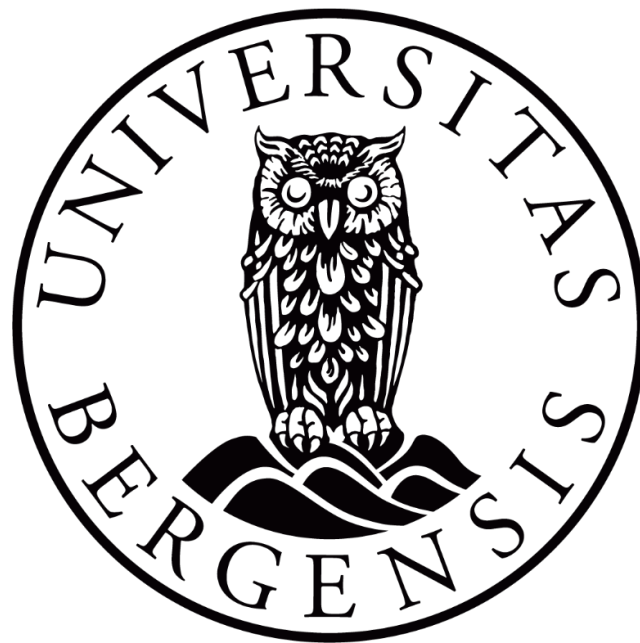


# Hvordan opplever lærere på videregående skole bruk av mappevurdering i naturfag?

En kvalitativ studie av læreres erfaringer med vurderingsformen

Thor Modalsli Sleveland



Masteroppgave i kjemididaktikk

Kjemisk institutt

Det matematisk-naturvitenskaplige fakultet

UNIVERSITETET I BERGEN

JUNI 2024

## Forord

Denne masteroppgaven symboliserer avslutningen på en femårig studieperiode som har vært preget av lærerike øyeblikk og uforglemmelige minner.

Jeg vil starte med å takke skolen, lærere og elever som har vært involvert i dette masterprosjektet. En ekstra stor takk gis til de tre lærerne som sammen bidro med verdifulle resultater i studien. Uten dere hadde det aldri blitt noe av denne oppgaven.

Videre vil jeg takke veilederen min, Matthias Stadler. Han har fulgt meg tett i prosjektet og gitt meg både jevn og god oppfølging. Matthias har vært fleksibel og rask å få tak i på e-post, og oppgaven hadde ikke blitt slik den er uten all veiledningen han har gitt.

Til slutt vil jeg gi en takk til medstudentene mine på lektorstudiet, særlig Johanne og Elise Marie som har skrevet en lignende masteroppgave. Vi har hatt felles veiledning med Matthias, og de to har gitt meg nyttige tips og triks underveis.

## Sammendrag

Denne studien tar utgangspunkt i tre naturfagslærere på en videregående skole som har brukt mappevurdering i litt over ett år. Gjennom kvalitative intervjuer har lærerne beskrevet hvordan de har gjennomført mappevurderingen og hvordan vurderingsformen har påvirket de selv og elevene.

Mappevurdering er en vurderingsform som kjennetegnes av at elevene arbeider med flere oppgaver i løpet av et skoleår og til slutt velger de som skal inngå i mappen og bli vurdert. Bruk av mappevurdering kan legge til rette for verdifulle konsepter som dybdelæring, formativ vurdering og metakognisjon. Vurderingsformen kan fremme dybdelæring fordi oppgavene er åpne og er formulert slik at elevene må trekke inn forskjellige elementer i besvarelsen sin. Tilbakemeldingene som elevene får underveis kan legge til rette for en formativ vurderingssituasjon. Elevenes aktive rolle og tilstedeværelsen av refleksjonsnotat kan fremme metakognisjon.

Mappevurdering resulterer i at lærere får et annet veilednings- og vurderingsarbeid enn med vanlige skoleprøver, noe lærerne i studien foretrekker. Deres totale arbeidsmengde har imidlertid økt, særlig fordi de legger opp til tett oppfølging med elevene. Det kan dessuten ta lenger tid å vurdere mappetekster enn skoleprøver. Elevene har også vært positive til mappevurdering. De uttrykker at det er mindre stressende og at de i større grad får vist hva de kan. Lærerne opplever at elevene har blitt mye bedre til å forstå sammenhenger på tvers av skolefaget.

Til slutt i studien drøfter jeg hvordan lærerne kan redusere arbeidet tilknyttet veiledning og vurdering samt øke elevenes utbytte. To tiltak blir diskutert: Bruk av responsgrupper og bruk av flere innleveringer. Jeg diskuterer også hvordan lærere kan møte utfordringene med språkmodeller.

# Innholdsfortegnelse

<b>Forord</b> .....	<b>i</b>
<b>Sammendrag</b> .....	<b>ii</b>
<b>Innholdsfortegnelse</b> .....	<b>iii</b>
<b>Innledning</b> .....	<b>1</b>
<b>Teori</b> .....	<b>2</b>
<i>Læring</i> .....	2
Forutsetninger for læring .....	2
Dybdelæring og overflatelæring .....	3
Kompetanse i å lære.....	4
<i>Vurdering</i> .....	5
Hvorfor vurdere elever? .....	5
Hvordan vurdere elever?.....	6
Sluttvurdering og undervisvurdering .....	6
Vurdering for læring og formativ vurdering .....	7
Egenvurdering og hverandrevurdering.....	8
Vurdering i skolen.....	9
<i>Mappevurdering</i> .....	9
Mappe som undervisningsstrategi.....	10
Mappe som vurderingsredskap .....	12
<i>Språkmodeller i skolesystemet</i> .....	14
<i>Læreres tilfredshet i skolen</i> .....	16
<i>Oppsummering</i> .....	17
<i>Forskningsspørsmål</i> .....	17
<b>Metode</b> .....	<b>19</b>
<i>Oppstart av forskningsprosjektet</i> .....	19
<i>Kvalitativ tilnærming</i> .....	20
<i>Intervju</i> .....	20
Begrunnelse for intervju.....	21
Utvalgsmetode .....	21
Intervjuguide.....	22
Gjennomføring av intervjuene .....	23
<i>Analyse av datamaterialet</i> .....	24
Transkribering .....	25
Koding.....	25
<i>Kvalitet</i> .....	27
Generalisering .....	28
Validitet .....	29
Reliabilitet .....	30
<i>Etiske betraktninger</i> .....	31
<i>Oppsummering</i> .....	32
<b>Resultat</b> .....	<b>33</b>

<i>Gjennomføring</i> .....	33
Innhold i mappene .....	34
Veiledning.....	36
Vurdering.....	37
<i>Lærernes opplevelse</i> .....	37
Arbeidsmengde.....	38
Annerledes veilednings- og vurderingsarbeid .....	41
<i>Elevstrategier</i> .....	42
Komme i gang tidlig eller sent?.....	43
Bruk av skriveøktene på skolen .....	43
Status førsteutkast.....	44
Innlevering andreutkast.....	44
<i>Elevresultater</i> .....	45
Arbeidsvaner, motivasjon og interesse .....	45
Faglig utbytte .....	45
Emosjonelt .....	47
Hva sier elevene? .....	48
<i>Andre læreres holdning</i> .....	48
<i>Oppsummering</i> .....	50
<b>Diskusjon</b> .....	<b>52</b>
<i>Hvordan påvirkes lærerens undervisningspraksis av mappevurdering?</i> .....	52
Undervisning i vanlige skoletimer .....	52
Veiledning i skriveøktene .....	54
<i>Hvordan påvirkes lærerens vurderingspraksis av mappevurdering?</i> .....	56
Elevenes arbeidsprosess og lærerens arbeidsmengde .....	56
Andre påvirkninger mappene har på lærerens vurderingspraksis .....	58
<i>Hvordan påvirkes elevene av mappevurdering?</i> .....	60
<i>Mappevurdering og språkmodeller</i> .....	63
<i>To tiltak for å redusere lærerens arbeidsmengde og øke læringsutbyttet</i> .....	67
Bruk av responsgrupper.....	67
Flere og hyppigere innleveringer .....	70
<b>Konklusjon</b> .....	<b>73</b>
<b>Referanser</b> .....	<b>74</b>
<b>Vedlegg</b> .....	<b>I</b>
<i>Vedlegg 1: Intervjuguide for Cecilie, Theo og Anita</i> .....	I
<i>Vedlegg 2: Intervjuguide for oppfølgingsintervju med Cecilie</i> .....	III

## Innledning

Skoleprøver er en mye utbredt vurderingsform i mange fag. Forskning har vist at slike prøver kan ha negative effekter på elevenes læringsprosesser og føre til overflatelæring (NOU 2014: 7). I en tid hvor samfunnet og arbeidslivet etterspør andre kompetanser enn tidligere kan det være på sin plass å justere vurderingspraksisen i norsk skole. Den nye læreplanen setter blant annet fokus på dybdelæring (Utdanningsdirektoratet, 2019a). Det vil si at elevene gradvis utvikler kunnskap og varig forståelse av begreper, metoder og sammenhenger i fag og på tvers av fag (Utdanningsdirektoratet, 2019b).

Mappevurdering er en relativt ny vurderingsform som setter læring i sentrum. Ved bruk av mappevurdering skriver elever tekster og gjør eventuelt andre arbeider i løpet av skoleåret. Elevene er med på å bestemme hvilke bidrag som skal bli en del av den endelige mappen, og de kan gjøre endringer i tekstene gjennom hele skoleåret.

Resultatet er at elevene arbeider i en kontinuerlig prosess hvor de stadig tilfører verdi til mappen. Dermed blir ikke elevene testet i hva de kan på et bestemt tidspunkt, som ved mange andre vurderinger, men de blir istedenfor vurdert basert på utviklingen de har hatt i løpet av et helt skoleår.

Bruk av mappevurdering er med andre ord en ypperlig måte å vurdere elevene på og som i tillegg støtter den nyeste læreplanen. I denne studien skal jeg undersøke hvordan tre naturfagslærere på videregående trinn opplever mappevurdering. Dette tas stilling til ved hjelp av den generelle problemstillingen:

*Hvordan opplever lærere bruk av mappevurdering i naturfag på videregående skole?*

## Teori

I denne delen av masteroppgaven legges det frem relevant forskningslitteratur for å kunne ta stilling til oppgavens problemstilling. Hovedtemaene som det settes lys på er læring, vurdering, mappevurdering, språkmodeller og læreres tilfredshet i skolen.

Læring og vurdering henger tett sammen: Kompetansemålene i læreplanen bestemmer hva elevene skal lære, og det er vurderingene som gir svar på i hvilken grad elevene har lært innholdet i læreplanen. Mappevurdering er et alternativ til andre vurderingsformer og gir læreren et ekstra verktøy i verktøykassen. Endringer i lærerens arbeidspraksis, for eksempel av å ta i bruk nye vurderingsformer, påvirker lærerens tilfredshet. Derfor blir lærertilfredshet tatt opp. Det legges også frem forskning på språkmodeller og hvilken innvirkning de har på skolens vurderingssystem.

## Læring

Ludvigsen-utvalget (NOU 2014: 7) har lagt frem mye tidligere forskning på læring. I denne første delen av teorien presenteres et utvalg av forskningen. Det blir først gjort rede for sju forutsetninger for at god læring skal finne sted. Deretter blir definisjoner på dybdelæring og overflatelæring lagt frem. Til slutt presenteres konseptet med å ha kompetanse i å lære.

### *Forutsetninger for læring*

Ludvigsen-utvalget (NOU 2014: 7) presenterer sju sentrale forutsetninger for god læring basert på tidligere forskning. Funnene er hentet fra omfattende forskningsoppsummeringer og har sterk støtte på tvers av undersøkelser (NOU 2014: 7). Forutsetningene må ivaretas på en god måte for at de skal fremme læring, men også andre forhold kan være avgjørende for læring hos elevene (NOU 2014: 7). Forutsetningene kommer til å bidra i studien ved å gi grunnlag for om mappevurdering legger til rette for læring.

Den første forutsetningen er at elevene er aktive i læringsprosessene. For å kunne utvikle kompetanse innenfor et kunnskapsområde må elevene forstå fakta, idéer og begreper og kunne knytte dette til det de kan fra før. Deretter organiseres kunnskapen og det bygges relasjoner mellom den slik at elevene kan hente kunnskapen frem og bruke den i konkrete situasjoner. Prosessen med å utvikle kompetanse krever at elevene er aktive i

læringssituasjonen (NOU 2014: 7). Den neste forutsetningen handler om at elevene må delta i kommunikasjon og samarbeid. Kunnskap produseres via kommunikasjon, og lærere bør derfor legge til rette for faglig og læringsfremmende kommunikasjon i klasserommet. Elever kan dessuten bli motiverte og engasjerte av å samarbeide med andre elever (NOU 2014: 7).

Den tredje forutsetningen for god læring er at elevene får hjelp til å forstå sammenhenger. Undervisning som stimulerer elevene til å se sammenhenger øker muligheten for at elevene kan ta i bruk den opparbeidede kompetansen i nye situasjoner (NOU 2014: 7). Ludvigsen-utvalget (NOU 2014: 7) sier imidlertid at forskning har vist at mange elever synes det kan være vanskelig å se slike sammenhenger, men at læreren kan bidra ved å for eksempel koble læringsarbeidet til virkelighetsnære problemstillinger. Dette gjør dessuten læringsprosessen mer meningsfull for elevene (NOU 2014: 7).

To andre forutsetninger for læring som Ludvigsen-utvalget (NOU 2014: 7) presenterer er at læreren må tilpasse opplæringen og at elevene må bli utfordret. Læreren må altså både ta hensyn til at elevene har individuelle forskjeller, men samtidig ta hensyn til at hver og en av elevene må få muligheten til å bli utfordret ved å jobbe med oppgaver som ligger like over deres kapasitet. En sjette forutsetning er at elevene har klare oppfatninger av hva de skal lære. De må i tillegg vite hva som forventes av dem, de må få tilbakemeldinger på arbeidet de gjør underveis og de må få råd om hvordan de kan forbedre seg for videre læring (NOU 2014: 7). Den sjuende forutsetningen som legges frem er gode relasjoner mellom lærer og elev, et trygt klassemiljø og motiverte elever (NOU 2014: 7).

Nedenfor forklares dybdelæring og overflatelæring. Dybdelæring bærer preg av flere av forutsetningene for læring beskrevet ovenfor.

### *Dybdelæring og overflatelæring*

Dybdelæring har fått sitt innpass i læreplanverket fordi dagens samfunn og arbeidsliv krever at barn og unge utvikler evne til å være reflekterte og kritiske (Utdanningsdirektoratet, 2019a). Utdanningsdirektoratet (2019b) definerer dybdelæring som "*det å gradvis utvikle kunnskap og varig forståelse av begreper, metoder og sammenhenger i fag og mellom fagområder*". Det innebærer at elevene analyserer, løser



problemer og reflekterer over egen læring, noe som forutsetter at elevene er aktive i læringsprosessen (NOU 2014: 7). Videre er det fordelaktig at elevene sprer arbeidet ut over en lengre tidsperiode fremfor å gjøre alt arbeidet i løpet av noen få arbeidsøker. Sistnevnte bidrar til å oppnå varig forståelse, som inngår i dybdelæring (Pashler, 2007).

Overflatelæring vil på den andre siden si at elever memorerer fakta og prosedyrer uten å forstå hva de gjør (NOU 2014: 7, s. 36). Elevene jobber da med nytt fagstoff uten å relatere det til tidligere kunnskap, og de har problemer med å løse oppgaver som ikke er helt like de oppgavene de allerede har gjort (NOU 2014: 7). Taasen et al. (2004) påstår at lærebøker legger opp til overflatelæring i tilfeller der spørsmålene i slutten av hvert kapittel ber elevene gjengi informasjon.

Nedenfor beskrives hvordan elever kan opparbeide seg kompetanse i å lære. Flere av tiltakene som legges frem for å få slik kompetanse er tilstedeværende ved mappevurdering, som det kommes tilbake til senere.

### *Kompetanse i å lære*

Skolen er den primære læringsarenaen for barn og unge. Mye av det som læres i skolen, vil være viktig for personer å kjenne til også lenge etter skolegangen tar slutt. Videre forventes det at skolen tar ansvar for å gi elevene verktøy som de kan bruke til å lære senere i livet, såkalt livslang læring (NOU 2014: 7). I løpet av skoleårene skal altså elevene lære seg å lære.

Innenfor forskning på læring er metakognisjon, selvregulert læring og læringsstrategier begreper som brukes til å forklare kompetanse i å lære (NOU 2014: 7). Læringsstrategier foregår underveis i læringsprosessen og kan for eksempel være at eleven tar et skritt tilbake fra en matematikkoppgave og tenker over om det er bedre og mer effektive måter å løse oppgaven på (NOU 2014: 7). Både metakognisjon og selvregulert læring viser til elevens evne til å reflektere over – og aktivt forsøke å kontrollere og påvirke – egen læring (NOU 2014: 7). Elever med svake skoleprestasjoner har gjerne svakere tro på mestring og sliter dermed mer enn andre elever med å komme i gang å planlegge, strukturere og ferdigstille en oppgave (NOU 2014: 7). Videre er selvregulering knyttet til elevenes evne til å fortsette arbeidet i tilfeller hvor de møter på utfordringer og står fast (NOU 2014: 7).

Ludvigsen-utvalget (2014: 7) legger frem viktige tiltak lærere kan forholde seg til for at undervisningen skal stimulere til kompetanse i å lære. Et tiltak er å la elevene være med i vurderingsarbeidet, både i form av egenvurdering og hverandrevurdering. Egenvurdering og hverandrevurdering blir omtalt senere. Et annet tiltak er å la elevene delta i å utarbeide kriterier for hva som kjennetegner et godt arbeid. Tiltakene resulterer i at elevene får en mer aktiv rolle i klasserommet. Læreren får på den andre siden en støttende og veiledende rolle, men det blir likevel behov for struktur og klasseledelse (NOU 2014: 7).

Nedenfor blir vurderingsaspektet ved skolen presentert. Dybdelæring har fått sin plass i den nye læreplanen, og som profesjonsutøver bør lærere holde seg oppdatert på hvordan de kan planlegge undervisning og vurdering som settes dybdelæring i fokus.

## Vurdering

Vurdering er en viktig del i skolen og opplæringen. Nedenfor blir relevante begreper som underveisvurdering, formativ vurdering, egenvurdering og hverandrevurdering forklart. Det gis også noen korte svar på hvorfor og hvordan elever bør vurderes. Trådene samles mot slutten når vurderingspraksis, hovedsakelig i norsk skole, beskrives.

### *Hvorfor vurdere elever?*

Elever har av forskrift til opplæringsloven (2006, § 3-2) rett til underveisvurdering, sluttvurdering og dokumentasjon av opplæringen. Formålet med vurderingen er å fremme læring, bidra til lærelyst underveis og gi informasjon til elevene – både underveis og ved avslutning av opplæringen i faget – om deres kompetanse (Forskrift av opplæringsloven, 2006, § 3-3). Vurderinger gir dessuten essensiell informasjon til læreren. Vurderingene kan nemlig brukes til å kontrollere om læringsaktivitetene læreren har gjennomført faktisk har medført læring (William, 2011). Læreren kan dermed bruke vurderingene til å endre på og forbedre undervisningsopplegget for de kommende skoleårene.

Videre gir vurderinger informasjon til skolen og samfunnet om i hvilken grad eleven har opparbeidet seg kompetanse i de ulike fagene eleven har hatt. Sluttvurderinger, som jeg kommer tilbake til senere, gir slik informasjon. Karakterene elevene får er basert på hvor godt de har prestert på vurderingene, og de brukes til å søke skoleplass på videregående

og eventuelt høyere utdanning. Dette avgjør i neste omgang hvilke karrieremuligheter elevene får når de blir eldre. Vurderinger i skolen er derfor viktige for å gi elevene et bevis på deres kompetanse.

### *Hvordan vurdere elever?*

Som en profesjonell yrkesutøver forventes det at en lærer har kjennskap til hvordan god vurderingspraksis kan oppnås (Hopfenbeck, 2016). Det gir elevene mulighet til å oppleve de beste vurderingssituasjonene. Forskningen til Hopfenbeck (2016) viser blant annet at elever rundt omkring i verden har lignende preferanser angående vurderingssituasjonen: De foretrekker forutsigbarhet og at vurderingen oppfattes som rettferdig.

Videre forteller Hopfenbeck (2016) at god undervisningspraksis innebærer å vite hva man bør unngå grunnet at feilslått vurderingspraksis kan medføre angst og tap av mestringsopplevelse hos elevene. Læreren risikerer dessuten at elevene mister sin indre motivasjon, som man nettopp vil at elevene skal ha, hvis elevene opplever at vurdering blir brukt i form av kontroll eller sanksjoner (Hopfenbeck, 2016). Bruk av vurderinger som stresser elever, som er av lav kvalitet eller som blir gitt i feil sammenhenger er eksempler på feilslått vurderingspraksis (Hopfenbeck, 2016).

Nedenfor forklares sentrale begreper innenfor vurdering: Sluttvurdering, underveisvurdering, vurdering for læring, formativ vurdering, egenvurdering og hverandrevurdering. Læreren bør tenke på alle seks begrepene når en vurdering står for tur.

### *Sluttvurdering og underveisvurdering*

I skolen må lærere velge mellom ulike typer vurderinger. Som påpekt tidligere, følger det av opplæringslovens § 3-2 (2006) at elever har rett på både sluttvurdering og underveisvurdering i alle fagene de har.

Jeg bruker forskriften til opplæringsloven § 3-10 og § 3-14 til å definere begrepene. I § 3-14 står det at formålet med sluttvurderingen er å gi informasjon om kompetansen til en elev ved avslutningen av opplæringen i et fag. På studiespesialiserende på videregående skoler omfatter dette standpunkt karakterer og eksamens karakterer (Forskrift til opplæringsloven, 2006, § 3-14). Ettersom sluttvurderingen skal teste hva elevene

innehar av kompetanse, er det en form for vurdering av læring (Michaelson & Johansen, 2007).

Underveisvurdering omfatter all vurdering som skjer før avslutning av opplæringen (Forskrift til opplæringsloven, 2006, § 3-10). Forskriften har tydelige krav til underveisvurderingen. Elevene skal nemlig delta i vurderingsarbeidet og reflektere over egen læring, de skal forstå hva som forventes av de, de skal få beskjed om hva de mestrer og de skal få veiledning om hvordan de kan videreutvikle kompetansen sin (Forskrift til opplæringsloven, 2006, § 3-10). Vi kan si at formålet med sluttvurdering er å bedømme kompetanse, mens formålet med underveisvurdering er å bygge kompetanse.

### *Vurdering for læring og formativ vurdering*

Ovenfor ble de to typene vurdering kjapt beskrevet. Når hensikten med en vurdering er å fremme elevens læring, kan vi kalle det vurdering for læring (William, 2011).

Underveisvurdering, men ikke sluttvurdering, går dermed under konseptet vurdering for læring. William (2011) skiller videre mellom vurdering for læring og formativ vurdering. Han mener en formativ vurdering kun er de vurderingssituasjonene som faktisk fører til læring. En vurdering som har som hensikt å fremme læring, men i realiteten ikke gjør det, vil i så fall ikke være en formativ vurdering.

William (2011) bruker tilbakemelding til å forklare skillet mellom begrepene, og sier at formative vurderinger gir informasjon i form av tilbakemeldinger som eleven og læreren kan bruke for videre læring. De beste tilbakemeldingene er de som gir eleven informasjon om hva som var riktig svar, de som har med noen forklaringer og de som har spesifikke punkter eleven kan følge for å forbedre seg (William, 2011). Distinksjonen mellom vurdering for læring og formativ vurdering, slik William presenterer det, er interessant. Den får frem viktigheten av at tilbakemeldingene elevene får på en vurdering faktisk fører til videre læring, og at det altså ikke er en selvfølge at elevene lærer av å få tilbakemeldinger.

Nedenfor beskrives egenvurdering og hverandrevurdering. De innebærer at elevene er aktive i vurderingsarbeidet og trekkes frem av Black og William (2009) som én av fem kjernestrategier i formativ vurdering.

## *Egenvurdering og hverandrevurdering*

Elever kan vurdere eget arbeid, kalt egenvurdering, men de kan også vurdere andre sitt arbeid, kalt hverandrevurdering (Utdanningsdirektoratet, 2022b). Utdanningsdirektoratet (2022b) legger frem et forslag på hverandrevurdering. Det går på å la elevene finne to "stjerner" og ett "ønske" i hverandres tekster eller besvarelser. Elevene får da beskjed om noe som er bra og noe som kunne vært bedre med teksten deres. Læreren kan også la elevene sammenligne og rangere et par tekster som læreren deler ut (Utdanningsdirektoratet, 2022b). Da kan elevene oppdage hva som skiller to tekster fra hverandre i kvalitet, noe de kan bruke når de skriver sine egne tekster.

Elever som involveres i vurderingsarbeidet lærer bedre og blir mer motiverte (Utdanningsdirektoratet, 2022b). Som beskrevet tidligere, er et av målene med skolegangen at elevene skal lære seg å lære (NOU 2014: 7). Læreren kan legge til rette for slik læring med å inkludere elevene i vurderingsarbeidet, da dette gir elevene trening i å vurdere eget arbeid og reflektere over egen læring (Utdanningsdirektoratet, 2022b). Dette skyldes at elevene da blir mer bevisst på hvor de står i læringsprosessen, hvor de skal og hvordan de på best mulig vis kan komme seg dit (Utdanningsdirektoratet, 2022b). Videre viser studier at hverandrevurdering kan forbedre elevers skriveferdigheter (Cho & Cho, 2011; Kurihara, 2017). Studier viser også at det kan være mer lærerikt å gi kommentarer enn å få kommentarer (Baker & Lundstrom, 2009; Cho & Cho, 2011). Basert på sistnevnte funn bør lærere vurdere å bruke hverandrevurdering fremfor at læreren selv gir tilbakemelding på vurderingene.

Selv om hverandrevurdering kan fremme læring, forutsetter metoden et trygt klasse miljø og at elevene har tydelige vurderingskriterier som de kan ta utgangspunkt i når de vurderer hverandres arbeid (Utdanningsdirektoratet, 2022b). Videre sier Brøyn (2022) at elevene må ha kompetanse i å vurdere eget arbeid for at egenvurdering skal ha en positiv effekt (Brøyn, 2022). Hun nevner ikke hverandrevurdering, men det er tenkelig at tilsvarende vil være gjeldende hvis elevene skal vurdere hverandres arbeid. Brøyn (2022) påpeker at det er læreren sin oppgave å lære elevene å vurdere eget arbeid.

Nedenfor gis et raskt blikk på vurdering i skolen, og i den sammenheng legges det frem noen forskningsbaserte anbefalinger.

## Vurdering i skolen

Det er ikke en ny tanke at vurdering kan fremme læring (Wiliam, 2011). Tidligere ble vurdering i all hovedsak brukt til å måle elevenes faglige oppnåelse, men på slutten av 1980-tallet kom idéen om læringsstøttende vurderingsformer (Wiliam, 2011). Dette resulterte i mye forskning i 1990-årene (Wiliam, 2011). Wiliam (2011) konkluderer med at det fremdeles kreves forskning på feltet, men at forskning og empiri så langt tyder på at elever har store fordeler av at vurderingen integreres i lærerens undervisningspraksis.

Det nye vurderingssynet har også funnet veien til Norge (Hopfenbeck, 2016; NOU 2015: 8). I 2015 kom Ludvigsen-utvalget blant annet med følgende anbefalinger (NOU 2015: 8):

1. Lærere bør legge til rette for elevenes dybdelæring, progresjon og aktive rolle i undervisningen
2. Lærere bør i samarbeid videreutvikle metoder for vurdering
3. Elevene bør i større grad involveres i vurderingsarbeidet

Utvalget forklarer viktigheten av at elevene inntar en aktiv rolle tilknyttet arbeidet etter en vurdering (NOU 2015: 8, s. 81):

Å få vurdere eget arbeid og egen fremgang og reflektere over læringsarbeidet sammen med medelever og lærere, kan bidra til at elevene utvikler et bevisst forhold til sin egen læring og lærer å ta en aktiv rolle i undervisningen. Dette er relatert til elevenes utvikling av metakognisjon og selvregulert læring.

Ifelebuegu (2023) foreslår at vurderingssystemet – globalt sett – bør endres fra at vurderinger inneholder gjenkallingsspørsmål (eng: recall questions), som de ofte gjorde før, til at spørsmålene istedenfor fokuserer på de øvre delene i Blooms taksonomi:

*Applying, analyzing, evaluating og creating*. Begrunnelsen er blant annet at slike spørsmål etterspør kompetanser som elevene trenger i dagens arbeidsliv.

Mappevurdering, som beskrives nedenfor, er blant forslagene til Ifelebuegu (2023) for å integrere de øvre delene av Blooms taksonomi i vurderingene.

## Mappevurdering

En mappe er en systematisk samling av bidrag som viser en elev sin innsats, fremskritt og prestasjoner (Lam, 2018; Taube, 2000). Taube (2000) mener denne definisjonen ikke er utfyllende nok, og inkluderer at mappen må omfatte elevmedvirkning med tanke på

valg av innhold, utvalgsriterier og vurderingsriterier, samt at den viser elevens selvrefleksjoner. Eleven skal altså kunne være med å bestemme hva mappen skal bestå av, men forfatteren sier at læreren ofte hjelper til i denne prosessen (Taube, 2000).

Mappen kan introduseres i forskjellige fag, men i hovedsak brukes den innenfor språkfag, matematikk, naturfag og samfunnsfag (Lam, 2018; Taube 2000). Ettersom det er lite forskning på mappevurdering i naturfag, særlig i Norge, kommer det til å refereres til en del forskning om mappevurdering i andre fagområder, særlig tilknyttet skiving og lesing.

Mapper har to hovedbruksområder. Den kan brukes som et redskap i skoletimene til å fremme læring, men den kan også brukes som grunnlag for vurdering (Taasen et al., 2004; Taube, 2000). Nedenfor følger en utredelse av de to bruksområdene.

### *Mappe som undervisningsstrategi*

Som undervisningsstrategi har mappevurdering mange kjennetegn som ofte trekkes frem som læringsfremmende (Taube, 2000): elevaktivitet, delaktighet, ansvar for egen læring og metakognitiv utvikling. Metakognitiv utvikling oppnås ved at elevene får erfaring i å vurdere og reflektere over sitt eget arbeid (Michaelsen & Johansen, 2007). I tillegg trekker Michaelsen og Johansen (2007) frem tilpasset opplæring, samarbeid og utvikling av hensiktsmessige arbeidsvaner som grunner til å bruke mapper i undervisningen. En annen fordel med mapper er konseptet "*writing to learn*". Prinsippet bygger på, som det fremkommer av navnet, at elever lærer av å skrive tekster. Rivard (1994) forteller at *writing to learn* kan brukes i naturfag og kan fremme naturfaglig læring dersom skriveoppgavene er godt tilknyttet læringsmålene, dersom elevene har nok metakognitiv kompetanse til å skrive oppgavene og dersom oppgavene bygger på konseptuell forståelse.

Læreren får en endret rolle i klasserommet når mappevurdering tas i bruk (Taube, 2000). Elevene er mer aktive i læringsprosessen og har mer påvirkning over – og ansvar for – egen læring. Dette resulterer i at læreren blir en medhjelper som går rundt og veileder elevene (Taube, 2000). Taube (2000) forteller at lærere kan oppleve denne rollefordelingen som utfordrende, og at dette kan ta noe tid å venne seg til. Taube (2000) foreslår å ha classesamtaler med elevene ved oppstart av mappeprosjektet for å snakke om endringene det kommer til å medføre, som for eksempel den nye rollefordelingen.

Når elevene får være med å bestemme hva de skal skrive om, og dermed kan skrive om noe som virkelig engasjerer dem, fører det til at elevene ofte har lyst til å jobbe med og forbedre mappetekstene (Taube, 2000). På den måten gjør mappearbeidet elevene mer motiverte og engasjerte i læringsprosessene (Taube, 2000). Når derimot andre personer bestemmer hva elevene skal gjøre, blir innsatsen overflatisk (Taube, 2000). Derfor mener Taube (2000) det blant annet er avgjørende at elevene får være med å ta ansvar i læringssituasjonen, for eksempel med å utforme vurderingskriterier.

Taasen et al. (2004) hevder at bruk av mapper forutsetter et læringsmiljø med innebygde mekanismer for utfordring og støtte, både fra læreren og medelever. Taube (2000) beskriver en slik mekanisme, nemlig bruk av responsgrupper. I responsgrupper – som Taube (2000) legger frem som et kjennetegn ved mappevurdering – skal elevene undersøke, diskutere og vurdere hverandres arbeider. Responsgruppene kan være store eller små, og formålet er å få til et elevsamarbeid der elevene kan lære av hverandre (Taube, 2000). Denne måten å lære på har sin opprinnelse i Vygotsky (1986) sitt sosiokulturelle læringsyn hvor læring skjer gjennom bruk av språk og deltakelse i en sosial kontekst. En studie av studenter viser for øvrig at å bruke medstudenter som et verktøy for læring ved bruk av mappevurdering var betydelig mer effektivt enn å øke antall arbeidstimer og å redusere klassestørrelsen (Taasen et al., 2004).

Taube (2000) legger frem fire eksempler på skoler hvor mappevurdering har blitt prøvd ut. Det første eksempelet er fra en barneskole i Sverige hvor mappevurdering hadde blitt brukt i både språkfag, matematikk, naturfag og samfunnsfag. Læreren beskrev bruken av mappevurdering som svært vellykket og at "*mappevurdering er en utmerket måte å lære barn å lære på*" (Taube, 2000, s. 68). Læreren oppdaget en økt ansvarsfølelse hos elevene og at elevene tok læringen mer alvorlig. Dessuten merket læreren at elever med lese- og skrivevansker ble bedre til å skrive.

Det andre eksempelet er fra en 2. klasse i Finland (Taube, 2000). Der hadde læreren innført "*ukas arbeid til mappen*" for å hjelpe elevene med å komme i gang med egenvurderingen. Læreren fortalte at mappevurderingen hadde gitt henne en helt ny innsikt i hvordan elevene tenkte og lærte, noe som hadde gjort henne mer bevisst på hvordan hun selv kunne hjelpe elevene. Læreren påpekte at hun mente det er viktig at



skoler våger å stole på elevene og gi de mer ansvar, og at læreren da kan støtte elevene fremfor å styre dem.

Neste eksempel er fra en 6. klasse i Finland, på samme skole som klassen beskrevet ovenfor (Taube, 2000). I klassen ble mappevurdering brukt til å utvikle skrivekompetanse. Læreren la frem at et viktig innslag var bruk av samtaler mellom elev-elev og elev-lærer. Når elevene hadde skrevet ferdig et utkast, skulle de nemlig gå sammen i par og lese opp utdragene for hverandre, la medeleven gjenfortelle og til slutt diskutere om noe kunne utdypes eller forbedres. Læreren forteller at samtaler mellom elevene førte til at de lærte av hverandre, at de begynte å ta ansvar for egen læring og at enkelte fikk inspirasjon og skriveglede.

Det siste eksempelet er fra en ungdomsskole i Ski hvor mappevurdering ble brukt i norsk (Taube, 2000). Mappevurderingen var en sentral del av skoletimene, og hver uke ble mange skoletimer brukt til å skrive mappetekster. Læreren brukte skoletimene til å gå rundt og snakke med hver elev. Hun forteller at samtaler reduserte arbeidsmengden tilknyttet vurdering siden hun da hadde fått god innsikt i innholdet i tekstene. Læreren brukte responsgrupper. Enkelte av elevene likte ikke responsgruppene og ville helst ha lærerrespons. Andre elever tok selv initiativ til å gå i responsgrupper ved behov. Læreren fikk inntrykk av at elevene var positive til mappevurderingen, og mange av elevene uttrykte at de likte å se tilbake på alt de hadde fått gjort. Klassemiljøet hadde blitt betraktelig bedre i løpet av året, og læreren tror mappevurdering kan ha bidratt til dette fordi elevene lærer å samarbeide og dermed respektere hverandre.

### *Mappe som vurderingsredskap*

Mapper kan også brukes til vurdering. Michaelsen og Johansen (2007) forteller at en fordel med å bruke mapper til vurdering er at den viser et større spekter av elevens kunnskaper og ferdigheter enn vanlige skoleprøver. Dette kan skyldes den åpne formen i oppgavetekstene og at det er elevene selv som setter rammen for hvor dypt de vil dykke ned i oppgavene. Det skyldes også at elevene arbeider i en prosess og at mappevurderingen dermed viser utviklingen elevene har hatt fra ett tidspunkt til et annet (Taube, 2000). Dette er i motsetning til tradisjonelle vurderinger som typisk blir brukt til å teste hva elevene kan på et bestemt tidspunkt (Taube, 2000). Taasen et al. (2004)

forklarer at hensikten med en vurdering skiftes ved bruk av mappevurdering, fra at det før var fokus på testing og karaktersetting til at det nå blir fokus på læring.

Både mappevurdering og vanlige skoleprøver har som hensikt å motivere elevene til å arbeide jevnt gjennom skoleåret (Taasen et al., 2004). En studie fra universitetsnivå viser likevel at studentene arbeidet mer jevnt og trutt da de hadde mappevurdering enn de ellers hadde gjort (Wittek, 2003). Med mappevurdering legges det opp til at elevene skal arbeide videre med oppgavene, for eksempel i form av refleksjonsnotater eller av at elevene reviderer på tekstene etter hvert som de lærer mer fagstoff i senere kapitler. (Taube, 2000). Ved bruk av vanlige skoleprøver er elevene derimot ferdige med lærestoffet etter de har levert inn prøvene.

Et annet kjennetegn ved mapper som vurderingsform er at det er mindre stressende for elevene (Taube, 2000). Videre viser studien til Wittek (2003) at studentene i større grad følte at de fikk vist hva de kunne med mappevurdering enn om de hadde hatt en avsluttende eksamen. Studentene fikk dessuten utviklet refleksjonsevner og skriveferdigheter via arbeid med mappen (Wittek, 2003). Selv om studentene synes det var mye arbeid, ønsket de at mapper skulle utgjøre mesteparten eller alt av arbeidet fremover på studiet (Wittek, 2003).

Taube (2000) trekker det frem som en fordel at mappevurdering tester et relativt bredt kompetansespekter og derav er en vurderingsmetode som gir valide målinger.

Forfatteren hevder en svakhet kan være at to lærere ikke nødvendigvis ville gitt lik bedømmelse av innholdet i en elevs mappe, og at påliteligheten dermed er usikker (Taube, 2000). Taube (2000) foreslår riktignok å la elevene være involvert i prosessen med å utarbeide vurderingskriterier. Involveringen gir elevene mulighet til å endre eller bestemme vurderingskriteriene og kan dermed resultere i at to lærere vurderer like mappetekster ulikt.

Det faktum at elevene er med på å forme vurderingskriteriene og at dette diskuteres i klasserommet med læreren – for eksempel ved at det diskuteres hva som kjennetegner gode tekster – gjør målene med mappetekstene og veien dit mer synlig for elevene (Taube, 2000). I tillegg analyserer og vurderer elevene mappetekstene, noe som gjør veien enda synligere. På denne måten blir det en tydelig sammenheng mellom undervisningen og vurderingen i klasser som bruker mappevurdering (Taube, 2000).

Michaelsen og Johansen (2007) tar opp spørsmålet om hvorvidt mappevurdering tar for mye av lærerens tid. De sier at mangel på tid ofte er en hindring til å gjøre endringer i skolen (Michaelsen & Johansen, 2007). I så fall er det en ulempe hvis oppstart av mappevurdering krever mer tid enn læreren har tilgjengelig. Michaelsen og Johansen (2007) mener derimot at mappevurdering bør være en integrert del av det daglige arbeidet på skolen, og at det derfor ikke er noe ekstra arbeid som læreren må finne tid til. Likevel viser for eksempel studien til Wittek (2007) at arbeidsmengden var stor for studentveilederen. Blant årsakene til den økte arbeidsmengden var kontinuerlig oppfølging med studentene – hvor studentene forventet veiledning og engasjement – og tillaging av mappeoppgaver (Wittek, 2007).

Det finnes ikke én korrekt måte å gjennomføre mappevurdering på, hverken som undervisningsstrategi eller som vurderingsredskap (Taube, 2000). Videre behøver ikke mappevurdering å være den eneste vurderingsmetoden (Taube, 2000). Mappevurdering tester ulike kompetanser enn en del andre vurderinger, så det kan hende at en god kombinasjon er det som fungerer best.

Mappevurdering innebærer at elevene skriver tekster i løpet av skoleåret. ChatGPT er en relativt ny språkmodell som kan generere lange og gode tekster på få sekunder. Nedenfor presenteres forskning på hvordan språkmodellene har hatt innflytelse i skolen og hvordan lærere kan ta hensyn til kunstig intelligens.

## Språkmodeller i skolesystemet

Språkmodeller er dataprogram som kan holde en samtale med et menneske (Rudolph et al., 2023). De kan også besvare en rekke spørsmål. Språkmodellene kan fortelle meg Serbias tre største byer, men de kan også gi meg et treningsopplegg for hvordan jeg kan legge opp treningen til å løpe et maraton. ChatGPT er en språkmodell som har fått stor innflytelse fordi den er gratis å bruke. Den er dessuten en av de mer avanserte språkmodellene som finnes (Rudolph et al., 2023).

I skolesystemet kan språkmodeller som ChatGPT oppfattes både som en trussel og som et verktøy (Ifelebuegu, 2023; Rudolph et al., 2023). Noen fordeler med språkmodellene er at de gir umiddelbar respons, de kan justere svarene sine og de kan gi ytterligere informasjon – for eksempel gi praktiske eksempler – basert på spørsmålene de stilles

(Chen et al., 2022; Ifelebuegu, 2023). I en studie med studenter kom det frem at en annen fordel er at man kan stille språkmodellene "dumme" spørsmål som man helst ikke vil at læreren skal vite om (Chen et al., 2022). En mer konkret fordel er at språkmodellene kan gi elever idéer til hva de kan skrive om (Rudolph et al., 2023).

Det finnes også potensielle trusler vedrørende språkmodeller i skolen. De kan besvare avkrysningsspørsmål, de kan besvare kortsvar-oppgaver og de kan generere lengre tekstsvaer. Elever kan dermed få ferdige, utfyllende svar. Språkmodellene blir dessuten bedre og bedre for hver uke som går, og Rudolph et al. (2023) sier at forskning påstår at det kan være vanskelig å se forskjell på en tekst skrevet av en elev og en tekst skrevet av ChatGPT. Det er mulig å overvåke bruk av kunstig intelligens i klasserommet, men ikke hjemme på elevens soverom. Den enkle tilgangen til språkmodeller medfører et behov for en omveltning i skolens vurderingspraksis (Ifelebuegu, 2023).

Språkmodellenes kontinuerlig utvikling stiller spørsmål til om det fremdeles er mulig å ha innleveringsoppgaver ettersom lærere ikke med sikkerhet kan vite hvem som har skrevet oppgaven; er det eleven deres, eller er det en datamaskin? Usikkerheten var også relevant før fordi foreldre eller personer på nettforum kunne ha hjulpet med innleveringsoppgaver. Problematikken er likevel på sitt høyeste i dag ettersom det er mye enklere for en elev å spørre ChatGPT om skrivehjelp enn å spørre mor eller far.

Ifelebuegu (2023, s. 3) legger frem noen anbefalinger tilknyttet språkmodeller i skolen basert på annen forskning. Elevene bør få opplæring i fordelene og ulempene med språkmodellene slik at modellene kan brukes til å fremme læring (Ifelebuegu, 2023). Videre bør skolen lage retningslinjer for hva språkmodellene kan brukes til og hva de ikke kan brukes til (Ifelebuegu, 2023). En annen anbefaling er å endre på vurderingssystemet slik at det i bedre grad tester kompetanser som kreves i dagens samfunn. Spørsmål som krever kompleks anvendelse er gunstige for elevens fremtid samtidig som de vanskeliggjør muligheten for at språkmodeller kan besvare oppgavene (Ifelebuegu, 2023). Et eksempel som legges frem er å formulere oppgavetekster hvor elevene skal ha et kritisk blikk ettersom språkmodellene foreløpig ikke er i stand til å generere slike svar av god kvalitet (Ifelebuegu, 2023).

Ifelebuegu (2023) foreslår tre vurderingspraksiser som kan være gunstige for å etterspørre slike kompetanser. De tre praksisene er kontekstuelle problembaserte

vurderinger, gruppebaserte vurderinger og mappevurdering. Selv om mappevurdering resulterer i at elevene *kan* bruke språkmodeller, argumenterer Ifelebuegu (2023) for at mappevurdering har unike kjennetegn som gjør det vanskelig – eller umulig – å bruke språkmodeller til tekstgenerering. Argumentene innebærer at mappene er svært individualiserte, at de bygger på elevenes personlige liv og at oppgavene er komplekse. Nedenfor legges det frem teorier og forskning som kan beskrive læreres tilfredshet på skolen.

## Læreres tilfredshet i skolen

Forskning innenfor mappevurdering viser at det er en tidkrevende vurderingsform (Witteck, 2007). Samtidig viser undersøkelser at lærere allerede har mye arbeid å ta hånd om i løpet av arbeidsdagen (Ellingsen & Johnsgård, 2024; Mejlbo, 2024a). Elevene kan få større læringsutbytte av arbeid med mappene, men kan også lærerne få noe til gjengjeld?

I Herzbergs motivasjonsteori utviklet Frederick Herzberg (1985) faktorer – motivasjonsfaktorer og hygienefaktorer – som kan knyttes til motivasjon, tilfredshet og misnøye hos ansatte (Waage, 1985, s. 79). Motivasjonsfaktorene omfatter a) tilfredsstillelse ved å fullføre et arbeid, løse problemer og se resultater, b) anerkjennelse for vel utført arbeid, c) arbeid og aktivitet i seg selv, d) ansvar, e) forfremmelse og f) vekst (Waage, 1985). Noen av faktorene fremkom i TALIS-undersøkelsen fra 2018 der unge lærere i OECD-land ble spurt om hva som motiverte de til å bli lærere (OECD, 2019). De mest populære svaralternativene var at lærere har mulighet til å påvirke utviklingen til barn og unge og at lærere kan bidra i samfunnet. Rundt 70 % av lærerne svarte at sikkerhet tilknyttet jobbmuligheter og stabil inntekt gjorde læreryrket fristende.

Selv om en jobb som lærer kan virke fristende, ser vi stadig nyhetsoppslag om lærere som mistrives og som er på utkikk etter andre jobber. Blant årsakene til at lærere prøver lykken innenfor andre yrker er lange kvelder på jobb og mye helgearbeid (Jelstad, 2023; Molvær, 2023). Årsakene omhandler lærerens arbeidsforhold og utgjør én av hygienefaktorene til Herzberg, som kan brukes til å forklare misnøye på arbeidsplassen (Waage, 1985, s. 79). Videre fremkommer det i en ny undersøkelse fra Utdanningsforbundet at åtte av ti lærere oppgir at de har flere arbeidsoppgaver nå enn

for tre år siden (Mejlbo, 2024a). I samme undersøkelse sier sju av ti lærere at de har fått mer ansvar og ni av ti at de utfører oppgaver utenfor deres ansvarsområde (Ellingsen & Johnsgård, 2024; Mejlbo, 2024a). En undersøkelse utført av EU (2013) viser lignende resultater. I undersøkelsen er ekstraarbeid som ikke er tilknyttet læringsprosessen én av de fem hovedfaktorene som blir trukket frem som mest negativt med lærerprofesjonen (EU, 2013).

Lederen i Utdanningsforbundet frykter at den lange rekken med arbeidsoppgaver medfører at enda flere slutter i jobben som lærer (Mejlbo, 2024b). Nå er nemlig læreryrket mer dynamisk, mer utfordrende og mer krevende enn det noen gang har vært (OECD, 2018). Det forventes at lærere kontinuerlig utvikler og tilpasser sine ferdigheter for å gi elevene kompetansen de trenger for å kunne lykkes videre i livet (OECD, 2018). Videre forventes det at mangfoldet av elever i et klasserom skal få tilrettelagt opplæring, samt at læreren skal samarbeide med en rekke aktører for å sikre elevenes sosiale, akademiske og emosjonelle utvikling (OECD, 2018).

## Oppsummering

Mappevurdering er en vurderingsform som setter fokus på elevenes læring, og mange av de forskningsbaserte læringsforutsetningene som Ludvigsen-utvalget (NOU 2014: 7) presenterer er tilstedeværende ved bruk av mappevurdering. Elevene arbeider med mappeoppgaver over en lang tidshorisont og får veiledning fra læreren – og gjerne medelever – underveis. Videre legger mappene til rette for at elevene må reflektere og være aktive i læringsprosessen. Forskning viser at mappevurdering er tidkrevende for læreren, men samtidig kan lærere få motivasjon og glede av å se elevene lykkes.

## Forskningsspørsmål

Basert på forskningslitteraturen ovenfor kan problemstillingen konkretiseres i følgende forskningsspørsmål:

- Hvordan opplever lærere bruk av mappevurdering som vurderingsmetode i naturfag?
- Hvordan arbeider elevene med mappeoppgaver sett fra lærerens perspektiv?
- Hvordan opplever lærere at mappevurdering påvirker elevenes utbytte?

Det første forskningsspørsmålet omfatter både hvordan lærerne i studien opplever mappevurdering og hvilket syn de mener andre lærere på skolen har til vurderingsformen. Det andre forskningsspørsmålet belyser forskjellige strategier elevene følger når de arbeider med mappeoppgavene. Det tredje forskningsspørsmålet tar opp hvordan elevstrategiene påvirker deres utbytte av mappevurdering. I studien besvares alle forskningsspørsmålene fra et lærerperspektiv. Johanne og Elise Marie (2024) undersøker mappevurdering på samme skole, men fra et elevperspektiv. Deres resultater kommer til å bli oppsummert i resultatdelen. Selv om det typisk er noe overlapp mellom lærer- og elevperspektiver, vil det være nyttig å kunne sammenligne funnene.

## Metode

I denne delen av masteroppgaven skal jeg beskrive og begrunne valgene jeg har gjort underveis – fra oppstart av prosjektet og frem til strukturering av resultatet. Aller først skal jeg gi en beskrivelse av hvordan forskningsprosjektet ble til. Deretter går jeg i detalj på hvordan jeg gjennomførte de kvalitative intervjuene og hvordan jeg analyserte datamaterialet. Til slutt vurderer jeg kvaliteten på forskningen og drøfter etiske aspekter.

### Oppstart av forskningsprosjektet

De tre lærerne i studien, heretter kalt Anita, Cecilie og Theo, jobber ved en videregående skole. De startet opp med mappevurdering i naturfag skoleåret 2022/2023. Lærerne ville prøve noe annet enn vanlige to-timers skoleprøver fordi de mente at slike vurderinger ikke fører til like mye læring som andre alternativer. Etter å ha lest og hørt om mappevurdering vekket dette interesse hos lærerne, og valget på ny vurderingsform endte derfor der. De tre lærerne la en slagplan for hvordan de ville gjennomføre mappevurderingen det inneværende skoleåret og fulgte planen uten direkte påvirkning fra forskning.

Det var først godt ut på våren 2023 at Cecilie tok kontakt med masterveilederen min, en kjemididaktiker på universitetet. Anita, Cecilie og Theo er nemlig kjent med veilederen gjennom tidligere samarbeid. Cecilie sa at det hadde vært interessant å forske på bruk av mappevurdering i naturfag, men at de ikke hadde tid til å gjøre det selv. Hun foreslo derfor at mappevurderingen de hadde brukt kunne være et utgangspunkt i en masteroppgave. Dette forslaget fikk gjennomslag, og i den sammenheng valgte jeg å fokusere på lærernes perspektiv tilknyttet mappevurderingen. To andre lektorstudenter, Johanne og Elise Marie (2024), er med i samme prosjekt og skriver en masteroppgave sammen. De har fokus på elevene, altså elevenes erfaringer og opplevelser med mappevurdering.

Ettersom Anita, Cecilie og Theo kun har gjennomført mappevurdering i ett helt skoleår og masterprosjektet skal stå ferdig innen sommeren 2024, var det begrenset hvor mye datamateriale det var mulig å få tilgang til. Av den grunn er datamaterialet i hovedsak basert på skoleåret 2022/23. Jeg hadde for øvrig et oppfølgingsintervju med Cecilie på våren 2024, og da fikk jeg en oppdatering på mappevurderingen i det aktuelle skoleåret.



Videre hadde jeg tilgang til mappetekstene elevene hadde skrevet det første året, men grunnet oppgavens tidshorisont ble det bestemt å ikke trekke de inn i arbeidet.

Det ble holdt et møte med de tre lærerne før prosjektet startet opp. Der diskuterte vi oss frem til aktuelle mål med masterprosjektet. Siden det er relativt lite forskning på feltet, hadde det vært interessant å få frem hvilke fordeler og ulemper som finnes i vurderingsformen. Videre ønsket lærerne at jeg skulle undersøke deres mappeopplegg og finne eventuelle forbedringsmuligheter slik at mappevurderingen kan videreutvikles og etableres som en aktuell vurderingsform på skolen. De tre lærerne uttrykte dessuten usikkerhet rundt mappevurdering hos andre lærere på skolen. Derfor ønsket de at masterprosjektet skulle bidra til lette på denne usikkerheten og få flere av kollegaene interesserte i mappevurdering.

Nedenfor følger en begrunnelse for valg av kvalitativ metode som forskningsdesign og deretter en fremleggelse av hvordan og hvorfor intervju ble benyttet til datainnsamling.

## Kvalitativ tilnærming

Jeg har fokus på lærernes opplevelse av mappevurdering, som beskrevet ovenfor og i problemstillingen:

*Hvordan opplever lærere bruk av mappevurdering i naturfag på videregående skole?*

Jeg ønsker ikke å ramse opp en liste med opplevelser og telle opp kvantitativt hvor mange lærere som kan si seg enige i de forskjellige opplevelsene. Formålet med studien er istedenfor å kartlegge hvordan et lite utvalg naturfagslærere på en videregående skole gjennomfører mappevurdering og undersøke hvilke opplevelser de har omkring vurderingsformen. For å kunne finne ut av dette må jeg sette meg inn i lærernes tanker og erfaringer. Jeg skal undersøke hva lærerne trekker frem som positivt med mappevurdering, men også hva de synes er utfordrende eller negativt. Valg av forskningsdesign havnet derfor på kvalitative metoder fordi det er slik jeg får tak i lærernes perspektiv på eller oppfatning av mappevurdering (Nilssen, 2012, s. 25).

## Intervju

Nedenfor følger en utredelse av hvorfor og hvordan intervju ble benyttet til datainnsamlingen. For å kunne observere hvordan mappevurderingen foregikk i praksis

ble det på et tidspunkt vurdert å bruke deltakende observasjon som et supplement til intervjuene. På grunn av forskningens tidshorisont ble ikke dette gjennomført.

### *Begrunnelse for intervju*

Jeg har brukt intervju til å samle inn data for å kunne belyse studiens problemstilling. Valget endte på intervju fordi jeg hadde et ønske om å utvikle en forståelse av hvordan lærerne på skolen opplever mappevurdering og hvilke tanker de har omkring vurderingsformen. Thagaard (2018) legger frem at intervjuer bidrar til å oppnå slik forståelse. En annen forklaring på hvorfor jeg valgte intervju var at jeg hadde begrenset tilgang til informanter og var derfor avhengig av å få mest mulig datamateriale fra det tilgjengelige utvalget (Thagaard, 2018).

Intervjuer er godt egnet når jeg skal forske på ulike aspekter av menneskelig erfaring (Kvale & Brinkmann, 2015, s. 134). Jeg utarbeidet en intervjuguide og gjennomførte halvstrukturerte intervjuer. Halvstrukturerte intervju gjør meg i stand til å planlegge tema jeg ønsker å finne ut mer om samtidig som jeg lar døren stå åpen for at informantene kan bringe frem nye sider ved saken som jeg ikke hadde tenkt på forhånd. Jeg kan stille oppfølgingsspørsmål, spørre om å få eksempler, la informantene forklare noe som først virket uklart og så videre. Denne muligheten og friheten finner man i intervju. Den halvstrukturerte formen henger sammen med at forskningsspørsmålene er formulert ved bruk av ordet *hvordan*, noe som er karakteristisk for prosjekt hvor intervju benyttes (Kvale & Brinkmann, 2015).

### *Utvalgsmetode*

Firestone (1993) hevder at det er utfordrende å si konkret hvem populasjonen omfatter i kvalitative studier. I dette forskningsprosjektet var det imidlertid kun tre aktuelle personer å intervju. Dette skyldes at det kun var tre naturfagslærere på skolen som brukte mappevurdering skoleåret 2022/23. Selv om jeg hadde hatt ønske av å intervju flere personer, ville det ikke vært flere å velge mellom. Jeg mener det derfor ikke gir mening å snakke om at populasjonen består av andre personer enn Theo, Anita og Cecilie. Ettersom alle tre godtok å stille på intervju, har i så fall hele populasjonen blitt intervjuet. Det er derfor ikke aktuelt å presentere utvalgsmetoder for dette prosjektet.

Det ble likevel vurdert å intervju andre naturfagslærere på skolen. Tanken var å intervju et par lærere i naturfag som ikke hadde tatt i bruk mappevurdering for å få

deres synspunkter. Dette ville bidratt til å gi mer dybde i datamaterialet, men det ville også fått frem årsakene til at de ikke bruker mapper. Anita, Theo og Cecilie endte opp med å presentere mulige årsaker i intervjuene deres, men det hadde også vært interessant å høre forklaringene direkte fra andre lærere. Cecilie ga beskjed til de andre naturfagslærerne og spurte om noen hadde mulighet til å delta i intervjuer, men det var dessverre ingen som hadde anledning den perioden intervjuene ble gjennomført. Derfor ble ingen andre naturfagslærere intervjuet.

### *Intervjuguide*

Jeg benyttet intervjuguide til alle intervjuene. Først ble den samme intervjuguiden brukt til de tre enkelte intervjuene med Cecilie, Anita og Theo. I et kort oppfølgingsintervju med Cecilie fikk jeg svar på noen få detaljspørsmål som hadde dukket opp ved analysering av de første intervjuene. En ny intervjuguide ble utarbeidet til dette intervjuet.

### *Den første intervjuguiden*

Den første intervjuguiden ble laget i samarbeid med veilederen min (se vedlegg 1). Vi hadde et kort møte høsten 2023 hvor vi diskuterte aktuelle temaer for intervjuene, blant annet lærernes arbeidsmengde, hvilke strategier lærerne så at elevene brukte og hvilke faglige resultater elevene satt igjen med. Etterpå laget jeg et forslag til intervjuguide som veilederen hjalp med å strukturere. Intervjuguiden besto av fem hoveddeler: 1) bakgrunnsspørsmål, 2) planlegging av mappene, 3) gjennomføring, 4) resultater og 5) konklusjon. Se vedlegg 1 for en liste over alle spørsmålene i intervjuguiden.

Intervjuguiden hadde det Kvale og Brinkmann (2015) kaller eksplorative formål. Jeg innledet hoveddel 2, 3 og 4 med et åpent spørsmål. Dette gjorde jeg for å få læreren til å fortelle om hen sin erfaring og opplevelse med mappevurdering uten at jeg ledet vei. Etter det innledende spørsmålet hadde jeg punktlistene med undertema som jeg ønsket å få frem i intervjuene. De var enten formulert som spørsmål, stikkord eller korte setninger. Jeg valgte å ikke ha en for detaljert intervjuguide av primært to årsaker. For det første ville jeg ikke låse meg til spørsmål grunnet at jeg ikke visste helt hva lærerne kom til å svare. På den måten blir intervjuguiden mer fleksibel, og jeg kan bedre få frem lærernes faktiske opplevelser rundt mappevurderingen. For det andre ønsket jeg at intervjuguiden skulle være oversiktlig slik at jeg kunne fokusere på å lytte til hva som ble sagt underveis

i intervjuene fremfor å bruke tiden på å finne frem i intervjuguiden. Den fleksible formen gjorde intervjuet halvstrukturert (Andersen, 2020).

Bakgrunnsspørsmålene i intervjuguiden ble ikke brukt noe særlig senere i forskningsprosjektet. De ble inkludert i tilfellet de var relevante og for å få en myk og rolig start på intervjuene. I konklusjonsspørsmålene lot jeg lærerne trekke frem hva de var mest fornøyde med og hvor de så størst forbedringspotensial. Dette var nyttig for meg da jeg skulle over til å analysere datamaterialet.

### *Den andre intervjuguiden*

Intervjuguiden som ble brukt til oppfølgingsintervjuet med Cecilie, se vedlegg 2, var delvis hypotesetestende (Kvale & Brinkmann, 2015). Da jeg transkriberte og kodet de tre første intervjuene, noterte jeg ned uklarheter som jeg støtte på underveis. Et eksempel er hvilket tidspunkt elevene skulle levere andreutkastet, da dette ikke hadde kommet helt tydelig frem. Slike spørsmål var konkrete og la til rette for korte svar.

Intervjuguiden var dessuten delvis eksplorativ (Kvale & Brinkmann, 2015). I de første intervjuene kom det frem at Cecilie, som er den eneste som har mappevurdering dette skoleåret, hadde planer om å legge inn begrepstester i mappevurderingen. Jeg benyttet derfor muligheten til å stille et åpent spørsmål om hvordan dette hadde foregått.

Avslutningsvis hadde jeg det åpne spørsmålet "hvordan opplever du mappevurdering?". Jeg hadde utarbeidet en skissert struktur til resultatdelen før oppfølgingsintervjuet, men jeg var ikke helt bestemt på hvilke elementer jeg ville trekke frem med tanke på hvordan lærerne opplevde mappevurderingen. Jeg inkluderte derfor dette spørsmålet i ønske om å få en form for oppsummering av hva de viktigste opplevelsene var, slik at jeg kunne bekrefte at jeg hadde trukket ut de riktige delene av de første tre intervjuene.

### *Gjennomføring av intervjuene*

Det ble totalt gjennomført fire intervjuer i forskningsprosjektet. De tre naturfagslærerne – Anita, Cecilie og Theo – ble først intervjuet på høsten 2023. Intervjuene ble holdt etter hverandre på samme dag og hadde en varighet på omtrent én time. Våren 2024 ble Cecilie kontaktet for et oppfølgingsintervju. Dette intervjuet endte etter 30 minutter. Det ble brukt en lydopptaker i alle intervjuene for å gjøre det mulig å analysere datamaterialet. Jeg har i tillegg fått svar på enkelte detaljer om mappevurderingen fra

Cecilie via spørsmål på e-post. Dette foregikk etter oppfølgingsintervjuet, og spørsmålene var korte nok til at det ikke var nødvendig å planlegge et femte intervju.

### *De tre første intervjuene*

Jeg startet intervjuene med å fortelle at datamaterialet kom til å bli anonymisert og at lydopptaket kom til å bli slettet. Etter noen kjappe bakgrunnsspørsmål var intervjuene for fullt i gang.

Det varierte hvor lenge lærerne svarte på det åpne og innledende spørsmålet til hver hoveddel, men det var i hovedsak mellom to til fem minutter. Etterpå gikk intervjusamtalen over til å omhandle underpunktene i intervjuguiden. Ettersom intervjuet var halvstrukturert, var det umulig å stille nøyaktig de samme spørsmålene til hver lærer og i nøyaktig samme rekkefølge. De trakk frem ulike aspekter underveis, og jeg stilte dermed ulike oppfølgings spørsmål. Dessuten var ikke underpunktene i intervjuguiden formulert som fullstendige spørsmål, og dermed ble også ordleggingen min rundt spørsmålene ulik i hvert intervju. Som en følge av dette har jeg ikke fått tre perspektiver – ett fra hver lærer – til hver minste detalj. I resultatdelen er det derfor tilfeller hvor det jeg legger frem kun er basert på én eller to av informantene, ikke alle tre.

Jeg gjennomførte de tre første intervjuene etter hverandre med en liten innlagt pause. Jeg kunne brukt pausen til å gjøre endringer i intervjuguiden, men jeg synes den fungerte bra slik som den var. Basert på svarene fikk jeg inntrykk av at spørsmålene jeg stilte var tydelige, og dessuten var intervjuguiden oversiktlig for min egen del og lett å manøvrere i.

### *Oppfølgingsintervjuet*

I det siste intervjuet fikk jeg oppklart de hypotesetestende spørsmålene jeg hadde med. Videre fikk jeg en bekreftelse på at det jeg hadde planlagt å trekke frem i resultatdelen faktisk var det Cecilie anså som det viktigste med mappevurderingen. Jeg kunne selvsagt ha stilt Theo og Anita tilsvarende spørsmål, men de tre lærerne har samarbeidet mye rundt mappevurderingen og var stort sett enige på det overordnede nivået.

## **Analyse av datamaterialet**

Nedenfor er en beskrivelse av hvordan jeg transkriberte og kodet intervjumaterialet. Det legges også frem begrunnelser for valgene jeg foretok i denne prosessen.

## Transkribering

Etter intervjuing følger transkribering (Kvale & Brinkmann, 2015). Jeg startet transkriberingen med å høre gjennom intervjuene på lav hastighet, gjerne halvert hastighet, slik at jeg var i stand til å skrive like raskt som lydfilen spilte av. Jeg satte pause på lydfilen og hoppet tilbake noen sekunder ved behov. Målet med denne første transkriberingen var å få skrevet ned mesteparten av det som ble sagt, altså å produsere grovtranskripsjoner.

I neste runde med transkribering hørte jeg gjennom lydfilene på nytt, men denne gangen på normal hastighet. Da sammenlignet jeg grovtranskripsjonene med lydfilene og flettet inn ord og leddsetninger som jeg ikke allerede hadde fått skrevet ned. Jeg sjekket også at tidsstemplene var riktige, og jeg strukturerte transkripsjonene med avsnitt der informanten hadde lange innspill. Bruk av avsnitt gjorde det mye lettere å lese og analysere tekstfilene.

I en siste runde gjorde jeg en del mindre endringer i transkripsjonene for å gjøre det mer muntlig og virkelig. Endringene medførte dessuten at minst mulig informasjon gikk tapt da lydfilene ble slettet. Jeg la for det første inn ord som *hmm*, *ehm* og så videre fra lydfilen. Videre symboliserte jeg pauser i intervjusamtalen på minst tre til fire sekunder med [*Tenker litt*]. Det ble behov for bedre tegnsetting da transkripsjonene fikk en mer muntlig form. En tredje endring jeg gjorde var derfor å sette komma i tilfeller hvor jeg tolket det læreren sa som en rekke med leddsetninger eller ufullstendige setninger. En siste endring var å sette anførselstegn rundt eksempler som lærerne hadde sagt i intervjuene, for eksempel kommentarer de hadde sagt eller skrevet til elevene.

## Koding

Før jeg startet å kode, hørte jeg gjennom alle lydfilene enda en gang og skrev et fortløpende sammendrag til hver. Sammendragene var omtrent én A4-side lange og besto av organisering, arbeidsmengde, utbytte, endret undervisning, samarbeid og språkmodeller. Jeg valgte denne inndelingen fordi den dekket det meste som ble sagt i intervjuene og fordi det var lite overlapp mellom temaene.

En kombinasjon av deduktiv og induktiv koding ble anvendt da jeg skulle over til å kode datamaterialet (Hsieh & Shannon, 2005). Jeg laget et forslag til kodesystem i NVivo 12, som jeg benyttet i kodearbeidet, før jeg begynte å kode intervjuene. Det besto av

kategorier og koder basert på sammendragene jeg hadde skrevet. Videre var spørsmålene som ble stilt i intervjuene delvis basert på tidligere forskning, og forskning hadde dermed påvirkning på datamaterialet som ble samlet inn. Ettersom kodesystemet dessuten ble utarbeidet i forkant av selve kodingen, oppfatter jeg dette som deduktiv koding (Hsieh & Shannon, 2005). Samtidig var ikke alle intervju spørsmålene forskningsbasert. I tillegg laget jeg nye koder underveis da jeg møtte på ytringer som ikke passet med de eksisterende kodene. Dermed mener jeg at kodeprosessen også bar preg av induktiv koding (Hsieh & Shannon, 2005).

Etter å ha kodet omkring 20 minutter av det første intervjuet oppdaget jeg imidlertid at kodene jeg hadde ikke fungerte helt optimalt. Årsaken var at det ble mye overlapp; både mellom de seks kategoriene og mellom kodene. Dette var ikke et problem da jeg skrev sammendragene, trolig fordi de ikke var altfor detaljerte. Jeg prøvde å løse problemet med å lage et helt nytt kodesystem. Det nye systemet var mer detaljert enn det første og besto til slutt av følgende kategorier:

1. Bakgrunnsinformasjon
2. Hvordan strukturere og lage mappeoppgaver
3. Hvordan veilede elevene
4. Hvordan vurdere
5. Elevstrategier
6. Resultater elever
7. Resultater lærer
8. Endret undervisning
9. Oppsummering intervju
10. Mappe i andre fag
11. Justeringer
12. Video og podkast
13. Hva tenker andre lærere

De første ni kategoriene ble laget i forkant av kodingen. Kategori 2-7 har mellom åtte og ti koder, hvorav noen koder ble laget underveis. De fire siste kategoriene ble laget mens jeg kodet fordi det dukket opp nye tema i intervjuene som ikke var dekket i de eksisterende kodene eller kategoriene. De har relativt få koder, og noen tema var dessuten omtalt i så

liten grad i intervjuene at alle ytringene endte opp i én samlet kode. I kategori 12 kodet jeg det som ble sagt med tanke på valg av oppgaveform ettersom elevene hadde mulighet til å velge mellom å skrive tekst eller å levere en podkast eller video.

Tabell 1 nedenfor viser kodene jeg brukte i kategorien *hvordan vurdere*. Til høyre i tabellen er korte ytringer som kodene inneholdt.

Tabell 1: Oversikt over kodene som ble brukt til *hvordan vurdere* og tilhørende eksempler

Kode	Eksempel på ytring
2. utkast	"Jeg liker å sammenligne første og andre utkast [...] Og da ser jo jeg om de har gjort store endringer."
Arbeidsmengde mer	"Det er selvfølgelig mye arbeid når jeg skal rette de, men det er heller ikke noe som haster."
Arbeidsmengde mindre/uendret	"Har elevene gjort en god jobb med førsteutkastet, så blir det mindre og for meg som lærer å gjøre med andreutkastet."
Innlevering 2. utkast	"Vi anbefalte elevene om å levere så fort som mulig mens de hadde teksten på en måte friskt i minnet"
Karakter 2. utkast	"De har fått karakter på andreutkastet" "Jeg prøvde å ikke gi karakter på andreutkastet."
Rettferdig vurdering	"Vi hadde jo oppe 2 partier i naturfageksamen, [...] og resultatene samsvarte egentlig veldig godt med mappekarakterene"
Samarbeid andre & språkmodeller	"Altså én del av det vi har gjort, er at vi har prøvd å lage oppgavene såpass åpne at det blir enklere å se om for eksempel de brukte ChatGPT"
Vurderingskriterier	"Jeg har en rubrikk i teams som på en måte går på det generelle: Generelt språk, kilde, innhold, fagbegrep og sånn."

Koden *2. utkast* er ganske generell og omfatter det meste som ble sagt med tanke på hvordan lærerne vurderte andreutkastene. Hvis noen ytringer ikke ble tatt med i denne koden, skyldtes det at de istedenfor ble tatt med i en av de mer spesifikke kodene i tabellen.

## Kvalitet

Nedenfor vurderer jeg kvaliteten på forskningsprosjektet med tanke på generalisering, validitet og reliabilitet.



## *Generalisering*

Innenfor forskning betyr generalisering å bruke forskningsresultatene til å si noe om andre kontekster enn den konteksten hvor forskningen fant sted. Firestone (1993) hevder at det er utfordrende å generalisere innenfor kvalitative studier, men forfatteren legger likevel frem tre perspektiver som bør tas stilling til: 1) ekstrapolering fra utvalg til populasjon, 2) analytisk generalisering og 3) "case-to-case overføring".

Firestone (1993) sier for øvrig at det raskt blir vanskelig å ekstrapolere i kvalitativ forskning. Dette skyldes både utfordringer tilknyttet hvem populasjonen består av, utvalgsmetodene som har blitt lagt til grunn og den lille utvalgsstørrelsen (Firestone, 1993). I dette forskningsprosjektet kan det tenkes at populasjonen består av Anita, Cecilie og Theo ettersom det kun var de tre naturfagslærerne som hadde erfaring med mappevurdering skoleåret 2022/23. Siden utvalget i så fall tilsvarer populasjonen, gir det ikke mening å ekstrapolere fra utvalg til populasjon. Det er likevel viktig å presisere at selv om hele populasjonen har blitt intervjuet, er populasjonen liten. Derfor er også bredden på datamaterialet smalt. Dette samsvarer med Kvale og Brinkmann (2015) sin anbefaling om å intervju 15 ± 10 personer, da dette skaper et bredere datamateriale.

Analytisk generalisering dreier seg om å generalisere fra nye forskningsfunn til allerede eksisterende forskning (Firestone, 1993). Tidligere bruk av mappevurdering viser at det er normalt å skille mellom mappevurdering som undervisningsstrategi og mappevurdering som vurderingsverktøy (Michaelsen & Johansen, 2007; Taasen et al., 2004; Taube, 2000). Dette skillet fant også sted på skolen som inngikk i denne studien. På skolen var det derimot størst fokus på vurderingsaspektet, mens det i tidligere eksempler ofte har vært størst fokus på å bruke mapper som en undervisningsstrategi. Ellers har mange av kjennetegnene ved vurderingsformen – det man kan forvente å oppnå ved bruk av mappevurdering – vært å finne ved mappevurderingen i denne studien. Det har blant annet vært fokus på elevaktivitet, ansvar for egen læring, metakognitiv utvikling og dybdelæring. Videre har funn fra studien vist at språkmodellene både bringer gevinster og trusler inn i skolehverdagen, slik blant annet Ifelebuegu (2023) beskriver.

Case-to-case overføring er en mer vanlig generaliseringsmetode innenfor kvalitativ forskning som dreier seg om at leseren gjør sammenligninger innad i et datamateriale og på tvers av flere datamaterialer (Firestone, 1993). Som forsker tenker jeg derfor at jeg bør

legge til rette for at leseren har mulighet til å foreta en slik generalisering. Jeg har gjort dette ved å gi detaljerte beskrivelser av hvordan forskningsprosjektet ble gjennomført og ved å forsøke å legge frem et så tykt datamateriale i resultatdelen som mulig uten at det går på bekostningen av strukturen og oversikten. Jeg har prøvd å få frem nyanser mellom informantene der hvor det finnes slik at det blir tydelig hvilke forskjeller man finner mellom de tre lærerne. Handlingene beskrevet gjør forskningsprosessen mer transparent, og som et resultat mener jeg dermed at leseren har bedre muligheter til å sette seg inn i forskningen og gjennomføre case-to-case overføring.

### *Validitet*

Innenfor kvalitativ forskning kan validitet sies å omhandle hvor ærlig og rikt et datamateriale er, i hvilken grad triangulering finner sted og objektiviteten til forskeren (Cohen et al., 2011). Maxwell (1992) presenterer flere typer av validitet innenfor kvalitativ forskning, blant annet deskriptiv validitet og fortolkningsvaliditet.

Deskriptiv validitet omhandler hvorvidt mine beskrivelser og fremlegginger av datamaterialet faktisk er riktige (Maxwell, 1992). For å styrke deskriptiv validitet forutsettes derfor at jeg blant annet husker riktig angående det som ble sagt i intervjuene (Maxwell, 1992). Jeg brukte en lydopptaker til å bidra på dette området, noe som Maxwell (1992) legger frem som én metode for å styrke deskriptiv validitet. Videre hørte jeg gjennom lydfilene gjentatte ganger ved transkribering for å forsikre meg om at alle ordene var fanget opp, noe jeg tenker bidrar til å opprettholde deskriptiv validitet.

Maxwell (1992) omtaler også fortolkningsvaliditet; forskerens evne til å tolke en situasjon "riktig". Dette er imidlertid ikke enkelt ettersom mennesker er forskjellige og tolker ulikt. Nilssen (2014) forklarer at tolkningsprosessen starter etter analysering av datamaterialet og innebærer å skape mening i funnene, men at det ofte er overlapp mellom tolkning og analysering. Jeg er enig i overlappet hun beskriver, og tenker at tolkningene mine gjorde sitt innpass allerede i intervjusammenheng. Da måtte jeg gjøre kjappe vurderinger av hva lærerne svarte og avgjøre hva jeg ønsket å stille oppfølgingsspørsmål til. Et grep jeg gjorde for å bedre oppfatte eller tolke det som ble sagt i intervjuene var å stille oppklarende spørsmål (Kvale & Brinkmann, 2015). I noen få tilfeller, omkring tre-fire ganger i løpet av hvert intervju, stilte jeg spørsmål til det lærerne sa – rett og slett fordi jeg ikke klarte å tolke det som ble sagt. Da fikk jeg en annen beskrivelse som jeg forsto.

Videre hadde jeg oppsummeringsspørsmål i alle intervjuene for å bidra til at jeg tolket ytringene slik lærerne ville tolket sine egne ytringer.

Jeg tenker det subjektive kjennetegnet i kvalitativ forskning som Nilssen (2012) beskriver medfører at det blir vanskelig å tolke "riktig". Samtidig tenker jeg subjektiviteten resulterer i at det er flere veier til mål i kvalitative studier. Som Kvale og Brinkmann (2015) ordlegger seg, så kan et tenkt kriterium for validitet være at funnene er troverdige; at de gir mening. Dette tenker jeg skyldes nettopp det faktum at forskere har en subjektiv rolle i forskningen. Det vil være forskjellige måter å tolke resultater på, men den ene tolkningen er ikke nødvendigvis bedre enn en annen.

### *Reliabilitet*

Innenfor kvalitative studier er kvalitet noe vi kan gjenkjenne når det foreligger, men det er vanskelig å forklare hva det er og hvordan det oppnås (Corbin & Strauss, 2015). Cohen et al. (2011) prøver likevel å legge frem noen tiltak angående reliabilitet. På den ene siden kan strukturerte intervju være bra fordi da kan de samme spørsmålene stilles til hver informant (Cohen et al., 2011). På den andre siden sier forfatteren at mer åpne intervjuer er bra for å få frem informantenes unike perspektiver (Cohen et al., 2011). Jeg tenker at forskningens problemstilling avgjør hvilken metode som velges. I dette forskningsprosjektet vil jeg få frem lærernes opplevelser av mappevurdering, og derfor kreves intervjuer som legger til rette for at lærerne kan fortelle. Dette kan ikke oppnås i like stor grad ved bruk av strukturerte intervjuer.

Videre kan reliabiliteten til en studie styrkes ved at intervjueren er dyktig (Cohen et al., 2011). Det legges ikke frem en forklaring, men jeg tolker dette i den forstand at dyktige intervjuere har noen form for lik forståelse av hvordan intervjuer bør gjennomføres. I intervjuene stilte jeg blant annet oppfølgingsspørsmål og oppklarende spørsmål ved behov, jeg prøvde å stille tydelige spørsmål, jeg unngikk ledende spørsmål og jeg forsøkte å være tålmodig og gi informantene tid til å svare. Kvale og Brinkmann (2015) presenterer kjennetegn på en dyktig intervjuer, og noen av tiltakene jeg la frem omfattes av kjennetegnene. Tiltakene kan derfor være med på å styrke reliabiliteten i forskningen. Reliabilitet kan være relevant også med tanke på transkribering og koding (Cohen et al., 2011; Braun & Clarke, 2022). Ved transkribering prøvde jeg å trekke ut det aller meste som fantes i lydfilene fra intervjuene, som beskrevet tidligere. I tillegg hørte jeg gjennom

lydfilene etter transkriberingen var fullført for å dobbeltsjekke at jeg ikke hadde misforstått enkelte ord. Transkripsjoner vil likevel ha en subjektiv dimensjon, men Cohen et al. (2011) forklarer at det gjerne er viktigere at transkripsjonene har verdi for forskeren enn at de er fullstendig pålitelige.

Braun og Clarke (2022) fremmer noen kjennetegn på "bra koding", blant annet at kodene oppstår på et organisk vis. De sier ikke eksplisitt at det styrker reliabilitet, men jeg tolker det slik. Som beskrevet tidligere laget jeg koder basert på intervjuene i forkant av kodeprosessen, jeg begynte kodingen forfra igjen og jeg slo sammen koder og laget nye ved behov. Alle endringene jeg gjorde var basert på datamaterialet foran meg og erfaringen jeg fikk underveis, og alle endringene faller innenfor Braun og Clarke (2022) sin betydning av organisk koding.

Som en oppsummering på validitet og reliabilitet kan jeg referere til Braun og Clarke (2022) som hevder at den kvalitative forskerens subjektivitet er hens primære verktøy og en forutsetning for forskning, ikke et problem som bør kontrolleres. De setter ordene *accuracy* og *reliability* i anførselstegn og poengterer problematikken med å skape (kvalitativ) kunnskap som er så objektiv som mulig (Braun & Clarke, 2022). Jeg har prøvd å vurdere validiteten og reliabiliteten i dette forskningsprosjektet, men subjektiviteten min vil alltid spille en rolle i funnene jeg presenterer.

## Etiske betraktninger

I kvalitative studier er forskeren avhengig av at andre personer har gitt plass til forskeren i livet deres og har satt av tid til å fortelle om deres tanker og opplevelser, noe som danner grunnlag for etiske betraktninger. Vivi Nilssen (2014) beskriver noen etiske hensyn som foreligger i kvalitative studier. For det første finnes det etiske retningslinjer utarbeidet av Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH) som forskning i Norge skal ivareta. Retningslinjene omfatter krav om respekt for menneskeverdet og respekt for integritet, frihet og medbestemmelse (Nilssen, 2014). UiB (2023) har et eget sett med personvern i forbindelse med forskning på universitetet, deriblant RETTE for dokumentasjon av forskningsprosjekter, som følger NESH sine retningslinjer. Dette prosjektet har blitt registrert i systemet til UiB av veilederen min og følger dermed de angitte retningslinjene.

Videre presenterer Nilssen (2014) andre etiske aspekter som må ivaretas. Det ene er informert samtykke. Veilederen min sendte ut skjema før oppstart av prosjektet som de tre lærerne skrev under på. På den måten sikrer jeg at lærerne deltar frivillig i prosjektet og er så godt som mulig kjent med forskningens hensikt (Nilssen, 2014). Lærerne hadde mulighet til å si nei, noe som Nilssen (2014) poengterer at er viktig. Jeg har i tillegg beskyttet forskningsdeltakerne på flere vis. Jeg har holdt konfidensiell informasjon til meg selv og de som inngår i forskningen og jeg har anonymisert lærernes navn. Jeg benyttet meg av en lydopptaker til alle intervjuene for å kunne foreta en grundig analyse av datamaterialet. I den sammenheng hadde det blitt avtalt med lærerne at opptakene kom til å bli slettet når de ikke lenger var av verdi og senest ved slutten av prosjektet.

## Oppsummering

I masterprosjektet har kvalitative intervjuer blitt benyttet til datainnsamling. Tre av intervjuene ble gjennomført høsten 2023, og ett intervju ble gjennomført våren 2024. Intervjuene var halvstrukturerte, og det ble brukt intervjuguider og lydopptaker til samtlige intervjuer. Datamaterialet fra de fire intervjuene ble transkribert og kodet, og funnene trekkes frem i resultatdelen nedenfor. Studiens validitet og reliabilitet, muligheter for generalisering og etiske aspekter har blitt tatt stilling til.

## Resultat

Forskningsspørsmålene for denne masteroppgaven er:

- Hvordan opplever lærere bruk av mappevurdering som vurderingsmetode i naturfag?
- Hvordan arbeider elevene med mappeoppgaver sett fra læreres perspektiv?
- Hvordan opplever lærere at mappevurdering påvirker elevenes utbytte?

På de neste sidene presenterer jeg funnene fra intervjuene. Da brukes følgende inndeling: 1) Gjennomføring, 2) lærernes opplevelse, 3) elevstrategier, 4) elevresultater og 5) andre læreres holdning.

Tabell 2 nedenfor viser sammenhengen mellom datamaterialet fra kodingen og inndelingen nevnt ovenfor. Funnene omkring lærernes opplevelse utgjør en relativt stor andel i studien og kan knyttes til flere av kategoriene, som vi ser av tabellen. Deler av ytringene i kategoriene 2, 3 og 4 – nærmere bestemt det som omhandler lærerens arbeidsmengde, se for øvrig tabell 1 – blir også behandlet sammen med lærerens opplevelse.

Tabell 2: Sammenheng mellom inndelingen i resultatet og datamaterialet i NVivo.

Inndeling resultat	Tilknyttede kategorier fra kodeprosessen (se side 26)
Gjennomføring	2, 3, 4
Lærernes opplevelse	7, 8, 10, 11
Elevstrategier	5
Elevresultater	6
Andre læreres holdning	13

## Gjennomføring

Jeg skal starte med å beskrive hvordan de tre lærerne på skolen har gjennomført mappevurdering. Jeg har valgt å inkludere gjennomføringen fordi den er nødvendig for å kunne forstå lærernes uttalelser. Uten nok detaljer er det ikke mulig å vite hva de mener.

Lærerne har hatt best erfaring med mappevurdering på studiespesialiserende. De har også prøvd det på yrkesfaglige studieretninger, men der har opplegget fungert dårligere. I intervjuene trekkes det i hovedsak frem tre mulige forklaringer. Timetallet er mye større på studiespesialiserende enn på andre studieretninger. Av den grunn er læreplanene svært forskjellige. I tillegg har elevene en annen holdning til naturfag på studiespesialiserende. Lærerne utelukker ikke mappevurdering på andre studieretninger, men foreslår at opplegget må tilpasses faget og læreplanen.

De tre lærerne er nå på deres andre år med mappevurdering, men intervjuene baseres i all hovedsak på erfaringene deres fra det første året med den nye vurderingsformen. Ettersom de har samarbeidet mye med hverandre, er det generelt lite forskjeller med tanke på hvordan de har gjennomført mappevurderingen. Jeg har likevel vist frem forskjeller i de tilfellene hvor de eksisterer.

### *Innhold i mappene*

Mappen som elevene leverer består av tre elementer: mappetekster, forsøksrapporter og refleksjonsnotater. Lærerne har delt inn naturfaget i seks hovedtema, men elevene skriver kun mappetekst til fire av de. Dette skyldes at det er satt av fire naturfagsvurderinger i prøveplanen på skolen. Elevene leverer to utkast per tekst, og mappetekstene utgjør den klart største delen av den ferdige mappen. Klima og miljø er det ene temaet som ikke dekkes direkte av mappeoppgavene. De forteller at dette gjerne er et oppsummerende tema som trekker inn en del fra det elevene har lært tidligere. Siden det kun er fire vurderinger i prøveplanen, har de derfor valgt å ikke ha en mappeoppgave til dette temaet.

Det andre temaet som ikke dekkes av mappetekstene, er naturvitenskapelig metode. Dette blir likevel dekket av at elevene gjør forsøk til hvert av de seks hovedtemaene. Elevene velger selv ut én rapport til hvert tema som de vil inkludere i mappen. Det fremkommer i intervjuene at noen av elevene velger å inkludere en rapport som dekker et annet tema enn det de har skrevet om i mappeteksten for å vise større faglig bredde. Andre elever velger den forsøksrapporten som de fikk minst kommentarer på og som dermed krever minst endringer. Til slutt skal mappene inneholde refleksjonsnotater. Elevene skriver ett refleksjonsnotat per halvår og får i notatet mulighet til å reflektere over det de har lært det siste halvåret, hva de synes var vanskelig og hva de synes var

ekstra interessant. Jeg kommer senere tilbake til at elevene dessuten kan bruke refleksjonsnotatene til å kommentere det de skrev i tidligere utkast og som de i ettertid har funnet ut at var upresist eller galt.

Ett av kjennetegnene med mappeoppgavene er at de har en åpen oppgaveformulering. Lærerne begrunner avgjørelsen med at det gir rom for at elevene da kan skrive om et vidt spekter av tema, slik at alle elevene kan finne noe å skrive om som de interesserer seg for. Elevene fikk følgende oppgavetekst tilknyttet arv, evolusjon og bioteknologi:

I denne oppgaven skal dere fordype dere i et selvvalgt tema innenfor temaet arv, evolusjon og bioteknologi. Dere skal lage en problemstilling knyttet til de to kompetansemålene gitt under. Ta utgangspunkt i én eller flere metoder knyttet til bioteknologi, og drøft faglige og etiske problemstillinger knyttet til denne bruken av bioteknologi.

I dokumentet fikk elevene også oppgitt de to relevante kompetansemålene samt ni eksempler på bruksområder innen moderne bioteknologi. Noen utvalgte eksempler er genmodifisering av levende organismer, stamcelleforskning, kunstig befruktning og xenotransplantasjon.

Denne mappeoppgaven var en av de siste oppgavene som elevene fikk. Theo forklarer at de ikke kan forvente at elevene klarer å lage egne problemstillinger de første gangene. Til de første tekstene har derfor lærerne enten gitt elevene en ferdig problemstilling eller latt elevene velge mellom noen problemstillinger.

I tillegg til at elevene har valgfrihet omkring hva de skriver om, har de dessuten valgfrihet angående oppgaveform. Elevene kan nemlig velge mellom å skrive tekst eller å spille inn og levere enten en video eller en podkast. I intervjuet med Cecilie ble det sagt at elevene som velger video eller podkast gjerne gjør dette fordi de vil gjøre noe annet enn å skrive tekst, noe som de allerede har gjort mye av tidligere i skolegangen deres. Mye av prosessen er imidlertid lik uavhengig av hvilken oppgaveform elevene velger. De har samme oppgavetekst som utgangspunkt, de får ganske lik veiledning underveis og vurderingskriteriene er tilnærmet like. Forskjellene er at elever som leverer en video eller podkast som andreutkast typisk leverer en disposisjon i form av stikkord som førsteutkast. De skriver derfor ikke en tekst som de leverer og får tilbakemeldinger på,



men de skriver istedenfor en disposisjon over det de har tenkt å si og hvilke fagbegreper de har tenkt å bruke. Jeg kommer i liten grad til å trekke inn bruk av video eller podkast senere ettersom det viste seg å være små forskjeller mellom oppgaveformene. Dessuten var det skriving av mappetekster som ble hovedfokuset i intervjuene.

### *Veiledning*

Elevene får utdelt mappeoppgaven idet de begynner med et nytt tema. Etter utlevering foregår det meste av undervisningen som normalt, men et par uker før innleveringsfrist av førsteutkastet får elevene naturfagstimene på skolen til å arbeide med utkastet. Jeg kommer ofte til å omtale de skoletimene som skriveøkter. Mye av veiledningen foregår i løpet av de to ukene med skriveøkter. Videre får elevene veiledning i form av tilbakemeldinger på førsteutkastet. Da får elevene kommentarer direkte i teksten og en generell oppsummering av det læreren mener at de bør jobbe videre med.

Elevene får i tillegg en oversikt over hvor godt førsteutkastet er på de ulike vurderingskriteriene. De tre lærerne har gjort noen justeringer på vurderingskriteriene fra det første året med mappe og frem til i dag. Foreløpig vurderer de utkastene basert på struktur, problemstilling, fagterminologi, språk, sammenhenger og kildebruk.

Naturfagslærerne har valgt å dele inn kriteriene i fire nivåer: Utmerket, bra, ganske bra og dårlig. Den firedelte skalaen resulterer i at elevene får en god indikasjon på hvor bra førsteutkastet er uten at det blir helt tydelig hvilken karakter utkastet tilsvarer. Lærerne har nemlig vært konsekvente på å ikke gi elevene karakter på førsteutkastet for å dempe litt karakterpresset som de opplever på videregående skole. Denne tråden gjenopptas senere under elevresultater.

Det blir ikke satt av tid i skoletimene til å arbeide med andreutkastet, men elevene kan spørre lærerne om det skulle være noe de lurer på. Innleveringsfristen på det andre utkastet har vært noe varierende. Det første året med mappevurdering hadde alle fire andreutkastene én samlet innleveringsfrist i mai. Det kom frem i intervjuene at elevene hadde blitt oppfordret til å gjøre endringer på førsteutkastet og levere andreutkastet ganske snart etter de fikk tilbake det første utkastet, men at det likevel var en del elever som hadde utsatt arbeidet til godt ut på våren. Jeg kommer tilbake til mulige forklaringer på denne utsettelsen under elevstrategier. Implikasjonen det andre året var imidlertid klar: Elevene skulle ikke ha mulighet til å utsette arbeidet med alle fire andreutkastene.

Det andre året har derfor andreutkastene på høsthalvåret én samlet innleveringsfrist før jul og andreutkastene på vårhalvåret én samlet frist i mai.

### *Vurdering*

Ved vurdering av andreutkastene bruker lærerne de samme kriteriene som ved veiledning på førsteutkastet. Det er likevel noen små forskjeller. For det første får elevene denne gangen en karakter på utkastet de har levert. Theo unnlot å gi karakterer det første året, men både han og elevene synes det var vanskelig å forholde seg til. Nå gir derfor alle lærerne karakter på andreutkastet. Elevene får karakteren med kommentarer i teksten fortløpende etter hvert som lærerne vurderer. For det andre får elevene færre kommentarer i dokumentet på andreutkastet. Dette skyldes at elevene er ferdige med temaet etter innlevering av andreutkastet. Alle lærerne poengterer likevel at de gir kommentarer på det som elevene kan få bruk for i senere oppgaver. Eksempler som de trekker frem er struktur på teksten og kildeføring. De vurderer også forsøksrapportene og refleksjonsnotatet som elevene skriver, men det er mappetekstene som veier mest når karakteren på mappene skal settes.

Nedenfor bygges det videre på gjennomføringen som har blitt presentert. Det beskrives mer detaljert hvordan mappevurderingen har hatt innvirkning på lærerne, både med tanke på arbeidsmengde og det faktum at veilednings- og vurderingsarbeidet skiller seg fra en del andre vurderingsformer.

### **Lærernes opplevelse**

I forkant av oppfølgingsintervjuet med Cecilie hadde jeg en grov skisse av hvordan denne delen av resultatet kom til å se ut basert på de tre første intervjuene. Jeg valgte likevel å direkte stille henne spørsmålet: "Hva tenker du er det viktigste med din opplevelse av mappevurdering?", og forklarte at jeg så for meg at dette kom til å utgjøre én av resultatdelene. Jeg ga henne tid til å tenke, og etter litt betenkningstid var det arbeidsmengde og at hun generelt har en positiv innstilling til mappevurdering som ble trukket frem. Hun sa at det var vanskelig å nevne noe annet konkret, og jeg har selv valgt å bruke denne todelte strukturen. I de fire intervjuene ble det også trukket frem mange andre fordeler med mappevurderingen slik de tre lærerne ser det, men de angår fordeler hos elevene og blir nevnt etterpå under elevresultater.

## *Arbeidsmengde*

Alle tre lærerne uttrykker at det samlet sett er mer arbeid med mappevurdering enn med for eksempel tradisjonelle, skriftlige naturfagsprøver. På de neste sidene skal jeg med utgangspunkt i intervjuene beskrive arbeidsmengden ved tillaging, veiledning og vurdering av mappetekstene. Jeg skal også presisere valg lærerne har gjort underveis for å redusere arbeidsbelastningen.

## *Tillaging av mappeoppgaver*

Da de startet opp med mappevurdering det første året og startet helt fra begynnelsen av, sier samtlige at det tok betydelig tid å utarbeide oppgavene. Theo uttaler at "*det å lage til oppgavene og det å lage til rapporter og forsøksbeskrivelser og den type ting, det var mye arbeid*". Cecilie utdyper at "*jeg ser for meg at det er mye arbeid i starten med å lage gode oppgaver, og jeg har brukt mye tid på å strukturere: Hva er det som skal inn? Hvilke forsøk skal vi gjøre? Hvordan skal man utforme et refleksjonsnotat?*". Før læreren kan ta i bruk mappevurdering må hen både formulere oppgavetekster, forsøksbeskrivelser og refleksjonsnotat, ha tenkt gjennom hvordan alle de ulike delene skal vurderes, når de skal ha innleveringsfrist og hvordan veiledningen skal foregå. Dette tar tid.

Theo uttrykker at det tok mye tid i starten og at de var klar over denne arbeidsbelastningen. Han fortsetter imidlertid med å si at det i år to er klart mindre arbeid på denne fronten "*fordi vi gjorde et såpass grundig arbeid da vi begynte med det*". Sitatet kan gi inntrykk av at Theo har mappevurdering i år, noe som ikke er tilfellet. Han og Anita har imidlertid hjulpet Cecilie med å videreutvikle vurderingsformen og vet derfor at arbeidsmengden har blitt redusert. Cecilie sier at de brukte et par timer på å lage oppgavene i fjor, men at de i år kun har brukt en halvtime på å finne ut av hva de vil endre på. Anita sier noe tilsvarende, og legger til at det nå ville gått raskere for henne å lage mappeoppgaver til en ny klasse enn det ville gjort å lage en ny naturfagsprøve. Cecilie legger videre frem en god forklaring på denne tidsbesparelsen:

Jeg tenker at når vi har gjort det et par år, så vil man jo ha en bank med oppgaver som man kan dele ut. Hvis de er gode, så er det jo lite arbeid. Vi kan kanskje lage nye oppgaver innimellom eller endre på tema for å få litt variasjon, men ellers er det jo oppgaver som kan gis år etter år.

Hun poengterer at dette selvsagt ikke er mulig med skriftlige to-timers prøver fordi hun ikke kan gi samme prøve til ulike klasser. Anita er enig og uttaler at denne reduserte arbeidsmengden ikke er å finne ved bruk av vanlige skoleprøver. Hun forklarer at det alltid vil være noen elever som har eldre søsken eller venner som har hatt prøven tidligere, og at en prøve derfor ikke gis på nytt i en senere anledning.

### *Veiledning*

Alle tre lærerne er enige i at arbeidsmengden tilknyttet veiledning av elevene er stor. Ettersom elevene selv kan velge hva de vil skrive om, gitt at de klarer å trekke det inn i en faglig kontekst, er resultatet at alle elevene gjerne ender opp med å skrive helt forskjellige oppgaver. Noen ganger skriver eleven om noe som læreren har mye kompetanse innenfor. Andre ganger skriver eleven om noe som læreren ikke har hørt noe særlig om fra før og som man ikke finner i lærebøkene. Læreren må da lese seg opp på det fremmede temaet for å kunne gi veiledning. Både Cecilie og Anita sier at de foretrekker denne variasjonen og mener at det er fordelaktig for elevenes læring. Anita er tydelig i sin tale: *"Jeg synes jo det er veldig gøy å på en måte kunne gå lenger inn i tema som ikke jeg kan så masse om nødvendigvis"*.

Denne variasjonen medfører likevel mer arbeid fra lærernes side underveis i prosessen. Cecilie sier at hun typisk leser seg opp på det faglige mellom skoletimene. Anita gjør det annerledes: *"Jeg prøver å unngå å gi meg selv veldig masse lekser med at jeg skal sjekke opp dette, og dette og dette. For det fungerer ikke på en måte"*. Hun sier at hun da raskt blir overarbeidet som lærer, så derfor *"må man få gjort ting med elevene i timen"*. Anita beskriver at hun søker opp kilder og finner svar på spørsmålene fortløpende i skoletimene sammen med elevene. Selv om hun gjør dette raskere enn elevene, sier hun likevel at det går masse tid til å hjelpe én elev. Hun avslutter med å si at det kan være vanskelig å finne balansen med hvor mye tid hun skal bruke på hver enkelt elev.

Like etterpå i intervjuet får Anita frem et viktig poeng. Hun forteller at i periodene når elevene arbeider med mappeoppgavene på skolen, altså i forkant av innleveringsfrist til førsteutkastet, vil det være et par uker hvor hun ikke trenger å planlegge et undervisningsopplegg, lage Powerpoint-presentasjoner og finne oppgaver som elevene skal arbeide med. Hun sier at *"da opplever jeg at det er en god periode der jeg har lite arbeid med faget fordi jeg bare går til timen og på en måte veileder i timen"*. For Cecilie

blir ukene noe mer travle siden hun svarer på elevspørsmål og finner ut av det faglige utenom skoletimene, men også hun slipper det normale forberedelsesarbeidet i forkant av skoletimene.

Det kan tyde på at en god andel av arbeidet tilknyttet veiledningen skjer i form av å gi tilbakemeldinger på førsteutkastene. Anita gir et estimat på at hun bruker omkring 20-30 minutter på å gi kommentarer på førsteutkastene, som for øvrig er på omtrent 1000 ord. Hun beskriver at det som tar tid, er å gi tilbakemeldinger som på den ene siden er generelle, for eksempel i form av "*hvordan er språket – er det presist nok?*", "*hvordan er kildebruken – er den riktig?*" og "*hvordan er strukturen og innholdet?*". På den andre siden ønsker Anita at tilbakemeldingene skal være konkrete nok til at elevene forstår dem og skjønner hvor de skal begynne å jobbe frem mot andreutkastet.

Selv om lærerne bruker en del tid på å gi tilbakemelding på det første utkastet, forteller de i intervjuene om tiltak de har gjort for å redusere arbeidsmengden. De har for eksempel tatt i bruk et tillegg i Microsoft Word som gjør det mulig å forhåndsskrive kommentarer som ofte går igjen. Et eksempel fra Cecilie er: "*Kildene dine er ikke ført riktig. Typiske ting som gjør at de blir dårlig ført er at du mangler forfatter, årstall eller komma*". På den måten kan de raskt legge inn de ferdige kommentarene underveis. I tillegg er de helt tydelige på det vi kan kalle øvelse gjør mester. Alle tre har nemlig klart lagt merke til at det går mye raskere å veilede elevene – både i skoletimene og på førsteutkastet – og vurdere det endelige utkastet når de har fått erfaring med det.

### *Vurdering*

Vurdering av andreutkastene fungerer i stor grad på tilsvarende vis som veiledning av førsteutkastene. Selv om det ikke gis like mange kommentarer den andre gangen, må utkastene likevel leses gjennom og endringer må fanges opp. Totalt sett på de to utkastene sier Anita at "*det blir jo på en måte dobbelt, eller litt dobbelt opp, i hvert fall én og en halv sammenlignet med prøver*". Theo uttrykker noe lignende, og tilføyer: "*har elevene gjort en god jobb med førsteutkastet, så blir det mindre for meg som lærer å gjøre med andreutkastet*". Han forklarer at hvis elevene har levert et tilnærmet ferdig produkt som førsteutkast og han da har gitt en grundig tilbakemelding, vil det være lite endringer for elevene å foreta frem til det endelige utkastet. Som et resultat sier Theo at

det da også går raskt å vurdere andreutkastet, naturlig nok siden det ikke har blitt gjort store endringer.

### *Annerledes veilednings- og vurderingsarbeid*

Lærere er forskjellige og har ulike preferanser. Av de tre lærerne som har prøvd ut mappevurdering i naturfag, var likevel samtlige positive til vurderingsformen og til resultatene som elevene satt igjen med. Alle tre har planer om å fortsette med mappevurdering og videreutvikle den. Anita, Cecilie og Theo legger frem mange argumenter for hvorfor de foretrekker mapper fremfor andre vurderingsformer. Noen av argumentene har tilknytning til lærerens arbeidsdag og det læreren sitter igjen med av å bruke mappevurdering. Andre argumenter handler om hvorfor mappevurdering kan være gunstig for elevene. I de neste avsnittene presenterer jeg kun førstnevnte. Elevenes mulige gevinster av mappevurdering tas opp senere under elevresultater.

Et argument som gjentar seg i alle de tre intervjuene, er at lærerne synes det er kjekt å veilede elevene mens de arbeider med førsteutkastene på skolen. De får da tid til å gå rundt og snakke med hver enkelt elev, noe Cecilie trekker frem som en gyllen mulighet til å bygge gode relasjoner til elevene.

Lærerne uttrykker at de får mindre tid til å gjennomgå fagstoffet ettersom det går med noen skoletimer til å jobbe med mappeoppgavene, og resultatet er at de ikke setter av tid til repetisjon ved slutten av temaene. Mappene har altså en påvirkning på lærerens undervisningspraksis. Vurderingsformen kan også påvirke oppgavene elevene arbeider med i skoletimene. Cecilie forteller at hun typisk lar elevene arbeide med diskusjonsoppgaver hvor de skal forklare fagstoff og knytte det til hverdagslige eksempler. Dette er oppgaver som forbereder elevene på å skrive mappeoppgaver.

Som nevnt må lærerne typisk bruke litt tid på å sette seg inn i det faglige som elevene skriver om, men både Anita, Cecilie og Theo er positive til dette. Theo sier til og med at *"det er jo noen av de tingene jeg har tatt inn i undervisning på det temaet ved senere tidspunkt"*. Cecilie påpeker at *"det er kjempespennende å lese meg opp på ting utenfor pensum, spesielt i kjemi som er mitt hovedfag, der kan jeg gjerne hjelpe de"*. Hun forteller at det blir litt vanskeligere når elevene skriver oppgave til temaer tilknyttet biologi eller fysikk, men at det også går greit.

De tre lærerne er enda mer positive til vurderingsaspektet. Anita legger frem at "det er mye kjekkere å rette ulike tekster enn å rette samme prøve". Theo formulerer seg som følger: "Det å lese den samme definisjonen av for eksempel global oppvarming 26 ganger etter hverandre, det er anstrengende når en sitter og retter". Han sier at dette var én av årsakene til at han tok i bruk mappevurdering; for å gjøre vurderingssituasjonen mer interessant for han som lærer. Anita er dessuten "litt redd for at det bare går på autopilot med skriftlige prøver, at man kanskje er litt for rask med å sette karakterer av og til". Hun ser for seg at hun klarer å være mer skjerpet og konsentrert når hun skal vurderer mappetekstene ettersom alle er forskjellige. Selv om de tre lærerne foretrekker variasjon i elevbesvarelsene, uttrykker de at lærere er forskjellige og at andre kan foretrekke å vurdere besvarelser som har mer til felles.

Cecilie presenterer et annet poeng som hun nå synes er bedre ved vurderingssituasjonen. Da hun brukte skriftlige prøver i naturfag, forteller hun at "så sitter jeg én eller to uker før og må lage prøven, også skal den ikke være for vanskelig og ikke være for lett. Og hvordan var arbeidsmengden her og der og sånn? Også skal jeg lage løsningsforslag". Nå slipper hun å "styre med dette underveis" siden oppgaven blir gitt ut til elevene når de begynner på et nytt tema. Hun konkluderer med at hun bruker mindre tid på å tenke på vurderinger i naturfag og at det er mindre stressende, noe hun synes er deilig.

De to neste overskriftene – elevstrategier og elevresultater – gir en beskrivelse av hvordan elevene arbeider med mappeoppgavene og hvorfor mappevurdering kan være gunstig for elevene.

## Elevstrategier

I intervjuene ble det tydelig at lærerne ser at elevene tar i bruk ulike strategier når de skal skrive mappeoppgavene. Jeg har valgt å dele inn elevstrategiene i fire stadier. Først skal jeg beskrive i hvilken grad elevene arbeider med mappeoppgavene i forkant av de to ukene med skriveøkter på skolen. Deretter legger jeg frem i hvilken grad elevene tar nytte av veiledningen de får i skriveøktene. Videre skal elevene levere et førsteutkast. Da viser det seg at noen elever leverer en tilnærmet ferdig tekst, mens andre elever kun har skrevet noen få setninger. Til slutt skal jeg beskrive hvordan elevene arbeider videre med kommentarene de får på førsteutkastet når de gjør endringer frem mot det endelige

utkastet. Det er viktig å presisere at dataene som legges frem her er basert på lærernes erfaringer, meninger og observasjoner – ikke på hva elevene selv har sagt.

### *Komme i gang tidlig eller sent?*

Fra oppgavene deles ut ved starten av et nytt tema og frem til det settes av tid til å arbeide med førsteutkastet på skolen, er det flere elever som i mellomtiden har begynt å tenke gjennom hva de vil skrive om og hvilken problemstilling de vil ha. Før elevene skulle begynne å arbeide på skolen, uttaler Cecilie at "*da hadde jo veldig mange bestemt seg for temaet allerede. De hadde begynt å tenke, de hadde lest litt på stoffet som var relevant, og de hadde tenkt på det ene og det andre*". Hun sier at selv om den første økten på skolen kun er et par uker før innleveringsfristen, begynner ikke elevene på null. Dette kan skyldes at lærerne henter til at elevene kan skrive om det ene og det andre underveis i undervisningen og at elevene av og til får de siste minuttene av en skoletime til å komme i gang med eller fortsette arbeidet med mappeoppgavene. Likevel er det elever som ikke har kommet særlig i gang i forkant av skriveøktene på skolen. Lærerne legger raskt merke til dette, som beskrevet nedenfor.

### *Bruk av skriveøktene på skolen*

Når startskuddet går og elevene skal arbeide med mappene på skolen, er den elevgruppen som forberedte seg godt raskt i gang og fortsetter arbeidet der de har kommet. Den andre elevgruppen bruker lenger tid på å komme i gang, noe som lærerne raskt oppdager. Theo får frem at det er ekstra synlig hvem som ikke har kommet i gang fordi oppgavene er åpne, noe som forutsetter at elevene må arbeide selvstendig. I de siste mappeoppgavene skal elevene dessuten formulere problemstilling. Da er det viktig å ta en runde i klasserommet i starten av de to ukene for å passe på at alle elevene får formulert problemstillinger og skriver om noe som er relevant, presiserer Anita.

Uavhengig av hvordan de har forberedt seg, trekker Cecilie frem at elevene bruker de to ukene med veiledningstimer bra og at de er flinke til å stille spørsmål. Spørsmålene kan for det meste kategoriseres i tre grupper, nemlig spørsmål tilknyttet struktur, faglig innhold og kildebruk. Noen eksempler som lærerne gir i intervjuene er "*er dette en god forklaring?*", "*hvor kan jeg finne relevante kilder på dette her?*", "*kan du lese denne setningen?*" og "*har du et eksempel på dette?*". Samtidig forteller lærerne at elevene er svært forskjellige med tanke på hvor selvstendige de er. Noen av elevene spør om lite



veiledning og ønsker å utforske på egenhånd. Andre er mer avhengig av å dobbeltsjekke at valgene de gjør underveis er fornuftige.

### *Status førsteutkast*

Etter skriveøktene er det tid for innlevering av førsteutkastet. Theo forteller at noen elever leverer førsteutkastet et par dager i forveien, men at de aller fleste leverer like før innleveringsfristen. På samme måte som arbeidsprosessen blant elevene kan deles i to, nemlig om de kommer i gang tidlig eller sent, kan også mengden de leverer på førsteutkastet grupperes i to. Noen av elevene skriver et utfyllende utkast og *"et tilnærmet ferdig produkt"*, som Theo formulerer det. Han kan da gi grundige tilbakemeldinger som elevene kan arbeide videre med når de skal gjøre endringer til det siste utkastet. Andre elever er mer usikre på hvordan de skal begynne og har kanskje bare levert tre setninger. Cecilie sier at hun da ikke har mulighet til å gi noen særlige kommentarer, men at hun gjerne skriver *"kanskje du skal skrive om [dette] temaet, da bør du innom disse og disse fagbegrepene"*.

Anita presiserer at det er viktig å poengtere veldig tydelig til elevene at *"ja, det er kun et førsteutkast, men tenk at det på en måte er en tekst som du er fornøyd med"* slik at hun kan gi gode tilbakemeldinger. Hun fortsetter at man mister litt poenget med denne type vurdering hvis elevene ikke benytter seg av muligheten til å levere et godt førsteutkast. Elevene får dermed lite tilbakemeldinger og mindre muligheter til læring. Theo hevder også at elevene lærer mer av å skrive et fullstendig førsteutkast, blant annet fordi elevene da må arbeide med teksten over flere arbeidsøkter.

### *Innlevering andreutkast*

Alle lærerne oppfordrer elevene til å gjøre endringene ferdige og levere andreutkastet istedenfor å utsette alt til siste frist. Noen elever tenker at de vil bli ferdige og gjør derfor raskt endringer på utkastet og leverer, forteller Theo. Han forteller også at det alltid er noen som utsetter alt til siste innleveringsfrist, og da hopper det seg også opp for hans egen del. En årsak til at elevene utsetter å levere andreutkastet blir beskrevet av Theo. Han sier at fordelene med å vente er at elevene da får mer faglig kunnskap innenfor andre tema som de kan trekke inn i oppgavene. Theo og de andre lærerne sier imidlertid at dette enkelt kan løses ved at elevene leverer andreutkastet tidlig og bruker

refleksjonsnotatene til å drøfte hvordan tekstene deres kunne ha vært bedre etter de fikk mer faglig kompetanse.

## Elevresultater

På de neste sidene skal jeg legge frem elevenes faglige utbytte av mappevurdering og hvordan mappevurdering har påvirket elevene med tanke på arbeidsvaner, motivasjon, naturfaglig interesse og stress. Til slutt presenteres de viktigste funnene fra elevintervjuene som Johanne og Elise Marie (2024) gjennomførte i deres masteroppgave.

### *Arbeidsvaner, motivasjon og interesse*

I intervjuene tok jeg opp elevenes arbeidsvaner, faglige utbytte, motivasjon og interesse. Lærerne hadde merket noen endringer i både arbeidsvaner, motivasjon og interesse etter de tok i bruk mappevurdering, men Theo sier at dette er veldig individuelt. Han sier imidlertid at han har møtt på elever året etter han hadde naturfag med de, og at elevene har begynt å snakke om det de skrev om i mappeoppgavene og fremdeles er engasjerte i diskusjonen. Dette har han ikke opplevd tidligere. Både Anita og Cecilie er enige i at valgfriheten i oppgavene og det at elevene kan velge tema selv gjør at de blir mer engasjerte i det de skriver om. Anita forteller at hun har hatt elever som har kommet bort og sagt at "*denne oppgaven var veldig gøy å skrive*". Hun har heller ikke hatt slike erfaringer før hun tok i bruk mappevurdering.

I det andre intervjuet med Cecilie forteller hun at noen av klassene fikk muligheten til å velge mellom mappevurdering og skriftlig skoleprøve på den ene vurderingen de skulle ha. Da dukket det opp et interessant utfall. I hennes klasse, som allerede hadde hatt mappevurdering et par ganger, var det ingen som ville ha en tradisjonell skriftlig prøve fremfor mappevurdering. Mulige årsaker til dette blir presentert nedenfor basert på datamaterialet fra elevintervjuene. Cecilie forteller at det i de andre klassene var rundt halvparten av elevene som valgte skoleprøver.

### *Faglig utbytte*

Det kommer tydelig frem i intervjuene at lærerne har lagt merke til at elevene i større grad klarer å se sammenhenger i faget når de har hatt mappevurdering. De presiserer at de kun har prøvd mappevurdering i noen få klasser og at det også kan være andre faktorer som spiller inn, men de forklarer at elevene så faget fra en annen vinkel enn

tidligere klasser har gjort. Cecilie forteller at dette også kom tydelig frem i standpunktkarakterene og at det var et helt annet nivå enn det hun har hatt tidligere i naturfag, til tross for at hun sier at hun er veldig streng når hun setter karakterene.

Theo hadde for øvrig noen av elevene hans oppe til muntlig praktisk eksamen. Etter kun to elever fikk han høre uoppfordret fra sensoren at "*oi, de er mye flinkere på å se liksom det store bildet*". Theo sier at elevene kunne forklare hvorfor matsvinn er dårlig og koble dette til utslipp, frakt og "*hele den bøttebilletten*". Som hintet til tidligere mener Anita at dette skyldes at elevene får arbeide med temaet over to utkast. Cecilie tror også at elevene lærer mer når de får bearbeide ting over lenger tid og får tilbakemeldinger på det de har skrevet.

Selv om det kan tyde på at elevene ble bedre til å se sammenhenger, la lærerne også merke til at elevene ikke var fullt så gode på faktakunnskaper. Cecilie forklarer at ettersom elevene ikke blir spurt om å gjengi definisjoner i mappeoppgavene, er det naturlig at de ikke har like god kontroll på dette som tidligere elever. Sensoren på eksamen trakk også frem denne manglende kompetansen. Anita har en verdifull kommentar til dette: "*Jeg ønsker å trene opp elevene sine ferdigheter til å på en måte bearbeide stoff, se sammenhenger og formidle på en gode måte heller enn at de skal huske hva disse fagbegrepene betydde på stående fot*". Hun fortsetter som følger:

Dette er jo elever som går på studiespesialiserende og som sannsynligvis skal inn i høyere utdanning. Jeg tror de vil ha bedre forutsetninger for å lære seg akademisk skriving og tilpasse deg til de kravene som blir stilt senere når vi har hatt mappevurdering.

Cecilie er likevel ikke tilfreds med begrepsforståelsen blant elevene og sier at de må lære seg noen definisjoner. For å bedre situasjonen har Cecilie i det andre året prøvd å gjennomføre begrepstester med elevene. Cecilie forteller at planen er å ha én begrepstest per tema. Da blir det satt av to skoletimer hvorav den første brukes til begrepsøving. Den andre skoletimen brukes til en flervalgstest uten karakter for å unngå ekstra arbeid. Cecilie forteller at elevene likevel bør gjøre det omtrent like bra på begrepstesten som på mappetekstene, selv om de ikke får karakter på begrepstestene. I oppfølgingsintervjuet sier Cecilie at hun har prøvd begrepstest et par ganger og at det fungerte ganske bra. Mange av elevene øvde bra og tok testen seriøst, men det var også

noen elever som ikke tok testen like høytidelig. Cecilie uttaler at de fremdeles endrer på begrepstestene og at de til neste år har planer om å erstatte begrepstestene med korte videoer som elevene skal spille inn hvor de forklarer tre til fem begreper.

### *Emosjonelt*

Ved siden av faglig utbytte er elevenes emosjonelle status en av de viktigste begrunnelsene for bruk av mappevurdering, ifølge lærerne. Cecilie sier det slik: *"Det jeg tenker at kanskje er det viktigste med den måten å vurdere og jobbe med faget på, er jo at det blir mindre stress for elevene, at man klarer å senke skuldrene litt"*. Theo sier også at han har fått tilbakemelding fra elever som sier at de likte bedre å jobbe på den måten og at de følte seg mindre stresset. Anita trekker dessuten frem karakterpress og generelt stress i hverdagen som årsaker til å bruke mappevurdering.

Det er nemlig flere elementer ved mappeoppgavene som kan føre til at elevene blir mindre stresset enn ellers. For det første har elevene to utkast på hver oppgave. Theo forteller at det er elever som har fått seg en aha-opplevelse av at *"det må ikke være perfekt, jeg får tilbakemeldinger på det jeg har"*. Cecilie sier også at elevene ikke har like stort press på seg når de skal skrive det første utkastet siden de *"ikke må prestere her og nå"*. De kan arbeide i sitt eget tempo og de kan skrive om det de interesserer seg for, forteller Cecilie.

For det andre får elevene tid på skolen til å arbeide med oppgavene. Både Cecilie og Anita trekker frem at naturfag ikke er et stort fag for elevene på studiespesialiserende, og de vil derfor ikke at elevene skal stresse med at det er mye arbeid og mange vurderinger i faget. For Cecilie er det derfor viktig at de får tid på skolen til å jobbe. Dessuten sier hun at dette er nødvendig for at hun skal kunne hjelpe elevene og gi de veiledning underveis.

Et tredje grep lærerne har tatt i bruk for å gjøre vurderingssituasjonen mindre stressende for elevene, er å ikke gi de karakter på førsteutkastene. De mener at dette er med på å dempe karakterpresset som elevene føler på når de går på videregående. Theo prøvde dessuten å ikke gi karakter på andreutkastet heller, men fant ut at det ble utfordrende både for han selv og for elevene.

## *Hva sier elevene?*

Johanne og Elise Marie (2024) har intervjuet et utvalg av elevene til Anita, Theo og Cecilie i forbindelse med masteroppgaven deres. Av de tretten elevene som ble intervjuet, var det elleve som likte mappevurdering fremfor vanlige prøver. Begrunnelsene var stort sett basert på tre prinsipper: 1) elevene hadde mer valgfrihet, 2) det var mindre stressende og 3) de fikk bedre vist hva de kunne i vurderingssammenheng. Prinsippene omtales mer detaljert i påfølgende avsnitt.

Mange av elevene hadde sagt at de likte mapper bedre fordi de fikk mulighet til å bestemme mer selv. Dette gjaldt både med tanke på oppgaveform – tekst, podkast eller video – og innhold, altså hva tekstene skulle omhandle. Det at de ikke ble tvunget til hva de skulle gjøre, var noe de tydelig verdsatte. Videre nevnte mange at det var mindre stress med mappevurdering. Her var det særlig to argumenter, som i en viss grad overlapper med hverandre. For det første sier elevene at de hadde god tid til å gjøre oppgavene ettersom de fikk utlevert oppgavene lenge før innleveringsfristen. De kunne derfor gjøre det når de fikk tid. For det andre slapp elevene problematikken med at dagsform kan ha stor påvirkning på et prøveresultat. Hvis elevene er syke når det er skoleprøve, er det ikke mye de kan gjøre. Dagsform er derimot ikke et problem med mappevurdering. Elevene kommenterte dessuten at de følte de bedre fikk vist frem det de kunne når det var mappevurdering. En prøve har fastsatte spørsmål, mens mappeoppgavene er åpne. Elevene kan da trekke inn detaljer som de muligens ikke ville blitt spurt om på en prøve.

Det er to elever som foretrekker skoleprøver. De endte likevel opp med gode karakterer, men trakk frem at det negative med mappevurdering er at de ikke liker å skrive tekster. Videre er skoleprøver mer forutsigbare fordi de kun må øve på og pugge det som står i læreboken.

## **Andre læreres holdning**

Det ble ikke mulig å intervju naturfagslærere på skolen utenom de tre som allerede anvender mappevurdering. På de tre intervjuene spurte jeg likevel om hvilken holdning andre naturfagslærerne på skolen har til mappevurdering. Da var det i hovedsak tre perspektiver som kom frem: De opplever at andre lærere er skeptiske til arbeidsmengden, de er skeptiske til elevenes læringsutbytte og de er skeptiske til

hvordan mappevurdering kan fungere når kunstig intelligens er tilgjengelig for elevene. Perspektivene utdypes nedenfor.

Arbeidsmengden er som påpekt tidligere større enn for skriftlige prøver, særlig de første årene. Selv om det er mer arbeid, får Cecilie frem et viktig poeng: *"Jeg kan gjerne bruke tid på vurderingsarbeid så lenge det også fører til læring hos elevene. Jeg hadde ikke giddet å bruke så mye tid på det hvis de ikke lærte noe av det"*. Andre lærere er likevel skeptiske til elevenes læringsutbytte. Elever med mappevurdering blir nemlig ikke testet i nøyaktig det samme som elever med skriftlige prøver. Som sagt tidligere kan det tyde på at elevene blir bedre til å se sammenhenger i faget av å ha mappevurdering, men at de blir dårligere på faktakunnskaper sammenlignet med elever som har hatt skoleprøver.

Videre er de andre lærerne usikre på om mappevurdering lar seg bruke etter kunstige intelligenser kom på banen. Anita, Theo og Cecilie sier at de ikke kan være helt sikre på om tekstene som elevene leverer har blitt skrevet av de selv eller om de har brukt for eksempel ChatGPT. De mener likevel det ikke er noe problem. De blir kjent med elevene og får oversikt over hvor langt de har kommet på oppgavene i skriveøktene på skolen. Det gir de en god indikasjon på hvilken kvalitet elevene klarer å opprettholde i mappetekstene. Cecilie sier at det kan hende elever slipper unna med å kopiere et par setninger fra en språkmodell, men at det ikke har noen innvirkning på hverken lærdom eller karakter.

Elevene får tid på skolen til å arbeide, de kan skrive om noe de engasjerer seg for, lærerne går rundt og gir veiledning og elevene kan levere to utkast. Derfor mener lærerne at det blir enklere for elevene å sette i gang med skrivningen og at elevene har mindre grunn til å bruke språkmodeller. Videre blir elevene spurt om å trekke inn egne erfaringer eller tanker, noe som gjør det vanskeligere å bruke kunstige intelligenser. Cecilie forteller at hun har hatt innleveringer med mer lukkede spørsmål før. Da har hun opplevd enkelte elevbesvarelser med 100 prosent avskrift. Dette har ikke vært tilfellet med mappevurdering, forteller hun.

Anita har et godt motargument til de som er skeptiske: *"Jeg tenker jo at vi kan ikke droppe å ha innleveringsoppgaver bare fordi det finnes gode språkmodeller etter hvert"*. Hun tenker at det istedenfor er viktig å ta en prat med elevene før de begynner med første mappeoppgave for å diskutere hva ChatGPT kan brukes til og hva den ikke bør

brukes til. Hun sier at språkmodeller er kommet for å bli og at det derfor er viktig at elevene er klar over hva de kan brukes til. Anita legger frem noen eksempler på hva ChatGPT kan brukes til. Elevene kan spørre ChatGPT om forslag til problemstillinger, om stikkord til hvordan det er lurt å strukturere en tekst eller om å få hjelp med å forstå en kilde på nett. Hun er tydelig til elevene om at språkmodeller ikke må brukes til å generere ferdig tekst som de kopierer inn i mappetekstene deres. Anita forteller at de også har brukt tid på å jobbe med hvordan elevene skal ha med språkmodeller i kildelisten dersom de har brukt slike. Elevene skal da ta med spørsmålene som de har stilt, noe som Anita sier at gir henne en pekepinn på hvor mye de i utgangspunktet kan om fagstoffet. Som en avslutning på spørsmålet sier Anita at *"så ja, jeg tenker at vi må fortsette med innleveringer selv om det er verktøy som finnes, og så må vi jobbe med å bruke de på en god måte"*.

Selv om noen av lærerne på skolen er skeptiske, sier Cecilie at det er mange andre lærere på skolen som bruker tilsvarende vurderingsformer i andre fag. Hun trekker frem norsk som et eksempel og sier at lærerne der har hatt mye fokus på prosesskriving og karakterdempning; to fokusområder som hun også ser i mappevurdering. Med prosesskriving mener hun at elevene skriver to utkast og får tilbakemelding fra det ene utkastet til det neste. Resultatet er at elevene kan bruke veiledningen til å lære av sine feil og videreutvikle sin kompetanse.

## Oppsummering

Oppsummert har lærerne opplevd lite endring i undervisningstimene etter de tok i bruk mappevurdering. Mappene ble i hovedsak brukt som et vurderingsverktøy, og det er stort sett kun skoletimene i forkant av innleveringsfristen på førsteutkastet som har hatt store endringer. Skriveøktene ble satt av til at elevene skulle få tid til å arbeide med og få veiledning på mappeoppgavene. Lærerne opplever veiledningsarbeidet i øktene som interessant, særlig fordi de får mer tid til å snakke med elevene og fordi de får spørsmål som de selv må lese seg opp på. Etter innlevering av førsteutkastet gir læreren tilbakemelding på de varierte besvarelsene. Selv om dette er tidkrevende hvis utkastene er lange, oppfordres elevene til å bruke god tid på førsteutkastene slik at de får størst mulig utbytte av tilbakemeldingene. Da blir dessuten vurderingsarbeidet tilknyttet andreutkastet redusert. Arbeidsmengden er generelt stor tilknyttet mappevurderingen,

men lærerne opplever arbeidet som mer spennende og dermed fordelaktig. Videre ser de tydelige reduksjoner i arbeidsmengden fra det første til året det andre, både ved planlegging, veiledning og vurdering av oppgavene.

Lærerne opplever at elevene blir bedre til å se sammenhenger i faget av å skrive mappetekster, men at de ikke er like gode på begreper som elever som har vanlige skoleprøver. De har valgt å prioritere at elevene blir i stand til å bearbeide fagstoff, se sammenhenger og formidle fagstoff på en god måte. Dette mener de mappevurdering kan bidra til. Ved bruk av mappevurdering er elevene aktive i læringssituasjonen, de får erfaring i å reflektere rundt egen læring ved å skrive refleksjonsnotater, de lærer i en mer kontinuerlig prosess, vurderingssituasjonen kan oppfattes som mindre stressende og de kan bli mer engasjerte og motiverte av å få oppgaver der de får mulighet til å skrive tekster om noe som de interesserer seg for.

Mappevurdering har med andre ord svært mange fordeler for elevene. En ulempe kan være tilgjengeligheten på språkmodeller, noe som gjør andre lærere på skolen skeptiske til denne vurderingsformen. Anita, Cecilie og Theo legger imidlertid frem argumenter for at det fremdeles er mulig å anvende mapper som et vurderingsverktøy. De uttrykker dessuten at språkmodellene kan utnyttes til å gi elevene veiledning underveis, altså at språkmodellen kan være en form for hjelpelærer.

Nedenfor blir funnene fra studien diskutert i lys av forskningslitteratur for å besvare studiens tre forskningsspørsmål.



## Diskusjon

I denne delen kommer jeg til å diskutere de viktigste funnene fra intervjuene opp mot teori og forskning. Jeg diskuterer først mappevurderingens påvirkning på læreren sin undervisnings- og vurderingspraksis. Dette er basert på Taube (2000) sin inndeling om at mappevurdering kan brukes både i undervisningen og som vurderingsredskap. Videre kommer jeg til å ta stilling til hvordan elevene påvirkes av mappene. Hovedvekten av diskusjonen er tilknyttet elevenes læringsutbytte, men deres opplevelse av stress kommer også til å bli berørt. Deretter ser jeg på løsninger som lærere kan ta i bruk for å ta hensyn til det faktum at elevene har enkel tilgang til språkmodeller. Til slutt vurderer jeg to tiltak for å redusere læreres arbeidsmengde og øke elevenes læringsutbytte tilknyttet mappevurdering: Bruk av responsgrupper og bruk av flere innleveringer.

### Hvordan påvirkes lærerens undervisningspraksis av mappevurdering?

Fra eksemplene av andre lærere i Norden som har brukt mappevurdering, særlig eksemplene av Taube (2000), formidles det at mappevurdering i stor grad ble brukt som en måte å undervise på. Mappearbeidet ble også vurdert, men en stor andel av tiden i skoletimene ble brukt til at elevene arbeidet med mappetekstene. Elevene skrev mange tekster underveis og plukket til slutt ut de tekstene som de var mest fornøyde med og som mappen skulle bestå av. På denne skolen var det annerledes. I intervjuene kom det nemlig frem at mange av skoletimene var mer eller mindre uendret etter de gikk over til bruk av mappevurdering. Alle tre lærerne sa seg enige i dette. Likevel har det trolig vært noen endringer, og i denne første delen av diskusjonen blir to aspekter trukket frem: 1) Undervisning i "vanlige" skoletimer og 2) veiledning i skriveøktene før innlevering av førsteutkastet.

#### *Undervisning i vanlige skoletimer*

I intervjuene sa samtlige lærere at de ikke hadde merket noen særlige endringer i undervisningen deres etter oppstart av mappeprosjektet. Dette kan skyldes min formulering av spørsmålet eller deres tolkning. Uansett vil mappevurderingen til en viss grad påvirke de ellers "vanlige" skoletimene, det vil si skoletimene frem til elevene arbeider med førsteutkastene på skolen. Nedenfor diskuterer jeg tre slike påvirkninger: Tidsbruk, type oppgaver og bruk av begrepstester.

Den ene endringen som til dels fremkommer i intervjuene er tilknyttet tidsbruk. Lærerne uttrykker at de får mindre tid til å gå gjennom pensum ettersom det settes av to uker med skoletimer til å arbeide med mappetekstene. Dette kan ses på som en ulempe, særlig fordi det allerede kan oppleves som travelt å komme gjennom alle kompetansemålene i læreplanen. Ingen av lærerne i studien sa noe om denne tidsmangelen under intervjuene, men det er noe jeg hører jevnlig fra lærere på andre skoler. Som et resultat sier lærerne at de kun har tid til å gjennomgå pensum én gang. Det blir altså ikke satt av tid til at læreren har repetisjonsøkter. Elever som er borte vil dermed gå glipp av gjennomgangen og må lære fagstoffet på egenhånd. I tillegg kan det hende elever glemmer det som har blitt gjennomgått eller at de mister oppmerksomheten enkelte ganger. Det vil dermed være negativt for elevene at pensumet ikke repeteres. På den andre siden er elevene aktive i læringssituasjonen og foretar en repetisjon på egenhånd når de arbeider med mappetekstene. Dette bidrar til læring (NOU 2014: 7). Istedenfor at læreren formidler repetisjonen, er det elevene som inntar en aktiv rolle ved å se tilbake på det som er aktuelt for tekstene de arbeider med. En annen forskjell er tilknyttet oppgavene elevene arbeider med i skoletimene. Særlig Cecilie uttrykte at elevene ikke arbeider med det vi kan kalle gjenkallingsspørsmål, for eksempel navnssetting av salter. Hun forklarer at elevene ikke får slike spørsmål i mappeoppgavene, og derfor nedprioriterer hun det i skoletimene. Elevene bruker istedenfor tid på å diskutere og forklare naturfag med hverandre og til å knytte fagstoff til eksempler. Læringsaktiviteten er gunstig for læring både fordi elevene samarbeider og kommuniserer og fordi de kobler læringsarbeidet til noe de har kjennskap til fra før (NOU 2014: 7). I tillegg forbereder diskusjonsoppgavene elevene på å skrive mappetekster siden det gir elevene erfaring i å arbeide på tilsvarende vis som når de skriver mappetekster. Dette skaper en overensstemmelse mellom det elevene gjør i skoletimene og det de blir vurdert etter i mappetekstene, noe som Taube (2000) beskriver ofte er tilfellet.

En siste side ved undervisningstimene jeg vil trekke frem, er bruk av begrepstester. Cecilie setter av én dobbeltime per tema til begrepstestene. Målet er at elevene skal bli bedre på begrepsforståelse ettersom dette var noe lærerne hadde oppdaget at elevene var svakere på enn tidligere skoleår. Valget kan dessuten forsvares ved hjelp av

dybdelæring. Utdanningsdirektoratet (2019b) sin definisjon på dybdelæring innebærer nemlig å gradvis utvikle kunnskap og varig forståelse av blant annet begreper og sammenhenger i fag. Kombinasjonen av begrepstester og mapper legger til rette for en slik gradvis utvikling. Elevene utvikler kunnskap rundt begreper i begrepstestene samtidig som de blir bedre til å se sammenhenger på tvers av naturfaget i mappeoppgavene. Cecilie forteller at begrepstestene kan være tester med flervalgsspørsmål som retter seg selv, at elevene kan spille inn en video på noen få minutter hvor de forklarer et utvalg begreper eller at elevene kan levere et kort dokument hvor de forklarer tre til fem begreper som de mener er relevante til det de skriver om i mappetekstene. Cecilie mener en slik gjennomføring resulterer i at vurderingene – begrepstester, korte videoer eller korte dokumenter – hverken blir arbeidskrevende for læreren eller stressende for elevene.

Mange av skoletimene er altså ganske like som for et par år siden. Elevene gjør fremdeles forsøk og leverer rapporter, læreren går fremdeles gjennom pensum og elevene arbeider med fagstoffet individuelt eller med andre elever. Dette var ikke tilfellet i eksemplene til Taube (2000), men der ble mapper mye brukt til å utvikle skrive- og leseferdigheter. Man kan lære naturfag av å skrive tekster, som Rivard (1994) beskriver, men det vil også være behov for andre læringsaktiviteter. Det kan likevel tyde på at mappevurderingen "skjuler seg i klasseromsluften" og at elevene er mer bevisst på vurderingen som kommer, slik Taube (2000) beskriver. Mot slutten av temaet, når andre klasser repeterer og forbereder seg til å ha kapitellprøve, kommer mappevurderingen mer fysisk til syne og elevene legger inn en siste innsats for å skrive ferdig utkastet. Denne prosessen diskuteres nedenfor.

### *Veiledning i skriveøktene*

Elevene får tid på skolen til å arbeide med mappetekstene. Dette foregår i de to siste ukene før frist for innlevering av førsteutkastene, og i prøveplanen er dette snakk om perioden når elevene ellers skulle hatt en prøve. En god andel av elevene har begynt å gjøre opp noen tanker og har begynt å skrive før dette tidspunktet, men som blant annet Theo uttrykker, er det tydelig hvilke elever som enda ikke har kommet i gang med mappeoppgavene. Nedenfor diskuteres tre aspekter rundt skriveøktene: 1) lærerens

påvirkning av at elevene har ulike evner i å komme i gang med skriveoppgavene, 2) lærerens rolle i klasserommet og 3) lærerens arbeidsmengde tilknyttet veiledning.

Vanskelighetene med å komme i gang kan skyldes at elevene har lite kjennskap med å sette seg ned og strukturere og skrive en lang tekst. Læreren må derfor være tilgjengelig for elevene og støtte de som synes det er krevende (Taube, 2000). Det kan bli utfordrende hvis en stor andel av klassen ikke har kommet skikkelig i gang. Derfor kan det være fornuftig, slik Taube (2000) foreslår, å ha en klassesamtale med elevene ved oppstart av mappeprosjektet hvor det tas opp sider ved mappevurderingen som skiller seg fra andre vurderinger. Da kan elevene bli mer oppmerksomme på forskjellene og det faktum at det forutsettes en grad av selvstendighet hos elevene. Å diskutere oppgaveteksten fra starten av kan dessuten hjelpe elevene med å ha mappetekstene i bakhodet underveis i skoletimene slik at de med jevne mellomrom kan tenke over om det de lærer er relevant for teksten de skal skrive. Cecilie forteller at hun jevnt og trutt hjelper elevene med denne tankeprosessen i skoletimene ved å hinte til områder som de kan skrive om.

Lærerens rolle i klasserommet er et annet punkt å belyse. Med mappevurdering sier ikke læreren at elevene skal gjøre oppgave én til åtte på side 34 i læreboka. De åpne oppgavetekstene resulterer i at elevene mer eller mindre selv velger hva de skal jobbe med, og eventuelt om de vil lage en video eller podkast istedenfor å skrive tekst. Istedenfor at læreren sier hva elevene skal gjøre, blir det til at læreren spør elevene hva de arbeider med (Taube, 2000). Læreren får dermed mindre kontroll og en helt annen posisjon i klasserommet. Taube (2000) uttrykker at denne rollefordelingen kan være vanskelig for læreren. Anita, Cecilie og Theo synes derimot ikke at dette var tilfellet. De kunne gå rundt og snakke med elevene og bruke tid til å bygge sterkere elevrelasjoner. Dette kan være særlig relevant når elevene går i første klasse på videregående skole og ikke har fått tid til å binde sterke bånd til noen lærere enda.

Et siste aspekt jeg vil diskutere er elevenes behov for veiledning og lærerens tilhørende arbeidsmengde. I intervjuet forteller lærerne at noen elever arbeider selvstendig og skriver mesteparten av mappeteksten uten behov for noe særlig veiledning. Andre elever er derimot mer usikre og liker å dobbeltsjekke valgene de tar med læreren. Noen spørsmål er raske å besvare, men andre er mer tidkrevende. Ettersom oppgavene er

åpne og elevene velger å skrive om noe de interesserer seg for, er det mulig at læreren får spørsmål som hen ikke vet svaret på. Samtlige lærere forteller at dette forekommer. Da har læreren to alternativer. Læreren kan enten sette seg ned og prøve å finne frem til et svar sammen med eleven. Da får læreren spørsmålet kjapt ut av tankene, men ulempen er at dette tar noe tid, noe som kan resultere i at lærere ikke får tid til å ta en runde rundt til alle elevene i løpet av skriveøkten. Læreren kan alternativt utsette spørsmålet til etter skoletimen, men da blir arbeidsdagen lenger. Samme dilemma utspringer når læreren får spørsmål fra elevene på ettermiddagene eller i helgene; skal læreren svare på spørsmålet på fritiden eller utsette det til neste naturfagstime? I intervjuene fremkommer det at Cecilie og Anita har ulike preferanser på dette området, men at begge er fornøyde med deres gjennomføring. Anita forklarer at hun bestemt velger å kun gi veiledning i skoletimene for å unngå å bli overarbeidet. Overarbeid er nemlig én av årsakene til at mange lærere forlater yrket, og er derfor noe å være ekstra oppmerksomme på (Jelstad, 2023; Molvær, 2023).

Etter skriveøktene leverer elevene utkastet sitt og får tilbakemelding fra læreren. Tilbakemeldingene brukes til å ferdigstille det endelige utkastet, som i sin tur blir vurdert. Nedenfor diskuterer jeg hvordan vurderingsarbeidet påvirker læreren. Ettersom veiledningen av førsteutkastene og vurderingen av andreutkastene foregår ved bruk av samme vurderingskriterier, kommer jeg til å behandle de to situasjonene samlet.

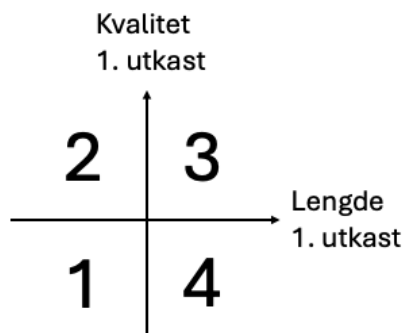
## Hvordan påvirkes lærerens vurderingspraksis av mappevurdering?

Mappevurdering gir læreren en ny måte å drive med vurderingsarbeid. Fra intervjuene er det tydelig at veiledning på førsteutkastet og vurdering av andreutkastet er det som er mest tidkrevende ved mappevurderingen. Theo uttrykker at arbeidsmengden hans avhenger av hvordan førsteutkastene ser ut, noe som dannet utgangspunkt for figur 1 nedenfor. Jeg diskuterer hvilke typer førsteutkast som minimerer lærerens totale arbeidsmengde. Til slutt bringer jeg frem andre faktorer med mappevurdering som har påvirket lærernes vurderingsarbeid, deriblant deres positive holdning til at elevbesvarelsene blir mer varierte og dermed mer interessante å vurdere.

### *Elevenes arbeidsprosess og lærerens arbeidsmengde*

Mye av det som tar tid med å vurdere mappetekstene, er at læreren både må gi veiledning på førsteutkastet og så på nytt lese gjennom og vurdere andreutkastet. Theo

forteller riktignok at det går raskere å vurdere andreutkastet hvis elevene har levert et tilnærmet ferdig produkt som førsteutkast. Dette skyldes at det har blitt gjort få endringer i teksten ettersom utkastet var såpass bra. Theo sin kommentar påpeker hvordan lærerens arbeidsmengde påvirkes av elevenes arbeidsprosess. Figur 1 nedenfor kan brukes til å vise denne sammenhengen. Horisontal akse viser lengden på førsteutkastet som elevene har levert mens vertikal akse viser kvaliteten. Med kvalitet menes i hvor stor grad eleven blir foreslått å gjøre større endringer i mappeteksten. God kvalitet vil altså si lite behov for endringer. Lærerne sa ikke eksplisitt at kvaliteten varierer i elevutkastene, men ved analysering av datamaterialet tolket jeg at dette er tilfellet. Det er dessuten vanlig at elevene i en klasse leverer tekster med ulik kvalitet.



Figur 1: Ulik status på førsteutkastene

Nede til venstre (1) har vi elevene som har skrevet et kort førsteutkast og hvor innholdet er svakt. Dette er gjerne elever som aldri kom skikkelig i gang med oppgaven og som fremdeles ikke har bestemt seg for tema eller problemstilling. Lærerne forteller at det går raskt å gi veiledning på utkastet, rett og slett fordi det er lite tekst å gi tilbakemelding på. Det tar derimot lang tid å vurdere det endelige utkastet siden læreren ikke har et førsteutkast å sammenligne med. Det samme gjelder elever som befinner seg oppe til venstre (2). De har også et kort utkast, men de har en plan for resten av teksten. Læreren kan raskt gi veiledning, men må bruke mer tid på vurdering av det endelige produktet.

Oppe til høyre (3) finner vi elevene som har et langt førsteutkast og hvor kvaliteten er så bra at det nærmest kunne vært et andreutkast. Slike elever har satt seg ned og brukt god tid på å skrive et utfyllende utkast. Læreren må bruke noe tid på å lese gjennom utkastet og gi veiledning, men eleven kommer ikke til å gjøre store endringer etterpå ettersom utkastet allerede var så pass bra. Det går derfor raskt å vurdere det endelige utkastet.

Nede til høyre (4) er elevene som har skrevet et langt førsteutkast, men hvor teksten av diverse årsaker har lav kvalitet. Det kan hende strukturen er dårlig, at eleven har brukt for avanserte kilder, at eleven ikke har svart på oppgaven eller av andre årsaker har et dårlig utkast. Læreren må først bruke mye tid på å gi veiledning på utkastet, men siden elevene må gjøre store endringer frem til det endelige utkastet, må læreren i praksis vurdere andreutkastet uten noe å ta utgangspunkt i. Resultatet er at læreren får mye arbeid tilknyttet begge utkastene.

Ved å ta utgangspunkt i lærerens arbeidsmengde vil det altså være best hvis elevene skriver et godt og langt førsteutkast. Tilsvarende vil være best for elevene sitt læringspotensial, noe som diskuteres senere. Det er med andre ord en vinn-vinn-situasjon at elevene bruker god tid på førsteutkastet. Læreren bør derfor legge til rette for at flest mulig elever ender opp med utkast som havner under denne kategorien. Et par slike tiltak diskuteres mot slutten i diskusjonsdelen, nemlig bruk av responsgrupper og bruk av hyppigere og flere innleveringer.

### *Andre påvirkninger mappene har på lærerens vurderingspraksis*

Lærerne uttrykker at de liker arbeidet med å se gjennom og kommentere utkastene. En klar gjenganger i alle intervjuene var at tekstene ble mer varierte, og at dette var noe samtlige foretrakk. Når elevene arbeider med åpne oppgaver, er det åpenbart at produktene de leverer inn vil variere i innhold. Den ene eleven kan ha skrevet om avokadoens reise rundt jorden og en annen elev kan ha skrevet om hvordan fredagstacoen kan bli mer bærekraftig. Det er eleven sin kreativitet som setter grensen. Videre har lærerne typisk hatt krav i oppgavene om å knytte inn personlige opplevelser, noe også forskningslitteratur foreslår (Ifelebuegu, 2023; Taasen et al., 2004; Taube, 2000). Dette fører til at elevene får en mer personlig tilknytning til mappene, og resultatet er en ytterligere spredning mellom tekstene. Selv om alle lærerne sier de foretrekker denne variasjonen når de skal vurdere besvarelsene, kan dette være svært individuelt. Cecilie sier også dette, og hevder at andre lærere på skolen kan foretrekke å vurdere elevbesvarelser som har mer til felles.

Til tross for at lærerne forteller at deres egen arbeidsmengde har økt av å bruke mappevurdering, foretrekker de likevel vurderingsarbeidet med mappevurdering fremfor andre alternativer. Dette er interessant og kan knyttes til Herzbergs motivasjonsteori

(Waage, 1985), blant annet ved at lærerne kan få økt motivasjon og økt trivsel av å se elevenes gode resultater. Anita sier hun nå bruker omtrent 50 prosent mer tid på vurdering. Hvis vi setter tall på det, kan vi si at Anita heller vil bruke 30 minutter på å vurdere en mappetekst enn å bruke 20 minutter på å vurdere en skriftlig prøve. Det betyr at hun er villig til å bruke 50 prosent mer tid på vurderingsarbeid dersom hun får mulighet til å lese gjennom varierte mappetekster enn mer monotone skoleprøver. Selv om lærerne ikke har noe imot den økte arbeidsmengden, kan arbeidsmengden være en avgjørende faktor som gjør at andre lærere unngår mappevurdering. Cecilie, Anita og Theo har allerede gjort grep for å redusere arbeidsmengden. Senere diskuteres flere muligheter som kan redusere lærers arbeidsmengde.

Vurderingsarbeidet har økt for lærerne, men Cecilie sier likevel at hun bruker mindre tid på å tenke på vurderinger fordi hun slipper å tenke gjennom om prøven er vanskelig nok, om den dekker kompetansemålene, om den er passelig lang og så videre. Med mappevurdering er det elevene selv som bestemmer hvor høyt taket er og hvor mye de vil gjøre ut av oppgaven. Lærerne må likevel passe på at oppgavetekstene er bra. De forteller at de brukte god tid på oppgaveutforming det første skoleåret for at de skulle bli av bra kvalitet, blant annet ved å formulere oppgavene slik at de trekker inn flere kompetansemål i samme oppgave. Dette bygger opp under forutsetningene som Rivard (1994) legger frem for at elever kan lære naturfag gjennom konseptet "writing to learn".

På den andre siden kan det være mer utfordrende å vurdere mappetekstene enn det er å vurdere en skriftlig prøve. Besvarelser i skoleprøver er gjerne korte og inneholder riktige eller feil svar. I mappetekstene må lærerne vurdere andre ferdigheter enn det som er direkte tilknyttet naturfag. Innholdet er selvsagt viktig, men det samme er blant annet struktur og bruk og oppføring av referanser. Det blir mer komplekse vurderingskriterier, noe som påvirker både lærer og elev. Taube (2000) påstår at dette munner ut i en noe upålitelig vurderingsform fordi ulike lærere kan vurdere samme mappetekst ulikt.

Samtidig foreslår Taube (2000) at elevene skal være med på å utarbeide vurderingskriteriene. Dette kan bidra til at elevene er mer klar over hvordan tekstene blir vurdert og hva som forventes av de, men samtidig kan det forsterke Taube (2000) sitt poeng om upålitelighet ettersom læreren sine vurderingskriterier dermed avhenger av



elevenes innspill. En løsning kan være at læreren har størst medvirkning på vurderingskriteriene, men at elevene kommer med forslag etter å ha diskutert i grupper.

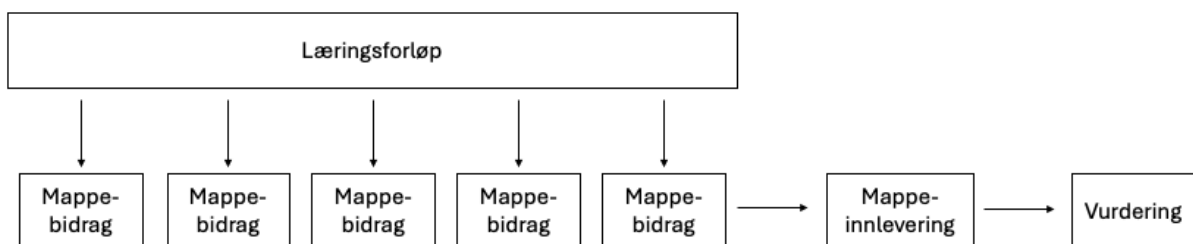
Mappene har i større grad påvirket lærerne i vurderingssammenheng enn de har i undervisningen. Vurderingsarbeidet har blitt annerledes både med tanke på hvordan det gjennomføres og arbeidsmengden det krever. Det kan likevel tyde på at lærerne har funnet en måte å integrere mappene på i skolehverdagen uten at de har medført en for stor omveltning for lærerne. Hvordan har elevene taklet den nye vurderingsformen, og hvilke kompetanser sitter de igjen med i ettertid? Dette blir diskutert nedenfor basert på datamaterialet fra lærerintervjuene.

## Hvordan påvirkes elevene av mappevurdering?

Lærerne har vært fornøyde med mappevurdering, men hvordan har elevene opplevd det? Skolen skal produsere mennesker med kompetanser som kreves i dagens samfunn og arbeidsliv. Nedenfor diskuteres hvorvidt mappevurdering kan bidra i prosessen. Jeg kommer først til å fortsette tråden som lærerne tok frem i intervjuene da de fortalte at mappevurdering legger til rette for at elevene får arbeide i en prosess. Til diskusjonen kommer jeg til å bruke figur 2 og 3 under. De er inspirert av Taasen et al. (2004, s. 49).



Figur 2: Tidsforløp tradisjonell vurdering



Figur 3: Tidsforløp mappevurdering

Figur 2 viser hvordan en typisk fremdriftsplan ser ut (Taasen et al., 2004). Først er det undervisning i skoletimene. Etter gjennomgang av pensumet har elevene en prøve som danner utgangspunkt for en vurderingssituasjon. Denne prosessen gjentar seg frem til skoleåret er ved sin slutt.

Med mappevurdering er prosessen annerledes, som vist i figur 3. Det er skoletimer som før, men elevene har fått oppgaveteksten ved oppstarten av et nytt tema. Elevene kan dermed arbeide med mappene kontinuerlig gjennom skoleåret, som lærerne forklarer. Dette skjer ved at læreren henter til eller kommer med forslag til idéer som elevene kan skrive om underveis, noe særlig Cecilie forteller at hun gjør. Videre gjør elevene forsøk og oppdager at forsøket kan knyttes inn i mappeteksten de skal skrive. I tillegg forteller Cecilie at de av og til har tid igjen etter et forsøk eller mot slutten av en skoletime, og at elevene typisk får den resterende tiden til å fortsette der de har kommet med mappeteksten. Gjennom slike handlinger får elevene stadig bragt frem mappeteksten og skrevet ned tankene sine. Selv om det ikke settes av tid i fremdriftsplanen til å arbeide med mappene ved starten av et nytt tema, får elevene likevel litt tid her og der. Videre arbeider elevene med førsteutkastet på skolen før innlevering og får veiledning og tilbakemelding av læreren. Samlet sett bringes altså mappetekstene frem mange ganger, enten rent fysisk eller på et kognitivt nivå. Dette samsvarer med studiene til Wittek (2003) om at studentene arbeidet mer jevnt og trutt når de hadde mappevurdering istedenfor avsluttende eksamen. Forskning har dessuten vist at å hente frem noe gjentatte ganger bidrar til å styrke hukommelsen (Pashler, 2007). Dette er ikke et nytt fenomen, og det er snarere allmennkunnskap at man husker bedre av å jobbe med noe gjentatte ganger. Det skyldes også at mennesket da bygger sterkere relasjoner mellom begreper og dermed er i stand til å bedre forstå begrepene. Prosessjobbingen kan være en forklaring på at Theo sine elever begynte å snakke med han om – og husket – mappetekstene de hadde skrevet året før da de hadde naturfag sammen.

En annen forklaring på at elevene var engasjerte i samtalene med Theo kan være at de fikk lov til å skrive om det som interesserte de aller mest. Med mappevurdering er oppgavene åpne og det legges opp til stor grad av valgfrihet. Elevene kan dermed lære naturfag samtidig som de skriver om noe som engasjerer dem. Oppgavene er formulert slik at elevene ikke skal gjengi tekst fra en lærebok, men at de istedenfor skal skrive en sammenhengende tekst og knytte inn flere forskjellige elementer. Derfor viser de et større spekter av elevens ferdigheter enn vanlige skoleprøver (Michaelsen & Johansen, 2007).

Lærerne forteller dessuten at de opplever at elevene ble bedre på å forstå sammenhenger på tvers av faget etter de startet med mappevurdering.

Mappeoppgavene legger til rette for dybdelæring fordi elevene da i større grad forstår naturfag og er i stand til å anvende det de har lært i praktiske kontekster (NOU 2014: 7). Motsetningen er overflatelæring der elevene pugges definisjoner eller prosedyrer (NOU 2014: 7). Dette skillet beskrives også av Taasen et al. (2004) som sier at hensikten med vurdering skiftes fra et testfokus til et læringsfokus når mappevurdering anvendes.

Selv om mappeoppgavene legger til rette for dybdelæring (NOU 2014: 7), har noen elever bedre forutsetninger for læring enn andre. Dette kan knyttes til figur 1 ovenfor. De elevene som leverer et kort utkast, altså 1 og 2 i figuren, får mindre tilbakemelding enn elevene som har levert lengre utkast. Elevene får også utbytte av tilbakemeldingene fra et kort utkast, men læreren kan ikke gi tilbakemelding og veiledning på noe elevene ikke har skrevet om enda. Både Wiliam (2011) og Ludvigsen-utvalget (NOU 2014: 7) presiserer tilbakemeldingens viktige rolle for elevers læring. Anita forteller at elevene blir oppfordret til å skrive lange utkast av denne årsak, men at noen elever likevel ikke tar det like seriøst ettersom de kan utsette arbeidet og heller levere alt eller mesteparten når det er innleveringsfrist på det andre utkastet. En slik holdning reduserer mulighetene elevene har til å få verdifulle tilbakemeldinger. Det resulterer dessuten i at elevene ikke arbeider like kontinuerlig som de bør gjøre. Samlet sett er det derfor best at elevene bruker god tid på førsteutkastene.

I mappene foretar elevene skrivevalg i tillegg til at de leverer refleksjonsnotater én gang i halvåret. Elevene må dessuten plukke ut de forsøksrapportene de vil skal inngå i mappen. Som Cecilie sier, velger noen elever de rapportene som krever minst arbeid. Andre elever velger de rapportene som viser kompetanse på andre fagområder enn det mappetekstene viser. Det er uklart om valgene elevene foretar underveis påvirkes av deres holdninger og kompetanser eller omvendt. Elever som velger "feil", for eksempel slik Cecilie beskriver, bør få tilbakemelding på at de har gjort et dårlig valg. Elevene kan dermed lære og utvikle metakognisjon og selvregulering (Michaelsen & Johansen, 2007).

En siste påvirkning som trekkes frem er tilknyttet stressnivået til elevene. Lærerne forteller at dette er en av hovedårsakene til at de tok i bruk mapper i naturfag. Elevene uttrykte også at mappevurdering er mindre stressende, noe som forsterker Taube (2000)

sine erfaringer. Elevene la frem flere forklaringer. De har blant annet to innleveringer på seg til å skrive mappetekstene og må derfor ikke prestere på første innlevering. Som beskrevet ovenfor, kan denne avslappede holdningen riktignok medbringe ulemper i det lange løp. Videre sier elevene at de liker at de får tid på skolen til å skrive. Lærerne sier at dette er avgjørende for at de skal få mulighet til å veilede elevene med førsteutkastet. I tillegg sier Anita og Theo at de ikke vil gi elevene for mye hjemmearbeid fordi elevene allerede har mye annet stress i skolen og hverdagen, og at dette er en annen årsak til at elevene må få tid på skolen til å skrive. Hverken elevene eller lærerne nevnte forutsigbarhet i intervjuene, med forutsigbare vurderinger kan være noe elever foretrekker og kan gjøre elevene mindre stresset (Hopfenbeck, 2006). Mappevurdering kan sies å være forutsigbar ettersom elevene både får utlevert oppgavene og vurderingskriteriene ved oppstarten av et nytt tema. Elevene som var negative til mappevurdering hevdet at skoleprøver er mer forutsigbare, men det stiller jeg meg uenig til. Elevene vet ikke på forhånd hvordan prøvespørsmålene ser ut, mens med mappevurdering spiller læreren i praksis med åpne kort.

Elever har lett tilgang til språkmodeller som ChatGPT. Internett kan slås av og skjermer kan kontrolleres under en skoleprøve for å eliminere muligheten til å få skrivehjelp fra språkmodellene. Dette er ikke mulig med mappevurdering. Nedenfor diskuteres hvilken innflytelse – både på godt og vondt – språkmodellene har på mappevurdering.

## Mappevurdering og språkmodeller

Det er klart at språkmodeller kan oppfattes som en trussel til mappevurdering. De blir bedre for hver uke som går, og teksten som genereres blir stadig mer troverdig. Hvordan kan skolen og lærere ta grep for å gjøre det mulig å fortsette med skriveoppgaver?

Anita fikk frem et godt poeng mot slutten av intervjuet. Hun fortalte at lærere ikke kan slutte med å ha innleveringsoppgaver fordi språkmodeller eksisterer. Dyr tilpasser seg endringer i klimaet og økosystemet, og på samme måte må lærere tilpasse seg endringer som skjer i skolen. Dette kan være store eller små endringer. Språkmodellenes innvirkning på skolen og elevene har riktignok vært enorm fordi ChatGPT er gratis og lett tilgjengelig. De evner dessuten å generere tekst som kan anses som noe elevene selv har skrevet (Rudolph et al., 2023). Videre er språkmodellene en trussel for konseptet *writing to learn*. Elevene lærer av å skrive tekster, men samtidig er det en risiko at elever

kan velge å la ChatGPT være forfatteren. Språkmodellene er kommet for å bli, som Anita sier. I de følgende avsnittene legger jeg frem noen grep som lærere kan ta i bruk for å tilpasse mappene til den nye språkmodell-hverdagen.

Selv om bruk av mappevurdering fører til at elevene kan arbeide med oppgavene utenom skolen og dermed få bistand fra språkmodeller, er samtidig mappevurdering et av forslagene til Ifelebuegu (2023) for å ta hensyn til språkmodellenes eksistens. Dette skyldes at lærere kan legge inn krav i oppgavene som resulterer i at det blir vanskeligere å generere ferdige tekstsvare for eksempel ChatGPT. Oppgavene kan kreve at elevene skal inkludere noe de har lært fra et av forsøkene de har gjennomført eller inkludere noe de har observert eller erfart fra hverdagen. Elevene kan også bli spurt om å komme med deres meninger og ha et kritisk blikk på det de skriver om, da dette kan være noe som språkmodellene foreløpig har vanskeligheter med å gjøre (Ifelebuegu, 2023).

Videre er det ikke slik at mappene kun må bestå av tekster. Elevene på skolen skriver allerede forsøksrapporter og refleksjonsnotatet som inngår i mappen. De kan i tillegg levere noe praktisk, som for eksempel å ta bilde av et bærekraftig måltid de har laget alene eller med familien og ha med forklaringer på hva det er som gjør måltidet bærekraftig. Et annet eksempel er å velge ut tre faktorer som påvirker livsstilen vår, for eksempel søvn, kosthold og stressnivå. Elevene loggfører de tre faktorene i løpet av en tidsperiode på én uke og har med en forklaring på hvordan deres livsstilsvaner påvirker helsen og hverdagen deres. Slike arbeidsoppgaver, som krever at elevene skal gjøre noe praktisk, minimerer mulighetene for å lene seg på språkmodellene.

Lærerne kan altså gjøre visse grep for å tilpasse oppgaveteksten slik at det blir vanskelig å bruke språkmodeller til å generere gode tekstsvare. Likevel blir språkmodellene mer intelligente for hver dag som passerer, så læreren havner i så fall i en kontinuerlig kamp mot maskinene. Det kan dessuten være problematisk å legge inn for mange krav i mappeoppgavene. Oppgavene blir da mer avgrenset, noe som reduserer elevenes muligheter til å skrive om det de virkelig interesserer seg for. Da kan de bli mindre engasjerte og mindre motiverte til å jobbe med oppgavene (Taube, 2000). Dette har videre en negativ effekt på elevenes forutsetninger for å lære (NOU 2014: 7).

En bedre løsning kan være å fokusere på de positive sidene ved språkmodellene. De kan være en god ressurs for elevene hvis de blir utnyttet på riktig vis (Chen et al., 2022;

Ifelebuegu, 2023; Rudolph et al., 2023). Anita hadde noen gode poeng i intervjuet angående dette, nemlig at språkmodellene kan brukes til å komme med idéer til hva elevene kan skrive om, til å formulere problemstillinger, til å forklare et avsnitt elevene ikke forstår eller til å hjelpe med å strukturere en tekst. Forslagene har likheter med de som fremkommer av forskningen til Chen et al. (2022) og Rudolph et al. (2023). Dersom en elev synes det er vanskelig å bestemme seg for noe konkret å skrive om, kan det være bra at språkmodeller gir forslag til eleven. Dette er dessuten noe lærerne på skolen uttrykker at de hjelper elevene med i tilfeller hvor elevene står fast. Språkmodellen kan dermed komme med forslag istedenfor at læreren gjør det og dermed bistå i lærerens arbeidsoppgaver.

Lærerne uttrykker nemlig at det kan ta lang tid å veilede elevene i skriveøktene som de får på skolen. Lærerne får liten tid per elev hvis 30 elever skal få veiledning i løpet av en dobbelttime. En mulighet kan være å overlate noe av veiledningen til språkmodellene, men at elevene i så fall får opplæring i hva språkmodellene kan hjelpe med. Denne tråden tas opp i neste avsnitt. Da får læreren mer tid til å hjelpe elevene med områder der språkmodellene er mindre pålitelige. Elevene kan også spørre læreren dersom de er skeptiske til rådene og svarene de får fra en språkmodell.

Videre kan det være viktig at læreren har kontroll på at elevene ikke lar språkmodellene ta for mange valg for dem. Det er mulig at språkmodellene kommer til å ha en stor rolle i arbeidslivet fremover og at vi kan støtte oss på de, men det vil uansett være avgjørende at elevene utvikler evne til å blant annet være kritiske og problemløsningsorienterte (NOU 2014: 7). Da fungerer det ikke å overlate all skrivingen til en datamaskin. En mulighet er å sette av et par skoletimer til at elevene i grupper diskuterer språkmodellenes fordeler og ulemper i skolen. Dette bør gjøres før elevene starter opp med mappearbeidet. Det kan gjøres innad i klassen, men det kan også gjennomføres for alle elevene på skolen, for eksempel ved å sette av et par timer i starten av skoleåret. Da får elevene opplæring i å bruke en språkmodell, på tilsvarende vis som en snekker får opplæring i å bruke en stikksag.

Et annet argument for at språkmodeller ikke setter en stopper for mappevurderingen, som samtlige lærere fortalte om i intervjuene, er at elevene får tid til å arbeide på skolen og i den sammenheng får veiledning av læreren. Da blir det enklere å komme i gang med

skrivningen og mindre behov for å ty til en språkmodell. I tillegg får læreren et overblikk over hva elevene skriver om, hvor langt de har kommet i løpet av skriveøktene og hvilken kvalitet det er på teksten. Læreren kan dermed oppdage om elevene plutselig har fått et mye lengre og bedre utkast ved innlevering enn de hadde i skoletimene.

Elevene skriver dessuten om noe som interesserer dem, noe som de tre lærerne hevder at gjør elevene mer engasjerte og motiverte til å arbeide med mappetekstene. Taube (2000) legger frem en lignende tanke. Cecilie tror det faktisk at elevene er engasjerte resulterer i at de i mindre grad bruker språkmodeller fordi de selv har lyst til å utforske og finne ut mer om det de skriver om. Hun følger opp med å si at hun har hatt innleveringsoppgaver med lukkede oppgavetekster, og der har hun fått inn tekster med 100 prosent avskrift. Dette har ikke vært tilfellet med mappeoppgavene.

Dersom læreren har prøvd å ta hensyn til språkmodeller og eventuelt andre former for skrivehjelp, men fremdeles opplever at elever leverer inn arbeid som ikke er deres eget arbeid, er en løsning å ha en kort vurderingssamtale med de aktuelle elevene. Læreren kan stille spørsmål til teksten eleven har levert og kontrollere om eleven faktisk har forstått det faglige innholdet. Dette kan ikke garantere at eleven ikke har brukt en språkmodell, men det gir læreren en god indikasjon. Det kan også tenkes at elevene kun har brukt språkmodeller til et par setninger i oppgaven deres, og at dette ikke blir fanget opp av læreren. Lærerne kommenterte dette kort i intervjuene og sa at det kan være tilfellet. Likevel har trolig ikke et par setninger en avgjørende effekt på hverken karakteren elevene ender opp med eller læringsutbyttet de får av å skrive en mappeoppgave, slik Cecilie forklarer. I intervjuet forteller Cecilie at hun ikke ser for seg at en av elevene hennes har gått opp en hel karakter av å muligens ha brukt en språkmodell til å generere et par setninger.

Forskning tyder på at språkmodeller ikke er en direkte trussel for mappevurdering, men at skolen som helhet bør gjøre grep for å gi elevene opplæring i bruk av en språkmodell (Ifelebuegu, 2023). Språkmodellene kan dermed brukes som et verktøy for læring og bli en samarbeidspartner for elevene istedenfor å være en skrivemaskin som gir elevene mulighet til å slippe unna skolearbeid.

Nedenfor samles trådene i to tiltak som kan integreres i mappevurderingen for å både redusere lærerens arbeidsmengde og øke elevenes læringsutbytte. Det ene er bruk av responsgrupper. Det andre er bruk av flere og hyppigere innleveringer.

## To tiltak for å redusere lærerens arbeidsmengde og øke læringsutbyttet

Responsgrupper er en vanlig del av mappevurdering ifølge Taube (2000), men det er et konsept lærerne i studien ikke har prøvd. Nedenfor diskuterer jeg hvordan responsgrupper kan integreres og hva de kan oppnå. Jeg diskuterer også bruk av flere innleveringer tilknyttet mappetekstene. Særlig sistnevnte bygger på figur 1 presentert tidligere om elevenes forskjellige status på førsteutkastene de leverer.

### *Bruk av responsgrupper*

Responsgrupper, som beskrevet av Taube (2000), er en læringsaktivitet hvor elevene lærer av hverandre ved å undersøke, diskutere og vurdere hverandres mappeoppgaver. Denne praksisen følger Vygotsky (1986) sin læringsteori, og samsvarer med Ludvigsen-utvalget (NOU 2014: 7) sine forutsetninger for læring. Aktiviteten gir elevene mulighet til å være aktive i og reflektere rundt vurderingsarbeidet, noe elevene har rett til (Forskrift til opplæringsloven, 2006, § 3-10). Responsgruppene kan variere i størrelse og elevene kan gå sammen i gruppene enten ved behov eller ved bestemte tidspunkter (Taube, 2000). I de påfølgende sidene diskuteres elevenes læringsutbytte av å bruke responsgrupper. Det diskuteres også i hvilken grad responsgrupper kan redusere lærerens arbeidsmengde, som Theo, Cecilie og Anita uttrykker er det mest negative med mappevurderingen.

Elever lærer mer når de selv er involvert i vurderingsarbeidet, blant annet fordi de får trening i å reflektere over egen læring (NOU 2014: 7; Utdanningsdirektoratet, 2022b). Dette kan føre til metakognisjon og selvregulert læring, og kan i neste omgang medføre at elevene får kompetanse i å lære (NOU 2014: 7). Det kan med andre ord tyde på at egenvurdering eller hverandrevurdering er en læringsfremmende aktivitet. Jeg har selv opplevd at dette ikke alltid er tilfellet. I matematikk på videregående har jeg erfart elever som gang på gang har fått i oppgave å vurdere sine egne prøver. Når jeg spør hva de fikk ut av det, har de generelt ikke noe bra svar. Det gir inntrykk av at elevene kun får beskjed om å sammenligne besvarelsen sin med et løsningsforslag og deretter sette en karakter på besvarelsen. Det kan tyde på at elevene ikke har fått opplæring i å vurdere eget arbeid



og at elevene derfor hverken vet hva de skal se etter i besvarelsene eller hvor de skal begynne. I slike tilfeller vil ikke læring finne sted, som Brøyn (2022) forklarer.

Ved bruk av responsgrupper i mappevurdering vil det derfor være avgjørende at elevene får tydelige beskjeder angående hva de skal gjøre. Elevene kan for eksempel gå sammen to og to i responsgrupper etter de har levert førsteutkastene. Det er nødvendig at dette skjer ganske snart etter innlevering, slik at de ikke glemmer tekstene de har skrevet. De leser først hverandres tekster og diskuterer innholdet: Hva er godt forklart? Hva er vanskelig å forstå? Er det noe som mangler i teksten, noe eleven savner? Deretter kan elevene gå mer i detalj og vurdere tekstene opp mot vurderingskriteriene som læreren bruker. Dette blir mer overkommelig dersom læreren og elevene på forhånd har sett på et par mappetekster og vurdert de i fellesskap. Da blir dessuten elevene mer bevisst på hvordan læreren vurderer tekstene. Elevene kan dessuten rangere tekstene etter kvalitet og gi begrunnelser etter Utdanningsdirektoratet (2022b) sin anbefaling. Det kan likevel være noe utfordrende for elevene de første gangene, men som Brøyn (2022) forteller, er elevvurdering en læringsaktivitet elevene ikke kan mestre fra dag én; det krever øving.

Studier viser at det kan være mer lærerikt å gi kommentarer enn å få kommentarer (Baker & Lundstrom, 2009; Cho & Cho, 2011). Derfor er det fordelaktig at elevene tar jobben i responsgruppene seriøst. Dermed kan elevene både lære av kommentarene de selv har fått og lære av kommentarene de har gitt til sine medelever. De kan bruke kommentarene de fikk til å direkte forbedre utkastet sitt. Videre kan de lære av styrkene og svakhetene til medelevens utkast, og bruke lærdommen fra dette i sin egen tekst. Elevene kan dessuten få inspirasjon av hverandres tekster, slik Taube (2000) forteller fra skolen i Finland.

Dersom responsgruppene fungerer bra, kan elevene gjøre endringer på utkastet basert på tilbakemeldingene de har fått og deretter levere et nytt utkast som læreren gir ytterligere tilbakemelding på. Da er utkastet som læreren får inn forbedret, og læreren finner færre elementer i teksten som trenger konstruktiv tilbakemelding. Et annet alternativ kan være at elevene gir hverandre tilbakemelding i et word-dokument uten å gjøre endringer etterpå. Læreren kan da bruke samme word-dokument og gi tilbakemelding på det som ikke allerede har blitt kommentert, eller eventuelt å kommentere en elevkommentar som læreren er uenig i. Resultatet av de to alternativene

er at elevene totalt sett får mer veiledning fordi de i tillegg får veiledning fra medelever. Dette er gunstig for videre læring (NOU 2014: 7; Wiliam, 2011). I tillegg får læreren mindre vurderingsarbeid hengende over seg fordi elevene bidrar i arbeidet.

Figur 1 viser at elevene på skolen leverer utkast med ulik lengde og kvalitet. I responsgruppene kan det derfor være lurt å la elever som har levert lignende utkast gå i lag. Det er potensielt mye lærdom i vente for de som har utfyllende utkast. Gruppene med korte utkast har ikke mulighet til å gi noe særlig respons til hverandre. I slike elevgrupper kan elevene hjelpe hverandre med å komme videre med skriveingen og bruke resten av tiden på å fortsette skrivearbeidet. Det er også mulig at elevene skriver mer på utkastene når de vet at deres medelever skal lese det de har skrevet. Jeg har selv erfart dette på universitetet. Videre kan det være en fordel at elever som har skrevet mappetekster om samme tema samarbeider, spesielt de første gangene. Da blir det enklere for elevene å sette seg inn i medelevens tekst og vurdere den. Med mer erfaring er det mulig at elevene også kan vurdere tekster som omhandler andre temaer. En tredje faktor å ta hensyn til – som Utdanningsdirektoratet (2022b) trekker frem som en forutsetning for at hverandrevurdering skal fremme læring – er klassemiljø. Elevene må være komfortable med at de gir tilbakemelding på hverandres tekster.

På universitetsnivå kan studenter få i oppgave å gi tilbakemelding på hverandres tekster. Dette skjer blant annet i stor grad på juridisk fakultet ved Universitetet i Bergen. Der gir studentene grundig, skriftlig tilbakemelding på to andre studenttekster og får selv tilbakemelding fra to medstudenter. Etterpå er det obligatorisk seminar på universitetet hvor studentene diskuterer tekstene og arbeider videre med dem. Denne prosedyren gjennomføres omkring ti ganger per halvår. Jeg har fått inntrykk av at prosessen oppleves som lærerik, både i form av å gi og å få tilbakemeldinger. Dette samsvarer med tidligere forskning (Baker & Lundstrom, 2009; Cho & Cho, 2011; Wiliam, 2011).

Det kan stilles spørsmål til om videregående elever er voksne nok til å gi gode tilbakemeldinger til sine medelever. Situasjonen beskrevet ovenfor viser imidlertid at studenter fint mestrer slike arbeidsoppgaver. Ved studiestart er studenter ofte 19 år og har nylig fullført videregående skole. Ettersom studenter evner å gi konstruktive og informative tilbakemeldinger, er det høyst mulig at elever på videregående trinn – hvor de kun er noen få år yngre – også kan gi tilbakemeldinger av god kvalitet. Dessuten gir det

elevene erfaring i å vurdere andres arbeid slik at de er bedre rustet hvis de tar høyere utdanning og får slike arbeidsoppgaver der.

Responsgrupper kan fremme læring på ulike vis og har potensiale for å redusere lærerens arbeidsmengde. Nedenfor diskuterer jeg et annet tiltak som kan vurderes for å redusere arbeidsmengden og fremme læring, nemlig bruk av flere og hyppigere innleveringer.

### *Flere og hyppigere innleveringer*

Et tiltak som kan redusere arbeidsmengden er sannsynligvis å øke antall innleveringer. Nedenfor vurderes hvordan dette kan oppnås gjennom følgende forløp: 1) innlevering av problemstilling, 2) innlevering av disposisjon, 3) innlevering av førsteutkast og til slutt 4) innlevering av andreutkast. Tanken oppsto delvis fra intervjuene der det ble sagt at elever som lager video eller podkast leverer en disposisjon som førsteutkast.

Foreløpig leverer elevene et førsteutkast helt mot slutten av temaet de arbeider med. Deretter leverer de andreutkastet mot slutten av terminen. Det kan være mindre stressende for elevene å ha få innleveringer. Elevene har da færre arbeidskrav å gjøre i naturfag og får mer tid til andre skolefag og fritidsaktiviteter. Dette – at mappevurdering oppfattes som mindre stressende enn alternative vurderingsformer – er også noe lærerne verdsetter, og er en av hovedgrunnene til at de i det hele tatt tok i bruk mappevurdering. Derfor bør ikke de ekstra innleveringene legge flere krav på elevenes skuldre. Samtidig vil elevene bli lite påvirket hvis innleveringene er små og nærmest kan gjøres i skoletimene. Bruk av flere innleveringer har noen likheter med tiltaket "ukas arbeid til mappen" som læreren i Finland hadde satt i spill (Taube, 2000). Der hjalp læreren elevene med å komme i gang med å vurdere eget arbeid. På noe tilsvarende vis kan flere innleveringer hjelpe elevene med å komme i gang med skrivearbeidet.

Det fremkom i intervjuene at det er stor variasjon med tanke på førsteutkastene som elevene leverer. Dette er tidligere vist i figur 1. Variasjonen kan skyldes at elever er forskjellige og har ulike arbeidsvaner. Noen elever er tidlig ute med mappeoppgavene og har allerede begynt å skrive før skriveøktene på skolen. Elevene kan dermed bruke veiledningen i skriveøktene til å forbedre utkastet deres. Andre elever utsetter skrivingen siden innleveringsfristen er langt frem i tid. En konsekvens er at slike elever ikke kan få

veiledning i skriveøktene til å forbedre utkastet deres fordi de i stedet for må få veiledning med å i det hele tatt komme i gang.

Som beskrevet tidligere i forbindelse med figur 1, vil både læreren og elevene ha fordel av at elevene skriver lange førsteutkast med høy kvalitet. Læreren kan da gi tilbakemelding på førsteutkastet og har relativt lite vurderingsarbeid å gjøre på andreutkastet ettersom det blir små endringer mellom de to utkastene. Samtidig får elevene størst læringsutbytte av å bruke god tid på førsteutkastet siden det forutsetter at elevene arbeider mer jevnt og trutt med mappetekstene. Derfor kan det være lurt å prøve å få de aller fleste elevene til å levere slike gode førsteutkast.

Elevene kan få bedre førsteutkast av at læreren legger inn flere innleveringsoppgaver eller arbeidskrav tilknyttet mappetekstene slik at elevene må arbeide mer kontinuerlig. Det første arbeidskravet kan være at elevene skal levere et skriftlig forslag til en problemstilling. Dette bør skje tidlig, men samtidig ikke så tidlig at klassen nettopp startet med gjennomgangen av det nye temaet og elevene dermed har få alternativer å velge mellom. Når læreren får inn problemstillingene, vet hen at elevene har begynt arbeidet og har bestemt seg for det de ønsker å skrive om. Det gir dessuten læreren mulighet til å gi tidlig beskjed til elevene om problemstillinger som må omformuleres eller erstattes.

Det neste arbeidskravet kan være å levere inn en kort disposisjon – gjerne i stikkordformat – over hva førsteutkastet skal inneholde. Her kan relevante begreper inkluderes, et par eksempler eleven ser for seg å trekke inn og muligens noen få kilder. Disposisjonen bør vise at elevene har begynt å tenke gjennom hvordan de skal strukturere oppgaven. Læreren må bruke noe tid på å lese gjennom disposisjonene, men dette går raskt dersom de er i stikkordformat. Resultatene er at elevene blir presset til å komme i gang tidligere med skriveprosessen og at læreren får oversikt over at alle elevene er på riktig spor.

Det tredje arbeidskravet kan være at elevene leverer et førsteutkast, som før. Elevene har forhåpentligvis kommet langt på vei i skriveøktene av å bygge videre på disposisjonen, og utkastet de leverer blir sannsynligvis bra. Til slutt leverer elevene andreutkastet mot slutten av terminen etter å ha fått tilbakemeldinger fra læreren og eventuelt medelever.

Et slikt innleveringssystem medbringer flere arbeidsoppgaver for elevene og flere elementer som læreren må vurdere. Selv om læreren får flere innleveringer å vurdere, kan det totalt sett redusere arbeidsbelastningen. Dette skyldes at læreren får mer oversikt over elevenes skriveprosess og dermed kan unngå at elevene leverer lange førsteutkast som ikke har særlig bra kvalitet. Elevene får også flere oppgaver, men samtidig arbeider de mer kontinuerlig og får mer tilbakemelding underveis. Totalt sett får elevene muligheter for et enda bedre læringsutbytte (William, 2011).

En ulempe med innleveringssystemet er at elevene får færre områder de kan skrive om ettersom de må starte arbeidsprosessen tidligere enn de muligens ellers hadde gjort. Elevene kan bli mindre motiverte til å arbeide med oppgavene hvis de må velge mellom alternativer som ikke interesserer de (Taube, 2000). Samtidig får elevene mer tilbakemelding, mer kontinuitet i arbeidsprosessen og bedre forutsetninger for læring av det foreslåtte innleveringssystemet (NOU 2014: 7; Pashler, 2007; William, 2011). Dette er en avveiningssituasjon som må tas hensyn til. En mulighet kan være å starte et nytt tema med å vise frem forskjellige elementer som kommer til å bli trukket frem i de kommende ukene slik at elevene får oversikt over hva de kan velge å skrive om.

## Konklusjon

Mappevurdering er en relativt ny vurderingsform som blant annet fremmer dybdelæring, noe den nye læreplanen bærer preg av. I studien har jeg lagt frem både fordelene og ulempene med mappevurdering fra lærerens perspektiv. Forskning og lærernes erfaringer tyder på at fordelene er mange mens ulempene er beskjedne. Den største ulempen er tilknyttet lærerens arbeidsmengde, men dette er noe som kan tas hensyn til.

Som fremtidig lærer er mappevurdering en vurderingsform jeg kommer til å anvende. Vurderingsformer tester andre kunnskaper enn mange andre vurderingsformer, og den kan derfor brukes ved jevne mellomrom for å gi elevene en variert vurderingsopplevelse.

Studien har bidratt med å gi svar på spørsmål, men den har også resultert i ubesvarte spørsmål. I dette siste avsnittet presenteres tre forslag til videre forskning. En forskningsidé er å undersøke hvordan elevtekstene utvikler seg gjennom et skoleår, fra første til siste mappeoppgave. Erfaringer fra Taube (2000) og konseptet med *writing to learn* (Rivard, 1994) tilsier at elever utvikler skrive- og fagkunnskaper av å skrive tekster, men det hadde være interessant å utføre grundig forskning på utviklingene som finner sted i mappetekstene. En annen idé for videre forskning er å undersøke hvordan oppgaveteksten i mappeoppgavene påvirker elevenes læringsmuligheter, og dermed gi innsikt til lærere om hvordan mappeoppgavene bør utformes. En tredje mulighet er å forske på responsgrupper på videregående trinn og studere hvordan lærere kan legge til rette for at gruppene skal fremme læring og samtidig redusere lærerens arbeidsmengde. Det finnes forskning som tilsier at det fungerer bra på universitetsnivå og på barnetrinnet, men det har blitt gjort lite forskning på videregående skoler.

## Referanser

- Andersen, G. (2020, 16. april). *Kvalitative intervjuundersøkelser*. NDLA. <https://ndla.no/nb/subject:1:9bb7b427-3f5b-4c45-9719-efc509f3d9cc/topic:1:432baee9-5671-47ce-870e-48b8fc3b7a42/topic:1:1db7bf3c-3a7b-44af-b632-e3c5ff2a999e/resource:201ce19e-7011-49a6-b415-91fd42d5dfe9>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31. <https://doi.org/https://doi.org/10.1007/s11092-008-9068-5>
- Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3-26. <https://doi.org/10.1037/qup0000196>
- Brøyn, T. (2022, 13. april). *Egenvurdering – på godt og vondt*. Bedre Skole. <https://www.utdanningsnytt.no/bedre-skole-fagartikkel-laering/egenvurdering-pa-godt-og-vondt/317237>
- Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2022). Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success. *Information Systems Frontiers*, 25, 161-182. <https://doi.org/https://doi.org/10.1007/s10796-022-10291-4>
- Cho, Y. H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science*, 39, 629-643. <https://doi.org/https://doi.org/10.1007/s11251-010-9146-1>
- Cohen, L., Morrison, K., & Manion, L. (2011). *Research Methods in Education*. Routledge. <https://web.p.ebscohost.com/ehost/detail/detail?vid=0&sid=873684bf-3741-49ed-907e-3b4cd5b82ff0%40redis&bdata=JnNpdGU9ZWwhvc3QtbGl2ZQ%3d%3d#AN=548475&db=nlebk>
- Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory* (4. utg.). SAGE Publications.
- Ellingsen, R., & Johnsgård, C. (2024, 16. april). *Læreren Asgil har 4,2 minutter til hver elev i uka*. NRK. [https://www.nrk.no/tromsogfinnmark/undersokelse-fra-utdanningsforbundet\\_-laerer-har-fatt-flere-oppgaver-og-ansvar-1.16836399](https://www.nrk.no/tromsogfinnmark/undersokelse-fra-utdanningsforbundet_-laerer-har-fatt-flere-oppgaver-og-ansvar-1.16836399)
- European Commission, Directorate-General for Education, Youth, Sport and Culture, Udave, J., Carlo, A., Valette, S. (2013). *Study on policy measures to improve the attractiveness of the teaching profession in Europe: final report. Volume 1*, Publications Office. <https://data.europa.eu/doi/10.2766/40827>
- Firestone, W. A. (1993). Alternative Arguments for Generalizing from Data as Applied to Qualitative Research. *Educational Researcher*, 22(4), 16-23. <https://doi.org/10.2307/1177100>

Forskrift til opplæringslova. (2006). *Forskrift til opplæringsloven*. (FOR-2006-06-23-724). Lovdata. <https://lovdata.no/forskrift/2006-06-23-724>

Gjesholm, J., & Røren, E. M. M. (2024). *Videregående elevers opplevelse av mappevurdering i naturfag*. [Masteroppgave]. Universitetet i Bergen.

Hopfenbeck, T. N. (2016). *Å lykkes med elevvurdering*. Fagbokforlaget.

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9), 1277-1288. <https://doi.org/10.1177/1049732305276687>

Ifelebuegu, A. O. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning & Teaching*, 6(2), 385-392. <https://doi.org/https://doi.org/10.37074/jalt.2023.6.2.2>

Jelstad, J. (2023, 15. juni). *4 av 10 lærere tenker ofte på å slutte i jobben: Etter ni år i klasserommet takket Hilde for seg*. Utdanningsnytt. <https://www.utdanningsnytt.no/ks-laerernonnsoppgjor/4-av-10-laerere-tenker-ofte-pa-a-slutte-i-jobben-etter-ni-ar-i-klasserommet-takket-hilde-for-seg/352314>

Kurihara, N. (2017). Do Peer Reviews Help Improve Student Writing Abilities in an EFL High School Classroom? *TESOL Journal*, 8(2), 450-470. <https://doi.org/https://doi.org/10.1002/tesj.282>

Kvale, S., & Brinkmann, S. (2015). *Det kvalitative forskningsintervju*. Gyldendal.

Lam, R. (2018). *Portfolio Assessment for the Teaching and Learning of Writing*. SpringerBriefs in Education. <https://doi.org/10.1007/978-981-13-1174-1>

Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30-43. <https://doi.org/https://doi.org/10.1016/j.jslw.2008.06.002>

Maxwell, J. A. (1992). Understanding and Validity in Qualitative Research. *Harvard Educational Review*, 62(3), 279-300. <https://www.proquest.com/scholarly-journals/understanding-validity-qualitative-research/docview/212250067/se-2>

Mejlbo, K. (2024a, 13. mars). *Lærere får stadig flere oppgaver – mister motivasjonen*. Utdanningsnytt. <https://www.utdanningsnytt.no/helsesykepleier-laerer-laererollen/laerere-far-stadig-flere-oppgaver-mister-motivasjonen/393434>

Mejlbo, K. (2024b, 13. mars). *UDF-lederen frykter økt lærerflukt*. Utdanningsnytt. <https://www.utdanningsnytt.no/arbeidstid-arbeidstidsavtale-laerer/udf-lederen-frykter-okt-laererflukt/393555>

Nilssen, V. (2014). *Analyse i kvalitative studier. Den skrivende forskeren*. Universitetsforlaget.



- NOU 2014: 7. (2014). *Elevenes læring i fremtidens skole – Et kunnskapsgrunnlag*. <https://www.regjeringen.no/no/dokumenter/NOU-2014-7/id766593/sec1?q=2014>
- NOU 2015: 8. (2015). *Fremtidens skole – Fornyelse av fag og kompetanser*. <https://www.regjeringen.no/no/dokumenter/nou-2015-8/id2417001/>
- OECD (2018). *Teaching for the Future: Effective Classroom Practices to Transform Education*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264293243-en>
- OECD (2019). *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, OECD Publishing, Paris. <https://doi.org/10.1787/1d0bc92a-en>
- Opplæringsloven. (1998). *Lov om grunnskolen og den vidaregåande opplæringa* (LOV-1998-07-17-61). Lovdata. <https://lovdata.no/lov/1998-07-17-61>
- Pashler, H., Rohrer, D., Cepeda, N.J. et al. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review* 14, 187–193. <https://doi.org/10.3758/BF03194050>
- Rivard, L. P. (1994). A Review of Writing to Learn in Science: Implications for Practice and Research. *Journal of Research in Science Teaching*, 31(9), 969-983. <https://doi.org/10.1002/tea.3660310910>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Taasen, I., Havnes, A., & Lauvås, P. (2004). *Mappevurdering – av og for læring. Med eksempler fra helse- og sosialfag*. Gyldendal Akademisk.
- Taube, K. (2000). *Mappevurdering. Undervisningsstrategi og vurderingsredskap*. Tano Aschehoug.
- Thagaard, T. (2018). *Systematikk og innlevelse*. Fagbokforlaget. [https://issuu.com/fagbokforlaget/docs/systematikk\\_og\\_innlevelse\\_utdrag](https://issuu.com/fagbokforlaget/docs/systematikk_og_innlevelse_utdrag)
- UiB. (2023, 24. oktober). *Personvern i forskning ved UiB*. UiB. <https://www.uib.no/forskningsetikk/129857/personvern-i-forskning-ved-uib#oversikt-og-dokumentasjon-av-forsknings-og-studentprosjekter-ved-uib-rette>
- Utdanningsdirektoratet. (2019a, 18. november). *Hva er nytt i læreplanverket?* <https://www.udir.no/laring-og-trivsel/lareplanverket/stotte/hva-er-nytt-i-lareplanverket/>
- Utdanningsdirektoratet. (2019b, 13. mars). *Dybdelæring*. <https://www.udir.no/laring-og-trivsel/dybdelaring/>
- Utdanningsdirektoratet. (2022a, 4. februar). *Underveisvurdering*. <https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/underveisvurdering/>

Utdanningsdirektoratet. (2022b, 11. mai). *Involver elever og lærlinger i vurderingsarbeidet*. <https://www.udir.no/laring-og-trivsel/vurdering/underveisvurdering/involvering/#a177796>

Utdanningsforbundet. (2024). *Lærernes kjernekompetanse*. <https://www.utdanningsforbundet.no/var-politikk/utdanningsforbundet-mener/artikler/larerutdannede-larere-til-alle-elever/dette-kan-en-larer/>

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans.). MIT Press.

William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14. <https://doi.org/10.1016/j.stueduc.2011.03.001>

Wittek, L. (2003). *Mapper som lærings- og vurderingsform. Eksempler fra Universitetet i Oslo*. Unipub. <https://www.nb.no/items/bae15a03e445f01822498032f42279b7>

Waage, J. F. (1985). *Arbeidspsykologi*. Cappelen Damm.

# Vedlegg

## Vedlegg 1: Intervjuguide for Cecilie, Theo og Anita

### Bakgrunnsopplysninger

- Hvilken utdanning har du?
- Hvilke fag underviser du?
  - Hvor lenge har du undervist i fagene?
- Har du erfaring med mappevurdering fra før?
  - Hvilke? (*Hvor startet idéen?*)
  - Hatt det selv som elev?
  - Bruker det nå?

### Planlegging av mappene

Kan du fortelle hvordan du går frem for å planlegge mappeoppgavene?

- Designing av mappeoppgaver
- Opp mot kompetansemål og undervisning
- Hvordan ser du for deg å støtte elevene underveis?

### Gjennomføring

Hvordan foregår gjennomføringen av mappene?

- Arbeidsmengde for deg?
  - Omfang
  - Til planlegging, gjennomføring og evaluering
  - Sammenligning med vanlig undervisning og vanlige prøver
- Noen endringer i undervisningen?
- Hvordan følge med i elevenes prosess? Utfordrende?
- Samarbeid mellom elever eller bruk av internett?
  - Bruk av ChatGPT? Hvordan ta hensyn til dette?
- Hvordan gi tilbakemeldinger?
- Hvordan vurdere besvarelser og sette karakter?
  - Noen utfordringer?

## Resultater

Hvilke forskjeller har du lagt merke til etter du tok i bruk mappevurdering?

- Områder hvor elever utviklet seg (læringsutbytte, motivasjon, engasjement, arbeidsvaner)
- Noen forskjeller sammenlignet med vanlige prøver?
  - La du merke til en endret karakterfordeling på mappene sammenlignet med vanlige prøver?
- Hva synes elevene om metoden? (tilfredshet, rettferdighet, lærte de nok)
- Du har brukt mappevurdering i naturfag. Kan du se for deg å også bruke det i andre fag? Hvorfor/Hvorfor ikke?

Dette utdraget er hentet fra et refleksjonsnotat som en av elevene har skrevet:

*"En negativ ting med mappevurdering er at man kun går grundig gjennom det temaet man skriver/snakker om. Andre temaer som er i pensum lærer man ikke like godt, og noen rekker man ikke å gå gjennom i det hele tatt."*

Hva tenker du om dette?

## Konklusjon

- Hva vil du si at du er mest fornøyd med angående hvordan mappevurderingen gjennomføres per i dag?
- Tenker du at det er noe rom for forbedring? Hvor?
  - Knytte til eksamensresultater, forsto sammenhenger, ikke detaljer
- Tenker du å fortsette med bruk av mappevurdering? Hvorfor/Hvorfor ikke?

## Vedlegg 2: Intervjuguide for oppfølgingsintervju med Cecilie

### Oppfølgingsspørsmål

- Er det noen forskjeller med tanke på valg av oppgaveform (tekstoppgave eller video/podkast)?
  - Angående veiledning
  - Angående vurdering
- Når er innleveringsfristen på 2. utkastet?
  - Hva var den i fjor?
- Har du fått prøvd å bruke begrepstester?
  - Hvordan gikk det?
- Det virket som om noen lærere valgte å ikke bruke karakterer på mappeoppgavene i fjor. Hvordan har du gjort dette? Og hvorfor gjorde du det slik?
- Hvordan opplever du mappevurdering?
  - Arbeidsmengde
  - Hva sier/mener andre naturfagslærere?
  - Noe annet?