen. Jo nærmere punktene er linja, jo større styrke har ρ . Da pretesten og posttesten hadde like mange poeng og ble vurdert likt, anvender spredningsplottet testskårene direkte på aksene. Figur 10 viser posttestskår på x-aksen, plottet mot pretestskår på y-aksen.



Figur 10: Spredningsplott av pretest – og posttestskår

Figur 10 viser at punktene er forholdsvis spredt rundt linja, noe som tyder på at poengskårene på den ene testen ikke nødvendigvis korresponderer med skårene på den andre. Vi kan likevel tyde en svak helning oppover mot høyere. En slik helning indikerer positiv korrelasjon, noe som samsvarer med $\rho > 0$. En sammenligning av venstre øvre halvdel mot høyre nedre halvdel, viser at flere elever skårer lavt på pretesten og høyt på posttesten, enn motsatt. Dette er ikke overraskende ettersom $\bar{x}_\delta > 0$. Det virker også som at flesteparten av elevene som skårer rundt 50 % (ca. 15 p) riktig på pretesten, i større grad forbedrer skåren sin enn elever som skårer under 30 % (ca. 10 p). Spredningsplottet viser også at det var en relativt liten del av utvalget som skåret lavt (under 10 p) på *begge* testene. Avstanden fra punktene til regresjonslinja gir også et bilde av variasjonsbredden mellom pretest – og posttestskår, der større avstand er entydig med større variasjonsbredde.

Av de 78 elevene som utførte pre – og posttest var det 58 som leverte inn labheftet i etterkant. Da posttesten og labhefte hadde ulik poengfordeling ble poengskårene omgjort til prosent før korrelasjonsanalysen, og spredningsplottet mellom de to variablene er vist i Figur 11. X-aksen er tilpasset de faktiske skårene på labheftet, og da ingen elever havnet under 30% eller over 90 % skår på labheftet var det ikke hensiktsmessig å ta disse x-verdiene med.



Figur 11: Spredningsplott av labheftet og posttest

Korrelasjonskoeffisienten for labheftet-posttest ($\rho = 0.46$) hadde en ubetydelig økning på 0.02 sammenlignet med ρ for pretest-posttest ($\rho = 0.44$). Det er dessuten ingen hensikt i å sammenligne de to korrelasjonsanalysene da det var færre elever som leverte inn labheftet, i tillegg til at labheftet og posttesten ikke er relatert til hverandre på samme måte som pretesten og posttesten. Trenden i Figur 11 er den samme som for korrelasjonsanalysen mellom pretest-posttest der elevene forholdsvis holder seg på samme nivå. Det er selvsagt enkelttilfeller der elever har to ulike skåringsmønstre på labheftet og testen. Av elevene som leverte inn labheftet, skårer hovedvekten over 50% rett på labheftet med en sentrering rundt 60-80%. En sammenligning mot punktene i Figur 10 viser at flere av elevene som skåret lavt på posttesten var de samme elevene som ikke leverte inn labheftet til vurdering. Derfor er det vanskelig å si noe om skåringsnivået på labheftet sammenlignet med nivået på pretesten.

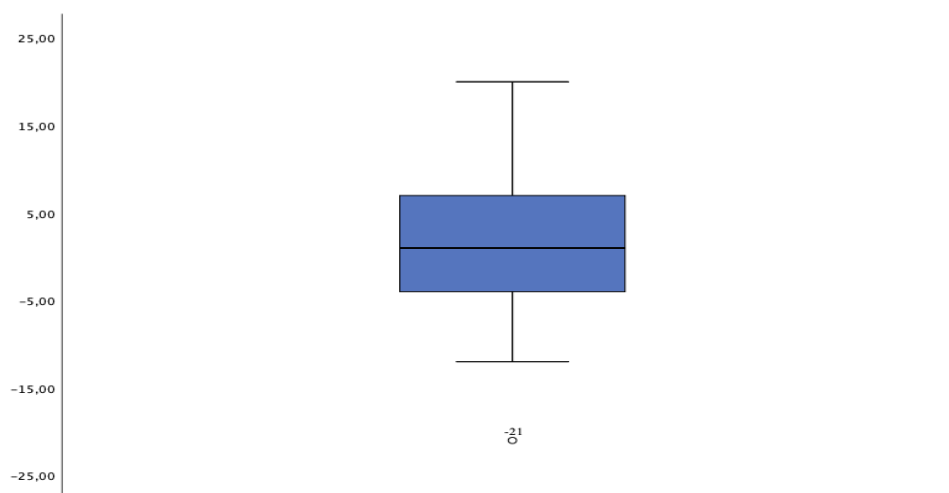
4.2 Statistisk inferens

For å kunne trekke slutninger angående *populasjonens* skårifferanse (μ_δ), ble det utført statistisk inferens av utvalgsdifferansen $\bar{x}_\delta=2.1$ som ble funnet i kapittel 4.1.2.

Konfidensintervall brukes til å estimere hvilke verdier μ_δ forventes å havne innenfor, og paret t-test sjekker om \bar{x}_δ er statistisk signifikant. Ettersom de statistiske prosedyrene forutsetter at datamaterialet er tilnærmet normalfordelt og ikke inneholder noen outliers (ekstreme verdier), ble disse antakelsene sjekket i SPSS, og de følgende resultatene presenteres som første del av kapitlet.

4.2.1 Outliers

For å bruke paret t-test må alle outliers fjernes fra datasettet da disse verdiene kan føre til et ukorrekt resultat. Utvalgets skårifferanser (δ) kartlegges i SPSS gjennom en visualisering av et boksploTT i Figur 12. Signifikante outliers vil bli tydelig markert ved et punkt utenfor selve boksploTTet.



Figur 12: BoksploTT av skårifferansen (δ) med outlier

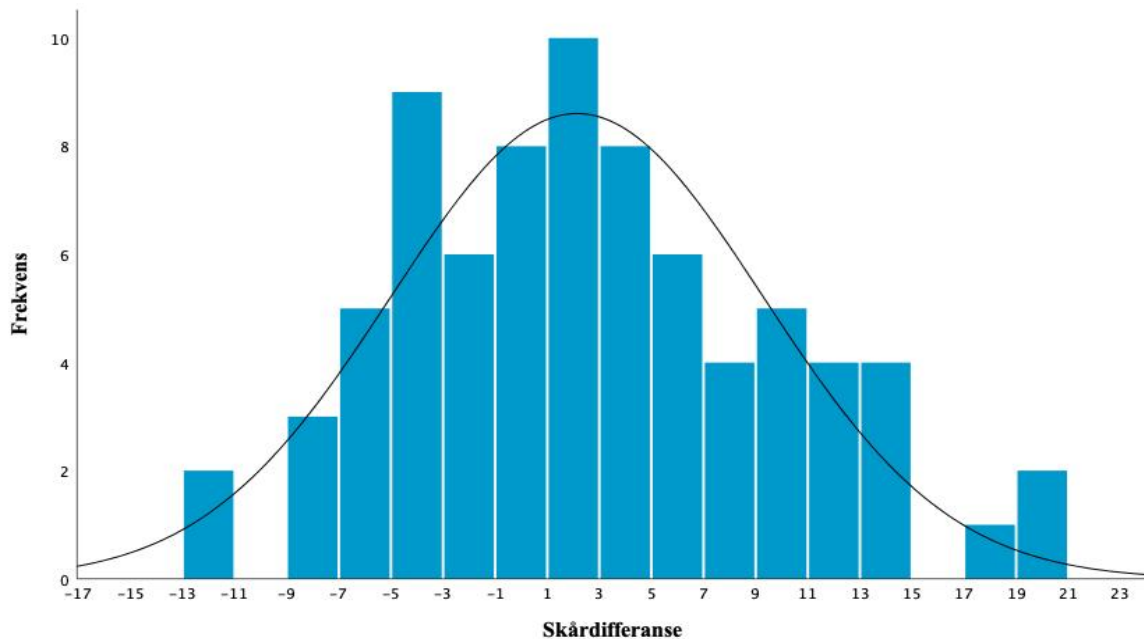
Den øverste og nederste viskeren i boksploTTet markerer henholdsvis 1. og 3. kvartil (Q) av datasettet, mens streken i midten representerer medianen (Q2). En verdi som avviker fra Q1 eller Q3 med mer enn 1.5 ganger differansen mellom Q1 og Q3, betegnes som en outlier. I dette tilfellet beregnes $\delta = -21$ til å være outlier og vil bli fjernet fra datasettet i forbindelse med videre analyser.

4.2.2 Normalfordelt data

Histogram

Skårddifferansene (δ) plottes mot frekvens i et histogram med en normalkurve i Figur 13.

Aksene øker med en verdi på 2. Ved en normalfordeling vil stolpene i histogrammet befinne seg under normalkurven.



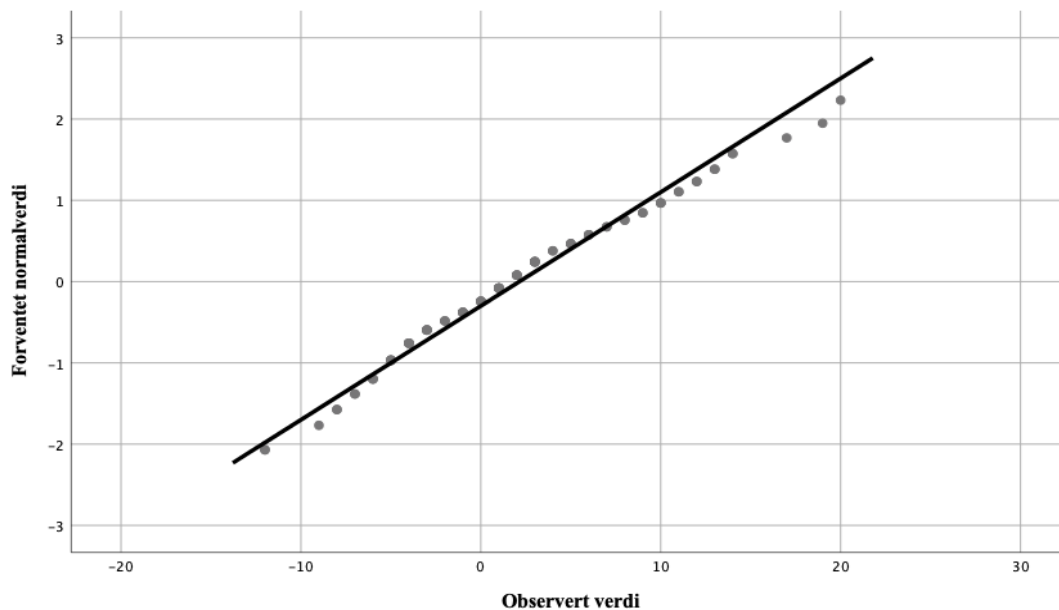
Figur 13: Histogram av skårddifferansen (δ)

Enkelte stolper i Figur 13 ligger utenfor normalkurven, og i ytterkantene observeres små avvik fra gjennomsnittet. Fordelingen har en svak tendens mot høyre og samsvarer med Tabell 5 som viser at flere elever hadde positiv, fremfor negativ skårendring. Ellers samler δ -verdiene seg hovedsakelig rundt utvalgsdifferansen på $\bar{x}_\delta = 2.1$. Fra y-aksen kan det leses at flest elever hadde en $\delta = 2$. Visuelt sett avviker histogrammet noe fra hvordan en normalfordeling ville ha sett ut. Derimot er ikke histogrammet alene et tilstrekkelig bevis for normalfordelt data.

Normalfordelt Q-Q plot

Et Q-Q plot sammenligner utvalgsfordelingen av δ med en normalfordeling ved å plote kvartiler mot hverandre. I Figur 14 plottes de observerte verdiene (δ) mot en forventet verdi i en normalfordeling (standardisert verdi), og det beregnes en lineær regresjonslinje. Dersom punktene former en omtrentlig rett linje med relativt små avvik fra regresjonslinja kan de to

fordelingene sies å være lik. Dette indikerer at datasettet er normalfordelt (Martin & Bridgmon, 2012).



Figur 14: Normalfordelt Q-Q plot av skår differansen (δ)

I høyre ytterkant svinger noen få punkter ut fra linja. Dette indikerer at datasettet inneholder «ekstreme» verdier som muligens ikke ville forekommet fra en normalfordeling. Likevel er avviket ikke særlig stort, og det peker på at utvalgsfordelingen ligner en normalfordeling ettersom de øvrige punktene danner en tilnærmet rett linje.

Skjevhet og kurtose

En mer håndfast analyse for normalitet er å se på verdiene av skjevhet og kurtose for δ , og disse presenteres i Tabell 9. Ratioen mellom skjevhet/kurtose og standardfeil (SE) gir en standardisert z-skår som sammenlignes med en kritisk z-verdi (z_k). Et signifikansnivå på $\alpha = 0.05$ gir en $z_k = 1.96$. Dersom den kalkulererte z-skåren ligger mellom ± 1.96 kan utvalgsfordelingen av δ sies å være normalfordelt.

Tabell 9. Skjevhet, kurtose, standardfeil og z-skår for δ

	Verdi	SE	Z-skår
Skjevhet	0.37	0.274	1.35
Kurtose	-0.30	0.54	-0.55

Skjevhet $0.37 > 0$ impliserer en svak, men høyreforskjøvet fordeling. Dette samsvarer med histogrammet i Figur 13 der de fleste δ -verdiene ligger til høyre for \bar{x}_δ . Kurtosen er svakt

negativ, noe som betyr at fordelingen er «flatere» enn en normalfordeling og ytterkantene er «rundere». I likhet med skjevheten ligger kurtosen relativt nær null og kan ikke sies å være et betydelig avvik fra en normalfordeling. Z-skåren for både skjevhet og kurtose ligger innenfor $-1.96 < z < 1.96$, og oppfyller dermed kriteriet for en tilnærmet normalfordeling. Med bakgrunn i sentralgrenseteoremet, denne analysen og de øvrige funnene i dette kapitlet er det rimelig å anta at **utvalgsfordelingen er tilnærmet normalfordelt.**

4.2.3 Paret t-test og konfidensintervall

Med bakgrunn i antakelsene om kontinuerlige variabler, ingen outliers og normalfordelt data gjennomføres en paret t-test for å undersøke om utvalgsdifferansen $\bar{x}_\delta = 2.1$ er statistikk signifikant. Det ble også generert et 95%-konfidensintervall for å få en indikasjon på forventet μ_δ -verdi, og resultatene fremlegges i Tabell 10. Nullhypotesen H_0 som antar ingen forskjell i skår mellom pretest og posttest ($\mu_2 - \mu_1 = 0$) forkastes på α -nivå dersom t-verdien er >1.99 eller < -1.99 , og ved en p-verdi < 0.05 .

Tabell 10. Resultater fra paret t-test og konfidensintervall

t-verdi	p-verdi	95% KI	Standardavvik
2.63	0.01	[0.52, 3.76]	6.4

Analyse av paret t-test gir grunnlag til å **forkaste H_0** på et $\alpha=0.05$ -nivå da t-verdi = 2.63 > 1.99 og p-verdi = 0.01 < 0.05 . I tillegg ligger verdien av H_0 (null) utenfor konfidensintervallet, noe som styrker grunnlaget til å forkaste H_0 og heller beholde den alternative hypotesen H_1 om *endring* i skår. Dette bekrefter at utvalgsdifferansen $\bar{x}_\delta = 2.1$ er **statistisk signifikant**, med forbehold om en 5% sjanse for å feilaktig forkaste nullhypotesen. Funnene fra korrelasjonsanalysene om at $\rho \neq \pm 1$ styrker også dette statistiske funnet.

Fra konfidensintervallet kan vi med en 95 % sikkerhet si at populasjonens gjennomsnittlige skår-differanse, populasjonsdifferansen μ_δ , befinner seg mellom 0.52 p og 3.76 p.

Ettersom poengskårene oppgis i naturlige tall avrundes intervallverdiene oppover. I praksis betyr dette at elever i *kjemi 2* kan forventes å oppnå **mellom 1 og 4** poeng mer på posttesten enn på pretesten. Ettersom datamaterialet er tilnærmet normalfordelt betyr et standardavvik på 6.4 p at 68% av populasjonen, altså 3624 elever, vil ha en δ mellom [-4.3, 8.5], og 27 % vil oppnå en $\delta = [-10.7, 14.9]$.

4.3 Temabasert poengskår

Elevenes skåringsmønster på hver oppgavene gir innsikt i deres kunnskap om de ulike fagbegrepene. Gjennom intervaller med 25% bredde, gir Tabell 11 en oversikt over elevenes skår på alle oppgavene på pretesten og posttesten. I poengsjiktet der flest elever skåret på hver oppgave, utheves prosentandelen med fet skrift. Som tidligere påpekt var oppgave 1 og 6 flervalgsspørsmål på pretesten, og ble vurdert som enten rett (100%) eller galt (0%).

Tabell 11. Skårfordeling på hver av oppgavene i pretesten og posttesten

Oppg.	Total poeng	Oppgavens tema	Pretestskår (%)				Posttestskår (%)			
			0-24	25-49	50-74	75-100	0-24	25-49	50-74	75-100
1	2	<u>Reaksjonsmekanismer:</u> Elektronoverføring i en reaksjon	17			83	38		37	25
2	3	<u>Reaksjonsmekanismer:</u> Forutsi produkt basert på reaksjonstype	9	33		55	6	7	8	79
3a	5	<u>Reaksjonsmekanisme:</u> addisjonsreaksjon	23	28	27	22	45	20	12	23
3b	5	<u>Reaksjonsmekanisme:</u> eliminajonsreaksjon	42	17	29	12	20	20	24	36
4	6	<u>Tolkning av ¹HNMR-spekter:</u> areal, kjemisk skift, hydrogenmiljø, n+1-regel	27	4	23	46	29	7	13	51
5	6	<u>Tolking av MS-spekter:</u> tegne struktur, isotoper, fragmentering	52	2	3	43	25	4	3	68
6	2	<u>OmkrySTALLISERING:</u> Prinsipper og gjennomføring	19			81	19		28	53

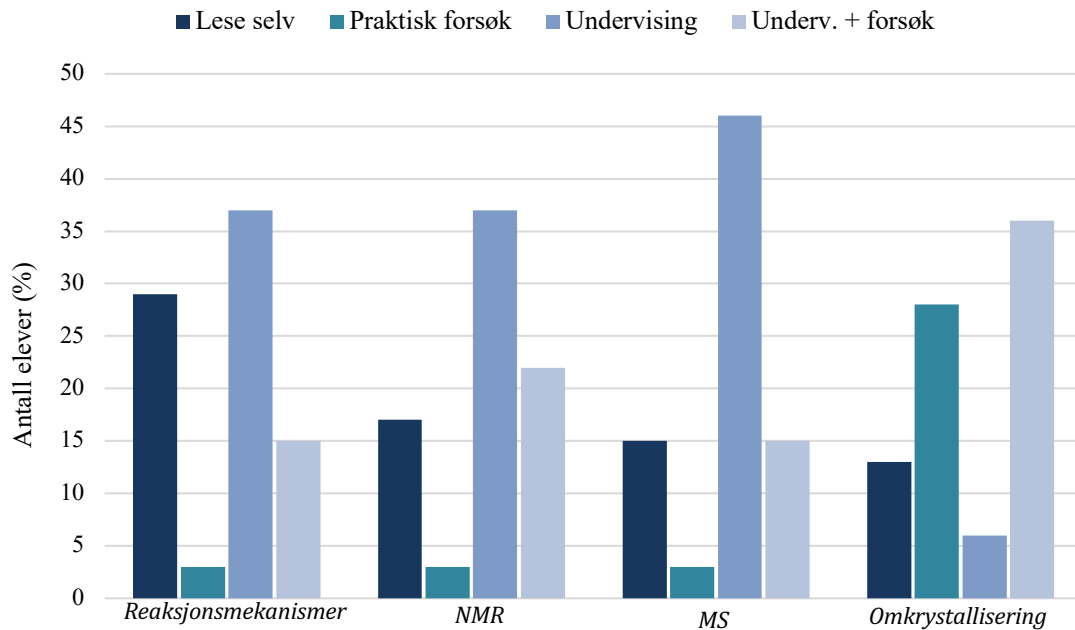
På spørsmålet om elektronpilens betydning i oppgave 1 krysset 83% av riktig svaralternativ. For samme oppgave på posttesten der elevene selv måtte tegne inn elektronpiler var det kun 25 % som mestret dette. I oppgave 2 hvor elevene skulle avgjøre reaksjonstype, er trenden den samme for begge testene: Majoriteten av elevene skårer i det høyeste poengintervallet og svært få svarer feil. På posttesten er det denne oppgaven elevene skårer best på. Samlet sett gjorde elevene det *dårligst* på oppgave 3a og 3b som omhandler reaksjonsmekanismer. På posttesten skårer de fleste elevene lavest på oppgave 3a om addisjonsreaksjon. Eliminajonsreaksjonen i 3b hadde en beskjeden økning i skår fra pre - til posttest.

For både 3a og 3b var det en ganske stor spredning i poengskår der minst 20 % av elevene skåret innenfor hvert intervall. Hva gjelder oppgave 4 som omhandler tolkning av NMR-spekter er det ikke stor forskjell mellom pretest og posttest i elevenes skåringsmønster. Det er en liten økning på 5% i det høyeste intervallet, men ellers er poengskårene relativt likt fordelt. Fra pretest til posttest er den en ganske tydelig forbedring i skår på oppgave 5 der elevene skulle tolke massespekter. Fra å være flest elever i det laveste poengsjiktet på pretesten, gikk det til å være flest elever i det øverste på posttesten. For begge testene er det få elever som skårer i de midterste poengsjiktet på oppgave 5, og elevene skårer enten over 75% eller under 25%. Hva gjelder oppgave 6 som omhandlet omkrystallisering, skåret majoriteten i det øverste poengsjiktet på begge testene. Fra flervalgsoppgaven på pretesten til det åpne spørsmålet på posttesten var det en nedgang på 28 % av elever som skåret i det høyeste intervallet. Ettersom oppgavene hadde ulik struktur der den ene var et flervalgsspørsmål, er det vanskelig å gjøre en real sammenligning.

For å oppsummere tabellen ser det ut som at de elevene samlet gjorde det best på oppgave 2 der de skulle bestemme reaksjonstype ut fra produkt. De gjorde det dårligst på oppgavene om reaksjonsmekanismer, hvor spredningen av poengskår også var størst. Den største forbedringen fra pretest og til posttest var i tilknytning til tolkning av MS i oppgave 5. Oppgave 4 derimot, tolkning av NMR-spekter, hadde liten endring i skår. Ellers gjorde elevene det bedre på begge flervalgsspørsmålene på pretesten sammenlignet med posttesten som hadde åpne spørsmål.

4.4 Elevevaluering av undervisningsmetoder

Etter endt opplegg ble elevene bedt om å krysse av for hvilken arbeidsmetode/undervisningsmetode de likte best for å lære om de ulike fagbegrepene. Elevene hadde også mulighet til å komme med egne forslag, men da det var svært få som benyttet seg av dette ble besluttet å ekskludere svarene fra stolpediagrammet i Figur 15. Under alternativet «praktisk forsøk» inngår selve syntesen av DBA, jobbing med labheftet og etterarbeid med NMR - og MS spektre. Med «undervisning» menes videoforelesningene, oppgavejobbingen og annen aktivitet som ble gjort i klasserommet. Figur 15 viser prosentandelen av elever som foretrakk hver av de ulike arbeidsmetodene.



Figur 15: Elevevaluering av undervisningsmetoder/arbeidsmetoder

De fleste elevene følte de lærte reaksjonsmekanismer best gjennom klasseromsundervisningen eller å lese på egenhånd. For både NMR og MS var *undervisning* i klasserommet tydelig foretrukket fremfor de andre arbeidsmetodene. 22 % av elevene evaluerte *undervisning + forsøk* som beste arbeidsmetode for å lære om NMR. For MS var det like mange (15%) som syntes *undervisnings + forsøk*, og *lese selv* var like nyttig. Omkrystallisering var det eneste emnet der elevene favoriserte praktisk forsøk som en del av undervisningen fremfor de andre metodene, og den mest prefererte arbeidsmetoden var *undervisning og forsøk kombinert*.

5 Diskusjon

Naturvitenskapelig *fagkunnskap* er en forutsetning for naturvitenskapelig allmenndannelse, som igjen er et overordnet mål med naturfagsundervisning i skolen. I dette kapitlet vil jeg drøfte ulike aspekt ved praktisk arbeid som undervisningsmetode for å formidle *fagkunnskap* til elever. Suksessen av en bestemt praktisk aktivitet må drøftes ut fra dens hensikt, og for å evaluere studiens labforsøk benyttes effektivitetsmodellen til Millar mfl. (1999). Dette henger igjen sammen med labforsøkets oppbygning som en *helhetlig forskningsprosess*, og slagkraften av denne typen forsøk for å lære fagkunnskap vil også bli diskutert.

5.1 Fra utvalg til populasjon

Gjennom signifikanstesting ble det bekreftet at elevene hadde en *statistisk signifikant forbedring* i poengskår etter labforsøket. Fra pretesten til posttesten var det en gjennomsnittlig økning på 2 poeng. Beregningen av et 95%- konfidensintervall indikerer at populasjonen i gjennomsnitt vil ha en forbedring på mellom 1 og 4 poeng. Da signifikansnivået var satt til $\alpha = 0.05$ innebærer dette at vi med 95 % sikkerhet kan si at dette resultatet ikke oppsto ved tilfeldigheter eller samplingsfeil, noe som helt klart styrker den statistiske validiteten. Det er viktig å understreke at den statistiske hypotesen som ble undersøkt i signifikanstesten ikke må forveksles med studiens forskningshypotese (Lund mfl., 2006). Med andre ord bekrefter ikke en forkasting av nullhypotesen, masteroppgavens forskningshypotese. Det er heller et bevis som styrker forskningshypotesens antakelse om at elevenes fagkunnskaper ble styrket som følge av undervisningsopplegget og labforsøket.

Én av de tre klassene i utvalget, K2, hadde en nedgang i skår fra pretest til posttest. K1 og K3 økte gjennomsnittskåren med mellom 1 og 2 poeng, mens K4 hadde en skårifferanse på 10 poeng. Skårifferansen på 10 poeng ligger utenfor konfidensintervallet (se Tabell 10), og er *over* ett standardavvik (6.4) fra gjennomsnittet. I en normalfordeling kan vi forvente at kun 13% av elevene oppnår et slikt resultat og det kan sies å være et noe «ekstremt» tilfelle. Samtidig må det tas høyde for at K4 i utgangspunktet hadde en lav skår på pretesten og dermed et større forbedringspotensial sammenlignet med de andre klassene. Da skoleklasser gjerne har en øvre grense på rundt 30 elever vil klassegjennomsnittet være lett påvirkelig fra veldig høye eller veldig lave skårifferanser. Det er ikke usannsynlig at andre skoleklasser i populasjonen også vil ha en gjennomsnittlig økning/nedgang i skår som ligger utenfor konfidensintervallets verdiområde, altså i ytterkantene av normalfordelingen.

Til tross for at skårddifferansen ble beregnet til å være statistisk signifikant ønsker jeg ikke å trekke noen slutninger angående generaliserbarheten av studiens funn. Som det ble diskutert i validitetskapitlet 3.6.1, er det usikkerhet knyttet til utvalgets representativitet, og i tillegg anses utvalgsstørrelsen på 78 elever som relativt liten for et slikt formål.

5.2 Styrket fagkunnskap som følge av labforsøk

Undervisningsoppleggets kjerne var selve labforsøket. Hvorvidt forsøket bidro til elevenes læring av fagbegrepene er et spørsmål om studiens indre validitet, og eksterne faktorer som kan ha påvirket utfallet må tas i betraktning. Som tidligere nevnt i validitetskapitlet foregikk det ingen undervisning i perioden mellom forsøket og posttesten, men elevenes eventuelle selvstudier er et usikkert element. Fra elevenes evaluering av arbeidsmetoder i Figur 15, kommer det fram at det var elever som leste på egenhånd i alle temaene, og det er da rimelig å anta at dette hadde en innvirkning på deres prestasjoner på posttesten. Elevevalueringen viste også at samlet sett var *praktisk forsøk* den arbeidsmetoden elevene foretrakk *minst* for å lære fagkunnskap.

I følge Hofstein og Lunetta (2004) foreligger det bevis som at elevenes prestasjoner på skriftlige prøver *i liten grad* korrelerer med deltakelse i praktiske labforsøk. TIMSS og PISA undersøkelsene i norsk skole viser det samme (P. M. Kind, 1999; Kjærnsli mfl., 2007). På bakgrunn av usikkerheten knyttet til studiens indre validitet, mangelfulle observasjoner og det faktum at svært få elever foretrakk praktisk forsøk som undervisningsmetode er det vanskelig å påvise noen sammenheng mellom labforsøket og testresultatene om forbedring i skår.

5.3 Styrket fagkunnskap som følge av undervisningsopplegg

Elevenes læring fra praktiske aktiviteter avhenger av deres teoretiske kunnskaper forut aktiviteten (Abrahams & Millar, 2008; Berry mfl., 1999; Séré, 2002). Studiens undervisningsopplegg, illustrert i Figur 7, innebar som første steg videoforelesninger og oppgaveregning i klasserommet før labforsøket. Dette ble gjort for å danne et teoretisk grunnlag om komponentene i aktiviteten, slik at elevene hadde mulighet til å knytte labforsøket tilbake til denne teorien. Der er imidlertid ikke gitt at elevene, til tross for undervisning, har lært fagstoffet godt nok til å ta det med seg inn i aktiviteten og bruke det til å reflektere over sine handlinger (Abrahams & Millar, 2008). Elevene evaluerte klasseromsundervisning til å være den beste metoden for å lære alle fagbegrepene, med

unntak av omkrystallisering der elevene foretrakk *undervisning + forsøk*. Dette alternativet var samlet sett favorisert som nummer to.

Korrelasjonsanalysen av labhefte-posttest og pretest-posttest gav resultater om en moderat, positiv lineær sammenheng mellom skårene. Dette tyder på at elever til en viss grad holder seg på samme faglige nivå i form av poengskår på pretesten/labheftet og posttesten. Samtidig var det i hovedsak elever med middels skår på pretesten som hadde den største forbedringen. Ut fra dette kan det løst spekuleres i om undervisningsopplegget/labforsøket var best egnet for elever med middels måloppnåelse. En perfekt positiv korrelasjon ($\rho=1$) ville ha indikert at ingen av elevene hadde endring i skår da skårverdiene ville ha økt/minsket i tandem. Avviket fra en lineær sammenheng bekrefter derimot at elevene har hatt endring i skår, og dette resultatet gir enda et påskudd til å akseptere forskningshypotesen. Når det skal sies er korrelasjon et statistisk mål som kun sammenligner tallverdier, og sier ikke noe om årsakssammenhenger mellom to fenomener.

Gjennom undervisningen, både når det gjelder tradisjonell tavleundervisning og praktisk arbeid, bør læreren bygge assosiasjoner til de teoretiske begrepene som skal læres (Ringnes & Hannisdal, 2014). Under labforsøket fikk elevene mulighet til å observere fargeforandringer i blandingen av aceton og benzaldehyd. Om de knyttet fargeforandringen i løsningen til reaksjonen mellom reaktantene (reaksjonsmekanismer) er vanskelig å si basert på studiens datamateriale. En slik kobling mellom observasjon og teori (se Figur 3) kan være mer kognitivt utfordrende for elevene enn læreren forventer (Abrahams & Millar, 2008; Millar, 2004). Derfor er det nødvendig at læreren *hjelper* elevene å lage denne koblingen. I et sosiokulturelt læringssyn innebærer dette at læreren veileder elevene gjennom stillasbygging av deres kunnskap. Det foreligger bevis for at elevens forståelse av naturvitenskap målt gjennom karakterer, styrkes ved interaksjon mellom lærer og elev (Cooper & Stowe, 2018).. Via labheftets refleksjonsspørsmål ønsket jeg å skape diskusjon både blant elevene seg imellom, men også mellom elevene og meg selv. Da en stor andel av elevene foretrakk *undervisning + forsøk* som læringsmetode peker det mot at undervisningsopplegget i større grad enn kun labforsøket hadde en innvirkning på elevenes testresultater.

5.4 Elevenes læring av fagbegrepene

I denne seksjonen diskuteres elevenes læring av de fem sentrale fagbegrepene: addisjonsreaksjon, eliminasjonsreaksjon, NMR, MS og omkrystallisering. Ettersom begrepene underbygges av ulik teori som i og for seg ikke er knyttet sammen, vil læringsutbyttet av undervisningsopplegget/labforsøket variere på tvers av begrepene. Resultatene i Tabell 4 viser at elevene i gjennomsnitt hadde en skår på henholdsvis 55% og 62% på pretesten og posttesten. Dette bedømmes til en *middels måloppnåelse* (karakteren 3-4). Snittkarakteren for *kjemi 2*-eksamen i 2018 var 3.8 (Utdanningsdirektoratet, 2018). Selv om testene i denne masteroppgaven ikke kan sammenlignes direkte med eksamen da omfanget er svært ulikt, kan det gi en indikasjon på at testene holdt seg på et *akseptabelt* faglig nivå. Samtidig skal ikke en «kapittel-test» ha samme vanskelighetsgrad som eksamen, og ettersom mine testresultater var lik eksamensresultatene, peker det mot at studiens tester var for krevende ut fra elevenes forutsetninger (undervisningsopplegget).

5.4.1 Reaksjonstyper og reaksjonsmekanismer

På både pretesten og posttesten omhandlet oppgave **1**, **2**, **3a** og **3b** reaksjonstyper og reaksjonsmekanismer. Tabell 13 viser at elevene i mye større grad behersker karakterisering av reaksjonstyper fremfor å skulle illustrere selve reaksjonsmekanismen. På oppgave **2** ble det kun etterspurt reaksjonstype og ferdig produkt, og flesteparten av elevene fikk full skår på disse oppgavene. Tabellen viser også en klar trend der oppgave **3a** og **3b** som omhandler reaksjonsmekanismer er de oppgavene elevene samlet gjør det dårligst på. Oppgave **1** på posttesten krevde at elevene tegnet inn elektronpiler, noe de åpenbart slet med. Dette kan igjen forklare deres utfordringer med oppgave 3 da forståelse for elektronpiler er en forutsetning for å mestre reaksjonsmekanismer. Elevene skåret litt høyere i oppgave **3a** (addisjonsreaksjon) på pretesten enn på posttesten. Mekanismen for addisjonsreaksjonen på pretesten var ganske lik eksemplene fra læreboken, mens posttesten tok for seg addisjonsreaksjonen mellom aceton og benzaldehyd fra labforsøket. Det var hele 45 % som ikke fikk til posttestoppgaven i det hele tatt, og det kan tenkes at elevene ble «overveldet» av å måtte jobbe med mer komplekse molekylstrukturer sammenlignet lærebokens forholdsvis enkle alifatiske hydrokarboner. Det skal nevnes at aceton-benzaldehyd reaksjonen heller ikke var ukjent for elevene da den ble presentert gjennom videoforelesingene, i tillegg til at samme oppgave ble gitt gjennom et refleksjonsspørsmål i labheftet.

Hva gjelder oppgave **3b** var resultatene motsatt sammenlignet med **3a**. Her gjorde elevene det best på posttesten. Denne oppgaven var en eliminasjonsreaksjon der reaksjonen skjedde i to steg (E1, uni-molekylær eliminasjon) i likhet med alle eksemplene i læreboken, mens pretesten opererte med en bi-molekylær eliminasjon (E2) som kun ble vist gjennom videoforelesningene. Fra disse resultatene kan det virke som at elevene har hentet kunnskap fra læreboken fremfor videoforelesningene da de gjorde det bedre på oppgavene som gjenspeilet lærebokens eksempler. Samtidig kommer det frem av elevenevalueringen i Figur 15 at den største andelen av elevene (37%) foretrakk *undervisning* for å lære om reaksjonsmekanismer. Under klasseromsundervisningen ble det foretatt oppgaveregning fra læreboken og det kan tenkes at det er dette elevene først og fremst sikter til fremfor videoforelesningene. Reaksjonsmekanismer er et relativt teoritungt emne som i starten kan være lettere å få forklart fremfor å skulle lese teorien selv. Cooper og Stowe (2018) hevder at dette temaet trolig er blant de vanskeligste innen organisk kjemi, ikke bare for elever på videregående nivå, men også i høyere utdanning.

Gjennom undervisningsopplegget ble elevene eksponert for kjemiske reaksjoner på både makro, mikro og representasjonsnivå. De utgjør «*Tre dimensjoner ved kjemifaget*» som elevene må mestre for å oppnå læring (Ringnes & Hannisdal, 2014). Barke mfl. (2009) rapporterer at elever sliter med overgangen mellom de tre nivåene, og da spesielt å konvertere fra makronivå til representasjonsnivå. Under labforsøket opplevde jeg at det var elever som ikke så sammenhengen mellom reaksjonslikningen i labheftet (representasjonsnivå) og stoffene de blandet sammen (makronivå) for å få produkt. Dette kom frem av elevenes henvendelser angående refleksjonsspørsmålet om addisjon av aceton og benzaldehyd. De skjønnte ikke hvorfor dette var et relevant spørsmål, og mislyktes i å bruke den oppgitte reaksjonslikningen for å løse problemet. Videre sier Barke mfl. (2009) at problematikken knyttet til overgangen mellom nivåene også skaper misoppfatninger blant elevene. Et eksempel på dette ble tydelig gjennom oppgave **1** på posttesten der flere elever tegnet en pil som indikerte at hydrogenatomet på ett av acetonets metylgrupper donerte elektroner til oksygenatomet i hydroksylgruppen for å lage en binding. Dette er ikke mulig da hydrogenatomet allerede har sitt ene elektron i en binding til karbonatomet. De feilaktige svarene kan skyldes manglende forståelse for elektronpilers betydning, hvordan de brukes eller hydrogenatomets elektronkonfigurasjon. Cooper og Grove i Cooper og Stowe (2018) påpeker at elever, og forståvidt studenter på universitetsnivå, har vansker med å tegne enkle Lewis strukturer og se sammenhengen mellom molekylstruktur og egenskaper. Ettersom

hydrogenatomet representeres gjennom symbolet «H» i oppgaveteksten er det naturlig at elevene assosierer «H» med ordet «hydrogen». Dette kan føre til at de mislykkes i å tenke på «H»-en som et atom på mikronivå, og derav overser dets struktur. Dersom elevens kunnskap om grunnleggende ideer fra generell kjemi er fraværende, vil de slite med å lære seg organisk kjemi og vil til en viss grad tvinges til å memorere fremfor å forstå interaksjoner mellom molekyler (Barke mfl., 2009).

5.4.2 NMR

Oppgave 4 dreide seg om tolkning av NMR-spekter. Oppgavene krevde at elevene måtte sammenkoble ulik informasjon, presentert i kapittel 2.6, for å løse oppgaven. Oppgave 4 ble vurdert ut fra elevenes evne til å bruke informasjon om kjemisk skift, hydrogenmiljø, areal og $n+1$ regelen (finstruktur). På både pretesten og posttesten mestret over halvparten av alle elevene disse oppgavene. Den mest gjennomgående feilen på begge testene var at elevene begrunnet antall signaler på spektret med antall karbonatomer i karbonkjeden fremfor å se på ulike typer protoner. På pretesten var det like mange signaler som karbonatomer, mens på posttesten var det ett karbonatom som ikke var bundet til noen hydrogenatomer. Likevel tilskrev enkelte elever et signal for dette karbonatomet, noe som gjenspeiler deres usikkerhet om hva som faktisk gir oppgav til signalene (hydrogenkjerner). Dette kan være en følge av at elevene ikke evnet å se at karbonatomet som et atom med et fullt valensskall (alle de ytterste elektronene er i binding). Igjen kan problematikken knyttet til å veksle mellom mikronivå og representasjonsnivå være en årsak til situasjonen beskrevet over. V. Kind (2004) poengterer at måten kjemiske formler skrives på kan være med på å skape en misoppfatning om at formelens første element er det «viktigste». For organiske forbindelser begynner formelen som regel med en «C» for karbonatomet, og dette kan ha forvirret elevene til å utelukkende se og tenke på karbonatomene i molekylet. For eksempel spurte et av pretestens oppfølgingsspørsmål etter hvilke hydrogenkjerner som gav opphav til de ulike signalene, og noen av elevene svarte med å skrive ned hele metylgruppen «CH₃».

Når det gjaldt å tolke finstrukturen ved hjelp av $n+1$ -regelen oppdaget jeg at dette var den vanskeligste informasjonen for elevene å tolke. Årsaken til at dette kan være forvirrende for mange elever, er at finstrukturen sier noe om antallet hydrogenatomer i nabogruppen til atomgruppen som gir opphav til signalet. Den mest gjennomgående feilen i bruken av $n+1$ -regelen var at elevene regnet med n som antall hydrogenatomer på den aktuelle atomgruppen i stedet for på nabogruppen. Taber (2002) mener at i forklaringen av atomets oppbygning, legges det ofte liten vekt på de elektrostatiske kreftene som virker mellom kjernen og

elektronene, noe som kan medføre at elever overser atomkjernenes rolle i NMR. Læreboken til elevene i utvalget tar for seg en rask forklaring av hvordan finstrukturen dannes og skriver at «hydrogenatomene...kan ha to ulike posisjoner, enten med eller mot det ytre feltet» (Brandt & Hushovd, 2011, s. 197). Dette kan skape en misoppfatning om at det er atomer som gir opphav til toppene, og ikke atomkjerners spinn som igjen danner et magnetisk moment. En grunnleggende forståelse av påvirkningskraften til et elektromagnetisk felt på en atomkjerne med spinn, er uten tvil fordelaktig i tolkningen av et NMR-spekter. Selv om teorien bak NMR ikke er direkte implementert i læreplanen som et kompetansemål, ble den likevel gjennomgått i videoforelesningene da jeg mener dette vil gagne elevenes læring av fagbegrepet. Elevevalueringen av arbeidsmetoder viste at *undervisning* ble foretrukket for å lære om NMR, noe som er forståelig da temaet kan være komplisert å lese på egenhånd.

5.4.3 Massespektrometri

Oppgave 5 dreide seg om massespektroskopi (MS) og resultatene i Tabell 11 viser en tydelig forbedring i skår fra pretest til posttest. For å finne strukturen til molekylet i pretesten, måtte elevene koble spesifikke isotoper av et halogen til et karakteristisk mønster i spektret. Posttesten hadde mest fokus på fragmentering av molekylet og hvordan dette hadde en sammenheng med toppene i spekteret. Selv om oppgavene på testene ikke var helt like ivaretok de til sammen den fagkunnskapen som, i henhold til læreplanen, kan forventes at elevene skal bruke i tolkning av et massespekter. I likhet med NMR mente elevene at de lærte om MS best gjennom klasseromsundervisning.

I et massespektrometer blir molekyler *ionisert* (mister/tar til seg elektroner) og disse molekylene splittes igjen i mindre fragment-*ioner* som en konsekvens av at *bindinger* brytes. Det var tegn til at flere elever manglet en forståelse for at det var *ioner* som gav opphav til toppene i massespekteret. På spørsmål om hvilket fragment som gav opphav til en topp i spekteret svarte noen av elevene «CH₃» i stedet for «CH₃⁺» eller «CH₃-ion». Barke mfl. (2009) påpeker at både elever og lærere har problemer med begrepet «ion» og at begrepet som regel blir introdusert i sammenheng med ioniske bindinger. Dette kan føre til en misoppfatning der elevene ikke klarer å knytte begrepet til andre situasjoner. De fleste elevene klarte å splitte molekylet i de riktige fragmentene i henhold til atommasse. I noen tilfeller var det derimot elever som vekslet mellom å tegne fragmenteringspilene (pilene som indikerer hvor bindingene brytes) gjennom bindingene langs karbonkjedene (korrekt) og tvers gjennom et karbonatom (feil). Dette ble spesielt gjort dersom elevene illustrerte molekylet

ved strekformel. Elevenes inkonsekvente inntegning av pilene viser usikkerhet knyttet til bindinger, molekylstruktur eller hva en strekformel representerer. Taber (2002, s. 118) rapporterer at elever noen ganger tenker på kjemiske bindinger som fysiske bånd mellom atomer i form av linjer, fjæringer eller at de er «limt» fast i hverandre. Han skriver videre at dette er en uheldig misoppfatning som en følge av hvordan av molekyler illustreres i lærebøker.

5.4.4 Omkrystallisering

Omkristallisering var gitt i oppgave 6 og var det siste spørsmålet på testene. Pretestens oppgave var formulert som et flervalgsspørsmål der elevene skulle krysse av for den påstanden som *ikke* var korrekt, mens posttesten var en åpen oppgave. Begge oppgavene krevde at elevene hadde kunnskaper om prinsippene bak omkrystallisering, noe over halvparten tilsynelatende hadde ifølge Tabell 13. Spesielt på pretesten gjorde elevene det bra, men som tidligere nevnt er risikoen med flervalgsspørsmål at elevene gjetter på et tilfeldig svar. Posttesten spurte etter hvilken temperatur etanolen burde ha dersom den skulle brukes til å vaske produktet DBA etter omkrystallisering, og hvorfor. I stedet for å svare på dette beskrev enkelte elever hele omkrystalliseringsprosessen. Dette tyder på at de enten ikke leste oppgaven nøye nok og bare antok at de skulle beskrive prinsippene bak opprensingsmetoden. En annen mulighet er at de misforstod begrepet *å vaske*. Det kan også tenkes at noen elever ville vise den kunnskapen de hadde om temaet selv om det ikke besvarte spørsmålet.

Omkristallisering baseres på at løseligheten til et stoff endres som følge av temperaturforandringer. De elevene som svarte at man burde vaske produktet med varm etanol og begrunner det med at man ønsker å fjerne urenheter har tydelig misforstått hvorfor vi varmer opp produktet i en løsning av etanol i utgangspunktet under omkrystalliseringen. Det var også elever som svarte at kald etanol egnet seg best, noe som i utgangspunktet er riktig, men begrunnet det med at urenheterne skulle krystalliseres og kastes. I sin doktoravhandling avdekket Ebenezer (1991) at elever ikke betrakter temperatur som en faktor som påvirker løseligheten av et stoff. Det er viktig å påpeke at av de fem fagbegrepene i denne masteroppgaven var omkrystallisering det eneste som var *direkte* i spill under labforsøket. Abrahams og Millar (2008) foreslår at det i enkelte tilfeller er viktig at de teoretiske ideene er i spill under den praktiske aktiviteten. Omkrystallisering var også det eneste temaet der elevene rangerte *undervisning + forsøk* og *praktisk forsøk* som de beste arbeidsmetodene. Elevenes evaluering sett sammen med at over halvparten av elevene fikk

full skår på oppgaven på posttesten tyder på at labforsøket, i kombinasjon med klasseromsundervisningen, hadde en positiv innflytelse på elevenes læring om omkrystallisering.

5.5 Evaluering og planlegging av labforsøk: syntese av DBA

Til nå har labforsøket og undervisningsoppleggets innvirkning på elevenes fagkunnskap blitt diskutert, både samlet og hver for seg. Videre vil problemstillingen belyses fra en litt annen vinkel gjennom Millar mfl. (1999) sin effektivitetsmodell i Figur 4. Abrahams og Millar (2008) utførte en case-studie der de brukte observasjon og intervju som metode og implementerte effektivitetsmodellen som et analytisk rammeverk i kombinasjon med kunnskapsdomenene (Figur 3) til Millar mfl. (1999). Til tross for at denne studien har en helt annen metodisk tilnærming og et annet type datamateriale ser jeg likevel nytte av å kunne bruke effektivitetsmodellen for å drøfte problemstillingen. Labforsøket *syntese av DBA* vurderes gjennom de fire stegene i modellen (boks A-D) med utgangspunkt i observasjoner fra selve forsøket og elevenes testresultater. Studiens datamateriale gir meg grunnlag for å si noe om hva elevene fysisk gjorde under forsøket (boks C) og hva elevene lærte om fagbegrepene (boks D).

BOKS A – Lærerens hensikt med timen: hva elevene skal lære

Hensikten med forsøket *Syntese av DBA* var å *hjelp*e elevene å lære fagkunnskap om følgende sentrale tema innen organisk kjemi: addisjonsreaksjon, eliminasjonsreaksjon, NMR, MS og omkrystallisering. Forsøket skulle gi elevene mulighet til å feste sine teoretiske kunnskaper til praktiske handlinger og hendelser. Tidligere i teorikapitlet ble det presentert fire mål med praktisk arbeid i naturfagundervisningen, og dette forsøket samsvarer med det første målet om å utvikle naturvitenskaplige fagkunnskaper.

BOKS B – Lærerens design av aktiviteten: hva elevene skal gjøre for å nå målene i boks A

Elevene skal gjennomføre en *helhetlig forskningsprosess*, som beskrevet i kapittel 1.2, for å fremstille, omkrystallisere og analysere spekter av DBA. Elevene skal følge et ferdigstilt labhefte der de under hvert steg skal besvare refleksjonsspørsmålet som er knyttet til handlingene i steget.

Ettersom problemstilling, fremgangsmåte og resultat var kjent for elevene på forhånd kan labforsøket sies å ha, etter Herrons (1971) klassifisering, *null frihetsgrader*. Valget om å

bruke en slik tilnærming ble tatt av flere årsaker. For det første ønsket jeg ikke at forsøket hensikt om å fremme fagkunnskap skulle bli overskygget av et fokus på problemløsning der elevene måtte utarbeide en egen fremgangsmåte. Det var viktig at tiden heller ble brukt til å reflektere over stegene i prosedyren. Hodson (1996) foreslår at i tilfeller der et spesifikt fenomen skal demonstreres eller observeres, kan det være fordelaktig å benytte seg av et forsøk med lav frihetsgrad. Dersom formålet imidlertid hadde vært at elevene skulle få kunnskap om naturvitenskapens egenart ville kanskje et forsøk av 2 eller 3 frihetsgrader vært å foretrekke.

For det andre er det urealistisk å forvente at elevene skal komme opp med en egen fremgangsmåte for en organisk syntese, eller «oppdage» hvilket produkt som dannes fra reaktantene. Abrahams og Millar (2008) understreker at mange lærere havner i fallgruven hvor de overvurderer elevenes evner i praktisk arbeid. Planleggingen av en organisk syntese ligger langt over hva læreplanen og kompetansemålene tilsier at elevene skal kunne. En slik kunnskap er visstnok kjent for personer som har studert kjemi på et høyere nivå, og det er nettopp her skillet går mellom elever i skolen og forskere innen naturvitenskapen (Millar, 2004). Selv om jeg har valgt å navngi forsøket *en helhetlig forskningsprosess* skal ikke dette tolkes analogt med hvordan forskere praktiserer for å oppnå kunnskap da det er noe ganske annet. Et annet element jeg valgte å utelate var skriving av labrapport som en del av etterarbeidet. Rapportskriving i forbindelse med praktisk arbeid er hyppig brukt i skolen både som et vurderingsgrunnlag, og fordi det ofte inngår i lærebøkens beskrivelse av en «naturvitenskapelig arbeidsmetode» (P. M. Kind, 2003). Dessverre har rapporter en tendens til å bli et «ritual» elevene følger uten mye tanker rundt innsamlet data (Lunetta mfl., 2007). Jeg fant ikke tid til å verken tilrettelegge for eller gjennomføre rapportskriving som ville ha gagnet elevenes læring og av den grunn ble dette ikke inkludert i *forskningsprosessen*.

BOKS C – Hva elevene faktisk gjorde under aktiviteten (fysisk)

Samtlige elever fullførte prosedyren og fikk fremstilt og omkrystallisert produktet. I den første klassen som gjennomførte brukte noen av elevene 96 % etanol til omkrystalliseringen i stedet for 70 % etanol. Dette resulterte at produktet ikke ble krystallisert, og det var dermed vanskelig å utvinne nok fast stoff til å beregne utbytte. For å unngå at dette gjentok seg laget jeg en 70%-blanding i forkant av forsøket som jeg spesifikt ba elevene bruke i de påfølgende klassene. De fleste spørsmålene elevene kom med omhandlet praktiske ting som hvor de kunne finne utstyr, og oppsettet på vakuumfiltreringen. Det ble observert at samtlige elever

jobbet med refleksjonsspørsmålene i labheftet under forsøket selv om det nødvendigvis ikke skjedde i kronologisk rekkefølge etter en utført handling. Mesteparten av elevenes spørsmål angående labheftet var knyttet til refleksjonsspørsmålet om reaksjonstype/reaksjonsmekanisme, og spørsmålet om «grumset» (bunnfallet) som ble dannet i aceton-benzaldehyd løsningen. Begge disse spørsmålene var knyttet til temaet *reaksjonsmekanismer* og jeg måtte ofte peke på reaksjonslikningen illustrert i labheftet for å hjelpe elevene å se sammenhengen til dannelsen av DBA. Andre spørsmål som ble spurt var hvorfor krystallene ble vasket med vann og hvorfor løsningen måtte avkjøles langsomt i omkrystalliseringsprosessen. Læreboken sa ingenting om dette så elevene måtte selv tenke ut et svar. Elevenes svar på hvorfor krystallene burde vaskes varierte alt i fra at krystallene måtte vaskes så de ble rene, til at vann var polart så dette ikke ville løse opp DBA.

Som siste del av forsøket analyserte elevene NMR og MS spekteret av sitt eget produkt. Dette etterarbeidet var siste steg i undervisningsopplegget og var også en del av den helhetlige forskningsprosessen. I fellesskap ble opphavet til signalene/toppene diskutert, og det var flere elever som foreslo at urenheter kunne skape «forstyrrelser» i spekteret. Enkelte NMR spekter gjenspeilet at produktet ikke fikk tørket skikkelig da det dukket opp signaler for vann og etanol. Elevene gav uttrykk for at de syntes det var greit å ha en felles gjennomgang da spektrene er noe mer kompliserte enn de fra læreboken.

BOKS D – Hva elevene faktisk lærte av aktiviteten (kognitivt)

Evaluering av elevenes læring tok utgangspunkt i resultatene fra *posttesten*. Ettersom læring av de fem fagbegrepene ble drøftet i kapittel 5.4, vil denne delen kun oppsummere de øvrige funnene for å få et mer sammensatt bilde av hva elevene, etter testen å bedømme, kan.

Ut av alle oppgavene skåret majoriteten høyest på oppgave 2 der de skulle skrive reaksjonstype og krysse av for det riktige produktet. Hele 61 av 78 elever fikk full skår på denne oppgaven. Oppgavene om reaksjonsmekanismer var desidert vanskeligst, spesielt addisjonsreaksjonen mellom reaktantene fra labforsøket. Over halvparten av elevene fikk til oppgaven om massespektrometri og 74 % klarte å finne rett molekylstruktur i henhold til spekteret. Hva gjaldt NMR svarte elevene totalt sett ganske bra. Utfordringen med NMR var anvendelsen av «n+1»-regelen. I det siste temaet, omkrystallisering, skåret over halvparten i det høyeste poengsjiktet, og de gale svarene var som regel et resultat av at elevene ikke forstod begrepet «å vaske» og forvekslet det med selve omkrystalliseringen.

Effektivitetsnivå

Effektivitetsmodellen består av to nivå, der samsvar mellom boks B & C og A & D henholdsvis tilsvarer effektivitetsnivå 1 og 2. Ut fra mine observasjoner, og det faktum at samtlige elever utførte prosessen i labheftet, antydes det en match mellom boks B og C. Dette innebærer at elevene fysisk gjorde det som var tiltenkt med utstyr og materiale for å gjennomføre forsøket. Effektivitetsnivå 1 kan derfor sies å være oppfylt i det Millar mfl. (1999) beskriver som *objekt*-domenet. Ut fra studiens datamateriale var det derimot ingen holdepunkt for å trekke den samme slutningen om elevenes mentale engasjement under forsøket. Abrahams og Millar (2008) argumenterer for at en aktivitet først og fremst må være på effektivitetsnivå 1 i objekt-domene før det i det hele tatt kan nå effektivitetsnivå 2 i idé-domenet. Å trekke en kobling mellom boks A og D er imidlertid litt mer komplisert enn for boks B og C, da dette handler om elevens mentale læringsutbytte av fagkunnskap.

Det er flere faktorer som kan hindre elevenes læring gjennom praktiske aktiviteter. Det kan blant annet være at elevene ikke er klar over læringsmålene med aktiviteten, at det er for *mange* mål, eller at læreren har urealistiske forventninger til elevene (Abrahams & Millar, 2008; Hart mfl., 2000; Séré, 2002). Både før og under undervisningsopplegget forsøkte jeg å formidle målene i boks A, men det ble klart for meg under forsøket at enkelte elever likevel ikke oppfattet sammenhengen mellom undervisningen de hadde hatt i forkant og labforsøket. Det advares mot å tillegge et praktisk forsøk for mange læringsmål da elever, og mennesker generelt, har en øvre grense for hvor mye informasjon som kan tas inn uten at det blir et kognitivt kaos (Hart mfl., 2000; Lunetta mfl., 2007). I etterkant har jeg vært innom tanken at det muligens var litt i overkant å skulle tillegge undervisningsopplegget og forsøket *fem* fagbegreper som læringsmål. Samtidig gikk alle fagbegrepene under ett og samme kapittel i læreboken, og det er ikke unormalt at tilsvarende mengde fagstoff blir testet i en tradisjonell skoleprøve. Siden aktivitetens mål var rettet mot læring av fagkunnskap, og forsøket i utgangspunktet innbefattet fem sentrale tema i organisk kjemi, virket det unaturlig å utelukke noen av dem. Dersom fokuset skulle ha vært på å lære kun ett av fagbegrepene, for eksempel reaksjonsmekanismer, ville det kanskje ikke egnet seg å gjennomføre omkrystallisering eller NMR og MS analyser, og forsøket burde nok hatt en annen oppbygning. Det er lærerens ansvar å designe en praktisk aktivitet som ikke ligger utenfor elevenes proksimale utviklingssone da dette medfører at elevene ikke har noe annet valg enn å følge eventuelle instruksjoner slavisk (Lunetta mfl., 2007). Fra et sosiokulturelt syn burde den praktiske aktiviteten ligge innenfor elevenes proksimale utviklingssone for optimal læring, og dette

krever at læreren er kjent med elevenes evner og tidligere kunnskaper. Ettersom jeg kom inn som forsker og lærer uten noe kjennskap til elevene ble forsøket designet uten å ta hensyn til disse elementene, noe som visstnok har sine svakheter. I følge Berry mfl. (1999) husker de færreste elever tidligere gjennomførte labforsøk, noe som igjen antyder at de ikke lærte stort av det heller. Forsøk av minneverdig karakter derimot, der det skjer noe «utenom det vanlige», eller som foregår i uvante omgivelser kan bidra til at elevene husker det i ettertid (Abrahams & Reiss, 2012). Studiens labforsøk ble utført i Forskningsparken, og til tross for manglende bevis, vil jeg foreslå for at dette skapte en annen setting med mer «forsker»-preg enn elevenes ordinære klasserom. Samtidig er det en risiko for at elevenes husk av forsøket begrenses til de praktiske omstendigheter og hva de faktisk lærte eller skulle lære.

Jeg ser det som problematisk å skulle trekke en konklusjon om forsøkets effektivitet på nivå 2 basert på poengskårene fra posttesten da læring er en kompleks prosess som krever at elevene lagrer kunnskapen i langtidsmindet (Ringnes & Hannisdal, 2014). Fra et kognitivt konstruktivistisk ståsted innebærer dette at elevene har endret sine skjema gjennom enten akkomodasjon eller assimilasjon, og slik mental aktivitet er som sagt vanskelig å oppdage gjennom en skriftlig test.

5.6 Forsøk som en helhetlig forskningsprosess

Hvorvidt studiens labforsøk i form av en *helhetlig forskningsprosess* egner seg til å lære fagkunnskap fremfor et annet type forsøk, drøftes mot litteratur og tidligere forskning på området.

Essensen i den *helhetlige forskningsprosessen* var at elevene skulle skape sitt *eget* produkt som de skulle renske og analysere. Av grunner jeg nevnte i forrige kapittel måtte forsøket gjennomføres under kontrollerte forhold med null frihetsgrader. Derfor var det viktig å gi elevene den «friheten» som var mulig ut fra rammene som var gitt. I stedet for å gi elevene handlingsrom når det kom til planlegging av forsøk eller utforming av prosedyre, fikk de frihet gjennom det å skape noe *eget*. Elevene kan ha følt et visst ansvar for å bearbeide produktet sitt slik at de fikk et akseptabelt utbytte med minst mulig urenheter, som igjen ville blitt avdekket gjennom NMR og MS analyse. Å fremkalle en følelse av eierskap under et labforsøk kan bidra til å styrke elevenes motivasjon og læring (Berry mfl., 1999; Enghag & Niedderer, 2008). Det kan tenkes at elevenes visitt til Instituttet for kjemi der de fysisk fikk se NMR og MS maskinene, samt snakke med en kjemiker, påvirket deres eierskapsfølelse rundt forsøket. Ut fra innsamlet datamateriell kan dette imidlertid verken avkreftes eller bekreftes.

Den *helhetlige forskningsprosessen* innebærer ikke bare selve forsøket, men også relevant underveis – og etterarbeid. Köller, Olufsen, Stofanovska og Petruševski (2015) foreslår at læring fra praktisk arbeid kan forekomme dersom elevene får tid og mulighet til diskusjon og refleksjon. Refleksjonsspørsmålene i labheftet er ment som et verktøy for å fremme diskusjoner. Labforsøk har ofte en annen sosial kontekst enn det vanlige klasserommet, og gir elevene anledning til produktivt samarbeid (Hofstein & Lunetta, 2004). Dette sosiale aspektet er også noe som karakteriserer naturvitenskapen og måten naturvitenskapelig kunnskap dannes på (Lederman mfl., 2014). Elevene jobbet i par under labforsøket, og det ble observert flere diskusjoner rundt spørsmålene både innad og på tvers av lab-parene. Det var imidlertid vanskeligere å tilrettelegge for faglige diskusjoner mellom elevene og meg selv, da mesteparten av tiden gikk med til å assistere elevene i selve gjennomførelsen. En reell utfordring for lærere er å hjelpe elevene sine til å ta kontroll over egen læring. Det er lærerens oppgave å støtte elevene i deres læringsprosess gjennom tilbakemeldinger og faglige samtaler (Hofstein & Lunetta, 2004). Gjennom etterarbeidet med analysespektrene fikk elevene mulighet til å diskutere eventuelle feilkilder i fellesskap. Da disse spektrene også gjenspeiler de andre fagbegrepene, var etterarbeidet et viktig steg i å hjelpe elevene koble sammen domenet av objekter og ideer (se Figur 3).

Jeg vil argumentere for at det er en vesentlig forskjell mellom å gjennomføre *syntese av DBA* og en *helhetlig forskningsprosess*. Ved å kun utføre syntesen ville forsøket i større grad fått et preg av «kokebok» stil. *Forskningsprosessen* innebar elementer som fremmet refleksjon, diskusjon og eierskapsfølelse og kan på den måten skilles fra et tradisjonelt «kokebok»-forsøk.

5.7 Svakheter med studien

I metodekapitlet ble studiens validitet og reliabilitet diskutert i form av studiens gyldighet og pålitelighet. I dette kapitlet vil jeg trekke frem konkrete utfordringer knyttet til innsamlet datamateriale og gjennomføringen av undervisningsopplegget.

En summativ vurdering av elevenes læring basert på kun skriftlige tester har sine svakheter. Testene gir meg et begrenset innblikk i elevenes tankegang og hvilke kunnskaper fra labforsøket eller undervisningsopplegget de brukte for å løse problemene. Dette kunne for eksempel ha vært avdekket ved å gjennomføre dybde intervju av et utvalg elever. En slik metode ville også gitt et bedre grunnlag enn kun observasjon til å si noe om elevenes tanker under selve labforsøket, og hvordan de knyttet det til tidligere klasseromsundervisning. En

annen utfordring med testene, og for så vidt labheftet, var at enkelte av elevene ikke hadde norsk som morsmål. Denne språkbarrieren kan ha påvirket deres prestasjoner på testene/labheftet i og med at de krevde både lesing og en god del skriving.

Studiens funn viser at 7 av 13 elever som fikk over 80% riktig på pretesten hadde en negativ endring i skår. Funnet betyr nødvendigvis ikke at elevene gjorde det dårlig på posttesten, men er en konsekvens av at de åpenbart har en større fallhøyde og mindre et forbedringspotensial enn elever som i utgangspunktet skåret lavt på pretesten. For å unngå dette kunne testene hatt et annet design, for eksempel med oppgaver av variert vanskelighetsgrad. Da studiens hovedfokus var på elevenes *skårifferanse*, og ikke den enkelte skåren, ville slike oppgaver ikke ha skadet testenens hensikt og heller gitt elever med høy måloppnåelse et større spillerom for forbedring.

Det var ønskelig at undervisningsopplegget, inkludert gjennomføring av testene, skulle gjennomføres så likt som mulig på tvers av klassene i utvalget. Alle utførte pretesten på morgenen før labforsøket, og posttesten omtrent en ukes tid etter. Den største forskjellen, som jeg tror kan ha hatt en betydelig for elevens prestasjoner, var at i klasse K2 ble posttesten gjennomført som en *del* av en større test innenfor temaet organisk kjemi. Denne andre testen var laget av deres faglærer og ble brukt som underveisvurdering i faget. K2 var den eneste klassen som hadde en gjennomsnittlig nedgang i skår. Da jeg forhørte meg med faglæreren om testresultatene, fortalte læreren at flere av elevene gikk tom for tid på prøven, og at enkelte hadde blitt slitne slutten og av den grunn ikke gjennomført. For ordens skyld ønsker jeg å informere om at posttesten ble gjennomført slik etter elevenes eget ønske da de hadde prøver i andre fag de måtte ta hensyn til.

I ettertid oppdaget jeg at alternativene i evalueringsskjemaet over arbeidsmetoder ikke var spesifikke nok. Det burde ha kommet klarere frem at evalueringen gjaldt arbeidsmetoder i den hensikt å fremme *fagkunnskap*. Det kan tenkes at elever koblet alternativet «praktisk arbeid» som arbeidsmetode opp mot å lære praktiske ferdigheter. Alternativet «undervisning» skulle vært presisert til at det innebar alt som skjedde i klasserommet på skolen, inkludert oppgaveregningen. Da det ble åpenbart at flere elever benyttet seg av egenstudier innser jeg at alternativene «klasseromsundervisning + lese selv» og «praktisk forsøk + lese selv» med fordel skulle ha blitt inkludert. Selv om elevene fikk mulighet til å skrive opp en egen arbeidsmetode var det ytterst få som benyttet seg av dette.

Til slutt vil jeg bemerke at bruk av en kontrollgruppe kunne vært nyttig for å styrke den indre validiteten. Kontrollgruppen kunne enten ha blitt eksponert for tradisjonell tavleundervisning uten praktiske innspill, eller ha utført et annet type forøk for å lære de sentrale fagbegrepene. Ut fra prosjektets rammer var det imidlertid ikke mulig å ta i bruk en slik gruppe, og denne beslutningen er redegjort for i kapittel 3.

6 Avslutning

Som siste kapittel i denne masteroppgaven presenteres studiens hovedfunn og implikasjoner for fremtidig forskning.

6.1 Konklusjon

Formålet med denne masteroppgaven er å se på hvilke innvirkninger et praktisk undervisningsopplegg har på elevers faglige prestasjoner. Studiens forskningshypotese «*Et undervisningsopplegg som innebærer et praktisk forsøk i form av en helhetlig forskningsprosess, vil styrke elevenes fagkunnskaper om sentrale begrep i organisk kjemi*», ble undersøkt gjennom resultater fra en pretest og en posttest, henholdsvis før og etter det praktiske forsøket.

Det ble avdekket en statistisk signifikant *forbedring* i poengskår fra pretest til posttest. Altså gjorde utvalget det i gjennomsnitt bedre *etter* å ha gjennomført labforsøket. Videre er spørsmålet om dette resultatet skyldes labforsøket utelukkende, eller labforsøket som en del av undervisningsopplegget. Forskningshypotesen antar sistnevnte, og de statistiske funnene om en signifikant, positiv økning kan betraktes som et bevis for at elevenes fagkunnskaper ble styrket.

Abrahams og Millar (2008) påpeker at læring fra praktisk arbeid kan styrkes dersom elevene på forhånd introduseres for de teoretiske ideene aktiviteten er ment å fremme. I tillegg er det anbefalt å diskutere teorien i etterkant slik at elevene får mulighet til å bearbeide inntrykkene fra aktiviteten (Hofstein & Lunetta, 2004). Disse elementene var en del av undervisningsopplegget gjennom videoforelesninger og etterarbeidet med elevenes analysespekter i fellesskap. I en rapport fra *European Comission* sitt forskningsprosjekt om praktisk arbeid i naturfag, konkluderer Séré (2002, s. 638) med at «*The view of labwork serving conceptual knowledge exclusively must be abandoned*». Dette utsagnet støttes opp av elevens lave rangering av «*praktisk forsøk*» som læringsmetode sammenlignet med de øvrige alternativene. Samlet sett tyder det på at elevenes forbedring av testresultater kan forklares gjennom undervisningsopplegget i sin helhet, fremfor labforsøket alene.

Suksessen av en praktisk aktivitet bedømmes ut fra om den samsvarer med de respektive læringsmålene (Sjøberg, 2011). I denne studien ble det praktiske arbeidet gjennomført som en *helhetlig forskningsprosess* (labforsøk + etterarbeid), med hensikt å styrke fagkunnskaper om spesifikke begreper. Hodson (1996) foreslår at under omstendigheter der bestemte deler av

læreplanen skal læres, er det fordelaktig å tilskrive forsøket en lav grad av frihet. På bakgrunn av at labforsøket innebar en organisk syntese så jeg det nødvendig å allokere en stegvis lab-prosedyre. Dette gav elevene mer tid til å kunne reflektere over de respektive stegene, et moment som trekkes frem som viktig i en læringsprosess (Köller mfl., 2015). Selv om handlingsfriheten til elevene under forsøket var noe begrenset, tilrettela *forskningsprosessen* for eierskapsfølelse gjennom å la elevene skape, bearbeide og analysere eget produkt. Å implementere et slikt element kan i følge Enghag og Niedderer (2008) være med på å øke elevens motivasjon og interesse, noe om ofte har en positiv korrelasjon med læring. Hart mfl. (2000) gjennomførte en studie med åpne forsøk og konkluderte med at den typen forsøk var uegnet til å lære elevene naturvitenskapelig fagkunnskap. Det kan altså se ut som at den *helhetlige forskningsprosessen* var passende til sitt formål med å lære fagkunnskaper. Denne tilnærmingen til praktisk arbeid kan dessuten brukes som inspirasjon til videre utvikling og gjennomføring av praktiske undervisningsopplegg i *kjemi 2*.

Til slutt vil jeg trekke frem lærerens fornemme rolle i å styrke elevenes fagkunnskaper fra praktisk arbeid. Læreren skal hjelpe elevene å lage en kobling mellom teori og praksis. En viktig del av dette er å kartlegge elevenes hverdagsforestillinger og misoppfatninger. Videre bør læreren utfordre eleven i hans eller hennes proksimale utviklingszone. Læreren må imidlertid ikke ha urealistiske forventninger som overskrider denne sonen da dette kan hemme læringsprosessen (Millar, 2004). I gjennomsnitt fikk elevene 18.5 poeng av 30 mulige rett på posttesten, noe som tilsvarer 62 %. Dette resultatet kan sies å reflektere middel måloppnåelse og ligger på omtrent samme nivå som snittkarakteren i *kjemi 2*-eksamen fra 2018 (Utdanningsdirektoratet, 2018). Fra egne erfaringer som lærer i skolen opplever jeg at elever i større grad skårer bedre på kapitteltester enn store prøver (heldagsprøver, eksamen osv.), og siden elevene i denne studien skåret på det samme snittnivået som på eksamen, kan det vitne mot at mine tester var noe krevende for elevene ut fra deres forutsetninger fra undervisningsopplegget. Samtidig reflekterer mine tester læreplanens kompetansemål som igjen beskriver hva som kan forventes at elevene mestrer.

6.2 Veien videre

Det har blitt gjort mye forskning på læring fra praktisk arbeid i naturfag opp gjennom årene. Hodson 1990 i Abrahams og Millar (2008, s. 1946) skrev at praktisk arbeid i skolen var dårlig gjennomtenkt, forvirrende og lite produktivt i forhold til elevenes læringsutbytte. Det er gått 29 år siden Hodsons bekymringsmelding, og mye har skjedd siden både innenfor skolen og forskning

. Det har blitt et større fokus på blant annet utforskende arbeid, hvordan elever lærer i praktiske sammenhenger og lærerens rolle. Likevel rapporterte Hofstein og Lunetta i senest 2004 at det fremdeles er et stort avvik mellom hvordan faglitteraturen anbefaler at praktisk arbeid skal foregå i undervisningen, og hvordan det faktisk praktiseres.

Gjennom arbeidet med denne masteroppgaven har jeg oppdaget flere utfordringer knyttet til læring gjennom praktisk arbeid. For det første stilles det høye krav til lærerens faglige og didaktiske kompetanse som planlegger av aktiviteten, veileder under gjennomføring og stilasbygger for elevenes faglige utvikling. En praktisk aktivitet må samtidig passe innenfor skolens budsjett, laboratoriefasiliteter og tidsrammer. For videre forskning kunne det vært interessant å undersøke bruk praktiske tilnærminger for å fremme fagkunnskap fra en annen vinkel – nemlig lærerens.

Referanseliste

- Abrahams, I. (2009). Does practical work really motivate? A study of the affective value of practical work in secondary school science. *International journal of science education*, 31(17), 2335-2353.
- Abrahams, I. & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, 30(14), 1945-1969.
- Abrahams, I. & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, 49(8), 1035-1055.
- Arnesen, N. & Ødegaard, M. (2010). Hva skjer i naturfagklasserommet? – resultater fra en videobasert klasseromsstudie; PISA+. *Nordina: Nordic Studies in Science Education*, 6(1), 16-32.
- Barke, H. D., Hazari, A. & Yitbarek, S. (2009). *Misconceptions in chemistry: Addressing perceptions in chemical education*.
- Berry, A., Gunstone, R., Loughran, J. & Mulhall, P. (2001). Using laboratory work for purposeful learning about the practice of science. I *Research in science education—past, present, and future* (s. 313-318): Springer.
- Berry, A., Mulhall, P., Gunstone, R. & Loughran, J. (1999). Helping students learn from laboratory work. *Australian Science Teachers' Journal*, 45(1), 27-31.
- Bjørndal, C. R. P. (2017). *Det vurderende øyet : observasjon, vurdering og utvikling i pedagogisk praksis* (3 utg.). Oslo: Gyldendal akademisk.
- Brandt, H. & Hushovd, O. T. (2011). *Kjemi 2 : studiespesialiserende utdanningsprogram*. Oslo: Aschehoug.
- Bybee, R., McCrae, B. & Laurie, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 46(8), 865-883.
- Christoffersen, L. & Johannessen, A. (2012). *Forskningsmetode for lærerutdanningene*. Oslo: Abstrakt forl.
- Cohen, L., Manion, L. & Morrison, K. (2018). *Research methods in education* (8 utg.). London, England: Routledge.
- Cooper, M. M. & Stowe, R. L. (2018). Chemistry Education Research—From Personal Empiricism to Evidence, Theory, and Informed Practice. *Chemical reviews*, 118(12), 6053-6087.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4 utg.). Thousand Oaks, California: Sage publications.
- Dysthe, O. (2008). Klasseromsvurdering og læring. *Bedre skole*, 4(8), 16-23.
- Ebenezer, J. V. (1991). *Students' conceptions of solubility: a teacher-researcher collaborative study* (Doktoravhandling). University of British Columbia. Hentet fra <https://open.library.ubc.ca/media/download/pdf/831/1.0100936/1>
- Eikemo, T. A. & Clausen, T. H. (2012). *Kvantitativ analyse med SPSS : en praktisk innføring i kvantitative analyseteknikker* (2 utg.). Trondheim: Tapir akademisk forl.
- Engh, R. (2014). Elevvurdering. I M. B. Postholm & T. Tiller (Red.), *Profesjonsrettet pedagogikk 8-13* (s. 130-146). Oslo: Cappelen Damm Akademisk.
- Enghag, M. & Niedderer, H. (2008). Two Dimensions of Student Ownership of Learning During Small-Group Work in Physics. *International Journal of Science and Mathematics Education*, 6(4), 629-653.

- Hart, C., Mulhall, P., Berry, A., Loughran, J. & Gunstone, R. (2000). What is the purpose of this experiment? Or can students learn something from doing experiments? *Journal of Research in Science Teaching*, 37(7), 655-675.
- Herron, M. D. (1971). The nature of scientific enquiry. *The school review*, 79(2), 171-212.
- Hodson, D. (1990). Practical work in science: Time for a reappraisal. *Studies in Science Education*, 19(1), 175-184.
- Hodson, D. (1996). Practical work in school science: exploring some directions for change. *International Journal of Science Education*, 18(7), 755-760.
- Hofstein, A. & Lunetta, V. N. (1982). The Role of the Laboratory in Science Teaching: Neglected Aspects of Research. *Review of Educational Research*, 52(2), 201-217.
- Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty - first century. *Science Education*, 88(1), 28-54.
- Imsen, G. (2005). *Elevens verden - Innføring i pedagogisk psykologi* (4 utg.). Oslo: Universitetsforlaget.
- Kind, P. M. (1999). *Hva i all verden gjør elevene i realfag? : praktiske oppgaver i matematikk og naturfag*. Oslo: Universitetet i Oslo, Institutt for lærerutdanning og skoleutvikling.
- Kind, P. M. (2003). Praktisk arbeid og naturfagvitenskapelig allmenndannelse. I D. Jorde & B. Bungum (Red.), *Naturfagdidaktikk: Perspektiver, forskning, utvikling* (s. s. 226-244). Oslo: Gyldendal akademisk.
- Kind, V. (2004). *Beyond appearances: Students' misconceptions about basic chemical ideas*. Durham: Durham University.
- Kjærnsli, M. & Jensen, F. (2016). *Stø kurs: Norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015*: Universitetsforlaget.
- Kjærnsli, M., Lie, S., Olsen, R. V. & Roe, A. (2007). *Tid for tunge løft: norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006*: Universitetsforlaget.
- Kolstø, S. D. & Knain, E. (2011). *Elever som forskere i naturfag* (2 utg.). Oslo: Universitetsforlaget.
- Kristensen, Ø. & Wikan, A. (2016). *Sannsynlighetsregning og statistikk for høyere utdanning*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Kunnskapsdepartementet. (2018a). *Overordnet del - verdier og prinsipper for grunnopplæringen*. Hentet fra <https://www.udir.no/laring-og-trivsel/lareplanverket/overordnet-del/>
- Kunnskapsdepartementet. (2018b). *Fornyer innholdet i skolen. Pressemelding 132-18*. Hentet fra <https://www.regjeringen.no/no/aktuelt/fornyer-innholdet-i-skolen/id2606028/?factbox=factbox2606066>
- Köller, H.-G., Olufsen, M., Stofanovska, M. & Petruševski, V. (2015). Practical work in chemistry, its goals and effects. I I. Maciejowska & B. Byers (Red.), *A Guidebook of Good Practice for the Pre-service Training of Chemistry Teachers* (s. 85-106).
- Lederman, N. G., Antink, A. & Bartos, S. (2014). Nature of science, scientific inquiry, and socio-scientific issues arising from genetics: A pathway to developing a scientifically literate citizenry. *Science & Education*, 23(2), 285-302.
- Lund, T., Fønnebø, B. & Haugen, R. (2006). *Forskningsprosessen*. Oslo: Unipub.
- Lunetta, V. N., Hofstein, A. & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. I S. K. Abell & N. G. Lederman (Red.), *Handbook of research on science education* (s. 393-441): Lawrence Erlbaum Associates.
- Lyngsnes, K. M. & Rismark, M. (2014). *Didaktisk arbeid* (3 utg.). Oslo: Gyldendal akademisk.

- Martin, W. E. & Bridgmon, K. D. (2012). *Quantitative and statistical research methods : from hypothesis to results*. San Francisco, Calif: Jossey-Bass.
- McComas, W. F. (1998). The principal elements of the nature of science: Dispelling the myths. I *Nature of science in science education*. Hingham, MA, USA: Kluwer Academic Publishers.
- Millar, R. (2004). The role of practical work in the teaching and learning of science. *High school science laboratories: Role and vision*, 1-24.
- Millar, R., Le Maréchal, J.-F. & Tiberghien, A. (1999). Mapping'the domain: Varieties of practical work. I J. Leach & A. Paulsen (Red.), *Practical work in science education* (s. 33-59). Roskilde/Dordrecht, Nederland: Roskilde University Press/Kluwer.
- Millar, R., Tiberghien, A. & Le Maréchal, J.-F. (2002). Varieties of labwork: A way of profiling labwork tasks. I D. Psillos & H. Niedderer (Red.), *Teaching and learning in the science laboratory* (s. 9-20). Dordrecht: Kluwer Academic Publishers.
- Mohrig, J. R., Alberg, D. G., Hofmeister, G. E., Schatz, P. F. & Hammond, C. N. (2014). *Laboratory techniques in organic chemistry : supports inquiry-driven experiments* (4 utg.). New York: W. H. Freeman and Company.
- NESH. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi*. Hentet fra <https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>
- Overton, T., Potter, N. & Leng, C. (2013). A study of approaches to solving open-ended problems in chemistry. *Chemistry Education Research and Practice*, 14(4), 468-475.
- Quale, A. (2003). Konstruktivisme i naturvitenskapen: kunnskassyn og didaktikk. I D. Jorde & B. Bungum (Red.), *Naturfagdidaktikk - Perspektiver, forskning, utvikling* (s. 86-104). Oslo: Gyldendal Akademisk
- Ringnes, V. & Hannisdal, M. (2000). *Kjemi i skolen : undervisning og læring*. Kristiansand: Høyskoleforl.
- Ringnes, V. & Hannisdal, M. (2014). *Kjemi fagdidaktikk: kjemi i skolen* (3 utg.). Oslo: Cappelen Damm Akademisk.
- Roberts, S. G. (2002). *SET for success. The supply for people with science, technology, engineering and mathematics skills*. London: HM Treasury.
- Ross, S. M. (2014). *Introduction to probability and statistics for engineers and scientists*: Academic Press.
- Séré, M. G. (2002). Towards renewed research questions from the outcomes of the European project Labwork in Science Education. *Science Education*, 86(5), 624-644.
- Sjøberg, S. (2011). *Naturfag som allmenndannelse* (3 utg.). Oslo: Gyldendal Norsk Forlag.
- Taber, K. (2002). *Chemical misconceptions: prevention, diagnosis and cure, Volume I: Theoretical Background*. London: Royal Society of Chemistry.
- Thagaard, T. (2018). *Systematikk og innlevelse : en innføring i kvalitative metoder* (5 utg.). Bergen: Fagbokforlaget.
- Utdanningsdirektoratet. (2006). *Læreplan i kjemi- programfag i utdanningsprogram for studiespesialisering (KJE1-01)*. Hentet fra <https://www.udir.no/k106/KJE1-01>
- Utdanningsdirektoratet. (2013). *Læreplan i naturfag (NAT1-03)*. Hentet fra <https://www.udir.no/k106/NAT1-03>
- Utdanningsdirektoratet. (2018). *Karakterstatistikk for videregående opplæring skoleåret 2017-18*. Hentet fra <https://www.udir.no/tall-og-forskning/finn-forskning/tema/karakterer/karakterstatistikk-for-videregaaende-opplaring-skolearet-2017-18/>
- Utdanningsdirektoratet. (2019). *Eksamensveiledning - om vurdering av eksamensbesvarelser*. Hentet fra <https://sokeresultat.udir.no/eksamensoppgaver.html?k=kjemi%202&start=1&r2=%C>

[7%82%C7%82456b73616d656e737665696c65646e696e676572&r2val=Eksamensveiledninger](https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/naturfag-rettleiende-nasjonale-kjenneteikn-pa-maloppnaing-for-standpunktvrdering-etter-10.-trinn-skolearet/)

Utdanningsdirktoratet. (2016). *Naturfag: kjenneteikn på måloppnåing*. Hentet fra <https://www.udir.no/laring-og-trivsel/vurdering/sluttvurdering/naturfag-rettleiende-nasjonale-kjenneteikn-pa-maloppnaing-for-standpunktvrdering-etter-10.-trinn-skolearet/>

Vedlegg 1: Samtykkeskjema

Til deg som tar Kjemi 2!

Jeg holder på med et mastergradsprosjekt for Universitetet i Tromsø som omhandler praktisk arbeid i organisk kjemi. Jeg ønsker å se på om labforsøk som illustrerer en helhetlig forskningsprosess kan være med på å hjelpe forståelsen for sentrale begreper i organisk kjemi. Begrepene som skal fokuseres på er: addisjonsreaksjon, eliminasjonsreaksjon, NMR, MS og omkrystallisering.

I forbindelse med dette vil det bli delt ut to prøver; *én før* og *én etter* labforsøket for sammenligning. Disse prøvene tester både kunnskap og forståelse for de aktuelle begrepene.

I tillegg skal det skrives en liten rapport fra forsøket som også vil bli brukt som data. Under forsøket vil jeg notere ned observasjoner og muligens bruke dette som data dersom det er relevant.

Resultatene fra prøvene og lab rapporten vil behandles konfidensielt og vil anonymiseres.

Det er kun jeg og din faglærer som vil ha tilgang på testene og rapportene.

Det er frivillig å delta og du kan når som helst trekke deg dersom du ønsker det.

Jeg håper du vil delta i prosjektet mitt!

Med vennlig hilsen

Janita Fosslund Hansen

Ta kontakt på: janitah_94@hotmail.com dersom du lurer på noe!

Dato:

Deltakers signatur:

Vedlegg 2: Godkjenning fra NSD

1.5.2019

Meldeskjema for behandling av personopplysninger

NSD NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Forskningsprosess i organisk kjemi 2

Referansenummer

323496

Registrert

18.09.2018 av Janita Yu Fosslund Hansen - jha123@post.uit.no

Behandlingsansvarlig institusjon

UiT Norges arktiske universitet / Fakultet for humaniora, samfunnsvitenskap og lærerutdanning / Institutt for lærerutdanning og pedagogikk

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Hans-Georg Køller, hans.koller@uit.no, tlf: 77645509

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Janita Fosslund Hansen, janitah_94@hotmail.com, tlf: 97977282

Prosjektperiode

01.09.2018 - 01.06.2019

Status

14.11.2018 - Vurdert

Vurdering (1)

14.11.2018 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet med vedlegg den 14.11.2018, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD ENDRINGER

Vedlegg 3: Vurderingskriterier for pretest og posttest

Retningslinjene som ble brukt i vurderingsarbeidet av testene.

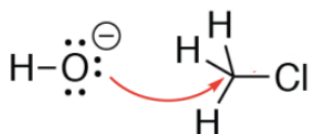
Kompetansemål	Kompetansenivå	Kjennetegn på måloppnåelse
<i>Gjøre rede for reaksjonstypene addisjon og eliminasjon</i>	Lav	- Eleven gjenkjenner i lav grad karakteristikker ved en reaksjonstype - Eleven skiller i lav grad mellom reaksjonstypene
	Middels	-Eleven gjenkjenner og skiller reaksjonstypene - Eleven anvender delvis teori i forklaringer på reaksjonstype
	Høy	-Eleven gjenkjenner og skiller reaksjonstypene -Eleven begrunner reaksjonstype presist ut fra teori og forklarer relevante sammenhenger -Eleven forutser produkt basert på reaksjonstype og begrunner dette med teori
<i>Forklare reaksjonsmekanismen ved addisjon og eliminasjon</i>	Lav	-Eleven illustrerer reaksjonsmekanismene med moderate mangler - Eleven ser i lav grad kobling mellom teori og stegene i mekanismen
	Middels	-Eleven illustrerer reaksjonsmekanismene delvis korrekt - Eleven kan delvis begrunne stegene i mekanismen fra teori
	Høy	-Eleven illustrerer en korrekt mekanisme der elektronoverflytting og formalladninger er presisert -Eleven viser en teoretisk forståelse for stegene i mekanismen
<i>Gjøre rede for omkrystallisering</i>	Lav	-Eleven reproduserer noe fagstoff og kan til en viss grad gjenkjenne det som en separasjon metode
	Middels	-Eleven kan forklare hovedprinsippet bak omkrystallisering med delvis bruk av teori -Eleven kan forklare stegene i prosessen
	Høy	-Eleven kan forklare prinsippene bak omkrystallisering med bakgrunn i teori om løselighet og temperatur -Eleven kan forklare stegene i prosessen med bakgrunn i relevant teori
<i>Tolke enkle massespektre (MS) og ¹H-NMR-spektre</i>	Lav	-Eleven kan til en viss grad gjenkjenne og skille mellom spektrene -Elevene viser lav grad forståelse for hvordan spektrene skal tolkes
	Middels	-Eleven kan til en viss grad gjenkjenne og skille mellom spektrene -Elevene viser lav grad forståelse for hvordan spektrene skal tolkes
	Høy	-Eleven tolker spektrene korrekt ut fra teori og kjemiske tabeller - Eleven viser god forståelse for sammenhengen mellom molekylets struktur og dets respektive spekter, og kan begrunne dette ut fra hydrogenmiljø, n+1 regelen, kjemisk skift og areal (gjelder NMR). -Eleven har forståelse for hvilke innvirkninger isotoper har i et MS-spekter.

Vedlegg 4: Pretest

I den originale testen var periodesystem, tabell og kjemisk skift og forekomst av isotoper vedlagt. Denne versjonen har også et annet format for å passe inn i vedlegget, men innholdet i oppgavene er helt likt.

OPPGAVE 1

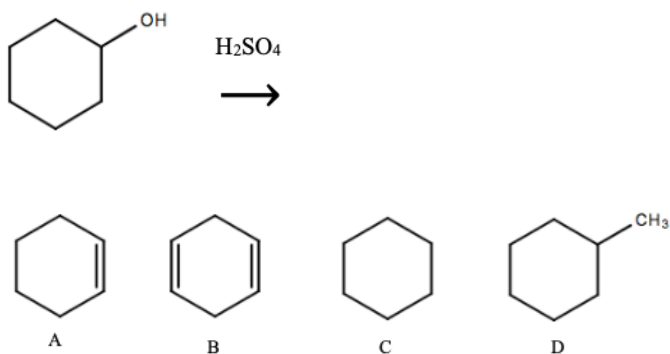
Hva indikerer den kurvede pilen



- H-atomet donerer et elektronpar for å lage en binding til C- atomet
- O-atomet donerer et elektronpar for å lage en binding til C- atomet
- C-atomet donerer et elektronpar for å lage en binding til O-atomet
- H-atomet donerer et elektronpar for å lage en binding til O-atomet

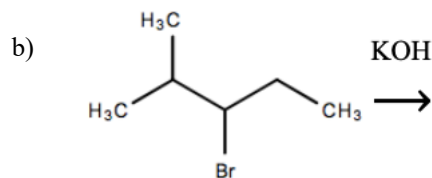
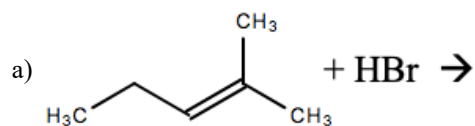
OPPGAVE 2

Hva er utfallet av reaksjonen? Og hvilken type reaksjon har funnet sted?



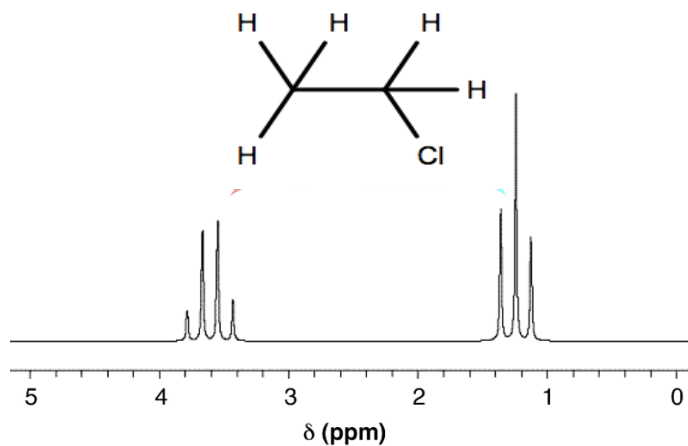
OPPGAVE 3

Hvilken reaksjon kan skje her og hva blir produktet? Gi forslag til reaksjonsmekanismen.



OPPGAVE 4

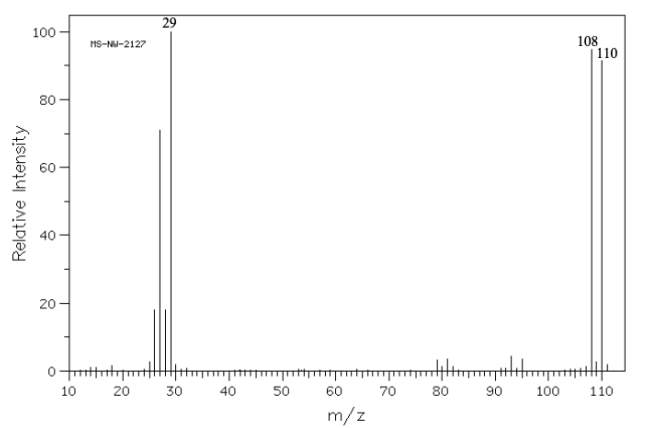
Figuren viser strukturen og $^1\text{H-NMR}$ -spekteren til kloreten. Se vedlagt tabell over kjemisk skift.



- Forklar hvorfor vi ser to grupperinger av signaler og hvilke H-atomer som gir opphav til signalene?
- Hvorfor er finstrukturen splittet slik?

OPPGAVE 5

Figuren viser et MS-spekter av et haloalkan (alkan med et halogen). Molekylionet har en masse =108 u



- Forklar hvilket halogen som er i forbindelsen og hvordan man kan finne dette ut.
- Hvilken forbindelse hører dette spekteret til?

OPPGAVE 6

Hvilken påstand om omkrystallisering er IKKE riktig

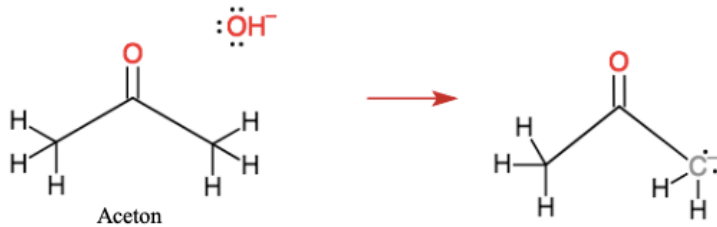
- Ønsket stoff og eventuelle urenheter må ha lik løselighet ved både høy og lav temperatur
- Et viktig aspekt i omkrystallisering er valg av løsemiddel
- Løsningen skal avkjøles sakte etter oppvarming for å fremme rene krystaller
- Ønsket stoff og eventuelle urenheter må ha ulik løselighet ved enten høy eller lav temperatur

Vedlegg 5: Posttest

I den originale testen var periodesystem, tabell og kjemisk skift og forekomst av isotoper vedlagt. Denne versjonen har også et annet format for å passe inn i vedlegget, men innholdet i oppgavene er helt likt.

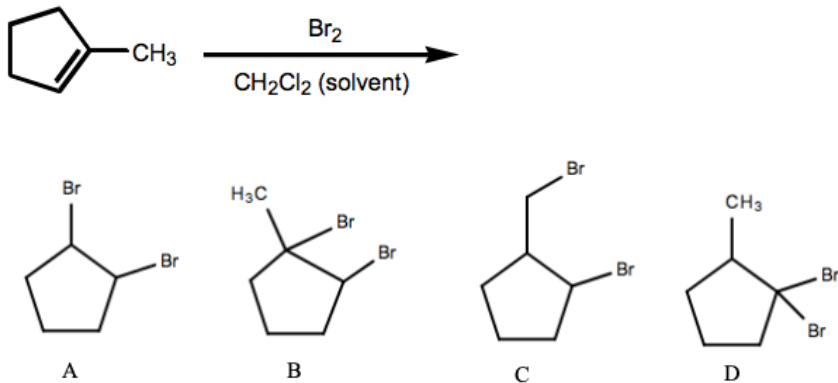
OPPGAVE 1

Tegn inn krummet pil som passer med reaksjonen



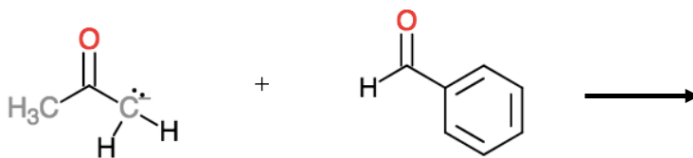
OPPGAVE 2

Hva er produktet av denne reaksjonen? Og hvilken reaksjon er det som skjer

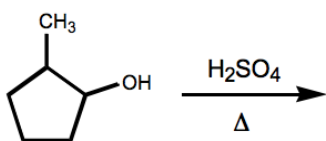


OPPGAVE 3

- a) Her ser vi et steg i reaksjonsmekanismen i syntesen av DBA. Hva dannes i det neste steget av mekanismen og hvordan dannes det? Vis med å bruke krumme piler.

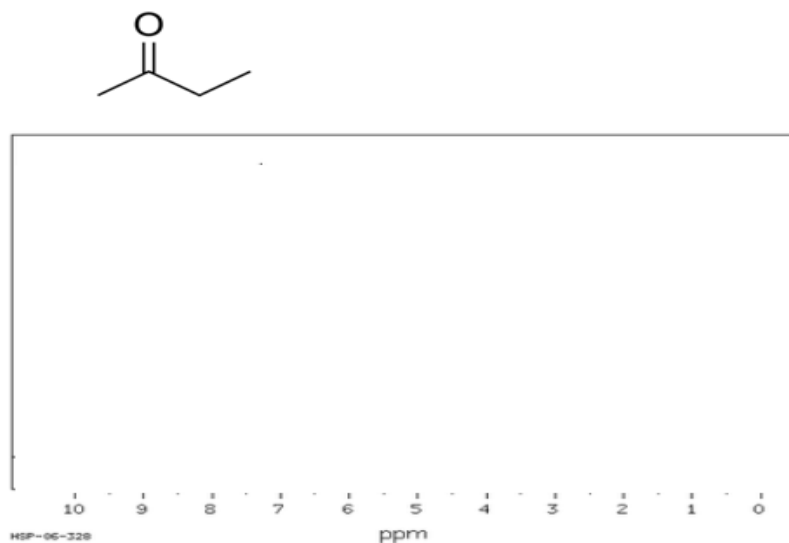


- b) Hvilken reaksjon kan skje her og hva blir produktet? Gi et forslag til reaksjonsmekanismen. Δ er et symbol for varme.



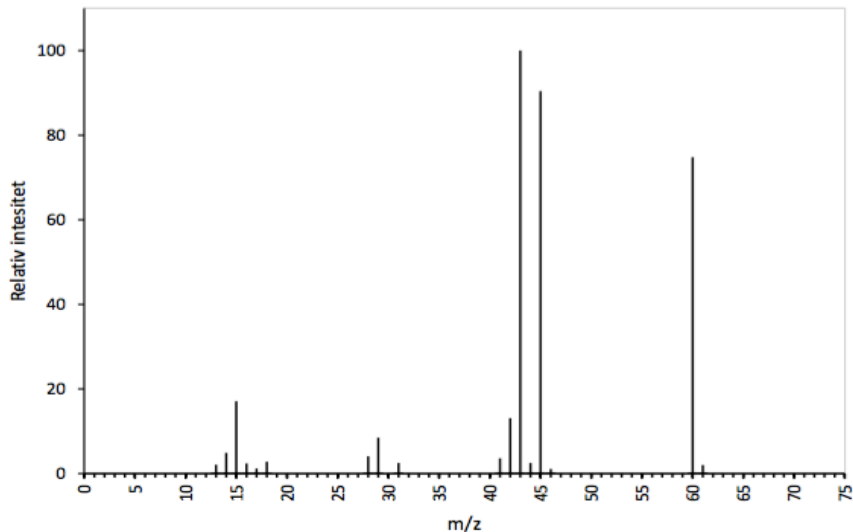
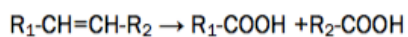
OPPGAVE 4

Foreslå hvordan NMR-spektrum til butanon vil se ut. Tegn signaler, finstruktur og forklar hvilke hydrogenatomer som tilhører hvilke topper. Hva er forholdet mellom arealet under signalene?



Oppgave 5

Et alken reagerer med KMnO₄ og gir to produkt R₁-COOH (1) og R₂-COOH (2). Massespekteret for produktet er gitt under.



Hvilket produkt har dette massespekteret opphav fra? Molekylion: m/z = 60. Tegn strukturen til molekylet og foreslå mulige fragmenteringer.

OPPGAVE 6

Som siste steg i en omkrystalliseringsprosess av DBA ønsker du å *vaske* produktet ditt. Du har to løsninger med etanol. En løsning med temperatur like under kokepunkt og en løsning du har avkjølt på isbad en god stund. Hvilken av de to bør du velge til å vaske med og hvorfor?

EVALUERING

Hvilken arbeidsmåte/undervisningsmetode følte du at du lærte best av i de aktuelle temaene. Sett X

Tema	Arbeidsmåte/Undervisningsmetode				
	Lese på egenhånd	Praktisk forsøk	Undervisning i klasserommet	Undervisning + praktisk forsøk	Annet (skriv selv)
Reaksjonsmekanismer					
NMR					
MS					
OmkrySTALLISERING					

Vedlegg 6: Labhefte

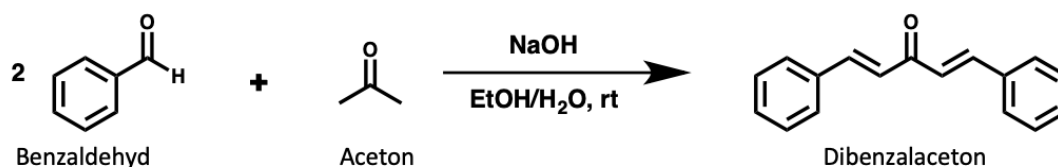
Labheftets format er modifisert til å passe inn i vedlegget. Spørsmålene og informasjonen er uendret

Introduksjon



Dibenzalacetone (DBA) er en organisk forbindelse som brukes i solkrem. DBA absorberer UV-stråling fra sola og beskytter huden vår. DBA fremstilles i en syntese hvor benzaldehyd og aceton inngår i en reaksjon med hverandre.

Reaksjonslikning



Sikkerhetsanalyse

Briller, hansker og lab frakk skal brukes til all tid for å unngå søl i øyne og ellers på huden. **NB!** Syntesen gjøres i avtrekksskap – ikke innhaler noen av stoffene!

Stoffer	Farer	Forholdsregel
Aceton	Svært brannfarlig	Hold unna varme, varmeoverflater og åpen flamme
	Kan skape irritasjon på hud og øyne	Bruk briller, hansker og lab-frakk
Benzaldehyd	Brennbar	Hold unna åpen flamme
	Kan skape irritasjon på hud og øyne	Bruk briller, hansker og lab-frakk
Dibenzalacetone	Kan skape irritasjon på hud og øyne	Bruk briller, hansker og lab-frakk
Etanol	Svært brannfarlig	Hold unna varme, varmeoverflater og åpen flamme
	Kan skape irritasjon på hud og øyne	Bruk briller, hansker og lab-frakk
Natriumhydroksid	Etsende ved øye eller -hud kontakt	Bruk briller, hansker og lab-frakk

DEL 1 – SYNTSE AV DBA

Fremgangsmåte

1. Vei ut 2.5 g NaOH. Løs dette i 25 mL vann og 20 mL 96 % etanol i en 100 mL erlenmeyerkolbe. Kjøl løsningen ned til romtemperatur og sett på røring med magnetrører.

Spm. 1: Den sterke basen NaOH brukes som en katalysator i reaksjonen. Hva gjør en katalysator?

Svar 1:

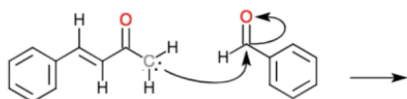
2. Vei opp 2.92 g benzaldehyd og 0.73 g aceton og løs dette i ca. 4 mL 96% etanol. **Vei ut aceton sist da den fordampes fort!**

Spm. 2: Hvilken rolle har etanol i denne syntesen?

Svar 2:

3. Tilsett halvparten av benzaldehyd-aceton blandingen til løsningen som ble laget i punkt 1 og la det røre i 10 minutter.

Spm. 3: Her er et steg av reaksjonsmekanismen som inngår i dannelsen av DBA. Hvilken reaksjon har skjedd så langt med reaktantene? Og hvilken reaksjon er det som skjer i dette steget?



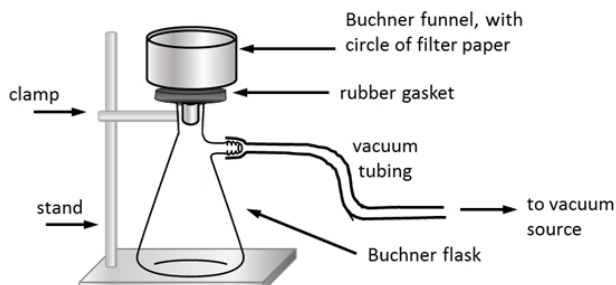
Svar 3:

4. Tilsett så resten av benzaldehyd-aceton blandingen og la det røre i nye 10 minutter

Spm. 4: Hva er det som gjør at løsningen får en gul farge? Hvorfor blir løsningen «grumsete» etterhvert?

Svar 4:

5. Bruk en büchner-trakt og vakuump-filtrer løsningen.



Spm. 5: Hva er det som blir igjen på filterpapiret?

Svar 5:

6. Vask krystallene på filtrerpapiret med 50 mL destillert vann 3 ganger. Kast filtratet i sugeflasken i vasken

Spm. 6: *Hvorfor vasker vi krystallene med vann?*

Svar 6:

DEL 2 – OMKRYSTALLISERING

For å øke renheten i produktet skal det **omkrystalliseres**.

Spm. 7: *Hva er prinsippene bak omkrystallisering?*

Svar 7:

7. Ha krystallene i et 250 mL begerglass og tilsett 150 mL 70% etanol og rør rundt med en glasstav. Ha oppi et par kokesteiner for å forhindre foss-koking.

Spm. 8: *Hvilken rolle har etanol i dette steget?*

Svar 8:

8. Plasser begerglasset på en varmeplate og dekk til med et urglass. Varm opp løsningen gradvis til kokepunkt. Dersom ikke alt løses opp under oppvarmingen tilsett litt mer 70 % etanol.

Spm. 9: *Hvorfor dekker vi begerglasset med et urglass?*

Svar 9:

9. Når alt fast stoff er løst opp tas begerglasset av varmeplaten og løsningen kjøles til romtemperatur før det settes i et isbad.

Spm. 10: *Hvorfor er det viktig å avkjøle løsningen langsomt og ikke bare sette det rett på isbad?*

Svar 10:

10. Når løsningen er godt under romtemperatur og krystaller har blitt dannet vakuumpiltreres løsningen. Tørk krystallene på filter papiret 2-3 minutter med vakuumsuget på.
11. **Vei et urglass** og skrap forsiktig produktet over på urglasset fra filtrerpapiret. Vei så urglasset med produktet på og noter hvor mye produktet veier (g). Husk å trekk fra vekten av urglasset.

DEL 3 - ANALYSE AV PRODUKTET DBA

Gjør klar produktet til analyse med NMR og MS.

1. Ta en liten spatelspiss av produktet (DBA) i et prøverør som utdeles.
2. Merk prøverøret med gruppe nummer

Labansvarlig vil kjøre analyse på produktet gjennom en NMR-maskin og MS-maskin. Dere vil få deres respektive NMR/MS spekter tildelt når analysen er ferdig.

DEL 4 – RESULTATER OG DISKUSJON

Resultater

Vekt: Hvor mye veide produktet du fikk etter omkrystallisering?

Prosentutbytte: Hvor mye produkt fikk du i forhold til teoretisk utbytte.? Husk at du først må finne begrenset reaktant!

Tolkning av spektre: Legg ved en utskrift av spektrene hvor du svarer på følgende spørsmål:

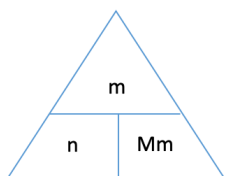
1. Tolkning av NMR: hvilke topper representerer hvilke atomer/atomgrupper?
2. Tolkning av MS: gir forslag på fragmenter som kan ha vært opphav til toppene

Diskusjon og konklusjon

1. Svar på kontrollspørsmålene i dette lab heftet.
2. Diskuter hvilke feilkilder som kan oppstått under forsøket og som kan ha påvirket utbyttet av produktet.
3. Diskuter eventuelle avvik på MS og NMR spekteret

FORMLER

Mol – masse - molarmasse



Teoretisk utbytte

For å beregne prosent utbytte må det først regnes ut *teoretisk utbytte*. Altså hvor mye produkt vi ville ha fått dersom reaksjonen var helt fullstendig og ingenting gikk tapt. For å kunne finne dette må det først avklares hva som er *begrenset reaktant*.

$$\text{Formel for teoretisk utbytte (masse): } = n_{\text{begrenset reaktant}} \times Mm_{DBA}$$

Prosentutbytte

$$\text{Formel for prosent utbytte: } \frac{\text{masse av faktisk utbytte (g)}}{\text{masse av teoretisk utbytte (g)}} \times 100 \%$$