



Uit

NORGES  
ARKTISKE  
UNIVERSITET

Institutt for lærerutdanning og pedagogikk

## Hvordan vurderer lærerne praktisk arbeid i naturfag?

*En kvantitativ analyse av lærernes formative vurderingspraksis av praktisk arbeid i naturfagundervisningen.*

—

**Karoline Belt**

*Master i lærerutdanning 5.-10.trinn*

*LRU – 3908 Naturfagdidaktikk*

*Mai 2017*







## Sammendrag

Denne masteroppgaven er en kvantitativ studie som er gjennomført blant naturfaglærere i grunnskolen. Forskningsspørsmålet er: ”På hvilken måte og i hvor stor grad bruker naturfaglærere formativ vurdering i den praktiske delen av naturfagundervisningen i grunnskolen?”. Jeg har også med følgende tilleggsspørsmål: ”Påvirkes lærernes vurderingspraksis av (1) undervisningserfaringen i faget; (2) klassetrinn det undervises på og (3) antall studiepoeng i naturfag?”.

Oppgaven bygger på en deduktiv tilnærming hvor jeg gjennom ulike teorier rundt formativ vurdering har utformet en spørreundersøkelse. Undersøkelsen ble gjennomført blant naturfaglærere som deltok på en konferanse i Tromsø og til sammen 61 lærere responderte på spørreskjema.

God formativ vurdering skal besvare tre viktige spørsmål: (1) ”Hvor skal eleven?”, (2) ”Hvor er eleven nå i forhold til målene?” og (3) ”Hvordan kan eleven jobbe videre for å oppnå ønsket mål?” (Hattie & Timperley, 2007). Resultatene viste at mange lærere i stor grad bruker ulike aspekter knyttet til formativ vurdering i praktisk arbeid i naturfag. Særlig viste lærerne seg å ha fokus på læringsmål, som knytter seg mot det første formative spørsmålet. I tillegg er elevenes rolle viktig i formativ vurdering (Black, Harrison, Lee, Marshall & Wiliam, 2003; Wiliam, 2011). Her viste det seg at de fleste lærerne i liten grad la til rette for elevmedvirkning i vurderingsarbeidet i praktisk arbeid i naturfag.

Resultatene i denne undersøkelsen kan tyde på at lærere med lengre undervisningserfaring i større grad bruker formativ vurdering i praktisk arbeid i naturfag. Det viste seg at lærere på småtrinnet (1.-4. trinn) skåret signifikant lavere på vurderingspraksis enn lærere på mellomtrinnet og ungdomstrinnet. Det ble ikke gjort tydelige funn på variasjoner i vurderingspraksis for lærere med mellom 0 og 60 studiepoeng i naturfag.



## Forord

Denne masteroppgaven markerer avslutningen på integrert master i lærerutdanning 5.-10. trinn ved UiT. Det har vært noen fine og spennende år. Takk til medstudenter for et godt studiemiljø.

Arbeidet med masteroppgaven har vært svært lærerikt på flere måter. Jeg har lært mye om prosessen rundt å skrive en masteroppgave. I tillegg har jeg lært svært mye om vurdering, et tema jeg også ønsker å ha fokus på når jeg begynner å jobbe som lærer. I alt har arbeidet med oppgaven vært en positiv opplevelse som jeg vil ta med meg videre i yrkeslivet.

Takk til alle lærerne som tok seg tid å besvare spørreundersøkelsen. Dere gjorde det mulig for meg å foreta mange interessante kvantitative analyser. Takk til venner og familie for støtte og hjelp. Jeg vil spesielt takke Roy for korrekturlesing.

Jeg vil rette en stor takk til min dyktige veileder Magne Olufsen. Uten din hjelp hadde jeg ikke klart å fullføre masteren på den tiden jeg hadde til rådighet. Takk for all tid du har brukt på nettopp min oppgave og takk for alltid raske tilbakemeldinger. Jeg vil også takke hele naturfagseksjonen for mye god undervisning gjennom årene.

Sist, men ikke minst, vil jeg takke min kjære mann, Per Inge. Takk for alle oppmuntringer, gode innspill, middag og husvask. Du utgjorde det lille ekstra for at jeg skulle bli ferdig i tide!

Tromsø, mai 2017

Karoline Belt



# Innholdsfortegnelse

<b>SAMMENDRAG</b> .....	<b>I</b>
<b>FORORD</b> .....	<b>III</b>
<b>1 INNLEDNING</b> .....	<b>1</b>
1.1 BAKGRUNN .....	1
1.2 FORSKNINGSPØRSMÅL.....	2
<b>2 TEORI</b> .....	<b>3</b>
2.1 PRAKTISK ARBEID I NATURFAG .....	3
2.2 HVORFOR ER VURDERING VIKTIG?.....	4
2.3 HVA ER FORMATIV VURDERING?.....	5
2.4 HVORDAN BRUKE FORMATIV VURDERING I UNDERVISNINGEN? .....	7
2.5 ELEVMEDEVIRKNING I VURDERINGSARBEIDET .....	10
2.6 VURDERING I PRAKTISK ARBEID I NATURFAG.....	11
2.7 BETYDNINGEN AV UNDERVISNINGSERFARING, KLASSETRINN OG STUDIEPOENG .....	14
2.8 TEORETISK RAMME FOR STUDIEN.....	15
<b>3 METODE</b> .....	<b>16</b>
3.1 FORSKNINGSDESIGN .....	16
3.2 BEGRUNNELSE FOR VALG AV METODE .....	16
3.3 DATAINNSAMLING .....	17
3.3.1 Utvalg.....	17
3.3.2 Spørreundersøkelsen.....	18
3.4 DATAANALYSE.....	20
3.4.1 Innmating og koding av data .....	21
3.4.2 Datareduksjon.....	21
3.4.3 Variansanalyser.....	23
3.5 VALIDITET OG RELIABILITET .....	24
3.5.1 Validitet.....	24
3.5.2 Reliabilitet.....	26
3.6 FORSKNINGSETIKK.....	28
<b>4 RESULTAT</b> .....	<b>29</b>
4.1 OVERSIKT OVER RESPONDENTENE I SPØRREUNDERSØKELSEN .....	29
4.2 LÆRERNES FORMATIVE VURDERINGSPRAKSIS I PRAKTISK ARBEID I NATURFAG .....	31
4.2.1 Generelle spørsmål om vurderingen.....	31
4.2.2 "Hvor skal eleven?" .....	32

4.2.3 "Hvor står eleven og hvordan jobbe videre mot målet?" .....	33
4.2.4 Nivå av tilbakemelding .....	33
4.2.5 Egen- og hverandrevurdering .....	34
4.3 VURDERINGSPRAKSIS KNYTTET TIL HVOR LENGE LÆRERNE HAR JOBBET SOM NATURFAGLÆRER .....	36
4.3.1 År som naturfaglærer og god formativ vurdering .....	36
4.3.2 År som naturfaglærer og nivå av tilbakemelding .....	37
4.3.3 År som naturfaglærer og tilrettelegging av egen- og hverandrevurdering .....	38
4.4 VURDERINGSPRAKSIS KNYTTET TIL HVILKET KLASSETRINN LÆRERNE UNDERVISER PÅ.....	40
4.4.1 Klassestrinn og god formativ vurdering .....	40
4.4.2 Klassestrinn og nivå av tilbakemelding .....	41
4.4.3 Klassestrinn og tilrettelegging av egen- og hverandrevurdering .....	42
4.4.4 Klassestrinn og skriftlig tilbakemelding .....	43
4.5 VURDERINGSPRAKSIS KNYTTET TIL ANTALL STUDIEPOENG LÆRERNE HAR I NATURFAG .....	45
4.5.1 Studiepoeng og god formativ vurdering .....	45
4.5.2 Studiepoeng og nivå av tilbakemelding .....	45
4.5.3 Studiepoeng og tilrettelegging av egen- og hverandrevurdering .....	46
<b>5 DISKUSJON.....</b>	<b>48</b>
5.1 BRUK AV UNDERVEISVURDERING.....	48
5.2 BRUK AV VURDERING FOR Å TILPASSE UNDERVISNINGEN.....	48
5.3 HVORDAN LÆRERNE BRUKER FORMATIV VURDERING I PRAKTISK ARBEID I NATURFAG .....	49
5.3.1 "Hvor skal eleven?" .....	49
5.3.2 "Hvor er eleven i forhold til målene?" .....	51
5.3.3 "Hvordan jobbe videre for å oppnå ønsket mål?" .....	51
5.3.4 Bruk av de ulike tilbakemeldingsnivåene .....	52
5.4 BRUK AV EGEN- OG HVERANDREVURDERING.....	53
5.5 BRUK AV MUNTLLIG OG SKRIFTLIG TILBAKEMELDING .....	54
5.6 UTFORSKENDE ARBEIDSMÅTER SOM EN ARENA FOR FORMATIV VURDERING .....	56
5.7 BETYDNING AV ÅR I JOBB, KLASSETRINN OG ANTALL STUDIEPOENG .....	58
5.7.1 Vurderingspraksis knyttet til antall år i jobb som naturfaglærer .....	58
5.7.2 Vurderingspraksis knyttet til hvilket klassestrinn lærerne underviser på .....	59
5.7.3 Vurderingspraksis knyttet til antall studiepoeng i naturfag blant lærere.....	60
5.8 SVAKHETER MED STUDIEN .....	60
<b>6 AVSLUTNING.....</b>	<b>62</b>
6.1 KONKLUSJON.....	62
6.2 VIDERE FORSKNING .....	63



<b>LITTERATURLISTE</b> .....	<b>65</b>
<b>VEDLEGG</b> .....	<b>68</b>
VEDLEGG 1: SPØRREUNDERSØKELSE .....	68



# 1 Innledning

Målet med denne masteroppgaven var å lære mer om formativ vurdering i naturfag. Mye forskning både nasjonalt og internasjonalt tar for seg formativ vurdering og dens betydning for elevenes læring (Black et al., 2003; Black, Harrison, Lee, Marshall & Wiliam, 2004; Throndsen, Hopfenbeck, Lie & Dale, 2009). Kunnskap om hvordan en kan bruke vurdering i undervisning for å forbedre læringen til elevene tenker jeg er en svært viktig del av lærerens profesjonsutvikling. Naturfag skiller seg ut fra mange andre fag med at det i tillegg til å lære kunnskap om naturfag også skal læres praktiske ferdigheter i faget. Jeg ønsket derfor å få en større forståelse av hvordan formativ vurdering gjøres i praktisk arbeid i naturfag. Dette ble gjort gjennom en spørreundersøkelse blant naturfaglærere i grunnskolen.

## 1.1 Bakgrunn

Iver etter å forstå mer av hvilken betydning vurdering har i skolen har hatt stor vekst i Norge de siste årene. Både nasjonal og internasjonal forskning, revisjoner av forskrift til opplæringsloven om vurdering både i 2007 og 2009, prosjektet *Bedre vurderingspraksis* og Utdanningsdirektoratets satsning på vurdering for læring har vært med å bidra til denne veksten. Også Stortingsmelding 22 peker på viktigheten av å anvende kunnskap om vurdering for å fremme elevenes læring, hvor blant annet en forbedret tilbakemeldingspraksis er et av tiltakene meldingen foreslår (Kunnskapsdepartementet, 2011).

I løpet av årene på lærerutdanningen har jeg utviklet en stor interesse for vurderingens betydning i undervisningen. Jeg har erfart at hos de forskjellige skolene og lærerne har vurdering svært ulik betydning i deres undervisningspraksis. I naturfagdidaktikk på det nest siste semesteret i utdanningen leste jeg flere artikler om formativ vurdering og hvordan dette kan øke elevenes læringsutbytte i naturfag. Særlig ble et prosjekt kalt *The King's Medway, Oxfordshire Formative Assessment Project* (KMOFAP) en stor inspirasjon. Prosjektet hadde som intensjon å hjelpe lærere til å transformere formative vurderingsformer fra forskningen inn i deres egen praksis i klasserommet (Black et al., 2003).

Formativ vurdering skal være med å støtte elevenes læring. Den formative vurderingens betydning kan gjerne belyses ut fra et sitat av musikeren George Harrison: *If you don't know where you're going, any road will take you there*. I undervisningssammenheng betyr det at lærer og elever må ha kjennskap til hva målene for undervisningen er og hvordan man kan nå disse målene. Dersom dette ikke er tilfelle vil hvilken som helst undervisning være god nok, enten man når målet eller ikke.

## **1.2 Forskningsspørsmål**

Min opplevelse av naturfagundervisningen er at det kan være vanskelig å vurdere elevene i praktisk arbeid. Jeg har erfart at elevenes praktiske kompetanse i mange tilfeller stort sett vurderes gjennom rapporter eller spørsmål på en prøve. Jeg ønsket derfor gjennom min masteroppgave å lære mer om hvordan en kan bruke formativ vurdering gjennom det praktiske arbeidet i naturfag. Samtidig var jeg interessert i å lære mer om hvordan naturfaglærere i dag praktiserer formativ vurdering i praktisk arbeid i naturfag. Forskningsspørsmålet mitt ble derfor:

**På hvilken måte og i hvor stor grad bruker naturfaglærere formativ vurdering i den praktiske delen av naturfagundervisningen i grunnskolen?**

I tillegg ønsket jeg å undersøke i hvor stor grad noen faktorer var med å påvirke deres vurderingspraksis og tok derfor med følgende underspørsmål:

*Påvirkes lærernes vurderingspraksis av (1) undervisningserfaringen i faget; (2) klassetrinn det undervises på og (3) antall studiepoeng i naturfag?*

## 2 Teori

### 2.1 Praktisk arbeid i naturfag

I dette prosjektet er det lagt vekt på praktisk arbeid i naturfag. Problemstillingen ønsker å besvare i hvor stor grad lærerne vurderer formativt i den praktiske delen av faget. Det er viktig å ha en felles forståelse for hva praktisk arbeid i naturfag innebærer og til dette er det tatt i bruk en definisjon basert på internasjonal litteratur om naturfagundervisning, som beskriver praktisk arbeid følgende:

*Any type of science teaching and learning activity in which students, working either individually or in small groups, are involved in manipulating and/or observing real objects and materials (...) as opposed to virtual objects and materials such as those obtained from a DVD, a computer simulation, or even from a text-based account.*  
(Millar, 2011, referert i Abrahams & Reiss, 2012, s. 1036)

Denne definisjonen av praktisk arbeid omhandler arbeid hvor elevene, både alene og i grupper, behandler og/eller observerer virkelige objekter og materialer. Dette er en nokså vid definisjon av begrepet. Praktisk arbeid handler ikke om *hvor* det blir utført, men *hvordan*, og dekker derfor bredere enn labarbeid (Abrahams & Reiss, 2012).

Praktisk arbeid har en sentral rolle i naturfag. Dette kommer klart til uttrykk i læreplan i naturfag både gjennom formålet med faget og i flere av kompetansemålene (Kunnskapsdepartementet, 2013). Å undersøke hvordan lærere vurderer i den praktiske delen av naturfaget er derfor sentralt for å kunne si noe om lærernes vurderingspraksis for faget i sin helhet. Til tross for den sentrale rollen praktisk arbeid har fått i undervisningen har det vist seg vanskelig å påvise læringseffekt av slikt arbeid (Hodson, 1990; Hofstein & Lunetta, 1982; Kind, 1999). Det hevdes likevel at det er åpenbart viktig for læring i naturfag at det foregår noe praktisk arbeid (Kjærnsli, Lie, Olsen & Roe, 2007). I tillegg legger det til rette for motivasjon og interesse for faget:

*Det viktigste funnet er at det å fokusere på anvendelser av lærestoffet i praksis ser ut til å fremme både interessen for og verdsettingen av naturfag. (...) Når det gjelder å øke interessen for faget generelt, og spesielt å fremme rekrutteringen til realfagene, framstår det imidlertid som et viktig punkt. (Kjærnsli et al., 2007, s. 120)*

Kjærnsli et al. (2007) skriver også at motivasjon og engasjement ofte blir regnet som en viktig drivkraft for læring. Millar (2011) hevder også at en av de største fordelene med praktisk arbeid er at det bidrar til økt interesse i naturfag og støtter læringen. Derfor bør naturfaglærere bruke praktisk arbeid i naturfag som en naturlig del av undervisningen. God vurdering i det praktiske delen av naturfaget vil derfor også være viktig.

## **2.2 Hvorfor er vurdering viktig?**

Robin Millar (2013) hevder at vurdering har en fundamental rolle i skolen og skriver at vurderingen har som oppgave å tydeliggjøre den læringen som er tiltenkt i enhver gitt situasjon. Han utdyper dette videre og skriver at vurderingsarbeidet skal brukes til å operasjonalisere ønsket utfall av undervisningen. Det vil si at vurderingen har en sentral rolle i elevenes læring. Det heter seg at vi ønsker at elevene skal forstå hva vi underviser om, men siden vi ikke kan observere forståelse direkte må vi tyde dette ut fra hva vi faktisk kan observere: Det elevene sier, skriver og gjør. Dersom vi ønsker å vurdere om elevene forstår X, må vi gi eleven oppgaver som bare en person som forstår X vil klare å løse (Millar, 2013). Vurderingen er derfor sentral for å kunne avgjøre om elevene har oppnådd ønsket læring, noe også den britiske vurderingseksperten Dylan Wiliam fastslår:

*It is only through assessment that we can find out whether a particular sequence of instructional activities has resulted in the intended learning outcomes.*

(Wiliam, 2011, s. 3)

Wiliam (2011) hevder at uansett hvor nøye læreren planlegger sin undervisning vil ikke elevene lære akkurat hva læreren hadde som intensjon den timen. Dette er fordi elevene møter undervisningen med ulike erfaringer og forutsetninger og oppfatter derfor undervisningen ulikt. Læreren kan altså ikke forutse hva elevene vil lære gjennom planleggingen av



undervisningen. Derfor, skriver Wiliam, er vurdering en veldig sentral prosess i effektiv undervisning.

### 2.3 Hva er formativ vurdering?

Formativ vurdering har med tiden blitt et omfattende og kanskje noe uklart begrep i pedagogiske og didaktiske diskusjoner. Det er derfor viktig å klargjøre begrepets betydning i dette prosjektet. Jeg vil nedenfor presentere tre kjente begreper som brukes ulikt, men som likevel alle sier noe om den formative vurderingens natur; vurdering som støtter læring, vurdering for læring og underveisvurdering.

Historisk sett er det særlig to begreper i vurderingen som det har vært naturlig å skille; den formative og den summative vurderingen. Det er innledningsvis viktig å klargjøre forskjellen mellom disse. Millar (2013) skriver at vurderingen er summativ når den har som intensjon å *sertifisere* læring. Dette kan være for eksempel en prøve med karakter etter avsluttet emne. Til motsetning vil vurderingen virke formativ når intensjonen er å *støtte* læring. Det ikke er selve vurderingsformen og teknikken som brukes som avgjør om vurderingen er formativ eller summativ, men hva som er formålet bak og bruken av den (Dysthe, 2008). Vurderingen ikke kan kalles formativ før den faktisk blir brukt til å tilpasse undervisningen ut fra elevenes behov (Black & Wiliam, 1998). At vurderingen skal kalles formativ vurdering krever altså at intensjonen og bruken av den skal være med å støtte læringen ved å tilpasses til hver enkelt elev.

*The Assessment Reform Group* mente at begrepet formativ vurdering var brukt på så mange ulike måter og hadde en så vid betydning at det ikke lengre var til hjelp (Broadfoot et al., 1999). De foretrakk begrepet *assessment for learning*, oversatt *vurdering for læring*, som de ga følgende definisjon:

*The process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there. (Broadfoot et al., 2002).*

Som vi vil se senere i teorikapitlet ligger denne definisjonen av vurdering for læring nært til andre definisjoner og konseptualiseringer av formativ vurdering. Det er spesielt enighet rundt tre faktorer som er viktige for å fremme elevenes læring: (1) De må vite hva de skal lære, (2) de må vite hvor de er i læringsprosessen og (3) de må vite hvordan de skal oppnå ønsket læring (Broadfoot et al., 2002; Hattie & Timperley, 2007; Wiliam, 2011).

Til tross for at formativ vurdering og vurdering for læring har mange likhetstrekk argumenterer blant annet Black og Wiliam og deres kollegaer at det fins en viktig forskjell mellom disse. De forklarer dette med at vurdering for læring handler om formålet med vurderingen og formativ vurdering forteller noe om funksjonen som vurderingen faktisk har. Det vil si at vurdering for læring har som formål å støtte og forbedre elevens læring og at vurderingen kalles formativ når den faktisk blir brukt for å møte elevenes læringsbehov (Black et al., 2004). Black og Wiliam (2009) gir derfor denne definisjon av formativ vurdering, eller formativ praksis som de her bruker:

*Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (Black & Wiliam, 2009).*

Målet med formativ vurdering er med andre ord å synliggjøre elevenes læring for å kunne forbedre undervisningen i neste steg. Også i definisjonen til Black og Wiliam har både lærer, elev og medelev en sentral rolle hos elevens læring.

I Norge er begrepet *underveisvurdering* mye brukt når en snakker om vurdering i skolen. Underveisvurdering har en sentral plass i det norske skoleverket og det er også pålagt at elever og lærlinger skal få denne typen vurdering i opplæringen. Forskrift til opplæringslova (2006) § 3-11 sier om underveisvurdering:

*Underveisvurdering i fag skal brukast som ein reiskap i læreprosessen, som grunnlag for tilpassa opplæring og bidra til at eleven, lærlingen og lære kandidatene aukar kompetansen sin i fag. Underveisvurderinga i fag, i orden og i åtferd skal givast løpande og systematisk og kan vere både munnleg og skriftleg (2006).*

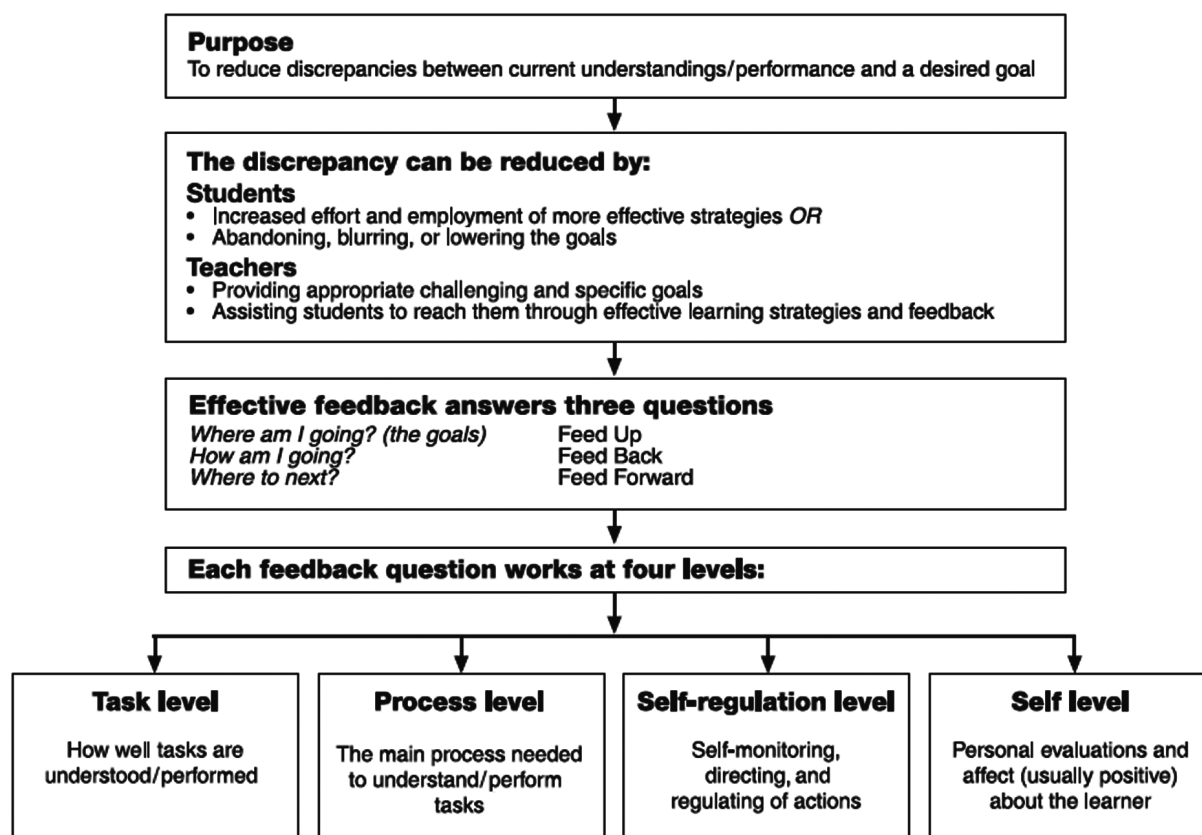
Lærere er derfor pålagt å gjøre vurderinger av elevene gjennom hele læreprosessen. Videre står det også at vurderingen skal inneholde informasjon om hvilken kompetanse eleven innehar samt rettleiding til hvordan eleven kan utvikle kompetansen i faget. Underveisvurderingen er derfor også nært tilknyttet formativ vurdering.

Formativ vurdering, slik den brukes i denne oppgaven, legger vekt på at vurderingen skal fortelle noe om hvor eleven skal, hvor eleven står i forhold til hva som skal kunnes og samtidig gi informasjon om hva som må gjøres for at eleven skal bli bedre. Samtidig vil det også legges vekt på elevenes rolle i vurderingen, noe som vil belyses ytterligere i kapittel 2.5.

## **2.4 Hvordan bruke formativ vurdering i undervisningen?**

Prosjektet ”Forskning på individuell vurdering i skolen” (FIVIS) har gjort en omfattende studie av lærernes vurderingskompetanse i skolen (Sandvik & Buland, 2014). FIVIS fant i intervjuer ut at elever på alle trinn opplever vurdering som viktig, også for egen læring. De ønsker gode, tydelige tilbakemeldinger på hva de er gode til, hva som kan bli bedre, og hvordan. FIVIS så også at lærerne selv anså tilbakemeldinger som viktige i læringsprosessene. Det ligger altså et ønske om at god formativ vurderingspraksis skal finne sted både blant elever og lærere i den norske skolen.

Tilbakemelding er en av de faktorene som har størst betydning for læring og prestasjon (Hattie & Timperley, 2007) og er derfor aktuell å belyse når vi snakker om formativ vurdering. Hattie og Timperley (2007) påpeker at tilbakemeldingen er konseptualisert som informasjon gitt av en agent (for eksempel lærer, medelev, foreldre, seg selv eller erfaring) og omhandler aspekter av en persons prestasjon eller forståelse. Selv om Hattie og Timperley bruker begrepet *feedback* (tilbakemelding) og ikke *formativ vurdering* har den store likhetstrekk med annen teori og forskning rundt formativ vurdering og den anses derfor som svært aktuell med tanke på forskningsspørsmålet. Hovedformålet med feedback er å redusere avviket mellom nåværende kompetanse og målet med læringen (Hattie & Timperley, 2007, s. 86), noe som også samsvarer med annen teori om formativ vurdering (Black & Wiliam, 1998, 2009; Broadfoot et al., 2002; Millar, 2013; Wiliam, 2011). På bakgrunn av tidligere forskning har Hattie og Timperley utviklet en modell for effektiv tilbakemelding som vises i figur 2.1.



Figur 2.1: Modell for effektiv tilbakemelding. Hentet fra Hattie og Timperley (2007, s. 87).

I modellen til Hattie og Timperley (figur 2.1) finner vi igjen de sentrale spørsmålene som god formativ vurdering skal besvare: ”Hva er målet?”, ”Hvordan ligger eleven an i forhold til målet?” og ”Hvordan nærme seg målet ytterligere?”. Disse tre spørsmålene vil alltid være sentrale for å kunne redusere forskjellen mellom nåværende kompetanse og målet for hva eleven skal kunne, og er derfor som påpekt tidligere viktige for å kunne utføre god formativ vurdering. I tillegg utdyper Hattie og Timperley at disse tre spørsmålene kan besvares på fire ulike nivå, som også er avgjørende for effektiviteten. Disse fire nivåene finner vi nederst i modellen i figur 2.1 og sier noe om tilbakemeldingen gis (1) direkte tilknyttet oppgaven eleven jobber med (oppgavenivå), (2) tilknyttet prosessen eleven har vært gjennom for å løse oppgaven (prosessnivå), (3) slik at eleven får mulighet til å rette seg selv og justere seg selv (selvreguleringsnivå), eller (4) på et personlig nivå som ikke er direkte tilknyttet utførelsen av oppgaven (personlig nivå). Tilbakemelding på det personlige nivået har vist seg å være lite effektivt for elevenes læring og er derfor ikke formative (Hattie & Timperley, 2007; Sadler, 1998). Det har likevel vist seg at lærere generelt gir mye ros til elevene (Thronsen et al.,

2009). Den tilbakemeldingen som har vist seg å være mest læringsfremmede for elevene, og derfor også formativ, er tilbakemelding som kombinerer de tre første nivåene; *task level*, *process level* og *self-regulation level* (Hattie & Timperley, 2007).

Lærerne i KMOFAP-prosjektet så at hvordan man gir tilbakemelding til elevene også har betydning for elevenes læring. Lærerne la spesielt vekt på tre viktige punkt i tilbakemeldingen. Den første var at oppgavene skulle oppmuntre elevene til å forstå de viktigste egenskapene de hadde lært. Det andre punktet gikk ut på at kommentarer skulle være spesifikke. De skulle gi god informasjon om hva som var gjort bra, hva som trengte å forbedres og hvordan de skulle forbedre seg. Til slutt så lærerne også at det var viktig å gi elevene mulighet til å følge opp kommentarene og arbeide med dem.

Det er ikke bare Hattie og Timperley som har forsøkt å kaste lys over hva god formativ vurderingspraksis er. Wiliam (2011) utdyper to viktige egenskaper når en skal bruke vurdering som støtte til læring. Den ene er at vurderingen må gi svar på mer enn bare informasjon om hvor langt det er mellom nåværende og ønsket kompetanse, men også gi forslag til aktiviteter som kan forbedre kompetansen. Dette samsvarer også med hva Forskrift til opplæringslova (2006) sier om underveisvurdering, som skrevet tidligere. For det andre må eleven selv engasjere seg i tiltak for å forbedre læring. Det vil si å gjennomføre de tiltakene læreren foreslår, spørre en medelev om hjelp eller reflektere over ulike måter å forbedre seg selv på. Altså er elevenes involvering i prosessen svært viktig for at den formative vurderingen skal fungere optimalt (Wiliam, 2011). For å imøtekomme dette kravet har Wiliam foreslått at formativ vurdering kan konseptualiseres som fem nøkkelstrategier som vist i tabell 2.1.

Tabell 2.1: Fem nøkkelstrategier for formativ vurdering. Hentet fra Wiliam (2011, s. 12)

	Where the learner is going	Where the learner is now	How to get there
<b>Teacher</b>	Clarifying, sharing and understanding learning intentions and success criteria	Engineering effective discussions, tasks and activities that elicit evidence of learning	Providing feedback that moves learning forward
<b>Peer</b>		Activating students as learning resources for one another	
<b>Learner</b>		Activating students as owners of their own learning	

Tabell 2.1 viser at også Wiliam er opptatt av at formativ vurdering skal besvare de tre spørsmålene; hva eleven skal kunne, hvor eleven er på nåværende tidspunkt og hvordan eleven skal oppnå målet. Samtidig viser også figuren at både eleven selv og medelevene har en stor rolle i elevens vurderingsarbeid.

## 2.5 Elevmedvirkning i vurderingsarbeidet

Både Hattie og Timperleys modell for effektiv tilbakemelding (Hattie & Timperley, 2007) og Williams tabell om fem nøkkelstrategier for formativ vurdering (Wiliam, 2011) forteller at eleven har en viktig rolle i vurdering av eget arbeid og vil være med å påvirke egen læring. For at elevene skal kunne bli gode til å evaluere eget arbeid har følgende fire punkter vist seg å ha betydning; (1) elevene får tydelige mål; (2) elevene involveres i å lage kriteriene for vurderingen; (3) elevene får tilbakemelding; og (4) elevene får mulighet til å korrigere seg selv (Bruce, 2001). Det har de siste årene vært mye fokus på å gi elevene læringsmål, noe vi kan se ved mange norske skoler i dag. Å skrive mål for timen på tavla før undervisningen starter er blitt vanlig praksis hos mange lærere. Det kan likevel diskuteres om målene blir tydelige nok for elevene for at de skal kunne gjøre gode formative vurderinger av eget arbeid. Å involvere elevene i å lage kriterier har vist seg å føre til at elevene forstår tydeligere hva de faktisk må jobbe med (Bruce, 2001). Tilbakemeldingen er sentral i formativ vurdering, men den er også viktig for elevenes muligheter for egenvurdering. Tilbakemeldingene behøver ikke bare å komme fra lærer, men kan også komme fra medelever. Vi skal se i neste avsnitt hvorfor tilbakemeldinger fra medelever kan vise seg å være spesielt effektive.



Lærerne i KMOFAP-prosjektet så mange fordeler med at elevene fikk mulighet til å vurdere seg selv og hverandre (Black et al., 2003). Konklusjonen var at egenvurdering hjalp elevene til å se eget arbeid opp mot de faglige målene og det stimulerte samtidig metakognitiv tenking. For hverandrevurdering så de at elevene la mer innsats i eget arbeid når de visste at en medelev skulle vurdere det i etterkant og en så at elevene aksepterte kritikk fra medelever i større grad enn fra læreren. I tillegg fikk de en vurdering på et språk de selv var komfortabel med og de fikk gjennom samtaler med medelever styrket sine egne meninger. Lærerne i prosjektet konkluderte med at kriteriene for hva som skulle læres måtte synliggjøres for elevene til enhver tid, ved at de visste hva som var målet med undervisningen og hva som skulle til for å oppnå dette. I tillegg mente de at elevene burde lære seg gode ferdigheter i å vurdere hverandre, både fordi dette kan ha stor egenverdi i seg selv og fordi det kan hjelpe elevene til å utvikle objektivitet med tanke på egenvurdering. Til sist er det også viktig at elevene blir oppmuntret til, og lærer å bruke målene deres aktivt slik at de kan selv vurdere arbeidet sitt underveis og på den måten hjelpe seg selv videre.

Studier har vist at egenvurdering og vurdering av medelever relativt lite utbredt blant norske lærere og elevene opplever i mindre grad å være ressurser for hverandres læring (Sandvik & Buland, 2014). Rapporten fra FIVIS forteller at elevene ønsket en god vurderingspraksis med tydelige mål og kriterier, der de selv var aktive i egen og andres læringsprosess. Dette anså elevene som viktig for egen læring, motivasjon og mestring (Sandvik & Buland, 2014).

## **2.6 Vurdering i praktisk arbeid i naturfag**

Mens praktiske ferdigheter i naturfagundervisningen helt klart blir ansett som viktig, er det mangel på klarhet i hva disse ferdighetene faktisk er og hvordan de kan bli vurdert på best mulig måte (Abrahams, Reiss & Sharpe, 2013). Noe som kan vanskeliggjøre vurderingsarbeidet i praktisk arbeid i naturfag er at denne delen av undervisningen innehar mange ulike mål. Hofstein og Lunetta (2004) skriver at hovedmålene med praktisk arbeid er å øke elevenes (1) forståelse av naturvitenskaplige konsepter; (2) interesse og motivasjon; (3) praktiske ferdigheter i naturfag og ferdigheter rundt problemløsning; (4) naturvitenskaplige tenkemåte; og (5) forståelse av naturvitenskapens egenart. I tillegg skal praktisk arbeid blant annet lære elevene å knytte teori til praksis, lære elevene ferdigheter på lab og støtte læring i

naturfag (Abrahams et al., 2013; Millar, 2011; Ottander & Grelsson, 2006). Dersom naturfaglærere skal ha mulighet til å nå samtlige mål sammen med elevene vil vurderingen i praktisk arbeid være sentral. Det er også viktig at elevene bruker praktisk arbeid i naturfag til å fremme tenking, stille spørsmål og forme teorier (Osborne & Dillon, 2010). Praktisk arbeid har den fordelen at den gir muligheter for at elevene aktivt deltar i egen læring gjennom blant annet diskusjoner og samarbeid, noe også Black og Harrison fremhever:

*An assessment for learning approach fits well with practical work since it calls on the learner to be actively responsible in their learning and to work collaboratively. Such an approach also highlights the importance of classroom talk (Osborne & Dillon, 2010, s. 194).*

Black og Harrison tar også opp diskusjonen rundt praktisk arbeid som viktig for elevenes tenking og læring, men for at dette skal fungere formativt må lærer planlegge diskusjoner med spørsmål som både får frem forståelse og misforståelse hos elevene (Osborne & Dillon, 2010).

I Sverige har det blitt gjennomført en case-studie som hadde som mål å få en bedre forståelse av hvilken rolle labarbeid har i naturfagundervisningen, med en vektlegging på vurdering. Uformelle diskusjoner med elevene tilknyttet labarbeid ble av lærerne i case-studien ansett som svært viktig. Likevel viste det seg at lærerne stort sett vurderte elevenes kompetanse i labarbeid ut fra en labrapport (Ottander & Grelsson, 2006). Svakheten med denne typen vurdering er at dette gjør det umulig å vurdere elevenes praktiske evner i et forsøk, noe som var et av hovedmålene med labarbeid i følge de svenske lærerne (Ottander & Grelsson, 2006).

I analysen av forventet læringsutbytte fant case-studien ut at lærere og elever ofte hadde ulik oppfatning av hva som ble vurdert i labarbeidet. Dette ble forklart blant annet ved at fokuset i undervisningen ofte var på hva elevene skulle gjøre, ikke på læringsmålene (Ottander & Grelsson, 2006). I følge studiene så alle lærerne viktigheten av å synliggjøre forventet læringsutbytte av ulike labøvelser for elevene. Som et resultat av studien anbefalte forfatterne at mål, undervisning og vurderingskriterier for eksperimentelle forsøk må tas hensyn til når man planlegger labarbeid. Det er også viktig at målene for hva elevene skal kunne, samt hva som vurderes, er synlig for elevene og at instruksjonene er formulert slik at disse målene kommuniseres til elevene (Ottander & Grelsson, 2006).

I følge en britisk studie ble det konkludert med at praktisk arbeid i naturfag ofte blir sett på som noe man må gjennomføre rutinemessig fordi det er forventet at man skal gjøre det, ikke for at elevene skal lære noe (Nott & Wellington, 1999). Nott og Wellington (1999) skriver også at dersom det ikke tydeliggjøres hvilke praktiske ferdigheter en ønsker at elevene skal oppnå kan vurderingen bare bli en prosess hvor elevene lærer hvordan de kan få gode karakterer i stedet for å lære hvordan de kan forbedre sine praktiske ferdigheter og øke forståelsen av vitenskapen.

Vurdering av praktisk arbeid kan gjøres på ulike måter. Abrahams et al. (2013) skiller mellom “Direct Assessment of Practical Skills” (DAPS) og “Indirect Assessment of Practical Skills” (IAPS). DAPS viser til hvordan eleven vurderes ut ifra demonstrasjoner av sine praktiske ferdigheter, mens IAPS refererer til vurdering som sier noe om elevens kompetanse ut ifra datamateriale og/eller rapport laget ut fra det praktiske arbeidet eleven foretok. Hver av disse tilnærmingene har sine fordeler og ulemper, men for mye fokus på IAPS reduserer sannsynligheten for at praktisk arbeid vil bli undervist og lært så bra som det kan gjøres (Abrahams et al., 2013). Dette fordi IAPS bare vurderer hva elevene kan *om* praktisk arbeid og hvordan det utføres, men viser ikke til de faktiske ferdighetene elevene har i praktisk arbeid, som kommer frem gjennom DAPS (Reiss, Abrahams & Sharpe, 2012).

Det fins ulike måter å undervise praktisk arbeid i naturfag, fra lukkede forsøk der elevene gjerne jobber ut fra en ”oppskrift”, til mer åpne arbeidsmetoder der elevene selv er med å forme undervisningen. Dersom praktisk arbeid skal bidra til at elevene utvikler kunnskap og forståelse av vitenskapen og kunne forstå vitenskapelige prosesser er det viktig at også arbeidsmetoden retter seg mot dette. Internasjonalt snakker vi da om begrepet *inquiry-based teaching*. I *inquiry-based teaching* i naturfagundervisningen utfører elevene mange av de samme aktivitetene og tankeprosessene som de gjør i vitenskapen og utvikler kunnskap og forståelse av vitenskapelige ideer, samt forståelse av hvordan vitenskapen fungerer (Council, 2000). I Norge snakker vi gjerne om utforskende arbeidsmåter i naturfag. Utforskende arbeidsmåter har tre klare kjennetegn sammenlignet med tradisjonell undervisning: (1) arbeidet bygger på spørsmål som er formulert innledningsvis, (2) elevene bruker data og informasjon til å utvikle, etterprøve og velge mellom mulige svar og (3) elevene arbeider med å vurdere og å videreutvikle kunnskap i en utforskende prosess (Knain & Kolstø, 2011). I utforskende arbeidsmåter får også lærerne mer en rolle som veileder i stede for en formidler

(Knain & Kolstø, 2011). *Inquiry-based teaching*, eller utforskende arbeidsmåter, legger i stor grad til rette for formativ vurdering i praktisk arbeid. Vurderingen av utforskende arbeidsmåter skal måle både elevenes konseptuelle forståelse i naturfag, evne til å gjennomføre vitenskapelige undersøkelser og forståelsen av disse, noe som best kan gjøres gjennom lærerens kontinuerlige vurdering av elevenes forståelse (Council, 2000).

## **2.7 Betydningen av undervisningserfaring, klassetrinn og studiepoeng**

Som ett av mine forskningsspørsmål ønsker jeg å undersøke om det fins noen sammenhenger mellom lærernes undervisningserfaring, hvilket klassetrinn de jobber på og hvor mange studiepoeng de har i naturfag. Etter et søk i litteraturen finner jeg ingen tidligere forskning på om lærernes undervisningserfaring har betydning for deres vurderingspraksis. Det fins likevel forskning som antyder at lærernes undervisningserfaring har betydning for undervisningskvaliteten (Berliner, 2001; Blömeke, Olsen & Suhl, 2016; Clotfelter, Ladd & Vigdor, 2007), noe som kan knyttes opp mot blant annet lærernes vurderingspraksis.

Prosjektet FIVIS (Sandvik & Buland, 2014) fokuserte ikke på naturfaglærere, men det er naturlig å tenke at flere av lærerne som deltok i prosjektet hadde naturfag som ett av sine undervisningsfag og en kan derfor anta en viss overføringsverdi av resultatene. FIVIS gjorde blant annet et kvantitativ undersøkelse av lærere for å finne ut av deres vurderingspraksiser. De fant ut at 3 av 32 bakgrunnsvariabler ga signifikante forskjeller i lærernes vurderingspraksis. Disse bakgrunnsvariablene var "Skoleslag", "Fordypning i fag" og "Type lærerutdanning". For bakgrunnsvariabelen "Skoleslag" så FIVIS at lærere som arbeider på barnetrinnet skårer signifikant lavere på vurderingspraksis enn lærere på ungdomsskoler og videregående skoler. Også resultater fra TIMMS 2015 viser at undervisningsvurdering blir mer vektlagt på barnetrinnet enn ungdomstrinnet både i Norge og internasjonalt (Bergem, Kaarstein & Nilsen, 2016). Rapporten til FIVIS så også at lærernes faglige og fagdidaktiske kompetanse har betydning for hvordan læreren arbeider med vurdering. For bakgrunnsvariabelen "Type lærerutdanning" så de at lærere med fagutdanning fra universitet/høgskole med praktisk-pedagogisk utdanning skårer signifikant bedre enn allmennlærere. I tillegg så de at lærere med fra ett til tre års utdanning i faget skårer bedre enn lærere med bare et halvt års utdanning i faget. Rapporten konkluderte derfor at lærere med svak faglig og fagdidaktisk kompetanse synes å ha utfordringer med å forstå og

operasjonalisere vurdering for læring godt nok (Sandvik & Buland, 2014). Også resultater fra TIMMS 2015 viser at læreres fagspesialisering i naturfag og faglig og fagdidaktisk trygghet som lærer har en sterk og positiv sammenheng med deres undervisningskvalitet (Bergem et al., 2016).

## **2.8 Teoretisk ramme for studien**

Som teoretisk ramme for prosjektet har jeg valgt å bruke Hattie og Timperleys modell for effektiv tilbakemelding (2007) som vist i figur 2.1. Hovedfokuset vil ligge på de tre spørsmålene (Hva skal læres? Hva kan eleven? Hva er veien videre?) som formativ vurdering bør inneholde. I tillegg ønsker jeg å undersøke de fire nivåene som spørsmålene kan virke på og dens betydning for den formative vurderingen. Hattie og Timperleys modell ble brukt som utgangspunkt i utarbeidelse av spørreskjema til prosjektet. I tillegg er modellen brukt aktivt i analysedelen ved at jeg ut fra spørsmålene i spørreundersøkelsen laget konstrukter på bakgrunn av modellen. Til slutt ønsket jeg også å undersøke i hvor stor grad lærerne la til rette for egen- og hverandrevurdering. Disse spørsmålene ble ikke utformet på bakgrunn av modeller, men anses som viktige å ta med i undersøkelsen. Dette begrunnes i at egen- og hverandrevurdering har vist seg å være sentral i god formativ vurdering (Black et al., 2003; Hattie & Timperley, 2007; Wiliam, 2011).

## **3 Metode**

### **3.1 Forskningsdesign**

Denne masteroppgaven er en kvantitativ studie som bygger på en deduktiv tilnærming. Studien er basert på tversnittundersøkelser av naturfaglærernes formative vurdering i den praktiske delen av naturfaget. Det ble tatt utgangspunkt i teorier om formativ vurdering og utformet en spørreundersøkelse som ble gitt ut på en konferanse for naturfaglærere i Nord-Norge.

### **3.2 Begrunnelse for valg av metode**

Jeg ønsket i denne masteroppgaven å få et innblikk i hvordan og i hvor stor grad naturfaglærere vurderer formativt i den praktiske delen av naturfag. Jeg ønsket også å se om det var en sammenheng i besvarelsene knyttet til lærernes undervisningserfaring, antall studiepoeng i naturfag og hvilket klassetrinn de jobbet på. En kvantitativ forskningsstrategi er ofte deduktiv, hvor forskeren stiller spørsmål og avleder hypoteser fra ett eller flere relevante teoretiske perspektiver (Ringdal, 2013). Med utgangspunkt i det teoretiske perspektivet var det derfor naturlig å bruke en kvantitativ spørreundersøkelse for å undersøke i hvor stor grad naturfaglærere bruker formativ vurdering. Jeg ønsket også å undersøke ulikheter og sammenhenger mellom ulike grupper lærere og dette peker også klart mot en kvantitativ forskningsstrategi, hvor et ønske om årsaksforklaringer er sentralt (Ringdal, 2013).

I startfasen av arbeidet med masteroppgaven fikk jeg en unik mulighet til å delta på en realfagskonferanse sammen med mange naturfaglærere i fylket. Her fikk jeg muligheten til å gjennomføre datainnsamling til masteroppgaven blant deltakerne på konferansen. Det ble da også naturlig å velge spørreundersøkelse som metode, da det ga meg mulighet til å innhente mye data over kort tid. Spørreundersøkelsens format er også påvirket av de rammene som var satt på konferansen, ved at datainnsamlingen ikke skulle ta lang tid.



### **3.3 Datainnsamling**

#### **3.3.1 Utvalg**

I kvantitative undersøkelser ønsker en at enhetene symboliserer et representativt utvalg av populasjonen (Ringdal, 2013). I denne masteroppgaven er det naturfaglærere i Norge jeg ønsker å studere og disse betraktes derfor som populasjonen. På grunn av muligheten jeg fikk til å gjennomføre undersøkelsen blant deltakerne på konferansen var det også naturlig at disse ble mitt utvalg av enheter, eller respondenter. Lærerne på konferansen kom i hovedsak fra skoler i Troms fylke, men også lærere fra Nordland fylke deltok. Totalt 61 av 85 registrerte deltakende lærere på konferansen besvarte spørreundersøkelsen, noe som gir et nettoutvalg med svarprosent på 72 %.

For at enhetene man undersøker skal kunne kalles et representativt utvalg må sammensetningen av ulike egenskaper i utvalget tilsvare sammensetningen i populasjonen (Johannessen, Tufte & Christoffersen, 2016, s. 243). Metoden som gir størst sannsynlighet for at utvalget er representativt er en sannsynlighetsutvelgelse, men dette forutsetter også at det skjer en tilfeldig trekking av utvalget. Det ble ikke foretatt et sannsynlighetsutvalg av populasjonen naturfaglærere i Norge i denne spørreundersøkelsen, men utvalget ble gjort på bakgrunn av tilgjengelighet på konferansen. Likevel kan det argumenteres for at utvalget kan være representativt for naturfaglærere i Norge.

Det kan kanskje tenkes at lærerne som deltok på konferansen er mer motiverte enn de lærerne som ikke deltok. Dette forutsetter at det var opp til hver enkelt av lærerne å delta på konferansen. Vi vet likevel at noen skoler sendte alle naturfaglærerne på konferansen og en kan derfor anta at lærere med ulik grad av motivasjon også var representert på konferansen. Vi vet også at det var lærere både fra distriktsskoler og byskoler, noe som også representerer en virkelighetsnær variasjon av naturfaglærere. Det ble ikke spurt om kjønn i spørreundersøkelsen, men på konferansen observerte jeg at begge kjønn var godt representert og at ulike kjønn også hadde varierende alder.

I spørreundersøkelsen ble lærerne spurt hvor lenge de har jobbet som naturfaglærer, antall studiepoeng i naturfag og ved hvilket klassetrinn de hadde mest erfaring. Som vi skal se senere var alle gruppene representert bra i spørreundersøkelsen. Det ble ikke funnet data som viser fordelingen av naturfaglærernes undervisningserfaring i Norge og det er derfor

vanskelig å si noe om utvalget er representativt for populasjonen. Når det gjelder antall studiepoeng i naturfag hos lærere i Norge har én av fire lærere mer enn 60 studiepoeng i naturfag, ifølge en rapport fra Statistisk sentralbyrå (Lagerstrøm, Moafi & Revold, 2014). I min spørreundersøkelse svarte 18 % at de hadde over 61 studiepoeng i naturfag, noe som er litt lavere enn populasjonen. Det er likevel vanskelig å gjøre en sammenligning av dataene da tallene fra Statistisk sentralbyrå bruker intervallet ”60 studiepoeng eller mer” mens denne spørreundersøkelsen bruker intervallet ”over 61 studiepoeng”. Dette medfører en overlapping for lærere med nøyaktig 60 studiepoeng. Tall fra statistisk sentralbyrå sier at 30 % av lærere i Norge har mellom 30 og 60 studiepoeng i naturfag (Lagerstrøm et al., 2014) sammenlignet med 32 % i min spørreundersøkelse. Dette kan antyde en likhet mellom utvalget og populasjonen. Det ble ikke funnet tall på hvor mange naturfaglærere det fins i Norge fordelt på klassetrinn, men tall hentet fra Statistisk sentralbyrå viser at elevmengden er fordelt 70% på barnetrinn og 30 % på ungdomstrinn (Statistisk sentralbyrå, 2017). I min undersøkelse svarte 61 % at de har undervist mest på barnetrinnet og 39 % har undervist mest på ungdomstrinnet. Det vil være vanskelig å sammenligne disse tallene da de er basert på henholdsvis elevgrupper og lærere. Det kan likevel gi en indikasjon på at det fins flere naturfaglærere på barnetrinnet enn på ungdomstrinnet i populasjonen, basert på antall elever.

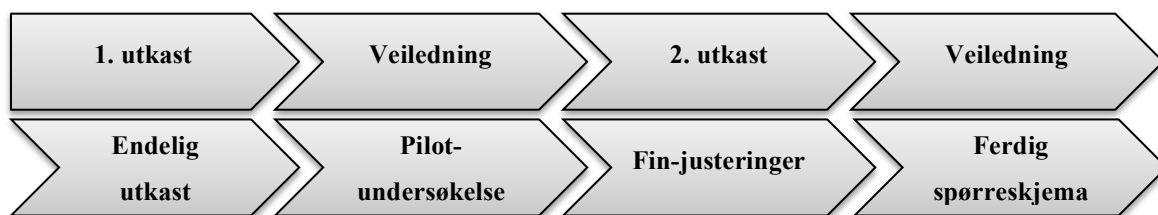
Det er selvfølgelig vanskelig å slå fast om utvalget i spørreundersøkelsen er representativt for naturfaglærere i Norge. Det er også vanskelig å gjøre noen statistisk generalisering på bakgrunn av resultatene mine. Men det vi kan si er at enhetene i spørreundersøkelsen representerer et bredt utvalg av naturfaglærere, både fordi de kommer fra ulike skoler i Troms og Nordland, men også på bakgrunn av undervisningserfaring, antall studiepoeng og klassetrinn. Dette er med å styrke validiteten i datamaterialet og vil kanskje kunne gi noen antydninger til trender blant naturfaglærere i Norge.

### **3.3.2 Spørreundersøkelsen**

Spørreskjemaet ble laget i papirutgave og respondentene besvarte denne skriftlig. Grunnen til dette var at denne metoden var best gjennomførbart i den type setting som datainnsamlingen pågikk i. Spørreskjemaet er betegnet som et prekodet spørreskjema, da det har oppgitte svaralternativer på alle spørsmålene. Fordelen med å bruke prekodete spørreskjema er at det blir enkelt for respondenten å fylle ut skjemaet. Samtidig blir prosessen med registrering av svar enklere for forskeren. Ulempen med prekodete spørreskjema er at det ikke er mulig å

fange opp informasjon utover det som spørres om i spørreskjemaet og det kan virke hemmende for respondentene. Spørsmålene i spørreundersøkelsen består av lukkede spørsmål hvor respondentene hovedsakelig avga svar på en likert-skala hvor verdiene ”aldri”, ”nesten aldri”, ”iblant”, ”nesten alltid” og ”alltid”. Målenivået for disse spørsmålene er på ordinalnivå, som kjennetegnes ved at svaralternativene er gjensidig utelukkende samtidig som at de har en logisk rangering (Johannessen et al., 2016).

Spørreskjemaet ble i hovedsak utarbeidet på bakgrunn av Hattie og Timperleys (2007) modell for effektiv tilbakemelding (teoretisk ramme). Utgangspunktet for utforming av spørreskjemaet var undersøkelsens problemstilling og det var viktig at spørsmålene ble formulert slik at de ga svar på den (Johannessen et al., 2016). Dette var derfor også hovedfokuset i prosessen. Spørreskjemaet ble til gjennom flere prosesser, noe som illustreres i figur 3.1.



Figur 3.1: Utarbeidelse av spørreskjema.

Det er flere faktorer en må tenke på når en skal utforme et spørreskjema. Spørsmålene skal være så konkrete som mulig for at det skal være enklere å besvare og at dataen blir enklere å tolke. Samtidig er det viktig å finne fram til relevante spørsmål som kan besvare det som skal undersøkes (Johannessen et al., 2016). Derfor var utformingen av spørreundersøkelsen en svært viktig prosess i denne masteroppgaven. Mellom de ulike utkastene av spørreskjemaet fikk jeg god veiledning med tanke på hvilke type spørsmål jeg stilte, spørsmålsformulering og utforming av svaralternativer. For å sikre meg at spørsmålene var konkrete og enkle å forstå ble det også gjennomført en pilotundersøkelse på to lærere. Disse lærerne var naturfaglærere med ulik bakgrunn og deltok ikke på konferansen. Tilbakemeldingene jeg fikk fra pilotundersøkelsen var stort sett positive og de siste finjusteringene ble gjort før ferdig spørreskjema kunne klargjøres for gjennomførelse.

Spørsmålene som ble stilt i spørreundersøkelsen er utarbeidet på bakgrunn teoretisk perspektiv om formativ vurdering (Hattie & Timperley, 2007; Wiliam, 2011). De første

spørsmålene i spørreundersøkelsen er ment som overordnende spørsmål som sier noe generelt om lærernes vurdering i praktisk naturfag. De neste spørsmålene i spørreundersøkelsen er utarbeidet på bakgrunn av de tre sentrale spørsmålene (Hvor skal jeg? Hvor er jeg nå? Hvordan kommer jeg videre?) som god formativ vurdering bør besvare (Hattie & Timperley, 2007; Wiliam, 2011). Det ble også utarbeidet noen spørsmål på bakgrunn av de fire ulike nivåene (oppgavenivå, prosess, selvregulering og personlig nivå) tilbakemeldingen kan virke på (Hattie & Timperley, 2007). Formålet med disse spørsmålene var å kartlegge om lærerne ga tilbakemelding på flere av nivåene eller om noen av disse ble utelukket i vurderingsarbeidet. Til slutt ønsket jeg også å spørre lærerne om de legger til rette for egen- og hverandrevurdering blant elevene. Dette fordi elevmedvirkning i vurderingsarbeidet har vist seg å ha stor effekt for elevenes læring (Black et al., 2003; Bruce, 2001; Wiliam, 2011). Det siste spørsmålet; ”Hvor ofte får elevene mulighet til å jobbe videre med tilbakemeldinger på eget arbeid?” ønsket jeg å stille på grunn av viktigheten av at elevene får mulighet til å arbeide videre med tilbakemeldinger (Black et al., 2003; Bruce, 2001).

Definisjon på praktisk arbeid ble også presentert for deltakerne i forskningsprosjektet for å sikre en felles forståelse for hva praktisk arbeid innebærer.

### **3.4 Dataanalyse**

I dataanalysen ble det brukt statistikkprogrammet *IBM SPSS Statistics Versjon 24*. Ulike analyseverktøy ble benyttet i arbeidet med datamaterialet, som deskriptiv statistikk, t-tester, krysstabeller og variansanalyser. Deskriptiv statistikk ble brukt for å sjekke datamaterialet for feil og å beskrive datamaterialets karakteristikk gjennom for eksempel minimumsverdier og maksimumsverdier, gjennomsnitt, standardavvik, kurtosis og skjevhet (skewness). Det ble utført deskriptiv statistikk på bakgrunnsvariablene for å beskrive utvalget, samt på de enkelte variablene i spørreundersøkelsen for å få undersøke data i sin helhet i forkant av videre undersøkelser. Resultatene i denne oppgaven er basert på deskriptive analyser og variansanalyser og jeg har derfor utelatt metodebeskrivelse av andre analyseverktøy som ble brukt i prosessen.

### 3.4.1 Innmating og koding av data

Det ble foretatt en manuell innmating av datamaterialet fra de skriftlige besvarelsene til SPSS. Hver besvarelse ble dobbeltsjekket for å sikre korrekt innmating av data. På forhånd ble alle spørsmålene og svaralternativene i spørreundersøkelsen kodet og deretter overført til kodeskjema i SPSS. I alt har jeg studert 23 variabler fordelt på 61 enheter.

### 3.4.2 Datareduksjon

Det var spesielt tre faktorer jeg ønsket å se nærmere på i etterkant av de deskriptive analysene; (1) i hvor stor grad lærerne i spørreundersøkelsen praktiserte formativ vurdering; (2) i hvor stor grad lærerne benyttet seg av de ulike nivåene i Hattie og Timperleys modell (2007) for effektiv tilbakemelding; og (3) i hvor stor grad lærerne la til rette for egen- og hverandrevurdering. Disse egenskapene kan bestemmes ut fra flere variabler i spørreskjemaet og jeg valgte derfor å slå sammen flere av spørsmålene til tre ulike konstrukt. Dette ble gjort ved å regne ut gjennomsnittsskåren av hva hver enkelt respondent hadde svart på spørsmålene tilknyttet konstruktene. Utarbeidelsen av de tre konstruktene vil presenteres i de neste avsnittene.

#### Formativ vurdering

Som skrevet tidligere skal god formativ vurdering besvare tre spørsmål; Hvor skal eleven?; Hvor er eleven nå? og Hvordan nå dit?(Hattie & Timperley, 2007; Wiliam, 2011). For å undersøke om deltakerne oppfyller disse kravene slo jeg sammen følgende spørsmål fra spørreundersøkelsen til det nye konstruktet *god formativ vurdering*:

- (1) *Hvor ofte gir du elevene læringsmål?*
- (2) *Hvor ofte gir du elevene kjennetegn på måloppnåelse?*
- (3) *Hvor ofte gir du elevene eksempler på måloppnåelse?*
- (4) *Hvor ofte lar du elevene bidra med å bestemme læringsmål?*
- (5) *Hvor ofte gir du tilbakemelding på hva eleven har forstått/lært av de aktuelle læringsmålene?*
- (6) *Hvor ofte gir du tilbakemelding på hva eleven ikke har forstått/lært av de aktuelle læringsmålene?*
- (7) *Hvor ofte gir du konkrete tilbakemeldinger på hva eleven må gjøre for å forbedre seg?*

Konstruktet forteller oss gjennomsnittlig skåre for alle nevnte spørsmål fra hver enkelt respondent i spørreundersøkelsen. Her sier de fire første spørsmålene noe om det første spørsmålet i god formativ vurdering; Hvor skal eleven?. Spørsmål fem og seks sier noe om det neste spørsmålet i god formativ vurdering; Hvor er eleven nå?. Spørsmål 7 tilsvarer det siste spørsmålet i god formativ vurdering; Hvordan nå dit?. Cronbach's alfa mellom disse spørsmålene er 0,83. Forklaring og vurdering av Cronbach's alfa vil bli gjort i kapittel 3.5.2 om reliabilitet. Det ble foretatt deskriptiv statistikk for det nye konstruktet for god formativ vurdering og finner at N=61 med en minimumsverdi på 1,00 og en maksimumsverdi på 4,43. Gjennomsnittsverdien for alle respondentene er 3,37 og standardavviket er 0,61.

### **Fire nivåer av tilbakemelding**

I følge Hattie og Timperleys modell (2007) kan hvert av de tre spørsmålene for god tilbakemelding fungere på de fire nivåene oppgavenivå, prosessnivå, selvreguleringsnivå og personlig nivå. Som skrevet i teorikapitlet bør samtlige av de tre første nivåene virke sammen for å oppnå en effektiv tilbakemeldingspraksis. Samtidig er tilbakemelding på et personlig nivå ikke særlig læringsfremmende. Jeg ønsket derfor å lage en ny variabel for faktoren *fire nivåer* ved å legge sammen følgende spørsmål fra spørreundersøkelsen:

- (1) *Hvor ofte gir du tilbakemelding på elevens svar på oppgaven og/eller forståelse av oppgaven?*
- (2) *Hvor ofte gir du tilbakemelding på elevens fremgangsmåte for å løse oppgaven?*
- (3) *Hvor ofte gir du tilbakemelding som oppmuntrer eleven til selv å reflektere over oppgaven og avdekke eventuelle feil/mangler?*
- (4) *Hvor ofte gir du tilbakemelding i form av ros, som f.eks. "bra", "du er flink", "bra jobbet"?*
- (5) *Hvor ofte gir du tilbakemelding i form av kritikk, som f.eks. "ikke bra", "dårlig"?*

Da de to siste spørsmålene her vil ha motsatt effekt enn de foregående spørsmålene lagde jeg motsatte verdier for disse spørsmålene før de ble tatt med i den nye variabelen. En høy gjennomsnittsscore for denne variabelen kan derfor tolkes som at lærerne er gode på å gi tilbakemelding på flere nivåer og på en måte som virker læringsfremmede. Cronbach's alfa for disse spørsmålene er 0,44. Det ble foretatt deskriptiv statistikk for det nye konstruktet hvor

N=60 med en minimumsverdi på 2,40 og en maksimumsverdi på 5,00. Gjennomsnittsverdien for alle respondentene er 3,68 og standardavviket er 0,47.

### **Egen- og hverandrevurdering**

Hattie og Timperleys modell (2007) sier ikke noe om egenvurdering blant elever eller elevmedvirkning i vurderingsarbeidet. Jeg ønsket likevel å undersøke dette i spørreundersøkelsen da dette er viktig for å få til god formativ vurdering (Black et al., 2003; Bruce, 2001; Wiliam, 2011). Jeg slo sammen de to spørsmålene som gikk på egen- og hverandrevurdering i spørreundersøkelsen:

*(1) Hvor ofte legger du til rette for at elevene vurderer seg selv?*

*(2) Hvor ofte legger du til rette for at elevene vurderer hverandre?*

Cronbach's alfa for disse spørsmålene er 0,68. Det ble foretatt deskriptiv statistikk for det nye konstruktet og fant at N=60 med en minimumsverdi på 1,00 og en maksimumsverdi på 4,00. Gjennomsnittsverdien for alle respondentene er 2,82 og standardavviket er 0,61.

### **3.4.3 Variansanalyser**

For å kunne sammenlikne flere variabler ble det gjennomført variansanalyser (One-Way ANOVA) i SPSS. Til forskjell fra t-tester hvor man sammenligner resultater fra to ulike grupper, gir ANOVA mulighet for å sammenlikne gjennomsnittskårer fra flere enn to grupper. I denne oppgavens tilfelle gir det meg mulighet til å sammenlikne gjennomsnittskårer for de ulike variablene fordelt på respondentenes klassetrinn, studiepoeng og undervisningserfaring. Variansanalyser sammenligner variansen mellom de ulike gruppene med variansen i hver enkelt gruppe. Ut av ANOVA får vi en F-ratio som representerer variansen mellom gruppene dividert på variansen innen hver gruppe. En høy verdi av F indikerer at det er større varians mellom gruppene enn det er innen hver av gruppene (Pallant, 2010). Dersom F er signifikant (p-verdien er mindre enn .05) kan vi med statistisk sikkerhet si at det er forskjeller mellom gruppene. Ikke-signifikans antyder at forskjellene kan være tilfeldige.

Det er en forutsetning for ANOVA at variansen i gruppene er homogene. Dette kan vi teste ved en Levene's test, som tester nullhypotesen om at variansen i gruppene er like. Dersom Levene's test er signifikant (p-verdien er mindre enn .05) vil det si at variansen for gruppene

er signifikant forskjellig fra hverandre og vi har da brutt antakelsen om homogen varians. For å gå videre i analysen må vi korrigere dette problemet. I tilfellene hvor nullhypotesen er brutt har jeg i resultatene mine valgt å korrigere F-ratio ved å gjennomføre en Welch's F test, som er ansett å være robust når homogenitet i variansen er brutt.

ANOVA forteller oss om det er signifikante forskjeller mellom gruppene, men den sier ikke noe om hvilke grupper som er signifikant forskjellig fra hverandre. Til dette kan en gjennomføre en post-hoc test. Avhengig av om det er lik varians i gruppene eller om antakelsen om homogen varians er brutt er det benyttet to ulike post-hoc tester. Med lik varians er det brukt Tukey's HSD test, som forutsetter homogen varians, og i de tilfellene hvor antakelsen om homogen varians er brutt er det brukt Games-Howell test, som ikke forutsetter lik varians.

En forutsetning for ANOVA er at en har uavhengige og randomiserte utvalg, normalfordeling og homogen varians i gruppepopulasjonene. Det er likevel gjort beviser på at ANOVA er så robust at skjevhet, kurtosis og ikke-normalitet har liten effekt på resultatene i en two-tailed test (Field, 2013, s. 444). ANOVA er også robust ved brudd på homogen varians, men forutsetter at størrelsen på gruppene er like (Field, 2013, s. 445). I denne undersøkelsen har vi sett at gruppestørrelsene er ganske like, med noe variasjon ved fordelingen av antall studiepoeng.

Det har blitt diskutert om parametriske tester (som ANOVA) kan brukes med data basert på en Likert-skala. Ekspertene har likevel fastslått at dersom utvalgsstørrelsen er stor nok (minst 5-10 observasjoner per gruppe) og dataen er tilnærmet normalfordelt, kan man bruke parametriske tester med Likert-skala (Sullivan & Artino Jr, 2013). Undersøkelser har også vist at parametriske tester generelt er mer robuste enn ikke-parametriske tester selv når statistiske forutsetninger, som normalfordeling av data, er brutt (Sullivan & Artino Jr, 2013).

## **3.5 Validitet og reliabilitet**

### **3.5.1 Validitet**

Dataen som samles inn er ikke selve virkeligheten, men en representasjon av den. Det er derfor viktig å vurdere hvor godt datamaterialet representerer fenomenet. I forskningsverden brukes begrepet validitet, eller gyldighet (Johannessen et al., 2016, s. 66).



*Begrepsvaliditet* går ut på om vi faktisk måler det teoretiske begrepet vi ønsker å måle (Ringdal, 2013, s. 96). I mitt tilfelle handler det om at spørsmålene i spørreundersøkelsen faktisk måler begrepene i Hattie og Timperleys modell (2007) om effektiv tilbakemelding, som jeg har valgt som teoretisk perspektiv for undersøkelsene. Spørsmålenes gyldighet kan selvsagt diskuteres da de kun tar utgangspunkt i egen forståelse av de teoretiske begrepene og ikke har tidligere undersøkelser å støtte seg på. I etterkant av undersøkelsen oppdaget jeg at noen av spørsmålene ikke ga svar på det jeg faktisk undersøkte og disse ble vektlagt mindre i analysene. I tillegg kan noen av spørsmålene i spørreundersøkelsen tolkes ulikt av ulike deltakere, noe som svekker validiteten til disse spørsmålene. Dette gjelder kanskje spesielt spørsmålene om i hvor stor grad lærerne ga muntlig eller skriftlig tilbakemelding i praktisk arbeid, noe som kan tolkes ulikt av den enkelte lærer. Det ble likevel jobbet grundig med forståelse av teoretisk perspektiv i forkant og underveis av utformingen av spørreskjema for å sikre at spørsmålene måler begrepene riktig, noe som er med å styrke validiteten.

Det er vanskelig å ha fullt samsvar mellom teoretisk og operasjonell validitet. Derfor kan vi også undersøke *innholdsvaliditeten* (content validity). Innholdsvaliditet går ut på om målet dekker de viktigste aspektene ved det teoretiske begrepet (Ringdal, 2013, s. 98). Som vi har sett i ulike teorier har formativ vurdering en bred betydning og det kan være krevende å dekke alle aspektene i en kort spørreundersøkelse. Det er likevel noen elementer som viser seg å være sentrale for god formativ vurdering. Den skal si noe om hvor eleven skal, hvor eleven står og hvordan eleven kan komme seg videre (Hattie & Timperley, 2007; Wiliam, 2011). Det er også disse elementene jeg særlig valgte å undersøke nærmere i spørreundersøkelsen. Jeg kan likevel ikke konkludere med at undersøkelsen dekker alle aspektene ved formativ vurdering, noe som svekker validiteten i datamaterialet.

I kvalitative undersøkelser vil det være aktuelt å diskutere *statistisk validitet* og *ytre validitet*. Statistisk validitet sier noe om det er mulig å generalisere resultatene fra utvalg til populasjon (Johannessen et al., 2016, s. 389). Vi må da vurdere om utvalget er representativt for populasjonen. Dette ble diskutert i kapittel 3.3.1 og en kan med forsiktighet si at utvalget kan gi antydninger for hele populasjonen, men å generalisere resultatene vil være vanskelig med mitt utvalg. Samtidig hadde jeg en svarprosent på over 70 % av bruttoutvalget. Dersom en kan argumentere for at utvalget er representativt for populasjonen er dette også med å styrke den statistiske validiteten, da bortfallet av respondenter er relativt lite. Det er likevel

usikkerhet rundt hvorfor noen lærere valgte å ikke besvare undersøkelsen og dersom bortfallet av respondenter viser seg å være systematisk vil dette helt klart ha vært med å påvirke resultatenes gyldighet.

Den ytre validiteten forutsetter statistisk validitet og sier noe om i hvilken grad resultatene kan overføres i rom og tid (Johannessen et al., 2016, s. 389). Dette handler om å vurdere hvorvidt resultatene kan si noe om hvordan naturfaglærere i andre deler av landet praktiserer formativ vurdering. Skolesystemet i Norge har en lang tradisjon rundt enhetsskolen og det er fornuftig å tenke at dette også påvirker at lærerne i landet tenker likt rundt undervisning og vurdering. Vi har også nasjonale retningslinjer når det kommer til utdanning av lærere, noe som kan styrke tanken om at lærere tenker likt rundt dette. Det er likevel viktig å ta hensyn til at det kan finnes ulike vurderingskulturer rundt om i landet og siden undersøkelsen kun er basert på lærere i Troms og Nordland kan det tenkes at lærere fra andre kanter i landet vurderer ulikt.

Det kan være vanskelig å bestemme om indikatorer er valide eller ikke. I noen tilfeller kan det være aktuelt å bruke sunn fornuft – *face validity* (Johannessen et al., 2016, s. 67). På bakgrunn av den teorien som er fremlagt om formativ vurdering og de viktigste elementene som formativ vurdering skal inneholde er det dekningsfor å kunne si at dataen i denne oppgaven er valide. Dette kan begrunnes i at dataen forteller noe om i hvor stor grad læreren praktiserer formativ vurdering på bakgrunn av de tre nevnte spørsmålene som god formativ vurdering skal inneholde. Samtidig tar undersøkelsen opp noen flere aspekter ved formativ vurdering, som hvordan type vurdering elevene får (tilbakemelding på flere nivåer styrker læringsutbytte), og elevmedvirkning i vurderingsarbeidet.

### **3.5.2 Reliabilitet**

I tillegg til å vurdere dataens validitet er det også i forskning viktig å vurdere dens reliabilitet, eller pålitelighet. Reliabiliteten knyttes til nøyaktigheten av data, hvilken data som brukes, måten den samles inn på og hvordan de bearbeides (Johannessen et al., 2016, s. 36). Dataens reliabilitet påvirkes av tilfeldige målefeil og en ønsker selvsagt å forhindre for mange tilfeldige målefeil da høy reliabilitet er en forutsetning for høy validitet (Ringdal, 2013, s. 96). Det er flere måter å styrke reliabiliteten på, blant annet basert på *allmenn kildekritikk*, *test-retest-teknikk*, *interreliabilitet* og *intern konsistens* (Johannessen et al., 2016; Ringdal, 2013).

Dataen som ble brukt i denne undersøkelsen er samlet inn på bakgrunn av et selvlaget spørreskjema og det ble ikke funnet tilsvarende undersøkelser hvor det kunne være mulig å sammenligne resultatene for å styrke reliabiliteten. Derfor ble arbeidet med spørreskjema spesielt viktig i denne sammenhengen. Spørreskjema ble, som vist i kapittel 3.3.2, utarbeidet gjennom flere prosesser hvor både egen tilegnelse av teori, veiledning og pilotundersøkelser ble brukt til å forbedre kvaliteten. Dette kan sies å styrke reliabiliteten til datamaterialet. Dataen ble samlet inn på en konferanse hvor jeg selv fikk muligheten til å presentere prosjektet for respondentene. Det var også muligheter for å stille spørsmål til meg om eventuelle uklarheter i spørsmålene. Når det gjelder måten dataen ble samlet inn på fikk lærerne utdelt spørreskjemaet på slutten av et foredrag og brukte noen minutter på å svare på spørreskjemaet i pausen før neste foredrag. Det kan tenkes at noen av respondentene ikke brukte nok tid til å tenke gjennom alle spørsmålene da de besvarte spørreskjemaet og på den måten ikke krysset av på det riktige svaret for dem. Dette vil eventuelt være med å svekke reliabiliteten.

Ved å gjennomføre en test-retest, hvor en måler det samme to ganger og undersøker korrelasjonen mellom svarene, kan en styrke reliabiliteten. Dette var ikke aktuelt å gjennomføre i mitt arbeid da det sannsynligvis ville være svært vanskelig å få respondentene til å besvare undersøkelsen på nytt ved en senere anledning. Dette var i tillegg ikke hensiktsmessig med tanke på tidsaspektet. Det har ikke blitt gjort samme undersøkelse av andre forskere og det er derfor ikke mulig å styrke reliabiliteten ved hjelp av interreliabilitet.

I tverrsnittsundersøkelser kan det være aktuelt å vurdere reliabiliteten ved å måle graden av intern konsistens mellom indikatorene. Intern konsistens måles statistisk med Cronbach's alfa og varierer fra 0 til 1. Jo sterkere sammenhenger det er mellom indikatorene, jo høyere er reliabiliteten. En indeks har en tilfredsstillende reliabilitet dersom alfa har en verdi over 0,70 (Ringdal, 2013, s. 98). Som beskrevet i kapittel 3.4.2 valgte jeg å slå sammen flere spørsmål som jeg mente besvarte det samme. Cronbach's alfa for *god formativ vurdering* var på 0,83, noe som er tilfredsstillende for denne variabelen. Dette styrker derfor reliabiliteten til undersøkelsene basert på denne variabelen. For spørsmålene om de *fire nivåene av tilbakemelding* var cronbach's alfa 0,44. Dette er en del lavere enn hva som kan anses som tilfredsstillende, noe som svekker reliabiliteten til undersøkelsene basert på denne variabelen. Jeg ønsker likevel å bruke denne variabelen i videre undersøkelser fordi den kan gi noen

antydninger til hvordan lærerne skårer selv om reliabiliteten er svekket. Det vil tas hensyn til dette i diskusjonen. Cronbach's alfa for variabelen egen- og hverandrevurdering er 0,68 og nærmer seg tilfredsstillende verdi på 0,70. Jeg mener derfor at denne verdien styrker reliabiliteten for undersøkelser basert på denne variabelen.

### **3.6 Forskningsetikk**

Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH) har vedtatt nasjonale retningslinjer for å fremme god og ansvarlig forskning. Retningslinjene ble sist oppdatert i 2016 og er delt i seks deler fordelt på 46 punkter (NESH, 2016). Disse retningslinjene er blitt vurdert og tatt hensyn til underveis i forskningsprosessen. Jeg vil her ikke nevne alle retningslinjene men fokusere på de punktene jeg anser som viktigst for denne forskningen.

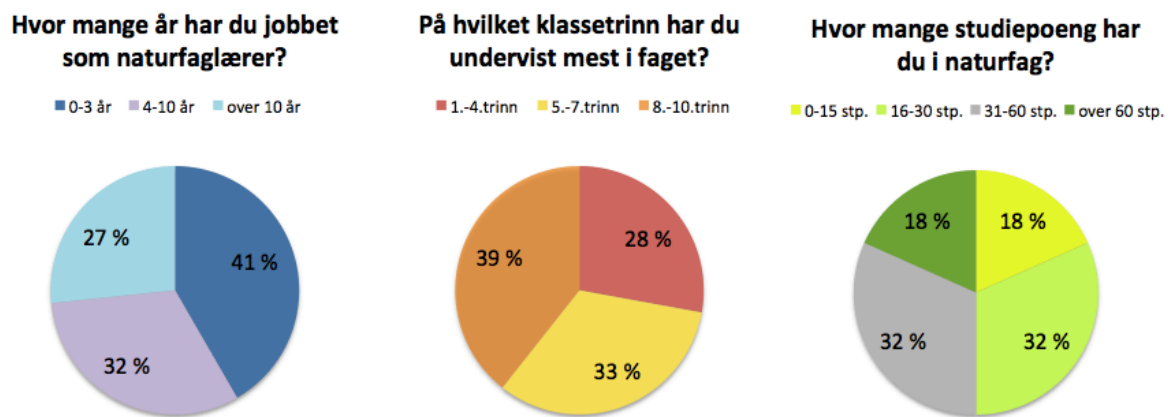
Et viktig prinsipp i samfunnsforskningen er å sikre at den som er gjenstand for forskning får god informasjon om hva som er forskningens hensikt, samt informert og fritt samtykke til å delta i forskningen. Dette ble sikret ved at jeg personlig fortalte litt om prosjektet i forkant av utdelingen av spørreskjema, samt at det ble laget et lite informasjonsskriv vedlagt på spørreskjema. Det ble også opplyst muntlig om frivillig deltakelse. Forskeren er også pliktig til å respektere informantens privatliv og deltakerne skal selv kunne bestemme hvilke opplysninger om seg selv som skal deles. Dette ble sikret ved at deltakerne stod fritt til å velge å ikke krysse av dersom det var noen spørsmål de ikke ønsket å besvare. Jeg så at lærerne i enkelte tilfeller hadde unnlatt å krysse av på én av bakgrunnsspørsmålene i spørreundersøkelsen. Det er usikkert om dette var en bevisst handling, men dette ivaretar respekten for fritt samtykke.

Gjennom spørreundersøkelsen var det ikke mulig at enkeltpersoner kunne identifiseres og innebar ikke behandling av personopplysninger. Prosjektet var derfor ikke meldepliktig. Prosjektet har derfor også sikret konfidensialitet i behandling av informasjon da dataen allerede var anonymisert da de leverte inn spørreskjema. Jeg som forsker har ingen mulighet til å spore spørreskjema tilbake til respondenten.

## 4 Resultat

### 4.1 Oversikt over respondentene i spørreundersøkelsen

Det ble spurt tre bakgrunnsspørsmål til respondentene i spørreundersøkelsen; hvor mange år de har jobbet som naturfaglærer; hvilket klassetrinn de har jobbet mest ved og hvor mange studiepoeng de har i naturfag. Jeg ønsket å undersøke om dette er faktorer som er med å påvirke lærernes vurderingspraksis. Jeg vil først presentere en oversikt over hvordan respondentene fordelte seg på disse tre spørsmålene.



Figur 4.1: Oversikt over hvordan respondentene fordeler seg på spørsmålene om år som naturfaglærer, klassetrinn og antall studiepoeng i naturfag.

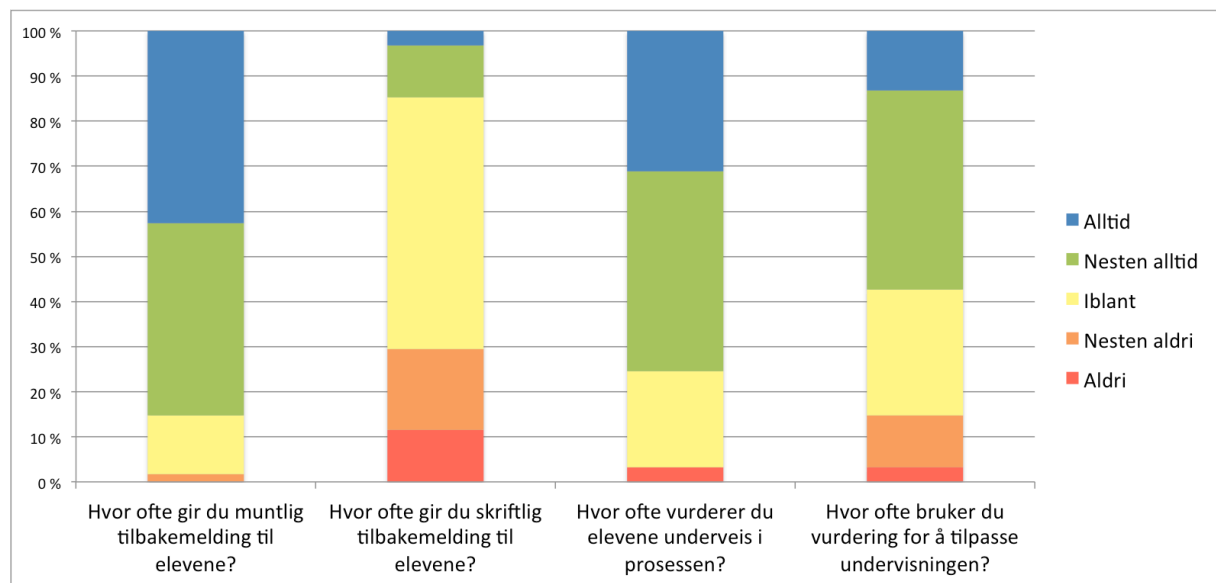
Spørreundersøkelsen hadde i utgangspunktet fire verdier for bakgrunnsvariabelen ”Hvor mange år har du jobbet som naturfaglærer?”. Da svarprosenten for verdien ”over 21 år” var mye lavere enn de andre verdiene (fire hadde svart at de hadde jobbet over 21 år), ble verdiene ”11-20 år” og ”over 21 år” slått sammen til en ny verdi ”over 10 år” for å få en jevnere fordeling i analysene. Man kan anta at lærere som har jobbet mer enn 10 år som naturfaglærer i stor grad er etablert i yrket og at det ikke er store skiller mellom lærere som har jobbet 11-20 år og de som har jobbet over 21 år, og jeg tenker derfor at det er greit å slå sammen disse verdiene. Fordelingen av hvor mange år respondentene har jobbet som naturfaglærer er forholdsvis jevnt fordelt, med en liten overvekt av lærere som har jobbet 0-3 år som naturfaglærer. Respondentene har også fordelt seg forholdsvis jevnt for bakgrunnsvariabelen ”På hvilket klassetrinn har du undervist mest i faget?”, men litt flere deltakere i spørreundersøkelsen svarer at de har jobbet mest på ungdomstrinnet. For

bakgrunnsvariabelen "Hvor mange studiepoeng har du i naturfag?" er respondentene jevnt fordelt for verdiene "0-15 stp." og "over 60 stp." (begge 18 %). De har også fordelt seg jevnt mellom verdiene "16-30 stp." og "31-60 stp." (begge 32 %). Det er derfor litt ujevn fordeling for denne bakgrunnsvariabelen.

## 4.2 Lærernes formative vurderingspraksis i praktisk arbeid i naturfag

Dette kapitlet gir en oversikt over frekvensfordelingen i flere av spørsmålene i spørreundersøkelsen. I disse analysene er det ikke tatt hensyn til antall år lærerne har jobbet, klasstrinn og studiepoeng. Resultatene sier noe om hva naturfaglærere generelt vektlegger når de foretar formativ vurdering i den praktiske delen av naturfaget.

### 4.2.1 Generelle spørsmål om vurderingen

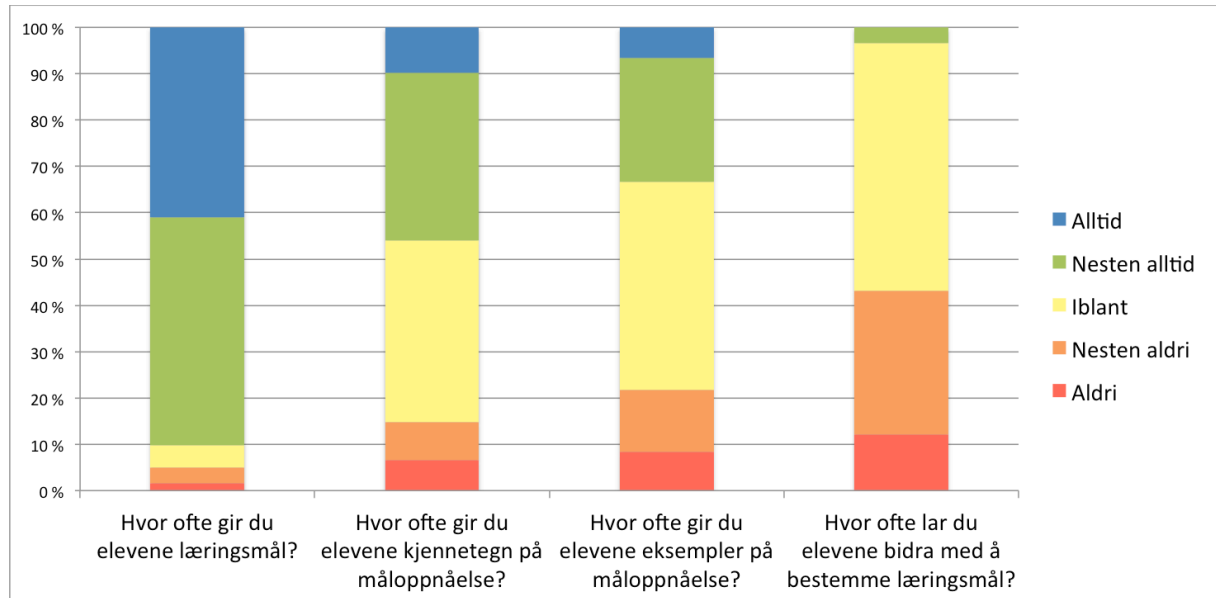


**Figur 4.2:** Frekvensfordeling av respondentenes besvarelser om muntlig og skriftlig tilbakemelding, undervisvurdering og bruk av vurdering for å tilpasse undervisningen i praktisk arbeid i naturfag.

Figur 4.2 viser at lærerne i mye større grad vurderer elevene muntlig enn skriftlig i den praktiske delen av naturfaget. Nesten 90 % svarer at de alltid eller nesten alltid vurderer elevene muntlig, mens bare rundt 15 % svarer at de alltid eller nesten alltid vurderer elevene skriftlig. Det er også en relativt stor andel som sier at de aldri eller nesten aldri vurderer elevene skriftlig. Figur 4.2 viser også at lærerne i stor grad vurderer elevene underveis i prosessen. Omtrent 75 % svarer alltid eller nesten alltid på dette spørsmålet. Litt færre lærere bruker vurdering for å tilpasse undervisningen (nesten 60 % svarer alltid eller nesten alltid).

### 4.2.2 "Hvor skal eleven?"

Det ble stilt fire spørsmål knyttet til lærernes bruk av læringsmål i forbindelse med praktisk arbeid i naturfag, som sier noe om hvilken informasjon elevene får knyttet til hvor de skal i undervisningsprosessen.



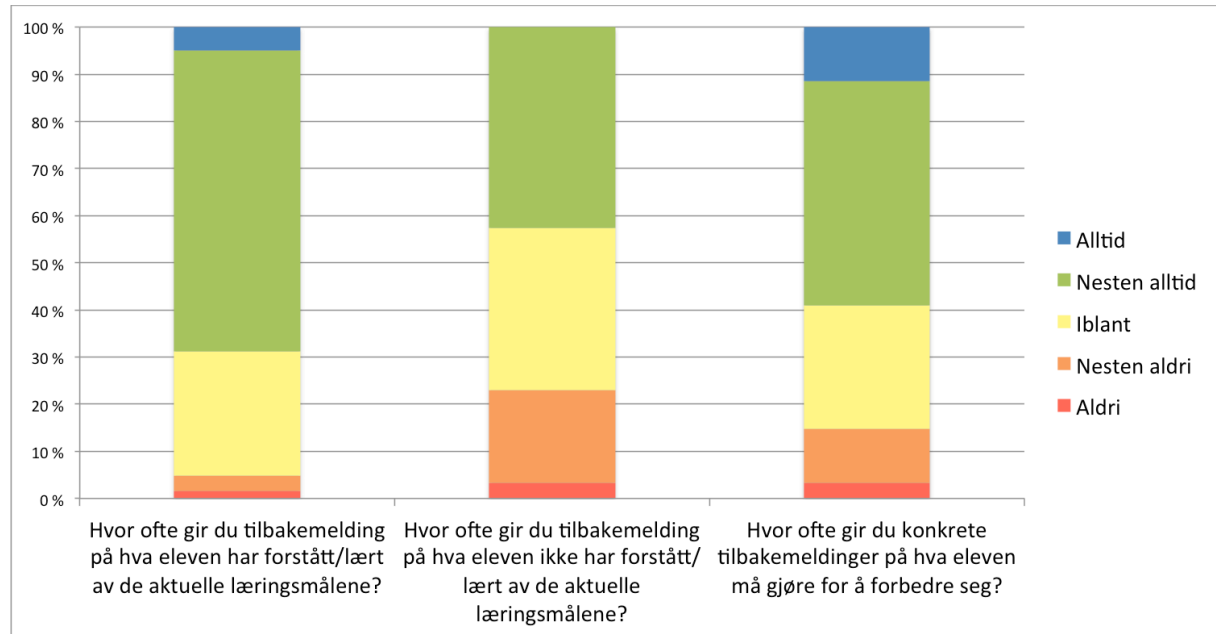
Figur 4.3: Frekvensfordeling av hvordan respondentene bruker læringsmål i praktisk arbeid i naturfag.

Omtrent 90 % av respondentene i spørreundersøkelsen svarer alltid eller nesten alltid på spørsmålet "Hvor ofte gir du elevene læringsmål?", noe som tyder på at dette er et stort fokus hos de fleste av lærerne. Resultatene viser videre at respondentene i mindre grad tydeliggjør læringsmålene med å gi elevene kjennetegn på måloppnåelse (omtrent 45 % svarer alltid eller nesten alltid på dette spørsmålet) og eksempler på måloppnåelse (litt over 30 % svarer alltid eller nesten alltid på dette spørsmålet). Resultatene viser også at svært få lærere ofte lar elevene bidra med å bestemme læringsmålene og over 40 % svarer at de aldri eller nesten aldri gjør dette.



### 4.2.3 "Hvor står eleven og hvordan jobbe videre mot målet?"

Det ble stilt noen spørsmål knyttet til hvordan lærere gir tilbakemelding på hvor elevene er i prosessen og hvordan de skal gå videre for å nå målet.

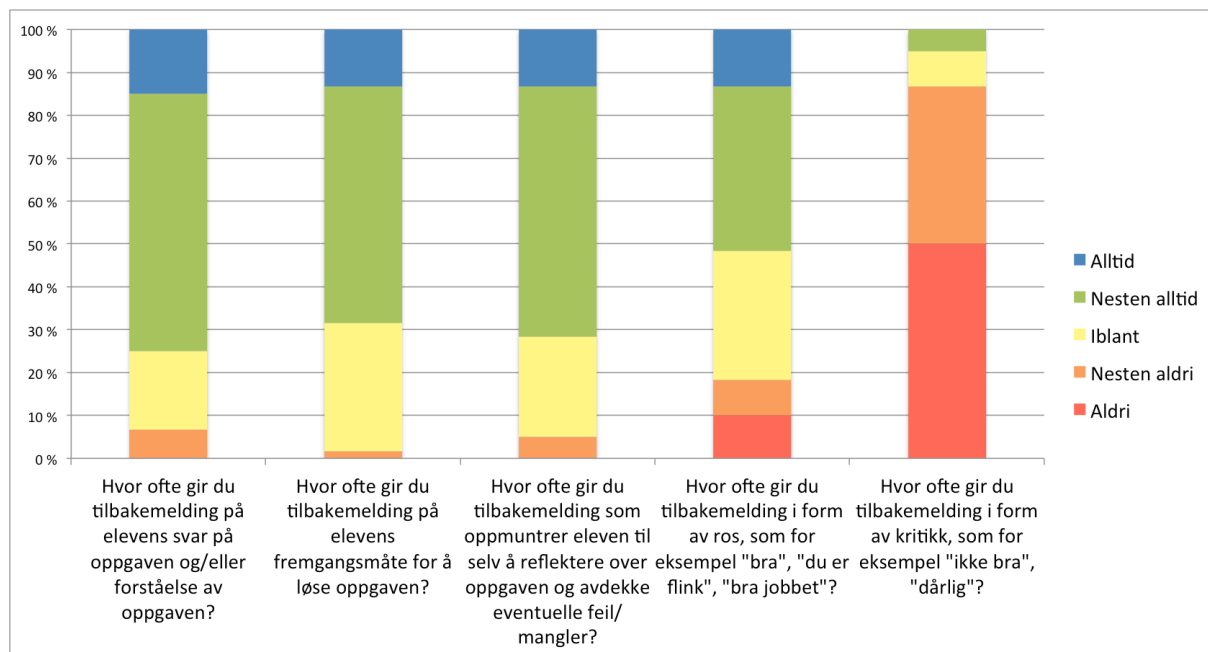


Figur 4.4: Frekvensfordeling av hvordan respondentene gir tilbakemeldinger på hvor eleven står og hvordan eleven skal forbedre seg i praktisk arbeid i naturfag.

Figur 4.4 viser at omtrent 70 % av lærerne svarer at de alltid eller nesten alltid gir tilbakemelding knyttet til hva eleven har forstått/lært. Dette knyttes opp mot det formative spørsmålet "Hvor er jeg nå i prosessen?". Litt færre gir tilbakemeldinger på hva eleven *ikke* har forstått/lært, noe som kan bety at lærerne har større fokus på å gi positive tilbakemeldinger enn negative. Omtrent 60 % har besvart at de alltid eller nesten alltid gir konkrete tilbakemeldinger på hva eleven må gjøre for å forbedre seg. Dette knyttes opp mot det siste formative spørsmålet "Hvordan skal jeg gå videre?".

### 4.2.4 Nivå av tilbakemelding

Spørreundersøkelsen inneholdt også noen spørsmål som kan relateres til de fire ulike nivåene tilbakemeldingen kan forekomme ved; *task level*, *process level*, *self-regulation level* og *self level*. Figur 4.5 viser hvordan respondentene besvarte spørsmålene i spørreundersøkelsen som var knyttet til disse.



**Figur 4.5: Frekvensfordeling av hvilke av de ulike nivåene av tilbakemelding respondentene bruker i praktisk arbeid i naturfag.**

Figur 4.5 viser at lærerne gir ganske like svar for hvor ofte de gir tilbakemelding på *task level*, *process level* og *self-regulation level* til elevene. Omtrent 70 % av respondentene svarer ”alltid” eller ”nesten alltid” og svært få svarer ”aldri” eller ”nesten aldri” på noen av disse spørsmålene. Dette kan tyde på at lærerne i stor grad anvender alle disse nivåene når de gir tilbakemelding på praktisk arbeid i naturfag. Det ble også spurt om hvor ofte lærerne ga tilbakemelding i form av ros og kritikk (*self level*). Her ser vi at mange svarer at de gir tilbakemelding i form av ros (over 50 % svarer alltid eller nesten alltid). Til motsetning er det svært få som gir tilbakemelding i form av kritikk og halvparten av respondentene svarer at de aldri gir kritikk. Tilbakemelding på *self level* har viser seg å være brukt relativt mye i lærernes tilbakemeldingspraksis.

#### 4.2.5 Egen- og hverandrevurdering

Det ble også spurt om hvor ofte lærerne legger til rette for egen- og hverandrevurdering. Figur 4.6 viser hvordan respondentene har besvart disse spørsmålene i spørreundersøkelsen.



**Figur 4.6: Frekvensfordeling av hvordan respondentene har besvart spørsmålene for om de legger til rette for egen- og hverandrevurdering i praktisk arbeid i naturfag.**

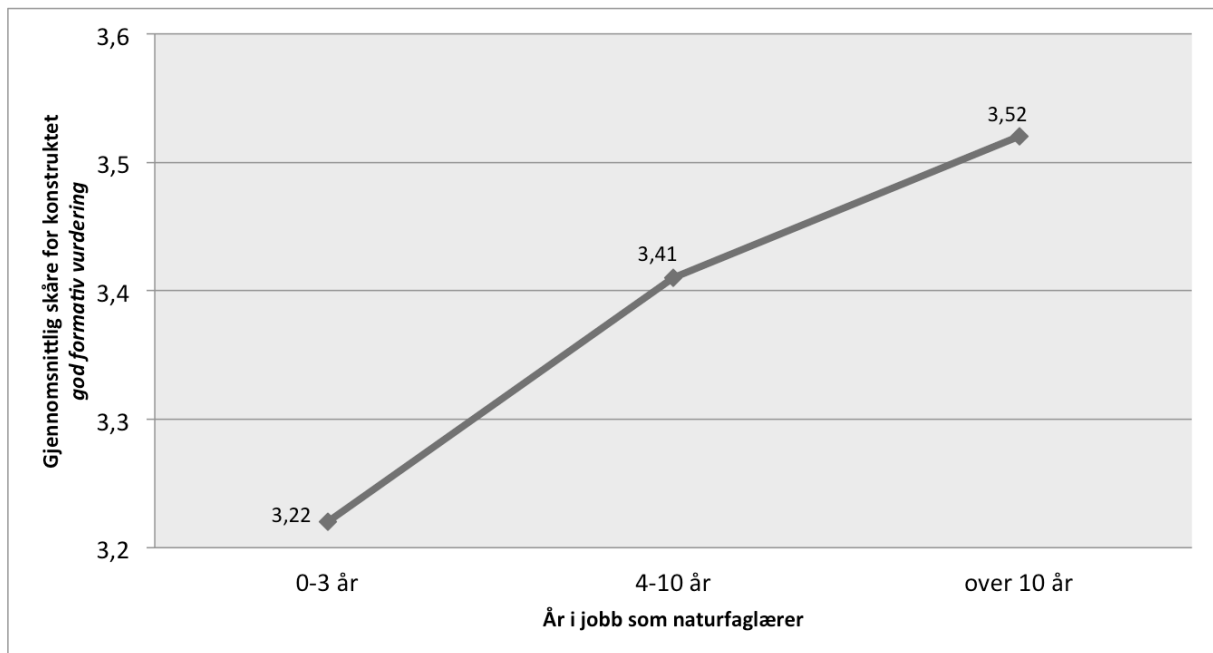
Figur 4.6 viser at egen- og hverandrevurdering har fått liten plass i vurderingsarbeidet i praktisk arbeid i naturfag blant respondentene i spørreundersøkelsen. Mange svarer ”iblant”, men svært få ser ut til å bruke egen- og hverandrevurdering jevnlig.

### **4.3 Vurderingspraksis knyttet til hvor lenge lærerne har jobbet som naturfaglærer**

T-tester at besvarelsene på mange av spørsmålene i spørreundersøkelsen antydte at de som har jobbet lenge som naturfaglærer skåret generelt høyere enn lærere som har jobbet kort tid. Dette ble undersøkt nærmere med variansanalyser i SPSS for de tre konstruktene *god formativ vurdering*, *nivå av tilbakemelding* og *egen- og hverandrevurdering*.

#### **4.3.1 År som naturfaglærer og god formativ vurdering**

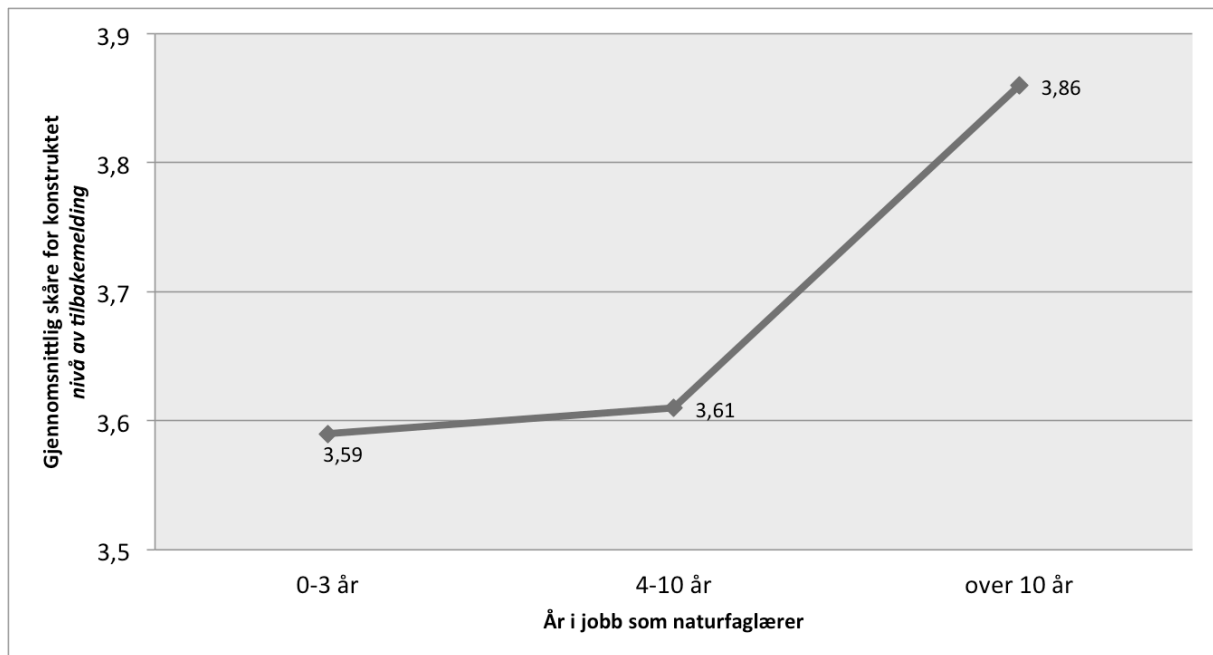
Om lærerne gir god formativ vurdering er blant annet tolket som at de gir tilbakemelding både i form av læringsmål, konkrete tilbakemeldinger og konkrete fremovermeldinger. Spørsmålene som angikk dette ble samlet i konstruktet *god formativ vurdering*. For å undersøke om det er sammenhenger mellom hvor mange år respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre på konstruktet *god formativ vurdering* ble det foretatt en variansanalyse for vilkårene 0-3 år, 4-10 år og over 10 år. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,671. Det vil si at signifikansen er større enn grensen på 0,05 og nullhypotesen om at variansen er homogen i alle grupper beholdes. ANOVA-analysen viser at det ikke var en signifikant effekt for hvor mange år i jobb som naturfaglærer og gjennomsnittlig svar på konstruktet *god formativ vurdering* på  $p < 0,05$  for de tre betingelsene [ $F(2,57)=1,264$  ,  $p=0,290$ ]. Det vil si at det ikke ble vist noen signifikante forskjeller på hvor lenge respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre på konstruktet *god formativ vurdering*. Grafen i figur 4.7 viser likevel at gjennomsnittlig skåre øker etter hvor mange år de har jobbet som naturfaglærer. Det kan antyde at lærere som har jobbet lengre som naturfaglærere også er bedre på formativ vurdering i praktisk naturfag, men som dataen viser er forskjellene på gjennomsnittsverdiene relativt små.



Figur 4.7: Gjennomsnittlig skåre på konstruktet *god formativ vurdering* i forhold til antall år som naturfaglærer.

#### 4.3.2 År som naturfaglærer og nivå av tilbakemelding

Konstruktet *nivå av tilbakemelding* sier noe om i hvor stor grad lærerne gir tilbakemeldinger både på oppgavenivå, prosessnivå, selvreguleringsnivå og personlig nivå. For å undersøke om det var sammenhenger mellom hvor mange år respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre for konstruktet *nivå av tilbakemelding* ble det foretatt en variansanalyse for vilkårene 0-3 år, 4-10 år og over 10 år. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,768. Det vil si at signifikansen er større enn grensen på 0,05 og nullhypotesen om at variansen er homogen i alle grupper beholdes. ANOVA-analysen viser at det ikke var en signifikant effekt for hvor mange år i jobb som naturfaglærer på gjennomsnittlig svar for spørsmålene om ulike nivå av tilbakemelding på  $p < 0,05$  for de tre betingelsene [ $F(2,56)=1,863$  ,  $p=0,165$ ]. Det var altså ingen signifikante forskjeller på hvor lenge respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre på konstruktet *nivå av tilbakemelding*. Tar likevel med grafen som illustrerer denne sammenhengen (figur 4.8) da det også i dette tilfelle antyder at de som har jobbet lenger som naturfaglærere skårer høyere enn de med kortere arbeidserfaring. I dette tilfelle vises den største forskjellen for de lærerne som har jobbet over 10 år sammenlignet med lærere som har jobbet inntil 10 år, men forskjellene i gjennomsnittsskåre er fortsatt små.

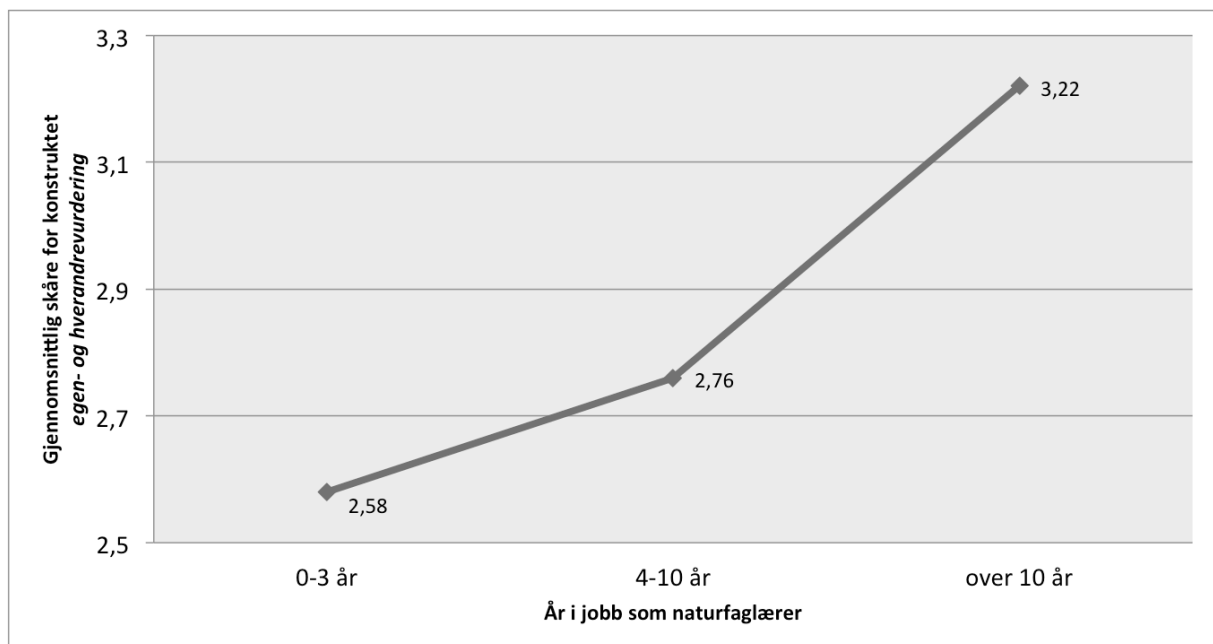


Figur 4.8: Gjennomsnittlig skåre for konstruktet *nivå av tilbakemelding* i forhold til antall år som naturfaglærer.

#### 4.3.3 År som naturfaglærer og tilrettelegging av egen- og hverandrevurdering

Jeg ønsket også å undersøke om lærerne legger til rette for egen- og hverandrevurdering i praktisk arbeid i naturfag og spørsmålene som angikk dette ble derfor samlet i konstruktet *egen- og hverandrevurdering*. For å undersøke om det var sammenhenger mellom hvor mange år respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre på konstruktet *egen- og hverandrevurdering* ble det foretatt en variansanalyse for vilkårene 0-3 år, 4-10 år og over 10 år. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,43. Det vil si at signifikansen er større enn grensen på 0,05 og nullhypotesen om at variansen er homogen i alle grupper beholdes. ANOVA-analysen viser at det var en signifikant effekt for antall år i jobb som naturfaglærer og gjennomsnittlig skåre for konstruktet *egen- og hverandrevurdering* på  $p < 0,05$  for de tre betingelsene [ $F(2,56)=6,139$ ,  $p=0,004$ ]. Post hoc-sammenlikninger ved bruk av *Tukey HSD test* viser at gjennomsnittlig resultat for 0-3 år ( $M=2,58$ ,  $SD=0,58$ ) var signifikant forskjellig fra de over 10 år ( $M=3,22$ ,  $SD=0,61$ ). Testen viser også at gjennomsnittlig resultat for 4-10 år ( $M=2,76$ ,  $SD=0,51$ ) var signifikant forskjellig fra de over 10 år. Gjennomsnittlig resultat for 0-3 år var derimot ikke signifikant forskjellig fra 4-10 år. Samlet viser resultatene at det er en forskjell på hvor mange år respondentene har jobbet som naturfaglærer og hvordan de legger til rette for egen- og hverandrevurdering blant elevene. Resultatene viser derimot bare en signifikant forskjell mellom de to første årsintervallene, 0-3 år og 4-10 år, sammenlignet med de lærerne som har

jobbet over 10 år som naturfaglærer. Resultatene antyder derfor at de som har jobbet over 10 år som naturfaglærer er flinkere å legge til rette for egen- og hverandrevurdering blant elevene enn lærere som har jobbet mindre enn 10 år. Forskjellene på hvor mange år respondentene har jobbet som naturfaglærer og gjennomsnittlig skåre for egen- og hverandrevurdering er illustrert i figur 4.9. Selv om forskjellene ikke var signifikante mellom lærere som har jobbet 0-3 år og de som har jobbet 4-10 år, ser vi likevel at gjennomsnittlig skåre også øker mellom disse vilkårene.



Figur 4.9: Gjennomsnittlig skåre for konstruert *egen- og hverandrevurdering* i forhold til antall år som naturfaglærer.

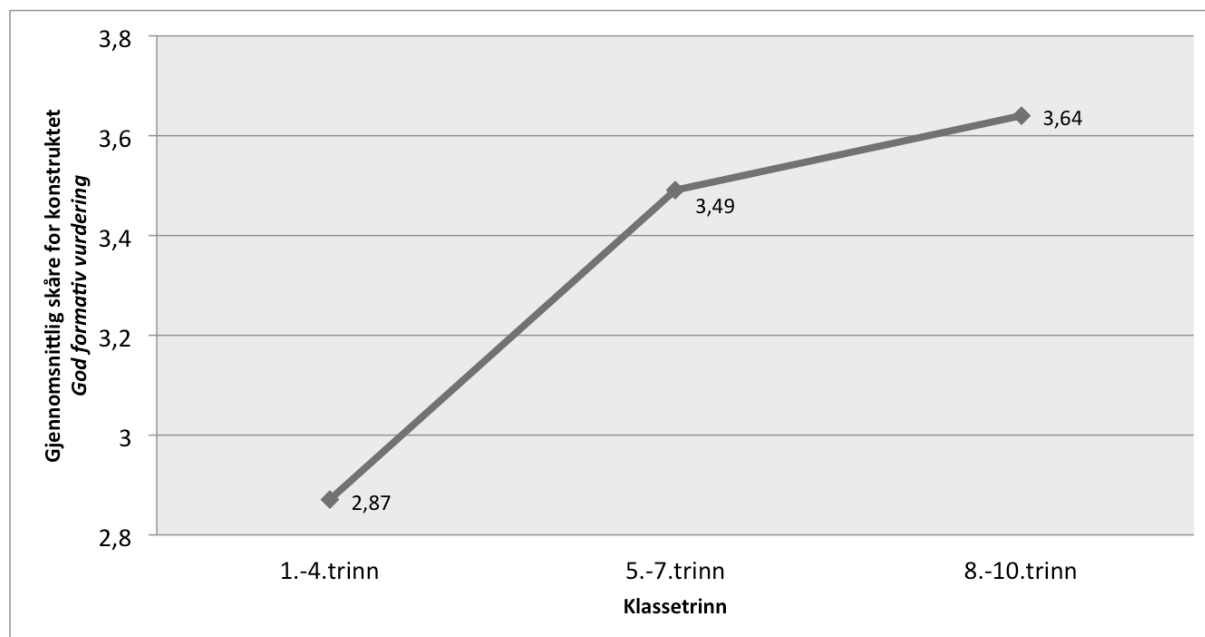
## 4.4 Vurderingspraksis knyttet til hvilket klassetrinn lærerne underviser på

Gjennom analyser ved t-tester ble det også gjort flere funn som kunne tyde på at det var forskjeller på hva respondentene hadde svart i spørreundersøkelsen etter hvilket klassetrinn de har undervist mest på. Det ble også her gjort variansanalyser for å undersøke dette nærmere.

### 4.4.1 Klassetrinn og god formativ vurdering

For å undersøke om det var sammenhenger mellom hvilket klassetrinn respondentene har jobbet mest på og gjennomsnittlig skåre på konstruktet *god formativ vurdering* ble det foretatt en variansanalyse for vilkårene 1.-4. trinn, 5.-7. trinn og 8.-10. trinn. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,029. Da signifikansnivået er mindre enn 0,05 oppfyller vi ikke kravet om homogenitet i variansen og vi må korrigere F ved å foreta en *Welch F test*. Welch F test viser at det var en signifikant effekt for hvilket klassetrinn de jobbet på og gjennomsnittlig svar for konstruktet *god formativ vurdering* på  $p < 0,05$  for [Welch's  $F(2, 31,76) = 7,445, p = 0,002$ ]. Det vil si at minst én av vilkårene er signifikant forskjellig fra de andre. Post hoc-sammenlikninger ved bruk av *Games-Howell* viser at gjennomsnittlig skåre for 1.-4. trinn ( $M = 2,87, SD = 0,76$ ) var signifikant forskjellig fra 5.-7. trinn ( $M = 3,49, SD = 0,50$ ). Testen viser også at gjennomsnittlig skåre for 1.-4. trinn var signifikant forskjellig fra 8.-10. trinn ( $M = 3,64, SD = 0,37$ ). Resultatene viser derimot ingen signifikant forskjell mellom 5.-7. trinn og 8.-10. trinn. Resultatene viser derfor at lærere som underviser på de to øverste trinnene skårer signifikant høyere på konstruktet *god formativ vurdering* enn lærere som underviser på det laveste trinnet (1.-4. trinn). Grafen som vist i figur 4.10 illustrerer også disse forskjellene.

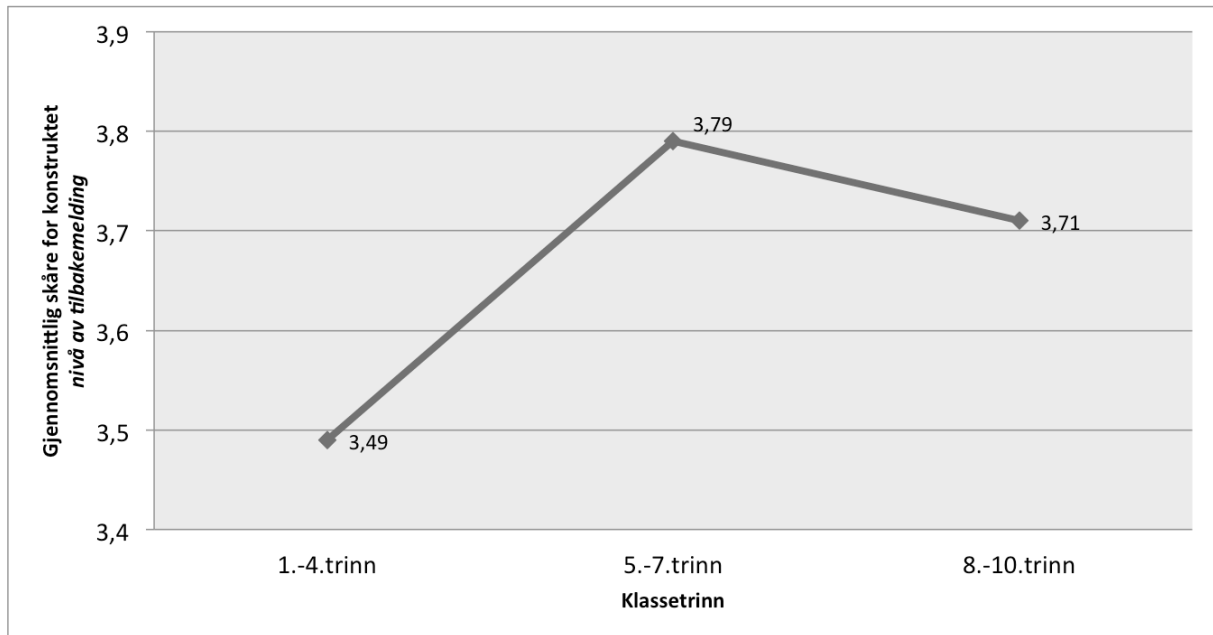




Figur 4.10: Gjennomsnittlig skåre for konstruktet *god formativ vurdering* i forhold til klassetrinn.

#### 4.4.2 Klassetrinn og nivå av tilbakemelding

For å undersøke om det var sammenhenger mellom hvilket klassetrinn respondentene har jobbet mest på og konstruktet *nivå av tilbakemelding* ble det foretatt en variansanalyse for villkårene 1.-4. trinn, 5.-7. trinn og 8.-10. trinn. Undersøkte homogenitet i variansen ved Levene's test og fikk en p-verdi på 0,73. Det vil si at signifikansen er større enn grensen på 0,05 og nullhypotesen om at variansen er homogen i alle grupper beholdes. ANOVA-analysen viser at det ikke var en signifikant effekt for hvilket klassetrinn respondentene har jobbet mest på og gjennomsnittlig skåre for konstruktet *nivå av tilbakemelding* på  $p < 0,05$  for de tre betingelsene [ $F(2,57) = 2,00$ ,  $p = 0,145$ ]. Det vil si at det ikke fins noen signifikant forskjell på hvilket klassetrinn respondentene har jobbet mest på og gjennomsnittlig skåre på konstruktet *nivå av tilbakemelding*. Grafen i figur 4.11 viser likevel at lærere på 1.-4. trinn skårer litt lavere enn lærere på de øverste trinnene.

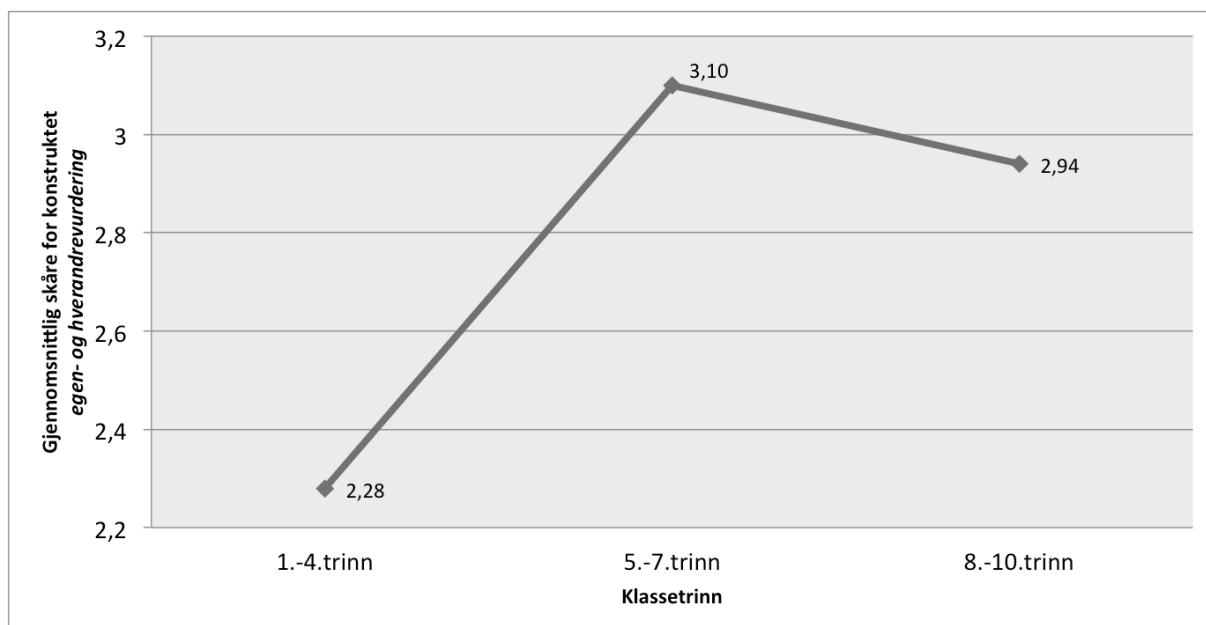


Figur 4.11: Gjennomsnittlig skåre for konstruktet *nivå av tilbakemelding* i forhold til klassetrinn.

#### 4.4.3 Klassetrinn og tilrettelegging av egen- og hverandrevurdering

For å undersøke om det var sammenhenger mellom hvilket klassetrinn respondentene har jobbet mest på og gjennomsnittlig skåre på konstruktet *egen- og hverandrevurdering* ble det foretatt en variansanalyse for vilkårene 1.-4. trinn, 5.-7. trinn og 8.-10. trinn. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,010. Da signifikansnivået er mindre enn 0,05 oppfyller vi ikke kravet om homogenitet i variansen og vi må korrigere F ved å foreta en *Welch F test*. Welch F test viser at det var en signifikant effekt for hvilket klassetrinn de jobber på og gjennomsnittlig skåre på konstruktet *egen- og hverandrevurdering* på  $p < 0,05$  for [Welch's  $F=(2, 31,62)=7,825$ ,  $p=0,002$ ]. Det vil si at minst én av vilkårene er signifikant forskjellig fra de andre. Post hoc-sammenlikninger ved bruk av *Games-Howell* viser at gjennomsnittlig skåre for 1.-4. trinn ( $M=2,28$ ,  $SD=0,71$ ) var signifikant forskjellig fra 5.-7. trinn ( $M=3,10$ ,  $SD=0,48$ ). Testen viser også at gjennomsnittlig skåre for 1.-4. trinn var signifikant forskjellig fra 8.-10. trinn ( $M=2,94$ ,  $SD=0,40$ ). Resultatene viser derimot ingen signifikant forskjell mellom 5.-7. trinn og 8.-10. trinn. Samlet viser analysen en signifikant forskjell for hvilket klassetrinn respondentene jobber på og gjennomsnittlig skåre for konstruktet *egen- og hverandrevurdering* mellom 1.-4. trinn og de to øverste trinnene. Resultatene viser ingen signifikant forskjell mellom de to øverste trinnene. Disse resultatene samsvarer med resultatene for forskjellene på klassetrinn og god formativ vurdering som vist i kapittel 4.4.1, og tilbakemeldingens nivå som vist i kapittel 4.4.2. Resultatene antyder derfor

at det er en forskjell på hvilket klassetrinn lærerne har jobbet mest som naturfaglærer, særlig mellom det laveste trinnet og de to øverste trinnene. Gjennomsnittlige skåre på spørsmålene om egen- og hverandrevurdering og hvilket klassetrinn de underviser på er illustrert i grafen i figur 4.12.



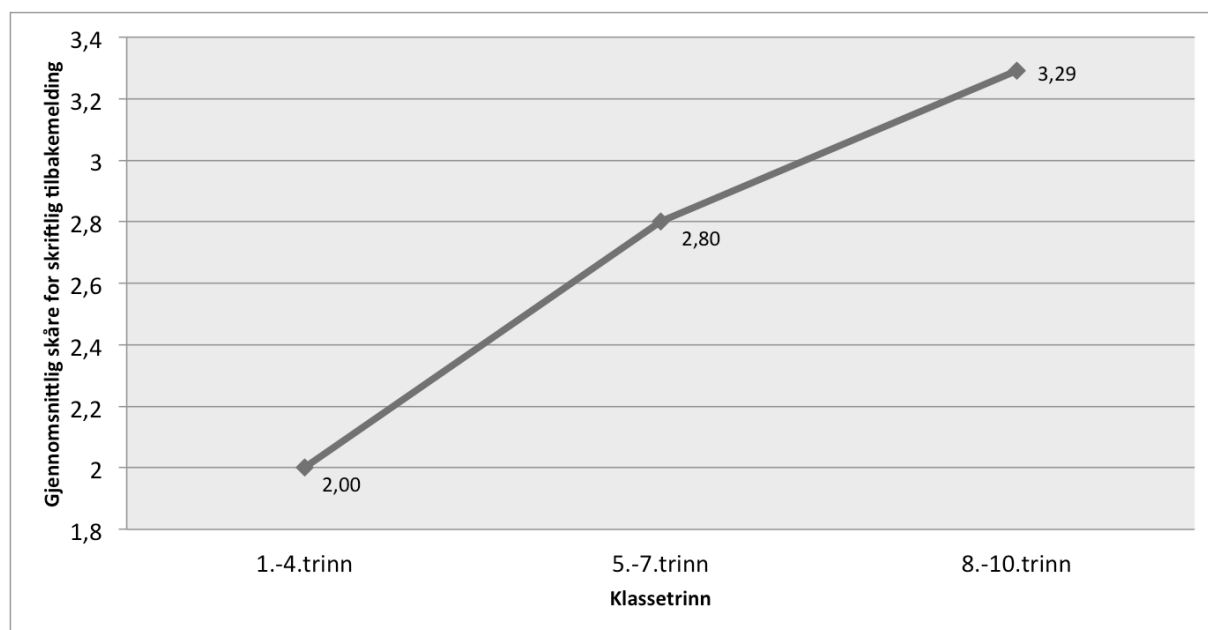
Figur 4.12: Gjennomsnittlig skåre for konstruktet *egen- og hverandrevurdering* i forhold til klassetrinn.

#### 4.4.4 Klassetrinn og skriftlig tilbakemelding

Figur 4.2 viste at lærere i stor grad brukte muntlig tilbakemelding fremfor skriftlig tilbakemelding i praktisk arbeid i naturfag. Det ble undersøkt om det er forskjeller mellom hvilket klassetrinn respondentene jobber på og i hvor stor grad de foretar muntlig og skriftlig tilbakemelding i praktisk naturfag. Variansanalyser viste ingen signifikante forskjeller på muntlige tilbakemeldinger, men skriftlige tilbakemelding ga derimot signifikante svar.

Det ble foretatt en variansanalyse for å sammenligne effekten av hvilket klassetrinn de jobbet med og gjennomsnittlig svar på hvor ofte de ga skriftlig tilbakemelding for vilkårene 1.-4. trinn, 5.-7. trinn og 8.-10. trinn. Undersøkte homogenitet i variansen ved *Levene's test* og fikk en p-verdi på 0,385. Det vil si at signifikansen er større enn grensen på 0,05 og nullhypotesen om at variansen er homogen i alle grupper beholdes. ANOVA-analysen viser at det var en signifikant effekt for hvilket klassetrinn de jobbet på og gjennomsnittlig skåre på hvor ofte de ga skriftlig tilbakemelding på  $p < 0,05$  for de tre betingelsene [ $F(2,58)=14,12$ ,  $p < 0,001$ ]. Post

hoc-sammenlikninger ved bruk av *Tukey HSD test* viser at gjennomsnittlig skåre for 1.-4. trinn ( $M=2,00$ ,  $SD=0,87$ ) både var signifikant forskjellig fra 5.-7. trinn ( $M=2,80$ ,  $SD=0,83$ ) og 8.-10. trinn ( $M=3,29$ ,  $SD=0,62$ ). Forskjellene var nesten signifikant for 5.-7.trinn og 8.-10.trinn ( $p=0,095$ ). Resultatene viser altså signifikant forskjell i hvilket klasstrinn respondentene har jobbet mest på og hvor ofte de gir skriftlig tilbakemelding. Resultatene er illustrert i figur 4.13. Grafen viser en tydelig øking i hvor ofte lærere gir skriftlig tilbakemelding i forhold til hvilket klasstrinn de jobber på.



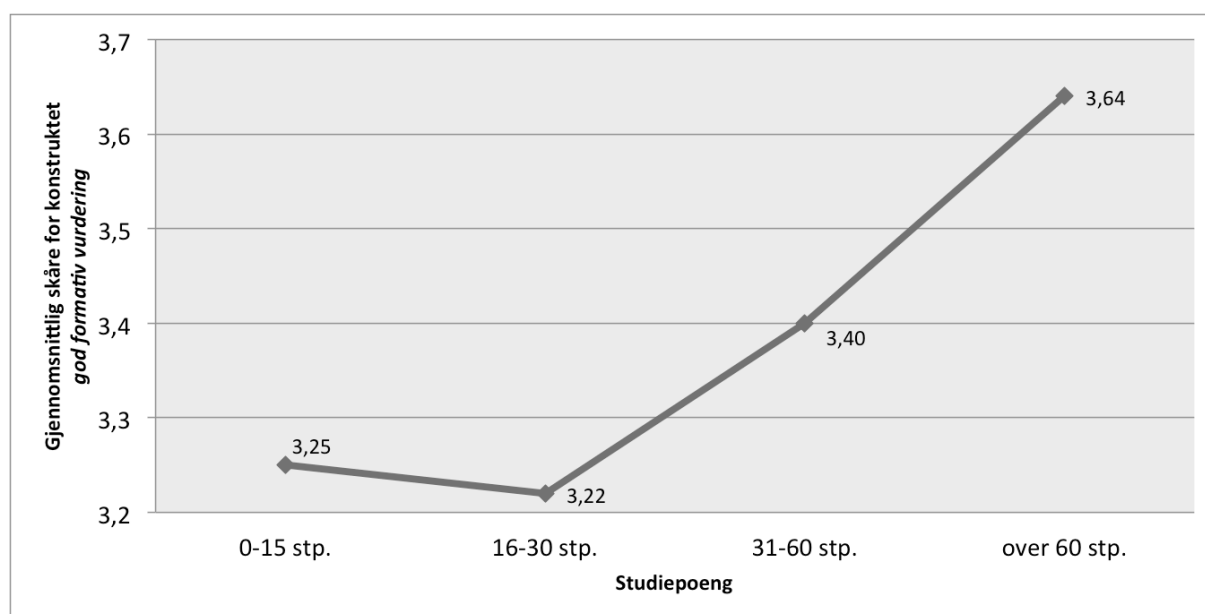
Figur 4.13: Gjennomsnittlig skåre for hvor ofte respondentene gir skriftlig tilbakemelding i forhold til klasstrinn.

## 4.5 Vurderingspraksis knyttet til antall studiepoeng lærerne har i naturfag

Det ble ikke gjort noen tydelige funn på om antall studiepoeng i naturfag påvirker i hvor stor grad lærerne foretar formativ vurdering i den praktiske delen av naturfaget. For gjennomsnittlig skåre konstruktene *god formativ vurdering* og *nivå av tilbakemelding* viser grafene at lærere med over 60 studiepoeng i naturfag scorer høyere. Likevel viser ikke svarene i ANOVA-analysene signifikante svar. Det var ingen tydelige trender for antall studiepoeng og om lærerne legger til rette for egen- og hverandrevurdering.

### 4.5.1 Studiepoeng og god formativ vurdering

Figur 4.14 illustrerer hvordan gjennomsnittlig skåre for konstruktet *god formativ vurdering* fordeler seg på antall studiepoeng i naturfag.



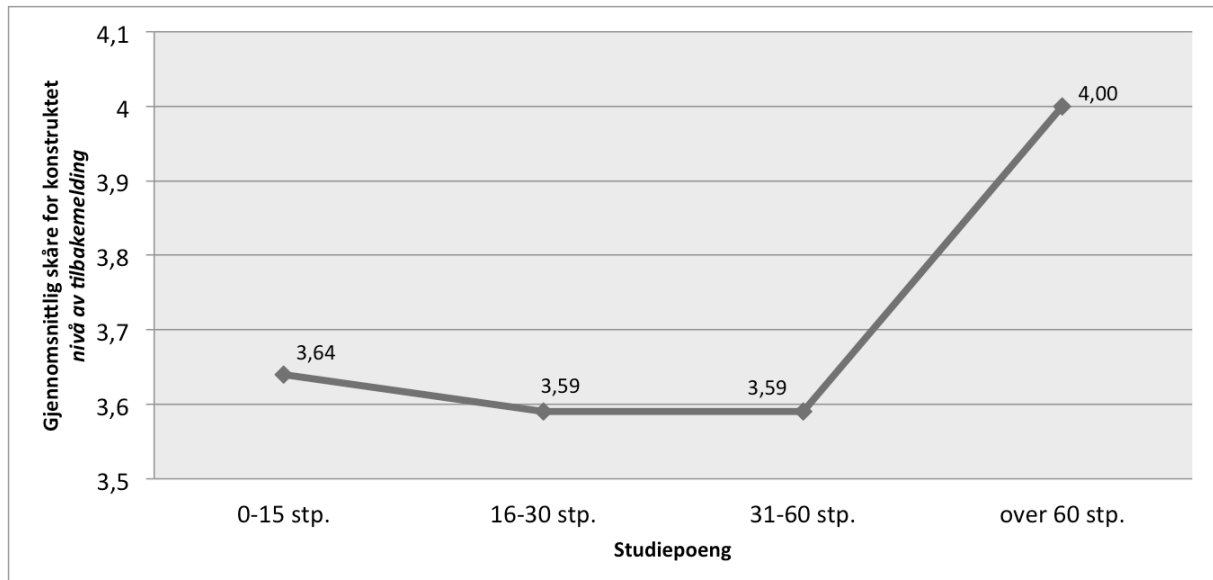
Figur 4.14: Gjennomsnittlig skåre for konstruktet *god formativ vurdering* i forhold til antall studiepoeng.

Figur 4.14 viser en antydning til at grafen stiger etter hvor mange studiepoeng lærerne har i naturfag fra 16-30 studiepoeng og videre. Likevel ga ikke ANOVA-analysene signifikante svar.

### 4.5.2 Studiepoeng og nivå av tilbakemelding

Figur 4.15 illustrerer hvordan gjennomsnittlig skåre for konstruktet *nivå av tilbakemelding* fordeler seg på antall studiepoeng i naturfag. Det er nesten ikke forskjeller i gjennomsnittlig

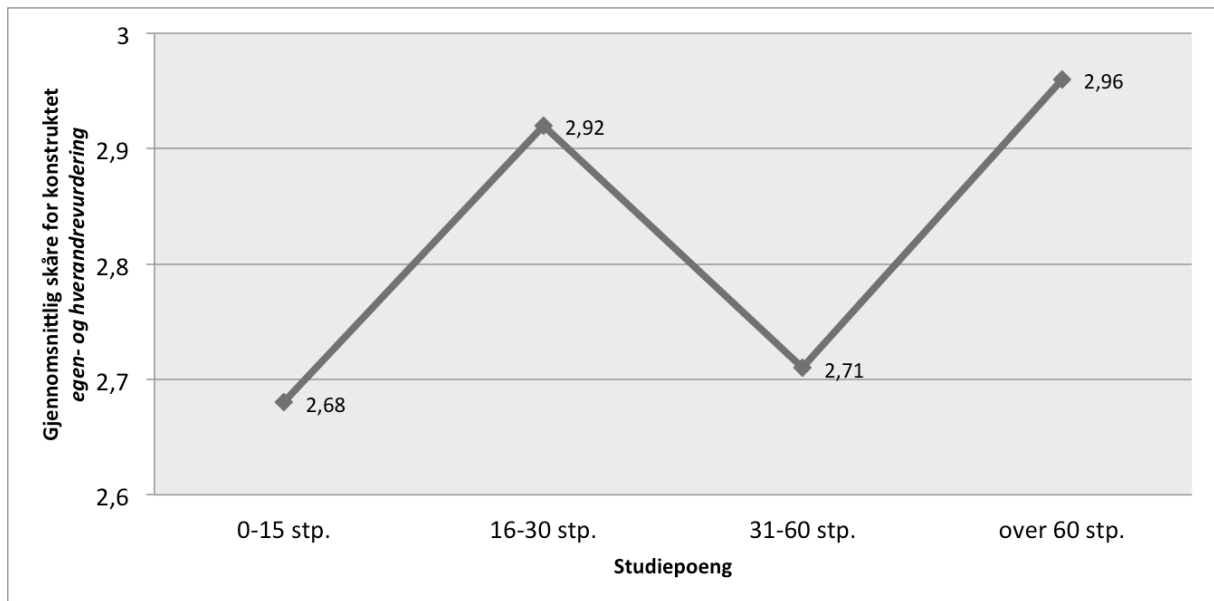
skåre for de tre første vilkårene 0-15 stp., 16-30 stp. og 31-60 stp. For de lærerne som har over 60 studiepoeng i naturfag er det en klar øking i gjennomsnittsskåre. ANOVA-analysene viste derimot ingen signifikant forskjell, noe som kan bety at resultatene er tilfeldige.



Figur 4.15: Gjennomsnittlig skåre for konstruert nivå av tilbakemelding i forhold til antall studiepoeng.

#### 4.5.3 Studiepoeng og tilrettelegging av egen- og hverandrevurdering

Det ble heller ikke funnet noen signifikante forskjeller på gjennomsnittlig skåre på konstruert egen- og hverandrevurdering og hvor mange studiepoeng lærerne har i naturfag. Figur 4.16 viser i tillegg ingen trender og respondentenes svar synes tilfeldige i forhold til antall studiepoeng. På bakgrunn av denne undersøkelsen kan man derfor ikke anta at antall studiepoeng i naturfag påvirker hvordan lærerne legger til rette for egen- og hverandrevurdering.



Figur 4.16: Gjennomsnittlig skåre for konstruktet *egen- og hverandrevurdering* i forhold til antall studiepoeng.

## 5 Diskusjon

### 5.1 Bruk av underveisvurdering

Resultatene viste at omtrent 75 % av lærerne svarte at de alltid eller nesten alltid vurderer elevene underveis i prosessen i praktisk arbeid i naturfag (figur 4.2). Svært få (mindre enn 5 %) av lærerne svarer at de aldri eller nesten aldri gir underveisvurdering. *Forskrift til opplæringslova* (Forskrift til opplæringslova, 2006) forteller at lærere er pålagt å vurdere elevene underveis. Derfor har de sannsynligvis god kjennskap til begrepet underveisvurdering. Den formative vurderingen skal støtte elevenes læring ved å si noe om hvordan de ligger an i læringsprosessen (Broadfoot et al., 2002; Hattie & Timperley, 2007; Millar, 2013; Wiliam, 2011). Dette kan bare gjøres ved å vurdere elevene underveis og underveisvurdering blir derfor en forutsetning for å praktisere formativ vurdering. Dersom vurderingen ikke gjøres underveis vil den bare kunne virke summativt fordi eleven ikke vil få mulighet til å jobbe videre med tilbakemeldingene og forbedre seg. Det er derfor positivt at 75 % av lærerne i spørreundersøkelsen svarer at de stort sett gir denne typen vurdering, noe som tyder på at mange lærere legger til rette for formativ vurdering i praktisk arbeid i naturfag.

### 5.2 Bruk av vurdering for å tilpasse undervisningen

Litt under 60 % av lærerne svarer at de alltid eller nesten alltid bruker vurdering for å tilpasse undervisningen i praktisk arbeid i naturfag (figur 4.2). Nesten 30 % svarer iblant på dette spørsmålet. For at vurderingen skal kalles formativ må den bli brukt til å tilpasse undervisningen ut fra elevenes behov (Black & Wiliam, 1998). Basert på lærernes svar på dette spørsmålet kan det tenkes at mange lærere ikke alltid tilpasser undervisningen i praktisk arbeid, noe som hindrer vurderingen å fungere formativt. Det er ikke spurt om hvilken type praktisk arbeid lærerne i spørreundersøkelsen benytter seg av i naturfag. Undervisningsmetode kan ha betydning for hvorvidt det lar seg gjøre å tilpasse undervisningen. I *inquiry-based teaching* har vurderingen en naturlig plass og lærerens veiledende rolle i denne type undervisning gir gode muligheter for formativ vurdering (Knain & Kolstø, 2011). Vurderingens naturlige plass i denne type undervisning kommer også tydelig frem i 5E-modellen. I 5E-modellen for utforskende arbeidsmåter inngår



komponentene engasjere, utforske, forklare og utvide, og vurderingen inngår i hele forløpet (Bybee, Taylor, Gardner, Westbrook & Landes, 2006). Derfor kan det tenkes at inquiry-based undervisning og bruk av 5E-modellen legger godt til rette for formativ vurdering i praktisk arbeid i naturfag. Vurdering trenger ikke være et like naturlig element i kokebokforsøk, hvor undervisningen gjerne baserer seg på ferdige opplegg som elevene skal gjennomføre. Derfor kan det kanskje være vanskeligere å tilpasse undervisningen i slike forsøk enn i utforskende arbeidsmåter. Litt færre enn halvparten av lærerne i spørreundersøkelsen bruker stort sett ikke vurdering for å tilpasse undervisningen i praktisk arbeid i naturfag. En antakelse kan derfor være at det ofte brukes undervisningsmetoder i praktisk arbeid som ikke i like stor grad legger til rette for formativ vurdering, som for eksempel kokebokforsøk.

### **5.3 Hvordan lærerne bruker formativ vurdering i praktisk arbeid i naturfag**

Som det ble utdypet i teorikapitlet bør den formative vurderingen besvare følgende tre spørsmål; ”Hvor skal eleven?”; ”Hvor er eleven nå i forhold til målene?” og ”Hvordan jobbe videre for å oppnå ønsket mål?” (Broadfoot et al., 2002; Hattie & Timperley, 2007; Wiliam, 2011). Jeg ønsker i dette kapitlet å legge frem hvordan lærerne i spørreundersøkelsen har besvart spørsmålene knyttet til dette, for å kunne si noe om deres formative vurderingspraksis. I tillegg vil jeg se på hvordan lærerne besvarer spørsmålene knyttet til de fire ulike nivåene av tilbakemelding, som kan si noe om hvilken type tilbakemelding elevene ofte får. Disse fire nivåene ble omtalt i teorikapitlet som oppgavenivå, prosessnivå, selvreguleringsnivå og personlig nivå (Hattie & Timperley, 2007).

#### **5.3.1 ”Hvor skal eleven?”**

En uklarhet rundt hva som skal vurderes i praktisk arbeid kan gjøre det vanskelig for lærere å foreta gode formative vurderinger. Praktisk arbeid skal blant annet lære elevene å knytte teori til praksis, gi interesse og motivasjon for faget, lære dem ferdigheter på lab, støtte læring i naturfag, utvikle kunnskap og forståelse av vitenskapen og forstå vitenskapelige prosesser (Abrahams et al., 2013; Hofstein & Lunetta, 2004; Millar, 2011; Ottander & Grelsson, 2006). Det er derfor helt avgjørende at målene for hva elevene skal lære kommer tydelig frem i det praktiske arbeidet. Det er tydelig at det å gi læringsmål har stort fokus blant naturfaglærere. Omtrent 90 % svarer at de alltid eller nesten alltid gir elevene læringsmål i praktisk arbeid i naturfag (figur 4.3). Å gi elevene god informasjon om hva som er målet med læreprosessen er

avgjørende for den formative vurderingen (Broadfoot et al., 2002; Hattie & Timperley, 2007; Wiliam, 2011). Resultatene tyder derfor på at lærerne i stor grad oppfyller dette kravet i formativ vurderingspraksis. Dette kan være avgjørende i praktisk arbeid, da det har vist seg å være vanskelig å påvise læringseffekt av praktisk arbeid (Hodson, 1990; Hofstein & Lunetta, 1982; Kind, 1999). På grunn av dette er det kanskje spesielt viktig å bruke læringsmål i praktisk arbeid. Tydelige mål gir et godt grunnlag for at det praktiske arbeidet hele tiden har et fokus på læring og ikke bare aktivitet. At de fleste av lærerne i undersøkelsen har fokus på læringsmål er et godt utgangspunkt for å praktisere formativ vurdering.

I spørreundersøkelsen ble det stilt noen utdypende spørsmål rundt lærernes bruk av læringsmål i praktisk arbeid (figur 4.3). Under halvparten av lærerne svarer at de alltid eller nesten alltid gir elevene kjennetegn på måloppnåelse og eksempler på måloppnåelse i praktisk arbeid, og nesten ingen svarer at de alltid eller nesten alltid lar elevene bidra med å bestemme læringsmål. Hattie og Timperley (2007) skriver i sin utvidede forklaring av modellen i figur 2.1 at lærerne bør løfte frem suksesskriteriene for læringsmålene samt vise eksempler på høy og lav måloppnåelse. I tillegg argumenterer de for at elevene bør være med å bestemme læringsmål. Dette løftes også fram hos Utdanningsdirektoratets hjemmesider for vurderingspraksis (Utdanningsdirektoratet, 2016). Resultatene i spørreundersøkelsen viser at færre lærerne har fokus på å tydeliggjøre læringsmålene ytterligere ved å fokusere på kjennetegn på måloppnåelse og å vise eksempler. Dette kan bety at lærerne ikke i tilstrekkelig grad oppfyller kravet om at den formative vurderingen skal være tydelig på hva eleven skal lære. Det kan være vanskelig å vise eksempler på tidligere arbeid i praktisk arbeid i naturfag dersom vurderingen av elevene gjøres muntlig, men det er helt klart muligheter for å samtale med elevene om hva som skal til for å få en god måloppnåelse. Å inkludere elevene i å lage læringsmål vil kanskje også bidra til å øke motivasjonen for det praktiske arbeidet ytterligere, noe mange lærere anser som et viktig mål med praktisk arbeid (Kjærnsli et al., 2007). Det hevdes i tillegg at elevene forstår tydeligere hva de faktisk må jobbe med når de involveres i å lage kriterier (Bruce, 2001). Dette kan føre til at elevene lettere kan jobbe mot målene også i den praktiske delen av faget.

En svensk case-studie så at lærere og elever ofte hadde ulik oppfatning av hva som faktisk ble vurdert i praktisk arbeid (Ottander & Grelsson, 2006). De forklarte dette blant annet med at fokuset var på hva som skulle gjøres, ikke hva som skulle læres. I og med at lærerne i spørreundersøkelsen har et stort fokus på å gi elevene læringsmål også i praktisk arbeid kan

det tyde på at de ønsker å ha et læringsfokus, ikke bare aktivitetsfokus, i praktisk arbeid. Men dersom målene ikke er tydelige nok kan lærere og elever likevel få en ulik forståelse av målene for undervisningen, noe som vil få konsekvenser for om elevene lærer det som er intensjonen i aktiviteten.

### **5.3.2 "Hvor er eleven i forhold til målene?"**

En viktig forutsetning for formativ vurdering er at elevene får tilstrekkelig med informasjon om hvor de er i læringsprosessen (Broadfoot et al., 2002; Hattie & Timperley, 2007; Wiliam, 2011). Nesten 70 % av lærerne i undersøkelsen svarer at de alltid eller nesten alltid gir tilbakemelding til elevene på hva de har lært (figur 4.4). Dette er positivt med tanke på den formative vurderingen. Likevel viser resultatene at kun omtrent 40 % av lærerne alltid eller nesten alltid gir tilbakemeldinger på hva eleven *ikke* kan eller har forstått i prosessen. For at elevene skal få en helhetlig vurdering av hvor de er i læringsprosessen kan det antas at elevene må få begge typer tilbakemelding. Resultatene kan derfor antyde at dette ikke alltid er tilfellet i praktisk arbeid i naturfag. Det kommer ikke frem i undersøkelsen hvorfor mange lærere unnlater å gi tilbakemeldinger på hva eleven ikke har forstått. Som lagt frem i teorikapitlet anser mange lærere praktisk arbeid i naturfag som en arena hvor elevene kan utvikle en positiv holdning til naturfag (Kjærnsli et al., 2007; Millar, 2011). En mulig årsak til at lærerne i mindre grad vektlegger tilbakemeldinger på det eleven ikke har forstått kan derfor være at de ønsker å holde et positivt blikk i praktisk arbeid i naturfag for å styrke elevenes motivasjon i faget. Som resultatene til FIVIS (*Forskning på individuell vurdering i skolen*) viste ønsker elever gode og tydelige tilbakemeldinger både på hva de er gode til og hva som kan bli bedre (Sandvik & Buland, 2014). Også lærere i KMOFAP-prosjektet kom frem til at det var viktig at tilbakemeldingen inneholdt god informasjon om både hva som var gjort bra og hva som måtte forbedres (Black et al., 2003). Dersom lærere i stor grad bare fokuserer på positive tilbakemeldinger i praktisk arbeid i naturfag kan dette være med på å forhindre at elevene får tilstrekkelig informasjon om hvor de er i læringen. Dette kan føre til at det blir vanskeligere for elevene å oppnå ønsket læring i undervisningen og at den formative vurderingen ikke blir tilstrekkelig.

### **5.3.3 "Hvordan jobbe videre for å oppnå ønsket mål?"**

I formativ vurdering bør tilbakemeldingene gi god veiledning på hvordan elevene kan forbedre seg (Black et al., 2003; Broadfoot et al., 2002; Forskrift til opplæringslova, 2006).

Resultatene i FIVIS viste også at elevene ønsker å få konkrete tilbakemelding på hvordan de kan bli bedre (Sandvik & Buland, 2014). I spørreundersøkelsen svarte omtrent 60 % av lærerne at de alltid eller nesten alltid gir konkrete tilbakemeldinger på hva elevene må gjøre for å forbedre seg (figur 4.4). Det betyr også at omtrent 40 % av lærerne i undersøkelsen ikke jevnlig gir denne typen tilbakemelding i praktisk arbeid i naturfag. En mulig konsekvens dersom elevene ikke får konkrete tilbakemeldinger om hvordan de kan forbedre seg, kan være at de ikke vet hvordan de skal lære det som forventes at de skal lære. I praktisk arbeid kan dette føre til at eleven ikke vil klare å gjennomføre den praktiske aktiviteten på en måte som bidrar til læring for eleven. Det er vanskelig å si noe om hvorfor mange lærere ikke fokuserer på konkrete tilbakemeldinger i praktisk arbeid. En mulig årsak kan være at lærere har størst fokus på aktivitet og mindre fokus på læring i denne delen av faget, noe som også viste seg i den svenske case-studien som ble lagt frem i teorikapittelet (Ottander & Grelsson, 2006). En annen mulig årsak kan være at lærere synes det er vanskelig å gi gode tilbakemeldinger til elevene i praktisk arbeid på grunn av rammefaktorene. Kanskje har lærerne nok med å holde oversikt over at elevene faktisk gjør det som er planlagt og ha fokus på HMS på laben.

### **5.3.4 Bruk av de ulike tilbakemeldingsnivåene**

Vurdering som kombinerer tilbakemeldinger på selve oppgaven, prosessen og elevens mulighet for selvregulering er mest ønskelig i effektiv tilbakemeldingspraksis (Hattie & Timperley, 2007). Resultatene i undersøkelsen viste at lærerne ga ganske like svar for i hvor stor grad de ga tilbakemelding på oppgaven, prosessen og selvreguleringsnivå (rundt 70 % svarer alltid eller nesten alltid for alle tre spørsmålene). Se figur 4.5. At svarene er ganske like kan tyde på at elevene får tilbakemelding i en kombinasjon av disse tre nivåene. Tilbakemeldinger på både oppgaven og prosessen i praktisk arbeid i naturfag kan være med å styrke elevenes forståelse av naturvitenskapen både som produkt og prosess, noe som også læreplan i naturfag ønsker å oppnå ved hjelp av forskerspiren. Hattie og Timperley (2007) peker på mange ulike aspekter ved tilbakemelding på selvreguleringsnivå som gir gode gevinster for effektiv tilbakemelding. De skriver at dersom lærere gir tilbakemelding til elevene som bygger på selvregulering kan det føre til at elevene blant annet øker evnen til egenvurdering, øker innsatsen for å gjøre noe med tilbakemeldingen, øker tilliten til at tilbakemeldingen er korrekt og øker attribusjon i faget. I tillegg påpeker de at denne typen tilbakemelding gjør at eleven blir bedre på å selv søke hjelp når det er behov (Hattie & Timperley, 2007). I praktisk arbeid i naturfag vil dette være til stor hjelp i den formative

vurderingen, da dette øker elevenes selvstendighet i vurderingsarbeidet og læreren er frigjort til å hjelpe og vurdere andre elever, eller fokusere på aktiviteten som skal gjøres. At mange lærere gir tilbakemelding på alle tre nivåene kan derfor være med å bidra til at praktisk arbeid fører til økt læring i naturfag blant elevene.

Tilbakemelding på personlig nivå har vist seg å være lite effektivt for elevenes læring (Hattie & Timperley, 2007). Omtrent halvparten av lærerne i undersøkelsen svarer at de alltid eller nesten alltid gir ros til elevene. Omtrent 30 % svarer iblant. 85 % svarer at de aldri eller nesten aldri gir kritikk. Resultatene bekrefter tidligere studier som viser at lærere generelt gir mye ros til elevene (Thronsen et al., 2009). Selv om Hattie og Timperley (2007) påpeker at denne typen tilbakemelding ikke har stor læringseffekt kan det kanskje virke motiverende for elevene å få ros. Praktisk arbeid kan være med å fremme interessen for og verdsettingen av naturfag (Kjærnsli et al., 2007). På grunn av dette har kanskje lærere et større behov for å gi elevene gode opplevelser i praktisk arbeid, noe som positive tilbakemeldinger i form av ros kan gi. Dette ble også diskutert i kapittel 5.3.2. Likevel poengterer Hattie og Timperley (2007) at tilbakemeldinger på personlig nivå er svært avhengig av hvordan eleven oppfatter denne. For eksempel vil ros til elever med lave forventninger til egen prestasjon vil virke negativt på elevene fordi de ikke tror på den. I tillegg vil negative tilbakemeldinger til elever som har høye forventninger til egen prestasjon også virke negativt for dem og ikke være læringsfremmede (Hattie & Timperley, 2007). Dette er viktig at lærerne tar hensyn til, også i den praktiske delen av faget.

#### **5.4 Bruk av egen- og hverandrevurdering**

Studier har vist at elevenes egenvurdering og vurdering av medelever generelt er lite utbredt blant norske lærere (Sandvik & Buland, 2014). Resultatene i denne undersøkelsen viste at under 20 % av lærerne svarer at de alltid eller nesten alltid legger til rette for egenvurdering i praktisk arbeid (figur 4.6). 5 % svarer at de alltid eller nesten alltid legger til rette for hverandrevurdering. Det blir altså ikke lagt stor vekt på egen- og hverandrevurdering i praktisk arbeid i naturfag blant lærerne i undersøkelsen. Elevene har en viktig rolle i formativ vurdering, noe også lærerne i KMOFAP-prosjektet fant ut. De så blant annet at elevene la mer innsats i arbeidet dersom de skulle bli vurdert av medelever og samtidig fikk de også tilbakemelding på et språk de selv forsto godt, noe som bidro til at de tydeligere forsto hva

som var bra og mindre bra i eget arbeid. (Black et al., 2003). Det kan være flere grunner til at lærerne har lite fokus på egen- og hverandrevurdering i praktisk arbeid i naturfag. En mulig forklaring kan være tiden. Praktisk arbeid har mange mål, deriblant å fange elevenes motivasjon og interesse for faget, gi de en større forståelse av vitenskapen og lære dem praktiske ferdigheter i naturfag (Abrahams et al., 2013; Hofstein & Lunetta, 2004; Kjærnsli et al., 2007; Millar, 2011). At det er mange ulike mål i praktisk arbeid kan påvirke hvordan det vurderes. Dårlig tid kan bety at læreren ikke setter av nok tid i praktisk arbeid til formativ vurdering. Kanskje går dette spesielt ut over egen- og hverandrevurdering, som vil kreve mer tid enn tilbakemeldinger fra lærer. En annen grunn kan være at noen lærere ikke tør å la elevene ta stor del i vurderingsarbeidet i praktisk arbeid. Dette kan være fordi de er redde for å miste kontrollen i undervisningen, for eksempel på grunn av usikkerhet med å lede ved praktisk arbeid, eller usikkerhet med å undervise på laben.

## **5.5 Bruk av muntlig og skriftlig tilbakemelding**

Resultatene i undersøkelsen viste at lærerne vektlegger muntlig tilbakemelding i større grad enn skriftlig tilbakemelding i praktisk arbeid i naturfag. 85 % svarer at de alltid eller nesten alltid gir muntlige tilbakemeldinger til elevene, mens bare 15 % svarer at de alltid eller nesten alltid gir skriftlig tilbakemelding til elevene (figur 4.2). 30 % svarer i tillegg at de aldri eller nesten aldri gir skriftlig tilbakemelding. At lærere i stor grad vektlegger muntlig vurdering fremfor skriftlig vurdering kan være med å legge til rette for god formativ vurdering i praktisk arbeid i naturfag. For det første gir muntlig tilbakemelding mulighet for dialog. Å ha en god diskusjon med elevene i forbindelse med praktisk arbeid er viktig (Osborne & Dillon, 2010; Ottander & Grelsson, 2006). I forbindelse med slike dialoger i praktisk arbeid kan lærer vurdere og gi tilbakemeldinger til elevene som sier noe om hva de har forstått knyttet til læringsmålene. I tillegg kan lærer gi tilbakemeldinger som fører læringen deres fremover. Muntlig tilbakemelding legger også til rette for at lærerne kan vurdere elevene underveis, noe som underbygger resultatene om underveisvurdering som diskutert i kapittel 5.1. Dette gir et godt grunnlag for at klasseroms diskusjonene kan fungere formativt i praktisk arbeid. Gjennom muntlige vurderinger er det også muligheter for å tilpasse undervisningen etter elevenes behov, noe som er en forutsetning for formativ vurdering (Black & Wiliam, 1998). Skriftlige vurderinger i praktisk arbeid vil ofte ikke kunne finne sted før aktiviteten er ferdig og vil derfor i mindre grad kunne fungere formativt ved at undervisningen tilpasses elevenes

behov. Derfor vil det ofte være både naturlig og mest hensiktsmessig å vurdere praktiske aktiviteter muntlig fremfor skriftlig.

Muntlige vurderinger i praktisk arbeid vil i større grad legge til rette for “Direct Assessment of Practical Skills” (DAPS). Denne type vurdering sier noe om elevens faktiske ferdigheter i det praktiske arbeidet (Reiss et al., 2012). DAPS er særlig ønskelig for å kunne vurdere elevens praktiske ferdigheter i naturfag. For at en direkte vurdering av elevens praktiske ferdigheter skal finne sted er det avgjørende at elevene blir vurdert muntlig. Resultatene kan derfor peke på at det blir tilrettelagt for denne type vurdering hos naturfaglærere. Skriftlige vurderingsformer vil i større grad legge til rette for “Indirect Assessment of Practical Skills” (IAPS) som gir et innblikk i hvilken kompetanse eleven har rundt det praktiske arbeidet (Reiss et al., 2012). Denne måten å vurdere på kan være fordelaktig for å si noe om elevens kunnskaper og forståelse rundt for eksempel naturvitenskapelige tenkemåter. Abrahams et al. (2013) skriver at en bør ha størst fokus på vurdering ved DAPS i praktisk arbeid i naturfag. Likevel har vi i den norske læreplan mange mål som går på kunnskaper om naturfag og ferdigheter rundt det å skrive rapport (Kunnskapsdepartementet, 2013). Det er derfor viktig at også denne typen vurdering blir tatt med i praktisk arbeid i naturfag. At flere lærere svarer at de aldri eller nesten aldri vurderer elevene skriftlig kan bety at ikke alle læreplanmålene blir oppfylt. Det kan tenkes at en kombinasjon av muntlig og skriftlig vurdering i praktisk arbeid er fordelaktig for å dekke de norske læreplanmålene.

Resultatene viste en signifikant forskjell i hvor stor grad lærere på ulike klassetrinn vurderte elevene skriftlig i praktisk arbeid i naturfag og bruken av skriftlige tilbakemeldinger var økende ved høyere klassetrinn. Det er naturlig å tenke at lærere på ungdomstrinnet har et større fokus på skriftlig vurdering generelt på grunn av krav om karaktersetning. Egne erfaringer har også vist at mange lærere på ungdomstrinnet føler et press på å vurdere elevene skriftlig for å kunne gi en god begrunnelse for karakteren. Samtidig er det også logisk å tenke at elever på de lavere trinnene får mindre skriftlig vurdering på grunn av modenhet, leseferdigheter, erfaringer med vurdering og så videre.

## 5.6 Utforskende arbeidsmåter som en arena for formativ vurdering

I forskningsspørsmålet mitt har jeg stilt spørsmål knyttet til praktisk arbeid i naturfagundervisningen. Praktisk arbeid kan gjennomføres på ulike måter og ulike arenaer. For eksempel kan det bety forsøk på lab, klasseromsforsøk eller ekskursjoner i uteområdet. Det kan legges opp til såkalte ”kokebokforsøk”, eller være mer åpne forsøk, som for eksempel utforskende arbeidsmåter. Selv om jeg ikke har tatt opp temaet utforskende arbeidsmåter i forskningsspørsmålet mitt synes jeg likevel at det er interessant å se på i forbindelse med formativ vurdering. Jeg ønsker derfor i dette delkapitlet å vurdere om lærernes svar kan si noe om det legges til rette for denne typen undervisning og trekke frem fordelene undervisningsmetoden kan gi for formativ vurdering i naturfag.

Knain og Kolstø (2011) skriver at utforskende arbeidsmåter gir rom for både underveis- og sluttvurdering, samt vurdering for og av læring. Dette gir læreren et godt utgangspunkt for å benytte mange ulike vurderingsverktøy. Resultatene i undersøkelsen kan tyde på at lærere gir elevene vurdering underveis i praktisk arbeid og at de i stor grad har fokus på vurdering som er med på å støtte læringen til elevene i form av læringsmål og konkrete tilbakemeldinger. I tillegg vektlegger de fleste lærerne muntlig tilbakemelding, noe som kan være med på å fremme formativ vurdering i praktisk arbeid. Dette er også elementer som vektlegges i utforskende arbeidsmåter. I utforskende arbeidsmåter arbeider elevene med å finne svar på konkrete spørsmål (Knain & Kolstø, 2011). Det betyr at elevene må vite hva som er målet med undervisningen, noe som er viktig i formativ vurdering. At lærerne i undersøkelsen hadde fokus på dette legger til rette for denne delen av utforskende arbeidsmåter. I tillegg arbeider elevene mye selvstendig i utforskende arbeidsmåter og må selv gjøre vurderinger underveis i prosessen (Knain & Kolstø, 2011). Dette legger i stor grad til rette for elevmedvirkning i vurderingsarbeidet, som blir ansett som viktig i formativ vurdering (Black et al., 2003; Wiliam, 2011). Da elevene ofte arbeider sammen i grupper vil det også legges til rette for at elevene vurderer hverandre i større grad. Resultatene viste at få lærere ofte la til rette for egen- og hverandrevurdering i praktisk arbeid i naturfag. Det kan bety at ikke så mange lærere bruker utforskende arbeidsmåter i undervisningen, eller at dette elementet mangler.

Det som skiller utforskende arbeidsmåter fra andre type praktisk arbeid er at elevene i stor grad gjør og tenker på samme måte som i naturvitenskapen (Council, 2000). Dette er også en



viktig del av fagområdet forskerspiren (Kunnskapsdepartementet, 2013). Å arbeide på denne måten inkluderer elevene på en annen måte enn såkalte ”kokebokforsøk” og utfordrer elevene på nye måter. Lærerens rolle i utforskende arbeidsmåter vil i tillegg bidra til gode muligheter for formativ vurdering. I og med at elevene får en fremtredende rolle i gjennomføringen av undervisningen i utforskende arbeidsmåter, vil læreren innta en veilederrolle i undervisningen (Knain & Kolstø, 2011). Det betyr at læreren får frigjort mye tid til å gi elevene nyttige tilbakemeldinger underveis i prosessen. At mange lærerne i undersøkelsen ga formative tilbakemeldinger til elevene i form av konkrete tilbakemeldinger på hvor de er i prosessen og hvordan de bør arbeide videre, åpner mulighetene for utforskende arbeidsmåter i undervisningen.

## 5.7 Betydning av år i jobb, klassetrinn og antall studiepoeng

Som en del av forskningsspørsmålet ønsket jeg å undersøke om antall år i jobb som naturfaglærer, hvilket klassetrinn det undervises på og antall studiepoeng læreren har i naturfag har betydning for deres vurderingspraksis. Resultatene vil diskuteres i dette kapitlet.

### 5.7.1 Vurderingspraksis knyttet til antall år i jobb som naturfaglærer

Jeg har ikke funnet tidligere forskning på sammenhengen mellom lærerens undervisningserfaring og vurderingspraksis. Resultatene viste heller ingen signifikante forskjeller på hvor lenge respondentene hadde jobbet som naturfaglærer og gjennomsnittlig skåre for konstruktet *god formativ vurdering*. Dette kan tyde på at hvor lenge man har jobbet som naturfaglærer ikke er en avgjørende faktor for i hvor stor grad man praktiserer formativ vurdering til elevene i praktisk arbeid. Det var heller ingen signifikante forskjeller på hvor lenge de har jobbet som naturfaglærer og gjennomsnittlig skåre på konstruktet *nivå av tilbakemelding*. Også her kan resultatene tyde på at lærernes erfaring ikke er avgjørende for vurderingspraksis. Grafene i figur 4.7 og figur 4.8 viste likevel en økende trend i gjennomsnittlig svar på konstruktene *god formativ vurdering* og *nivå av tilbakemelding* etter hvor lenge de har jobbet som naturfaglærer. For konstruktet *egen- og hverandrevurdering* viste derimot resultatene i undersøkelsen at lærere som har jobbet over 10 år skårer signifikant høyere enn lærerne med kortere erfaring, noe som tyder på at hvor lenge de har jobbet som naturfaglærer har noe å si for i hvor stor grad de legger til rette for egen- og hverandrevurdering. Grafen i figur 4.9 var også økende fra lærere med kort erfaring til lærere med lengre erfaring.

Alle grafene for konstruktene *god formativ vurdering* (figur 4.7), *nivå av tilbakemelding* (figur 4.8) og *egen- og hverandrevurdering* (figur 4.9) viser en økende gjennomsnittlig skåre etter hvor lenge lærerne har jobbet som naturfaglærer. Selv om ikke alle ANOVA-analysene viste signifikante forskjeller er det interessant at samtlige resultater viser at lærere med lengre erfaring skårer bedre enn lærere med færre antall år i jobb. Dette kan tyde på lærernes erfaring likevel kan ha betydning for deres vurderingspraksis. Som nevnt i teorikapitlet viser tidligere studier at lærerens erfaring har betydning for deres undervisningspraksis (Berliner, 2001; Blömeke et al., 2016; Clotfelter et al., 2007). Dersom en forutsetter at vurderingspraksis

knyttet opp mot undervisningspraksis kan en konklusjon om at lærerens erfaring har betydning for den formative vurderingen antas å være sannsynlig.

Det kan være flere årsaker til at antall år i jobb er av betydning for lærernes vurderingspraksis. For det første kan det tenkes at lærere med lengre erfaring har en større trygghet i å undervise i praktisk arbeid, noe som kan føre til at de i større grad kan bruke tid på vurdering underveis. Med tanke på alle de ulike hensiktene og målene med praktisk arbeid i naturfag vil det være naturlig å tenke at erfaring har noe å si for at læreren skal klare å omfavne flere av de ulike målene med praktisk arbeid etter hvert som læreren opparbeider seg praktisk undervisningserfaring. Det kan da også tenkes at lærerne i større grad tør å slippe til elevene i vurderingsarbeidet. Lærere som har jobbet lenge med naturfag har kanskje også gjort de samme forsøkene mange ganger og på den måten føler seg tryggere i å la elevene medvirke i vurderingsarbeidet i større grad. Av den grunn er det også naturlig å tenke at antall år i jobb som naturfaglærer er av betydning for hvorvidt læreren legger til rette for egen- og hverandrevurdering i praktisk arbeid i naturfag.

### **5.7.2 Vurderingspraksis knyttet til hvilket klassetrinn lærerne underviser på**

I denne undersøkelsen var det ikke store forskjeller på lærernes vurderingspraksis på mellomtrinnet og ungdomstrinnet, men for konstruktene *god formativ vurdering* og *egen- og hverandrevurdering* skåret lærere på småtrinnet signifikant lavere enn lærere på mellomtrinnet og ungdomstrinnet. Grafen i figur 4.11 viste også at lærere på småtrinnet skåret litt lavere enn lærere på høyere trinn for konstruktet *nivå av tilbakemelding*, selv om resultatene ikke var signifikante. Funnene kan sies å støtte rapporten fra FIVIS som fant ut at lærere på barnetrinnet skåret signifikant lavere på vurderingspraksis enn lærere på ungdomsskoler og videregående skoler (Sandvik & Buland, 2014). Vi må da ta hensyn til at lærere på mellomtrinnet i min undersøkelse havner i begge grupper. Grad av modenhet hos elevene kan være en viktig årsak til hvorfor lærere på de minste trinnene ikke har like stort fokus på læringsmål, konkrete tilbakemeldinger på hvor de er i læringsprosessen og hvordan de skal arbeide videre, og egen- og hverandrevurdering i praktisk arbeid. Det kan også tenkes at lærere ved de minste trinnene har hovedfokus på at det praktiske arbeidet skal være motivasjon og interesse for faget og at de derfor har mindre fokus på vurdering i praktisk arbeid. Ut fra resultatene er det derfor en antagelse at det er mindre fokus på vurdering på

lavere trinn og at hvilket klassetrinn lærerne underviser på er av betydning for lærernes vurderingspraksis i praktisk arbeid i naturfag.

### **5.7.3 Vurderingspraksis knyttet til antall studiepoeng i naturfag blant lærere**

Rapporten fra FIVIS viste at lærernes faglige og fagdidaktiske kompetanse har betydning for hvordan læreren arbeider med vurdering og at lærere med fagutdanning i tillegg til praktisk-pedagogisk utdanning skåret signifikant bedre enn allmennlærere på vurderingspraksis (Sandvik & Buland, 2014). Ut fra disse resultatene kan det være naturlig å tenke at lærere med flere studiepoeng i naturfag skårer bedre på formativ vurdering i spørreundersøkelsen. Også resultatene fra TIMMS viser at lærenes undervisningspraksis i stor grad påvirkes av lærernes fagdidaktiske ferdigheter (Bergem et al., 2016). Likevel viser resultatene fra undersøkelsen ingen signifikante forskjeller på antall studiepoeng i naturfag for konstruktene *god formativ vurdering*, *nivå av tilbakemelding* og *egen- og hverandrevurdering*. Ut fra grafene i figur 4.14, 4.15 og 4.16 ser vi likevel en antydning til at lærere med over 60 studiepoeng i naturfag skårer høyere for disse konstruktene. Lærere med fagutdanning i tillegg til PPU vil ha minst 60 studiepoeng i sine fag, men ofte også langt over 60 studiepoeng som var øvre grense i denne spørreundersøkelsen. Resultatene i denne spørreundersøkelsen støtter delvis resultatene fra FIVIS og det kan tenkes at lærere må ha en betydelig andel studiepoeng i faget for at de skal være bedre på vurdering. Gjennomsnittlig skåre fordelt på studiepoeng vil av den grunn utgjør små forskjeller i min spørreundersøkelse da verdiene er mellom 0 og 60 studiepoeng. Det ser ikke ut som at vurderingspraksisen påvirkes i særlig grad når læreren har under 60 studiepoeng, men kanskje ville resultatene vært signifikante dersom undersøkelsen også spurte etter lærere med flere studiepoeng i naturfag.

## **5.8 Svakheter med studien**

Avslutningsvis i diskusjonen ønsker jeg å drøfte noen av svakhetene med min studie. Validitet og reliabilitet ble diskutert i metodekapittelet, men jeg vil her diskutere noen konkrete utfordringer med datamaterialet mitt som kan ha betydning for hvorvidt jeg har klart å svare på forskningsspørsmålet mitt.

Da jeg begynte å analysere spørreundersøkelsen opp mot forskningsspørsmålet mitt oppdaget jeg fort at det kunne bli vanskelig å konkret knytte spørsmålene opp mot det praktiske

arbeidet i naturfag, noe som var fokuset i forskningsspørsmålet. Jeg savnet noen spørsmål som direkte omhandlet praktisk arbeid, for å kunne si noe konkret om lærernes vurderingspraksis der. Spørsmål om hvordan de gjennomførte praktisk arbeid, hvor ofte det ble gjort og hva de fokuserte på i denne type undervisning ble hengende over meg i analyse og diskusjon. Selv om hvert spørsmål i spørreundersøkelsen poengterte at det var knyttet til praktisk arbeid, er det vanskelig å si om lærerne har klart å skille deres vurderingspraksis i praktisk arbeid fra deres generelle vurderingspraksis i naturfag. Dette var likevel noe jeg ikke fikk gjort noe med i etterkant av undersøkelsen og jeg har valgt å gå ut i fra at lærerne har klart å svare spesifikt for praktisk arbeid.

Gjennom arbeidet med masteroppgaven min har jeg lært mye om formativ vurdering. Det ble også klart for meg underveis i prosessen hvor omfattende formativ vurdering faktisk er og hvor mange elementer forskningen anser som viktig i denne type vurdering. Spørsmålene i spørreundersøkelsen har forsøkt å ta for seg ulike elementer rundt formativ vurdering for å fange bredest mulig. Likevel er det et viktig poeng at det var jeg personlig som laget spørsmålene ut fra hvordan jeg tolket teorien om formativ vurdering. Konsekvensene her er at jeg i etterkant også har funnet mangler i spørreskjema for å kunne gi et helhetlig svar på i hvor stor grad lærere gjennomfører formativ vurdering. I tillegg ser jeg at det kunne vært aktuelt å ha et mindre fokusområde i forskningsspørsmålet mitt for å kunne gi en mer ryddig analyse og presentasjon av resultat og diskusjon. Det kunne for eksempel vært lurt å bare se på de tre formative spørsmålene; ”Hvor skal eleven?”; ”Hvor er eleven nå?”; og ”Hvordan kan eleven jobbe videre for å nå målet?”. Eventuelt bare fokusere på elevens rolle i vurderingen.

## 6 Avslutning

### 6.1 Konklusjon

For å besvare forskningsspørsmålet ”På hvilken måte og i hvor stor grad bruker naturfaglærere formativ vurdering i den praktiske delen av naturfagundervisningen i grunnskolen?” ble det gjennomført en spørreundersøkelse blant 61 naturfaglærere i grunnskolen. Det ble stilt ulike spørsmål som skulle si noe om lærernes formative vurderingspraksis i den praktiske delen av naturfag. Det viste seg at mange lærere svarte at de alltid eller nesten alltid brukte vurdering underveis i praktisk arbeid, noe som kan bety at lærerne har et fokus på formativ vurdering. Omtrent 85 % av lærerne svarte også at de alltid eller nesten alltid ga muntlig tilbakemelding til elevene, noe som åpner opp for gode klasseromsdialoger i forbindelse med praktisk arbeid. Dette kan også være med på å tilrettelegge for at det gjøres formative vurderinger underveis i undervisningen.

Det viste seg at lærerne i spørreundersøkelsen i stor grad var opptatt av å gi elevene læringsmål i praktisk arbeid, noe som kan bety at de har som mål at praktisk arbeid skal gi læring i naturfag. At elevene får god informasjon om hva som er målet med undervisningen er en viktig del av den formative vurderingspraksisen (Hattie & Timperley, 2007; Wiliam, 2011). Likevel viste det seg at lærerne i mye mindre grad ga elevene konkrete kjennetegn og eksempler på god måloppnåelse, noe som kan føre til at elevene ikke får tilstrekkelig informasjon om hvordan de skal nå målet på best mulig måte. Det viste seg videre at mange lærere også ga konkrete tilbakemeldinger på hvor eleven var i læringsprosessen og hvordan de kunne gå videre for å oppnå målet, noe som betyr at mange lærere gir elevene formativ vurdering i praktisk arbeid.

En viktig del av formativ vurderingspraksis er elevmedvirkning (Black et al., 2003; Hattie & Timperley, 2007; Wiliam, 2011). Det viste seg at få lærere i denne studien hadde et stort fokus på egen- og hverandrevurdering i praktisk arbeid i naturfag. Både på spørsmål om det ble lagt til rette for egenvurdering og hverandrevurdering i praktisk arbeid, og på spørsmål om de lot elevene bidra i å bestemme læringsmål, var det få lærere som ofte hadde denne praksisen.

For å besvare underspørsmålet ”Påvirkes lærernes vurderingspraksis av (1) undervisningserfaringen i faget; (2) klassetrinn det undervises på og (3) antall studiepoeng i naturfag?” ble det laget konstrukt for *god formativ vurdering, nivå av tilbakemelding og egen- og hverandrevurdering*. Målet var å sammenligne besvarelsene for flere av spørsmålene i spørreundersøkelsen opp mot de ulike lærergruppene. Til tross for at ikke alle variansanalysene viste signifikante forskjeller tyder likevel dataen på at lærerne blir bedre på vurdering etter hvor lang arbeidserfaring lærerne har i naturfag. En grunn til dette kan være at lærerne blir tryggere på gjennomføring av praktisk arbeid etter hvert som erfaringen kommer og derfor i større grad kan fokusere på formativ vurdering. Når det gjaldt hvilket klassetrinn lærerne jobbet på viste det seg at lærere på de laveste trinnene (1.-4. trinn) skåret lavere enn lærere på mellomtrinnet og ungdomstrinnet på de ulike konstruktene. En av grunnene kan være mindre fokus på vurdering i praktisk arbeid hos de yngste elevene. Det var lite som tydet på at det var noe sammenheng i hvor mange studiepoeng lærerne hadde i naturfag. Resultatene kunne likevel antyde at lærere med over 60 studiepoeng i naturfag skåret høyere og det kan antas at studiepoeng har betydning for lærernes formative vurderingspraksis når det er snakk om å inneha et betydelig antall studiepoeng.

Det er vanskelig å gjøre en kort oppsummering av i hvor stor grad lærerne bruker formativ vurdering i praktisk arbeid i naturfag. Dette er fordi formativ vurdering er en omfattende prosess i undervisningen og inneholder mange ulike komponenter. Formativ vurdering skal blant annet si noe om både hvor eleven skal, hvor eleven er i prosessen og hvordan eleven kan nå målene med undervisningen. I og med at praktisk arbeid i naturfag ofte også inneholder mange ulike mål kan det derfor være vanskelig å vurdere i hvor stor grad den formative vurderingen finner sted ved de ulike delene av praktisk arbeid. Naturfaglærere viser likevel at de bruker mange ulike komponenter i formativ vurdering også i praktisk arbeid i naturfag, men at elevmedvirkning i vurderingen brukes i mindre grad.

## **6.2 Videre forskning**

Denne studien bygger på en kvantitativ og anonym spørreundersøkelse, noe som gjør det vanskelig for meg å si noe om bakgrunnen for hvorfor lærerne har svart slik de har gjort. Det ville derfor vært interessant å gjennomføre noen kvalitative intervju for å undersøke dette nærmere. Spørsmål som hvorfor lærerne legger mest vekt på muntlig tilbakemelding og

hvordan denne tilbakemeldingen gis, hvordan de vurderer elevene underveis og hvordan de ser på praktisk arbeid i naturfag har dukket opp i arbeidet med datamaterialet. En dypere forståelse for deres formative vurderingspraksis ville derfor vært svært interessant å se nærmere på.

Noe som hadde vært svært interessant å se nærmere på var elevmedvirkning i vurderingsarbeidet. Jeg er nysgjerrig på om lærerne i større grad legger til rette for dette i ”vanlig” naturfagundervisning eller om dette er noe lærere gjennomgående bruker mindre tid på. Videre undersøkelser på hvordan elevmedvirkning i vurderingsarbeidet kan finne sted i naturfagundervisningen, både i ”vanlig” og praktisk undervisning hadde vært lærerikt. I tillegg ville det vært interessant å se på hvilken betydning denne type vurdering har å si for elevenes læring og motivasjon i naturfag.

Til tross for at spørsmålene i spørreundersøkelsen var siktet mot praktisk arbeid i naturfag var det ingen av spørsmålene som direkte knytter seg mot praktisk arbeid. Det vil derfor være interessant å forske nærmere på denne delen av naturfaget. Både hvilken betydning praktisk arbeid har for elevenes læring og motivasjon, men også mer konkret hvilken vurderingsformer som er hensiktsmessig for at elevene skal lære bedre. Både en ny spørreundersøkelse med spørsmål mer siktet mot praktisk arbeid, eller intervju/observasjon av praktisk arbeid kan være gode forskningsstrategier.

Denne studien hadde naturfaglærere som forskningsobjekter. Jeg vil til slutt foreslå å undersøke formativ vurdering og praktisk arbeid i naturfag ut fra et elevperspektiv. Hva elevene tenker rundt disse temaene kan være svært interessant og nyttig informasjon for naturfaglærer å inneha når vurdering og undervisning skal planlegges.



## Litteraturliste

- Abrahams, I. & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, 49(8), 1035-1055.
- Abrahams, I., Reiss, M. J. & Sharpe, R. M. (2013). The assessment of practical work in school science. *Studies in Science Education*, 49(2), 209-251.
- Bergem, O. K., Kaarstein, H. & Nilsen, T. (Red.). (2016). *Vi kan lykkes i realfag. Resultater og analyser fra TIMMS 2015*. Oslo: Universitetsforlaget.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International journal of educational research*, 35(5), 463-482.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*: McGraw-Hill Education (UK).
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, 86(1), 8-21.
- Black, P. & Wiliam, D. (1998). Inside The Black Box. *Phi Delta Kappan*, 80(2), 139.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5.
- Blömeke, S., Olsen, R. V. & Suhl, U. (2016). Relation of Student Achievement to the Quality of Their Teachers and Instructional Quality *Teacher Quality, Instructional Quality and Student Outcomes* (s. 21-50): Springer.
- Broadfoot, P., Daugherty, R., Gardner, J., Gipps, C., Harlen, W., James, M. & Stobart, G. (1999). *Assessment for learning: Beyond the black box*. Cambridge, UK: University of Cambridge School of Education.
- Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M. & Stobart, G. (2002). *Assessment for learning: 10 principles*. Cambridge, UK: University of Cambridge School of Education.
- Bruce, L. B. (2001). Student self-assessment: Making standards come alive. *Classroom Leadership*, 5(1), 1-6.
- Bybee, R. W., Taylor, J. A., Gardner, A., Westbrook, A. & Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness.
- Clotfelter, C., Ladd, H. & Vigdor, J. (2007). How and Why do Teacher Credentials Matter for Student Achievement? : National Bureau of Economic Research, Inc.
- Council, N. r. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*: National Academies Press.
- Dysthe, O. (2008). Klasseromsvurdering og læring. *Bedre skole*, 3, 16-23.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage Publications
- Forskrift til opplæringslova. (2006). Hentet fra <https://lovdata.no/dokument/SF/forskrift/2006-06-23-724>.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Hodson, D. (1990). A critical look at practical work in school science. *School Science Review*, 71(256), 33-40.
- Hofstein, A. & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of educational research*, 52(2), 201-217.

- Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty - first century. *Science education*, 88(1), 28-54.
- Johannessen, A., Tufte, P. A. & Christoffersen, L. (2016). *Introduksjon til samfunnsvitenskapelig metode*. Oslo: Abstrakt forlag.
- Kind, P. M. (1999). *Hva i all verden gjør elevene i realfag?: praktiske oppgaver i matematikk og naturfag*: Universitetet i Oslo, Institutt for lærerutdanning og skoleutvikling.
- Kjærnsli, M., Lie, S., Olsen, R. V. & Roe, A. (2007). Tid for tunge løft. *Norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006*.
- Knain, E. & Kolstø, S. D. (Red.). (2011). *Elever som forskere i naturfag*. Oslo: Universitetsforlaget.
- Kunnskapsdepartementet. (2011). *Motivasjon - Mestring - Muligheter*. (St. meld. nr. 22, 2010-2011). Hentet fra <https://www.regjeringen.no/no/dokumenter/meld-st-22-2010--2011/id641251/sec1>.
- Kunnskapsdepartementet. (2013). *Læreplan i naturfag (NAT1-03)*. Oslo: Utdanningsdirektoratet Hentet fra <https://www.udir.no/kl06/NAT1-03>.
- Lagerstrøm, B. O., Moafi, H. & Revold, M. K. (2014). *Kompetanseprofil i grunnskolen*. Oslo: Statistisk sentralbyrå.
- Millar, R. (2011). Practical Work. I J. Osborne & J. Dillon (Red.), *Good practice in science teaching: What research has to say* (s. 108-134). Maidenhead: Open University Press.
- Millar, R. (2013). Improving science education: Why assessment matters *Valuing assessment in science education: Pedagogy, curriculum, policy* (s. 55-68): Springer.
- NESH. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, jus og humaniora*. fra Den nasjonale forskningsetiske komite for samfunnsvitenskap og humaniora <https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>
- Nott, M. & Wellington, J. (1999). The state we're in: issues in key stage 3 and 4 science. *School Science Review*, 81, 13-18.
- Osborne, J. & Dillon, J. (2010). *Good practice in science teaching: What research has to say: What research has to say*: McGraw-Hill Education (UK).
- Ottander, C. & Grelsson, G. (2006). Laboratory work: the teachers' perspective. *Journal of Biological Education*, 40(3), 113-118.
- Pallant, J. (2010). *SPSS Survival Manual* (4 utg.). UK: McGraw-Hill Education.
- Reiss, M., Abrahams, I. & Sharpe, R. (2012). Improving the assessment of practical work in school science: Gatsby Charitable Foundation, London, <http://www.gatsby.org.uk/~-/media/Files/Education/Improving%20the%20assessment%20of%20practical%20work%20in%20school%20science.ashx> (last accessed 19 March 2013).
- Ringdal, K. (2013). *Enhet og mangfold: samfunnsvitenskapelig forskning og kvantitativ metode*: Fagbokforlaget.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in education: principles, policy & practice*, 5(1), 77-84.
- Sandvik, L. & Buland, T. (2014). *Vurdering i skolen. Utvikling av kompetanse og fellesskap. Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)*: Trondheim: NTNU, program for lærerutdanning, og SINTEF.
- Statistisk sentralbyrå. (2017). *Nøkkeltall for utdanning*. Hentet fra <http://www.ssb.no/utdanning/nokkeltall>.
- Sullivan, G. M. & Artino Jr, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of graduate medical education*, 5(4), 541-542.
- Thronsen, I., Hopfenbeck, T. N., Lie, S. & Dale, E. L. (2009). *Bedre vurdering for læring: Rapport fra "Evaluerer av modeller for kjennetegn på måloppnåelse i fag"*. *Better*

*Assessment of Learning. Report from " Evaluation of Models for Achievement Criteria in Subjects" ], Universitetet i Oslo.*

Utdanningsdirektoratet. (2016). Jobb med læreplan, mål, kjennetegn og kriterier. Hentet fra <http://www.udir.no/laring-og-trivsel/vurdering/undervisvurdering/mal-og-kriterier/>

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14.

# Vedlegg

## Vedlegg 1: Spørreundersøkelse

### **Kjære naturfaglærer!**

Mitt navn er Karoline Belt og jeg studerer på lærerutdanningen ved Universitetet i Tromsø. I tilknytning til min masteravhandling i naturfagdidaktikk ønsker jeg å finne ut hvordan lærere jobber med formativ vurdering (vurdering som støtter læring) i den praktiske delen av naturfaget. Praktisk arbeid i naturfag vil si all type undervisning og aktiviteter hvor elever, individuelt eller i grupper, blir involvert i å jobbe med og/eller observere virkelige objekter og materialer, som for eksempel eksperimenter på laboratorium, lærerdemonstrasjoner og undersøkelser ute i naturen.

Jeg ønsker i denne forbindelse å spørre deg som lærer om din praksis knyttet til dette og vil sette stor pris på om du kunne gi noen minutter av din tid for å svare på en kort spørreundersøkelse.

#### 1. Hvor mange år har du jobbet som naturfaglærer?

- 0-3 år                       4-10 år                       11-20 år                       over 21 år

#### 2. På hvilket klassetrinn har du undervist mest i faget?

- 1.-4. trinn                       5.-7. trinn                       8.-10. trinn                       Videregående skole

#### 3. Hvor mange studiepoeng har du i naturfag? 1 vekttall tilsvarer 3 studiepoeng.

- 0-15 stp.                       16-30 stp.                       31-60 stp.                       Over 61 stp.

#### 4. I praktisk arbeid i naturfag, hvor ofte...

	Aldri	Nesten aldri	Iblant	Nesten alltid	Alltid
...gir du muntlig tilbakemelding til elevene?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du skriftlig tilbakemelding til elevene?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...vurderer du elevene underveis i prosessen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...bruker du vurdering for å tilpasse undervisningen?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

#### 5. I praktisk arbeid i naturfag, hvor ofte...

	Aldri	Nesten aldri	Iblant	Nesten alltid	Alltid
...gir du elevene læringsmål?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du elevene kjennetegn på måloppnåelse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du elevene eksempler på måloppnåelse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...lar du elevene bidra med å bestemme læringsmål?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

...gir du tilbakemelding på hva eleven har forstått/lært av de aktuelle læringsmålene?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du tilbakemelding på hva eleven <i>ikke</i> har forstått/lært av de aktuelle læringsmålene?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du konkrete tilbakemeldinger på hva eleven må gjøre for å forbedre seg?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...justerer du oppgavene ut fra kompetansen eleven viser?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## 6. I praktisk arbeid i naturfag, hvor ofte...

	Aldri	Nesten aldri	Iblant	Nesten alltid	Alltid
...gir du tilbakemelding på elevens svar på oppgaven og/eller forståelse av oppgaven?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du tilbakemelding på elevens fremgangsmåte for å løse oppgaven?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du tilbakemelding som oppmuntrer eleven til selv å reflektere over oppgaven og avdekke eventuelle feil/mangler?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du tilbakemelding i form av ros, som f.eks. ”bra”, ”du er flink”, ”bra jobbet”?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...gir du tilbakemelding i form av kritikk, som f.eks. ”ikke bra”, ”dårlig”?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## 7. I praktisk arbeid i naturfag, hvor ofte...

	Aldri	Nesten aldri	Iblant	Nesten alltid	Alltid
...legger du til rette for at elevene vurderer seg selv?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...legger du til rette for at elevene vurderer hverandre?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...får elevene mulighet til å jobbe videre med tilbakemeldinger på eget arbeid?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Takk for ditt bidrag!

Med vennlig hilsen  
Karoline Belt