

Lise Vikan Sandvik, Henning Fjørtoft, Gustaf B. Skar,
Marita Byberg Johansen, Anne Holten Kvistad, Olav
Dalsegg Tokle, Stine Aarønes Angvik, Bodil
Svendsen, Runa Næss Korshavn, Egil G. Gjølme

Standpunktvurdering – praksis og prosesser i LK20

Trondheim 30.01.2026

FORORD

Denne rapporten presenterer resultater fra prosjektet *Standpunktvurdering – praksis og prosesser i LK20*, gjennomført av Skrivesenteret, NTNU, Institutt for lærerutdanning, på oppdrag for Utdanningsdirektoratet. Formålet med prosjektet har vært å belyse hvordan standpunktvurdering forstås og praktiseres i ulike fag og skoleslag etter innføringen av LK20/LK20S, og å bidra med kunnskap som kan støtte videre utvikling av vurderingsarbeidet i norsk skole.

Vi vil rette en stor takk til lærere og skoleledere som har deltatt i intervjuer og spørreundersøkelser, og som velvillig har delt sine erfaringer og refleksjoner. Deres bidrag har vært avgjørende for gjennomføringen av studien.

Prosjektet er gjennomført i perioden november 2023 til november 2025, og bygger på en kombinasjon av kvalitative intervjuer, dokumentanalyse og nasjonale spørreundersøkelser.

Prosjektet er gjennomført av følgende deltakere:

- Prosjektleder: Professor Lise Vikan Sandvik
- Prosjektgruppe: Professor Henning Fjørtoft, professor Gustaf B. Skar, universitetslektor Marita Byberg Johansen (også koordinator i prosjektet), førstelektor Anne Holten Kvistad (også koordinator i prosjektet), universitetslektor Olav Tokle, universitetslektor Stine Aarønes Angvik, førsteamanuensis Bodil Svendsen, universitetslektor Runa Næss Korshavn, dosent Egil G. Gjølme

Alle i prosjektgruppen har deltatt i datainnsamling, analyse av data og i arbeidet med formidling og rapportering av funn.

Vi vil takke Utdanningsdirektoratet for nyttige tilbakemeldinger i arbeidet med rapporten.

Forfatterne står selv ansvarlige for rapportens innhold og tolkninger.

Trondheim, januar 2026

Lise Vikan Sandvik

Rapport

Standpunktvurdering – praksis og prosesser i LK20

FORFATTER(E)

Lise Vikan Sandvik, Henning Fjørtoft,
Gustaf B. Skar, Marita Byberg
Johansen, Anne Holten Kvistad, Olav
Dalsegg Tokle, Stine Aarønes Angvik,
Bodil Svendsen, Runa Næss Korshavn,
Egil G. Gjølme

Trondheim 30.01.2026

NTNU
Nasjonalt senter for skriveopplæring
og skriveforskning
<https://skrivesenteret.no/>

OPPDRAUGSGIVER

Utdanningsdirektoratet

ISBN 978-82-7923-107-3



Standpunktvurdering – praksis og prosesser i LK20

© 2026 av Lise Vikan Sandvik et al. er lisensiert med en Creative Commons lisens CC BY 4.0. Du har lov til å kopiere, distribuere og spre verket i hvilket som helst medium eller format til et hvilket som helst formål, inkludert kommersielle. Du kan remixe, endre, og bygge videre på materialet til et hvilket som helst formål, inkludert kommersielle. Du må oppgi korrekt kredittering oppgi en lenke til lisensen, og indikere om endringer er blitt gjort. Se fullstendige lisensbetingelser på <https://creativecommons.org/licenses/by/4.0/deed.no>

INNHold

1. SAMMENDRAG	6
2. INNLEDNING	7
2.1 FORMÅLET MED OPPDRAGET OG VALG AV FAG I STUDIEN	7
2.2 PROBLEMSTILLING OG FORSKNINGSSPØRSMÅL	7
3. KUNNSKAPSGRUNNLAG	9
3.1 VURDERING OG STANDPUNKTVURDERING	9
3.2 KVALITETSKRITERIER FOR STANDPUNKTVURDERING	11
3.3 BETYDNINGEN AV FAGKOMPETANSE I VURDERING	12
3.4 SAMARBEID OG PROFESJONSFELLESSKAP	12
3.5 ALIGNMENT OG ACCOUNTABILITY I VURDERINGSSYSTEMER	13
3.5.1 Alignment i vurderingssystemer	13
3.5.2 Accountability i vurderingssystemer	14
3.5.3 Samspillet mellom alignment og accountability	14
4. METODE	16
4.1 DATAMATERIALE	16
4.2 DET KVALITATIVE DATAMATERIALET	17
4.3 DET KVANTITATIVE DATAMATERIALET	18
4.3.1 Datainnsamling og rekruttering	18
4.3.2 Spørreskjema utvikling og pilotering	19
4.3.3 Svarskala og tolkning av gjennomsnittsskårer	19
4.3.4 Analyse	19
4.3 METODISKE BEGRENSNINGER	20
4.4 VALG AV FAG I STUDIEN	21
4.4.1 Norsk	21
4.4.2 Matematikk	22
4.4.3 Naturfag	22
4.4.4 Musikk	22
4.4.5 Helse- og oppvekstfag (HO)	22
4.4.6 Samfunnsfag	23
4.4.7 Mat og helse	24
4.4.8 Kroppsøving	24
5. FUNN	25
5.1 BRUK AV VURDERINGSTEKSTENE OG KJENNETEGN PÅ MÅLOPPNÅELSE	25
5.1.1 Bruk av tekstene om underveis- og standpunktvurdering	25
5.1.2 Bruk av kjennetegn på måloppnåelse	27
5.1.3 Bruk av vurderingstekstene i skolens kvalitetsarbeid	28
5.1.4 Likheter og forskjeller mellom fag i bruken av vurderingstekstene og kjennetegn på måloppnåelse	29
5.2 PRAKSIS OG PROSESSER I STANDPUNKTVURDERING	30
5.2.1 Forholdet mellom underveis- og standpunktvurdering	30
5.2.2 Forståelse og bruk av karakterskalaen	31
5.2.3 Vurderingsgrunnlag og vurderingsformer	33
5.2.4 Samarbeid og profesjonsfellesskap i vurderingsarbeidet	34
5.2.5 Skoleledelsens rolle i kvalitetssikring av standpunktvurdering	36
5.2.6 Elevinvolvering i vurderingsarbeidet	37
5.2.7 Press, normendringer og karakterøkning	37

5.2.8 Likheter og forskjeller mellom fag i standpunktvurdering	38
5.3 OPPSUMMERING AV SENTRALE FUNN	39
6. DISKUSJON	41
6.1 SPENNINGER MELLOM NASJONALT, LOKALT OG INDIVIDUELT NIVÅ	42
6.2 SPENNINGEN MELLOM NASJONALE INTENSJONER OG LOKALE PRAKSISER	43
6.3 PROFESJONELT SKJØNN I MØTE MED ANSVARLIGGJØRING	44
6.4 FAGLIG AUTONOMI OG BEHOVET FOR FELLES STANDARDER	45
6.5 HÅNDBARHET SOM FORUTSETNING FOR VURDERINGSKVALITET	46
6.6 RETTFERDIGHET OG LEGITIME FORVENTNINGER FRA SAMFUNNET	47
6.7 AVSLUTTENDE REFLEKSJONER: STANDPUNKTVURDERING SOM SYSTEMISK PRAKSIS	48
7. IMPLIKASJONER OG ANBEFALINGER	50
8. LITTERATUR	53
VEDLEGG 1 NORSK	59
VEDLEGG 2 MATEMATIKK	63
VEDLEGG 3 NATURFAG	68
VEDLEGG 4 MUSIKK	73
VEDLEGG 5 HELSE- OG OPPVEKSTFAG	77
VEDLEGG 6 SAMFUNNSFAG	83
VEDLEGG 7 MAT OG HELSE	87
VEDLEGG 8 KROPPSØVING	91
VEDLEGG 9 LEDELSE	96
VEDLEGG 10 RAPPORT FRA SPØRREUNDERSØKELSE 1 OM VURDERINGSTEKSTENE OG STANDPUNKTVURDERING	101
VEDLEGG 11 RAPPORT FRA SPØRREUNDERSØKELSE 2 OM PRESS OG TILLIT I STANDPUNKTVURDERING	127

1. SAMMENDRAG

Denne studien har undersøkt praksis og prosesser knyttet til standpunktvurdering etter innføringen av LK20/LK20S, med utgangspunkt i lærerintervjuer, skolelederintervjuer, dokumentanalyse og nasjonale spørreundersøkelser. Formålet har vært å belyse hvordan vurderingstekstene om underveis- og sluttvurdering og kjennetegn på måloppnåelse forstås og brukes, og hvordan standpunktvurdering beskrives og begrunnes i og på tvers av fag og skoleslag.

Studien viser at kvalitet i standpunktvurdering utvikles i et samspill mellom nasjonale rammer, lokale strukturer og lærernes profesjonelle skjønn. Funnene viser at lærere i stor grad kjenner til vurderingstekstene i læreplanverket, og at tekstene oppfattes som legitime og relevante. Samtidig har tekstene begrenset funksjon som operative verktøy i det daglige vurderingsarbeidet. I praksis er det kompetansemålene og lokale faglige fortolkninger som i størst grad styrer vurderingsarbeidet. Kjennetegn på måloppnåelse brukes i mange fag, men er ofte lokalt tilpasset.

Det er betydelig variasjon i praksis mellom fag, lærere og skoler. Noen former for variasjon er faglig legitime og er knyttet til fagenes egenart og lokale rammebetingelser. Mer problematiske former for variasjon skyldes fravær av strukturer for samarbeid, der lærere kan utvikle og kalibrere felles forståelse. Årsakene til variasjonen ser ut til å være knyttet til ulikheter i lokale strukturer, som tilgang til faglige samarbeidsarenaer og profesjonsfellesskap. Dette er særlig tydelig i praktisk-estetiske fag og i yrkesfag, der vurderingsgrunnlaget ofte er mer situasjonsavhengig og krevende å beskrive med felles forankrede kriterier. Studien viser også at skoleledelsens prioritering og tilrettelegging har stor betydning for i hvilken grad vurderingsarbeidet preges av kollektive refleksjoner fremfor individuelle praksiser. Ved skoler som har etablerte strukturer for samarbeid om vurdering, og der skoleledelsen prioriterer og støtter dette arbeidet, framstår derimot vurderingsarbeidet som mer samordnet og faglig forankret. Nasjonale føringer forutsetter derfor at profesjonelle strukturer og lokal kapasitet er nødvendig for å kunne overordnede krav til praksis i skolen.

Standpunktvurdering framstår som preget av vedvarende faglige dilemmaer. Lærere må balansere mellom å gi en helhetlig og faglig god vurdering, sørge for at vurderingen er rettferdig og til å stole på, og samtidig håndtere rammer som tid, ressurser og krav til dokumentasjon. Funnene viser tydelige spenninger mellom lærernes faglige skjønn og kravene om begrunnelser og innsyn i vurderingsarbeidet.

Et særlig interessant mønster er at lærere i liten grad opplever seg selv som mer generøse i vurderingen, samtidig som mange rapporterer at terskler for karaktersetting har blitt justert over tid. Dette kan peke mot gradvise endringer i felles praksiser og normer i fagmiljøene, heller enn individuelle valg hos enkeltlærere, samtidig som variasjonen i svar viser at utviklingen ikke er entydig.

Et gjennomgående funn er at kvalitet i standpunktvurdering i stor grad er et systemansvar, ikke et individuelt anliggende. Samtidig forutsetter systemet høy profesjonell kompetanse og faglig dømmekraft hos lærerne. Utvikling av denne kompetansen i lokale sammenhenger framstår som viktig for kvalitet og likeverdighet. Samspillet mellom nasjonale rammer, lokale strukturer og lærernes profesjonelle dømmekraft framstår som avgjørende for å sikre likeverdig praksis. Det er derfor behov for å styrke strukturer som støtter felles forståelse, profesjonsfellesskap og praksisnær støtte til vurderingsarbeidet.

2. INNLEDNING

2.1 Formålet med oppdraget og valg av fag i studien

NTNU-ILU har fått i oppdrag av Utdanningsdirektoratet å belyse praksis og prosesser knyttet til standpunktvurdering i tråd med de nye læreplanene. Prosjektet har flere målsetninger: Det skal kaste lys over bruken av tekster om underveis- og sluttvurdering i læreplanene, samt kjennetegn på måloppnåelse. Det skal undersøke hvordan lærere anvender tekstene i vurderingsarbeidet og om de opplever dem som nyttige. Det skal også undersøke hvordan lærere og skoleledere forstår og praktiserer underveis- og sluttvurdering i læreplanene. Kunnskap fra oppdraget gi Utdanningsdirektoratet et godt utgangspunkt for å vurdere tiltak som kan støtte lokalt kvalitetsarbeid knyttet til standpunktvurderingen, samt vurdere eventuelle utprøvinger eller ytterligere kunnskapsbehov.

Oppdragsgiver ønsket at undersøkelsen skulle dekke fag der vurderingstekstene er fagspesifikke og fag der tekstene er generelle uten fagspesifisering. På bakgrunn av denne bestillingen valgte NTNU-ILU fag med ulike karakteristikk og vurderingsformer for å oppnå maksimal variasjon i undersøkelsen. Dette ga oss muligheten til å utforske ulike aspekter av vurderingspraksis og -prosesser i forskjellige kontekster. Noen fag har for eksempel et utpreget «spiralprinsipp», der kunnskaper og ferdigheter stadig kommer tilbake i mer avanserte former, mens andre fag er mer preget av å være organisert i moduler eller blokker. Fagene som ble inkludert i studien, var: *norsk, matematikk, naturfag, musikk og helse- og oppvekstfag (HO)*. De tre førstnevnte fagene er gjennomgående fra ungdomsskole til og med videregående skole. Ett av fagene, musikk, er et praktisk-estetisk fag. Ett av fagene, helse- og oppvekstfag, er et yrkesfag i videregående skole, som i motsetning til de andre fagene, har vurderingstekster som er generelle uten fagspesifisering. I tillegg har vi intervjuet skoleledere om skoleledernes erfaringer knyttet til problemstillingene. Vi sendte også ut en spørreundersøkelse til lærere i ungdomsskolen og i videregående skole.

Oppdraget ble utvidet med et tilleggsoppdrag våren 2025, og inkluderte fagene *samfunnsfag og mat og helse* på ungdomstrinnet, samt *kroppsøving* på ungdomstrinnet og i videregående skole. I dialog med oppdragsgiver utvidet vi undersøkelsen med spørsmål om opplevd karakterpress i fagene både i intervjuene og i en ny spørreundersøkelse til alle lærere i ungdomsskolen og i videregående skole i Norge.

Rapporten er bygd opp med et innledningskapittel, et kunnskapsgrunnlag, et metodekapittel, en sammenskriving av funn på tvers av fag, samt et drøftingskapittel og et kapittel med implikasjoner og anbefalinger. Som vedlegg finnes det rapporter som presenterer de kvalitative funnene for hvert av fagene i undersøkelsen (fagrapporter). Resultatene fra to spørreundersøkelsene finnes også som to vedlegg.

2.2 Problemstilling og forskningsspørsmål

Den overordnede problemstillingen for NTNU-ILUs undersøkelse har vært å undersøke praksis og prosesser i standpunktvurdering etter innføringen av LK20/LK20S.

For å operasjonalisere formålet, stilte vi derfor to forskningsspørsmål. For å presisere hva vi mener med de to spørsmålene, utdypet vi dem med underspørsmål.

Forskningsspørsmål 1: Hvordan benyttes vurderingstekstene i læreplanene og kjennetegnene på måloppnåelse?

- Hva kjennetegner bruken av tekstene om underveisvurdering, standpunktvurdering og kjennetegn på måloppnåelse i lærernes vurderingsprosesser?
- Hvordan påvirker vurderingstekstene og kjennetegnene på måloppnåelse lærernes tilnærming til vurderingsarbeidet?
- I hvilken grad inngår tekstene om vurdering til skolens arbeid med kvalitet i standpunktvurderingen?

Forskningsspørsmål 2: Hva kjennetegner praksis og prosesser i standpunktvurdering?

- Hvordan forstås og brukes karakterskalaen? (f.eks. av lærere med ulik utdanning, i ungdomsskole eller videregående skole)
- Hva blir vektlagt ved fastsetting av standpunktkarakter? Hva slags vurderingsformer danner grunnlag for standpunktvurdering? Hvilke ressurser (f.eks. kriterier, kjennetegn) anvender lærere i karaktersetting? Finnes det forskjeller mellom fagene i vektingen av ulike vurderingsaspekter?
- Hvordan er sammenhengen mellom underveis- og standpunktvurdering i praksis for lærere? Hvordan brukes underveisvurdering til å informere standpunktvurdering? Hvordan tar lærere hensyn til elevens utvikling over tid når de fastsetter standpunktkarakterer?
- Hvordan jobber lærere i samme fag med å utvikle felles forståelse for læreplan, forskrift og karaktersetting? Hvilke samarbeidsarenaer og ressurser bruker lærere for å utvikle en felles forståelse av vurderingspraksis og karaktersetting i faget? Hvordan bidrar samarbeidet til å sikre etterlevelse av regelverket?
- Hvordan legger skoleledelse til rette for å sikre kvalitet i vurdering? Hvordan støtter skoleledere arbeidet med å implementere regelverket for vurdering? Hva slags ressurser og veiledning tilbys til lærere for å sikre kvalitet i vurdering?

3. KUNNSKAPSGRUNNLAG

3.1 Vurdering og standpunktvurdering

I denne studien forstås vurdering som prosesser der lærere samler inn informasjon om elevers læring og trekker faglig begrunnede slutninger om elevenes kompetanse. Sluttvurdering, i form av standpunktkarakterer og eksamen, har stor betydning for elevers videre utdannings- og yrkesmuligheter. Det stilles derfor høye samfunnsmessige krav til kvalitet, legitimitet og rettssikkerhet i denne delen av vurderingssystemet.

Standpunktvurdering utgjør en sentral del av den norske sluttvurderingsordningen. I forskrift til opplæringslova § 9-6 slås det fast at standpunktkarakteren skal uttrykke elevens samlede kompetanse ved avslutningen av opplæringen. Kompetansebegrepet i LK20 er definert på følgende måte i LK20:

Kompetanse er å kunne tilegne seg og anvende kunnskaper og ferdigheter til å mestre utfordringer og løse oppgaver i kjente og ukjente sammenhenger og situasjoner. Kompetanse innebærer forståelse og evne til refleksjon og kritisk tenkning.
(Overordnet del av læreplanverket, 2.2)

Standpunktvurdering skal gjenspeile elevens helhetlige kompetanse i faget, og vil derfor kreve et mangfold av vurderingssituasjoner.

Internasjonalt finnes det betydelige variasjoner i hvordan sluttvurdering organiseres. Noen land baserer seg utelukkende på eksterne eksamener (f.eks. Irland), mens andre legger langt større vekt på standardiserte tester og opptaksprøver (f.eks. USA). Det norske systemet skiller seg ut ved at lærernes egne vurderinger har en svært sterk posisjon. Om lag 80 prosent av karaktergrunnlaget bygger på standpunkt, mens rundt 20 prosent kommer fra eksterne eksamener (Hovdhaugen et al., 2018). Dette bygger på en lang tradisjon der læreres profesjonelle nærhet til elevenes læringsprosesser anses som en styrke.

Denne tradisjonen innebærer at standpunktvurdering i Norge i stor grad er forankret i profesjonelt skjønn. Lærere fremheves som sentrale aktører i vurderingsarbeidet, og systemet hviler på en tillit til at lærere kan utøve faglige og etiske vurderinger på en forsvarlig måte. Samtidig har denne ordningen alltid vært diskutert, nettopp fordi den åpner for lokal variasjon i praksis mellom fag, skoler og lærere.

Forskning fra norsk kontekst viser at de fleste lærere opplever at de setter rettferdige karakterer basert på varierte vurderingsgrunnlag og i tråd med forskriftene. Selv om lærere forstår standpunktkarakterer primært som uttrykk for måloppnåelse, kan andre faktorer som elevenes innsats og deltakelse likevel påvirke vurderingen. Lærere rapporterer videre om spenninger mellom rollen som «trener» og «dommer», og nyutdannede lærere kan oppleve karaktersetting som særlig krevende, ikke minst i møte med foresatte (Birkelund & Midthaugen, 2019). I enkelte fag, som musikk og yrkesfag, pekes det på mangel på ressurser, svakere faglig infrastruktur og press om å ikke stryke elever (Fjørtoft & Morud, 2021; Vinge, 2014).

Evalueringen av fagfornyelsen (LK20) peker på særtrekk ved de praktiske og estetiske fagene som er relevante for forståelsen av vurderingsutfordringer i disse fagene. I Delrapport 1 beskrives hvordan det i forarbeidene til LK20 er formulert store forventninger til at de praktiske og estetiske fagene skal bidra med andre perspektiver og erfaringer enn øvrige skolefag, og dermed

representere «noe annet» i opplæringen (Borgen et al., 2023a). Samtidig peker evalueringen på at det tar tid før nye intensjoner blir realisert i praksis.

I Delrapport 2 understrekes det også at LK20 stiller høye forventninger til fagenes praktiske karakter, og at lærere, skoleledere og skoleeiere omtaler fagene som en arena der elever trives og får andre erfaringer enn i andre fag (Borgen et al., 2023b). Rapporten viser videre at læreplanene i de praktiske og estetiske fagene har blitt mer prosessorienterte og komplekse, og at det er forskjeller mellom fagene i hva som vektlegges i vurderingstekstene om underveis- og standpunktvurdering. Dette kan bidra til et tolkningsrom i vurderingsarbeidet, og dermed til variasjon i hvordan vurderingsgrunnlag etableres og hvordan kriterier forstås og praktiseres i fagene.

I evalueringen av fagfornyelsen i yrkesfag peker flere delrapporter på forhold som kan være relevante for vurderingspraksis og standpunkt. Delrapport 2 beskriver at læreplanene oppleves som bedre tilpasset arbeidslivet, men samtidig krevende å implementere, blant annet knyttet til endringer i tilbudsstruktur og overgangen fra vg1 til vg2 (Skålholt et al., 2023). Slike endringer kan påvirke hvordan opplæringen organiseres lokalt, og dermed også hvilke vurderingssituasjoner og hvilket vurderingsgrunnlag elevene får.

Delrapport 3 retter søkelyset mot overgangen til opplæring i bedrift og erfaringer fra læretiden, sett fra lærlinger, lærebedrifter og opplæringskontor (Skålholt et al., 2024). Rapporten synliggjør hvordan opplæringen i bedrift innebærer andre rammer, forventninger og praksiser enn i skolen. Dette kan ha betydning for sammenhengen mellom opplæring og vurdering, og kan gi et viktig bakteppe for drøftingen av vurderingsutfordringer i yrkesfaglige utdanningsprogram.

En viktig kontekst for standpunktvurdering er også utviklingstrekk som karakterpress og karakterinflasjon. Norske data viser en gradvis økning i standpunktkarakterer over tid. En slik karakterøkning kan ha flere mulige forklaringer, og innebærer ikke nødvendigvis karakterinflasjon. Utdanningsdirektoratet peker på at endringer i karaktergjennomsnitt kan henge sammen med blant annet endringer i elevgrunnlag, vurderingsformer, faglig innhold og praksiser for fastsetting av karakterer, i tillegg til mer generelle endringer i skole og samfunn (Utdanningsdirektoratet, 2025).

I forskningslitteraturen brukes karakterinflasjon ofte om situasjoner der karakterer over tid blir høyere uten at elevenes faktiske kompetanse har økt tilsvarende, eller der standardene for hva som regnes som måloppnåelse, gradvis forskyves (se f.eks. Parkes, 2013). Forskning peker på flere mulige mekanismer som kan bidra til slike forskyvninger, blant annet forhandlinger mellom elev og lærer, et ønske om å motivere elever, og en tendens til at tvilen kommer eleven til gode i grensetilfeller (Parkes, 2013; Rambøll, 2020). Nyere forskning på vurderingspraksiser viser også at læreres karakterpraksiser utvikles i et spenn mellom faglige standarder, lokale normer og krav om begrunnelse og dokumentasjon, og at dette kan bidra til endringer i hvordan karakterskalaen praktiseres over tid (Gamlem et al., 2024).

Sammen med sterkt karakterpress kan slike forhold påvirke læreres vurderingspraksis, særlig i fag der vurderingsgrunnlaget er komplekst og hvor karaktergrenser oppleves som vanskelige å trekke (Fjørtoft & Morud, 2021). Samlet tilsier dette at endringer i karakternivå bør forstås som et sammensatt fenomen, der både systemfaktorer, fagtradisjoner og profesjonelle normer kan spille inn.

Videre påvirker sluttvurderingsordningene undervisning og læring gjennom såkalte *washback*-effekter. I norsk sammenheng kan eksamen fungere som et viktig korrektiv ved å tydeliggjøre læreplanmål og bidra til felles forventninger. Sensorskolering og nasjonale rammer kan styrke tolkningsfellesskap på tvers av skoler (Federici et al., 2017; Hovdhaugen et al., 2022). Samtidig kan et sterkt eksamensfokus føre til en innsnevring av undervisningen, der lærere prioriterer det som anses som «eksamensrelevant» framfor fagets bredere kompetansemål. Forskning viser også systematiske forskjeller mellom standpunkt og eksamen etter blant annet kjønn, skoletyper og prestasjonsnivå, noe som indikerer at de to vurderingsformene ikke alltid måler samme type kompetanse (Hovdhaugen et al., 2018). Dette gjør washback-effekten sentral for å forstå spenninger og samsvar mellom ulike deler av sluttvurderingen. En nyere studie som sammenlignet skriftlige eksamensresultater i norsk på videregående skole viste at elever med dårlige karakterer eller som strøk, hadde økt risiko for å utvikle mentale helseproblemer i påfølgende år (Beck et al., 2023). Dette understreker viktigheten av å forstå hvordan sluttvurdering påvirker elevers livssituasjon.

Sammenfattende kan standpunktvurdering forstås som en ordning som kombinerer et sterkt profesjonelt skjønn med høye krav til legitimitet og rettssikkerhet. Denne dobbeltheten gjør standpunkt til et sentralt, men også utfordrende element i det norske vurderingssystemet.

3.2 Kvalitetskriterier for standpunktvurdering

For å analysere kvaliteten i standpunktvurdering benytter prosjektet et rammeverk basert på Utdanningsdirektoratets fire overordnede kvalitetskriterier: validitet, reliabilitet, rettferdighet og håndterbarhet (Utdanningsdirektoratet, 2024). Disse kriteriene er utviklet av Utdanningsdirektoratet som et rammeverk for sentralt gitt eksamen, men er også relevante som analytiske verktøy for standpunktvurdering, selv om de får noe annet innhold når vurdering foregår over tid og baseres på profesjonelt skjønn.

Validitet handler om hvorvidt vurderingsgrunnlaget gir et dekkende bilde av den kompetansen som faktisk skal vurderes. I standpunktvurdering innebærer dette at lærernes vurderingsgrunnlag, som oppgaver, aktiviteter og vurderingspraksis, må være tett knyttet til kompetansemålene i læreplanen (Kane, 2016). Et sentralt spørsmål er derfor hvordan lærere samler inn, tolker og vektet informasjon om elevenes utvikling gjennom året, og hvilke profesjonelle arenaer som støtter tolkningen av læreplanmålene.

Reliabilitet viser til graden av konsistens på tvers av lærere og skoler. I standpunktvurdering vil reliabiliteten aldri være like høy som ved standardiserte prøver, og noe variasjon er både forventet og legitimt (Parkes, 2013). Likevel forutsetter pålitelig standpunktvurdering at det finnes strukturer som støtter felles forståelse, for eksempel gjennom kollegial kalibrering, felles kjennetegn og profesjonsfaglige arenaer.

Rettferdighet er sentralt i vurderingsarbeidet for å sikre at alle elever behandles likt og vurderes på en likeverdig måte. Dette innebærer å unngå både bevisst og ubevisst forskjellsbehandling og sikre at vurderingen ikke favoriserer eller diskriminerer noen grupper av elever (Moss et al., 2008). Rettferdighet handler også om å sikre at vurderingsprosessen er åpen og tydelig for elevene, slik at de forstår hva som forventes av dem og hvordan de kan påvirke egen utvikling. Rettferdighet rommer også hensynet til elevers ulike forutsetninger, og læreres arbeid med tilpasset undervisning og vurdering.

Håndterbarhet viser til de praktiske og organisatoriske rammene for vurderingsarbeidet. Standpunktvurdering er tidkrevende og forutsetter tilgang til dokumentasjonsverktøy, tid til samarbeid og tilstrekkelige ressurser. Kriteriet er derfor sentralt i å forstå hvilke arbeidsbelastninger og ressursutfordringer lærere møter, og hvordan dette påvirker kvaliteten i arbeidet (Stobart, 2008). Dersom vurderingsarbeidet blir for omfattende, risikerer man at kvaliteten svekkes, enten gjennom overfladisk dokumentasjon eller ved at lærere må gjøre prioriteringer som snevrer inn vurderingsgrunnlaget.

Kriteriene må forstås i sammenheng, ikke som uavhengige størrelser (Messick, 1995; Kane, 2015). For eksempel kan et svært håndterbart system gå på bekostning av validitet dersom vurderingsformene blir for smale, mens et system som søker høy validitet gjennom mange og komplekse vurderingssituasjoner, kan bli lite håndterbart i praksis. I et profesjonsbasert system som det norske, vil spenningen mellom disse kriteriene være særlig tydelig.

I dette prosjektet brukes kvalitetskriteriene som et analytisk rammeverk for å undersøke hvordan lærere og skoleledere forstår og arbeider med kvalitet i standpunktvurdering, og hvordan erfaringer varierer mellom fag, skoler og nivåer.

3.3 Betydningen av fagkompetanse i vurdering

Det er bred enighet blant forskere om at høy faglig og fagdidaktisk kompetanse er avgjørende for læreres vurderingskompetanse (se f.eks. Fjørtoft & Sandvik, 2016; Schildkamp et al., 2020). Lærere med god fagkunnskap har bedre forutsetninger for å identifisere elevers misoppfatninger, tolke vurderingssituasjoner presist og gi faglig relevante tilbakemeldinger (Gottheiner & Siegel, 2012). FIVIS-prosjektet viste at manglende faglig trygghet svekker kvaliteten på vurdering og gjør det vanskelig å sette realistiske mål, planlegge undervisning og gi presise tilbakemeldinger (Sandvik & Buland, 2014).

Skar og Hopfenbeck (2021) peker på at kompetanseutvikling på vurderingsfeltet i stor grad har konsentrert seg om læringsfremmende vurdering. Standpunktvurdering, og spesielt kompetansen som kreves for å fastsette standpunkt karakterer, har fått mindre oppmerksomhet. Dette er en utfordring fordi standpunktvurdering både krever faglig innsikt, fortrolighet med læreplanen og evne til å tolke elevens samlede kompetanse på tvers av situasjoner. Forskningen viser også at læreres behov for kollegial støtte er særlig stort i fag med praktisk tilnærming, komplekse kompetansekrav eller store variasjoner i undervisningspraksiser.

3.4 Samarbeid og profesjonsfellesskap

Internasjonal forskning har vist at elevvurdering har en rekke utfordringer, og at elever ofte er prisgitt sine læreres tolking av kompetansemål og forskrifter (Brookhart et al., 2016). Forskning fremhever samtidig at læreres samarbeid og felles utforskning av praksis er en nøkkelfaktor i skoler som lykkes med elevenes læring (Darling-Hammond, 2017).

Begrepet *profesjonelle læringsfellesskap* har hatt stor innflytelse på forståelsen av hvordan lærere samarbeider til beste for elevenes læring (se f.eks. Stoll, 2006). I Overordnet del i læreplanverket (Kunnskapsdepartementet, 2017, overordnet del, del 3) løftes begrepene *profesjonsfaglig fellesskap* og *profesjonelt samarbeid* frem som sentrale prinsipper for utvikling av skolen. Forskning viser at når lærere deler en kollektiv forståelse av at samarbeid påvirker elevenes

læring, organiserer de og gjennomfører handlinger som faktisk gir bedre læringsutbytte (Goddard et al., 2004; Tschannen-Moran & Barr, 2004).

Skoleledelsen spiller en avgjørende rolle i å utvikle slike fellesskap. Effektiv ledelse handler både om å fremme læreres profesjonelle læring og å delta aktivt i den (Emstad & Birkeland, 2020; Lillejord, 2011). Dette krever faglig og relasjonell kompetanse, og en kultur der ledere og lærere samarbeider systematisk om å utvikle undervisningen. I LK20 understrekes lederens rolle i å legge til rette for læreres læring som en del av utviklingen av profesjonelle fellesskap (Kunnskapsdepartementet, 2017).

Kvalitetsutviklingsutvalget peker i NOU 2023:17 på at systematisk kvalitetsutvikling i skolen forutsetter tydelige strukturer og prosesser for samarbeid og dialog, både internt i skolen og mellom skoler og skoleeier. Utvalget anbefaler blant annet å utvikle mer varige samarbeidsstrukturer som kan støtte profesjonelle fellesskap, bidra til erfaringsdeling og sikre mer felles forståelse av praksis på områder som vurdering og standpunkt (Kunnskapsdepartementet, 2023).

Skar og Hopfenbeck påpeker at det fortsatt mangler klare forventninger og strukturer for hvordan skolene skal utvikle felles forståelse av standpunktvurdering. Mens mange skoleledere tilrettelegger for samarbeid innad på skolen, finnes det i mindre grad arenaer for samarbeid på tvers av skoler. Kunnskapen om hvordan slike samarbeidsarenaer bidrar til å etablere tolkningsfellesskap er begrenset, og forskning peker på behovet for mer systematisk utforskning av forskjeller mellom fag, særlig mellom fag med og uten eksamen.

3.5 Alignment og accountability i vurderingssystemer

Internasjonal forskning på implementering av utdanningspolitikk viser at kvaliteten i vurderingspraksis i stor grad avhenger av samspillet mellom ulike nivåer i utdanningssystemet. To sentrale begreper i denne litteraturen er *alignment* og *accountability* (Coburn et al., 2016; Darling-Hammond, 2017). Begrepene brukes til å analysere hvordan nasjonale føringer, lokale strukturer og profesjonell praksis enten forsterker eller undergraver hverandre. Samlet gir de et analytisk rammeverk for å forstå både stabilitet og variasjon i vurderingspraksis, samt de profesjonelle dilemmaene lærere står i.

3.5.1 Alignment i vurderingssystemer

Alignment viser til graden av *samstemthet* mellom ulike nivåer i utdanningssystemet (Coburn et al., 2016). I et vurderingsperspektiv handler dette om hvordan nasjonale læreplaner, forskrifter og vurderingsprinsipper henger sammen med lokal praksis i skoler og klasserom. Når det er sterk alignment, trekker styringssignaler, lokale strukturer og profesjonell praksis i samme retning, og gir lærere stabile rammer for vurderingsarbeidet. Svak alignment, derimot, åpner for fragmentering, lokale tolkninger og uforutsigbarhet.

På systemnivå handler alignment om hvordan læreplanverk, forskrift til opplæringsloven og nasjonale retningslinjer gir retning for vurderingsarbeidet. I norsk sammenheng legger LK20 vekt på dybdelæring, kompetanseutvikling og profesjonsfellesskap, samtidig som vurderingsforskriften stiller krav om bredt vurderingsgrunnlag og helhetlig vurdering av elevens kompetanse. Disse føringene er i stor grad åpne og profesjonsrettede, noe som gir handlingsrom, men som også forutsetter lokale prosesser for å omsette dem til operasjonell praksis.

På lokalt nivå innebærer alignment at skoleledelse og profesjonsfellesskap utvikler felles forståelser av hva kompetansemålene innebærer, hvordan vurderingskriterier bør utformes, og hvordan karakterskalaen skal tolkes. Coburn et al. (2016) viser at slik samstemthet sjelden oppstår gjennom enkeltstående tiltak, men utvikles over tid gjennom strukturerte samarbeidsarenaer, rutiner for kalibrering og kollektiv meningsskaping. Der slike strukturer mangler, blir vurderingspraksis i større grad overlatt til individuelle tolkninger.

På klasseromsnivå handler alignment om sammenhengen mellom lærerens daglige vurderingsarbeid og de kollektive forståelsene som er utviklet på skole- og systemnivå. Der koblingene er svake, vil lærernes praksis i større grad preges av personlige erfaringer, lokale tradisjoner og situasjonsbetingede hensyn. Det åpner for profesjonell autonomi, men også for utilsiktet variasjon.

3.5.2 Accountability i vurderingssystemer

Accountability viser til mekanismer som holder aktører i utdanningssystemet *ansvarlige* for kvaliteten i opplæringen og vurderingen. Dette omfatter både formelle og uformelle forventninger, kontrollmekanismer og insentiver (Darling-Hammond, 2017). I vurderingssammenheng dreier dette seg om hvordan lærere og skoler opplever krav til dokumentasjon, begrunnelse og etterprøvbarehet i karaktersetting.

I norsk sammenheng kommer accountability til uttrykk gjennom blant annet:

- krav i forskriften om bredt og relevant vurderingsgrunnlag
- skolens interne kvalitetssystemer og rutiner
- eksamen og sensorskolering
- tilsyn og nasjonale kvalitetsvurderingssystemer
- elevers og foresattes forventninger om rettferdighet og innsyn

Når accountability-mekanismene er utformet på en måte som støtter profesjonell praksis, kan de bidra til økt kvalitet, transparens og konsistens. Studier viser for eksempel at sensorskolering og felles tolkningsarbeid rundt eksamen kan styrke vurderingsfellesskap og gi større samsvar i karaktersetting (Federici et al., 2017; Hovdhaugen et al., 2022).

Samtidig viser forskning at sterk vektlegging av dokumentasjon, målbare resultater og ekstern kontroll kan føre til strategisk tilpasning i praksis (Darling-Hammond, 2017). Dette kan gi utslag i fenomen som karakterinflasjon, der lærere i tvilstilfeller systematisk lar tvilen komme eleven til gode, eller der vurderingsformer velges ut fra hva som er lettest å dokumentere fremfor hva som best fanger elevenes kompetanse.

3.5.3 Samspillet mellom alignment og accountability

Alignment og accountability er analytisk ulike, men tett sammenvevde begreper. Alignment handler om samstemthet og sammenheng i systemet, mens accountability handler om ansvarliggjøring, kontroll og legitimitet. I praksis påvirker de hverandre kontinuerlig.

Der det er svak alignment mellom nasjonale føringer, lokale strukturer og klasseromspraksis, vil accountability-mekanismer ofte få økt betydning. Krav til dokumentasjon og etterprøvbarehet blir da et forsøk på å kompensere for manglende felles praksis. Samtidig kan sterk ekstern accountability uten tilstrekkelig alignment føre til at lærere opplever økt press, redusert handlingsrom og svekket profesjonell autonomi.

Det teoretiske perspektivet på alignment og accountability er særlig relevant for å forstå funn fra dette prosjektet fordi det fanger spenninger mellom:

- systemets forventninger og lærernes praktiske muligheter
- nasjonale føringer og lokale vurderingstradisjoner
- profesjonelt skjønn og krav om dokumentasjon
- felles standarder og faglig autonomi

Perspektivet gjør det mulig å analysere både variasjon i vurderingspraksis, opplevde dilemmaer og behov for strukturer som støtter kollektive prosesser. Det gir også et nyttig rammeverk for å forstå hvorfor enkelte fag, som yrkesfag og praktisk-estetiske fag, er mer utsatt for uønsket variasjon enn andre: de har ofte svakere nasjonale føringer, mindre faglig infrastruktur og færre etablerte tolkningsfellesskap på tvers av skoler.

4. METODE

4.1 Datamateriale

For å svare på forskningsspørsmålene har vi samlet inn flere typer data. Vi har intervjuet lærere i de ulike fagene i studien, samt skoleledere. I tillegg har vi samlet inn relevante dokumenter om vurdering. Vi har også gjennomført to spørreundersøkelser.

Under er en oversikt over gjennomførte gruppeintervju. Tabellen viser antall lærere (eller skoleledere) som har blitt intervjuet per skole og fag. Vi har gjennomført 47 intervjuer (varighet 60 minutter per intervju), med i alt 147 deltakere. Det viste seg vanskelig å rekruttere musikk lærere på skolene, fordi det er få lærere som underviser i musikk per skole. Flere av musikk lærerne underviser dessuten musikk på flere skoler, og er derfor lite tilgjengelig for intervju. På grunn av et lite antall musikk lærere på de opprinnelige skolene, valgte vi å inkludere én skole til for å få flere lærerstemmer inn. På Skole 4U intervjuet vi derfor bare musikk lærere. Skole 2VGS hadde i utgangspunktet bare én naturfagslærer, men denne læreren var ikke tilgjengelig i perioden da intervjuene foregikk. Intervjuene med skoleledelse omfatter både rektor/skoleleder og øvrige lederroller med ansvar for pedagogisk utvikling og vurderingsarbeid (for eksempel avdelingsledere/inspektører og tilsvarende funksjoner), og det varierer mellom skolene hvordan slike lederfunksjoner er organisert og benevnt.

Tabell 1 Oversikt over gjennomførte gruppeintervjuer

Skole 1U	Skole 2U	Skole 3U	Skole 4U		Skole 1VGS	Skole 2VGS	Skole 3VGS	Skole 4VGS
3	4	4		Norsk	4	4	4	2
4	3	4		Matematikk	2	3	-	3
4	3	4		Naturfag	4	-	3	2
2	1	2	3	Musikk				
				Helse- og oppvekstfag	5	6	5	3
4	4	2		Kroppsøving	6	4	3	5
4	4	3		Samfunnsfag				
3	3	2		Mat og helse				
2	4	2		Ledelse	6	3	4	4

I tillegg har det blitt samlet inn dokumenter av ulik art, som blant annet lokalt utviklede kjennetegn på måloppnåelse i enkelte fag, lokalt utviklede vurderingsmatriser, planer for vurdering på skolenivå og lignende.

Vi ønsket også å gjennomføre observasjon av karaktermøter i de ulike fagene for å observere hvordan lærere og ledere samtaler og samarbeider om karaktersetting og vurderingsresultater. Da vi kontaktet skolene for å finne ut om slike observasjoner var mulig, viste det seg at ingen av skolene arrangerte slike møter systematisk. Det eneste møtet vi fikk tilgang til, var ett fagmøte i naturfag på en videregående skole.

For å svare på forskningsspørsmålene til de to delstudiene gjennomførte vi i tillegg to spørreundersøkelser for lærere. Formålet med den første undersøkelsen (Undersøkelse 1) var å kartlegge en større gruppe læreres arbeid med standpunktvurdering, både på ungdomstrinnet og i videregående skole. Vi fikk svar fra i alt 792 lærere. Den andre undersøkelsen (undersøkelse 2) bygger videre på hovedundersøkelsen som er presentert i den øvrige kvantitative rapporteringen (se vedlegg 10), men skiller seg fra denne ved å ha et særskilt fokus på press, tillit, karakterinflasjon og styringsrelasjoner i arbeidet med standpunktvurdering. Vi fikk svar fra i alt 1216 lærere på denne undersøkelsen (se vedlegg 11).

4.2 Det kvalitative datamaterialet

Det kvalitative datamaterialet består av semistrukturerte gruppeintervjuer med lærere i åtte fag på sju skoler samt intervjuer med ledelsen på alle skolene. I tillegg inngår feltnotater fra ett fagmøte i naturfag og ulike dokumenter knyttet til skolens vurderingspraksis. Analysearbeidet hadde som mål å identifisere både fagspesifikke og tverrfaglige mønstre i standpunktvurdering, og å knytte disse mønstrene til forskningsspørsmålene og det teoretiske rammeverket for studien.

Gruppeintervjuene med lærere og skoleledelse ble gjennomført som semistrukturerte intervjuer (Kvale & Brinkmann, 2009) basert på intervjuguider. For lærergruppene omfattet guidene spørsmål om bruk av vurderingstekstene, om sammenhengen mellom underveis- og standpunktvurdering, forståelse og bruk av karakterskalaen, hva som vektlegges ved fastsetting av standpunkt og hvilke vurderingsformer som inngår i vurderingsgrunnlaget, bruk av vurderingsressurser som kjennetegn på måloppnåelse og digitale verktøy, elevinvolvering, samarbeid i fagmiljøet og eventuelle moderasjonspraksiser, samt erfaringer med karakterpress og håndtering av press. For skoleledelsen omfattet guidene blant annet spørsmål om hvordan vurderingsarbeidet er organisert og fulgt opp ved skolen, hvilke strukturer som finnes for samarbeid og kvalitetssikring av standpunktvurdering, skoleledelsens rolle i å støtte lærere i vurderingsarbeidet (også ved klage- og konfliktsaker), tilgang til kompetanseutvikling og faglige nettverk, samt ledelsens erfaringer med dokumentasjonsbehov og etterprøvnbarhet i vurdering. Intervjuene hadde en fleksibel struktur, der rekkefølge og oppfølgingsspørsmål ble tilpasset samtalen og lokale forhold, samtidig som guidene sikret temamessig sammenlignbarhet på tvers av fag og skoler.

Transkribert materiale utgjorde grunnlaget for analysen, som ble gjennomført ved hjelp av tematisk analyse (Braun & Clarke, 2006). Denne tilnærmingen ble valgt fordi den gir en fleksibel metode for å identifisere, analysere og rapportere mønstre i kvalitativt materiale knyttet til en spesifikk tematikk. Metoden tillater også en prosess som samtidig er sensitiv for konteksten og variasjonene i dataene.

Første trinn i analysen var å organisere dataene. Dette innebar en grundig gjennomlesing av alle transkripsjoner for å bli kjent med materialet og få oversikt over innholdet. Notater og refleksjoner ble gjort underveis for å identifisere mulige mønstre eller temaer som kunne være relevante for forskningsspørsmålene. Dette arbeidet bidro til å skape en forståelse av det overordnede materialet før mer systematisk koding ble igangsatt.

Deretter ble materialet kodet ved å dele det inn i meningsbærende enheter. Disse enhetene bestod ofte av sitater eller korte beskrivelser som fanget essensen av deltakernes erfaringer og perspektiver. Kodene ble generert med utgangspunkt i både forskningsspørsmålene og

dimensjoner som validitet, reliabilitet, rettferdighet og håndterbarhet, samt begreper fra forskningen om profesjonsfellesskap og alignment, som fungerte som analytiske linser for å fange opp relevante fenomener.

For å ivareta fagenes egenart ble det gjennomført fagvise analyser før tverrgående syntese. To forskere analyserte hvert fag i fellesskap. Denne arbeidsformen gjorde det mulig å identifisere særlige vurderingsutfordringer i hvert fag og å forstå hvordan kompetansebegrepet i læreplanen realiseres fagdidaktisk. Dessuten kunne vi se hvordan ulike vurderingsformer (praktiske, skriftlige, muntlige, prosessuelle) påvirker standpunktvurderingen, og undersøke fagspesifikke dilemmaer knyttet til karaktersetning (f.eks. innsatsens rolle i kroppsøving, praksisvurdering i helse- og oppvekstfag og naturfagets praktiske versus teoretiske komponenter). Analysene i hvert fag ble i første runde skrevet fram i fagrapporter på over tyve sider hver, før de ble syntetisert ned til et format som er tilpasset denne rapporten. Hver fagrapport finnes som vedlegg (1-9).

Etter de fagvise analysene gjennomførte forskergruppen en systematisk sammenligning av temaer på tvers av fag. Slik kunne vi identifisere mønstre som går igjen i flere fag, synliggjøre variasjoner mellom fag, skoler og profesjonsfellesskap, undersøke hvilke praksiser som virker mest sårbare for uønsket variasjon og analysere hvordan fagets egenart former mulighetsrommet for vurdering. Det var gjennom denne delen av analysen at flere av de tverrgående funnene – som karakterpress, variasjon i tolkning av karakterskalaen, usystematisk samarbeid om standpunktvurdering og ulik grad av elevinvolvering – ble tydeligst. Alle temaer ble deretter analysert i lys av kvalitetskriteriene for vurdering og Coburns begreper om alignment og accountability. Dette bidro til å forstå hvordan nasjonale føringer (læreplan, forskrift, kjennetegn), lokale strukturer (profesjonsfellesskap, ledelse, kvalitetssystemer) og individuelle praksiser (lærerens skjønn, tolkning og vurderingsarbeid) samspiller i utviklingen av standpunktvurdering.

Dette perspektivet var særlig nyttig for å identifisere områder der manglende samsvar mellom nivåer skapte tolkningsrom, usikkerhet eller uønsket variasjon.

Avslutningsvis ble fagvise og tverrfaglige analyser integrert i en helhetlig syntese av funn. Her ble det brukt korte eksempler fra fagene for å illustrere variasjon og konkrete eksempler uten at enkeltfag dominerte analysen. Denne syntesen ligger til grunn for funnkapittelet i sluttrapporten.

4.3 Det kvantitative datamaterialet

Den kvantitative delen av studien består av to nasjonale spørreundersøkelser rettet mot lærere i ungdomsskolen og videregående opplæring. Spørreundersøkelsene er utformet for å belyse læreres erfaringer med standpunktvurdering, bruk av vurderingstekster og kjennetegn på måloppnåelse, samt opplevde forventninger og støtte i vurderingsarbeidet. Undersøkelsene fungerer som et supplement til den kvalitative delen av prosjektet ved å gi et bredere bilde av hvordan sentrale problemstillinger i standpunktvurdering kommer til uttrykk på tvers av fag og skoleslag.

4.3.1 Datainnsamling og rekruttering

Spørreundersøkelsene ble distribuert digitalt. Rekrutteringen skjedde ved at prosjektet sendte forespørsel til postmottak ved samtlige ungdomsskoler og videregående skoler i Norge, med anmodning om videresending til skolens lærere. Denne distribusjonsformen innebærer at prosjektet ikke har oversikt over hvor mange lærere som faktisk mottok lenken til undersøkelsen. Det er derfor ikke mulig å beregne svarprosent eller responsrate på en presis måte.

Spørreundersøkelse 1 hadde til sammen 792 respondenter, og spørreundersøkelse 2 hadde 1216 respondenter. Materialet gir et omfattende datagrunnlag, men må tolkes i lys av at deltakelse er basert på selvseleksjon.

Spørreundersøkelse 1 inkluderte flere fag enn prosjektets kjernefag. Hensikten med dette var å etablere et bredere sammenligningsgrunnlag for å forstå vurderingspraksiser på tvers av fag, og dermed sette prosjektets fag i en større kontekst. De fagene som inngår i prosjektets kvalitative del og i de fagvise analysene, utgjør likevel hovedvekten i rapportens drøftinger og konklusjoner.

4.3.2 Spørreskjemautvikling og pilotering

Spørreundersøkelsene ble utviklet på bakgrunn av prosjektets forskningsspørsmål og funn fra den innledende kvalitative datainnsamlingen. Før utsending ble undersøkelsene pilottestet for å sikre tydelighet i spørsmål, relevans for målgruppen og hensiktsmessig svarformat. Piloteringen bidro til justeringer av enkelte formuleringer, rekkefølge og språklig presisjon, slik at påstandene i større grad samsvarte med praksisnære vurderingssituasjoner slik lærere selv beskriver dem.

4.3.3 Svarskala og tolkning av gjennomsnittsskårer

Flertallet av spørsmålene i begge spørreundersøkelsene er formulert som påstander som respondentene vurderer på en sekstrinnsskala (Likert-skala) fra 1 = «helt uenig» til 6 = «helt enig». I funnkapittelet presenteres flere resultater som gjennomsnittsskårer. Disse uttrykker graden av enighet i utvalget, der høyere gjennomsnitt indikerer sterkere enighet og lavere gjennomsnitt indikerer sterkere uenighet. Som en praktisk tolkningsramme kan verdier fra 1 til 3 forstås som overveiende uenighet, mens verdier fra 4 til 6 kan forstås som overveiende enighet. En gjennomsnittsskår rundt 3,5 kan indikere et mer delt eller balansert svarbilde.

I enkelte analyser er svaralternativene også omtalt i form av «enig» og «uenig». Dette bygger på en rekoding av skalaen der 1–3 representerer uenighet og 4–6 representerer enighet. Formålet med denne inndelingen er å gjøre mønstre og hovedtendenser lettere tilgjengelig i en rapport med bred lesergruppe. Vi gjør oppmerksom på at gjennomsnittsskårer og gruppering i «enig/uenig» kan skjule variasjon i svarfordelingen, og der standardavviket er høyt eller svarbildet er delt, er dette kommentert i teksten.

4.3.4 Analyse

De lukkede spørsmålene ble analysert ved hjelp av deskriptiv statistikk (gjennomsnitt og fordelinger) for å gi oversikt over hovedtendenser i materialet. I tillegg ble det gjennomført korrelasjons- og regresjonsanalyser for å undersøke sammenhenger mellom sentrale variabler, for eksempel forholdet mellom opplevde krav, støtte fra skoleledelse, samarbeid i profesjonsfellesskapet og vurderingspraksiser knyttet til standpunkt. De åpne spørsmålene i undersøkelsene ble analysert tematisk og brukt som støtte til fortolkning av de kvantitative mønstrene.

Analysene er i hovedsak brukt for å belyse variasjoner og sammenhenger som kan gi økt forståelse av utfordringer i standpunktvurdering på tvers av fag og skoleslag. Resultatene må leses som uttrykk for læreres opplevelser og vurderinger, og inngår i rapporten som et supplement til de kvalitative funnene.

4.3 Metodiske begrensninger

Selv om metodetilnærmingen i denne studien har sikret en systematisk innsamling og analyse av et omfattende datamateriale, er det flere metodiske begrensninger som må tas i betraktning ved tolkning av funnene. Studien bygger på en kombinasjon av kvalitative intervjuer og kvantitative spørreundersøkelser, og gir et helhetlig grunnlag for å belyse standpunktvurdering fra flere vinkler. Samtidig innebærer både designet og datagrunnlaget avgrensninger knyttet til generaliserbarhet, datakvalitet og hvilke perspektiver som er inkludert.

Studien bygger hovedsakelig på intervjuer og spørreundersøkelser, og beskriver derfor læreres og skolelederes erfaringer, begrunnelser og selvrapporterte praksiser. I prosjektet var det opprinnelig planlagt observasjoner av karaktermøter, men skolene som deltok rapporterte i liten grad at de hadde faste møtearenaer for slike prosesser. Fraværet av observasjonsdata innebærer at rapporten ikke kan gi et direkte bilde av vurderingshandlinger slik de faktisk utspiller seg i praksis, men heller et bilde av hvordan aktørene beskriver og reflekterer over vurderingsarbeidet. Samtidig er dette i seg selv et relevant funn, ved at det peker på at standpunktvurdering i mange tilfeller ikke er institusjonalisert som en formalisert kollektiv prosess ved skolene.

Den kvalitative analysen innebærer en tolkningsprosess der forskernes egne faglige perspektiver kan påvirke kodingen og tematiseringen av materialet. Selv om det er benyttet forskertriangulering og fortløpende drøftinger for å redusere risikoen for skjevheter, kan kodingsprosessen og valg av analytisk ramme likevel ha styrt oppmerksomheten mot enkelte aspekter ved vurderingspraksisen, mens andre temaer kan ha fått mindre plass. Studien har videre stått i et metodisk spenn mellom behovet for bredde og ønsket om dybde, der det har vært krevende å gi like omfattende analyse av alle identifiserte temaer.

Datagrunnlaget innebærer også avgrensninger i hvilke aktørperspektiver som inngår. Studien baserer seg hovedsakelig på intervjuer med lærere og skoleledere, og inkluderer i liten grad andre sentrale aktører som elever og foresatte. Dette innebærer at funnene i hovedsak gir innsikt i profesjonsperspektivet på standpunktvurdering, mens erfaringer og opplevelser fra andre involverte parter ikke belyses i samme grad.

Når det gjelder spørreundersøkelsene, gir materialet verdifull innsikt i læreres erfaringer på tvers av fag og skoleslag, men har også klare begrensninger. Rekrutteringsformen gjør det ikke mulig å beregne responsrate, og utvalget kan derfor ikke uten videre forstås som representativt for lærerpopulasjonen. Dette innebærer at resultatene først og fremst må leses som uttrykk for mønstre og tendenser i et omfattende selvselektivt utvalg, og at generalisering til hele lærerpopulasjonen må gjøres med varsomhet.

I prosjektet var det opprinnelig planlagt observasjoner av karaktermøter som en supplerende datakilde. Skolene som deltok i studien rapporterte imidlertid i liten grad at de hadde faste møtearenaer for slike prosesser. Fraværet av observasjoner innebærer at rapporten i hovedsak bygger på intervjuer og spørreundersøkelser som datakilder til praksisbeskrivelser. Dette er en metodisk avgrensning som må tas med i vurderingen av hvor direkte funnene kan si noe om faktiske vurderingsprosesser og beslutningssituasjoner.

Samlet gir studien et solid kunnskapsgrunnlag for å belyse standpunktvurdering slik den forstås og erfares av lærere og skoleledere, og for å identifisere sentrale utfordringer og dilemmaer i vurderingsarbeidet. Samtidig må funnene tolkes i lys av at materialet i hovedsak bygger på

selvrapportering, at utvalget i spørreundersøkelsene ikke kan vurderes representativt i statistisk forstand, og at enkelte relevante perspektiver og datakilder ikke inngår.

4.4 Valg av fag i studien

Oppdragsgiver ønsket at studien skulle dekke både fag med fagspesifikke vurderingstekster og fag der vurderingstekstene er mer generelle. Fagene som er inkludert i studien representerer derfor et strategisk utvalg som gir bred variasjon i kompetansesyn, vurderingsformer og profesjonelle forutsetninger. Utvalget gjør det mulig å undersøke hvordan standpunktvurdering praktiseres i fag som har ulike tradisjoner, strukturer og utfordringer. Fagene som inngår er norsk, matematikk, naturfag, samfunnsfag, mat og helse, musikk, helse- og oppvekstfag (HO) og kroppsøving. Under beskrives hovedtrekk ved hvert fags egenart med relevans for standpunktvurdering. Fagutvalget er valgt for å gi variasjon i:

- type kompetanse (teoretisk, praktisk, estetisk, yrkesrettet)
- vurderingsformer (skriftlig, muntlig, praktisk, flerkomponent)
- fagtradisjoner og profesjonelle støtteordninger
- kjente vurderingsutfordringer som reliabilitet, tolkning av kompetanse og ressursgrunnlag

Denne variasjonen gir et bredt grunnlag for å analysere hvordan standpunktvurdering forstås og praktiseres på tvers av fag og skoler.

4.4.1 Norsk

Norsk er et gjennomgående fag med lange vurderingstradisjoner og tre separate standpunktkarakterer (hovedmål skriftlig, sidemål skriftlig og muntlig). Standpunktkarakterer i norsk er preget av kjønnsforskjeller: Jenter får stort sett bedre karakterer enn gutter (Blömeke et al., 2019). Sensorreliabiliteten i norsk skriftlig eksamen er stabilt lav (Bjørnson & Skar, 2021), som antyder at det kan være utfordringer i vurdering av norskfaglig skriving også i standpunktkarakterer. En annen utfordring for lærere er at norskfaget har hatt tre karakterer, og at undervisningstid må brukes til å skaffe et grunnlag for summativ vurdering. Forsøk med færre karakterer har vist at lærere bruker den frigjorte tiden til læringsfremmende tilbakemeldinger, og at dette kan åpne for større kreativitet og engasjement; imidlertid viste evalueringen også at elever gjerne vil ha karakterer (Seland et al., 2018).

Andre områder av norskfaget har betydelig mindre empirisk forskning. Læreplanen i norsk har for eksempel lenge inkludert skjønnlitterær lesing, men det finnes lite forskning på læreres vurderingspraksis i skjønnlitterær lesing. Et viktig spørsmål i leseundervisningen på de høyere trinnene er hvordan lærere skal angripe skjønnlitterære tekster, og, i forlengelsen av dette, hvordan elevenes læring skal vurderes. PISA-studiene av leseundervisningen i norskfaget viser for eksempel at svært få lærere leser hele romaner på åttende trinn (Frønes & Roe, 2020), noe som kan påvirke både elevenes læring og lærernes grunnlag for å sette standpunktkarakterer i tiende trinn. Med andre ord er konstruert for vurdering av skjønnlitterær lesing trolig uklart. På samme vis fins det lite forskning på muntlighet. Kaldahls (2020) doktorgradsavhandling viser at norsklærere i ungdomsskolen forstår muntlighetskonstruert på ulikt vis, noe som kan lede til mer urettferdig vurdering. Samlet viser forskningen at det kan være uklart hva slags læringsutbytte norskfaget skal lede til, og at det finnes iboende trusler mot både reliabilitet, validitet og rettferdighet i karaktersettingen. I tillegg vil norskfagets tre karakterer også bidra til å problematisere læreres

arbeidsbyrde og elevenes utbytte av undervisningen sett opp mot læreplanhistoriske og språklig-kulturelle forhold i samfunnet.

4.4.2 Matematikk

Matematikklærere oppgir at matematikk er et fag det er lett å vurdere elevens kompetanse i (Kristiansen, 2017). Dette støttes av undersøkelser som viser at sensorreliabiliteten for matematikkeksamener i både grunn- og videregående skole er høy (Bjørnson, 2020; Bjørnson & Skar, 2021). Årsaken til dette kan være at matematikkfaget historisk sett har vært preget av et formalistisk fagsyn (Morgan, 2016), der matematikk forenklet sett enten er riktig eller feil. Endringene i LK20 innebærer et bredere syn på matematisk kompetanse, med større vekt på resonnering, problemløsning og anvendelse. Dette gjør at vurdering av matematisk kompetanse, krever andre eller mer tilpassede vurderingsverktøy enn tidligere læreplaner. Foreløpig er det forsket lite på vurdering i matematikk etter innføringen av LK20, men en publisert masterstudie viser imidlertid at lærerne som inngikk i studien stort sett gjennomfører vurdering, spesielt summativ vurdering, på samme måte som tidligere (Fjellet & Ryer, 2022).

4.4.3 Naturfag

Naturfag blir ofte vurdert skriftlig, selv om faget også legger opp til at elevene skal få vise kompetanse gjennom muntlige, praktiske og utforskende arbeidsmåter (Utdanningsdirektoratet, 2015). Tradisjonelt har det i naturfag i liten grad vært gjort (formelle) vurderinger av praktiske og utforskende kompetanser i naturfag, men i større grad av produktene av naturfag (Svendsen et al., 2023). Dette gjelder både i grunnskolen og i videregående opplæring. I tråd med LK20 skal praktisk arbeid inkluderes i vurderingen i naturfag. I læreplanen til naturfag står det at «eleven skal få vise kompetansen sin på varierte måter» og «karakteren skal også være basert på kompetansen eleven har vist når eleven har arbeidet praktisk» (Utdanningsdirektoratet, 2020). Faget er modulbasert og dekker mange kompetanseområder, noe som gjør tolkning av kompetansemål og praktisk gjennomføring krevende.

4.4.4 Musikk

Vurdering i musikkfaget har alltid vært en utfordring på grunn av fagets mangfoldige kunnskapsområde og dets tilknytning til individuelle følelser og identitet. Vurdering er et sensitivt tema spesielt når det gjelder vurdering med karakter av elevers musikalske prestasjoner (Constantinou & Bresler, 2010; Ferm Almquist et al., 2017). Tradisjonelt har musikkklærere sett på musikk som en arena der barn og ungdom oppmuntres til å ha det gøy med musikk og lek. Dette har ført til at vurdering ikke alltid har vært en sentral del av lærerens oppgaver. En diskrepans eksisterer mellom den generelle oppfatningen av musikkfagets overordnede mål og læreplanenes mer spesifikke faglige mål, noe som har påvirket synet på vurdering i musikkfaget. En studie om vurdering i ungdomsskolens musikkfag (Vinge, 2014) belyser begrensninger knyttet til tid og vurdering av den brede kompetansen i læreplanen. Studien avdekker også hvordan karakterer fordeles, hvor elever som oppnår karakteren seks ofte har en ekstra dimensjon. Studien peker på de komplekse utfordringene med vurdering i musikkfaget, inkludert behovet for å balansere faglige mål med musikkens kreative dimensjon, samt utvikling av rettfærdige vurderingsmetoder for ulike elever.

4.4.5 Helse- og oppvekstfag (HO)

En generell utfordring i yrkesfag er at lærerne må ha kompetanse både fra et yrke og som lærere. Det har vært krevende å utforme etter- og videreutdanning innen vurdering som er tilpasset norske yrkesfaglæreres behov (Leonardsen, 2021; Sandal, 2023). Lærere på yrkesfag kan dessuten

mangle formell pedagogisk kompetanse (Ekren et al., 2018), noe som påvirker deres forståelse av læreplan og vurdering. Yrkesfaglærere må derfor forholde seg til ulike og til dels motstridende syn på læring og kompetanse, og må ofte forhandle mellom skolens og praksisarenaens syn (Leonardsen et al., 2022). Utdanningsprogrammet *helse- og oppvekstfag* leder til en rekke yrker innen helsesektoren, som helsefagarbeider, ambulansearbeider, helse- eller tannhelsesekretær, og hudpleier. Programmet kan også føre frem til arbeid i barnehage, skole og skolefritidsordning. Det er flest elever som går på yrkesfaglig utdanningsprogram for helse- og oppvekstfag. Skoleåret 21/22 gikk 27 prosent av elevene som går yrkesfaglig vg1 på helse- og oppvekstfag, med 81,9 prosent jenter og 18,1 prosent gutter. Samtidig er det det yrkesfaget hvor færrest får lærlingeplass - kun 69 prosent av elevene fikk lærlingplass i 2022 (Utdanningsdirektoratet, 2022). Lærere i HO må derfor forholde seg til svært ulike typer kompetanse. De må for eksempel både vurdere elevenes kunnskaper om grunnleggende emner som smittekjeder og antibiotika på den ene siden, og deres evne til å være i samspill med pasienter og brukere i ulike institusjoner på den andre (Fjørtoft & Sandvik, 2022). Det er derfor behov for mer kunnskap om standpunktvurdering i HO, når det på den ene siden er det yrkesfaget hvor færrest får lærlingplass, og det på den andre siden utdanner ungdom innenfor yrker det er et stort behov for i samfunnet, og hvor behovet vil øke ytterligere i årene som kommer (Helsedirektoratet, 2022).

4.4.6 Samfunnsfag

Læreplanen i samfunnsfag ble betydelig endret med Fagfornyelsen da de tre tidligere hovedområdene historie, geografi og samfunnskunnskap ikke lenger er konstituerende for faget i grunnskolen. Faget er nå organisert rundt kjerneelementene i faget. Under LK06 hadde faget mange kompetansemål og ble ansett som for omfattende (Sætra & Stray, 2019). I LK20 er det færre kompetansemål i faget. Lærere i samfunnsfag opplever en større autonomi etter innføringen av LK20 (Burner et al., 2024), men de uttrykker også at kompetansemålene er videre og mer abstrakt formulert enn tidligere, noe som kan skape utfordringer i kommunikasjon med elever. Læreplanen for samfunnsfag på ungdomstrinnet går over tre år, med standpunktvurdering etter 10. trinn. Læreplanen slår fast at samfunnsfag har et særlig ansvar for at elevene skal utvikle digitalt medborgerskap.

De tre tverrfaglige temaene, bærekraftig utvikling, demokrati og medborgerskap og folkehelse og livsmestring er også sentrale i samfunnsfaget. Det finnes lite kunnskap nasjonalt om vurdering i samfunnsfag på ungdomstrinnet (Skjæveland, 2020), men de få studiene som finnes, viser at det er store variasjoner mellom skoler både når det gjelder arbeidet med standpunktvurdering og hvilke ferdigheter og kunnskaper som ligger til grunn for karakterene (Dahl et al., 2015). En masteroppgave fra 2021 finner at vurdering i samfunnsfag i stor grad handler om reproduksjon av faktakunnskaper (Davidsen, 2021). Internasjonalt har utdanningsforskere undersøkt hvordan elever kan utvikle lesestrategier og kildekritiske holdninger i samfunnsfaglige spørsmål (Wineburg & McGrew, 2019). Slik forskning er relevant for den norske læreplanen siden Kunnskapsløftet har vektlagt grunnleggende ferdigheter i lesing i alle fag siden 2006, og den har også blitt anvendt av norske fagdidaktikere (Hatlen, 2020). Det er likevel uklart hvordan slike fagdidaktiske tanker preger vurderingspraksis i norsk skole. Et usikkerhetsmoment i vurdering i samfunnsfag er derfor hvordan lærerne forstår og anvender de ulike strukturene og innholdselementene i læreplanverket (kompetansemål, kjerneelementer, grunnleggende ferdigheter, tverrfaglige tema) i arbeidet med standpunktvurderingen.

4.4.7 Mat og helse

Sluttvurdering i mat- og helsefaget gis i form av en skriftlig standpunktkarakter på 10. trinn, eller det steget der faget blir avsluttet. Det er vårt inntrykk at faget på ungdomstrinnet ofte legges til 9. trinn. Mat og helse preges trolig av at mange lærere mangler formell kompetanse i faget. Ifølge SSB har hele 55 % av lærerne null studiepoeng i faget, og få har mer enn 60 studiepoeng, og mat og helse-lærere har også nest lavest andel med lærerutdanning (Stangebye Arnesen et al., 2023). Bottolfs (2020) fant i sin studie at lærere på barne- og ungdomstrinnet med formell kompetanse i faget (30 eller 60 studiepoeng) i større grad underviste i tråd med kompetansemålene, enn lærere uten formell kompetanse. Beinert et al. (2021) har også dokumentert et misforhold mellom læreplanen og praktisk utøvelse av faget blant lærere som underviser i mat og helse på 6. og 9. trinn. Vi registrerer at det er mange klagesaker knyttet til faget. De fagdidaktiske fagmiljøene ved universiteter og høyskoler er små, og antallet faglærere med førstekompetanse er trolig svært begrenset. Gitt at vurderingskompetanse består av både allmennpedagogiske og fagdidaktiske kompetanser, er det viktig å undersøke karaktersetting i dette faget. Våre litteratursøk viser at det er svært lite forskning på vurdering i mat og helse.

4.4.8 Kroppsøving

Praktisk-estetiske fag er blant fagene der vurderingsarbeidet ofte oppleves særlig krevende, og fagområdet er samtidig overrepresentert i klagestatistikken. En viktig ramme for dette er at lærere i praktisk-estetiske fag ofte fastsetter standpunktkarakterer for et stort antall elever, noe som kan gjøre det utfordrende å etablere et tilstrekkelig bredt og variert vurderingsgrunnlag for hver enkelt elev. I tillegg har tidligere undersøkelser pekt på at en relativt stor andel av undervisningen i praktisk-estetiske fag gjennomføres av lærere uten tilstrekkelig fagkompetanse, noe som kan påvirke vurderingssikkerhet og felles forståelser av kjennetegn på måloppnåelse.

Kroppsøvfaget karakteriseres som et av de vanskeligste fagene å fastsette rettferdige standpunktkarakterer i, også av kroppsøvingslærere selv (Prøitz & Borgen, 2010). Det har vært en innarbeidet praksis hos norske kroppsøvingslærere å bruke en tredeling (innsats/holdning, ferdigheter og kunnskap) når elevene vurderes i kroppsøving. Da Kunnskapsløftet ble innført i 2006, ble innsats fjernet som en del av vurderingsgrunnlaget i kroppsøving. Denne endringen bidro til å skape usikkerhet og frustrasjon blant lærere (Prøitz & Borgen, 2010). Lærernes usikkerhet bidro til at læreplanen i kroppsøving ble revidert i 2012, og at innsats igjen skulle telle med i vurderingen.

I LK20 er det presisert at «innsatsen til elevane er ein del av kompetansen i kroppsøving og er derfor ein del av grunnlaget for vurdering». Kroppsøving er dermed det eneste av fellesfagene i skolen der innsats inngår eksplisitt i vurderingsgrunnlaget. Det finnes imidlertid relativt lite kunnskap om hvordan innsats operasjonaliseres i vurderingsarbeidet (Aasland & Engelsrud, 2017). Innsats er noe lærere rapporterer at de «lett kan se», blant annet gjennom at elevene blir «drivende svette», at de forbedrer seg, og at de fremviser en positiv innstilling ved å bidra til å gjøre andre gode. Samtidig problematiseres det at elevs synlige aktivitet kan tillegges stor vekt i vurderingen av innsats, og det stilles spørsmål ved om dette innebærer en risiko for å innskrenke alternative måter å forstå elevs innsats på i kroppsøving.

5. FUNN

Dette kapitlet presenterer funnene fra studien og er organisert etter forskningsspørsmålene. Hensikten er å gi en tydelig og systematisk framstilling av sentrale mønstre i standpunktvurdering på tvers av fag, skoler og lærere. Analysen bygger på kvalitative intervjuer, fagvise rapporter og dokumentgrunnlag fra åtte fag, samt skoleledelse. Spørreundersøkelsene bidrar med et bredere bilde av i hvilken grad erfaringer og praksiser som beskrives i intervjuene også gjenfinnes blant lærere på nasjonalt nivå. Funnene beskriver i hovedsak selvrapporterte praksiser og begrunnelser fra lærere og skoleledere, og må tolkes i lys av dette.

Funnene presenteres tematisk i kapittel 5.1 og 5.2, i tråd med forskningsspørsmål 1 og 2. De fagvise rapportene i vedleggene (vedlegg 1–9) utdyper hvordan funnene kommer til uttrykk innenfor de enkelte fagene, og gir mer utførlige beskrivelser av fagspesifikke vurderingsutfordringer og praksiser.

For å synliggjøre variasjon og konkrete uttrykk for praksis benyttes korte eksempler fra ulike fag. Eksempelene er ikke direkte sitater, men kondenserte beskrivelser basert på mønstre i datamaterialet. De fungerer som illustrasjoner av hvordan praksiser, utfordringer eller valg kommer til uttrykk i bestemte fag. Ikke alle fag representeres i alle underkapitler, og fag trekkes særlig inn der de bidrar med relevante eksempler på de mønstrene som drøftes.

Resultatene fra spørreundersøkelsene presenteres både som gjennomsnittsskårer og som andeler som er «enig» eller «uenig». De fleste påstandene er besvart på en sekstrinnsskala fra 1 = helt uenig til 6 = helt enig. En høy gjennomsnittsskår uttrykker derfor sterkere enighet, mens en lav gjennomsnittsskår uttrykker sterkere uenighet. I enkelte framstillinger er svarene gruppert slik at 1–3 tolkes som «uenig» og 4–6 som «enig». En gjennomsnittsskår rundt 3,5 kan indikere et mer delt svarbilde. Resultatene fra spørreundersøkelsene er deskriptive og bør leses som mønstre og tendenser i materialet, i lys av metodiske begrensninger omtalt i kapittel 4.

I presentasjonen av spørreundersøkelsene brukes både gjennomsnittsskårer og andeler som er enige/uenige. Det er viktig å understreke at gjennomsnitt alene kan skjule variasjon i svarene. I flere påstander er det betydelig spredning mellom lærere, noe som betyr at selv om gjennomsnittet peker i én retning, kan det samtidig være en ikke ubetydelig andel som svarer motsatt. Derfor kommenteres fordeling av svar der dette er særlig relevant, og resultater som framstår delte eller polariserte, omtales som uttrykk for variasjon i praksis og erfaringer, snarere enn som entydige funn.

5.1 Bruk av vurderingstekstene og kjennetegn på måloppnåelse

Dette delkapitlet presenterer funn knyttet til hvordan lærere bruker vurderingstekstene i læreplanen, hvordan de bruker eller utvikler og tilpasser kjennetegn på måloppnåelse, og i hvilken grad disse tekstene inngår i skolens arbeid med kvalitet i standpunktvurderingen. Funnene presenteres samlet på tvers av fag, med fagvise eksempler der dette belyser variasjoner og likhetstrekk.

5.1.1 Bruk av tekstene om underveis- og standpunktvurdering

Lærere i alle de åtte fagene beskriver vurderingstekstene i læreplanen og forskriften som sentrale for å forstå det overordnede rammeverket for vurderingsarbeidet. Intervjuene viser at lærerne

opplever seg godt kjent med tekstene, og at de i hovedsak mener at de skal brukes i vurderingsarbeidet. Spørreundersøkelse 1 bekrefter dette gjennom høye skårer på kjennskap: for eksempel rapporterer lærerne samlet en gjennomsnittsskår på 5,02 for kjennskap til teksten om underveisvurdering og 5,12 for teksten om standpunktvurdering. Dette indikerer at tekstene er godt kjent og oppfattes som legitime styringsdokumenter i profesjonen.

Samtidig tyder funnene på at vurderingstekstene i begrenset grad fungerer som operative verktøy i lærernes daglige vurderingsarbeid. I praksis beskriver lærerne at tekstene hovedsakelig brukes som en overordnet ramme eller retning. Dette støttes av spørreundersøkelsen, der lærerne rapporterer moderat grad av bruk av tekstene i planlegging av undervisning og underveisvurdering. Kompetansemålene får klart høyest skår som praktisk styringsverktøy. Dette kan tolkes som en styrke ved at lærerne opplever læreplanverket som kjent og legitimt, men også som en svakhet i systemet: sentrale tekster om vurdering omsettes i liten grad direkte til praksis, og overlates i stor grad til lokale tolkninger. Slik kan funnene indikere et gap mellom intensjonene i nasjonale føringer og lærernes faktiske vurderingsarbeid.

I samfunnsfag og naturfag beskriver lærere vurderingstekstene som omfattende og bredt formulerte. Flere uttrykker at det kan være krevende å avklare hvilke deler av kompetansemålene som faktisk skal inngå i vurderingen, særlig når arbeidsmengden er stor og tidsrammene trange. I fagteam arbeider de derfor med å «skrelle ned» målene til mer vurderbare dimensjoner, ofte gjennom omfattende nedbryting til konkrete kriterier. Denne praksisen bidrar til lokal forankring og faglig eierskap, men innebærer samtidig at vurderingsgrunnlaget i stor grad formes lokalt og kan variere mellom skoler og fagmiljøer.

I mat og helse beskriver lærere at vurderingstekstene i LK20 og tilhørende begreper om prosesser, ernæring og bærekraft krever omfattende oversettelse til praktiske situasjoner. Lærerne forteller at de bruker tid på å gjøre språket mer konkret og forståelig for seg selv og elevene, og at de ofte må hente støtte fra kolleger eller eksterne ressurser for å sikre en felles forståelse. Denne praksisen kan bidra til faglig utvikling lokalt, men bidrar også til økt belastning og større variasjon i hvordan vurderingskriterier operasjonaliseres. I norsk og matematikk rapporterer lærere at vurderingstekstene i større grad brukes aktivt i planlegging og vurdering, men at også her skjer tolkningen først og fremst lokalt. Læreplanens begreper og formuleringer må operasjonaliseres i fagseksjonen for å kunne brukes i konkret vurdering. Det er dermed fagfellesskapet, snarere enn selve teksten, som i praksis fungerer som det viktigste verktøyet for å omsette nasjonale føringer til konkret vurderingsarbeid.

I helse- og oppvekstfag viser funnene at vurderingstekstene spiller en aktiv rolle i arbeidet med å vurdere elevenes kompetanse i praksis. Samtidig beskriver lærerne et tydelig spenn mellom læreplanens formuleringer og de konkrete situasjonene elevene møter i arbeidslivet. Vurderingsarbeidet kjennetegnes av kontinuerlige avveiiinger mellom forskriftsfestede krav om helhetlig kompetanse og praksisveileders vurderinger, som ofte er situasjonsbetingede og bygger på andre kriterier enn skolens.

Når lærerne blir spurt om hva som eventuelt kunne ha økt bruken av vurderingstekstene, gir spørreundersøkelse 1 et nyansert bilde. For teksten om fagets kjerneelementer heller lærerne samlet mot enighet i at mer tid ($M = 4,10$) og mer konkretisering ($M = 4,03$) kunne ha bidratt til økt bruk. For de øvrige tekstene ligger gjennomsnittene nær eller under midtpunktet på skalaen. Dette indikerer at verken mer tid, økt kunnskap eller ytterligere konkretisering alene trolig vil føre til vesentlig økt bruk av tekstene. Funnene peker derfor i retning av at utfordringen ikke primært

handler om tilgjengelighet eller forståelse, men om at tekstene i dagens form i begrenset grad oppleves som praktisk anvendbare i lærernes vurderingsarbeid.

Samlet viser funnene at vurderingstekstene har høy legitimitet og er godt kjent, men at det er et tydelig gap mellom nasjonale føringer og lokal vurderingspraksis. Dette gapet fylles i stor grad gjennom lokale tolkninger, profesjonsfellesskap og fagspesifikt utviklingsarbeid. Selv om vurderingstekstene brukes i alle fag, viser materialet samtidig at fagenes egenart påvirker hvordan tekstene tolkes og tas i bruk i praksis (jf. oppsummering i del 5.1.4).

5.1.2 Bruk av kjennetegn på måloppnåelse

Både intervjuene og spørreundersøkelsen viser at kjennetegn på måloppnåelse brukes av de fleste lærere, men i ulik grad og ofte i lokal bearbeidet form. Lærerne oppgir at de kjenner til teksten om kjennetegn relativt godt, men at den brukes mindre enn kompetansemålene, særlig i planlegging og undervisvurdering. I spørreundersøkelse 1 oppgir lærerne at de kjenner teksten om kjennetegn på måloppnåelse relativt godt ($M = 4,94$), men at de bruker den mindre enn kompetansemålene, med et gjennomsnitt på 4,46 når de spørres om bruk i planlegging av undervisning. Disse funnene tyder på at kjennetegnene er kjent og vurdert som relevante, men at de ikke brukes som et verktøy i det daglige vurderingsarbeidet.

En tydelig styrke i funnene er at kjennetegnene i praksis brukes som et redskap for å konkretisere kompetanse. I alle fag utvikles det lokale kjennetegn på måloppnåelse, ofte med utgangspunkt i nasjonale føringer, felles maler eller tidligere praksis. Disse lokale kjennetegnene oppleves som mer anvendelige enn de nasjonale tekstene fordi de er tettere knyttet til fagets konkrete oppgaver og undervisningsformer. Særlig der kjennetegnene er utviklet og revidert i fellesskap i fagseksjonene, beskriver lærerne at de gir økt trygghet i vurderingsarbeidet og gjør det lettere å gi presis og konstruktiv tilbakemelding til elevene.

I norsk beskriver flere lærere eksamensveiledninger som mer brukbare i vurdering av skriftlige tekster enn kjennetegn på måloppnåelse. I matematikk brukes kjennetegn lite systematisk. Flere matematikklærere mener at elevenes kompetanse kommer tydeligere til syne gjennom konkrete oppgaver enn gjennom verbale beskrivelser av måloppnåelse. Kjennetegn på måloppnåelse brukes likevel som støtte for å tolke elevenes løsningsstrategier og resonnering, særlig rundt karaktergrensene. Dette kan tolkes som en faglig begrunnet prioritering, men bidrar samtidig til større grad av individuelle vurderingspraksiser.

Funn fra samfunnsfag og naturfag viser også at kjennetegn på måloppnåelse i stor grad må oversettes og operasjonaliseres lokalt for å fungere som støtte i vurderingsarbeidet. I samfunnsfag rapporterer lærerne om god kjennskap til vurderingstekstene og kjennetegn på måloppnåelse, men flere beskriver dokumentet «kjennetegn på måloppnåelse» som vagt formulert og lite egnet å bruke direkte med elever. I praksis opplever lærerne at kompetansemålene i faget er komplekse og lite elevnære, og kjennetegnene gir derfor begrenset støtte uten lokal tilpasning og nedbryting i fagteam.

I naturfag er vurderingsutfordringene ofte knyttet til at faget rommer både teoretiske, praktiske og utforskende kompetanser. Lærerne beskriver at kjennetegn og vurderingskriterier blir særlig viktige som støtte når praktisk arbeid og utforskende aktiviteter skal inngå i vurderingsgrunnlaget. Samtidig viser funnene at også i naturfag varierer bruken av kjennetegn, og at slike verktøy ofte inngår i et lokalt tolkningsarbeid der lærerne forsøker å konkretisere hva kompetansemål betyr i ulike typer aktiviteter.

I alle de tre praktisk-estetiske fagene vi har undersøkt i den kvalitative delen av studien brukes kjennetegn som et hjelpemiddel for å synliggjøre progresjon og mestring i praktiske ferdigheter. Lærerne beskriver samtidig at det kan være krevende å utvikle presise og relevante kjennetegn for praktiske situasjoner. For eksempel forteller kroppsøvlingslærere at de utvikler kjennetegn for ulike aktiviteter, men at det er vanskelig å fange hele bredden i kompetansemålene gjennom slike beskrivelser. Her framstår kjennetegnene både som et støtteverktøy og som en potensiell begrensning.

I HO er det ikke utviklet nasjonale kjennetegn på måloppnåelse, men mange fylkeskommuner har nettverk som har utviklet lokale kjennetegn. Flere skoler har også utviklet egne kjennetegn. I vår studie rapporterer mange lærere at de opplever at lokalt utviklede kjennetegn gir nødvendig struktur, særlig i arbeidet med å vurdere elevenes progresjon. Samtidig beskriver de at kjennetegn kan bli for detaljerte eller for skjematiske i møte med arbeidslivets praksis, og de uttrykker et behov for fleksibilitet i bruk av verktøy når de vurderer elevenes kompetanse i komplekse og uforutsigbare praksissituasjoner. Dette peker på en grunnleggende spenning mellom behovet for standardisering og behovet for profesjonelt skjønn.

Spørreundersøkelsen viser at lærerne bare i begrenset grad mener de ville brukt teksten om kjennetegn mer dersom den var mer relevant, mer konkret eller dersom de hadde mer kunnskap om den. Gjennomsnittsskårene ligger her rundt 3–3,8. Dette tyder på at utfordringen ikke primært handler om manglende kompetanse eller forståelse, men om at bruken av kjennetegn i stor grad formes av lokale behov, tradisjoner og faglige vurderinger snarere enn av selve nasjonale tekstens utforming.

Samlet tyder funnene på at fagenes egenart har betydning for hvordan kjennetegn tas i bruk: I fag med sterke prøvetradisjoner (som norsk og matematikk) brukes kjennetegn mer som støtte for tolkning i grensetilfeller, mens de i fag med mer sammensatt vurderingsgrunnlag (som naturfag, samfunnsfag, praktisk-estetiske fag og yrkesfag) i større grad må konkretiseres lokalt for å fungere som vurderingsverktøy.

5.1.3 Bruk av vurderingstekstene i skolens kvalitetsarbeid

Skolenes arbeid med kvalitetsutvikling knyttet til standpunktvurdering framstår som ujevnt og preget av stor variasjon mellom skoler. Funnene viser at vurderingstekstene inngår i kvalitetsarbeidet ved mange skoler, men at graden av systematikk og struktur varierer betydelig. Spørreundersøkelse 1 viser at lærerne i stor grad opplever at skoleledelsen har vært tydelig på at vurderingstekstene skal brukes som grunnlag for kvalitetssikring av standpunktvurderingen, særlig når det gjelder kompetansemålene ($M = 4,94$) og fagets kjerneelementer ($M = 4,14$). Disse høye skårene kan tolkes som et uttrykk for at skolene formelt har forankret vurderingstekstene i sitt kvalitetsarbeid. Samtidig er skårene betydelig lavere når lærerne vurderer om de har fått konkret veiledning i hvordan tekstene kan brukes i praksis. Her ligger gjennomsnittene mellom 3,28 for veiledning knyttet til standpunktteksten og 3,62 for veiledning knyttet til kompetansemålene. Tallene tyder på at vurderingstekstene i varierende grad inngår i skolens systematiske kvalitetsarbeid. Samtidig viser intervjuene at det er store forskjeller mellom skoler i hvordan vurderingstekstene faktisk brukes og følges opp, fra skoler der dette er tydelig integrert i kollektive prosesser, til skoler der arbeidet i større grad blir overlatt til den enkelte lærer.

Når det gjelder bruken av tekstene i profesjonsfellesskapet, er lærerne tydeligst enige i at kompetansemålene brukes aktivt i kollegiale drøftinger og i arbeidet med å kvalitetssikre standpunktvurdering ($M = 4,86$). Også tekstene om kjerneelementer, undervisvurdering,

standpunktvurdering og kjennetegn inngår i disse diskusjonene, men med noe lavere gjennomsnittsskårer, omkring 3,9–4,1. Dette mønsteret viser at det især er kompetansemålene som fungerer som felles referansepunkt, mens de mer generelle vurderingstekstene i mindre grad preger det kollektive kvalitetsarbeidet.

Intervjuene gir et mer detaljert bilde av hvordan disse tallene kan forstås. Ved skoler med godt etablerte strukturer inngår vurderingstekstene som en integrert del av arbeidet i fagseksjoner og profesjonsfellesskap. Her beskriver lærerne at tekstene brukes i planlegging av undervisning, utvikling av vurderingssituasjoner og konkretisering av kriterier for måloppnåelse. Disse skolene har ofte utviklet rutiner for dokumentasjon, kalibrering og systematisk erfaringsdeling. Et eksempel er naturfagmiljøer der lærerne gjennom regelmessige seksjonsmøter utvikler felles kjennetegn og gjennomfører felles vurderingssituasjoner for å sikre større likhet i praksis. Dette framstår som uttrykk for god praksis der nasjonale føringer er tydelig integrert i skolens kvalitetsarbeid.

Ved andre skoler beskrives kvalitetsarbeidet som mer fragmentert. Her skjer diskusjoner om vurdering mer sporadisk og uten faste strukturer. Vurderingstekstene brukes i mindre grad systematisk, og flere lærere uttrykker at de savner både tid og organisatorisk støtte til felles faglige drøftinger. I slike sammenhenger overlates tolkningen av vurderingstekstene i stor grad til den enkelte lærer eller til små fagmiljøer. Dette bidrar til større grad av lokal variasjon i vurderingspraksis.

I yrkesfagene beskrives det også betydelig variasjon i hvordan vurdering i praksisfeltet inngår i skolens kvalitetsarbeid. Noen skoler har utviklet faste rutiner i samarbeid med praksisfeltet, inkludert felles skjemaer, forventningsavklaringer og jevnlig dialog. Andre skoler overlater i større grad vurderingen til praksisveiledere uten tilsvarende kvalitetssikrende strukturer. Dette kan bidra til ulik praksis og ulik forståelse av hva som regnes som tilfredsstillende måloppnåelse.

Skoleledelsens rolle framstår som en nøkkelfaktor. På skoler der ledelsen følger vurderingsarbeidet tett, deltar i faglige diskusjoner og systematisk legger til rette for felles utviklingsarbeid, beskrives vurderingspraksisen som mer forutsigbar og konsistent. Der ledelsen i større grad begrenser sin rolle til administrativ koordinering, framstår kvalitetsarbeidet som mer personavhengig og mindre forpliktende.

Samlet peker funnene på at vurderingstekstene brukes som utgangspunkt for kvalitetsarbeid ved mange skoler, men at det er stor variasjon i hvordan de faktisk omsettes til praksis. Der arbeidet er strukturert og ledelsesforankret, bidrar tekstene til økt felles forståelse og mer likeverdig vurdering. Der strukturer og støtte mangler, forsterkes variasjonen mellom lærere og mellom fag.

5.1.4 Likheter og forskjeller mellom fag i bruken av vurderingstekstene og kjennetegn på måloppnåelse

Funnene viser både tydelige likhetstrekk og fagspesifikke forskjeller i hvordan vurderingstekstene og kjennetegn på måloppnåelse tas i bruk. På tvers av fag rapporterer lærerne høy kjennskap til tekstene og oppfatter dem som viktige referanser for det overordnede vurderingsarbeidet. Samtidig beskrives tekstene i begrenset grad som operative verktøy som direkte styrer praksis i vurderingshverdagen. Dette innebærer at mye av arbeidet med å gjøre vurderingsgrunnlaget konkret og anvendelig skjer lokalt, gjennom profesjonelt skjønn og kollektive tolkninger.

Materialet tyder samtidig på at fagenes egenart har betydning for hvilken funksjon kjennetegn på måloppnåelse og vurderingstekster får i praksis. I fag med sterke prøvetradisjoner og relativt

standardiserbare kompetanseuttrykk, som norsk og matematikk, brukes kjennetegn ofte mer indirekte, for eksempel som støtte i vurdering av grensetilfeller eller som en overordnet referanse for kvalitet i elevprestasjoner. I fagene samfunnsfag og naturfag, der kompetansen kommer til uttrykk gjennom mer sammensatte og varierte arbeidsmåter, beskriver lærerne i større grad behov for å operasjonalisere kjennetegn lokalt gjennom konkrete kriterier, eksempler og vurderingssituasjoner. Tilsvarende beskriver lærere i praktisk-estetiske fag og yrkesfag at vurderingsgrunnlaget ofte må konkretiseres gjennom lokale tolkninger og felles rammer for kvalitet, nettopp fordi kompetanse uttrykkes gjennom praktiske, skapende eller praksisnære aktiviteter.

Samlet peker funnene på at vurderingstekster og kjennetegn fungerer som overordnede referanser på tvers av fag, men at de får ulik funksjon avhengig av fagets egenart og vurderingsformer: fra å være en bakgrunnsramme i noen fag til å fungere som et mer aktivt verktøy for konkretisering og kommunikasjon av vurderingsgrunnlag i andre.

5.2 Praksis og prosesser i standpunktvurdering

Dette delkapitlet beskriver hvordan lærere forstår og praktiserer sentrale elementer i standpunktvurdering, herunder bruk av karakterskalaen, hva som vektlegges i vurderingsgrunnlaget, sammenhengen mellom underveis- og standpunktvurdering, samarbeid om vurdering og skoleledelsens rolle i kvalitetsarbeidet. De kvantitative dataene gir et samlet bilde av mønstre, mens de kvalitative dataene utdyper hvordan disse mønstrene faktisk utspiller seg i praksis. Samlet viser materialet både områder der praksis er relativt konsistent, og områder preget av tydelig variasjon mellom fag og skoler.

5.2.1 Forholdet mellom underveis- og standpunktvurdering

Lærerne beskriver gjennomgående at underveisvurdering og standpunktvurdering henger tett sammen, og at det er viktig å følge elevenes utvikling over tid. Spørreundersøkelse 1 understøtter dette: lærerne er samlet sett enige i at de vektlegger om eleven har utviklet seg på de områdene de har gitt underveisvurdering i når de fastsetter standpunkt ($M = 4,12$). Samtidig viser materialet variasjon i hvordan denne koblingen forstås og praktiseres mellom fag og skoler. I fag som norsk og matematikk sier mange lærere at de bruker elevenes progresjon som et viktig element ved fastsetting av standpunkt, og at vurderinger tidlig i skoleåret revideres når elever viser forbedring. I andre fag (naturfag, samfunnsfag og HO) beskrives en tydeligere vektlegging av læringsutbytte i perioden nær standpunktsetting. Her opplever flere at vurderingene mot slutten av skoleåret veier tyngst. At lærerne samlet sett er tydeligst enige i påstanden om at elevens kompetanse mot slutten av opplæringen veier tungt ($M = 4,82$), støtter denne forståelsen. Dette kan tolkes som en styrke, fordi standpunktkarakterer i tråd med § 9-16 i Opplæringsforskriften skal være uttrykk for elevens kompetanse ved avslutningen av opplæringen. Samtidig kan det innebære en utfordring dersom vurderingspraksisen blir så sterkt konsentrert om prestasjoner helt mot slutten at vurderingsgrunnlaget blir mindre bredt, og at kompetanse som er vist underveis (og som er relevant for sluttkompetansen) i mindre grad inngår i den samlede vurderingen.

I de praktisk-estetiske fagene beskriver lærerne at underveisvurderingen i stor grad baseres på hyppige observasjoner og dialog i undervisningssituasjonen. Slike løpende vurderinger utgjør et viktig grunnlag for standpunkt, og kan sees som en styrke i form av helhetlig og prosessnær vurdering. Samtidig beskriver lærerne store utfordringer knyttet til dokumentasjon og

systematisering av denne typen vurderingsinformasjon. Dette gjør vurderingsgrunnlaget mer sårbart og mindre etterprøvbart, noe som kan utfordre både transparens og rettssikkerhet.

Flere lærere peker også på økende press fra elever og foresatte knyttet til hvordan enkeltvurderinger påvirker standpunktkarakteren. Elever etterspør i større grad detaljerte forklaringer på hvordan hver vurderingssituasjon «teller» i slutt karakteren. Enkelte lærere beskriver at elever i økende grad argumenterer for forbedring av karakterene ut fra enkeltprestasjoner, snarere enn en helhetlig forståelse av kompetanse. Dette oppleves som krevende for lærerne, og kan bidra til en dreining mot mer målbare og avgrensede vurderingssituasjoner på bekostning av helhetlige vurderingsformer.

Funnene peker på tydelige fagforskjeller i hvordan lærere vektlegger elevenes kompetanse over tid i arbeidet med standpunkt. I norsk og matematikk beskriver lærere ofte at tidligere vurderinger revideres relativt tidlig i skoleåret, og at det tidlig etableres en relativt stabil forståelse av elevenes nivå. Dette kan henge sammen med at fagene har hyppige vurderingssituasjoner, en etablert prøvetradisjon og et vurderingsgrunnlag som i større grad kan sammenlignes på tvers av tid og elever. I slike fag kan vurderingsarbeidet i større grad handle om å justere og stabilisere karakterbildet tidlig, samtidig som senere prestasjoner naturligvis fortsatt tas med i vurderingen.

I samfunnsfag og naturfag framtrer et noe annet mønster. Lærere i disse fagene beskriver oftere at det er elevenes kompetanseutvikling og læringsutbytte i perioden nærmere standpunktsetting som tillegges størst vekt, og at vurderingsgrunnlaget i større grad «bygges opp» utover året. Dette kan blant annet knyttes til at fagene rommer flere kompetanseområder som ikke nødvendigvis dekkes tidlig i skoleåret, og til at vurderingsgrunnlaget kan være mer sammensatt (for eksempel knyttet til utforskende arbeid, muntlige framstillinger, tverrfaglige tema eller tematisk organisering). Samlet kan dette innebære at vurderingspraksisen i større grad blir utviklings- og prosessorientert i samfunnsfag og naturfag, mens norsk og matematikk i større grad legger opp til tidlig stabilisering av vurderingsnivå og justering av tidligere vurderinger. Funnene peker dermed på et mønster der fagklynger skiller seg fra hverandre i hvordan vurderingsgrunnlaget etableres over tid: I noen fag stabiliseres vurderingsbildet tidlig, mens det i andre fag i større grad utvikles gradvis og knyttes tettere til sluttfasen av opplæringsløpet.

Samlet viser funnene at sammenhengen mellom underveis- og standpunktvurdering er godt forankret som prinsipp, men at praktiseringen preges av faglige tradisjoner, tidsnærhet til karaktersetting og økende ytre press. Dette skaper en spenning mellom idealet om helhetlig vurdering over tid og praksiser som i større grad konsentreres rundt avsluttende vurderingssituasjoner.

5.2.2 Forståelse og bruk av karakterskalaen

Både spørreundersøkelsene og intervjuene viser at det er betydelige variasjoner i hvordan skalaen brukes og forstås i praksis. Spørreundersøkelsen viser at lærerne i liten grad opplever at avstanden mellom karaktertrinnene er lik. Påstanden om at avstanden er lik har et samlet gjennomsnitt på 3,4, altså under «verken eller». På tvers av fagene beskrives særlig karakterene 3 og 4 som brede kategorier som rommer stor variasjon i kompetanse.

I flere av fagene beskriver lærere at karakterene 3 og 4 benyttes hyppigst, og at det ofte er krevende å trekke tydelige grenser mellom disse nivåene. Dette framstår både som en faglig utfordring og som en potensiell svakhet, da uklare grenser kan bidra til ulik praksis mellom lærere og skoler. Samtidig beskriver norsklærere at slike utfordringer i enkelte fagmiljøer møtes med

faglige diskusjoner og lokale avklaringer i fagseksjonene. Dette framstår som en styrke ved praksis, da kollektive prosesser bidrar til å redusere vilkårlig variasjon.

Et tydelig funn er at mange lærere benytter et mer finmasket vurderingssystem enn den formelle karakterskalaen tilsier. Flertallet av lærerne oppgir at de bruker pluss- og minusmarkeringer som et internt vurderingsverktøy. Gjennomsnittet på denne påstanden er høyt ($M = 4,74$), med særlig høy bruk blant musikk lærere på videregående ($M = 5,88$). Dette kan tolkes som en styrke ved lærernes vurderingskompetanse, ved at de søker større presisjon i vurderingen. Det kan også tyde på et behov for finere vurderingsnyanser enn hva den formelle karakterskalaen gir rom for. Samtidig innebærer det en svakhet i systemet, ettersom disse nyansene ikke synliggjøres i den offisielle karaktersettingen.

Lærerne rapporterer gjennomgående høy tillit til eget profesjonelt skjønn i karaktersettingen (spørreundersøkelse 2). Påstanden om at de setter høyere standpunktkarakterer nå enn tidligere uten at elevene har blitt faglig sterkere, får lav tilslutning ($M = 2,80$), noe som indikerer at lærerne i liten grad oppfatter egen praksis som preget av ettergivenhet. Samtidig mener mange at kravene for å bestå (karakter 2) er blitt lavere over tid ($M = 4,02$), og at terskelen for de høyeste karakterene også er senket ($M = 3,81$). Dette mønsteret varierer mellom fag og fagklynger, og forskjellene mellom fag er større enn forskjeller knyttet til kjønn og erfaring. Disse funnene støtter intervjuene, der lærerne beskriver spenninger knyttet til bruken av karakterskalaen, samtidig som de fastholder høy tillit til eget faglig skjønn.

I praktisk-estetiske fag, som kroppsøving og musikk, framstår det som særlig krevende å benytte de laveste karakterene. Mange lærere beskriver at karakteren 1 sjelden brukes i praksis. Elevers innsats og deltakelse oppleves ofte som tilstrekkelig til å løfte dem over stryknivå, selv når den faglige kompetansen vurderes som svak. I kroppsøving på videregående skiller dette faget seg ytterligere ut ved at elevenes utvikling fra eget utgangspunkt tillegges større vekt enn i de fleste andre fag ($M = 4,28$, sammenlignet med $3,26$ samlet). Dette kan tolkes som en styrke ved fagets dannelses- og mestringsfokus, men også som en utfordring for likebehandling på tvers av fag.

Tilsvarende mønstre beskrives i mat og helse og i yrkesfag. Den sjeldne bruken av karakteren 1 forklares ofte med fagets praktiske karakter: elever som møter opp og deltar i aktivitetene, vurderes som å ha nådd et minimumsnivå. Dette styrker opplevelsen av mestring hos elevene, men kan samtidig bidra til en utvanning av skalaens differensierende funksjon.

I matematikk beskriver flere lærere at karakterskalaen oppleves som klarere og mer konsistent enn i mange andre fag. Dette framstår som en relativ styrke ved fagets vurderingspraksis. Samtidig rapporteres det variasjon også her, særlig når det gjelder forståelsen av de høyeste karakterene. Lærerne beskriver at kravene til karakterene 5 og 6 kan tolkes ulikt mellom skoler og fagmiljøer, noe som påvirker opplevelsen av nivå og rettferdighet.

På tvers av fagene i det kvalitative materialet beskriver lærere et økende karakterpress fra elever og foresatte. Flere forteller om økte forventninger om innsyn, dokumentasjon og muligheter for å påvirke karakterfastsettingen, samt en økning i klager på standpunktkarakterer. Dette gjelder særlig i norskfaget, men omtales også i kroppsøving, mat og helse og naturfag på ungdomstrinnet. Enkelte lærere beskriver at dette presset påvirker praksis direkte. I tvilstilfeller opplever flere at de oftere lar tvilen komme elevene til gode.

De kvalitative intervjuene og de kvantitative dataene gir et sammensatt bilde av endringer i karakterpraksis. De fleste lærere rapporterer at de ikke selv opplever å sette høyere karakterer

enn tidligere. Samtidig mener en betydelig andel at vurderingspraksisen i fagmiljøene har beveget seg i en mer generøs retning over tid, og at tersklene for bestått og høye karakterer er senket. Dette peker mot at eventuelle endringer i karakterstandarder i stor grad skjer gjennom kollektive faglige normer og lokale vurderingskulturer, snarere enn gjennom individuell ettergivenhet hos den enkelte lærer.

Samlet sett viser funnene at karakterskalaen oppleves som et nødvendig og legitimt verktøy, men at den praktiseres på en måte som både er fagavhengig og kontekstavhengig. Variasjonen peker på en grunnleggende spenning mellom behovet for nasjonal likeverdighet og rommet for profesjonelt skjønn.

5.2.3 Vurderingsgrunnlag og vurderingsformer

Lærerne beskriver standpunktvurdering som en helhetlig vurdering av elevenes samlede kompetanse. Samtidig viser både spørreundersøkelsen og intervjuene at ulike deler av vurderingsgrunnlaget tillegges ulik vekt i praksis. Disse funnene peker på et tydelig spenn mellom idealet om bred, helhetlig vurdering og de prioriteringene som faktisk gjøres i en presset skolehverdag.

Spørreundersøkelse 1 gir et tydelig bilde av hva lærerne samlet sett legger vekt på når standpunktkarakterer fastsettes, og underbygger hovedtendensene fra intervjuene. Samtidig viser svarfordelingen at det er variasjon mellom lærere og fag i hvilke typer vurderingssituasjoner som vektlegges mest. Lærerne er i gjennomsnitt bare svakt enige i at standpunkt først og fremst er et gjennomsnitt av prestasjoner gjennom året ($M = 3,16$), noe som kan tolkes som en styrke ved praksis, ved at lærerne i begrenset grad praktiserer en mekanisk «poengsummering» av arbeid. Samtidig rapporterer de at større vurderingssituasjoner veier tyngre enn mindre oppgaver ($M = 4,15$), og at arbeid gjennomført på skolen tillegges større vekt enn hjemmearbeid og innleveringer ($M = 4,29$). Disse funnene antyder en prioritering av vurderingssituasjoner som lærerne selv kan kontrollere bedre med hensyn til gjennomføring og vurderingsgrunnlag.

Det mest markante funnet viser at nyere arbeid veier tyngst når læreren skal vurdere elevenes samlede kompetanse. Lærerne er svært tydelige på at elevenes kompetanse mot slutten av opplæringen tillegges størst vekt ($M = 4,82$). Dette indikerer en utbredt praksis der nyere prestasjoner får særlig stor betydning. Dette kan tolkes som en styrke ved at standpunkt i stor grad reflekterer elevenes kompetanse ved avslutningen av opplæringen, i tråd med § 9-16 i Opplæringsforskriften. Samtidig kan det innebære en risiko dersom vurderingsgrunnlaget i for stor grad avgrenses til prestasjoner nær standpunktsetting, slik at kompetanse som er vist underveis, og som er relevant for sluttkompetansen, tillegges mindre vekt. I slike tilfeller kan intensjonen om et bredt og variert vurderingsgrunnlag bli vanskeligere å ivareta, selv om sluttkompetansen fortsatt er det overordnede vurderingspunktet. Mønsteret støttes av at lærerne i betydelig grad vektlegger muntlig aktivitet og deltakelse i klasserommet ($M = 4,15$) og utvikling innen områder der det er gitt underveisvurdering ($M = 4,12$). Påstanden om at elevens individuelle utvikling bør veie tyngre enn oppnåelse av læreplanmål får derimot lavere tilslutning ($M = 3,26$). Dette kan tyde på at lærerne i begrenset grad støtter en vurdering der elevens personlige utgangspunkt tillegges større vekt enn kompetansemålene. Samtidig ligger gjennomsnittsskåren nær midtpunktet på skalaen, noe som kan indikere et mer delt syn eller en viss ambivalens: Lærerne kan vektlegge elevers utvikling som del av vurderingsgrunnlaget, men uten at dette forstås som et alternativ til vurdering opp mot kompetansemål.

Intervjuene gir innsikt i hvordan disse mønstrene faktisk utspiller seg i ulike fag. I norsk, matematikk og samfunnsfag beskriver lærerne at vurderingsgrunnlaget formelt sett er bredt og inkluderer både skriftlige og muntlige aktiviteter gjennom året. I praksis forteller de likevel at nyere og «store» vurderingssituasjoner ofte får størst betydning for standpunkt karakteren. Dette kan tolkes som et uttrykk for behovet for et «sikkert» vurderingsgrunnlag tett opp mot karakterfastsettingen.

I naturfag og i praktisk-estetiske fag kombineres skriftlige prøver med praktiske aktiviteter og lærerobservasjoner. Samtidig påvirkes denne balansen tydelig av rammebetingelser. Flere lærere beskriver at det er krevende å dekke alle kompetansemål innenfor tilgjengelig tid og ressurser. Dette gjelder særlig naturfag og samfunnsfag på ungdomstrinnet, og naturfag i videregående opplæring. Lærerne forklarer at enkelte kompetansemål, særlig knyttet til utforskende og praktisk arbeid, i mindre grad inngår i vurderingsgrunnlaget fordi de er tidkrevende å gjennomføre og vanskeligere å dokumentere. Dette bidrar til at tradisjonelle skriftlige prøver i praksis får større vekt enn det læreplanens intensjoner legger opp til. Dette framstår som en tydelig systemutfordring snarere enn et individuelt lærerproblem.

I praktisk-estetiske fag og i mat og helse beskrives vurdering av praktisk kompetanse som særlig krevende. Lærerne peker på at slike vurderingssituasjoner krever omfattende tilrettelegging, og at dokumentasjon av elevenes praktiske ferdigheter er arbeidskrevende. I mat og helse beskriver lærerne at praktiske økter er fagets kjerne, men at vurderingen i stor grad baseres på lærerens observasjoner i situasjonen. Dette gir rom for profesjonelt skjønn og helhetlig vurdering, men oppleves samtidig som krevende å kvalitetssikre og systematisere.

På tvers av fagene beskriver lærerne at framveksten av kunstig intelligens har gjort tradisjonelle skriftlige hjemmeoppgaver mindre egnet som vurderingsgrunnlag. Dette har bidratt til en tydelig dreining mot større vektlegging av prøver, muntlig arbeid og aktiviteter gjennomført i skoletiden. Denne utviklingen kan tolkes som en styrking av vurderingens validitet, ved at lærerne får bedre kontroll over hva som faktisk er elevens eget arbeid, men innebærer samtidig at bredden i vurderingsgrunnlaget kan bli snevrere enn tidligere.

Ressursrammer framstår som en gjennomgående strukturell utfordring. I fag som krever spesialrom, utstyr eller store arealer – som naturfag, kroppsøving og mat og helse – beskriver lærerne at mangel på tid, rom og utstyr direkte påvirker hvilke vurderingsformer de faktisk kan bruke. Dette bidrar til variasjon mellom fag og skoler, og kan begrense muligheten for å realisere læreplanens intensjoner om varierte og autentiske vurderingssituasjoner.

5.2.4 Samarbeid og profesjonsfellesskap i vurderingsarbeidet

Lærerne beskriver gjennomgående at samarbeid om vurdering er sentralt for å utvikle en felles forståelse av kompetanse og karaktersetting. Samtidig viser funnene at omfanget og kvaliteten på samarbeidet varierer betydelig mellom fag, skoler og nivåer. Dette peker mot et gjennomgående spenn mellom profesjonelle idealer og de strukturelle rammene for samarbeid i praksis.

Spørreundersøkelse 1 gir et nyansert bilde av hvordan samarbeidet faktisk foregår. Lærerne rapporterer i stor grad at de samarbeider om å lage oppgaver i fagene (gjennomsnitt $M = 4,34$) og at de har etablerte møter eller arenaer for felles arbeid med vurdering ($M = 4,14$). Dette tyder på at strukturer for samarbeid i stor grad er til stede i skolehverdagen. Samtidig rapporterer lærerne i lavere grad at de vurderer elevbesvarelser i fellesskap ($M = 3,75$). Dette peker mot at de mest direkte kvalitetssikrende samarbeidsformene er mindre utbredt og i større grad overlatt til lokale

initiativ. Påstanden om at samarbeid er viktig for at vurderingen skal være pålitelig og gyldig får svært høy tilslutning ($M = 4,59$), noe som indikerer bred faglig enighet om betydningen av profesjonelt fellesskap, uavhengig av hvor systematisk samarbeidet faktisk er organisert.

Samlet viser tallene et tydelig mønster: Arenaer for samarbeid finnes i stor grad, og lærerne deler en felles forståelse av at samarbeid er viktig, men de mest direkte kvalitetssikrende formene for samarbeid, som felles analyse av elevarbeid, er mindre systematisert og varierer mellom skoler. Dette kan tolkes som en styrke ved profesjonens vurderingsforståelse, men samtidig som en svakhet i de organisatoriske rammene.

Intervjuene utdyper dette bildet. Ved flere skoler foregår samarbeidet i faste seksjons- eller fagmøter, men lærerne beskriver at tiden ofte brukes til praktisk koordinering, deling av undervisningsopplegg, årsplaner eller organisatoriske oppgaver. Arbeid med felles analyse av vurderingsgrunnlag, utvikling av kriterier eller drøfting av konkrete elevprestasjoner får i mindre grad plass. Flere lærere uttrykker et ønske om mer systematisk arbeid med elevtekster, muntlige prestasjoner og praktisk kompetanse, men peker på at det mangler både tid og tydelige strukturer for slikt arbeid.

Også her gir spørreundersøkelse 2 et utfyllende bilde. Lærerne rapporterer i stor grad at de diskuterer standpunktkarakterer med kollegaer før de fastsettes ($M = 4,68$), og at de opplever å ha tilstrekkelig kompetanse til å sette faglig begrunnede standpunktkarakterer ($M = 5,53$).

Videre opplever lærerne i stor grad at skoleledelsen støtter dem i arbeidet med standpunktvurdering ($M = 4,73$). Samtidig er tilslutningen lavere til påstanden om at skoleledelsen følger med på vurderingene for å sikre faglig forsvarlig praksis (lavere gjennomsnitt i materialet), noe som peker mot at støtten i større grad oppleves som situasjonsavhengig enn systematisk. Dette utdyper det kvalitative bildet av et profesjonsfellesskap preget av uformelt samarbeid mellom lærere, men med svakere strukturell forankring fra skolens ledelsesnivå.

I samfunnsfag beskriver lærere at det er få faste arenaer for å drøfte konkrete elevbesvarelser eller utvikle felles tolkningspraksis. Det samme trekket går igjen i flere fag, og samarbeidet i de ulike fagene omtales som «overflatisk» og i stor grad avhengig av enkeltpersoners initiativ. I store fag som norsk og matematikk beskrives samarbeidet som mer strukturert, særlig knyttet til utvikling av vurderingssituasjoner og diskusjoner om karaktergrenser mellom nivåer. Lærerne opplever at dette bidrar til mer lik praksis i karaktersettingen på skolen.

I praktisk-estetiske fag og i mat og helse beskrives fagmiljøene ofte som små, og mange lærere underviser alene eller med få kolleger. Dette begrenser mulighetene for systematisk kalibrering. Lærerne peker på at de i større grad må utvikle kriterier, vurderingsopplegg og praksis på egen hånd, og flere beskriver at dette påvirker tryggheten i vurderingsarbeidet. I slike fag blir behovet for tydeligere støtte fra skoleledelse og organisatoriske strukturer særlig tydelig.

Yrkesfaglærere i helse- og oppvekstfag beskriver at samarbeid ikke bare handler om kollegialt samarbeid på skolen, men også om dialog med praksisfeltet. Flere peker på at vurderingen av elevenes helhetlige kompetanse i praksis varierer mellom praksissteder, noe som gjør det krevende å sikre samsvar mellom skolens vurdering og vurderingen i praksis. Der samarbeidet med praksisfeltet er systematisert, opplever lærerne større forutsigbarhet. Der det mangler faste rutiner, oppleves vurderingsarbeidet som mer sårbart.

På tvers av fag uttrykker mange lærere et ønske om mer tid til felles arbeid med vurdering, særlig i perioder der standpunktkarakterer skal fastsettes. Samtidig beskriver flere at det mangler tydelige,

formaliserte rutiner eller forventninger om slikt arbeid. Samarbeid om vurdering blir dermed i stor grad person- og fagavhengig, og mindre forankret i skolens overordnede strukturer.

5.2.5 Skoleledelsens rolle i kvalitetssikring av standpunktvurdering

Funnene viser at skoleledelsen har en sentral, men varierende rolle i kvalitetssikringen av standpunktvurdering. Både spørreundersøkelsen og intervjuene peker i samme retning: ledelsens prioriteringer, strukturering av tid og tydelighet i forventninger har stor betydning for hvordan vurderingsarbeidet faktisk organiseres og oppleves av lærerne. Dette fremstår som en overordnet styringsfaktor for kvalitet i praksis.

I spørreundersøkelse 2 ligger lærernes vurderinger av ledelsens tilrettelegging rundt midtpunktet på skalaen. Lærerne rapporterer i moderat grad at skoleledelsen legger til rette for samarbeid om standpunktvurdering ($M = 3,57$), arbeider systematisk med utvikling av vurderingskompetanse ($M = 3,62$) og tilrettelegger for kurs eller opplæring i regelverket for vurdering ($M = 3,35$). Skårer rundt midtpunktet indikerer at erfaringene er delte: mens noen lærere opplever tydelig prioritering og støtte fra ledelsen, opplever andre i liten grad systematisk oppfølging. Spredningen i svarene tyder på betydelige forskjeller mellom skoler. Dette viser både en styrke og en svakhet i systemet. Styrken ligger i at enkelte skoler har etablert praksiser som støtter kvalitet i vurderingen. Samtidig innebærer variasjonen at elevers vurderingsgrunnlag i stor grad påvirkes av lokal organisering og prioriteringer fra ledelsen, snarere enn av felles nasjonale rammer.

Intervjuene utdyper dette bildet. På enkelte skoler har ledelsen etablert tydelige strukturer for kvalitetssikring, blant annet gjennom faste møtearenaer for vurderingsarbeid, avsatt tid til profesjonsfaglig samarbeid og klare forventninger til bruk av kjennetegn, vurderingsgrunnlag og dokumentasjon. Lærere som arbeider i slike rammer, beskriver vurderingsarbeidet som mer forutsigbart og mindre personavhengig. De opplever også større grad av trygghet i karaktersettingen. Dette fremstår som et tydelig eksempel på ledelsens potensial til å styrke kvalitet og likeverdighet.

Eksempler fra matematikk viser hvordan ledelsesinitierte strukturer kan påvirke praksis: Ved en skole beskriver lærerne at ledelsen har satt av tid til felles gjennomgang av elevbesvarelser, og at dette har bidratt til en mer felles forståelse av nivåforskjeller, særlig i grensene mellom karakterene 4/5 og 5/6. Slike praksiser peker på hvordan ledelse kan bidra til økt konsistens og rettferdighet i vurderingen. Ved andre skoler beskrives en mer tilbakeholden rolle fra ledelsen. Her overlates vurderingsarbeidet i stor grad til den enkelte lærer eller seksjon, og det finnes få formaliserte arenaer for felles kalibrering. Lærere ved disse skolene beskriver større variasjon i vurderingspraksis mellom grupper og trinn, og en mer personavhengig vurderingskultur. Dette fremstår som en klar svakhet i systemet, da vurderingen i større grad blir sårbar for individuelle forskjeller.

Flere lærere beskriver også at skoleledelsen står i et økende press fra foresatte og elever, særlig i forbindelse med standpunktkarakterer og klagesaker. Dette presset oppleves som tidskrevende og påvirker lærernes vurderingspraksis, blant annet ved at det legger føringer på hvilke vurderingsformer som velges og ved å øke fokus på dokumentasjon. Dette kan tolkes som en styrking av rettssikkerheten for elever, men samtidig som en belastning som forskyver fokus fra pedagogisk vurderingsutvikling til formell dokumentasjon.

Kroppsøvlingslærere trekker særlig frem press knyttet til vurdering av innsats, der elever og foresatte i større grad utfordrer begrunnelsene for karakterer. Dette fører til et opplevd behov for

omfattende dokumentasjon. Flere faglærere peker også på at rammebetingelser som store elevgrupper, begrenset tilgang på egnede rom, utstyr og tid påvirker kvaliteten i vurderingsarbeidet. Dette gjelder særlig i naturfag, mat og helse og kroppsøving. Lærerne beskriver at ledelsen ofte har vilje til å støtte mer systematisk arbeid, men at ressursrammene i praksis begrenser handlingsrommet.

Samlet viser funnene at skoleledelsen har en nøkkelrolle i å redusere personavhengighet og øke kvaliteten i standpunktvurderingen, men at graden av systematikk, prioritering og ressursmessig støtte varierer betydelig mellom skoler. Funnene peker på et tydelig behov for å styrke skolens strukturer for systematisk arbeid med vurderingskompetanse og kalibrering, dersom målet er mer likeverdig og pålitelig standpunktvurdering på tvers av skoler.

5.2.6 Elevinvolvering i vurderingsarbeidet

Lærere beskriver at elevinvolvering er en viktig del av vurderingsarbeidet, men også at praksisen varierer betydelig. I flere fag, særlig norsk, samfunnsfag og matematikk, er det vanlig å gå gjennom vurderingskriteriene med elevene og gi dem innsikt i hva som kjennetegner ulike nivåer. I andre fag, som kroppsøving, musikk og mat og helse, beskrives elevinvolveringen som mer uformell og situasjonsbasert.

Mange lærere peker på at elevene i økende grad etterspør detaljert informasjon om hva som kreves for de ulike karakterene. Flere forteller at elever ønsker å «forhandle» om vurderinger, særlig når det nærmer seg tid for fastsetting av standpunktkarakterer. Lærerne beskriver dette som en effekt av økt karakterpress og sterkere bevissthet om betydningen av karakterer.

Lærere forteller at elever i større grad enn tidligere spør om hvordan hver vurderingssituasjon påvirker standpunktkarakteren og forventer fortløpende karakterprognoser. I praktisk-estetiske fag beskrives elevinvolveringen som utfordrende fordi vurderingen ofte skjer i situasjoner der muntlig dialog og observasjon er sentrale. Flere lærere uttrykker at de ønsker bedre verktøy for å involvere elevene på en mer systematisk måte.

På tvers av fag omtaler lærerne elevinvolvering som en positiv og ønsket del av vurderingsarbeidet, men peker samtidig på at forventningspress fra elever og foresatte gjør det krevende å balansere støtte og autonomi med profesjonelt skjønn.

5.2.7 Press, normendringer og karakterøkning

Et særlig interessant mønster trer fram når de kvantitative og kvalitative funnene ses i sammenheng. Spørreundersøkelse 2 viser at lærerne i liten grad mener at de selv bidrar til karakterinflasjon ($M = 2,80$), og at de i stor grad opplever at standpunktkarakterene de setter gjenspeiler elevenes faktiske kompetanse ($M = 5,39$). Samtidig rapporterer mange lærere at kravene for å bestå (karakter 2) er blitt lavere over tid ($M = 4,02$), og at terskelen for de høyeste karakterene også i noen grad oppleves som senket ($M = 3,81$). Resultatene må samtidig tolkes som uttrykk for læreres opplevelser og vurderinger, og ikke som direkte mål på endringer i elevenes faktiske kompetanse over tid.

Samlet sett kan dette peke mot opplevde endringer i vurderingsstandarder og terskler i ulike fag og skoler. Det er samtidig viktig å understreke at opplevd karakterøkning eller terskelendringer ikke nødvendigvis innebærer karakterinflasjon i streng forstand. Endringer i karakternivå kan ha flere forklaringer, og kan blant annet henge sammen med endringer i elevgrunnlag, opplæringsinnhold, vurderingsformer eller hvordan lærere fortolker kompetansemål og kjennetegn på måloppnåelse over tid. I dette materialet framstår endringene først og fremst som

et fenomen som forstås innenfor profesjonens egne vurderingsrammer, snarere enn som resultat av direkte press eller bevisst ettergivenhet hos enkeltlærere.

Denne tolkningen støttes av fagrapportene, der lærere beskriver fagspesifikke vurderingskulturer og lokale normer for hva som regnes som «tilstrekkelig» og «høy» måloppnåelse. I musikkfaget beskriver lærere at det er «enkelt å skille mellom lav, middels og høy kompetanse», men at karakterene i praksis ofte samler seg rundt midten av skalaen. Flere musikkklærere oppgir at det er «enklere å få fire i musikk enn i andre fag», og at øvre del av skalaen brukes mer aktivt i dette faget enn i mange teoretiske fag. Dette kan tyde på at vurderingsnormer og bruk av karakterskalaen varierer mellom fag, uten at dette nødvendigvis oppleves som «snill» vurdering.

I helse- og oppvekstfag beskriver lærerne et annet uttrykk for slike normer. Her oppleves avstanden mellom karaktertrinnene som ulik, særlig mellom 1 og 2 og mellom 4 og 5, og karakteren 2 bærer et særskilt ansvar fordi den symboliserer minimumskompetanse for yrkesdeltakelse. Dette kan bidra til en praksis der lærere er særlig tilbakeholdne med å gi bestått dersom eleven ikke vurderes å ha den nødvendige minimumskompetansen, uten at dette forstås som streng vurdering, men som en faglig og etisk forankret terskel for yrkesrelevans.

Også i praktisk-estetiske fag som kroppsøving og mat og helse viser datamaterialet tegn til fagspesifikke normer. De kvantitative funnene viser for eksempel at lærere i disse fagene i svært høy grad opplever avstanden mellom karaktertrinnene som lik, mens naturfag og matematikk ligger betydelig lavere på samme påstand (for eksempel naturfag i videregående: $M = 2,89$).

Kvalitative data fra flere fag viser dessuten at lærere ofte tar i bruk uformelle «glidekarakterer» (pluss, minus og skråstrek) for å håndtere grensetilfeller og kommunisere nyanser i elevenes kompetanse. I både musikkklærernes og samfunnsfaglærernes beskrivelser brukes karakterer som 4/3 og 4/5 både som pedagogisk støtte og som et uformelt signal til eleven om at de befinner seg i grenseland mellom to karakterer. Slike praksiser kan samtidig bidra til å etablere og finjustere lokale standarder over tid, uten at dette nødvendigvis tematiseres som avvik fra nasjonale føringer.

Samlet sett peker funnene på at opplevde endringer i karakternivå og terskler kan forstås som et sammensatt fenomen, der lokale vurderingskulturer, fagtradisjoner og profesjonsfellesskap spiller en viktig rolle. Dette gir grunnlag for å drøfte karakterøkning og normendringer som et profesjons- og systemfenomen, samtidig som det er behov for varsomhet med å tolke slike endringer entydig som karakterinflasjon. Et viktig spørsmål for videre arbeid er derfor hva slags former for profesjonelt samarbeid som kan bidra til felles standarder og likere vurderingspraksis, uten å redusere handlingsrommet for fagspesifikt skjønn.

5.2.8 Likheter og forskjeller mellom fag i standpunktvurdering

Funnene i del 5.2 viser både tydelige likhetstrekk og fagspesifikke mønstre i hvordan standpunkt fastsettes. På tvers av fag legger lærerne stor vekt på elevenes kompetanse ved avslutningen av opplæringen, og større, sammensatte vurderingssituasjoner tillegges ofte større betydning enn mindre arbeider. Samtidig viser materialet at fagenes egenart har betydning for hvordan vurderingsgrunnlaget etableres og vektas gjennom året.

Et gjennomgående mønster er at norsk og matematikk i større grad preges av hyppige vurderingssituasjoner og en vurderingspraksis der vurderingsbildet stabiliseres relativt tidlig, og tidligere prestasjoner brukes som aktive referansepunkter. I naturfag og samfunnsfag beskrives vurderingsgrunnlaget oftere som mer sammensatt og gradvis oppbygget gjennom året, og

perioden nærmere standpunktsetting tillegges ofte særlig vekt. I helse- og oppvekstfag framstår vurderingen tett knyttet til praksisnær kompetanse, og lærere beskriver hvordan karakteren 2 får en særlig betydning som terskel for minimumskompetanse.

Funnene peker også på fagspesifikke vurderingskulturer knyttet til bruken av karakterskalaen. I flere fag beskrives skalaen som rommelig og preget av glidende overganger, og praksiser som «glidekarakterer» brukes i daglig vurderingsarbeid. Samlet viser dette at variasjon i standpunktvurdering i stor grad er knyttet til fag og lokale vurderingskulturer, snarere enn til enkeltlæreres individuelle preferanser. En viss variasjon framstår som legitim og knyttet til fagets egenart, men funnene peker samtidig på behov for strukturer for kalibrering og felles standarder der variasjon kan få konsekvenser for rettferdighet og likebehandling.

5.3 Oppsummering av sentrale funn

Funnene i dette kapitlet viser at standpunktvurdering i skolen kjennetegnes av et komplekst samspill mellom profesjonelt skjønn, lokale praksiser og nasjonale føringer. På tvers av fag og skoleslag beskriver lærerne vurderingsarbeidet som både faglig meningsfullt og krevende. Det kvalitative materialet viser stor variasjon i hvordan vurderingsarbeidet organiseres lokalt, hvordan læreplanens føringer omsettes til praksis, og hvordan profesjonsfellesskapene fungerer som støtte i hverdagen. De kvantitative funnene gir et nasjonalt bakteppe som langt på vei underbygger dette bildet, samtidig som funnene viser tydelige variasjoner mellom fag og fagklynger.

I denne studien framstår variasjoner mellom fag og fagklynger som særlig betydningsfulle. Variasjonene henger ofte sammen med fagenes egenart, vurderingsformer og hvordan kompetansemålene lar seg operasjonalisere i praksis. Samtidig peker materialet på at forskjeller i lokale strukturer, som tilgang til samarbeidsarenaer, fagmiljøets størrelse og skoleledelsens prioriteringer, har stor betydning for hvordan vurderingsarbeidet utvikles og kvalitetssikres. Variasjon mellom enkeltlærere kan i større grad forstås som uttrykk for tolkningsrom og profesjonelt skjønn, men denne variasjonen kan bli mer problematisk når den forsterkes av svake kollektive strukturer og mangel på faglig kalibrering. Funnene tyder dermed på at ulikheter i standpunktvurdering i større grad bør forstås som et faglig og organisatorisk fenomen enn som et individuelt ansvar.

For det første viser funnene at det varierer i hvilken grad vurderingstekstene i læreplanen fungerer som arbeidsverktøy i lærernes daglige praksis. Lærere rapporterer høy kjennskap til tekstene, men de brukes primært som overordnede referanserammer. Kompetansemålene fremstår som den mest sentrale og praktisk anvendelige teksten, mens tekstene om underveis- og standpunktvurdering i mindre grad brukes aktivt i planlegging og vurderingsarbeid. Dette bidrar til at oversettelsen fra nasjonale føringer til konkret vurderingspraksis i stor grad skjer lokalt, gjennom faglige tradisjoner og kollektive og individuelle tolkninger. Samtidig tyder fagfunnene på at behovet for lokal operasjonalisering er særlig tydelig i fag med sammensatt vurderingsgrunnlag, som samfunnsfag og naturfag, der lærere beskriver at kjennetegn ofte må konkretiseres for å fungere som støtte i vurderingsarbeidet.

For det andre peker funnene på betydelig variasjon i forståelsen og bruken av karakterskalaen. Lærerne beskriver ikke skalaen som jevnt inndelt, men som rommelig og preget av glidende overganger, særlig på midten av skalaen. Mange benytter uformelle mellomkategorier i det daglige vurderingsarbeidet. De kvantitative funnene viser samtidig et interessant spenn: Lærerne

rapporterer høy tillit til eget profesjonelt skjønn og opplever i liten grad at de selv har blitt mer generøse i vurderingen, men en betydelig andel opplever likevel at tersklene for både bestått og høye karakterer har blitt lavere over tid. Dette peker mot opplevde normendringer og terskelforskyvninger som utvikles innenfor faglige kulturer og kollektive praksiser, og som kan arte seg ulikt mellom fag. I flere fag beskrives også økt press fra elever og foresatte og risiko for klager som en rammebetingelse som påvirker vurderingsarbeidet og kan bidra til mer risikoreduserende praksiser.

For det tredje viser materialet stor variasjon i hva som faktisk vektlegges i standpunktvurderingen. Lærerne legger gjennomgående størst vekt på elevenes kompetanse mot slutten av opplæringsløpet, og større, sammensatte vurderingssituasjoner tillegges større betydning enn mindre arbeider. Samtidig varierer dette mønsteret mellom fag. I norsk og matematikk beskriver flere lærere at tidligere vurderinger revideres relativt tidlig i skoleåret og at vurderingsbildet stabiliseres tidlig, mens lærere i naturfag og samfunnsfag oftere beskriver at vurderingsgrunnlaget bygges opp gjennom året, og at kompetanse vist nærmere standpunktsetting blir særlig viktig. Videre varierer betydningen av muntlig aktivitet, innsats, progresjon og arbeidsmåter mellom fag, og forskjellene framstår som særlig tydelige i praktisk-estetiske fag i studien (musikk, kroppsøving og mat og helse) og i helse- og oppvekstfag. Funnene antyder samtidig en spenning mellom vurderingens faglige bredde og krav til etterprøvnbarhet: I noen tilfeller beskriver lærere at vurderingsformer som er lettere å dokumentere får en mer framtrædende plass, selv når det oppleves at disse ikke fullt ut fanger kompetansens bredde.

For det fjerde viser funnene at samarbeid om vurdering oppleves som viktig for kvaliteten i arbeidet, men at praksisene er ujevnt utviklet. Mange lærere rapporterer samarbeid med kollegaer, men samarbeidet handler ofte om praktisk koordinering framfor systematisk faglig arbeid med vurderingsgrunnlag. Små fagmiljøer beskrives som særlig sårbare, og fravær av kollegial drøfting bidrar til større usikkerhet i vurderingsarbeidet. Spørreundersøkelsen viser at lærerne i større grad opplever støtte fra skoleledelsen i konfliktsituasjoner enn i systematisk, forebyggende arbeid med vurderingskvalitet. Samlet peker funnene derfor mot behov for mer varige strukturer som kan støtte kalibrering og utvikling av felles standarder, både internt på skolen og på tvers av skoler.

Samlet sett viser funnene at standpunktvurdering i norsk skole fortsatt hviler tungt på lærernes profesjonelle skjønn, og at kvaliteten i praksis i stor grad formes av lokale strukturer, faglige kulturer og skolens organisatoriske støtte. Funnene peker på et tydelig spenningsfelt mellom faglig autonomi og behovet for sterkere kollektive strukturer for kalibrering og kvalitetssikring. Variasjonene mellom fag og fagklynger framstår som mer forklaringskraftige enn individuelle kjennetegn ved lærere, noe som understreker betydningen av å forstå standpunktvurdering som en faglig og organisatorisk praksis snarere enn et individuelt ansvar.

6. DISKUSJON

Denne studien viser at standpunktvurdering i norsk skole i stor grad bygger på læreres profesjonelle skjønn, men at kvaliteten i praksis avhenger av hvordan nasjonale føringer, lokale strukturer og individuelle vurderinger henger sammen. Funnene peker ikke på ett enkelt problem, men på et sett av gjennomgående spenninger som har direkte betydning for likeverdighet, legitimitet og gjennomførbarhet i vurderingssystemet. Diskusjonen må også forstås i lys av utviklingstrekk over tid i vurderingssystemet, inkludert økt karakterpress og endringer i vurderingskulturer, som kan påvirke hvordan karakterskalaen praktiseres og hvordan vurderingsstandarter utvikles lokalt. I denne diskusjonen tydeliggjøres noen sentrale dilemmaer i dagens system for standpunktvurdering som har direkte relevans for videre utvikling av systemet. Fremfor å forstå utfordringene som uttrykk for manglende vilje eller kompetanse hos lærere, viser funnene at dilemmaene i stor grad handler om strukturelle rammer og systembetingelser som former lærernes handlingsrom. Diskusjonen bygger på aktørenes beskrivelser av praksis og erfaringer, og sier derfor mest om hvordan vurderingsarbeidet forstås, begrunnes og oppleves i skolen, ikke om alle detaljer i hva som faktisk skjer i vurderingssituasjoner.

For det første peker studien på et misforhold mellom nasjonale føringer og lokale forutsetninger. Læreplanverket forutsetter vurderingspraksiser som fanger kompleks og helhetlig kompetanse, men lokale rammebetingelser som tid, ressurstilgang og organisatorisk kapasitet gjør det vanskelig å realisere disse idealene fullt ut i praksis.

For det andre avdekkes svake koblinger mellom nivåene i vurderingssystemet. Det nasjonale rammeverket gir overordnede prinsipper, men gir begrenset operativ støtte til hvordan disse skal omsettes til konkret praksis på skolenivå. Dette skaper et stort lokalt handlingsrom, som på den ene siden gir rom for tilpasning i fagene, men som på den andre siden kan bidra til varierende praksis og ulik grad av kvalitetssikring.

For det tredje viser materialet en grunnleggende spenning mellom profesjonelt skjønn og økende krav til standardisering og rettsliggjøring. Lærernes skjønn står sentralt i vurderingsarbeidet og er en bærebjelke i det norske vurderingssystemet. Samtidig opplever lærerne et økende press knyttet til dokumentasjon, innsyn og klagehåndtering, noe som kan forskyve praksis mot mer kontrollerbare, men også mer avgrensede og mindre faglig helhetlige vurderingsformer.

For det fjerde peker studien på en spenning knyttet til vurderingsarbeidets omfang og håndterbarhet. Krav om bredt vurderingsgrunnlag, systematikk og dokumentasjon møter tidsmessige og organisatoriske begrensninger. Dette skaper en praktisk spenning mellom det vurderingsfaglig ønskelige og det som faktisk lar seg gjennomføre innenfor skolens rammer.

For det femte synliggjør studien et vedvarende spenningsfelt mellom rettferdighet og samfunnets forventninger. Elever og foresatte har legitime forventninger om rettferdig og etterprøvbar vurdering, samtidig som vurderingsarbeidet bygger på profesjonelt skjønn og faglige tolkninger som ikke alltid er fullt ut transparente. Dette skaper et grunnleggende dilemma mellom faglig begrunnet vurdering og opplevd rettferdighet.

Disse dilemmaene peker samlet mot at videre utvikling av standpunktvurdering ikke kan reduseres til individuelle kompetansetiltak alene. Snarere handler det om å styrke sammenhengen mellom nasjonale ambisjoner, lokale strukturer og lærernes profesjonelle handlingsrom, slik at systemet i større grad støtter den praksisen det forutsetter.

6.1 Spenninger mellom nasjonalt, lokalt og individuelt nivå

Standpunktvurdering er en bærebjelke i det norske skolesystemet. I internasjonal sammenheng skiller Norge seg ut ved å legge hovedansvaret for sluttvurdering til lærerne, fremfor standardiserte, eksterne prøvesystemer. Denne modellen bygger på et grunnleggende prinsipp om at lærere, gjennom sin daglige kontakt med elevene, har best forutsetninger for å vurdere elevenes samlede kompetanse over tid.

Systemet forutsetter høy grad av tillit til læreres profesjonelle skjønn. Samtidig viser funnene i denne studien at dette skjønnnet utøves innenfor et komplekst landskap av nasjonale føringer, lokale strukturer og individuelle fortolkninger. Diskusjonen må også forstås i lys av utviklingstrekk over tid i vurderingssystemet, inkludert økt karakterpress og endringer i vurderingskulturer. Utdanningsdirektoratet viser til at endringer i karakternivå over tid kan ha flere forklaringer, blant annet knyttet til elevgrunnlag, vurderingsformer, faglig innhold og praksiser for fastsetting av karakterer, i tillegg til mer generelle endringer i skole og samfunn (Utdanningsdirektoratet, 2025). Slike endringer innebærer ikke nødvendigvis karakterinflasjon i streng forstand, men aktualiserer spørsmål om hvordan vurderingsstandarter utvikles og vedlikeholdes i profesjonsfellesskap og lokale kontekster.

Vurderingssystemet kan forstås langs en vertikal akse fra nasjonale føringer som læreplaner og sentralgitt eksamen til lokalgitte eksamener og den enkelte lærers beslutninger om standpunktkaraktersetting. Det norske vurderingssystemet kan også forstås horisontalt: Mange føringer for karakterpraksis legges av lokale skolemyndigheter eller av ledere på den enkelte skole. For eksempel kan tilgang til etter- og videreutdanning eller rutiner for klagesaker være svært varierende fra kommune til kommune. Lokalt handlingsrom kan føre til lokal innovasjon og tilpasningsdyktighet, men kan også føre til uheldig variasjon i hvordan karakterer faktisk settes, spesielt hvis koblingene mellom nasjonal styring og lokal tolkningsfrihet er for svake (se Coburn, 2016).

Nyere vurderingsforskning peker også på at karakterpraksiser utvikles i et spenn mellom faglige standarder, lokale normer og krav om begrunnelse og dokumentasjon, og at dette kan bidra til gradvise endringer i hvordan karakterskalaen praktiseres over tid (Gamlem mfl., 2024). Dette understreker at variasjon i standpunktpraksis ikke bare handler om individuelle forskjeller mellom lærere, men også om hvordan normer, vurderingskriterier og praksiser etableres og forhandles i profesjonelle fellesskap. Dette samsvarer med funnene i kapittel 5, der lærerne i liten grad beskriver karakterendringer som et resultat av direkte press eller bevisst ettergivenhet, men heller som uttrykk for fagspesifikke normer og gradvise endringer i terskler og vurderingsstandarter.

Kvalitet i vurdering kan derfor ikke forstås som et spørsmål om den enkelte lærers kompetanse eller praksis alene, men må ses i et systemisk perspektiv der nasjonale føringer, lokale strukturer og profesjonsfellesskap legger viktige rammer for hvordan vurderingsarbeidet gjennomføres (Darling-Hammond, 2017). Samtidig peker tidligere forskning på at læreres fagkompetanse og vurderingsfaglige kompetanse kan ha betydning for kvaliteten i vurderingsarbeidet, selv om dette prosjektet ikke har undersøkt kompetanse som et eget tema. Et hensiktsmessig utgangspunkt for å utvikle standpunktkaraktersystemet er derfor å tenke i tre horisontale nivå:

- *Det nasjonale nivået* definerer rammene for vurdering gjennom læreplan, forskrift, vurderingstekster og kvalitetskriterier. Disse dokumentene uttrykker normative ideal om bred kompetanse, helhetlige vurderingsformer og et samspill mellom underveis- og standpunktvurdering.

- *Det lokale nivået* med skoleledelse, profesjonsfelleskap, samarbeidskultur og kvalitetssystemer oversetter og operasjonaliserer de nasjonale føringene. Forskning viser at styrken i denne oversettelsen er avgjørende for kvalitet og konsistens i vurderingsarbeidet (Schildkamp et al., 2020).
- *Det individuelle nivået* utgjør lærerens faglige kompetanse, praktiske dømmekraft, faglige tolkningsrammer og relasjonelle arbeid. Læreres skjønnsutøvelse er anerkjent som en nødvendig del av vurderingssystemet (Hovdhaugen et al., 2018), men blir samtidig påvirket av press fra både strukturelle og kulturelle faktorer.

Denne studien viser at koblingene mellom nivåene ikke alltid er like sterke. Mens det nasjonale nivået uttrykker tydelige ambisjoner om bred kompetanse, helhetlig vurdering og profesjonsfaglig samarbeid, varierer de organisatoriske og kulturelle rammene på lokalt nivå betydelig. Dette gir rom for både faglig tilpasning og lokal innovasjon, men åpner også for utilsiktet variasjon i hvordan standpunktvurdering forstås og praktiseres.

I et slikt system blir kvalitet i vurdering et kollektivt, ikke individuelt, anliggende. Kvalitet oppstår i samspillet mellom:

- nasjonale rammer og forventninger,
- lokale strukturer og kulturer,
- og lærernes profesjonelle dømmekraft.

Funnene viser at når koblingene mellom disse nivåene er svake, blir vurderingspraksis mer personavhengig, og risikoen for uønsket variasjon øker. Samtidig viser materialet at der skolene har tydelige strukturer, faste samarbeidsarenaer og en aktiv rolle for skoleledelsen, opplever lærerne større trygghet og mer lik praksis i vurderingsarbeidet.

6.2 Spenningen mellom nasjonale intensjoner og lokale praksiser

LK20 definerer et bredt og komplekst kompetansebegrep som skal ligge til grunn for både underveis- og standpunktvurdering. Dybdelæring, sammenheng mellom kunnskap og ferdigheter og elevenes evne til refleksjon og overføring, vektlegges. Standpunktvurdering skal i dette perspektivet speile elevenes samlede kompetanse ved avslutningen av opplæringen, basert på et bredt vurderingsgrunnlag. Evalueringen av fagfornyelsen peker samtidig på at det kan være krevende å realisere læreplanens intensjoner i praksis, særlig i praktiske og estetiske fag. Her beskrives det høye forventninger til at fagene skal bidra med andre perspektiver og læringsformer, samtidig som det rapporteres en viss treghet i å omsette intensjonene til praksis (Borgen mfl., 2023a; 2023b).

Denne studien viser at det ofte oppstår et gap mellom de nasjonale intensjonene i lov og rammeverk og de lokale realitetene. Lærerne beskriver et krysspress mellom idealene om bred kompetansevurdering og rammebetingelser som tid, gruppestørrelser, opplevde krav til dokumentasjon og etterprøvbarehet og tilgang til ressurser. I slike situasjoner blir vurderingsarbeidet i praksis en prioritering mellom det som er vurderingsfaglig ønskelig og det som er organisatorisk gjennomførbart. I tillegg viser nyere forskning at vurderings- og karakterpraksiser utvikles i et spenn mellom faglige standarder, lokale normer og krav om begrunnelse og dokumentasjon, noe som kan bidra til at nasjonale ideal om vurdering omsettes ulikt i lokale kontekster (Gamlem mfl., 2024).

Funnene i kapittel 5 viser at vurderingstekstene i læreplanverket i begrenset grad fungerer som konkrete arbeidsverktøy i lærernes daglige vurderingsarbeid, selv om lærerne i stor grad kjenner til dem. Dette kan tyde på at sentrale begreper som «samlet kompetanse», «bredt vurderingsgrunnlag» og «relevante vurderingssituasjoner» i stor grad operasjonaliseres lokalt, gjennom fagtradisjoner og profesjonelle tolkningsfellesskap. På den bakgrunn kan det være hensiktsmessig å vurdere om vurderingstekstene bør tydeliggjøres eller videreutvikles, særlig der intensjonene i LK20 forutsetter komplekse vurderinger. Samtidig peker funnene på at tekstendringer alene trolig ikke er tilstrekkelig, og at nytten av eventuelle revisjoner vil avhenge av hvordan de inngår i lokale støttestrukturer og profesjonsfellesskap.

Denne spenningen kan få betydning for vurderingens validitet. Når tids- og ressursrammer begrenser muligheten til å variere vurderingsformer, risikerer man at vurderingsgrunnlaget snevres inn til det som er mest effektivt å gjennomføre og dokumentere (Stobart, 2008). Dette kan føre til en forskyvning fra helhetlig kompetansevurdering til mer fragmenterte og instrumentelle vurderingsformer.

Samtidig viser funnene at dette ikke utelukkende kan forstås som manglende vilje eller kompetanse hos lærerne. Lokale strukturelle forhold spiller en avgjørende rolle: tilgang til spesialrom, utstyr, lærerkompetanse og tid til samarbeid påvirker direkte hvilke vurderingssituasjoner som faktisk lar seg realisere. Dette gjelder særlig praksisorienterte fag, der vurdering av praktiske ferdigheter krever mer omfattende tilrettelegging og dokumentasjon. I yrkesfag kan vurderingsarbeidet også påvirkes av strukturelle forhold i opplæringsløpet, særlig med tanke på overgangen til opplæring i bedrift og variasjon i vurderingssituasjoner mellom skole og læretid (Skålholt mfl., 2023; Skålholt mfl., 2024).

Spenningen mellom nasjonale intensjoner og lokale praksiser kan derfor forstås som et strukturelt dilemma snarere enn et individuelt problem. På den ene siden ønsker systemet vurdering som fanger kompleks kompetanse; på den andre siden stiller ikke alltid de organisatoriske rammene tilstrekkelig støtte til rådighet. Resultatet er at lærere i mange tilfeller må gjøre pragmatiske valg som kan svekke samsvaret mellom læreplanens intensjoner og vurderingspraksisens faktiske utforming.

Dette illustrerer at forbedring av standpunktvurdering i liten grad kan reduseres til spørsmål om individuelle ferdigheter, og i større grad handler om å styrke de strukturelle forutsetningene som gjør det mulig å praktisere vurdering i tråd med nasjonale mål. Dette samsvarer med analyser fra Utdanningsdirektoratet som peker på at endringer i vurderingspraksis og resultater må forstås i lys av systemfaktorer og endringer i rammebetingelser, ikke bare enkeltlæreres vurderinger (Utdanningsdirektoratet, 2025).

6.3 Profesjonelt skjønn i møte med ansvarliggjøring

Lærernes profesjonelle skjønn står sentralt i standpunktvurderingen. Samtidig viser funnene at dette skjønnet utøves i et landskap preget av økende opplevde krav til dokumentasjon, etterprøvnbarhet og eksplisitte begrunnelser. Det er viktig å presisere at regelverket i dag ikke stiller et generelt krav om dokumentasjon knyttet til underveisvurdering. Behovet for dokumentasjon aktualiseres likevel ofte i standpunktvurdering og ved klagesaker, der krav til begrunnelse og etterprøvnbarhet blir særlig tydelige, og der lokale rutiner fra skoleledelse eller skoleeier kan forsterke forventninger om skriftlige spor. Samlet kan dette bidra til at

dokumentasjon får en sentral plass i vurderingsarbeidet, også uten at det følger direkte av nasjonalt regelverk.

Nyere forskning viser at læreres vurderings- og karakterpraksiser utvikles i et spenn mellom faglige standarder, lokale normer og krav om begrunnelse og dokumentasjon, og at dette kan bidra til endringer i hvordan vurderingskriterier og karakterskala faktisk praktiseres over tid (Gamlem mfl., 2024). Dette understreker at dokumentasjonskrav ikke bare handler om administrative rutiner, men også om hvordan profesjonelle normer, forventninger og vurderingsstandarder utvikles i lokale fellesskap.

Disse forventningene skaper et grunnleggende profesjonelt dilemma som ikke kan reduseres til et spørsmål om «riktig» eller «gal» praksis, men som handler om balansering av legitime hensyn. På den ene siden er det både legitimt og nødvendig at vurderingssystemet stiller krav til sporbarhet, transparens og begrunnelser. Slike krav er forankret i elevers rettssikkerhet, i samfunnets behov for tillit til karaktersettingen og i idealet om rettferdighet i vurdering. Dette kan forstås som en form for nødvendig risikostyring: systemet må ha mekanismer som gjør vurderinger etterprøvbare og reduserer vilkårlighet.

På den andre siden viser både denne studien og tidligere forskning at et sterkt fokus på dokumentasjon kan få utilsiktede konsekvenser. Når lærere opplever at risikoen for klager og konflikter er høy, kan det oppstå en dreining i praksis der vurderingsformer velges ut fra hvor lett de lar seg dokumentere, snarere enn ut fra hva som best fanger elevenes faktiske kompetanse. Dette samsvarer med Moss med flere (2008) sin beskrivelse av en forskyvning fra læringsorientert vurdering til risikoreduserende praksiser.

Spenningen handler derfor ikke om at dokumentasjon i seg selv er uønsket, men om forholdet mellom ulike kvalitetskriterier. Når reliabilitet og etterprøvbarhet gis forrang, kan det gå på bekostning av validitet (om vurderingen faktisk fanger faglig bredde) og rettferdighet i betydningen meningsfull og situert vurdering. Dette gjelder særlig i fag der kompetansen i stor grad er praktisk, estetisk eller relasjonell, og vanskelig lar seg fullt ut fange i standardiserbare formater. Evalueringen av fagfornyelsen peker også på at praktiske og estetiske fag i LK20 kjennetegnes av skapende prosesser og sammensatte kompetanseuttrykk, noe som kan gjøre vurderingsgrunnlag og tolkning av kriterier særskilt krevende (Borgen mfl., 2023a; 2023b).

Det som trer fram er dermed ikke et enten–eller mellom profesjonelt skjønn og ansvarliggjøring, men et reelt og legitimt profesjonelt dilemma. Et bærekraftig system for standpunktvurdering forutsetter både høy grad av tillit til lærernes skønnsutøvelse og strukturer som ivaretar elevers rettssikkerhet og samfunnets behov for innsyn. Utfordringen ligger i å utvikle former for ansvarliggjøring som støtter profesjonelt skjønn, og som balanserer hensynene til validitet, reliabilitet, rettferdighet og håndterbarhet.

6.4 Faglig autonomi og behovet for felles standarder

Variasjon i vurderingspraksis er et gjennomgående funn i denne og tidligere studier (Federici et al., 2017; Skar & Hopfenbeck, 2021). En viss grad av variasjon er både uunngåelig og faglig legitim, gitt fagenes ulike egenart, arbeidsmåter og kompetansekrav. Samtidig blir variasjonen problematisk når den ikke springer ut av faglige hensyn, men av fravær av kollektive strukturer for kalibrering og felles forståelse.

Spenningen mellom autonomi og standardisering kan forstås som en avveining mellom to sentrale kvalitetskriterier: På den ene siden styrker faglig autonomi validiteten ved at lærere kan tilpasse vurderingene til fagets innhold og arbeidsformer. På den andre siden utfordrer manglende felles standarder reliabiliteten, ved at like elevprestasjoner kan bli vurdert ulikt på tvers av lærere, klasser og skoler.

Denne utfordringen fremstår særlig tydelig i fag uten nasjonale eksamener eller etablerte nasjonale tolkningsfellesskap, som kroppsøving, musikk, mat og helse og enkelte yrkesfag. I disse fagene er vurderingen ofte tett knyttet til praktisk utøvelse, situasjonsbundne ferdigheter og løpende observasjon. Slike kjennetegn styrker faglig relevans og validitet, men gjør samtidig vurderingene mer sårbare for ulik praksis og svakere reliabilitet. Evalueringen av fagfornyelsen peker også på at praktiske og estetiske fag ofte oppleves som «annerledes»-fag, preget av skapende prosesser, sosialt samspill og fysisk aktivitet, og at dette kan gi særskilte utfordringer knyttet til vurderingsgrunnlag og tolkning av læreplaner (Borgen mfl., 2023a; 2023b).

Forskning viser at vurderingskvalitet styrkes når lærere inngår i profesjonsfellesskap der de utvikler felles forståelse av kjennetegn på måloppnåelse og jevnlig drøfter konkrete elevprestasjoner (Brookhart et al., 2016). Våre funn viser imidlertid at tilgang til slike fellesskap varierer betydelig både mellom og innenfor skoler. Der fagmiljøene er små eller fragmenterte, blir kvalitetssikringen i større grad overlatt til den enkelte lærer.

Dette får også konsekvenser for rettferdighet. Uten felles kalibrering er det vanskeligere å sikre at elever vurderes likt på tvers av grupper og skoler. Samtidig kan overdrevent sterke standardiseringskrav svekke fagets egenart ved å presse komplekse ferdigheter inn i forenklete vurderingsformer. Det fremtrer dermed et grunnleggende dilemma: Et standpunktvurderingssystem trenger både faglig autonomi for å sikre validitet og faglig relevans, og kollektive strukturer for å sikre reliabilitet og rettferdighet. Vår studie viser at lokale profesjonsfellesskap er den mest avgjørende arenaen for å balansere dette forholdet. Behovet for mer systematiske strukturer for samarbeid og felles praksisutvikling løftes også i Kvalitetsutviklingsutvalgets sluttrapport, som peker på profesjonelle fellesskap som en forutsetning for kvalitetsutvikling i skolen (NOU 2023:17).

6.5 Håndterbarhet som forutsetning for vurderingskvalitet

Håndterbarhet fremstår i denne studien som et grunnleggende, men ofte undervurdert, kvalitetskriterium. Uten tilstrekkelig tid og ressurser svekkes forutsetningene for alle de øvrige kvalitetsdimensjonene: validitet, reliabilitet og rettferdighet. Flere studier viser at læreres arbeidsbelastning og rammebetingelser har direkte betydning for valg av vurderingsformer (Beinert et al., 2021; Stobart, 2008).

Våre funn bekrefter at lærere i praksis må gjøre løpende avveininger mellom faglig ideal og organisatorisk realitet. Når tiden er knapp, blir vurderingsformer valgt ut fra hva som er mulig å gjennomføre og dokumentere, snarere enn hva som best fanger bredden i elevenes kompetanse. Dette kan svekke validiteten ved at deler av kompetansebegrepet blir underrepresentert.

Utfordringen forsterkes av fagenes egenart. I fag som naturfag, kroppsøving og mat og helse krever vurdering tilgang til spesialrom, utstyr og ofte ekstra bemanning. I yrkesfag er vurderingsarbeidet tett knyttet til praksisfelt og samarbeid med eksterne aktører. Evalueringen av fagfornyelsen i yrkesfag peker på at overgangen til opplæring i bedrift kan innebære varierende

rammer og ulike vurderingssituasjoner, og at dette kan gi utfordringer for sammenheng og systematikk i vurderingsarbeidet (Skålholt mfl., 2023; Skålholt mfl., 2024). Disse fagene er derfor særlig sårbare for manglende ressurser. Når ressursene er begrensede, blir vurderingsgrunnlaget ofte snevrere og mer standardisert enn det læreplanens intensjoner tilsier.

Samtidig påvirker håndterbarheten også reliabiliteten. Kollegial kalibrering, felles analyse av elevarbeid og utvikling av felles kriterier er tidkrevende prosesser. Når slike arenaer ikke prioriteres eller ikke får tilstrekkelig tid, blir vurderingsarbeidet mer opp til den enkelte lærer, og risikoen for ulik praksis øker. Behovet for systematiske strukturer som gir rom for slikt samarbeid løftes også i Kvalitetsutviklingsutvalgets sluttrapport (NOU 2023:17). Håndterbarhet har også en klar kobling til rettferdighet. Dersom lærere ikke har tid til grundige begrunnelser, dialog med elever eller systematisk dokumentasjon, kan elevenes opplevelse av rettferdighet svekkes, også når den faglige vurderingen i utgangspunktet er forsvarlig.

Det fremkommer derfor et tydelig mønster. Håndterbarhet handler ikke bare om praktiske hensyn, men er en forutsetning for kvalitet. Når systemet legger høye forventninger til et bredt og variert vurderingsgrunnlag og til samarbeid og kalibrering, uten å sikre tilsvarende ressurser til arbeidet, oppstår det strukturelle spenninger mellom idealer og realiteter. Dette peker mot at forbedring av standpunktvurdering ikke primært handler om å stille strengere krav til lærere, men om å utvikle organisatoriske rammer som gjør kvalitetskravene realistisk gjennomførbare.

6.6 Rettferdighet og legitime forventninger fra samfunnet

Rettferdig vurdering er et grunnleggende normativt ideal både i norsk regelverk og i samfunnets forventninger til skolen. Rammeverket for vurdering legger vekt på likeverdighet, saklighet og at vurdering ikke skal være diskriminerende eller vilkårlig. Elever og foresatte har dermed en legitim forventning om at standpunkt karakterer er rettferdige, faglig begrunnede og konsistente. Slike forventninger er ikke problematiske i seg selv, de utgjør tvert imot en sentral forutsetning for vurderingens legitimitet.

Samtidig viser funnene i denne studien at det eksisterer et spenn mellom *legitime forventninger* og det som kan betegnes som mer urealistiske forventninger til standpunktvurdering. Mange elever og foresatte har begrenset innsikt i kompleksiteten i kompetansebegrepet, i det profesjonelle skjønnets rolle og i den usikkerheten som uunngåelig følger profesjonelle vurderinger. Dette skaper et vedvarende profesjonelt dilemma for lærere, særlig når karakterer får stor betydning for videre utdanning og livsmuligheter. I et slikt landskap kan også endringer i karakternivå over tid bidra til økt oppmerksomhet og forventningspress. Samtidig er det viktig å understreke at karakterutvikling kan ha flere forklaringer og ikke uten videre kan forstås som uttrykk for karakterinflasjon, men må ses i sammenheng med endringer i elevgrunnlag, vurderingsformer og praksiser for karakterfastsetting (Utdanningsdirektoratet, 2025).

Her er det nyttig å skille mellom ulike dimensjoner ved rettferdighet, slik nyere vurderingsteori gjør (Nisbet & Shaw, 2020). Rettferdighet handler ikke bare om teknisk kvalitet (reliabilitet og validitet), men også om opplevd rettferdighet hos elever og foresatte, og om vurderingens sosiale legitimitet i samfunnet. Disse dimensjonene sammenfaller ikke nødvendigvis.

Sett i lys av kvalitetskriteriene trer særlig rettferdighet og reliabilitet fram som et spenningsfelt. På den ene siden skal vurderingen være rettferdig i faglig forstand, basert på relevant og bred dokumentasjon over tid. På den andre siden skal den være konsistent og sammenlignbar på tvers

av elever, grupper og skoler. Samtidig skal den også fremstå som rimelig og forståelig for dem som berøres. Dette er tre delvis motstridende krav.

Dette dilemmaet blir særlig tydelig i fag der vurderingen er prosessuell, situasjonsavhengig og vanskelig å dokumentere fullt ut, som i kroppsøving, musikk, mat og helse og flere yrkesfag. I slike fag kan læreren ha et faglig forsvarlig vurderingsgrunnlag, samtidig som vurderingen fremstår lite transparent for elever og foresatte. Da utfordres den opplevde rettferdigheten, selv om vurderingen er faglig legitim. Læreren plasseres derfor i et krevende profesjonelt spenn mellom faglig integritet og relasjonell hensyntaken (Birkelund & Midthaugen, 2019; Fjørtoft & Morud, 2021). Læreren må samtidig ivareta tre former for rettferdighet: faglig, prosessuell og relasjonell.

Et grunnleggende dilemma trer dermed fram: Standpunktvurdering må både være teknisk forsvarlig og sosialt legitim, men det er ikke gitt at en vurdering som er faglig korrekt, også oppleves som rettferdig. Dette er et sentralt poeng hos Nisbet og Shaw (2020), som viser at rettferdighet i vurdering alltid må forstås i spenningsfeltet mellom standardisering, profesjonelt skjønn og samfunnets forventninger. Funnene i denne studien viser at dette spenningsfeltet er høyst reelt i arbeidet med standpunktvurdering i norsk skole.

Dette innebærer at rettferdighet i standpunktvurdering ikke kan forstås som et individuelt ansvar alene, men må ses som en systemisk egenskap, avhengig av tydelige nasjonale rammer, lokale strukturer for støtte og kollektive profesjonskulturer.

6.7 Avsluttende refleksjoner: Standpunktvurdering som systemisk praksis

Denne studien viser at standpunktvurdering i norsk skole ikke kan forstås som summen av enkeltlæreres praksiser alene, men som et resultat av et komplekst samspill mellom nasjonale føringer, lokale strukturer og individuelle profesjonsutøvere. Funnene tydeliggjør at kvalitet i vurdering oppstår, og noen ganger svekkes, i forbindelseslinjene mellom nivåene i systemet.

Et sentralt bidrag fra denne studien er å vise hvordan de etablerte kvalitetskriteriene i vurdering ikke opptrer som stabile idealer, men som størrelser i kontinuerlig spenn. Lærere og skoler forhandler i praksis mellom hensynet til faglig relevans og bredde i vurderingsgrunnlaget, behovet for sammenheng og forutsigbarhet i vurderingene, elevers og foresattes opplevelse av rettferdighet, og de konkrete rammene for tid, ressurser og kapasitet.

Disse spenningene forsterkes av fagenes egenart. Praksisnære, estetiske og yrkesrettede fag stiller særlige krav til vurderingsformer som er situerte, komplekse og mindre standardiserte enn i mer teoretisk orienterte fag. Samtidig stiller disse fagene høye krav til profesjonelt skjønn og lokal faglig forankring. Studien viser at der slike fag mangler tydelige nasjonale eller lokale tolkningsfellesskap, blir vurderingspraksisen mer sårbar for utilsiktede variasjoner.

Diskusjonen peker på at utfordringer i standpunktvurdering ikke kan forstås som et spørsmål om den enkelte lærers kompetanse eller vilje alene. Snarere synliggjøres hvordan systemets organisering legger sterke føringer for hva som er mulig å realisere i praksis. Når ambisjoner om helhetlig vurdering, likeverdighet og rettferdighet ikke følges av tilstrekkelige strukturer for samarbeid, kalibrering og profesjonell støtte, plasseres et uforholdsmessig stort ansvar hos den enkelte lærer. Behovet for mer systematiske strukturer for samarbeid og kollektiv profesjonsutvikling løftes også i Kvalitetsutviklingsutvalgets sluttrapport (NOU 2023:17).

Samtidig viser funnene et betydelig profesjonelt potensial. Der skolene har utviklet sterke fagmiljøer, faste arenaer for dialog og har en aktiv skoleledelse som støtter opp om vurderingsarbeidet, fremstår standpunktvurderingen som mer og mer forutsigbar, bedre begrunnet og med høyere faglig og sosial legitimitet.

Samlet sett viser studien at standpunktvurdering må forstås som en systemisk praksis, der kvalitet ikke primært er et resultat av individuelle vurderinger, men av hvordan systemet som helhet legger til rette for profesjonell dømmekraft, felles forståelse og bærekraftige arbeidsformer.

7. IMPLIKASJONER OG ANBEFALINGER

Funnene i denne studien viser at kvalitet i standpunktvurdering utvikles i et samspill mellom nasjonale rammer, lokale strukturer og lærernes profesjonelle skjønn. Implikasjonene nedenfor er formulert for å støtte mer likeverdig praksis på tvers av fag og skoler, og bygger på både det samlede kunnskapsgrunnlaget, de fagvise funnene og diskusjonskapet.

Anbefalingene må forstås som forslag basert på mønstre og tendenser i et sammensatt datamateriale, og ikke som uttrykk for entydige konklusjoner om årsaker eller effekter. Spørreundersøkelsene gir et bredt bilde av læreres erfaringer og oppfatninger, mens intervjuene gir dypere innsikt i begrunnelser og lokale kontekster. Samtidig varierer datagrunnlaget mellom fag, og enkelte fag omtales mer inngående gjennom de fagvise vedleggene. Anbefalingene er derfor formulert på systemnivå og peker på behov for strukturer og støtteordninger som kan bidra til kvalitet og likeverdighet i standpunktvurdering på tvers av fag og lokale kontekster.

Styrke felles strukturer for vurderingskvalitet

Det er behov for å utvikle og tydeliggjøre strukturer som støtter systematisk arbeid med standpunktvurdering på skole- og fagnivå. Funnene viser at der slike strukturer er svake, blir praksis mer personavhengig og sårbar for uønsket variasjon.

Anbefalinger:

- etablere faste og prioriterte arenaer for samarbeid og moderasjon av standpunktvurdering
- tydeliggjøre skoleledelsens ansvar for å lede og følge opp vurderingsarbeidet
- sikre at praktisk-estetiske fag og yrkesfag prioriteres på linje med øvrige fag i skolens kvalitetsarbeid

Utvikle og konkretisere kjennetegn på måloppnåelse

Funnene peker på et gjennomgående behov for mer presise, fagspesifikke og praksisnære kjennetegn på måloppnåelse, særlig i fag som forutsetter tilgang til utstyr, spesialrom eller praksisarenaer, og i fag der vurderingsgrunnlaget i stor grad bygger på praktisk utøvelse og situasjonsavhengige ferdigheter.

Anbefalinger:

- utvikle mer konkrete, fagspesifikke og praksisnære kjennetegn på måloppnåelse, særlig i praktisk-estetiske fag og yrkesfag
- sikre at kjennetegnene støtter både validitet og reliabilitet ved å være faglig forankret og gjenkjennelige på tvers av skoler
- utvikle støttemateriell som viser hvordan kjennetegn kan brukes i ulike vurderingssituasjoner

Styrke felles forståelse av prinsipper for standpunktvurdering

Selv om lærere i stor grad kjenner vurderingstekstene i læreplanverket, viser funnene betydelig variasjon i hvordan prinsippene for standpunktvurdering forstås og praktiseres.

Anbefalinger:

- legge til rette for systematisk arbeid med felles forståelse av grunnleggende prinsipper for standpunktvurdering
- integrere arbeid med vurderingstekster og prinsipper for standpunktvurdering i skolens øvrige kvalitetsutvikling, for eksempel gjennom etter- og videreutdanning, lokalt læreplanarbeid, sensorskolering og profesjonsfellesskap
- støtte utvikling av fagspesifikke eksempler som viser hvordan samlet kompetanse kan konkretiseres og vurderes i praksis
- vurdere å videreutvikle og tydeliggjøre vurderingstekstene i læreplanverket, spesielt knyttet til operasjonalisering av «samlet kompetanse» og «bredt vurderingsgrunnlag» i ulike fag

Styrke bredden i vurderingsgrunnlaget

Funnene tyder på at vurderingsgrunnlaget i flere fag fortsatt kan være preget av tradisjonell eksamenslogikk, blant annet i matematikk og naturfag, mens andre fag strever med å etablere og dokumentere praksisnær kompetanse.

Anbefalinger:

- styrke bruken av varierte underveisvurderingssituasjoner som grunnlag for standpunktvurderingen
- tydeliggjøre hvordan elevenes progresjon og kompetanseutvikling kan inngå i standpunktvurderingen, innenfor rammen av at standpunkt skal uttrykke elevens kompetanse ved avslutningen av opplæringen
- oppfordre til å vurdere bruken av tradisjonelle skriftlige prøver og heldagsprøver i det enkelte fag

Styrke elevmedvirkning i vurderingsarbeidet

Studien viser et ønske blant lærere om mer elevmedvirkning, men også at dette ofte oppleves som vanskelig å gjennomføre innenfor eksisterende rammer.

Anbefalinger:

- utvikle konkrete, fagspesifikke eksempler på hvordan elevmedvirkning kan integreres i vurderingsarbeidet
- prioritere støtteressurser i fag der behovet er særlig stort
- sikre at elevmedvirkning knyttes tydelig til læringsmål og vurderingskriterier

Styrke profesjonsfaglig samarbeid om vurdering

Funnene viser at samarbeid om standpunktvurdering ofte er tilfeldig organisert, og i stor grad avhengig av uformelle strukturer.

Anbefalinger:

- etablere faste strukturer for moderasjon og kalibrering av standpunktkarakterer
- utvikle rutiner for felles utvikling og drøfting av vurderingsoppgaver og kriterier
- styrke samarbeid både innenfor og på tvers av skoler, særlig i fag uten nasjonal eksamen

Se standpunktvurdering som et systemansvar

Funnene understreker at kvalitet i standpunktvurdering ikke kan reduseres til et individuelt anliggende, men må forstås som et ansvar på systemnivå.

Anbefalinger:

- styrke sammenhengen mellom nasjonale føringer, lokale strukturer og læreres praksis
- utvikle tiltak som ivaretar sentrale kvalitetsdimensjoner i vurdering
- videreutvikle standpunktvurdering som en kollektiv oppgave for profesjonen
- styrke lærerutdanningens ansvar for utdanning i vurdering blant lærerstudentene

8. LITTERATUR

- Beck, K. C., Røhr, H. L., Reme, B., & Flatø, M. (2023). Distressing testing: A propensity score analysis of high-stakes exam failure and mental health. *Child Development*, cdev.13985. <https://doi.org/10.1111/cdev.13985>
- Beinert, C., Palojoki, P., Åbacka, G., Hardy-Johnson, P., Engeset, D., Rudjord Hillesund, E., Selvik Ask, A. M., Øverby, N. C., & Nordgård Vik, F. (2021). The mismatch between teaching practices and curriculum goals in Norwegian Home Economics classes: A missed opportunity. *Education Inquiry*, 12(2), 183–201. <https://doi.org/10.1080/20004508.2020.1816677>
- Birkelund, I., & Midthaugen, P. (2019). Ny som kroppsøvlingslærer – hvordan oppleves det første året i yrket? *Nordisk tidsskrift for utdanning og praksis*, 13(1), Artikkel 1. <https://doi.org/10.23865/up.v13.1835>
- Bjørnson, J. K. (2020). *Matematikkeksamen gjennom tre år. Psykometrisk analyse av eksamen i matematikk for 10. Trinn fra årene 2017, 2018 og 2019*. Universitetet i Oslo, Institutt for lærerutdanning.
- Bjørnson, J. K., & Skar, G. B. (2021). *Sensorreliabilitet på skriftlig eksamen i videregående opplæring [Examiner reliability on written exams in upper secondary education]*. University of Oslo. <http://urn.nb.no/URN:NBN:no-88263>
- Blömeke, S., Skillinghaug, S., Blikstad-Balas, M., Eggen, P.-O., Fjørtoft, H., Gamlem, S. M., Prøitz, T. S., Tveit, S., Helgesen, R., Johannesen, S., Minken, M., Waage, A., & Walker, M. J. (2019). *Kunnskapsgrunnlag for evaluering av eksamensordningen*. Utdanningsdirektoratet. <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/Kunnskapsgrunnlag-for-evaluering-av-eksamensordningen/#>
- Borgen, J. S., Murtnes, Å., Bergsland, J. E., Bottolfs, M., Carlsen, K., Husebø, Ø., Ouff, S. M., Randers-Pehrson, A., Møller-Skau, M., Thorrud, S., Weum, K. B., & Ørb, T. (2023a). *De praktiske og estetiske fagene – fagtradisjoner, fornyelse og endringer i læreplaner for fagene frem mot LK20. Evaluering av fagfornyelsen i fire fag. Delrapport 1* (No. 130). Universitetet i Sørøst-Norge. https://openarchive.usn.no/usn-xmlui/bitstream/handle/11250/3097401/2023_130_Borgen_Delrapport%201_2utg.pdf?sequence=5&isAllowed=y
- Borgen, J. S., Murtnes, Å., Bergsland, J. E., Bottolfs, M., Carlsen, K., Husebø, Ø., Ouff, S. M., Randers-Pehrson, A., Møller-Skau, M., Thorrud, S., Weum, K. B., & Ørb, T. (2023b). *Praksis i de praktiske og estetiske fagene i LK20. Evaluering av fagfornyelsen i fire fag. Delrapport 2* (No. 131). Universitetet i Sørøst-Norge. https://openarchive.usn.no/usn-xmlui/bitstream/handle/11250/3097404/2023_131_Borgen_delrapport_2.pdf?sequence=5&isAllowed=y
- Bottolfs, M. (2020). Mat og helsefaget i dagens skole. *Norsk pedagogisk tidsskrift*, 104(2), 181–193. <https://doi.org/10.18261/issn.1504-2987-2020-02-08>

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Burner, T., Gustavsen, T. S., Husebø, Ø., Kacerja, S., Rubilar, K., Ruud, L. C., & Salvesen, S. (2024). *EvaFag 2025: Evaluering av fagfornyelsen*.
- Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and Accountability in Policy Design and Implementation: The Common Core State Standards and Implementation Research. *Educational Researcher*, 45(4), 243–251. <https://doi.org/10.3102/0013189X16651080>
- Constantinou, T. E., & Bresler, L. (2010). Assessment in Schools – Creative Subjects. I P. Peterson, E. Baker, & B. McGaw (Red.), *International encyclopedia of education* (s. 262–267). Elsevier. <https://www.sciencedirect.com:5037/science/chapter/referencework/abs/pii/B9780080448947003195>
- Dahl, T., Grut, G., & Østerås, A. K. (2015). *Kan vi stole på karakterene? Innhold og vurdering i samfunnsfag i norsk skole*. Utdanningsforskning.no. <https://utdanningsforskning.no/artikler/2015/kan-vi-stole-pa-karakterene-innhold-og-vurdering-i-samfunnsfag-i-norsk-skole/>
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, 40(3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>
- Davidsen, A. F. (2021). *Vurdering for djupnelæring?* [Master thesis, Høgskulen på Vestlandet]. <https://hvlopen.brage.unit.no/hvlopen-xmlui/handle/11250/2766842>
- Ekren, R., Holgersen, H., & Steffensen, K. (2018). *Kompetanseprofil for lærere i videregående skole. Hovedresultater 2017* (No. 19). Statistisk sentralbyrå. <https://www.ssb.no/utdanning/artikler-og-publikasjoner/kompetanseprofil-for-laerere-i-videregaende-skole>
- Emstad, A. B., & Birkeland, I. K. (2020). *Lærende ledelse*. Universitetsforlaget.
- Federici, R. A., Gjerustad, C., Vaagland, K., Larsen, E. H., Rønsen, E., & Hovdhaugen, E. (2017). Spørsmål til Skole-Norge våren 2017. I 114. Nordisk institutt for studier av innovasjon, forskning og utdanning NIFU. <https://nifu.brage.unit.no/nifu-xmlui/handle/11250/2447569>
- Ferm Almqvist, C., Vinge, J., Väkevä, L., & Zandén, O. (2017). Assessment as learning in music education: The risk of “criteria compliance” replacing “learning” in the Scandinavian countries. *Research Studies in Music Education*, 39(1), 3–18. <https://doi.org/10.1177/1321103X16676649>
- Fjellet, J. S., & Ryer, M. (2022). *Matematikklæreres vurderingspraksis på ungdomstrinnet-Hvordan har den endret seg etter LK20? (Master's thesis, Nord universitet)*. [Master's thesis]. Nord Universitet.
- Fjørtoft, H., & Morud, E. B. (2021). Assessment decision making in vocational education and training. *Studia Paedagogica*, 26(4), Artikkel 4. <https://doi.org/10.5817/SP2021-4-6>

- Fjørtoft, H., & Sandvik, L. V. (2016). *Vurderingskompetanse i skolen. Praksis, læring og utvikling*. Universitetsforlaget.
- Fjørtoft, H., & Sandvik, L. V. (2022). Fjørtoft, H., & Sandvik, L. V. (2022). Autentisk vurdering på yrkesfag. I L. V. Sandvik & H. Fjørtoft (Red.), (s. 145–160). Fagbokforlaget. I *Skoleutvikling i videregående opplæring* (s. 145–160). Fagbokforlaget.
- Frønes, T. S., & Roe, A. (2020). Hva skjer i norsktimene? Elevers oppfatning av norskundervisningen på skolen. I *Like muligheter til god leseforståelse?* (s. 196–221). Universitetsforlaget. <https://doi.org/10.18261/9788215040066-2020-08>
- Gamlem, S. M., Segaran, M., & Moltudal, S. (2024). Lower secondary school teachers' arguments on the use of a 26-point grading scale and gender differences in use and perceptions. *Assessment in Education: Principles, Policy & Practice*, 31(1), 56–74. <https://doi.org/10.1080/0969594X.2024.2325365>
- Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2004). Collective Efficacy Beliefs: Theoretical Developments, Empirical Evidence, and Future Directions: *Educational Researcher*. (Sage CA: Thousand Oaks, CA). <https://doi.org/10.3102/0013189X033003003>
- Gottheiner, D. M., & Siegel, M. A. (2012). Experienced Middle School Science Teachers' Assessment Literacy: Investigating Knowledge of Students' Conceptions in Genetics and Ways to Shape Instruction. *Journal of Science Teacher Education*, 23(5), 531–557. <https://doi.org/10.1007/s10972-012-9278-z>
- Hatlen, J. F. (2020). *Historikerens kode: Veien til historisk forståelse*.
- Helsedirektoratet. (2022). *Helsedirektoratets årsrapport 2022*. Helsedirektoratet. <https://www.helsedirektoratet.no/rapporter/helsedirektoratets-arsrapport-2022>
- Hovdhaugen, E., Flobakk-Sitter, F., Wollscheid, S., Fossum, L. W., & Korseberg, L. (2022). *Kartlegging av nordisk forskning på eksamen* (No. 18). NIFU.
- Hovdhaugen, E., Prøitz, T. S., & Seland, I. (2018). Eksamens- og standpunkt karakterer – to sider av samme sak? *Acta Didactica Norge*, 12(4), Artikkel 4. <https://doi.org/10.5617/adno.6276>
- Kaldahl, A.-G. (2020). The educational challenge of oracy - a rhetorical approach: Exploring and articulating the oracy construct in Norwegian schools. *Skriftserien*. <https://skriftserien.oslomet.no/skriftserien/article/view/707>
- Kane, M. T. (2016). Explicating validity. *Assessment in Education: Principles, Policy & Practice*, 23(2), 198–211. <https://doi.org/10.1080/0969594X.2015.1060192>
- Kristiansen, S. (2017). *Lærernes vurderingspraksis i matematikkfaget. En kvalitativ studie fra Nord-Troms av matematikklæreres oppfatninger av vurdering i faget* [Master's thesis]. UiT Norges arktiske universitet.
- Kunnskapsdepartementet. (2017). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Kunnskapsdepartementet.
- Kunnskapsdepartementet. (2023). *NOU 2023: 27 - Et nytt system for kvalitetsutvikling* (NOU No. 27). regjeringen.no. <https://www.regjeringen.no/no/dokumenter/nou-2023-27/id3013760/>
- Kvale, S., & Brinkmann, S. (2009). *Det kvalitative forskningsintervju*. Gyldendal Norsk Forlag.

- Leonardsen, J. K. (2021). One size fits nobody: En casestudie av yrkesfaglæreres deltakelse i kompetanse-utviklingsarbeid i videregående skole : [One size fits nobody: A case study of vocational education and training teachers' participation in professional development projects in upper secondary school]. *Nordic Journal of Vocational Education and Training*, 11(2), Artikkel 2. <https://doi.org/10.3384/njvet.2242-458X.211121>
- Leonardsen, J. K., Utvær, B. K. S., & Fjørtoft, H. (2022). The Five Faces of an Assessor: Conceptualizing the Enactment of Teacher Assessment Identity in Vocational Education and Training. *Educational Assessment*, 0(0), 1–17. <https://doi.org/10.1080/10627197.2022.2106967>
- Lillejord, S. (2011). Kunsten å være rektor. I J. Møller & E. Ottesen (Red.), *Rektor som leder og sjef. Om styring, ledelse og kunnskapsutvikling i skolen* (s. 284–301). Universitetsforlaget.
- Morgan, C. (2016). Studying the role of human agency in school mathematics. *Research in Mathematics Education*, 18(2), 120–141. <https://doi.org/10.1080/14794802.2016.1176595>
- Moss, P., Pullin, D. C., Gee, J. P., Haertel, E. H., & Young, L. J. (Red.). (2008). *Assessment, equity, and opportunity to learn*. Cambridge University Press.
- Nisbet, I., & Shaw, S. (2020). *Is Assessment Fair?*
- Parkes, J. (2013). Reliability in Classroom Assessment. I *SAGE Handbook of Research on Classroom Assessment* (s. 107–123). SAGE Publications, Inc. <https://doi.org/10.4135/9781452218649.n7>
- Prøitz, T. S., & Borgen, J. S. (2010). *Rettferdig standpunktvrdering – det (u)muliges kunst? Læreres setting av standpunktkarakter i fem fag i grunnopplæringen* (No. 16/2010). NIFU. <https://kudos.dfo.no/documents/12924/files/13059.pdf>
- Rambøll, M. C. (2020). *Vurdering i skolen*. Rambøll Management Consulting.
- Rammeverk for eksamen – LK20 og LK20S*. (u.å.). Hentet 2. november 2022, fra <https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-skriftlig-eksamen-i-lk20-og-lk20s/>
- Sandal, A. K. (2023). Vocational teachers' professional development in assessment for learning. *Journal of Vocational Education & Training*, 75(4), 654–676. <https://doi.org/10.1080/13636820.2021.1934721>
- Sandvik, L. V., & Buland, T. (2014). *Vurdering i skolen. Utvikling av kompetanse og fellesskap. Sluttrapport fra prosjektet «Forskning på individuell vurdering i skolen» (FIVIS)*. NTNU.
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Seland, I., Hovdhaugen, E., & Lødding, B. (2018). *Sluttrapport fra evaluering av forsøk med halvårsvurdering med én eller to karakterer i norsk*.
- Skar, G. B., & Hopfenbeck, T. N. (2021). *Kunnskapsgrunnlag om standpunktvrdering*. Skrivesenteret, NTNU.

- Skjæveland, Y. (2020). Samfunnsfagdidaktisk forskning med empiri frå norsk skule – eit forskingsoversyn. *Nordidactica. Journal of Humanities and Social Science Education*, 10(2020:4), 142–163.
- Skålholt, A., Aakernes, N., Andresen, S., Aspøy, T. M., Dahlback, J., Gleditsch, R. N., Johannesen, H. S., Lillebø, O. S., Lyckander, R. H., Lysvik, R. R., Samuelsen, Ø. A., Nyen, T., & Vagle, I. (2024). *Overgang til opplæring i bedrift. Tredje delrapport av evalueringen av fagfornyelsen yrkesfag* (No. 13). NIFU. <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/2024/evaluering-av-fagfornyelse-yrkesfag-rapport-3/>
- Skålholt, A., Aakernes, N., Andresen, S., Aspøy, T. M., Dahlback, J., Høst, H., Johannesen, H. S., Lillebø, O. S., Lyckander, R. H., Lysvik, R. R., & Nyen, T. (2023). *Evaluering av fagfornyelsen yrkesfag, delrapport 2* (No. 12). NIFU. <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/2024/evaluering-av-fagfornyelse-yrkesfag-rapport-3/>
- Stangebye Arnesen, H., Steffensen, K., Foss, E. S., Lervåg, M.-L., & Keute, A.-L. (2023). *Lærerkompetanse i grunnskolen. Hovedresultater 2021/2022* (No. 50/2022). Statistisk sentralbyrå.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional Learning Communities: A Review of the Literature. *Journal of Educational Change*, 7(4), 221–258. <https://doi.org/10.1007/s10833-006-0001-8>
- Svendsen, B., Juel, L. A., & Strømme, A. (2023). *Naturfag for lærere, 5-10*. Gyldendal.
- Sætra, E., & Stray, J. H. (2019). Hva slags medborger? En undersøkelse av læreres forståelse av opplæring til demokrati og medborgerskap. *Nordic Journal of Comparative and International Education (NJCIE)*, 3(1), 19–32. <https://doi.org/10.7577/njcie.2441>
- Tschannen-Moran, M., & Barr, M. (2004). Fostering Student Learning: The Relationship of Collective Teacher Efficacy and Student Achievement. *Leadership and Policy in Schools*, 3(3), 189–209. <https://doi.org/10.1080/15700760490503706>
- Utdanningsdirektoratet. (2015). *Naturfagene i norsk skole anno 2015—Faggjennomgang av naturfagene—Rapport fra ekstern arbeidsgruppe oppnevnt av Utdanningsdirektoratet*. Utdanningsdirektoratet. <https://www.udir.no/globalassets/filer/tall-og-forskning/forskningsrapporter/naturfag-rapport.pdf>
- Utdanningsdirektoratet. (2020). *Læreplanverket*. Utdanningsdirektoratet. [udir.no/laring-og-trivsel/lareplanverket/](https://www.udir.no/laring-og-trivsel/lareplanverket/)
- Utdanningsdirektoratet. (2022). *Utdanningsspeilet 2022*. Utdanningsdirektoratet. <https://www.udir.no/tall-og-forskning/publikasjoner/utdanningsspeilet/utdanningsspeilet-2022/>
- Utdanningsdirektoratet. (2025). *Lite endring i snittkarakterene i videregående*. <https://www.udir.no/tall-og-forskning/statistikk/statistikk-videregaende-skole/analyser/2025/karakterer-i-videregaende-skole-2025/>
- Vinge, J. (2014). *Vurdering i musikkfag – en deskriptiv, analytisk studie av musikk læreres vurderingspraksis i ungdomsskolen*. [Norges musikkhøgskole]. <https://nmh.no/forskning/publikasjoner/vurdering-i-musikkfag>

- Wineburg, S., & McGrew, S. (2019). Lateral Reading and the Nature of Expertise: Reading Less and Learning More When Evaluating Digital Information. *Teachers College Record*, 121(11), 1–40. <https://doi.org/10.1177/016146811912101102>
- Aasland, E., & Engelsrud, G. (2017). «Det er lett å se hvem av dere som har god innsats». Om elevers innsats og lærerens blikk i kroppsøving. *Journal for Research in Arts and Sports Education*, 1(3). <https://doi.org/10.23865/jased.v1.889>

VEDLEGG 1 NORSK

Marita Byberg Johansen og Anne Holten Kvistad

OPPSUMMERING

I denne fagrapporten presenteres et utvalg læreres erfaringer med vurderingsarbeid i norskfaget på ungdomstrinn og i videregående. Datagrunnlaget er fra gruppeintervju med 25 lærere fordelt på tre ungdomsskoler og fire videregående skoler. Under følger en kortfattet oppsummering av hva lærerne rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, og når det gjelder praksis og prosesser knyttet til standpunktvurdering ved skolene.

- Norsklærerne oppgir å ha kjennskap til vurderingstekstene i læreplanen og kjennetegn på måloppnåelse, men det varierer i stor grad hvordan disse dokumentene brukes.
- Vurderingspraksisen i norskfaget varierer i stor grad når det kommer til vurderingsgrunnlag for standpunktvurdering. Dette gjelder ikke bare fra skole til skole, men også innad i team på samme skole.
- Mange lærere uttrykker at norskfaget er et omfattende fag og at det er krevende å skulle sette tre standpunkt karakterer.
- På videregående styrer eksamen mye av arbeidet i faget, ifølge lærerne.
- Lærerne uttrykker noe ulik forståelse for karakterskalaen og avstand mellom skalasteg.
- Flere av norsklærerne uttrykker at det er satt av lite tid til formalisert samarbeid om vurdering.
- Mange norsklærere, særlig på videregående skole, opplever lite støtte fra skoleledelse når det gjelder vurderingsarbeid.
- Elevmedvirkning i vurderingsarbeidet blir nedprioritert på grunn av stor arbeidsbelastning i norskfaget, særlig i videregående skole.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Stort sett alle lærere i utvalget oppgir at de kjenner til vurderingstekstene i læreplanen, men det framkommer i lærernes svar at disse tekstene, med et par unntak, ikke tas fram jevnlig. En lærer uttrykker: «Jeg går ikke rundt og husker ordlyden, men jeg har det jo med meg. Jeg ser jo at jeg gjør mye av dette». Lærerne involverer elevene i liten grad i tolking av tekstene, med et par unntak.

Når det gjelder kjennetegn på måloppnåelse, ser lærerne ut til å bruke skjemaet som et utgangspunkt for å lage egne kriterier og/eller å formulere tilbakemeldinger til elever. Én lærer påpeker at alle formuleringer i kjennetegnene er formulert positivt, og at hen prøver å bruke lignende formuleringer. Det er ellers delte meninger om dokumentet. Noen synes det er «fint», mens andre påpeker at de skulle ønske at kjennetegnene var mer konkrete. På en skole oppgir lærerne å bruke eksamensveiledningen for skriftlig eksamen i norsk i vurderingssituasjoner, fordi kjennetegn på måloppnåelse oppleves som «for generelle». Det varierer fra skole til skole om lærerne viser fram og tolker dokumentet sammen med elevene.

HVA KJENNETEGNER PRAKSISER OG PROSESSER I STANDPUNKTVURDERINGEN?

Samlet sett er det mange lærere som trekker fram at norskfaget er omfattende, og det er travelt å rekke å få tilstrekkelig vurderingsgrunnlag til å sette tre karakterer. Lærerne uttrykker at de kjenner på tidspresset om å få et godt nok vurderingsgrunnlag med både muntlige og skriftlige oppgaver og to målformer. Flere av lærerne på videregående skiller mellom «uformelle» og «formelle» vurderingssituasjoner, altså kompetansen elevene viser i den daglige undervisningen kontra for eksempel prøvesituasjoner, presentasjoner og innleveringer. Alle lærerne i de videregående skolene trekker fram at det er en fordel å følge elevene i faget over flere år. Dette nevnes ikke av lærerne på ungdomstrinn.

På ungdomstrinn ser det ut til å variere hvordan lærerne tenker om standpunkt karakter. Noen uttrykker at de ser på hele semesteret og bruker ordet «snitt», mens andre er tydelige på at det er sluttkompetansen og heldagsprøver på tampen av skoleåret som skal telle mest: «Standpunkt er etter jul, før eksamen, så da er det sånn at man ser på det halvåret». Noen lærere påpeker at det kan være problematisk å sette standpunkt karakter når elever viser lavere kompetanse mot slutten av året, og særlig hvis de vet at regresjonen skyldes vanskelige forhold i hjemmet eller lignende. På videregående er alle lærerne enige om at eksamen styrer mye av arbeidet i faget, og noen lærere uttrykker at eksamensforberedelsene noen ganger kan komme i veien for annet arbeid.

Bruk av karakterskalaen og fastsetting av standpunkt karakter

Lærerne uttrykker noe ulik forståelse for karakterskalaen og avstand mellom skalasteg. Flere trekker fram utfordringer når det kommer til de laveste karakterstegene i standpunkt vurdering, da de vegrer seg for å sette lave karakterer på grunn av kjennskap til elevene. Det kan for eksempel være vanskelig å sette standpunkt når en elev tidligere har gjort det bedre i faget, og at synkende prestasjon mot slutten av året kanskje kan skyldes vanskelige hjemmeforhold. Noen føler et

generelt press om å la elever bestå i faget, og bruker begrepet «nåde-toer». En lærer i videregående påpeker at: «Det har jo ganske store konsekvenser om man består faget». Det er noe uenighet om karakteren 6. Noen mener at den «henger høyt», mens andre mener at det «florerer» med seksere sammenlignet med før. De fleste lærerne er enige i at det å gi «rene» karakterer er best, og at såkalte «glidekarakterer» eller «halvkarakterer» (f.eks. 4/5) kan forvirre elevene og gjøre det utfordrende for læreren når standpunkt skal settes. Likevel er det mange lærere som rapporterer å bruke plusser, minuser og skråstrek i vurderingsarbeidet. Praksis når det kommer til en slik utvidet karakterskala ser ut til å variere også innenfor team på samme skole.

Samlet sett ser det ut til at lærerne er opptatte av å legge til rette for varierte vurderingssituasjoner i faget, slik at elevene kan vise kompetanse på ulike måter. Det er imidlertid noen forskjeller mellom og innad på skolene når det gjelder vurderingspraksis. Noen lærere tematiserer hvordan kunstig intelligens har endret deres vurderingspraksis, og at skriftlige hjemmeoppgaver ikke lenger kan være vurderingsgrunnlag for standpunktkarakterer i faget. Andre er uenige i dette og vil ikke at juks fra enkelte elever skal legge føringer for deres praksis. Vi ser også at lærere på samme skole er uenige om bruken av skriftlige kunnskapsprøver som grunnlag for standpunktkarakter i norsk muntlig. Noen rapporterer å ha slike prøver, mens andre uttrykker at de mener at dette ikke er riktig praksis.

Det er store variasjoner når det kommer til hva lærere vektlegger i standpunktvurderingen. Forskjellene er også fremtredende mellom lærere på samme skole. Noen er tydelige på at det er sluttkompetansen som skal vurderes, og rapporterer å legge opp til store vurderingssituasjoner mot slutten av skoleåret (f.eks. heldagsprøver). Andre vektlegger i større grad «helheten» og «utvikling», og ser lenger tilbake i elevens læringsløp i norskfaget. Dette vil i noen tilfeller handle om det siste året, mens i andre tilfeller ser lærerne enda lenger tilbake. En av lærerne i videregående skole oppgir at hen tar hensyn til russetiden på vg3 når det skal settes standpunktkarakterer. Hvis en elev presterer mye lavere enn tidligere og læreren tenker at det skyldes russetiden, så kan hen velge å se helt bort ifra den aktuelle vurderingssituasjonen.

Det er også ulik praksis når det kommer til justering ved tvil om standpunktkarakterer. Noen opererer med såkalte «vippeprøver». På den ene videregående skolen har de fått beskjed av eksamensansvarlig om å alltid vippe karakterene opp dersom de står mellom to karakterer. En lærer på samme skole forteller at hen tidligere har justert de tre norskkarakterene opp mot hverandre, for eksempel at en av karakterene blir justert ned og den andre opp.

Samarbeid om vurdering

Flere av skolene oppgir at de har fagteam eller fagseksjonsmøter, og samlet sett oppgir lærerne at de generelt har god delingskultur og godt samarbeid med andre faglærere om blant annet å utvikle fagplaner, heldagsprøver og undervisningsopplegg. Flere nevner imidlertid at de gjerne skulle ha hatt mer tid til å jobbe sammen om vurdering og mer om vurdering på fellestid. Samarbeid om vurdering, for eksempel lesing av hverandres elevbesvarelser, skjer gjerne i uformelle settinger på arbeidsrommet. Da kan lærerne for eksempel spørre hverandre om «er dette en 3-er eller en 4-er»? På en av skolene sier en lærer at denne typen samkjøring tar tid dersom det skal gjøres mer systematisk, og at dette forutsetter at «alle må ønske det». På en av ungdomsskolene oppgir lærerne at de er sensor for hverandre på såkalt «prøvemuntlig», altså en

vurderingssituasjon som gjennomføres med like rammer som en muntlig eksamen. Ingen rapporterer at de har faste møter der standpunktkarakterer diskuteres.

Støtte fra skoleledelse

Samlet sett etterlyser lærerne mer samarbeidstid om vurdering. Lærernes opplevelser av støtte fra skoleledelse ser ellers ut til å variere. Mens det særlig på en av ungdomsskolene rapporteres om en skoleleder som har stort fokus på temaet vurdering, ser det ut å være en tendens til at lærerne på videregående i stor grad kjenner seg overlatt til seg selv i vurderingsarbeidet i norskfaget. Noen av lærerne på videregående rapporterer at ledelsen ikke har tilstrekkelig forståelse for norskfagets belastning når det kommer til vurdering. Dette gjelder for eksempel all dokumentasjonen lærerne må skaffe for å kunne håndtere klagesaker for tre standpunktkarakterer. På en skole sier lærerne at de trenger: «tre langsvar på hovedmål, tre langsvar på sidemål pluss to muntlig, og mye kortsvar som tilstrekkelig grunnlag for de tre karakterene». Videre sier lærerne at «vi har lyst til å kutte ut, men vi tør ikke», og de fremhever at norsk er det tyngste faget å drive med vurdering i. Noen av lærerne på videregående rapporterer om at de er med i fagnettverk om vurdering i fylket sitt, og at de gjennom dette samarbeidet har oppdaget at det er stort sprik i vurderingspraksis i norskfaget mellom skoler.

Elevinvolvering

Mange av lærerne uttrykker at elevinvolvering er tidkrevende, og at det ofte må nedprioriteres på grunn av kombinasjonen av mange elever og stor arbeidsbelastning i faget. Dette gjelder spesielt for de videregående skolene. På et par av skolene rapporterer lærerne at elevene selv har uttrykt at de har for lite medvirkning. Det at elever selv skal få velge arbeidsform og hvordan de skal bli vurdert, karakteriseres som utfordrende av flere lærere. Dette er både fordi det blir svært tidkrevende for læreren, og fordi lærerne opplever at det er behov for å ha mer kontroll på vurderingssituasjonene etter at KI-verktøy ble tilgjengelige. På noen skoler trekker lærerne fram at elevene involveres i hverandrevurdering, der de skal bruke kriterier for å vurdere andre. Dette gjøres for å bevisstgjøre elevene, og for å få elevene til å bruke kriterier mer aktivt. På en skole sier lærerne at bruk av kriterier og tydelige tilbakemeldinger fører til færre spørsmål fra elever og dermed mindre merarbeid.

VEDLEGG 2 MATEMATIKK

Olav Dalsegg Tokle og Bodil Svendsen

OPPSUMMERING

Under følger en kortfattet oppsummering av hva matematikklærere ved tre ungdomsskoler og tre videregående skoler rapporterer når det gjelder bruk av vurderingstekstene, kjennetegn på måloppnåelse, og praksis og prosesser knyttet til standpunktvurdering ved skolene. Fagrapporten skal bidra til å nyansere funnene i hovedrapporten og gi et mer konkret bilde av hvordan vurdering forstås og praktiseres i skolene.

- Lærerne oppgir at de er kjent med tekstene om underveis- og standpunktvurdering i læreplanen samt kjennetegn på måloppnåelse, men at få av dem bruker dem aktivt i vurderingsarbeidet sitt.
- Vurderingspraksisen blant matematikklærerne i studien er sterkt påvirket av vurderingspraksisen som historisk sett har vært gjeldende for matematikkeksamen, med individuelle skriftlige prøver der elevenes prestasjon blir skåret som en poengsum og transformert til en karakter ut fra hvor mange prosent av maksskår elevens poengsum tilsvarer.
- Lærerne samarbeider i noen grad med andre lærere om vurdering og grunnlaget som inngår i standpunktvurderingen, men i liten grad i fastsetting av elevenes standpunktvurdering.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Et kjennetegn er at både ungdomsskoler og videregående skoler benytter vurderingstekstene og kjennetegnene på måloppnåelse på ulike måter, med varierende tilnærminger og praksis.

Alle lærerne oppgir i intervjuene at de er kjent med vurderingstekstene, men bruken av tekstene varierer fra skole til skole og mellom lærere på samme skole. Ved flere skoler oppgir lærerne at de ikke bruker vurderingstekstene aktivt, men at de heller benytter egne vurderingskriterier eller andre ressurser som støtte i vurderingsarbeidet sitt. Dette begrunnes i at vurderingstekstene oppfattes som upresise og at det er for tidskrevende å benytte seg av dem i vurderingsarbeidet.

Flertallet av de som oppgir at de bruker vurderingstekstene sier de bruker kjennetegn på måloppnåelse i forbindelse med muntlige vurderingssituasjoner, mens enkeltlærere refererer til deler av teksten om underveis- og standpunktvurdering og sier at det har endret deres vurderingspraksis.

Ungdomsskole

På to av ungdomsskolene oppgir lærerne at de ikke bruker tekstene aktivt. En lærer uttaler at «tekstene drukner litt i læreplanen og alle tekstene som dukker opp». Lærerne opplyser at de bruker kjennetegn på måloppnåelse, eller egenutviklede vurderingskriterier utviklet med utgangspunkt i kjennetegn på måloppnåelse, i forbindelse med muntlig-praktisk eksamen.

En lærer uttaler: «Vi bruker indirekte teksten om standpunkt. Men vi ser ikke til teksten. Vi har jobbet som lærere i mange år, og bruker nok vår erfaring». Om kjennetegn på måloppnåelse uttaler en lærer: «Vi sitter ikke med dette når vi retter. Vi vet hva vi ser etter, og vi bruker det uten at det ligger foran oss».

Videregående skole

På den ene videregående skolen oppgir lærerne at de ikke bruker tekstene aktivt, og omtaler de som «hårete og upresise», men at de selv lager egne kjennetegn på måloppnåelse som «er mer konkrete kjennetegn inn mot ett tema». Vi tolker tema i denne sammenhengen som et matematisk tema eller deler av kjerneelementet *matematiske kunnskapsområder*.

Lærerne på en annen skole oppgir at vurderingstekstene brukes på fellesmøter og at kollegiet har jobbet med å tolke innholdet i tekstene for å utvikle felles forståelse ved skolen og ved tilfeller der elevene klager på standpunktvurderingen.

Den tredje skolen oppgir at kjennetegn på måloppnåelse brukes for eksempel i vurdering av prosjekter i matematikk, altså en vurderingssituasjon som ikke er en skriftlig prøve.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERING?

Alle matematikklærerne i studien oppgir at de vurderer elevenes kompetanse gjennom hele året. Eksempler på situasjoner lærerne bruker for å vurdere elevenes kompetanse, er observasjon, samtale og formelle vurderingssituasjoner, som for eksempel muntlige og skriftlige prøver.

Ingen av lærerne sier at de ser på standpunktvurdering som et gjennomsnitt av underveisvurdering, og lærerne er opptatte av at det er elevens kompetanse når standpunktvurderingen foretas som skal utgjøre elevens standpunktvurdering. Dette fører til at en stor vurderingssituasjon, som en heldagsprøve på slutten av skoleåret, i stor grad utgjør grunnlaget for elevenes standpunktvurdering. Lærerne oppgir videre at dersom resultatet på heldagsprøven ikke samsvarer med underveisvurderingen, kan underveisvurderingen inngå i standpunktvurderingen.

Ungdomsskole

Alle lærerne fra ungdomsskolen er samstemte om at elevene har mulighet til å utvikle kompetansen sin helt fram til standpunktvurderingen settes. I tillegg er lærerne samstemte om at en stor vurderingssituasjon på våren 10. trinn har stor betydning for elevenes standpunkt karakter.

Underveis i skoleåret forteller lærerne at de benytter tilbakemelding uten karakter i formelle vurderingssituasjoner. Om denne praksisen sier en av lærerne: «Det er klart, tilbakemeldingene vi gir nå er mye bedre enn hva de var når vi bare satte karakter». Lærerne forteller videre at de noterer karakterer for seg selv for å ha et system for å følge med på elevenes utvikling. En lærer sier: «Er jo fordi standpunkt ikke er et gjennomsnitt, men det du sitter igjen til med slutt», for å poengtere at elevene kan utvikle kompetansen sin underveis i året og at standpunktvurderingen skal uttrykke kompetansen på slutten av opplæringa. Samtidig sier lærerne at dersom resultatet på heldagsprøven ikke samsvarer med inntrykket læreren har fra før, kan underveisvurderingen

inngå som en del av grunnlaget for standpunktvurdering. En lærer sier om heldagsprøven: «Teller mye om du har gjort det bra».

Det blir fremhevet det som oppleves som et dokumentasjonskrav: «[det er] bra med dokumentasjonskrav, men frykten for å trække feil kan ødelegge for god undervisvurdering fordi det tar fokus bort fra vurdering for læring».

Videregående skole

Lærerne poengterer at det er kompetansen på slutten av året som er grunnlaget for standpunktvurderingen. Til å følge med på elevenes utvikling opplyser lærerne at de benytter egenutviklede skjemaer der de registrer informasjon om elevene.

En lærer sier: «Vi har heldagsprøve på slutten av året, og det liker jeg fordi det oppfyller kravet om at det er sluttkompetansen som skal vurderes og at elevene får mulighet til å forbedre det de har gjort tidligere», mens en annen lærer sier: «De skal ha lov til å tilegne seg kunnskapen på høsten, og som også kan vises på våren». Samtidig oppgir lærerne at dersom resultatet på heldagsprøven ikke samsvarer med inntrykket lærerne har av elevens kompetanse, kan undervisvurderingen inngå som grunnlag i standpunktvurderingen. Lærerne sier også: «Undervisvurderingen er hovedgrunnlaget. Selv om faget avsluttes med en heldagsprøve, så går det an å redde seg inn med arbeid underveis».

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Lærernes bruk av heldagsprøver samsvarer sterkt med hvordan eksamen har vært organisert i matematikkfaget, der matematisk kompetanse har vært tolket som en sum av poeng på deloppgaver.

Felles for lærerne er at de oppgir at de bruker ulike vurderingsformer underveis i skoleåret, men at en stor vurderingssituasjon på slutten av skoleåret utgjør en stor del av grunnlaget for standpunktkarakteren til eleven. Flere lærere oppgir at dersom resultatet på heldagsprøven ikke samsvarer med undervisvurderingen, blir det benyttet andre vurderingssituasjoner. Dette kan være muntlige prøver.

Når det gjelder bruken av karakterskalaen sier fem av lærerne at elevenes besvarelser blir vurdert ut fra et poengsystem og der elevenes poengsum transformeres til en karakter. Lærerne ved disse skolene omtaler skillene mellom de ulike trinnene på karakterskalaen som prosenttall, altså hvor mange prosent av oppnådd maksskår på prøven som kreves for de ulike karakterene.

Flere av lærerne forteller at de samarbeider med andre lærere i utvikling eller tilpasning av prøver, som for eksempel å bli enige om kriterier for poenggiving og grenser mellom de ulike karakterene i karakterskalaene. Samtidig oppgir lærerne i studien at samarbeid med andre lærere er begrenset til undervisvurdering og at de ikke samarbeider med andre lærere i fastsettelsen av elevenes standpunktkarakterer.

Alle lærerne vi har intervjuet oppgir at de benytter flere karakterer enn rene tallkarakter i undervisvurderingen, det gjelder også i halvårsvurderingen. Lærerne begrunner denne praksisen i at det øker elevens motivasjon.

Lærerne er i hovedsak positive til eksamen i matematikk, og de begrunner dette med at de ser på eksamen som en rettesnor for egen standpunktvurdering og at de opplever eksamen som en rettferdig vurderingssituasjon for elevene. Et fravær av eksamen vil, ifølge lærerne, kunne føre til at standpunktvurderingen gir et feilaktig bilde av elevens kompetanse, og at det ville blitt feilaktig variasjon mellom skoler. Dette bygger opp under det at lærerne i studien ser på en stor individuell vurderingssituasjon som et godt egna redskap til å vurdere elevenes matematiske kompetanse på, og at lærerne mener det skal være stort samsvar mellom standpunkt- og eksamenskarakterer. Lærerne påpeker også at eksamen er viktig for elevenes motivasjon og som forberedelse til videre studier.

Samarbeid om vurdering

Noen av lærerne sier at de samarbeidet med kolleger i vurderingsarbeidet, mens andre oppgir at de er helt alene i vurderingsarbeidet. Matematikklærerne som oppgir at de samarbeider med andre lærere i vurderingsarbeidet, forteller om et samarbeid som består i å utvikle prøver, å være sensorer for hverandre i muntlige prøvesituasjoner eller diskutere vurderingspraksis. Lærerne oppgir at de ikke samarbeider med andre lærere i selve fastsettelsen av standpunktkarakteren, og at dette er et ansvar som hviler på den enkelte faglærer.

Ungdomsskole

Lærerne i ungdomsskolen oppgir at de har samarbeidstid, gjerne på trinn, og at denne tida i noen grad brukes til å diskutere vurdering. Samtidig virker det som om det ikke er vanlig å samarbeide med lærere fra andre trinn, eller at dette bare skjer en sjelden gang.

Utover arbeid i faggrupper enten på trinn eller på skolene oppgir lærerne muntlig-praktisk eksamen i matematikk som en arena der de utvikler felles forståelse for læreplan, forskrift og karaktersetting. I tillegg arrangerer flere skoler «prøvemuntlig», der elever (og lærere) øver seg på hva som møter dem dersom elevene kommer opp i muntlig-praktisk eksamen i matematikk. Ved flere skoler deltar da lærere fra andre klasser som eksterne sensorer.

Videre oppgir lærerne i ungdomsskolen at ledelsen legger til rette for samarbeid i faggrupper på skolen, men at det meste av samarbeidet mellom lærerne er egeninitiert. Lærerne sier at dette samarbeidet i liten grad er knyttet til standpunktvurdering. Samtidig oppgir lærerne at det foregår lite samarbeid på tvers av trinn.

Videregående skole

Lærere på de videregående skolene nevner sensorskolering og deltakelse på nettverksdag i fylket som nyttige samarbeidsarenaer. Dette er arenaer der vurdering kan diskuteres på tvers av skolene. På en skole uttrykker lærerne at de samarbeider i fagteam der vurderingspraksis er et av temaene som diskuteres: «Vi snakker om det på matte fagteam. Vurderingspraksis og hvordan vi vurderer et tema». En annen lærer ved skolen poengterer imidlertid at lærerne ikke samarbeider om selve vurderingen: «Vi arbeider alene. Jeg kan ikke bidra inn i din standpunktvurdering, for jeg kjenner ikke eleven din. Du står alene».

På en annen skole sier lærerne at de har mange både formelle og uformelle samtaler med de lærerne de deler kontor med. Lærerne forteller også at de noen ganger har gjennomført noen spesifikke øvelser for å diskutere lærernes vurderingspraksis og å jobbe mot en lik

vurderingspraksis på skolen: «Planen var å dele ut tre ulike prøver, levere anonymt, ta opp prøven med karakter, for å se på forskjellene. Blir mer ufarlig. Det har vi gjort noen ganger, men for sjelden».

Lærerne i videregående skole forteller også at ledelsen legger til rette for samarbeid i faggrupper, men at de ikke legger seg opp i hvordan lærerens praksis om standpunktvurdering er.

Elevinvolvering

Matematikklærerne oppgir ulik grad av og ulike former for elevmedvirkning i vurderingsarbeidet. Grad av elevmedvirkning varierer fra tilnærmet fraværende til at elevene utvikler spørsmål til prøver og vurderer andre elevers prøver. Det kan også innebære at elevene får velge hvilken vurderingsform de vil vurderes etter.

Ungdomsskole

Lærerne fra ungdomsskolene oppgir ulik grad av elevmedvirkning fra skole til skole. Samtidig er eksemplene på elevmedvirkning i vurderingsarbeidet ulike fra skole til skole.

Lærerne ved en skole sier de legger opp til at elevene foretar egenvurderinger, men at disse ofte er overflatiske. En lærer uttaler om elevmedvirkning: «Skulle ønske jeg gjorde det mer. Det tar tid. Hvis elevene skal vurdere seg selv, så tar det så mye tid at det kan man ikke gjøre hver gang». I tillegg oppgir lærerne på skolen at elevene selv kan velge hvilken vurderingsform de ønsker å bli vurdert i underveis i skoleåret. På samme skole snakker også lærerne om elevmedvirkning i forbindelse med arbeidsmåte i faget, ved for eksempel å diskutere ulike løsninger og strategier.

Lærere ved en annen skole oppgir at de ikke involverer elevene i særlig grad i vurderingsarbeidet. En lærer sier: «Det er for lite egenvurdering. Jeg legger ikke opp til det», før en kollega fortsetter: «Vi kommer ofte dårlig ut når det gjelder elevinvolvering på Elevundersøkelsen. Vi er ikke så flinke til å ta de med på råd».

Videregående skole

Blant lærerne på videregående skole oppgir lærerne at det er ulik grad av elevmedvirkning på de ulike skolene.

På en skole sier lærerne at de arrangeres individuelle samtaler mellom elev og matematikklærer i det lærerne kaller «boost-samtaler» to ganger i året. Lærerne sier at de starter samtalen på følgende måte: «Hva kan vi gjøre for at du skal få mest mulig ut av faget?». Ved at lærerne spør hva elev og lærer kan gjøre for at eleven skal få mest mulig ut av faget, legger lærerne opp til elevmedvirkning av stor grad.

Lærerne ved en annen skole oppgir at elevmedvirkningen først og fremst er knyttet til prøver, og at elevene da vurderer sine egne prøver før de får karakter på prøvene.

Lærere ved den tredje skolen forteller at de har prøvd å involvere elevene i etterkant av prøvesituasjoner tidligere, men at de i dag ikke involverer elevene i særlig grad i vurderingsarbeidet. Årsaken til dette er at lærerne sitter med inntrykk av at elevene ikke er interesserte i å medvirke på denne måten.

VEDLEGG 3 NATURFAG

Bodil Svendsen og Stine Aarønes Angvik

OPPSUMMERING

Under følger en kortfattet oppsummering av hva som kjennetegner praksis og prosesser i et utvalg naturfaglæreres standpunktvurdering ved tre ungdomsskoler og fire videregående skoler når det gjelder bruk av vurderingstekstene, kjennetegn på måloppnåelse, og praksis og prosesser knyttet til standpunktvurdering ved skolene.

- Lærerne rapporterer at de er kjent med læreplanens tekster om underveis- og standpunktvurdering samt kjennetegn på måloppnåelse, men bare et fåtall av lærerne bruker disse aktivt i vurderingsarbeidet.
- Vurderingspraksisen blant naturfaglærerne i studien er preget av varierte vurderingssituasjoner med bruk av underveisvurdering. Dette betyr at lærerne innlemmer underveisvurderingen i grunnlaget til naturfaglæreres standpunktvurdering.
- Samarbeidet mellom lærere om fastsettelse av elevenes standpunkt karakter er begrenset. Lærere i ungdomsskolen samarbeid innad i, men ikke mellom trinn. På videregående skole samarbeides det lite om å sette standpunkt karakter i faget.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Alle vi har intervjuet lærerne oppgir at de kjenner til vurderingstekstene, men tekstene brukes lite i underveisvurderingen. Lærerne rapporterer at de i liten grad bruker vurderingstekstene i læreplanene som ressurser i vurderingsarbeidet ved skolene. I stedet foretrekker de fleste av lærerne å bruke kjennetegn på måloppnåelse fra eksterne ressurser som naturfag.no, ofte tilpasset til eget vurderingsarbeid. Selv om dette er en felles praksis som flere av lærerne bruker, samarbeider de ikke om å arbeide med tilpasning av kriteriene.

Generelt opplever lærerne at vurderingstekstene er lite konkrete i forhold til det faglige innholdet i faget, og derfor mener de tekstene er utfordrende å bruke aktivt i det daglige vurderingsarbeidet.

Ungdomsskole

Lærerne ved en skole forteller at de ikke bruker vurderingstekstene aktivt, og setter dette i sammenheng med at de ønsker mer styring fra ledelsen om og hvordan de skal arbeide mer inngående med vurderingstekstene. En lærer sier «... det er altfor travelt i skolehverdagen til å aktivt bruke disse tekstene».

På samme måte forteller lærerne ved en annen skole at de ikke bruker vurderingstekstene aktivt, og dette forklares med at de synes tekstene er formulert på en måte som gjør at lærerne ikke ser dem som hensiktsmessige å bruke som del av deres vurderingsarbeid. Innvendingen handler i hovedsak om at tekstene blir for lite konkrete.

Videregående skole

Lærerne på videregående skole sier at de kjenner til vurderingstekstene, men at kompetansemålene i faget definerer tema på en tilstrekkelig måte, også opp mot vurdering av kompetanse i faget.

Lærerne ved en skole synes at vurderingstekstene er for lite konkrete, men også for lite objektive, og de opplever at vurderingstekstene ikke kan brukes til å «måle noe». De savner tydeligere retningslinjer og konkrete forslag for hvordan dette kan gjøres i vurderingssammenheng i naturfag. Vurderingstekstene har likevel vært knyttet til overordnede diskusjoner rundt de tverrfaglige temaene og de grunnleggende ferdighetene.

En annen skole forteller at vurderingstekstene brukes på starten av året. Da viser de dem til elevene slik at elevene også skal være kjent med disse. Underveis i skoleåret brukes de ikke mer sammen med elevene.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERINGEN?

Samtlige lærere sier at de vurderer elevenes kompetanse gjennom hele året. Denne underveisvurderingen skaper et variert vurderingsgrunnlag, og består eksempelvis av individuelle og gruppebaserte fagsamtaler, skriftlige prøver og innleveringer, muntlige prøver og presentasjoner, og praktisk arbeid i klasserommet. Lærerne er tydelige på at de ikke gjør «gjennomsnittsmålinger» av elevenes kompetanse ved skoleårets slutt, men at de bruker en bred og variert underveisvurderingspraksis fortløpende og gjennom hele året, og at det er på dette grunnlaget standpunktkarakterene settes.

Lærerne opplever at underveisvurdering er en viktig, men utfordrende del av vurderingsarbeidet. Lærerne opplever press rundt dokumentasjonskravene rundt underveisvurderinger. Presset kommer spesielt fra elever og foresatte, særlig når det gjelder å vise at vurdering er gjennomført. Mange lærere bruker varierte vurderingsformer, som fagsamtaler, praktisk arbeid og kunnskapsprøver for å gi elevene mulighet til å vise progresjon og fagkompetanse fram mot standpunktvurderingen. Flere skoler prøver ut ulike strategier for å gjøre underveisvurderingen mer rettferdig og transparent, blant annet gjennom tiltak som anonym vurdering og tydeliggjøring av vurderingskriterier overfor elevene.

Det er imidlertid variasjoner i hvordan underveisvurderingen gjennomføres, basert på måten faget organiseres gjennom skoleåret. Naturfag undervises ofte i tematiske bolker, noe som medfører at elevenes kompetanse knyttet til enkelte emner vurderes tidligere på året. På den ene siden åpner dette for fordypning i et emne over lengre tid. På den andre kan det begrense elevenes mulighet til å forbedre seg i emner som blir undervist tidlig i skoleåret.

Ungdomsskole

Lærerne på en skole sier at vurderingsarbeidet skjer kontinuerlig, men peker på at det i naturfag er vanlig at noen emner avsluttes tidlig i skoleåret, fordi faget er inndelt og organisert i ulike «fagbolker». Dette fører til at enkelte tema er «ferdigvurdert» lenge før lærerne setter standpunktkarakter ved slutten av opplæringen i faget.

Ved en annen skole fremhever lærerne at de bruker mye mer tid på undervisvurdering nå enn tidligere, og de nevner spesielt at de bruker mye tid på dokumentasjon rundt undervisvurdering.

Lærerne på en tredje skole sier at det er vanlig ved denne skolen å bruke vurderingsformen prøvemuntlig for å «knytte trådene sammen» og styrke lærerens inntrykk av elevenes kompetanse fram mot standpunktvurderingen.

Videregående skole

Lærerne på to skoler rapporterer at de organiserer naturfagundervisningen i bolker og at noen emner som undervises på høsten ikke kommer tilbake på våren. For noen lærere oppleves dette som utfordrende, mens andre lærere synes dette er en god praksis. Lærerne er opptatt av at elevene skal oppleve en rettferdig vurderingspraksis, og de har prøvd ut anonym vurdering ved bruk av andre i kollegiet, som en av flere måter å vurdere på. Dette krever en del organisering, og gjøres ikke regelmessig.

På den tredje skolen forteller lærerne at elevene blir orientert om hva som skal telle med i standpunktvurderingen, og hvordan dette vil uttrykkes i standpunktkarakteren. Lærerne driver løpende undervisvurdering gjennom året, og de bruker varierte vurderingsformer for at elevene skal få vise sin kompetanse. Vurderingene gjennomføres både praktisk og teoretisk, og organiseres som en kombinasjon av både muntlige og skriftlige, individuelle og gruppebaserte vurderingssituasjoner. Elevenes arbeid i timene, som muntlig faglig aktivitet, oppgaver og forsøk, er også en del av det som vurderes underveis i læringsprosessen, sammen med for eksempel fremføringer, prosjekter, podkaster etc.

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Lærerne løfter fram at gjennom å bruke en systematisk og helhetlig tilnærming til undervisvurdering, kan læreren sikre at standpunktkarakteren oppleves rettferdig for eleven.

Lærerne legger vekt på at fastsetting av standpunktkarakter i naturfag skal henge tett sammen med en vurdering av elevens faglige utvikling gjennom året. Ved å inkludere en kombinasjon av individuelle aktiviteter og gruppeaktiviteter, samt skriftlige, muntlige og praktiske vurderingsformer, sikrer lærerne at standpunktkarakteren gjenspeiler elevens prestasjoner, sett i lys av kompetansemålene. Lærerne rapporterer om en kontinuerlig vurderingspraksis der varierte vurderingsformer brukes for å gi en rettferdig vurdering. Elevene får også jevnlig tilbakemeldinger og muligheter til å vise faglig forbedring underveis, noe som bidrar til en transparent og utviklingsfokuset vurderingsprosess.

De fleste naturfaglærerne vi har intervjuet, sier at de benytter hele tallkarakterer i undervisvurderingen og halvårsvurderingen. Noen lærere bruker også +/- på karakterene, og de begrunner denne praksisen med at det øker elevens motivasjon. Betegnelse lav, middels og høy måloppnåelse brukes av noen lærere i stedet for hele karakterer i vurderingssituasjoner underveis.

Eksamen blir sett på som en viktig kalibreringsmekanisme for å sikre rettferdige standpunktkarakterer, både innad og på tvers av skoler. Lærerne påpeker også at eksamen er viktig for elevenes motivasjon og som forberedelse til videre studier. Eksamen i naturfag vurderes av lærerne som et viktig verktøy både for å sikre rettferdighet i standpunktvurderingen og for å motivere elevene.

Ungdomsskole

På alle tre skolene sier lærerne de er bevisste på å vurdere elevenes kompetanse gjennom hele året, og de vektlegger å planlegge for varierte vurderingsformer. Standpunktvurderingen ses i lys av underveisvurderingen, og standpunktkarakteren settes ut fra samlet vurdering av elevens kompetanse.

Tilbakemeldinger gis jevnlig, både muntlig og skriftlig, for å støtte elevens utvikling. En lærer påpeker hvordan vurderingene reflekterer elevens innsats i praktiske og utforskende aktiviteter, samt evnen til å anvende fagbegreper og modeller.

Videregående skole

Lærerne på alle skolene rapporterer at ulike former for vurderinger gjøres kontinuerlig gjennom året, og standpunktkarakteren settes ut fra dette. Lærerne har mange ulike vurderinger, slik som fagsamtaler, skriftlige og muntlige vurderinger, digitale og praktiske vurderinger.

Elevene får tidlig informasjon om hvordan underveisvurderingen henger sammen med den avsluttende vurderingen. Lærerne er opptatt av å gi god faglig oppfølging og læringsfremmende tilbakemeldinger underveis i året, for slik å gi elevene mulighet til å bli bedre om det er noe de ikke mestrer.

Samarbeid om vurdering

Noen lærere sier at de samarbeider med kolleger i vurderingsarbeidet, mens andre oppgir at de er alene i vurderingsarbeidet. Oppsummert for skolene ser vi at alle savner en systematisk tilnærming til vurdering, med mer avsatt tid og flere arenaer til samarbeid innad i kollegiet. Lærerne opplever arbeidet med vurdering ved skolene på ulike måter.

Flere lærere ønsker å utvikle felles vurderingskriterier og arbeide sammen om å lage undervisningsopplegg knyttet til tverrfaglige temaer.

Lærere uttrykker behov for mer samarbeid og støtte fra skoleledelsen for å sikre en konsistent og rettferdig vurderingspraksis. Lærerne forteller at skoleledelsen tilrettelegger for samarbeid i faggrupper på skolen. Likevel er det meste av dette samarbeidet drevet av lærernes egen initiativ. De påpeker at samarbeidet sjelden handler om standpunktvurdering. Vurderingsarbeidet oppleves som en individuell oppgave. Selv om det finnes arenaer for samarbeid i faggrupper, oppgir lærerne at de får lite støtte fra både ledelsen og kolleger i dette arbeidet. Lærerne etterlyser mer samarbeidstid om vurdering.

Ungdomsskole

Lærerne skolene rapporterer om lite samarbeid i det samlede lærerkollegiet, og noe mer samarbeid innad i trinn, både formelt og uformelt. Innad i trinn sier lærere at det samarbeides mye med å lage felles undervisnings og vurderingsplaner, men de samarbeider ikke rundt selve vurderingen. Mellom trinnene er det ikke noe samarbeid. Dette begrunner de med at det er en travel skolehverdag, og slikt tar tid.

Videregående skole

Lærerne har samlet et tydelig ønske om mer tid og struktur for samarbeid rundt vurderingsarbeidet. Samlet sett kan det synes som at en mer helhetlig og styrt tilnærming fra ledelsen kan være nødvendig for å sikre at samarbeidet om vurdering blir effektivt og meningsfullt.

Lærerne sier det generelt er lite samarbeid om vurderingspraksis, men de er tydelige på at de ønsker mer tid til dette. Lærerne tar ofte initiativ til et samarbeid med kollegaer når de ser behov for det. Det rapporteres om få samarbeidsmøter for vurderingssamarbeid. Lærerne erkjenner at det er ulike oppfatninger rundt hva som er god vurderingspraksis, og de tenker dette kan bedres med økt lærersamarbeid.

Elevinvolvering

Naturfaglærerne mener det er en ulik praksis knyttet til elevmedvirkning, hvor elevene involveres i ulik grad og lærerne legger til rette for ulike former for elevmedvirkning i vurderingsarbeidet. Noen lærere sier at elevinvolvering er tilnærmet fraværende, mens andre forteller at elevene får velge hvilken vurderingsform de vil vurderes etter. Lærerne sier at det varierer i hvilken grad elevene inkluderes i vurderingsprosessen, og involveringen er stort sett knyttet til valg av vurderingsform.

På ungdomsskolenivå er graden av elevmedvirkning generelt lav, med noen lærere som lar elevene gi innspill på vurderingsformer. På videregående nivå ser det ut til at elevene involveres noe mer, men fremdeles i begrenset grad, og stort sett kun knyttet til valg av vurderingsform. Lærerne setter inn få systematiske tiltak for elevmedvirkning.

VEDLEGG 4 MUSIKK

Henning Fjørtoft og Marita Byberg Johansen

OPPSUMMERING

Under følger en oppsummering av intervjuer med åtte musikk lærerne ved fire ungdomsskoler og hva disse lærerne rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, og når det gjelder praksis og prosesser knyttet til standpunktvurdering ved skolene.

- De fleste lærerne sier at de kjenner til vurderingstekstene, men de er ikke i aktiv bruk i hverdagen. Lærerne bruker kjennetegn på måloppnåelse i liten grad, primært fordi de oppleves som overordnede og vage.
- Vurdering i faget er krevende på grunn av lavt timetall.
- Lærersamarbeid om vurdering i musikk varierer sterkt. På flere skoler beskriver lærerne få etablerte møtepunkter for vurderingssamarbeid, noe de knytter til fagets rammebetingelser og lave timetall. Enkelte opplever også at musikkfaget blir nedprioritert sammenlignet med andre praktisk-estetiske fag.
- Det varierer om elever blir involvert i vurderingsarbeidet i faget og hva de eventuelt involveres i.
- Det varierer hva lærerne vektlegger i og hvilke vurderingsformer som danner grunnlag for standpunktvurderingen.
- Generelt uttrykker lærerne at det er enkelt å skille grovt mellom lav, middels og høy kompetanse, men de opplever likevel at karakterene samler seg på midten av skalaen.
- Skoleledelsen er generelt ikke involvert i vurderingsarbeidet i musikkfaget.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE

Kompetansemålene er velkjente og regnes som viktige, men lærerne vektlegger gjerne kjerneelementene mer enn kompetansemålene. De fleste lærerne kjenner til vurderingstekstene, men bruker dem ikke aktivt i hverdagen. De fleste sier likevel at innholdet i tekstene i stor grad korresponderer med deres praksis. Noen bruker teksten om underveisvurdering som en påminner om hvordan faget skal forstås: «Tekstene om underveisvurdering samsvarer med min forståelse om hvordan man skal vurdere i faget». Lærerne bruker kjennetegn på måloppnåelse i liten grad, primært fordi de oppleves som overordnede og vage. Selv om kjennetegn på måloppnåelse kan være nyttig for lærerne, kan de ikke brukes mot elevene, da språket ikke er tilgjengelig. En lærer påpeker at vurderingstekstene er nyttige for lærerne, men at de «må bruke et annet språk ut mot elevene».

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERING?

De aller fleste musikk lærerne vi har snakket med, organiserer undervisningen i tematiske bolker. For eksempel kan de arbeide med temaet *musikk og samfunn* (protestsanger m.m.) før de går videre til *digital komponering* (f.eks. med GarageBand) og *utøvende musikk*. Noen lærere bruker også et spiralprinsipp, der ulike faglige områder berøres flere ganger med økende vanskegrad slik at elevens utvikling kommer til syne gjennom året.

Organiseringen i tematiske bolker gjenspeiler seg i bredden i vurderingssituasjoner i faget. Fremføring av musikk regnes som viktig, men noen lærere sier at ikke alle elever trenger å fremføre foran et publikum – det holder å fremføre for læreren for å vise sin kompetanse. En gruppe påpeker at samspill er «kjernen i faget» og at det veier tyngst fordi det er tidkrevende å utvikle individuelt engasjement og evne til å delta i samspill med andre. Musikk har imidlertid lavt timetall med bare 83 undervisningstimer på ungdomstrinnet. Lærerne påpeker at vurdering i faget må balansere mellom praktisk øving på samspill/musisering, innlæring av musikkfaglige begreper og utvikling av refleksjon og bevissthet hos den enkelte elev. Dette er krevende i praksis. Flere lærere beskriver vårsemesteret som et «vippesemester», der elevene stadig får fremovermeldinger og mulighet til å forbedre seg på det praktiske.

Lærerne søker å balansere de fire hovedområdene i faget, men ønsker samtidig å vektlegge mestring noe ulikt hos enkelte elever: «De som synger pent, må få uttelling for det. Noen ganger er det noen som har styrken sin i å reflektere over musikkens rolle i samfunnet, og det må de også få uttelling for.» Grunnlaget for vurderingen kan variere mellom elever, og noen lærere kan bruke grunnlag fra andre sammenhenger for å sette karakter i musikk: «for eksempel ved mye fravær så kan vi trekke inn ting fra andre settinger der de viser kompetanse, for eksempel valgfag *Produksjon for scene*».

Lærerne er uenige om læreplankravet om programmering i faget. Noen lærere påpeker at dette oppleves lite relevant for musikk, mens andre mener at dette er en del av faget og at det hører fremtiden til. Lærerne rapporterer også om utfordringer med tilgang til faglig relevant programvare for å øve på programmering i faget.

Bruk av karakterskalaen og fastsetting av standpunkt karakter

Generelt opplever lærerne at det er enkelt å skille grovt mellom lav, middels og høy kompetanse. Lærerne sier at de skal bruke hele karakterskalaen og sier at de vurderer «prosess fram til sluttprodukt» i de praktiske og utøvende aspektene av faget. De vektlegger betydningen av øvelse – læring i faget handler både om automatisering av ferdigheter og om kognitive aspekter av læring. Musikk lærerne opplever likevel at karakterene ofte samles rundt midten av skalaen: Det skal godt gjøres å få 1 eller 2 i musikk, men det er også krevende å få 6. Toppkarakterer krever «engasjement eller lidenskap» og det er ofte elever med annen musikk erfaring som skårer høyest. Noen lærere bruker glidekarakterer (med pluss, minus og skråstrek), mens andre markerer halvårskarakterer som «sterk» eller «svak» for å informere elevene om vippemuligheter. Atter andre bruker i hovedsak hele karakterer.

Lærerne nevner flere utfordringer i vurdering i musikk: antall elever per lærer, tilgang til kompetanse og utstyr, forskjeller i elevenes utgangspunkt, og synet på hva som skal ligge til grunn for karaktergrunnlaget. Flere lærere vi har snakket med, har mer enn 300 elever i musikk. En lærer sier: «Før jul satte jeg 327 terminkarakterer (karakterer til halvårsvurderingen). Jeg har 9 klasser og tre grupper i valgfag». Lærerne bruker derfor tabeller og matriser der det noteres karakterer og tilbakemeldinger fortløpende. Noen bruker mållark, vurderingskriterier, eller skjema med lav-middels-høy eller på begynnerstadiet-under utvikling-kompetent-fremragende. Skjema gjør det mulig å «sette raske vurderinger underveis». Programvare som OneNote og Teams brukes både til kommunikasjon med eleven og til dokumentasjon. Fremføringer av samspill filmes slik at læreren skal kunne observere flere elever samtidig. Dette er et eksempel på at teknologi kan anvendes for å gjøre vurderingen mer håndterbar for læreren.

Siden musikkfaget kombinerer praktisk-estetiske ferdigheter i musisering med musikkfaglig kunnskap, påvirkes vurderingen både av lærerens kompetanse og tilgang til lokaler og instrumenter.

Elevenes utgangspunkt varierer sterkt. Noen spiller musikk fra før, andre ikke. Noen elever har utfordringer med musikkfaglige begreper og har varierende utgangspunkt for å si noe om grunnelementene i musikkteori. Lærerne er tydelige på at de ikke skal regne gjennomsnittskarakterer og at manglende grunnlag i en tematisk bolk (f.eks. musikk og samfunn) ikke skal føre til trekk i standpunkt-karakter. Enkelte påpeker at det er utfordrende å vurdere alle kompetansemålene det siste halvåret før standpunkt.

Synet på hva som skal telle i karaktergrunnlaget varierer. Noen lærere ser på hele året som grunnlag for standpunkt-vurdering, mens andre sier at de «ser ikke på notatene fra høsten» når de setter karakter til våren. Argumentet er at elevene bør kunne «starte på nytt på våren». Det er ikke uvanlig at elever kan få en «second chance» til å vise kompetanse når de har øvd på et instrument, eller «vippeprøver» hvis de ligger mellom to karakterer.

I tillegg viser undersøkelsen at flere lærere uttrykker et ønske om at det legges til rette for fagfellesskap om vurderingsarbeid både internt og på tvers av skolene. Et slik fellesskap der lærerne kan få kalibrere og samkjøre om de slutningene de gjør seg om elevenes kompetanse er gyldige, vil være viktig med tanke på å ivareta validitet når det kommer til standpunkt-vurdering.

Samarbeid om vurdering

Lærersamarbeid om vurdering i musikk varierer sterkt på de fire skolene vi har besøkt, slik disse eksemplene viser:

- Skole 1 har to musikk-lærere. De har uformelle diskusjoner og har også uregelmessige møter med erfaringsutveksling med musikk-lærere fra andre skoler.
- Skole 2 og 3 har uformelle møter samt et digitalt diskusjonsforum for å dele undervisningsopplegg på egen skole.
- Skole 4 har flere musikk-lærere, men få møtepunkt og opplever at faget blir «glemt og nedprioritert», selv sammenlignet med andre praktisk-estetiske fag.

Faget har generelt lite tid til samarbeid, selv sammenlignet med andre praktisk-estetiske fag. Noen lærere har «ingen felles plattform for vurdering» og heller ikke noe nettverk utenfor skolen. Disse

lærerne etterlyser mer struktur i skolens arbeid med vurdering. Andre deler undervisningsopplegg på skolens digitale plattform, i chatkanaler for musikk lærere på samme skole og i uformelle samtaler. «De uformelle samtalene er viktige for å se om vi er samstemte,» sier en lærer. Det er vanlig å bruke kolleger som en ressurs: «Vi vurderer sammen hvis vi er i tvil på oppgaver og i standpunkt». En annen påpeker verdien av et fagfellesskap: «Det er veldig viktig at vi er flere musikk lærere. I forrige jobb var jeg alene på barnetrinn og var nyutdannet, så det å ha et fellesskap er kjempeviktig».

Flere musikk lærere etterlyser samarbeid på tvers av skoler. Enkelte påpeker at «kommunen setter i gang samarbeidsøkter med musikk lærere på andre skoler» for å dekke kompetansebehov som ikke fins på lærernes egen arbeidsplass. Atter andre sier at støtten til arbeidet kommer fra Udir, ikke fra skoleeier eller egen skole.

Støtte fra skoleledelse

Skoleledelsen er generelt ikke involvert i vurderingsarbeidet i musikkfaget. Mange av lærerne sier at de har full tillit fra ledelsen til å gjøre selvstendige vurderinger. Noen skoleeiere tilrettelegger imidlertid for nettverk på tvers av skoler for musikk lærere. En skole rapporterer om at kommunen tilrettelegger for nettverk på tvers av skolene, og at disse treffene oppleves som relevante. Lærerne ved skoler som ikke deltar i fagnettverk, uttrykker at det hadde vært nyttig å ha innsikt i hvordan andre lærere jobber med faget.

Elevinvolvering

Elevinvolvering i musikk skjer ved at elevene har ansvar for å dokumentere sin utvikling med loggskrivning. De får også ansvar for å være bandleidere. Elevinvolveringen skjer både underveis og i sammenheng med halvårs- og sluttvurderingen. Elevmedvirkningen tilpasses til elevene: De som allerede spiller et instrument, kan skape musikk selv. Noen elever får også velge vanskegrad på oppgaver, som igjen kan legge føringer for hva slags karakter de kan få.

Det kan være krevende for elever å forstå vurderingen i faget fordi de ikke alltid ser hva grunnlaget for lærerens vurdering er, eller at det ikke er samsvar med deres egne forventninger og karakteren som læreren setter. Noen lærere har fagsamtaler med elevene før standpunkt for å avklare gjensidige oppfatninger. Andre lærere åpner for at elevene kan inngå i selve karaktersettingen:

Etter egenvurderinga har vi en samtale der de får vite terminkarakter (karakter til halvårsvurderingen). Der har de et ord med i laget. Der hender det at jeg endrer oppfatning ut ifra noe som framkommer i samtalen. Den samtalen har vi også før standpunkt.

Det er likevel vanskelig å gjennomføre samtaler når timetallet er lavt. Noen lærere påpeker at elevinvolvering er lettere i fag med større timetall og i valgfag.

VEDLEGG 5 HELSE- OG OPPVEKSTFAG

Lise Vikan Sandvik og Henning Fjørtoft

OPPSUMMERING

Under følger en oppsummering av hva helse- og oppvekstlærere (HO) ved 4 videregående skoler rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, og når det gjelder praksis og prosesser knyttet til standpunktvurdering ved skolene.

- Lærere i HO kjenner til vurderingstekstene som er knyttet til kompetansemålene i læreplanen, men forholder seg hovedsakelig til kompetansemålene gjennom skoleåret. De opplever likevel at prinsippene er internalisert og reflektert i deres praksis.
- I dag utvikles kjennetegn på måloppnåelse lokalt på enkeltskoler eller i fagnettverk. Lærerne etterlyser nasjonale kjennetegn på måloppnåelse for yrkesfagene.
- En sentral utfordring i vurdering i HO er forskjeller i vurderingspraksis mellom skole og bedrift. Lærerne legger læreplanen til grunn for sine tilbakemeldinger, mens praksisveiledere ofte benytter sjekklister og gir generelt positive tilbakemeldinger. Dette skaper noen ganger forventninger hos elever som ikke samsvarer med skolenes vurderinger.
- Elevinvolvering står sentralt i HO, og lærerne understreker at de ønsker å sikre at elevenes stemmer og egenvurderinger tas på alvor.
- I fastsetting av standpunkt karakterer vektlegger lærerne elevens samlede kompetanse i lys av kompetansemålene i læreplanen, slik denne kompetansen har kommet til uttrykk gjennom opplæringen og vurderingsgrunnlaget. De tar hensyn til elevens prestasjoner i ulike vurderingssituasjoner, og lærerne sørger for at elevene har hatt flere muligheter gjennom et bredt spekter av vurderingsformer til å vise sin kompetanse.
- HO-lærerne uttrykker at det er ulik avstand mellom de ulike karaktertrinnene, spesielt mellom karakterene 1 og 2, samt 4 og 5. Dette skyldes de store konsekvensene en lav karakter kan ha for elevene. Karakteren 2 er spesielt viktig fordi den symboliserer et minimum av kompetanse som kreves for å kvalifisere seg til et yrke.
- Noen lærere opplever at det er vanskelig å avgjøre hvorvidt elever med høyt fravær skal få karakteren 1 eller IV (ikke vurderingsgrunnlag). For eksempel er det vanskelig å sikre rettfærdig bedømming når elever med høyt fravær likevel deltar i vurderingssituasjoner.
- Eksamen i programfag forstås som en kilde til kalibrering og styrking av tolkningsfellesskapet, samt en arena for mestring for elevene. Trekkordningen i fellesfag beskrives imidlertid som problematisk og et mulig hinder for læring og progresjon for enkelte elever.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Lærerne i HO oppgir at vurderingstekstene er godt kjent, men at disse i praksis brukes indirekte gjennom kompetansemålene og eget profesjonelt skjønn. I programfagene legger lærerne særlig vekt på å knytte vurderingssituasjoner til realistiske oppgaver fra praksisfeltet, og vurderingstekstene brukes som ramme for planlegging og refleksjon. En arena hvor alle vurderingstekstene imidlertid blir trukket frem, er i forbindelse med klagesaker. Når elever eller foresatte klager på standpunkt karakterer må lærerne fremlegge en formell begrunnelse for karakterfastsettingen. I slike tilfeller opplever lærerne at vurderingstekstene gir dem en struktur for å beskrive vurderingsprosessen, noe som styrker argumentasjonen og tydeliggjør hvilke vurderingskriterier som er blitt anvendt. Det er likevel variasjon i hvordan vurderingstekstene omsettes til praksis i den daglige undervisningen.

Mange lærere etterlyser tydeligere nasjonale kjennetegn på måloppnåelse for yrkesfag. De beskriver at dagens system gir stort rom for tolkning, og at dette kan føre til ulik praksis mellom skoler. Et slikt tolkningsrom oppleves både som en mulighet og en utfordring. Lærerne verdsetter muligheten til å tilpasse vurderingen til fagets egenart, men viser til at felles støtteverktøy kunne gi større trygghet i standpunkt vurderingen.

En stor utfordring som lærerne møter i HO-fagene, er forskjellene mellom vurderingspraksis på skolen og i praksisperiodene ute i bedrift. Praksisveiledere på arbeidsplassen gir ofte svært positive tilbakemeldinger til elevene basert på deres praktiske arbeid, men dette står ikke alltid i samsvar med vurderingen lærerne på skolen gir. En lærer beskriver det slik: «Praksisplassene gir gode tilbakemeldinger, men de kan ikke kompetansemålene. Da blir det vanskelig å få fram at eleven vurderes på kompetansemålene i faget.» Årsaken er at praksisveilederne ikke nødvendigvis kjenner til kompetansemålene i læreplanen. Dette skaper misforståelser, særlig når elevene forventer høyere karakterer basert på tilbakemeldingene de har fått i praksis.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERINGEN?

Lærerne beskriver underveisvurderingen som en integrert del av undervisningen, der mye skjer gjennom observasjon, samtaler og veiledning. Siden HO-elever ofte arbeider praktisk, oppstår mange vurderingsmuligheter i fortløpende situasjoner. Dette gir et bredt og rikt vurderingsgrunnlag, og lærerne opplever at det gir et godt bilde av elevenes utvikling over tid. Lærerne tilrettelegger for en helhetlig vurdering av elevens kompetanse, spesielt gjennom tverrfaglige opplegg som prøveeksamener, og en lærer uttrykker dette slik:

I programfag så bruker vi kompetansemålene i oppstart, underveis og sluttvurdering. [Vi b]ruker ulike metoder: podkast, gruppearbeid, fremlegg til ordinære prøver. Vi samarbeider mye når vi skal ha vurderinger. Tverrfaglig prøveeksamen er også med i grunnlaget for programfagene.

Disse tverrfaglige prøveeksamenene gir elevene muligheten til å vise sin kompetanse på tvers av fag, og brukes som en del av grunnlaget for programfagene. Dessuten er samarbeidet mellom lærerne om å utvikle og gjennomføre vurderinger viktig for å gjøre vurderingsprosessen mer

inkluderende, der ulike fagområder får innflytelse på vurderingen av elevens prestasjoner. Slik får elevene bedre muligheter til å vise sine ferdigheter i en rekke fag, og styrker lærernes evne til å gi en mer rettferdig standpunktvurdering ved slutten av året.

Samtidig opplever flere lærere utfordringer knyttet til elevenes forutsetninger. Mange elever kommer fra ungdomsskolen med svakt karaktergrunnlag eller manglende erfaring med vurderingsspråk. Dette gjør det nødvendig med tett oppfølging og hyppige faglige samtaler. Vurderingsarbeidet krever også omfattende dokumentasjon, og lærerne trekker frem tidspress som en begrensende faktor.

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Lærerne i helse- og oppvekstfag beskriver fastsetting av standpunktkarakter som en krevende prosess som forutsetter både et bredt vurderingsgrunnlag og et omfattende profesjonelt skjønn. De legger vekt på å vurdere elevens samlede kompetanse i lys av kompetansemålene og sørger for at eleven får flere muligheter til å vise kompetanse i ulike og varierte vurderingssituasjoner.

Lærerne benytter karakterskalaen fra 1 til 6, men det er variasjon i hvordan karakterene nyanseres og formidles. Noen lærere bruker såkalte glidekarakterer – pluss- og minusvarianter av karakterene (for eksempel 4+ eller 5–) – for å gi mer presise tilbakemeldinger underveis. Andre unngår glidekarakterer fordi de mener det kan skape uklarhet om hva som faktisk er elevens nivå i relasjon til kompetansemålene.

I tillegg benytter enkelte lærere kategoriene lav, middels og høy måloppnåelse i underveisvurderingen. Denne praksisen brukes særlig når lærerne ønsker å støtte en mer læringsorientert vurderingskultur og hjelpe elevene med å forstå vurderingsspråket uten å knytte alt til tallkarakterer. Lærere som bruker dette, opplever at det gjør vurderingen mer forståelig for elever som strever med fagbegreper eller som er i tidlige faser av kompetanseutvikling.

Variasjonen mellom glidekarakterer, hele karakterer og måloppnåelsesnivåer bidrar til at vurderingspraksisene kan fremstå ulike både for elever og foresatte. Mange lærere uttrykker derfor behov for tydeligere felles kjennetegn og mer enhetlige retningslinjer innen yrkesfagene, hvor vurderingsgrunnlaget er mangfoldig og situasjonsavhengig.

Diskusjoner om avstanden mellom karaktertrinn er også fremtredende. Flere lærere opplever at avstanden mellom karakter 1 og 2 og mellom 4 og 5 føles større enn mellom 3 og 4. Karakteren 3 beskrives som «vid», mens de høyeste karakterene forutsetter at eleven viser både teoretisk forståelse, praktiske ferdigheter og profesjonell refleksjon. Karakteren 2 trekkes frem som særlig viktig, fordi den representerer minimumskompetansen som kreves for å kunne fungere i yrket. Dette kan være krevende i programfag der eleven skal ut i ulike praksisinstitusjoner. Lærerne beskriver at det kan være utfordrende å vurdere hvor mange nye muligheter en elev skal få for å stå dersom de ikke mestrer situasjoner med ulike pasient- og brukergrupper.

Bruken av karakteren 1 og IV skaper også dilemmaer. Det kan være vanskelig å sikre rettferdig vurdering av elever med «kjempefravær», og som likevel deltar i vurderingssituasjoner. Lærerne har etterspurt kompetanse på dette på skolen, og kommenterer at de noen ganger er «(p)risgitt at man har gode kolleger» i «arbeid med forskrifter, lover og regler».

Ved en skole har ledelsen etablert retningslinjer for hvordan karakterskalaen skal brukes, men lærerne ser et behov for ytterligere felles kjennetegn og kriterier. Dette gjelder spesielt for yrkesfagene, der det ikke finnes standardiserte kjennetegn som kan sikre lik vurdering på tvers av skoler. Flere lærere trekker også frem de etiske og praktiske utfordringene ved å sette lave karakterer, på grunn av de store konsekvensene det kan ha for elevenes fremtid. Dette understreker behovet for både klare retningslinjer og bedre støtte for lærerne i vurderingsarbeidet.

Lærerne ser positivt på eksamen, spesielt fordi det gir en mulighet for kalibrering og kvalitetssikring gjennom sensorarbeid. Dette bidrar til å etablere et tolkningsfellesskap blant lærere og sensorer, og gir dem trygghet i vurderingsarbeidet. Eksamen blir også sett på som en motivasjonsfaktor for elevene, da den gir en klar målsetting for hva de skal vise av sin kompetanse.

Eksamen i programfagene blir sett på som en suksess. Lærerne opplever at elevene «gjør det veldig bra», blant annet fordi de gjennomfører eksamen ute i praksis: «Elevene er trygge, på en kjent plass». Læreren understreker at eksamen også har en skjerpande funksjon: Den krever at elevene forbereder seg, bruker hjelpemidler, og mestrer praktiske ferdigheter under tidspress. Dette kan være en nyttig erfaring, spesielt for de som skal ut i arbeidslivet.

Lærerne påpeker at fagprøven er sentral i yrkesfag og at elevene må lære å mestre at «ting kan gå galt eller godt» i livet. Det ligger også en trygghet i at noen andre enn læreren i klasserommet vurderer elevene. Selv om lærerne ser verdien av eksamen, er de åpne for alternative vurderingsformer som mappevurdering eller flere prøver gjennom året. Dette vil kunne gi elevene mulighet til å vise progresjon over tid og kan redusere stresset knyttet til én enkelt vurderingssituasjon. Lærerne mener også at flere og mer varierte vurderingssituasjoner kan gi et bredere bilde av elevens kompetanse og bidra til en mer rettferdig vurdering.

Lærerne opplever derimot utfordringer knyttet til trekkordningen i fellesfag. Denne ordningen kan oppfattes som «vanskelig» og kan i verste fall ødelegge for elevens videre progresjon mot opptak til VG2. Lærerne uttrykker bekymring for at eksamen «preger og hemmer læring» i disse fagene, og mener at eksamensformen kan være lite gunstig for noen elever.

Samarbeid om vurdering

Lærerne i HO rapporterer at de arbeider sammen både gjennom formelle strukturer, som faste møter, og gjennom uformelle samtaler. Samarbeidet har som mål å diskutere og utvikle vurderingspraksisen, spesielt med tanke på å sikre en felles forståelse av læreplanen, forskriften og vurderingskriteriene. Denne blandingen av strukturerte og uformelle arenaer gir lærerne mulighet til å justere sin praksis løpende og dele erfaringer fra undervisningen.

I intervjuene kommer det fram at flere lærere opplever utfordringer når det gjelder vurdering av Yrkesfaglig fordypning (YFF). Det er spesielt utfordrende med store variasjoner i hvordan praksisstedene vurderer og at det er ulike forventninger fra næringslivet når det gjelder elevens kompetanse. Dette fører til at vurderingene kan bli ujevne, noe som forsterker behovet for mer samarbeid om enhetlige kriterier for å sikre rettferdighet i vurderingsarbeidet. Det etterspørres mer standardiserte vurderingskriterier for å sikre at elevene vurderes likt, uavhengig av praksissted.

Samarbeidet internt på skolene, spesielt mellom lærere på VG1 og VG2, oppleves som viktig for å utvikle en felles forståelse og praksis for vurdering. De bruker flere ulike arenaer for samarbeid, som blant annet avdelingsmøter, gruppemøter, karaktermøter, og pedagogiske ledersamlinger. Disse møtene gir lærerne muligheten til å diskutere vurderingspraksis, karaktersetting og læreplanens krav. Dette samarbeidet sikrer at elevene møter en lik tilnærming til vurdering på tvers av trinnene. Likevel opplever flere lærere at samarbeidet ikke fungerer så godt som det burde, og at det oppleves spesielt utfordrende å samkjøre praksis i større nettverk eller på tvers av skoler. Flere av lærerne etterlyser hyppigere og mer strukturerte møtearenaer.

Støtte fra skoleledelse

Flere av lærerne ved HO opplever at skoleledelsen spiller en sentral rolle i å legge til rette for vurderingsarbeid. De setter opp faste møteplaner og legger til rette for samarbeid blant lærerne. Ved noen skoler har ledelsen har også ansatt avdelingsledere og pedagogiske ledere som fungerer som støttepersoner for lærerne, spesielt i arbeidet med vurdering og bruk av kompetansemål. Selv om skoleledelsen sørger for grunnleggende rammer, påpeker flere lærere at det fortsatt er behov for mer strukturert støtte og tydeligere retningslinjer for bruken av vurderingsverktøy og kjennetegn på måloppnåelse.

Teknologiske verktøy som Visma InSchool og Canvas muliggjør digital dokumentasjon og tilbakemelding og effektiviserer vurderingsprosessen. I tillegg finnes det fagnettverk som fremmer samarbeid og deling av beste praksis blant lærere på tvers av fag og avdelinger. Likevel er det flere lærere som påpeker behovet for mer støtte og opplæring fra skoleledelsen, spesielt for å sikre en enhetlig bruk av vurderingsverktøy og kriterier. De mener at det kan bidra til å redusere variasjoner i vurderingspraksis og øke kvaliteten på tilbakemeldingene til elevene.

Mange lærere deltar aktivt i fagnettverk på tvers av skoler. Disse nettverkene er viktige arenaer for å kalibrere vurderingskriteriene, dele erfaringer og sikre enhetlig bruk av karakterskalaen. Gjennom nettverksmøter kan lærerne diskutere utfordringer knyttet til vurderingsarbeid og dele beste praksis, noe som skaper en større grad av samsvar i vurderingspraksisen på tvers av skoler. Lærerne opplever at deltakelsen i fagnettverk bidrar til å heve kvaliteten på vurderingsarbeidet og gir dem en følelse av trygghet i fastsettingen av standpunkt karakterer.

Elevinvolvering

Lærerne på HO er opptatt av at arbeidet med elevinvolvering er en pågående prosess. Selv om de hele tiden prøver å involvere elevene i vurderingsprosessen, påpeker de at det er viktig å sikre at deres stemmer og egenvurderinger tas på alvor. De har stadige dialoger om hvordan elevene skal få reell innflytelse på hvilke vurderingsformer som brukes. I dette arbeidet mener de at de som lærere må sørge for at vurderingskriteriene er klare og forståelige, og at elevene får opplæring i hvordan de skal bruke disse til å vurdere seg selv og sitt arbeid. Dette sitatet er illustrerende for mangfoldet av metoder for elevinvolvering:

Elevene kan være med på å bestemme vurderingsform noen ganger. På skriftlige prøver kan jeg be dem om å vurdere seg selv og komme med karakterforslag. De fleste greier det godt, men noen er veldig optimistiske på egne vegne og disse er det fint å ta en samtale

med. De lurder på hvordan de ligger an til standpunkt og hva som skal til for å gå opp. [...] Det er viktig å snakke med elevene så de ikke får noen overraskelser slik at de vet hvordan de ligger an. Det kan jo gå litt tid mellom prøver. Jeg har ofte samtale med elevene etter en vurderingssituasjon, men også løpende underveis. Jeg er jo mye sammen med dem.

Elever som velger HO, kan ha svakt karaktergrunnlag fra ungdomsskolen, og flere av disse elevene opplever kompetansemålene som vanskelige å forstå. Det gjør det utfordrende for dem å vurdere seg selv på en realistisk måte. Lærerne peker på at det er nødvendig med mer veiledning og opplæring i hvordan både lærere og elever kan bruke vurderingskriterier for å skape en mer transparent og forståelig vurderingsprosess. Elever i HO kan ha ulike forutsetninger, og dette krever at lærerne jobber mer målrettet med å tilpasse vurderingspraksisen slik at den blir mest mulig inkluderende og forståelig for alle.

En annen utfordring som lærerne opplever i HO, er at enkelte elever kan ha stort sprik i sin kompetanse. De kan prestere svært godt på noen aspekter av faget, mens «andre deler står til en annen karakter», noe som gjør det vanskelig å gi en helhetlig vurdering, da forskjellige deler av elevenes kompetanse står til ulike karakterer. Dette understreker viktigheten av å bruke varierte vurderingsformer og å ha en tydelig kommunikasjon med elevene om hva som kreves for å oppnå høy måloppnåelse.

VEDLEGG 6 SAMFUNNSFAG

Anne Holten Kvistad og Stine Aarønes Angvik

OPPSUMMERING

I denne fagrapporten presenteres et utvalg læreres erfaringer med vurderingsarbeid i samfunnsfag på ungdomstrinn. Utvalget består av tolv lærere fordelt på tre skoler. Under følger en kortfattet oppsummering av hva lærerne rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, og når det gjelder praksis og prosesser knyttet til standpunktvurdering ved skolene.

- Lærerne har god kjennskap til vurderingstekstene i læreplanen og kjennetegn på måloppnåelse i samfunnsfag. Lærerne mener at dokumentet «kjennetegn på måloppnåelse» er vagt formulert og lite egnet for bruk med elever.
- Kompetansemålene i faget oppleves som komplekse, sammensatte og lite elevnære. Enkelte kompetansemål er så omfattende og ambisiøse at de blir utfordrende å bruke i vurderingsarbeid og for vanskelige å omsette til tydelige mål og kriterier som er forståelige for elever. Lærerne har et ønske om færre, enklere og tydeligere kompetansemål i faget.
- Standpunkt karakterene i faget har en tendens til å samle seg på midten av skalaen. Samfunnsfag er et sammensatt fag med flere fagdisipliner, og de fleste elever viser kompetanse i deler av faget.
- Elevmedvirkning i vurdering er ønsket, men lærerne uttrykker at det er vanskelig å få til på en systematisk måte på grunn av lavt timetall og en ambisiøs læreplan.
- Lærerne vektlegger et bredt vurderingsgrunnlag med varierte vurderingsformer. Kunstig intelligens stiller imidlertid krav til nytenkning når det gjelder bruk av skriftlige besvarelser som vurderingsgrunnlag.
- Lærerne framhever praksisen med «prøvemuntlig», der lærere på ulike trinn ofte er medsensorer for hverandre, som positivt for tolkningsfellesskapet på skolenivå. De trekker også fram at oppdrag som ekstern sensor på andre skoler anses som nyttig for kalibrering mellom skoler.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Lærerne på de tre skolene kjenner til tekstene og ser kompetansemålene, kjennetegn på måloppnåelse og vurderingstekstene i nær sammenheng.

Lærere opplever at kompetansemålene og kjennetegnene på måloppnåelse i samfunnsfag ofte er komplekse og til dels generelt formulert. Dette innebærer at lærerne må gjøre et omfattende tolknings- og oversettelsesarbeid for å konkretisere hva målene innebærer i faget, og for å gjøre dem om til forståelige læringsmål og vurderingskriterier i undervisningen. Slikt oversettelsesarbeid er en del av lærerprofesjonens ansvar, men kan samtidig være tidkrevende og krevende å gjennomføre uten støtte fra faglige fellesskap. Mange av lærerne mener at formuleringene er svulstige og lite egnet for direkte bruk med elever, og at det er nødvendig å oversette dem til et

mer elevvennlig språk når de arbeider med planlegging av undervisning og vurdering. Lærerne peker særlig på formuleringen om handlingsalternativer i kjennetegnene på måloppnåelse som uklar. For øverste nivå står det: «Eleven forklarer handlingsalternativer i fortid, nåtid og fremtid og konsekvenser av disse, og identifiserer nyanser.»

Vurderingstekstene om underveisvurdering oppleves som mer konkrete og anvendelige i praksis. De brukes i planlegging av undervisning og fungerer som en «rettesnor» eller en «huskeliste» for lærernes vurderingsarbeid. Samtidig er det variasjon i hvordan disse tekstene brukes. En lærer rapporterer for eksempel at hen heller foretrekker å se til forskrift til opplæringsloven, altså karakterbeskrivelsene og det som står om underveis- og sluttvurdering i loven.

Det er bred enighet blant lærerne om at kjennetegn på måloppnåelse burde finnes i alle fag, da dette kan bidra til mer helhetlig og lik vurderingspraksis. Samtidig etterlyses tydeligere og mindre sammensatte kompetansemål. Dette kan redusere arbeidsbelastningen for lærerne og skape mer konsistent undervisning og vurdering på tvers av skoler.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERINGEN?

Lærerne uttaler at samfunnsfag er et sammensatt fag med flere fagdisipliner, og at dette kan være utfordrende i arbeidet med standpunktvurdering i faget. Noen av lærerne poengterer at det kan være utfordrende å sette standpunkt karakter om man blir «kastet inn» som lærer på 10. trinn, og at det er en stor fordel å ha fulgt elevene helt fra 8. trinn. Lærerne poengterer at selv om visse tema avsluttes – for eksempel at man «gjør seg ferdig» med andre verdenskrig på 9. trinn – så skal elevene kunne trekke med seg kunnskapen videre i nye faglige sammenhenger. På slutten av 10. trinn legges det opp til mer komplekse vurderingssituasjoner som favner om flere kompetansemål.

Lærerne bruker et bredt spekter av vurderingsformer, men noen lærere påpeker at KI-teknologi utfordrer troverdigheten i noen av dem, spesielt skriftlige hjemmeoppgaver. På en av skolene har dette ført til en viss tilbakevending til mer kontrollerte prøver, selv om det er en bevissthet om deres begrensninger. Refleksjon, muntlige vurderingsformer og observasjon er sentralt i vurderingspraksisen i faget.

Bruk av karakterskalaen og fastsetting av standpunkt karakter

Lærere ved de tre skolene beskriver standpunktvurdering i samfunnsfag som en kompleks og krevende prosess. De opplever at en ambisiøs læreplan, kombinert med begrenset undervisningstid, gjør det utfordrende å vurdere elevenes samlede kompetanse på en presis og rettferdig måte. Lærerne legger vekt på at standpunkt karakteren ikke skal være et snitt, men et uttrykk for helhetlig kompetanse og evne til å se sammenhenger. Lærerne er opptatt av å fange opp faglig refleksjon gjennom hele året, både i muntlige og skriftlige sammenhenger, og gjennom observasjon i timene.

Lærerne peker på at samfunnsfag er et «overbelastet» fag med brede og sammensatte kompetansemål, i tillegg til at faget har fått et særlig ansvar for digitale ferdigheter. Det er enighet om at man ikke kan få til å vurdere alle kompetansemål mot slutten av 10. trinn. På en av skolene

forsøker lærerne å løse utfordringen ved å fokusere på kjerneelementene i vurderingssituasjoner mot slutten av året. Lærerne på en annen skole påpeker at de minner elevene på at vurdering skjer kontinuerlig, for eksempel ved leksegjennomgang og i gruppediskusjoner, og ikke bare i prøvesituasjoner. Det er enighet blant lærerne om at skriving som vurderingsgrunnlag må skje i kontrollerte former på skolen, og noen lærere påpeker at KI har bidratt til å gjøre det mer utfordrende å få til gode vurderingssituasjoner på tampen av året.

Alle de tre skolene har karakterdempet vurdering. Lærerne på alle skolene er enige om at standpunktkarakterene i samfunnsfag har en tendens til å samle seg rundt midten, og at karakterene 3 og 4 er «brede». Dette henger sammen med fagets egenart, med stor tematisk variasjon, noe som gjør at de fleste elever kan mestre noe og dermed lettere kan oppnå karakterene 3 og 4. På den tredje skolen har de et elevgrunnlag med mange som presterer godt. Der er det flere firere enn treere. På tvers av skolene er lærerne enige om at karakteren 1 brukes svært sjeldent eller aldri. Lærerne på to av skolene peker på at kravene i kompetansemålene er svært høye, noe som gjør det vanskelig å oppnå karakteren 6. En lærer uttrykte at «dersom man leser kompetansemålene i faget helt konkret, er det ingen som skal ha karakteren 6». På den tredje skolen svarer lærerne noe annerledes. På grunn av det høye sosioøkonomiske nivået i området der skolen ligger, er det mange som får karakteren 6. Lærerne ved skolen har fått tilbakemeldinger under muntlig eksamen om at det oppleves uvant for lærere fra andre skoler at så mange elever får karakteren seks i faget.

Det er uenighet om hvorvidt karakterskalaen brukes likt av lærerne. Flere opplever at kalibrering gjennom prøvemuntlig bidrar til en mer enhetlig praksis, men det finnes ulikheter, særlig knyttet til karaktersetting ved halvårsvurderingen. Noen lærere benytter betegnelser som «svake» og «sterke» karakter, mens andre mener skalaen bør begrenses til de seks satte trinnene. Ved en av skolene anvendes karakterskalaen likt på trinnet, men det er usikkerhet knyttet til hvorvidt praksisen er lik på tvers av hele samfunnsfagseksjonen. De mest krevende diskusjonene gjelder IV og karakterene 1 og 2, ettersom «kjennetegn på måloppnåelse» krever at elevene viser en del kompetanse for å oppnå karakteren 2.

Samarbeid om vurdering

Lærerne fremhever betydningen av kollegialt samarbeid i vurderingsarbeidet. De har alle medsensorordninger ved prøvemuntlig og jevnlig møtepunkter for planlegging av undervisning. Det varierer imidlertid hvor systematisk det arbeides med vurdering i fellesskap. Noen lærere opplever at vurdering i stor grad blir opp til den enkelte, mens andre har etablert ulike former for samarbeid, fra diskusjoner om vurderingskriterier til mer omfattende felles praksis. Et eksempel på mer omfattende samarbeid ved en av skolene er at alle elever på trinnet gjennomfører den samme prøven én gang i halvåret. Lærerne leser og diskuterer besvarelser opp mot «kjennetegn på måloppnåelse» og viser deretter en god besvarelse som eksempel for elevene.

På to av skolene fremhever lærerne at de er opptatte av at de skal ha faglig frihet og autonomi selv om de samarbeider om vurdering. Et eksempel på dette er at de sammen setter opp forslag til aktuelle vurderingsformer, og at lærerne kan velge innenfor dette utvalget.

Støtte fra skoleledelse

Når det gjelder støtte fra skoleledelsen i vurderingsarbeid, er lærerne enige i at prøvemuntlig har høy prioritet og at ledelsen på skolene legger til rette for dette. Lærerne mener det er viktig at man får være intern sensor for hverandre, fordi det bidrar til at man får en felles forståelse om vurdering i faget. Det å være sensor på eksamen ved andre skoler ser de også på som en viktig kalibreringsmulighet. På to av skolene er det lite samarbeid på tvers av skoler eller i fagnettverk. Unntaket er samlinger i forbindelse med muntlig eksamen. Den tredje skolen skiller seg ut. Her har ledelsen lagt opp til fagsamarbeid mellom skoler i kommunen for å diskutere vurdering. Dette handler ifølge lærerne om å trygges i et samfunn i endring, der de ser at det klages mer på karakterer enn tidligere.

Elevinvolvering

Elevene involveres i vurderingsarbeidet gjennom uformelle og formelle vurderingssamtaler der de får vite hvordan de ligger an og reflekterer over sin egen læring. Det framkommer i intervjuene at det er vanlig praksis at elever av og til får velge vurderingsform i faget, altså hvilket format de ønsker å vise sin kompetanse gjennom. Dette kan for eksempel være å velge mellom podcast, prøve og fagtekst. Lærerne trekker imidlertid fram dette som svært ressurskrevende og vanskelig med mange elever i klassene. Mange av lærerne uttrykker at de skulle ønske at de var bedre på å involvere elever i vurderingsarbeidet i faget, men at rammene gjør det utfordrende. En annen grunn til at det er lærerne som oftest velger vurderingsform, er at de ønsker å sikre et bredt vurderingsgrunnlag. Elevene får imidlertid gjerne velge tema selv når de skal skrive en fordypningsoppgave. Ifølge noen av lærerne er det utfordrende å involvere elever i å utforme vurderingskriterier i faget. Grunnen til dette er at kompetansemålene i læreplanen er for akademisk formulert.

Karakterpress

Det er noe ulik oppfatning om karakterpress ved de tre skolene. Ved den ene skolen merkes en holdningsendring blant elever etter pandemien. Lærerne mener at flere forventer karakteren 6 enn tidligere. Ved den andre skolen er det generelt lite fokus på karakterer. Lærerne på denne skolen tenker på karakterinflasjon mer som et fenomen man finner på videregående. Ved den tredje skolen, som ligger i et område med høyt sosioøkonomisk nivå, rapporterer lærerne at de kjenner på karakterpress fra foresatte med «tydelige formeninger om hva de unge håpefulle skal få». De snakker imidlertid hovedsakelig om at elevene forventer et visst karaktersnitt, og ikke en bestemt karakter i samfunnsfag. Lærerne opplever at skoleledelsen har tillit til deres vurderinger når det kommer henvendelser fra «sinte foreldre».

VEDLEGG 7 MAT OG HELSE

Marita Byberg Johansen og Henning Fjørtoft

OPPSUMMERING

I denne fagrapporten presenteres et utvalg læreres erfaringer med vurderingsarbeid i faget mat og helse på ungdomstrinn. Datagrunnlaget er fra tre gruppeintervju med åtte mat- og helselærere fra tre skoler.

Under følger en kortfattet oppsummering av hva lærerne rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, og når det gjelder praksis og prosesser knyttet til standpunktvurdering ved skolene.

- Lærerne opplever at faget er klageutsatt. Derfor er de opptatt av å synliggjøre sammenhengen mellom underveis- og sluttvurderingen og tydeliggjøre hva som er forventet av elevene i ulike planer og dokumenter. Når det gjelder vurderingstekstene i læreplanen, er lærerne delte i sin mening om disse. Kjennetegn på måloppnåelse brukes ikke.
- Lærerne samarbeider om planlegging og vurdering, men de har ikke nødvendigvis avsatt samarbeidstid i dette faget. Noen lærere opplever at mat og helse er mindre prioritert enn andre fag. Det varierer i hvilken grad elevene involveres i vurderingsarbeidet.
- Alle lærerne oppgir at de sjelden gir de laveste karakterene (1 og 2) i faget, men noen sier at de heller ikke gir like mange toppkarakterer som før på grunn av at læreplanen er «teoritung» og stiller høye krav til elevene.
- Lærerne har ulike erfaringer med karakterpress. Karakterpresset henger sammen med geografisk og sosioøkonomisk kontekst, foresattes forventninger og manglende kjennskap til læreplanendringer. I tillegg uttrykker noen lærere et generelt press om å sikre vurderingsgrunnlag.
- Lærerne peker på at de opplever at det er store forskjeller i opplæringa mellom skoler, blant annet fordi variasjoner i skolebudsjettene påvirker innkjøp og organisering.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Lærerne har delte meninger om vurderingstekstene. De fleste lærerne opplever disse tekstene som nyttige og mener at felles retningslinjer kan styrke felles bevissthet. Tekstene brukes av noen av lærerne i planlegging av opplæring og i begrunnelser av standpunktkarakterer. På en av skolene har lærerne inkludert tekstene om underveis- og sluttvurdering i årsplanen og gjennomgikk tekstene med elevene ved oppstart. Formålet med det er å bevisstgjøre elevene og forhindre unødvendige klager. Noen av lærerne opplever imidlertid at vurderingstekstene er for generelle.

Når det gjelder kjennetegn på måloppnåelse, er lærerne mer unisont kritiske til disse. Formuleringer som «kunnskap i faget» og «skapende arbeid» trekkes fram av noen lærere som vage og uklare. Lærerne er enige om at språket ikke er tilpasset elevene og at

karakterbeskrivelsene ikke er hensiktsmessige. En lærer sier: «Utrolig mange vanskelige ord som også er vanskelige å bryte ned for elevene. Vi bruker den ikke.» På en annen skole mener lærerne at kjennetegn på måloppnåelse ikke henger sammen med kompetansemålene og teksten om faget.

I tillegg til vurderingstekstene og kjennetegn på måloppnåelse, kommenterer lærerne på en av skolene andre aspekter ved læreplanen. Lærerne mener at det er uheldig med korte og abstrakte kompetansemål. Lærerne savner at læreplanen nevner konkrete eksempler på produksjonsmetoder og næringsstoffer, samt beskrivelser av fagspesifikk bruk av grunnleggende ferdigheter, som for eksempel å lese og anvende oppskrifter.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERING?

Alle lærerne peker på at faget er avsluttende på 9. trinn på de skolene som er med i studien, og dermed det første faget som elevene får standpunktkarakter i. Lærerne opplever faget som «klageutsatt». Lærerne er derfor bevisste på å synliggjøre sammenhengen mellom underveis- og sluttvurderingen og er opptatt av å tydeliggjøre hva som er forventet av elevene både gjennom veiledning av dem og i ulike planer og dokumenter de lager selv, som for eksempel vurderingsskjema. En lærer sier i intervju: «Jeg jobber nesten bedre med det [vurdering] i mat og helse enn i andre fag, egentlig. Fordi det er så kort tid, og det er i 9. trinn, uten standpunktkarakter i andre fag. Hvis en ikke er fornøyd med det, er det enklere å sende inn klage. Da tenker du over ting ekstra nøye.»

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Alle lærerne oppgir at de sjeldent gir de laveste karakterene (1 og 2). På noen av skolene snakker lærerne om at de opplever at kravene til elevene har økt med den nye læreplanen, blant annet med mer fokus på teori, slik at det ikke er like mange seksere som før. Flere av lærerne snakker også om at det er vanlig å «holde igjen» på karakteren som gis etter første halvår. Noen av lærerne oppgir å bruke glidekarakterer (med pluss, minus og skråstrek), enten muntlig eller skriftlig, som et signal til elevene på hvordan de ligger an.

Lærerne er opptatt av å gi elevene mange og varierte muligheter for å vise kompetanse, og det er rom for tilpassing dersom en vurderingssituasjon ikke fungerer optimalt for eleven: «Teoriprøve kjører vi digitalt, men jeg har elever som har så store lese- og skrivevansker at vi sitter sammen ved elevens maskin, og jeg leser spørsmål høyt og puncher svar for eleven. (...) Prøver å få frem elevens kompetanse på kjøkkenet hvis man stresser seg fryktelig opp for den fagsamtalen. Jeg har i noen få tilfeller kjørt en fagsamtale to ganger». På to av skolene uttrykker lærerne at det praktiske arbeidet teller mest, mens på en av skolene er lærerne innad uenige om vektingen mellom det praktiske og det teoretiske. En lærer uttrykker et ønske om tydeligere føringer for vekting mellom praktisk og teoretisk arbeid: «Jeg er enig i at det er et praktisk-estetisk fag, men vi må forholde oss til at læreplanen har ganske mange mål som du ikke får sett i elevens praktiske arbeid.» Lærerne opplever at det er vanskelig å bestemme standpunktkarakter hvis en elev for eksempel er god i det praktiske arbeidet, men svakere i teori.

Flere lærere peker på at de kommer tett på elevene på kjøkkenet og kan veilede dem og gi tilbakemeldinger fortløpende. Mange trekker fram at de observerer og gjerne noterer underveis i eller etter undervisningsøktene. En lærer peker også på at faget har større fokus på prosess enn resultat:

Det er viktig å snakke om vurdering i et praktisk-estetisk fag som mat og helse, fordi kompetansen som vises er ferskvare. Den vises her og nå, og så har du ikke noen bevis i ettertid. Noen flåser og sier sånn «ja, fordi de ble spist opp», men det er ikke bare det. Det er på en måte all handling som skjer underveis. Det er litt som i kroppsøving også. Hvis du snur deg vekk og ikke ser det spenst hoppet, så har det skjedd. Men det er også at vi er så tydelige underveis i vurderingen til elevene på at prosess er viktig. Du sier at kjøttkakene ser fine ut og smaker fint, men hvis de har slurvet med hygien og har kryssmittet ...

På to av skolene brukes «Mystery basket» som en større praktisk oppgave, der elevene skal lage et måltid basert på noen bestemte matvarer og kriterier som lærerne bestemmer. Formålet er at elevene skal få vise kompetanse gjennom både planlegging og lagging av et måltid. Oppgaven er gjerne lagt til slutten av det siste semesteret fordi lærerne er bevisste på at de skal sette standpunktkarakter basert på det eleven viser av samlet kompetanse.

Samarbeid om vurdering

Lærerne samarbeider både om planlegging og vurdering, og har i fellesskap utviklet flere dokumenter, blant annet årshjul og vurderingskriterier. Lærerne på to av skolene har ikke avsatt tid til samarbeid, og må dermed oppsøke hverandre utenom undervisnings- og møtetid. På en av disse skolene uttrykker lærerne at praktisk-estetiske fag er «stemoderlig behandlet» sammen med valgfag og andre fremmedspråk. På den tredje skolen har lærerne ukentlig fagteam-tid og samarbeider generelt tett: «Vi føler at vi er veldig like i hva vi ser etter og hva vi sier til elevene».

Alle lærerne har hatt mulighet til å delta i fagnettverk, men de har delte meninger når det gjelder om dette har vært fruktbart eller ikke. Gjennom deltagelse i fagnettverk har lærerne også blitt oppmerksom på at det er store forskjeller mellom skolene i forståelse av læreplan, vurderingspraksis, ressurstilgang og så videre. Flere av lærerne peker også på at deltagelse på fagdager i nettverk ikke prioriteres, fordi det oppleves som krevende å være borte fra undervisningen. På en av skolene oppgir lærerne at de har tilgang til en digital plattform som er tilgjengelig på tvers av skolene i kommunen, og som de kan bruke til å kommunisere og samkjøre med lærere på andre skoler.

Støtte fra skoleledelse

Lærerne har ulike opplevelser av i hvilken grad mat og helse er et prioritert fag eller ikke, både med tanke på interesse fra ledelsen og med tanke på å få avsatt tid til samarbeid. På to av skolene legges det til rette for samarbeid om vurdering, men ikke nødvendigvis i mat og helse. På den tredje skolen har de generelt hatt mye fokus på vurdering. Lærerne har avsatt tid jevnlig til samarbeidsmøter og føler seg av den grunn heldige sammenlignet med mat og helse-lærere på andre skoler.

Elevinvolvering

Lærerne er delte i hvor stor grad elevene involveres i vurderingsarbeidet. På to av skolene fyller elevene alltid ut egenvurdering eller logger der de må vurdere og reflektere over eget arbeid. Flere lærere sier også at de gjennomfører en karakterprat med elevene, der elevene blant annet får si hvordan de selv opplever at de ligger an i faget. En av lærerne sier om karakterpratene: «Det er overraskende mange som er mer enige med meg, enn før. Det kan være at jeg har blitt flinkere til å gi dem god tilbakemelding underveis, så de skjønner hva de er flinke til, og hva de kan jobbe mer med». På en av skolene sier lærerne at de gjennomfører karakterprat med elevene, men at de ellers er for dårlige på å involvere elevene i vurderingsarbeidet. De savner eksempler på hvordan både egenvurdering og hverandrevurdering kan gjennomføres.

Karakterpress

Lærerne har ulike erfaringer når det kommer til karakterpress. På to av skolene har de erfaring med at geografisk beliggenhet og sosioøkonomisk status i området kan påvirke i hvilken grad foresatte har forventning om høye karakterer eller ikke. Forventninger fra foresatte kan også henge sammen med at foresatte ikke kjenner endringer i læreplanen fra da de selv gikk på skolen eller at de forventer at innsats skal telle med i vurderingen: «Så lenge bollene glinser og hever så er det en 5-er, liksom. Vi måler ikke innsats slik som det gjøres på kjøkkenet hjemme». På den ene skolen opplever lærerne et press om å sørge for vurderingsgrunnlag, fremfor å gi «Ikke vurderingsgrunnlag» (IV) eller oppfordre foresatte til å søke om fritak. De knytter dette til å unngå frafall og å sikre gjennomføring for alle elever: «Jeg synes det presset har blitt større, på å få de gjennom (..) Jeg kjenner på at det, hvis jeg skal sette IV, føles som at jeg ikke har gjort jobben min. For det er mitt ansvar å skaffe grunnlag, og så har jeg ikke klart å skaffe det». De opplever at dette presset kommer fra flere kanter, både fra skoleledelsen, men også fra foresatte, kommunen og samfunnet generelt.

VEDLEGG 8 KROPPSØVING

Runa Næss Korshavn, Lise Vikan Sandvik og Egil G. Gjølme

OPPSUMMERING

Under følger en oppsummering av hva kroppsøvlingslærerne ved fire videregående skoler (18 lærere) og tre ungdomsskoler (10 lærere) rapporterer om når det gjelder praksiser og prosesser knyttet til standpunktvurdering og bruken av vurderingstekstene i læreplanen for kroppsøvfingsfaget.

- Skolene vektlegger innsats, sosial kompetanse og personlig utvikling i vurderingen, men uklare læreplanformuleringer gir stort tolkningsrom for lærerne, noe som gjør praksisene svært ulike både innad og på tvers av skoler.
- Kompetansemål og kjennetegn på måloppnåelse oppleves som generelle og vanskelige å bruke i praksis. Hvorvidt og hvordan tekstene benyttes, varierer mellom skolene og innad i kollegier.
- Skolenes organisering av vurderingsarbeidet varierer. Et par skoler har systematisk samarbeid og kalibrering, mens flere skoler preges av individuelle vurderingspraksiser, opplevd tidspress og manglende støtte fra ledelse.
- Lærerne vektlegger kartlegging av elevenes kompetanse med utstrakt bruk av egenvurderingsskjema, med mindre fokus på tilbakemeldinger til elevene og støtte over tid.
- Det er krevende å få elevene til å forstå hva som vurderes i faget.
- På de videregående skolene har karakterpresset i faget økt. På yrkesfag er det et press fra ledelsen om å gi ståkarakter, mens det på studiespesialiserende kommer fra elever og foreldre som ønsker høye karakterer. Når innsats vektlegges sterkt i læreplanen, blir høye karakterer tilgjengelige for flere og det blir vanskelig å argumentere imot.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

På tvers av skolene er det tydelig at vurderingstekstene i læreplanen og kjennetegn på måloppnåelse er kjent blant lærerne. Til tross for dette benyttes de i liten grad som et aktivt verktøy i vurderingsarbeidet. Tekstene oppleves av flere som språklig vanskelige, vage og lite konkrete, noe som gjør dem lite anvendelige i praksis. På tvers av skolene fungerer de først og fremst som utgangspunkt for faglige diskusjoner eller som kontrollverktøy i etterkant av karaktersetting. Mange lærere baserer vurderingen på kompetansemål og egen erfaring, og etterlyser tydeligere nasjonale føringer, for eksempel knyttet til hvordan formuleringer som «å delta», «til en viss grad» og «forutsetninger» skal forstås og operasjonaliseres. Det er på samme måte vanskelig å vurdere kompetanse knyttet til «å reflektere over», «delta» eller «gjennomføre» på en reliabel måte. Enkelte kollegier lykkes med å konkretisere tekstene og utvikle lokale kjennetegn på måloppnåelse, men flere skoler opplever at vurderingstekstene verken gir støtte eller retning.

Selv om noen lærere setter pris på handlingsrommet og metodefriheten tekstene gir, mener mange at tolkningsrommet er utfordrende og at det skaper usikkerhet og for stor variasjon. Lærerne etterlyser tydeligere tekster eller støtte til tolkningen av tekstene, og opplever det som utfordrende at det i stor grad overlates til den enkelte å operasjonalisere tekstene. Samlet viser funnene at vurderingstekstene i liten grad fungerer som operative verktøy per i dag, og at behovet for mer konkrete, tydelige og anvendbare beskrivelser er stort.

HVA KJENNETEGNER PRAKSISER OG PROSESSER I STANDPUNKTVURDERINGEN?

Lærerne er gjennomgående bevisste på sammenhengen mellom underveisvurdering og standpunktvurdering. De forsøker å arbeide med dette på en rettferdig og transparent måte med elevenes utvikling som utgangspunkt. Samtidig varierer praksisene betydelig mellom skoler og lærere. For samtlige lærere skjer vurderingen primært gjennom kontinuerlig observasjon i undervisningen. Store elevgrupper, mange «myke» kompetanser og varierende grad av struktur gjør det krevende å opparbeide et solid vurderingsgrunnlag. Noen skoler arbeider systematisk med faste vurderingssamtaler og egenvurderingsskjema og har tydelige planer som skaper stabilitet og progresjon i arbeidet, mens andre arbeider mer uformelt og tilfeldig.

På tvers av skolene trekkes individuell tilbakemelding frem som nøkkelen i arbeidet med å skape sammenheng mellom underveis- og standpunktvurdering, ofte gjennom halvårssamtaler og løpende dialog i timene. Samtlige lærere opplever det som utfordrende å gjøre vurderingskriteriene forståelige for elevene, og mange lærere beskriver at elevene ofte først sent i skoleløpet forstår hva de faktisk vurderes i. For å kommunisere og tydeliggjøre hva vurdering i faget handler om, bruker noen lærere kjennetegn på måloppnåelse, egenvurderingsskjema og tydelige læringsmål. Til tross for dette er det en felles utfordring å endre elevenes forståelse av faget. Manglende kontinuitet, som følge av at lærere ikke følger klasser over flere år, store elevgrupper og uklare læreplanformuleringer, forsterker utfordringen. Det bidrar til at sluttvurderingen ofte blir en helhetsvurdering basert på profesjonelt skjønn mer enn et fullstendig og tydelig dokumentert vurderingsgrunnlag.

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Karakterskalaen 1–6 oppleves som vanskelig å bruke i kroppsøving. En del av begrunnelsen handler om at retningslinjene og støtteverktøy som «kjennetegn på måloppnåelse» er uklare og ikke hjelper lærerne med å skille de ulike karaktertrinnene. Vurderingsgrunnlaget baseres derfor i stor grad på helhetsinntrykk og skjønn. Noen lærere utvikler egne kriterier, men karaktertrinnene, og da særlig midtsjiktet, beskrives som bredt og krevende å skille.

Det er forskjeller mellom skolene og lærerne når det gjelder bruken av karakterskalaen. Noen få lærere bruker hele skalaen, mens de fleste aldri setter karakterene 1 og 2. Flere forklarer at skalaen i praksis starter på 3 eller 4. Karakterene 4 og 5 dominerer på tvers av skoleslag og trinn, og beskrivelsen av hvordan 6-eren anvendes, varierer. Hos noen henger den fortsatt høyt, mens andre opplever at også den karakteren har økt i antall. Det forklares at særlig 6-eren er vanskelig å avgrense og begrunne, og at verbene i læreplanen gjør det utfordrende å identifisere og sette ord på forskjellen mellom nivåene på toppen av skalaen. Samarbeid og kalibrering oppleves som nyttig

der det forekommer, men er ved mange skoler uformelt og ujevnt forankret. Dette gjør at karaktersettingen og bruken av skalaen varierer både mellom skoler og lærere.

Ved de videregående skolene påvirkes bruken av karakterskalaen også av hvordan skoler forstår og praktiserer krav til vurderingsgrunnlag. På yrkesfag fører sterke føringer om at elever skal stå til at lærerne strekker skalaen nedover, mens på studieforberedende program fører forventninger om høye karakterer til at midtsjiktet utvides og at 5- og 6-områdene glir over i hverandre. Dette fører til økt behov for nasjonale avklaringer av hva som skiller nivåene i skalaen.

Samarbeid om vurdering

Intervjuene fra skolene viser at samarbeid om vurdering i kroppsøving varierer betydelig. Faget fremstår som mindre prioritert i skolens formelle strukturer sammenlignet med andre fag, og samarbeidet skjer i ulik grad og form.

Ved flere skoler beskrives samarbeidet som uformelt, gjerne gjennom samtaler på kontoret, i gangene eller når lærere underviser samtidig. Dette gir rom for praktiske avklaringer og sporadisk faglig refleksjon som flere lener seg på. Systematiske rutiner og arenaer hvor vurderingsarbeid kan foregå mer formelt, mangler ved flere skoler. Lærerne uttrykker behov for mer struktur og tid til å utvikle felles vurderingsforståelse knyttet til kompetansemål og ferdigheter i faget. Ulikheter i praksis oppdages sjelden, og samkjøring avhenger ofte av enkeltpersoners initiativ.

Samtidig er det enkelte skoler som skiller seg ut gjennom å ha et strukturert og systematisk samarbeid om vurdering. Dette gjøres blant annet gjennom å ha faste fagteam med strukturert drøfting av vurdering, kalibrering av karakterer og felles planarbeid. Disse skolene deltar også i regionale nettverk, noe de opplever gir faglig fellesskap og bevissthet om variasjoner mellom skoler. Der samarbeid om vurdering er strukturert og fungerer godt oppleves det som avgjørende for kvalitet og transparens i vurderingsarbeidet, både for lærerne og elevene.

Med unntak av et par skoler som var uenige om fagets innhold og vurdering, og opplevde dette som et konfliktfylt tema, var det et utbredt ønske om tydeligere prioritering av kroppsøving i skolens utviklingsarbeid og mer tid til samarbeid om vurdering.

Støtte fra skoleledelse

Det er store variasjoner i hvor mye skoleledelsen støtter vurderingsarbeidet i kroppsøving. Mange lærere opplever at faget har lav prioritet sammenlignet med andre fag, og at de står alene om å utvikle og begrunne egen vurderingspraksis. På flere skoler gis det lite konkret støtte og veiledning. Det er få strukturerte møteplasser og begrenset faglig oppfølging. Dette skaper stor grad av autonomi, men også risiko for ulik praksis, manglende kalibrering og faglig ensomhet. Flere lærere uttrykker et tydelig behov for mer systematisk støtte både lokalt og gjennom kommunale fagnettverk.

Skolene som har etablert et godt faglig samarbeid om vurdering, opplever imidlertid at ledelsen legger til rette for dette gjennom tilrettelegging for faste møtepunkter, felles faglige dager og en tydelig prioritering av vurdering i fellestid. Der lærerne får tid og rom til å møtes, rapporteres det om bedre samsnacking og mer likeverdig vurderingspraksis. Noen lærere opplever imidlertid også

at tillit fra ledelsen, gjennom at de får jobbe uten unødig innblanding, fungerer som en form for støtte, selv om andre opplever fravær av styring som manglende interesse.

Fagnettverk brukes ujevnt. Flere lærere deltar aktivt og opplever disse som nyttige for å diskutere læreplanen, dele erfaringer og kalibrere vurdering, særlig etter innføringen av LK20. Andre beskriver færre møter, liten kontinuitet eller begrenset utbytte av slike samlinger.

Samlet sett viser materialet at skoleledelsens rolle er avgjørende: Når vurdering i kroppsøving løftes inn i skolens profesjonsfellesskap gjennom planlagte og strukturerte møtepunkter, styrkes både vurderingspraksisen og den kollegiale tryggheten. Når slik støtte uteblir, forblir vurderingsarbeidet i stor grad personavhengig og fragmentert.

Elevinvolvering

Elevmedvirkning er et sentralt prinsipp i læreplanen, men intervjuene viser at elevinvolvering i kroppsøving ofte begrenser seg til informasjon og samtaler om karakterer, heller enn involvering i egen læringsprosess. Mange lærere bruker tid på å forklare hva som kreves for å oppnå ulike karakterer, særlig de høye, og opplever at elevene er sterkt karakterorienterte. Dette gjør at vurderingsdialogen ofte handler om dokumentasjon av kompetanse og måloppnåelse fremfor faglig utvikling.

Flere skoler bruker egenvurdering, særlig i forbindelse med halvårsvurderinger eller større prosjekter. Elevene reflekterer over egen innsats og utvikling, og lærerne opplever at dette bidrar til mer realistisk selvforståelse. Samtidig er det få steder hvor elevene faktisk deltar i samtaler om hva kompetanse i faget innebærer. Der praksisen er mer systematisk, kombineres egenvurdering med faste refleksjonssamtaler, tydelig gjennomgang av mål og vurderingskriterier, og regelmessig dialog i timene. Dette ser ut til å styrke elevenes forståelse av hva som vurderes, og bidrar til at innsats, refleksjon og læring får større plass.

På tvers av skolene beskriver lærerne det som et kontinuerlig arbeid å få elevene til å forstå hva som faktisk vurderes i kroppsøving. Begreper i læreplanen må forklares, læringsmål må brytes ned, og innsats må tydeliggjøres. Likevel oppfatter mange elever at tilstedeværelse er nok for å få en grei karakter i faget. Flere lærere peker på at elevenes og lærernes forståelse av fagets innhold ofte ikke stemmer overens, noe som gjør vurderingsdialogen krevende.

Samlet sett fremstår elevinvolvering som viktig, men fortsatt begrenset til å først og fremst handle om å dokumentere kompetanse fremfor å involvere elevene i refleksjoner om kompetanse, læring og vurdering i faget.

Karakterpress

Lærere ved alle skolene opplever et økende press i karaktersettingen. Dette kommer frem både på ungdomstrinn og i videregående, men fremstår som en mer gjennomgående og uttalt utfordring ved de videregående skolene.

At innsats er definert som en sentral kjernekompetanse i kroppsøving, bidrar til å utvide rommet for hva som kan gi grunnlag for høy måloppnåelse. Når innsats både forstås som en del av kompetansen og samtidig tolkes vidt i praksis, opplever lærerne at det blir vanskeligere å

argumentere mot elevers og foresattes forventninger om høye karakterer. Dette forsterker presset på karaktersettingen og gjør det utfordrende å opprettholde en tydelig og konsistent vurderingspraksis.

På yrkesfaglige studieprogram kommer presset primært fra ledelsen. De ønsker at lærerne skal strekke seg langt for å innhente tilstrekkelig vurderingsgrunnlag, slik at elevene står i faget. Lærerne utfordres på å legge til rette for dette, og tilbys også ekstra ressurser for å innhente et godt nok grunnlag. På yrkesfaglige program er imidlertid press fra foresatte og elever fraværende.

På studieforberedende program er situasjonen annerledes. Her uttrykker mange elever tydelige forventninger om høye karakterer for å komme inn på ønskede studier. Foresatte har også ambisjoner på vegne av egne barn, og kommuniserer og forsterker disse forventningene. Lærerne opplever dette som et press som utfordrer deres faglige vurderinger og karaktersetting. Dette presset forsterkes av at kompetansemålene oppleves som krevende å operasjonalisere, og av en karakterskala der grensene mellom nivåene er vanskelige å avklare. Lærerne beskriver at det krever tydelige begrunnelser og stor pedagogisk trygghet å sette både lave og høye karakterer i faget.

Enkelte lærere beskriver hvordan elever til tider «spiller på samvittigheten» deres, og at skolen i noen tilfeller forsterker presset ved å legge til rette for dialog med foresatte om karakterer også etter at elevene har fylt 18 år. I tillegg peker flere på at digitale plattformer, som Teams, har gjort det enklere å ta kontakt om karaktersetting. Samlet bidrar disse forholdene til et økt press på vurderingsarbeidet og knyttes av lærerne til flere klager og formelle saker etter semesterslutt, selv om omfanget varierer mellom skoler og over tid.

VEDLEGG 9 LEDELSE

Lise Vikan Sandvik og Stine Aarønes Angvik

OPPSUMMERING

Under følger en kortfattet oppsummering av hva skoleledelsen ved tre ungdomsskoler (8 ledere) og fire videregående skoler (17 ledere) rapporterer når det gjelder bruk av vurderingstekstene og kjennetegn på måloppnåelse, samt praksiser og prosesser knyttet til standpunktvurdering ved skolene. Lederintervjuene omfatter ikke bare rektor/skoleleder, men også andre lederroller med ansvar for pedagogisk utvikling og oppfølging av vurderingsarbeid, som for eksempel avdelingsledere/inspektører og tilsvarende funksjoner i skolens ledelsesteam. I videregående skole omfatter utvalget i tillegg ledere med særskilt ansvar for utdanningsprogram og faglig-pedagogisk oppfølging. Det er variasjon mellom skolene i hvordan lederfunksjonene er organisert og benevnt, og vi bruker derfor samlebetegnelsen «skoleledelse» om lederroller med personal- og/eller pedagogisk ansvar.

Fagrapporten skal bidra til å nyansere funnene i hovedrapporten og gi et mer konkret bilde av hvordan vurdering forstås og praktiseres ved skolene.

- Skolelederne rapporterer om variasjoner i tolkning og bruk av vurderingstekstene og kjennetegn på måloppnåelse.
- De ser behov for at ledere i skolen gir sterkere støtte til kollegialt samarbeid om tolkning av disse tekstene og om vurderingsarbeidet i skolen.
- For å sikre reliabilitet ser ledelsen behov for gode strukturer for kalibrering i vurderingsarbeidet, spesielt i ungdomsskolene, hvor slike prosesser er mindre systematiske sammenlignet med de videregående skolene.
- Skoleledere peker på tidspress og ressursmangel som barrierer for god vurdering. Systematisk støtte og tilrettelagt tid er avgjørende for å håndtere vurderingsarbeidet.
- Transparens er avgjørende for at elever og foresatte skal forstå vurderingsgrunnlaget og ha tillit til vurderingsarbeidet.
- Skolelederne er opptatt av å utvikle strukturer som involverer elevene aktivt i vurderingsarbeidet.

HVORDAN BENYTTES VURDERINGSTEKSTENE I LÆREPLANENE OG KJENNETEGNENE PÅ MÅLOPPNÅELSE?

Et særtrekk er at både ungdomsskoler og videregående skoler har varierende tilnærminger til bruken av vurderingstekstene og kjennetegnene på måloppnåelse. Det synes å være klare forskjeller mellom skoleslagene når det gjelder systematikk, implementering og forståelse av vurderingsarbeidet. Noen lærere bruker dem aktivt gjennom året, mens andre bare anvender dem ved eksamen, klagesaker eller ved behov for presiseringer. Kjennetegn på måloppnåelse brukes i varierende grad til lokal tilpasning, og lederne rapporterer at mange lærere opplever at de nasjonale kjennetegnene er for generelle. Flere skoler mangler felles føringer eller systemer som sikrer enhetlig bruk, og mange lærere lager egne lokale kriterier.

Ledere i ungdomsskolene beskriver et sprik i bruk av vurderingstekstene. Noen skoler arbeider systematisk med dem i kollegiet, mens andre kun har hatt sporadiske gjennomganger, gjerne i oppstarten av skoleåret. I flere tilfeller er vurderingstekstene først og fremst et støtteverktøy i spesifikke situasjoner, som i klagesaker og til muntlig eksamen. Lederne peker også på utfordringer knyttet til at kjennetegnene på måloppnåelse kan bli for generelle. Det kan hemme elevinvolvering, da elevene har problemer med å forstå og anvende kriteriene i egen læring. Mange skoler jobber med å utvikle lokale tilpasninger som bedre speiler elevenes progresjon.

Også i videregående opplæring er bruken av vurderingstekstene varierende. Flere skoler har utviklet egne systemer eller tilpassede kjennetegn, ofte for å tydeliggjøre hva som skiller de ulike nivåene i karakterskalaen. Lærere i enkelte fag samarbeider om felles vurderingspraksis, mens andre fagmiljøer arbeider mer individuelt. Vurderingstekstene brukes også i forbindelse med IV-vurderinger og klager. Noen skoler avlaster lærerne med dokumentasjon gjennom administrativ støtte.

HVA KJENNETEGNER PRAKSIS OG PROSESSER I STANDPUNKTVURDERING?

Ungdomsskolene beskriver ulike modeller for å sikre sammenheng mellom underveis- og standpunktvurdering. Noen skoler har tydelige strukturer og faste arenaer for samarbeid på trinn- og fagnivå, mens andre har mindre systematisk praksis. Variasjonene gjelder særlig bruken av felles vurderingsrutiner, graden av elevinvolvering og hvordan lærerne tolker og bruker vurderingstekstene. Ressurs- og tidsmangel løftes fram som en gjennomgående utfordring, spesielt når det gjelder å gjennomføre fagsamtaler og gi oppfølging som støtter elevenes progresjon. Flere skoleledere peker på at store forskjeller i fagmiljøenes størrelse påvirker mulighetene for kalibrering og faglig støtte.

Videregående skoler har generelt mer formaliserte strukturer for vurderingsarbeidet. Seksjonsmøter, årshjul og digitale systemer gir rammer for planlegging og oppfølging, og mange fagmiljøer arbeider systematisk med å utvikle vurderingssituasjoner som viser bredde i kompetanse. Samtidig er det fortsatt store forskjeller mellom fag, særlig når det gjelder praktisering av karakterskalaen og sammenhengen mellom underveis- og standpunktvurdering. Yrkesfaglige utdanningsprogram beskrives som mer praksisnære, mens studiespesialiserende fag i større grad bygger på tradisjonelle vurderingsformer. Flere skoleledere legger vekt på at kvaliteten i vurderingsarbeidet er avhengig av tid til samarbeid og kollegial støtte.

Bruk av karakterskalaen og fastsetting av standpunktkarakter

Skolelederne beskriver forståelsen og bruken av karakterskalaen, samt prosessen for fastsetting av standpunktkarakter, som et av de mest krevende områdene i vurderingsarbeidet. Selv om lærerne i både ungdomsskole og videregående skole har god kjennskap til regelverket, varierer tolkningen av nivåene og praksisen for fastsetting av karakterer betydelig mellom fag, lærere og skoler. Dette skaper utfordringer knyttet til både reliabilitet, rettferdighet og tillit til vurderingsarbeidet.

Digitale plattformer brukes i økende grad til dokumentasjon, kommunikasjon og strukturering av vurderingsarbeidet. Dette gir bedre oversikt for både elever og lærere, men skaper også nye

utfordringer, blant annet knyttet til bruk av KI. Mange lærere uttrykker tidspress, og vurderingsarbeidet beskrives som mer omfattende enn ressurstilgangen tillater.

Ungdomsskole

I ungdomsskolen oppstår variasjonene hovedsakelig knyttet til to forhold: forståelsen av nivåskiller i karakterskalaen og vurderingen av når vurderingsgrunnlaget er tilstrekkelig for karaktersetting. Flere skoleledere peker på at lærerne opplever det vanskelig å skille presist mellom mellomnivåene i skalaen, særlig mellom karakterene 3 og 4 og mellom 4 og 5. Lærerne legger ulik vekt på deler av kompetansen, og enkelte vurderer helhetlig progresjon, mens andre tar utgangspunkt i elevenes siste eller sterkeste prestasjoner. Dette fører til forskjeller i hvordan elevens kompetanse blir tolket og dokumentert.

Bruken av karakteren 1 og IV beskrives også som varierende. Noen lærere bruker IV tydelig og konsekvent når grunnlaget er for svakt, mens andre strekker seg langt for å gi elever nye muligheter før karakteren fastsettes. Dette kan gi elevene svært ulike rammebetingelser, avhengig av lærerens praksis eller skolens kultur for vurderingsgrunnlag. Flere skoleledere fremhever at lærere i liten grad får systematisk støtte i fagmiljøet til å utvikle en felles forståelse av karakterskalaen, særlig i fag der det kun er én lærer som underviser faget på skolen. Dette gjør kalibrering krevende og øker risikoen for individuell praksis.

Selve fastsettingen av standpunktkarakter oppleves som en helhetlig vurdering, men er samtidig sårbar når vurderingsgrunnlaget består av mange små oppgaver eller mangelfull dokumentasjon. Lærerne må i større grad støtte seg på profesjonelt skjønn, og noen opplever usikkerhet knyttet til hvilken kompetanse eleven faktisk har vist i relasjon til læreplanens bredde. Der vurderingsgrunnlaget er godt strukturert og dokumentert, beskriver skolelederne at prosessen er mer forutsigbar og oppleves som mer rettferdig for både elever og lærere.

Videregående skole

I videregående skole er utfordringene ofte mer komplekse, både fordi karakterskalaen tolkes ulikt og fordi konsekvensene av karaktersetting er større. Skolelederne beskriver at lærerne tolker nivåkravene for høyeste karakter ulikt. Spesielt karakter 6 oppfattes som krevende, og lærerne er usikre på hvor bred og dyptgående kompetansen må være for å kunne tilordne denne karakteren. Dette gjelder særlig i fag med stort teoretisk omfang og i programfag hvor kompetanse uttrykkes gjennom praktiske ferdigheter som er mer krevende å dokumentere på en konsistent måte.

Bruken av karakteren 1 og IV varierer også. I videregående skole har karakteren 1 direkte konsekvenser for elevenes videre muligheter, og flere lærere uttrykker tilbakeholdenhet med å bruke den. Andre lærere mener karakteren 1 er nødvendig for å sikre tydelighet og rettferdighet. Dette fører til en praksis der noen elever får ståkarakter på svakt grunnlag, mens andre får IV under lignende forhold. Skolelederne beskriver derfor et økende behov for mer systematisk kalibrering på tvers av fag og programområder.

Proessen for fastsetting av standpunktkarakter er mer formalisert i videregående enn i ungdomsskolen. Seksjonsmøter og faglige drøftinger brukes aktivt for å kvalitetssikre begrunnelser, vurdere elevers kompetanse i fellesskap og sikre at praksisen er i tråd med regelverket. Likevel beskrives praksisen som tidkrevende, og lærerne må ofte håndtere svært ulikt vurderingsgrunnlag mellom klasser, elever og programområder. I noen fag gir praktiske og prosessbaserte vurderingssituasjoner rike grunnlag, mens teoretiske fag kan ha mer fragmenterte vurderingsdata, noe som gjør helhetsvurderingen mer krevende.

Skolelederne fremhever at skolens vurderingskultur har stor betydning for kvaliteten på standpunktfastsettingen. Der det allerede finnes en tradisjon for felles drøfting av vurdering, sensorsamarbeid og gjennomgang av elevarbeid, oppleves arbeidet som mer forutsigbart og gir større tillit både blant lærere og elever. Der slike strukturer mangler, er det større risiko for individuell tolkning, ulik praksis og svakere dokumentasjon. Ledelsen peker på behovet for både bedre støtteverktøy og mer tid til samarbeid for å sikre en enhetlig og faglig troverdig standpunktvurdering i tråd med intensjonene i læreplanen.

Samarbeid om vurdering

Ungdomsskole

I ungdomsskolen foregår det meste av samarbeidet i trinnteam, tverrfaglige grupper og fellestid der lærerne drøfter elevgrupper, tilrettelegging og faglige utfordringer. Dette gir en god helhetlig forståelse av elevenes faglige og sosiale utvikling, men gir ikke alltid rom for dyp faglig drøfting om vurdering. Skolelederne beskriver at vurderingsfaglig samarbeid i enkeltfag blir begrenset, særlig når det kun er én lærer som underviser faget på skolen. Dette fører til at lærere ofte står alene i arbeidet med å tolke kompetansemål, utvikle vurderingssituasjoner og kalibrere karakternivåer. Likevel finnes det praksiser som styrker samarbeid. Flere ungdomsskoler jobber systematisk med felles planleggingsdager, fagsamlinger og erfaringsdeling der vurdering er et fast tema. Slikt arbeid bidrar til at lærerne i større grad utvikler en delt forståelse av vurderingstekstene og karakterskalaen, men slike tiltak er igjen avhengig av ledelsesforankring og prioritering av tid.

Ledelsen ved en av skolene beskriver klagesaker som en viktig mulighet for læring og utvikling i vurderingsarbeidet. Rektor uttrykker det slik:

Jeg elsker klager. Vil helst ha 5–10 klager, selv om vi ikke får det. Det strammer oss opp. Hadde heldigvis to i fjor. Jobber systematisk med innlevering av fagrapporter, jeg går igjennom dem, har gjennomgang mellom trinnene for å sjekke at alle kompetansemål dekkes i løpet av de tre årene.

Skolen bruker fagrapporter som et systematisk verktøy for å sikre at standpunktvurderingen bygger på et bredt og korrekt vurderingsgrunnlag. Rapportene gjennomgås av ledelsen og brukes både i klagebehandling og i kvalitetsarbeidet for å sikre at kompetansemålene dekkes på tvers av trinn. Lærerne deltar i gjennomgangen, noe som gir dem erfaring med å håndtere klager og reflektere over vurderingspraksis. Fagrapportene brukes også til å identifisere uvanlige mønstre i karaktersetting, som særlig høye eller lave resultater i enkeltklasser. Dette tas videre opp med fagmiljøene for å forstå årsakene og sikre at vurderingspraksisen justeres der det er behov.

Videregående skole

Videregående skoler har generelt større fagseksjoner og mer etablerte strukturer for samarbeid. Ved flere skoler brukes seksjonsmøter aktivt til å drøfte vurderingsgrunnlag, tolke læreplaner, analysere elevarbeid og forberede fastsetting av standpunktkarakter. I flere fagmiljøer er det utviklet rutiner der lærere gjennomgår vurderingskriterier i fellesskap, bruker eksempler på elevbesvarelser og diskuterer krevende vurderingssituasjoner. Dette gir mer enhetlig praksis og støtte for lærere med ulik erfaring.

Ledelsen ved en av skolene beskriver et systematisk arbeid med karakterskalaen i tett dialog med lærerne. De gjennomgår jevnlig hva som kjennetegner de ulike karaktertrinnene, og diskuterer konkret hva som skiller for eksempel en firer fra en sekser, med utgangspunkt i elevens kompetanse. Målet er å utvikle en felles forståelse av hva som uttrykker lav, middels og høy måloppnåelse i fagene. Ledelsen følger også med på karakterfordelingen og ser etter uvanlige mønstre. Dersom det oppdages det de kaller «avvik», som når en ny eller uerfaren lærer setter uvanlig høye karakterer, tas dette opp i dialog med læreren. I slike tilfeller kan ledelsen sette inn en erfaren lærer som veileder og diskuterer vurderingsgrunnlaget sammen med vedkommende. På denne måten arbeider skolen for å sikre en mer ensartet og faglig forankret bruk av karakterskalaen på tvers av fag og klasser.

Skolelederne peker særlig på sensorsamarbeid og felles vurderingsplanlegging og-gjennomføring som viktige tiltak for å sikre at lærerne tolker karakterskalaen likt. Slike tiltak beskrives som tidkrevende, men samtidig som de mest effektive for å styrke kvaliteten på vurderingsarbeidet. I programfag der vurderingen inkluderer praktiske ferdigheter eller komplekse prosesser, opplever lederne denne typen samarbeid som spesielt nødvendig. Til tross for mer utviklede strukturer uttrykker også ledere i videregående skole at tiden til samarbeid er knapp, og at vurderingsfaglig arbeid ofte må vike for andre oppgaver. Likevel fremheves samarbeid som en av de viktigste faktorene for å sikre at vurderingsarbeidet er rettfærdig, transparent og i tråd med læreplanens intensjoner.

Elevinvolvering

Skolelederne beskriver elevinvolvering som et område der det foregår mye utviklingsarbeid, men hvor praksisen fortsatt varierer betydelig mellom fag, lærere og skoler. Elevens rolle i vurderingsarbeidet påvirkes både av skolens kultur, vurderingsspråket, tilgjengelig tid og hvordan lærerne organiserer dialogen om læring. Mange skoler legger til rette for at elevene skal delta i vurderingsarbeidet gjennom samtaler, egenvurdering, refleksjonsoppgaver og bruk av mål- og kriterieskjema. Der dette arbeidet er integrert i undervisningen, opplever skolelederne at elevene får bedre forståelse av hva som forventes, og at de kan ta mer ansvar for egen faglig utvikling. Elevinvolvering bidrar også til at vurderingsgrunnlaget blir mer transparent, fordi elevene får bedre innsikt i hva som ligger bak vurderingen av deres kompetanse.

Samtidig rapporterer skolelederne om flere utfordringer. Det gjelder særlig elevenes forståelse av vurderingsspråket. Kompetansebegrepet og kjennetegn på måloppnåelse kan fremstå som abstrakte, og mange elever synes det er vanskelig å knytte kriteriene til konkrete eksempler i eget arbeid. Dette gjelder både i ungdomsskole og videregående, men særlig på lavere trinn. Lærerne bruker derfor mye tid på å oversette faglige standarder til elevvennlig språk og på å modellere hvordan man analyserer eget arbeid opp mot kompetansemålene.

Skolelederne viser til at elevinvolvering generelt styrker elevens innsikt i egen læring, men at praksisen fortsatt preges av store forskjeller. Mange ønsker derfor mer støtte til å utvikle felles rutiner, mer tilgjengelig vurderingsspråk og bedre eksempler som kan hjelpe elevene til å forstå hva slags kompetanse som skal vises i ulike fag.

VEDLEGG 10 RAPPORT FRA SPØRREUNDERSØKELSE 1 OM VURDERINGSTEKSTENE OG STANDPUNKTVURDERING

Av Gustaf B. Skar & Anne Holten Kvistad

INNLEDNING

I dette delkapittelet presenterer vi funn fra den kvantitative undersøkelsen av læreres arbeid med standpunktvurdering. Formålet med undersøkelsen var å kartlegge en større gruppe læreres arbeid med standpunktvurdering, både på ungdomstrinnet og i videregående skole.

Undersøkelsen omfatter flere fag enn de som undersøkes kvalitativt i prosjektet, noe som kan bidra til et mer helhetlig bilde av norske læreres vurderingspraksis. Vi hadde følgende forskningsspørsmål: *Hvilke vurderingstekster kjenner lærere til og hvilke vurderingstekster bruker lærere? Hva kjennetegner praksis i standpunktvurdering? Hvordan oppfatter lærere samarbeid og skoleledelsens tilbud og krav?* Når det gjelder det første forskningsspørsmålet, var oppdragsgiver i utgangspunktet interessert i undersøkelse av de to vurderingstekstene i læreplanen (tekst om standpunktvurdering og underveisvurdering) samt kjennetegn på måloppnåelse i de fagene som har dette. I den kvantitative undersøkelsen har vi valgt å inkludere kompetansemålene for fagene og kjerneelementene i tillegg. Resultatene fra undersøkelsen kan sammenfattes i punktene nedenfor:

- Den typiske læreren kjenner til og bruker teksten om fagets kjerneelementer, kompetansemålene, kjennetegn på måloppnåelse samt de to vurderingstekstene om standpunktvurdering og underveisvurdering i fagene. Den mest brukte teksten var kompetansemålene for fagene. De minst brukte var tekstene om standpunktvurdering og underveisvurdering. Undersøkelsen lyktes ikke i å klarlegge faktorer som ville ført til økt bruk; den typiske læreren mente at hverken mer tid, større relevans, økt grad av konklusjon eller mer kunnskap ville gjort at tekstene ble mer brukt.
- Skoleledere så samlet sett ut til å forvente bruk av de fem sentrale tekstene i undersøkelsen i vurderingsarbeidet, samtidig som lærerne samlet sett var uenige i at de hadde fått veiledning i hvordan de kan brukes disse tekstene. Vi så også at lærere brukte kompetansemål for å diskutere og kvalitetssikre vurdering sammen med kollegaer. Samlet sett ble ikke noen av tekstene mye brukt i kommunikasjon med foresatte.
- Lærerne oppfattet ikke at det er lik avstand mellom trinnene i karakterskalaen (altså at de er ekvidistante), men at noen skalasteg rommet «mer kompetanse». Det var også vanlig å bruke flere skalasteg enn kun karakterene 1–6 (ved hjelp av markeringer som *pluss*, *minus* og *skråstrek*).
- Lærerne var enige i at «store oppgaver» veide tyngre enn «små oppgaver» i forbindelse med standpunktvurdering. Det fremkom også at skoleoppgaver var viktigere enn hjemmearbeid, at elevens kunnskap mot slutten av året veide mest, og at muntlig aktivitet og elevens utvikling på områder som har vært tematisert i underveisvurdering var viktig for standpunktvurdering.
- Lærerne rapporterte samlet sett at de samarbeidet med kollegaer i vurderingsarbeid, men de opplevde ikke at skoleledelsen la til rette for systematisk arbeid.

METODE

Instrument: spørreskjema

For å svare på spørsmålene om hvilke deler av læreplanen og hvilke vurderingstekster lærere kjenner til og bruker til ulike formål, utviklet forskergruppen et spørreskjema. Dette skjemaet ble sendt til godkjenning av Utdanningsdirektoratet. Skjemaet inneholdt seks deler:

1. **Bruk av tekster i vurderingsarbeidet.** Denne delen besto av påstander relatert til deltakernes kjennskap til og bruk av tekstene: *fagets kjerneelementer* (FK), *kompetansemålene for faget* (KP), *teksten om underveisvurdering* (UV), *teksten om standpunktvurdering* (SP) og *teksten om kjennetegn på måloppnåelse* (KM). Vi ba lærerne om å indikere på en sekstrinnsskala («stemmer ikke i det hele tatt», «stemmer dårlig», «stemmer nokså dårlig», «stemmer nokså godt», «stemmer godt», «stemmer veldig godt») om de var kjente med tekstene, om de brukte tekstene for å planlegge undervisning, om de brukte tekstene for å støtte elever underveis i læringsprosessen, og om de brukte tekstene for å gjennomføre standpunktvurdering. Vi ba deltakerne om å bruke en annen sekstrinnsskala («helt uenig», «ganske uenig», «litt uenig», «litt enig», «ganske enig», «helt enig») for å indikere om de hadde brukt noen av tekstene mer dersom tekstene var mer relevante eller mer konkrete, eller om deltakerne hadde hatt mer tid eller bedre kjennskap til tekstene.
2. **Skolens arbeid med standpunktvurdering.** Vi ba lærerne om å bruke sekstrinnsskalaen fra «helt uenig» til «helt enig» for å indikere om skoleledelsen hadde gjort det tydelig at deltakeren skulle bruke de fem sentrale tekstene for å kvalitetssikre standpunktvurdering og om skoleledelsen hadde gitt veiledning i hvordan disse tekstene kunne brukes. Deltakerne indikerte med hjelp av samme skala om tekstene ble brukte for å diskutere og kvalitetssikre standpunktvurdering med kollegaer og om tekstene ble brukt i kommunikasjon med foresatte.
3. **Karakterskalaen.** Vi ba lærerne om å bruke skalaen fra «helt uenig» til «helt enig» for å vurdere to påstander om karakterskalaen. Den første påstanden dreide seg om avstand mellom karaktertrinnene: «Min oppfatning av karakterskalaen er at det er like stor avstand mellom alle karaktertrinnene (f.eks. at det er like stor avstand mellom 1 og 2 som mellom 3 og 4)». Den andre påstanden dreide seg om bruk av karakterer med markeringer: «Ved min skole er det vanlig å bruke pluss/minus, skråstrek, parenteser eller andre markeringer på tallkarakterer (for eksempel 4- på en svak firer eller 2/3 dersom et elevarbeid vipper mellom karakter 2 og karakter 3)».
4. **Arbeid med standpunktvurdering.** Vi ba lærerne om å bruke skalaen fra «helt uenig» til «helt enig» for å ta stilling til ni påstander om hva som ble vektlagt i forbindelse med fastsettelse av standpunkt karakter. Disse ni påstandene hadde stammen: «Når jeg fastsetter standpunkt karakterer vektlegger jeg ...» og var formulert slik:
 - a. et gjennomsnitt av hva eleven har fått til i løpet av året
 - b. «store» oppgaver i større grad enn «små» oppgaver
 - c. resultater fra prøver og arbeid som er gjennomført på skolen i større grad enn resultater på hjemmearbeid og innleveringsoppgaver
 - d. hva eleven kan på slutten av skoleåret i større grad enn hva eleven har prestert gjennom året
 - e. oppgaver hvor elever ikke har hatt hjelpemidler tilgjengelig
 - f. muntlig aktivitet og deltakelse i klasserommet i vesentlig grad
 - g. oppgaver hvor elever har arbeidet individuelt i større grad enn samarbeidsoppgaver
 - h. om eleven har utviklet seg fra sitt individuelle utgangspunkt i større grad enn om eleven når målene i læreplanen

- i. om eleven har utviklet seg på de områdene der jeg har gitt underveisvurdering
- 5. **Samarbeid og kompetanseutvikling om standpunktvurdering.** Vi ba lærerne om å ta stilling til syv påstander om samarbeid og kompetanseutvikling om standpunktvurdering. Deltakerne brukte skalaen fra «helt uenig» til «helt enig», og påstandene var følgende:
 - a. Jeg og mine kolleger samarbeider om å lage oppgaver i fagene
 - b. Jeg og mine kolleger vurderer elevbesvarelser i fellesskap
 - c. Jeg og mine kolleger har møter/arenaer for felles arbeid om vurdering i fagene
 - d. Samarbeid med andre faglærere er viktig for at min vurdering er pålitelig og gyldig
 - e. Skoleledelsen legger til rette for samarbeid om standpunktvurdering
 - f. Skoleledelsen jobber systematisk med utvikling av vurderingskompetanse blant lærerne
 - g. Skoleledelsen legger til rette for kurs/opplæring i regelverket for vurdering
- 6. **Om deg.** Den sjettede og avsluttende delen innhold spørsmål om deltakeres kjønn, alder, utdanning og erfaring. Vi spurte også om hvilket fylke deltakernes skole lå i, om skolen lå i byen, i forstad eller i distriktet.

Rekruttering

For å rekruttere deltakere til undersøkelsen, utarbeidet vi en liste over alle ungdomsskoler og videregående skoler i Norge. Vi sendte deretter en forespørsel til epostmottak på skoler på denne listen. I brevet skrev vi følgende:

Til skolens ledelse/postmottak: Vi ber om at du videresender denne e-posten til alle ungdomstrinns lærere/videregående lærere på din enhet/i din kommune/fylkeskommune.

Til deg som er lærer på ungdomstrinn eller i videregående.

På oppdrag fra Utdanningsdirektoratet gjennomfører NTNU en spørreundersøkelse om standpunktvurdering. Formålet med undersøkelsen er å bidra med mer kunnskap om praksiser rundt standpunktvurdering i skolen.

Vi håper du vil bidra ved å svare på spørreundersøkelsen.

Hilsen

Prosjektgruppa for prosjektet Forskning på standpunktvurdering
Nasjonalt senter for skriveopplæring og skriveforskning,
NTNU ILU

Fremgangsmåten gjør at vi ikke er i stand til å vite hvor mange som fikk skjemaet tilsendt (og dermed ikke heller hvor stor svarandelen er). Vår vurdering er at fremgangsmåten likevel var egnet, siden vi ikke hadde tilgang til en modell over populasjonen lærere og derfor ikke hadde anledning til å benytte en strategi som ville ha sikret at hele populasjonen var representert.

Gjennomføring

Vi sendte forespørsel til epostmottak den 12. mars 2024. Vi mottok 280 svar samme dato og siste svar den 6. juni 2024. Skjemaet var altså åpent i knappe tre måneder. Vi mottok 713 svar i mars 2024, 24 svar i april, 44 svar i mai og 11 svar i juni.

Deltakere

Til sammen svarte 792 lærere på spørreundersøkelsen. Disse representerte tolv fag som vi hadde spesifisert på forhånd (se tabell 1) samt «andre fag» (dvs. fag som vi ikke hadde spesifisert på forhånd). Eksempel på andre fag var akvakultur, biologi, entreprenørskap og bedriftsutvikling, markedsføring og innovasjon, psykologi og valgfag: innsats for andre. (Disse fagene blir ikke presentert separat grunnet få svar per fag.) Totalt deltok 279 lærere (35,2 %) fra ungdomstrinn og 513 lærere (64,8 %) fra videregående opplæring. Av lærerne på ungdomstrinn var 177 (63,4 %) kvinner og 98 (35,1 %) var menn. Fire personer (1,4 %) ønsket ikke å oppgi kjønn. Av lærerne på videregående var 288 (56,1 %) kvinner og 215 (41,9 %) var menn. Ti deltakere (2 %) oppga ikke kjønn.

Lærerne på ungdomsskolen var i gjennomsnitt 48,4 år ($SD = 10,3$) og lærerne i videregående var i gjennomsnitt 44,7 år ($SD = 10,8$). Ungdomstrinns lærerne hadde i gjennomsnitt tatt eksamen 16,8 år før undersøkelsen ($SD = 10,9$). Tilsvarende tall for lærere i videregående var 16,1 år ($SD = 10,3$). Av deltakende lærere fra ungdomstrinn hadde 38 (13,6 %) stilling som adjunkt, 127 (45,5 %) stilling som adjunkt med opprykk, 1 (0,36 %) var allmennlærer, 14 (5 %) lektorer, 92 (33 %) lektorer med opprykk, 1 (0,36 %) førskolelærer og 6 (2,2 %) hadde annen stilling. Lærerne på videregående hadde stilling som adjunkt ($N = 45$ [8,8 %]), adjunkt med opprykk ($N = 135$ [26,3 %]), allmennlærer ($N = 1$ [0,2 %]), lektor ($N = 26$ [5,1 %]), lektor med opprykk ($N = 286$ [55,8 %]) og annen stilling ($N = 20$ [3,9 %]). Av lærerne på ungdomstrinn svarte 139 (49,8 %) at skolen deres lå i distrikt, 60 (21,5 %) at den lå i forstad og 80 (28,7 %) at den lå i byen. Tilsvarende tall for videregående var henholdsvis 151 (29,4 %), 82 (16 %) og 280 (54,6 %). Lærerne representerte skoler av ulike størrelse. Deltakerne indikerte elevtall i hele hundretall. For ungdomstrinns lærerne var gjennomsnittet 314 elever ($SD = 202$), og for lærere i videregående var gjennomsnittet 698 ($SD = 441$).

Vi kan ikke si med sikkerhet hvor godt dette utvalget representerer ungdomstrinns- og videregående lærere i de tolv fagene generelt, men vi noterer at samtlige fylker er representert, og at deltakende lærere hadde variert utdanningsbakgrunn og erfaring.

Tabell 1

Antall deltakere fordelt på fag og trinn

Fag	U.trinn		VGO		Totalt	
	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Engelsk	34	4,3%	35	4,4%	69	8,7%
Fremmedspråk	11	1,4%	23	2,9%	34	4,3%
Helse og oppvekst	-	-	32	4,0%	32	4,0%
KRLE/religion	3	0,4%	16	2,0%	19	2,4%
Kroppsøving	18	2,3%	18	2,3%	36	4,5%
Kunst og håndverk	16	2,0%	1	0,1%	17	2,1%
Mat og helse	8	1,0%	1	0,1%	9	1,1%
Matematikk	50	6,3%	80	10,1%	130	16,4%
Musikk	37	4,7%	8	1,0%	45	5,7%
Naturfag	21	2,7%	35	4,4%	56	7,1%
Norsk	62	7,8%	99	12,5%	161	20,3%
Samfunnsfag	12	1,5%	33	4,2%	45	5,7%

Tabell 1

Antall deltakere fordelt på fag og trinn

Fag	U.trinn		VGO		Totalt	
	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Annet	7	0,9%	132	16,7%	139	17,6%
Sum	279	35,2%	513	64,8%	792	100,0%

*Merknad. U.trinn: Ungdomstrinn; VGO: Videregående opplæring.***Tabell 2**

Antall deltakere fordelt på fylke

Fylke	U.trinn		VGO		Totalt	
	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Agder	16	5,7 %	22	4,3 %	38	4,8 %
Akershus	24	8,6 %	24	4,7 %	48	6,1 %
Buskerud	8	2,9 %	17	3,3 %	25	3,2 %
Finnmark	8	2,9 %	6	1,2 %	14	1,8 %
Innlandet	21	7,5 %	30	5,8 %	51	6,4 %
Møre og Romsdal	8	2,9 %	31	6,0 %	39	4,9 %
Nordland	25	9,0 %	14	2,7 %	39	4,9 %
Oslo	31	11,1 %	51	9,9 %	82	10,4 %
Østfold	4	1,4 %	22	4,3 %	26	3,3 %
Rogaland	24	8,6 %	120	23,4 %	144	18,2 %
Telemark	2	0,7 %	6	1,2 %	8	1,0 %
Troms	12	4,3 %	9	1,8 %	21	2,7 %
Trøndelag	53	19,0 %	61	11,9 %	114	14,4 %
Vestfold	11	3,9 %	39	7,6 %	50	6,3 %
Vestland	32	11,5 %	61	11,9 %	93	11,7 %

Merknad. U.trinn = ungdomstrinn; VGO = videregående opplæring.

Analyse

For denne rapporten har vi valgt å analysere gjennomsnittsverdier og standardavvik for de tolv spesifiserte fagene, for «andre fag» og for totalen (dvs. uavhengig av fag og trinn). Siden begge skalaer var sekstrinnsskalaer, er det mulig å tolke verdier som overstiger fire som at gjennomsnittsdeltakeren mente at påstanden stemte eller at de var enige. Verdier lik tre og nedover kan tolkes som at gjennomsnittsdeltakeren var uenig. Vi har valgt å ikke gjennomføre statistiske tester på forskjeller mellom fagene eller forskjeller mellom de fem sentrale tekstene, fordi antallet tester ville vært så høyt at risikoen for å konkludere med signifikante forskjeller på feilaktig grunnlag ville vært svært stor. Analysen har dermed gitt en rekke deskriptive resultat, som vil bli presentert nedenfor. Det bør påpekes at vi nokså gjennomgående vil ta utgangspunkt i den typiske læreren (altså gjennomsnittsverdier) uten å inngående kommentere på forskjeller mellom fag. Dette har å gjøre med det at vi ikke har gjennomført signifikanstester.

RESULTAT

I dette resultatkapittelet vil vi presentere funnene fra de fem delene av spørreundersøkelsen som dreide seg om kjennskap til og bruk av de fem sentrale tekstene, deriblant vurderingstekstene i læreplanen. Vi vil gå gjennom del for del.

Bruk av læreplan og vurderingstekster

Vi spurte deltakere om de var kjente med tekstene *fagets kjerneelementer* (FK), *kompetansemålene for faget* (KP), *teksten om underveisvurdering* (UV), *teksten om standpunktvurdering* (SP) og *teksten om kjennetegn på måloppnåelse* (KM). Som det fremgår av tabell 4, ser det ut til å være slik på tvers av fag og trinn at lærerne mente de var kjent med tekstene. Ser vi på totalen, var det kompetansemålene som så ut til å være mest kjent blant deltakerne, mens teksten om kjennetegn på måloppnåelse var minst kjent. Det bør dog noteres at gjennomsnittsverdiene for samtlige tekster var godt over fire, som tilsvarer svaralternativet «stemmer nokså godt».

Tabell 3

Svar på påstanden: «Jeg er kjent med denne teksten...»

	FK	KM	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	5,21 (0,64)	5,5 (0,56)	5,24 (0,74)	5,29 (0,63)	5,38 (0,7)
Fremmedspråk	4,91 (0,83)	5,36 (0,67)	4,55 (1,21)	4,91 (0,83)	5,18 (0,75)
KRLE/religion	5 (0)	5,33 (0,58)	5 (1)	5,67 (0,58)	4,67 (0,58)
Kroppsøving	4,94 (0,8)	5,56 (0,51)	5,06 (0,87)	5,11 (0,76)	4,94 (0,94)
Kunst og håndverk	4,94 (1,06)	5 (1,26)	4,69 (1,49)	4,69 (1,45)	4,69 (1,35)
Mat og helse	5,13 (0,83)	5,63 (0,52)	4,63 (0,92)	5 (0,53)	5,5 (0,53)
Matematikk	5,02 (0,87)	5,42 (0,76)	4,6 (0,88)	4,86 (0,78)	4,82 (0,87)
Musikk	5,22 (1,06)	5,57 (0,65)	5,05 (1,15)	4,97 (1,09)	5,24 (0,89)
Naturfag	5 (0,89)	5,62 (0,67)	4,67 (1,15)	5 (1,1)	4,86 (0,91)
Norsk	5,31 (0,69)	5,65 (0,52)	5,26 (0,77)	5,24 (0,9)	5,29 (0,95)
Samfunnsfag	4,92 (0,67)	5,42 (0,67)	4,75 (1,06)	4,83 (0,83)	5,25 (0,62)
Annet	4,71 (1,8)	4,86 (1,86)	4 (1,63)	4,14 (1,68)	4,43 (1,72)
<i>VGO</i>					
Engelsk	5,26 (0,89)	5,77 (0,43)	5,26 (0,7)	5,31 (0,87)	4,97 (1,29)
Fremmedspråk	5 (1,17)	5,74 (0,45)	5,04 (1,07)	5 (1,21)	5,39 (0,84)
Helse og oppvekst	5,22 (0,66)	5,78 (0,42)	5,13 (0,83)	5,16 (0,85)	4,97 (1,38)
KRLE/religion	5 (1,15)	5,63 (0,62)	4,63 (1,02)	4,94 (1,06)	4,81 (1,22)
Kroppsøving	5,33 (1,24)	5,89 (0,47)	5,06 (0,94)	5,11 (0,96)	4,89 (1,23)
Kunst og håndverk	6 (0)	6 (0)	6 (0)	6 (0)	6 (0)
Mat og helse	5 (0)	5 (0)	5 (0)	5 (0)	5 (0)

Tabell 3

Svar på påstanden: «Jeg er kjent med denne teksten...»

	FK	KM	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Matematikk	4,95 (0,91)	5,68 (0,5)	4,9 (1,03)	5,06 (0,95)	4,4 (1,5)
Musikk	5,63 (0,52)	5,63 (0,52)	5,63 (0,52)	5,5 (0,76)	3,88 (2,47)
Naturfag	5,17 (0,82)	5,89 (0,32)	5,14 (0,85)	5,14 (0,88)	4,89 (1,21)
Norsk	5,25 (0,68)	5,71 (0,54)	5,29 (0,77)	5,33 (0,78)	5,26 (0,83)
Samfunnsfag	5,27 (0,84)	5,7 (0,47)	5,06 (0,97)	5,42 (0,61)	5,3 (0,68)
Annet	5,08 (0,91)	5,58 (0,63)	4,97 (0,98)	5,08 (1)	4,59 (1,59)
<i>Totalt</i>					
Alle deltakere	5,13 (0,87)	5,62 (0,62)	5,02 (0,96)	5,12 (0,93)	4,94 (1,22)

Merknad. FK = fagets kjerneelementer; KM = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

Tabell 4 presenterer verdier for svar på påstanden om at tekstene ble brukt for å planlegge undervisning. Igjen ser vi at gjennomsnittsverdiene var over fire for alle tekster. Kompetansemålene virket å være mest brukt. ($M = 5,49$; $SD = 0,79$). Minst brukt var teksten om standpunktvurdering ($M = 4,39$; $SD = 1,24$).

Tabell 4

Svar på påstanden: «Jeg bruker denne teksten for å planlegge undervisning...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	4,68 (0,88)	5,47 (0,71)	4,74 (0,9)	4,56 (1,02)	4,97 (0,97)
Fremmedspråk	4,18 (1,4)	5,45 (0,69)	4,27 (1,27)	3,91 (1,22)	4,82 (0,87)
KRLE/religion	5 (0)	5,67 (0,58)	4,67 (0,58)	4,67 (0,58)	4 (1)
Kroppsøving	4,5 (1,34)	5,56 (0,7)	4,5 (1,2)	4,39 (0,85)	4,44 (1,25)
Kunst og håndverk	4,56 (1,41)	5,31 (0,87)	4,69 (1,08)	4,63 (0,96)	4,56 (1,36)
Mat og helse	5,13 (0,83)	5,25 (0,89)	4,25 (1,49)	4,5 (0,93)	5,13 (0,83)
Matematikk	4,24 (1,02)	5,08 (1,01)	3,9 (1,13)	3,88 (1,19)	4,34 (1,12)
Musikk	4,7 (1,18)	5,49 (0,73)	4,43 (1,12)	4,08 (1,38)	4,65 (1,06)
Naturfag	4,48 (1,17)	5,62 (0,59)	4,14 (1,42)	4,14 (1,46)	4,14 (1,49)
Norsk	4,53 (1)	5,52 (0,76)	4,6 (1,22)	4,39 (1,4)	4,87 (1,21)
Samfunnsfag	4,5 (0,8)	5,17 (0,72)	4 (1,28)	3,92 (1,16)	4,58 (1,31)
Annet	4,71 (1,8)	4,71 (1,8)	4,14 (1,77)	4 (1,63)	4,29 (1,6)
<i>VGO</i>					
Engelsk	4,51 (1,2)	5,57 (0,74)	4,54 (1,09)	4,29 (1,27)	4,49 (1,34)
Fremmedspråk	4,35 (1,4)	5,26 (1,1)	4,61 (1,31)	4,35 (1,34)	4,43 (1,59)

Tabell 4

Svar på påstanden: «Jeg bruker denne teksten for å planlegge undervisning...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Helse og oppvekst	4,66 (1,18)	5,75 (0,51)	5 (0,88)	4,75 (1,08)	4,56 (1,78)
KRLE/religion	4,13 (0,96)	5,44 (0,73)	4 (0,97)	4,06 (1,12)	4,13 (1,36)
Kroppsøving	4,89 (1,23)	5,78 (0,55)	4,61 (1,38)	4,44 (1,42)	4,56 (1,62)
Kunst og håndverk	6 (0)	6 (0)	4 (0)	4 (0)	4 (0)
Mat og helse	5 (0)	5 (0)	5 (0)	5 (0)	5 (0)
Matematikk	3,96 (1,28)	5,43 (0,85)	4,15 (1,14)	4,09 (1,16)	3,71 (1,5)
Musikk	5,38 (0,52)	5,63 (0,52)	5,13 (0,83)	5 (0,93)	4,38 (2,2)
Naturfag	4,2 (1,05)	5,74 (0,44)	4,51 (1,04)	4,29 (1,23)	4,06 (1,53)
Norsk	4,52 (1,04)	5,65 (0,61)	4,71 (0,92)	4,62 (1,16)	4,92 (1,09)
Samfunnsfag	4,12 (1,43)	5,27 (1,26)	4,18 (1,45)	4,27 (1,55)	4,61 (1,46)
Annet	4,6 (1,19)	5,55 (0,69)	4,68 (1,12)	4,74 (1,2)	4,31 (1,63)
<i>Totalt</i>					
Alle deltakere	4,46 (1,17)	5,49 (0,79)	4,49 (1,15)	4,39 (1,24)	4,46 (1,42)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

I tabell 5 gjengis gjennomsnittsverdier for påstander om bruk av de fem tekstene for å støtte elever underveis i læringsprosessen. Med unntak av noen enkelttilfeller, var verdiene fire eller høyere på tvers av fag og trinn. Igjen var kompetansemålene den mest brukte teksten ($M = 5,17$; $SD = 0,98$). Den minst brukte teksten var den om standpunktvurdering ($M = 4,17$; $SD = 1,27$).

Tabell 5

Svar på påstanden: «Jeg bruker denne teksten for å støtte elever underveis i læringsprosessen...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	4,35 (1,18)	5,15 (0,96)	4,56 (1,08)	4,26 (1,05)	4,74 (0,99)
Fremmedspråk	4,27 (0,9)	4,91 (0,83)	3,82 (1,33)	3,73 (1,19)	4,64 (0,92)
KRLE/religion	4,33 (0,58)	4,67 (0,58)	3,67 (0,58)	4,33 (0,58)	4 (1)
Kroppsøving	4,33 (1,33)	5,5 (0,62)	4,67 (1,19)	4,22 (1,17)	4,33 (1,33)
Kunst og håndverk	4,25 (1,81)	4,81 (1,42)	4,56 (1,21)	4,25 (1,39)	4,81 (1,05)
Mat og helse	4,88 (0,83)	4,88 (0,83)	4,13 (1,36)	4,13 (0,99)	4,88 (0,83)
Matematikk	4,16 (1,23)	4,86 (1,23)	3,88 (1,12)	3,74 (1,23)	4,14 (1,14)
Musikk	4,27 (1,33)	4,95 (1,18)	4,22 (1,27)	3,81 (1,24)	4,43 (1,34)
Naturfag	4,24 (1,41)	5,1 (1,09)	4,1 (1,51)	3,67 (1,46)	4,05 (1,24)
Norsk	4,23 (1,34)	5,1 (1,11)	4,66 (1,27)	4,24 (1,43)	4,94 (1,2)
Samfunnsfag	4,33 (1,07)	5 (0,95)	4,25 (1,06)	4,17 (0,83)	4,5 (1,31)
Annet	4,14 (1,77)	4,29 (1,8)	3,57 (1,62)	3,57 (1,62)	3,86 (1,68)
<i>VGO</i>					
Engelsk	3,97 (1,48)	5,4 (0,74)	4,54 (1,27)	4 (1,5)	4,54 (1,52)

Tabell 5

Svar på påstanden: «Jeg bruker denne teksten for å støtte elever underveis i læringsprosessen...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Fremmedspråk	4,22 (1,51)	5,09 (0,73)	4,43 (1,56)	4,17 (1,34)	4,87 (1,22)
Helse og oppvekst	4,47 (1,29)	5,69 (0,64)	4,91 (1)	4,81 (1,06)	4,44 (1,79)
KRLE/religion	3,69 (0,95)	4,94 (1,12)	3,69 (1,14)	3,88 (1,15)	4,19 (1,42)
Kroppsøving	4,67 (1,14)	5,72 (0,46)	4,39 (1,2)	4,33 (1,37)	4,5 (1,65)
Kunst og håndverk	5 (0)	6 (0)	6 (0)	5 (0)	5 (0)
Mat og helse	5 (0)	6 (0)	5 (0)	5 (0)	5 (0)
Matematikk	3,94 (1,36)	5 (1,04)	4,08 (1,35)	3,86 (1,31)	3,76 (1,51)
Musikk	5 (0,53)	5,13 (0,83)	4,5 (1,2)	4,75 (0,71)	4,13 (2,03)
Naturfag	3,8 (1,35)	5,14 (0,88)	4,17 (1,15)	3,86 (1,06)	3,83 (1,52)
Norsk	4,15 (1,1)	5,28 (0,76)	4,64 (1,14)	4,34 (1,21)	4,82 (1,17)
Samfunnsfag	3,76 (1,52)	4,94 (1,32)	4,24 (1,39)	4,06 (1,41)	4,55 (1,54)
Annet	4,45 (1,28)	5,37 (0,79)	4,5 (1,23)	4,5 (1,2)	4,3 (1,66)
<i>Totalt</i>					
Alle deltakere	4,21 (1,3)	5,17 (0,98)	4,38 (1,26)	4,17 (1,27)	4,41 (1,44)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

I tabell 6 presenteres verdier for påstander relatert til hvilke tekster som ble brukt i forbindelse med standpunkt-vurdering. Kompetansemålene var mest brukt ($M = 5,51$; $SD = 0,79$), tett fulgt av teksten om standpunkt-vurdering ($M = 5,05$; $SD = 1,1$). Minst brukt var teksten om underveisvurdering ($M = 4,46$; $SD = 1,25$).

Tabell 6

Svar på påstanden: «Jeg bruker denne teksten for å gjennomføre standpunkt-vurdering...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	5,06 (0,92)	5,56 (0,61)	4,91 (1,03)	5,35 (0,81)	5,5 (0,66)
Fremmedspråk	4,82 (0,75)	5,55 (0,69)	4 (1,18)	5 (0,63)	5,09 (0,94)
KRLE/religion	5,33 (0,58)	6 (0)	4,33 (1,15)	5,67 (0,58)	5,33 (0,58)
Kroppsøving	4,61 (1,33)	5,5 (0,71)	4,61 (0,98)	5,11 (0,76)	4,94 (1,06)
Kunst og håndverk	4,56 (1,46)	5,31 (0,87)	4,56 (1,09)	4,94 (1)	4,81 (0,98)
Mat og helse	5 (0,93)	5,25 (0,89)	4,13 (1,36)	5,13 (0,83)	5,13 (0,64)
Matematikk	4,8 (1,09)	5,44 (0,76)	4 (1,16)	5 (0,95)	4,92 (1,18)
Musikk	4,65 (1,23)	5,38 (1,04)	4,27 (1,28)	4,76 (1,4)	5 (1,25)
Naturfag	4,62 (1,5)	5,43 (1,03)	3,81 (1,33)	4,81 (1,21)	4,81 (1,21)
Norsk	4,76 (1,15)	5,55 (0,67)	4,56 (1,15)	5,32 (0,97)	5,44 (0,97)
Samfunnsfag	4,92 (0,79)	5,42 (0,67)	4,67 (0,89)	5,08 (0,79)	5 (1,41)
Annet	4,71 (1,7)	4,86 (1,77)	4,29 (1,7)	4,29 (1,7)	4,14 (1,57)

Tabell 6

Svar på påstanden: «Jeg bruker denne teksten for å gjennomføre standpunktvurdering...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>VGO</i>					
Engelsk	4,63 (1,35)	5,51 (1,01)	4,49 (1,42)	4,86 (1,5)	4,74 (1,65)
Fremmedspråk	4,61 (1,47)	5,65 (0,57)	4,48 (1,68)	5,26 (1,29)	5,48 (1,12)
Helse og oppvekst	4,94 (1,08)	5,81 (0,47)	4,81 (1,38)	5,38 (0,98)	4,75 (1,78)
KRLE/religion	4,25 (1,24)	5,31 (1,08)	3,94 (1,48)	4,56 (1,36)	4,5 (1,51)
Kroppsøving	4,83 (1,15)	5,78 (0,43)	4,61 (1,38)	5 (1,14)	4,94 (1,55)
Kunst og håndverk	6 (0)	6 (0)	6 (0)	6 (0)	6 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	5 (0)	5 (0)
Matematikk	4,26 (1,33)	5,44 (0,78)	4,09 (1,36)	4,99 (1,04)	4,31 (1,48)
Musikk	5,25 (0,71)	5,75 (0,46)	5 (0,93)	5,25 (0,89)	4,25 (2,12)
Naturfag	4,17 (1,15)	5,57 (0,61)	4,37 (1,03)	4,94 (0,87)	4,37 (1,48)
Norsk	4,46 (1,14)	5,65 (0,58)	4,75 (1,04)	5,18 (1,01)	5,17 (1,01)
Samfunnsfag	4,48 (1,25)	5,24 (1,17)	4,36 (1,52)	5,24 (1,15)	5,03 (1,53)
Annet	4,63 (1,23)	5,46 (0,82)	4,61 (1,18)	4,93 (1,19)	4,48 (1,69)
<i>Totalt</i>					
Alle deltakere	4,61 (1,21)	5,51 (0,79)	4,46 (1,25)	5,05 (1,1)	4,85 (1,4)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

Vi ba deltakerne om å ta stilling til hva som ville fått dem til å bruke tekstene mer. I tabell 7 presenteres faktorer som ville fått deltakerne til å bruke teksten om fagets kjerneelementer mer. Sett under ett er det to faktorer, tid og konkretisering, som peker seg ut. Lærerne mente altså at hvis de hadde hatt mer tid, og hvis teksten var mer konkret, så ville den ha blitt brukt mer. Den typiske læreren var ikke enig i at bruken ville ha økt om teksten var mer relevant eller om hen hadde mer kunnskap om teksten.

Tabell 7

Svar på påstanden: «Jeg ville ha brukt teksten om *fagets kjerneelementer* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	3,47 (1,42)	3,74 (1,64)	2,91 (1,38)	3,88 (1,74)
Fremmedspråk	3,09 (1,22)	3,64 (1,29)	3,55 (0,93)	3,82 (0,87)
KRLE/religion	2,67 (0,58)	4 (1)	3,33 (2,52)	4 (1,73)
Kroppsøving	3,06 (1,86)	4,22 (1,9)	2,83 (1,62)	3,89 (1,88)
Kunst og håndverk	3,5 (1,32)	3,38 (1,31)	3,06 (1,44)	3,88 (1,75)
Mat og helse	3,25 (1,39)	4,5 (1,31)	2,5 (1,41)	3,38 (2,2)
Matematikk	3,64 (1,22)	4,62 (1,14)	3,38 (1,37)	4,64 (1,26)
Musikk	3,24 (1,42)	3,92 (1,67)	2,92 (1,42)	3,95 (1,56)
Naturfag	3,14 (1,53)	4,19 (1,57)	3,29 (1,35)	3,57 (1,57)
Norsk	3,08 (1,3)	4,18 (1,63)	3,02 (1,4)	3,69 (1,42)
Samfunnsfag	3,83 (1,47)	4,42 (1,16)	3,17 (1,47)	4,08 (1,24)

Tabell 7Svar på påstanden: «Jeg ville ha brukt teksten om *fagets kjerneelementer* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
Annet	4,14 (1,07)	4 (1,29)	4 (0,82)	4 (1,41)
<i>VGO</i>				
Engelsk	3,63 (1,46)	4,11 (1,37)	2,8 (1,43)	4,06 (1,71)
Fremmedspråk	3,48 (1,95)	3,87 (1,74)	2,87 (1,74)	3,61 (1,85)
Helse og oppvekst	2,97 (1,36)	3,78 (1,72)	3,22 (1,56)	3,41 (1,79)
KRLE/religion	3,5 (1,32)	4,38 (1,31)	3,81 (1,33)	4,75 (1,48)
Kroppsøving	3,44 (1,58)	4,11 (1,53)	2,28 (1,49)	4,17 (1,47)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	4 (0)
Matematikk	3,95 (1,42)	4,08 (1,48)	3,21 (1,38)	4,35 (1,61)
Musikk	2,75 (1,04)	2,88 (1,55)	1,63 (0,74)	3,38 (1,6)
Naturfag	3,63 (1,29)	4,54 (1,38)	3,23 (1,54)	4,14 (1,44)
Norsk	3,47 (1,44)	4,35 (1,66)	3,15 (1,46)	4,16 (1,57)
Samfunnsfag	3,85 (1,46)	4,06 (1,56)	2,64 (1,43)	4,21 (1,63)
Annet	3,64 (1,44)	3,97 (1,64)	2,99 (1,49)	3,97 (1,58)
<i>Totalt</i>				
Alle deltakere	3,5 (1,43)	4,1 (1,56)	3,05 (1,46)	4,03 (1,59)

Merknad. Relevant = «hvis teksten var mer relevant»; tid = «hvis jeg hadde mer tid»; kunnskap = «hvis jeg hadde mer kunnskap»; konkret = «hvis teksten var mer konkret».

I tabell 8–11 presenteres resultatene for de andre tekstene (kompetansemål [tabell 8], underveisvurdering [tabell 9], standpunktvurdering [tabell 10] og kjennetegn på måloppnåelse [tabell 11]). Med enkelte unntak ville ikke noen av faktorene relevans, tid, kunnskap og konkretisering bidratt til økt bruk. Med andre ord er eventuelt ikke-bruk knyttet til andre faktorer enn de vi foreslo, ifølge lærerne.

Tabell 8Svar på påstanden: «Jeg ville ha brukt *kompetansemålene* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	3,12 (1,49)	3,38 (1,65)	2,71 (1,49)	3,59 (1,69)
Fremmedspråk	2,73 (0,9)	3,09 (1,3)	3 (1,1)	3,82 (1,4)
KRLE/religion	2,33 (1,15)	3 (1)	2,67 (1,53)	3,67 (0,58)
Kroppsøving	2,67 (1,81)	3,17 (2,04)	2 (1,5)	3,39 (2,03)
Kunst og håndverk	3,38 (1,41)	3,13 (1,41)	2,75 (1,44)	3,56 (1,82)
Mat og helse	3,38 (1,51)	3,75 (1,83)	3 (1,69)	3,38 (1,92)
Matematikk	3,76 (1,46)	4,16 (1,5)	2,9 (1,36)	4,8 (1,29)
Musikk	3,19 (1,52)	3,27 (1,66)	2,73 (1,48)	3,76 (1,52)
Naturfag	3,19 (1,78)	3,43 (1,78)	2,24 (1,34)	3,38 (1,88)
Norsk	2,69 (1,5)	3,32 (1,76)	2,32 (1,4)	3,44 (1,68)
Samfunnsfag	3,08 (1,44)	3,75 (1,6)	2,67 (1,5)	4,5 (1,24)
Annet	3,29 (1,5)	3,29 (1,6)	2,71 (1,11)	3,43 (1,72)
<i>VGO</i>				
Engelsk	3,2 (1,75)	2,94 (2,04)	2,03 (1,6)	3,54 (2,21)
Fremmedspråk	3,35 (1,9)	3,52 (1,75)	2,43 (1,41)	3,83 (1,83)
Helse og oppvekst	2,72 (1,61)	2,72 (1,57)	2,44 (1,41)	2,94 (1,68)

Tabell 8Svar på påstanden: «Jeg ville ha brukt *kompetansemålene* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
KRLE/religion	3 (1,67)	3,38 (1,5)	2,38 (1,26)	3,81 (2,1)
Kroppsøving	3,22 (1,77)	2,89 (1,97)	1,89 (1,23)	4,06 (1,92)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	4 (0)
Matematikk	3,5 (1,53)	3,26 (1,68)	2,38 (1,48)	4,56 (1,68)
Musikk	2,63 (1,19)	2,63 (1,85)	2 (1,07)	3,38 (1,41)
Naturfag	3,26 (1,77)	3,14 (1,8)	2,2 (1,35)	3,83 (1,67)
Norsk	3,21 (1,72)	3,25 (1,8)	2,49 (1,53)	3,38 (1,74)
Samfunnsfag	3,15 (1,8)	3,3 (1,7)	1,97 (1,29)	3,79 (2)
Annet	3,37 (1,6)	3,23 (1,77)	2,47 (1,48)	3,75 (1,82)
<i>Totalt</i>				
Alle deltakere	3,21 (1,62)	3,29 (1,73)	2,44 (1,45)	3,77 (1,79)

Merknad. Relevant = «hvis teksten var mer relevant»; tid = «hvis jeg hadde mer tid»; kunnskap = «hvis jeg hadde mer kunnskap»; konkret = «hvis teksten var mer konkret».

Tabell 9Svar på påstanden: «Jeg ville ha brukt teksten om *undervisvurdering* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	3,56 (1,42)	3,56 (1,71)	3,18 (1,45)	3,74 (1,68)
Fremmedspråk	2,91 (0,83)	3,55 (1,37)	3,36 (1,21)	3,82 (1,25)
KRLE/religion	3,67 (1,15)	3,33 (1,53)	3 (2)	3,67 (1,15)
Kroppsøving	3,22 (1,66)	3,5 (1,58)	2,89 (1,53)	4 (1,85)
Kunst og håndverk	3,25 (1,61)	3,19 (1,56)	3,5 (1,75)	3,44 (1,59)
Mat og helse	3,75 (1,04)	4,38 (1,06)	4,38 (0,92)	4,25 (1,75)
Matematikk	3,74 (1,27)	4,48 (1,3)	3,7 (1,28)	4,36 (1,37)
Musikk	3,38 (1,44)	3,62 (1,55)	3,08 (1,34)	3,92 (1,53)
Naturfag	3,19 (1,54)	3,48 (1,69)	3,43 (1,94)	3,86 (1,49)
Norsk	2,94 (1,34)	3,5 (1,7)	2,84 (1,4)	3,4 (1,52)
Samfunnsfag	3,33 (1,44)	4 (1,35)	3,08 (1,51)	3,83 (1,47)
Annet	3,71 (1,8)	3,29 (1,89)	3,43 (1,51)	4,14 (1,77)
<i>VGO</i>				
Engelsk	3,57 (1,58)	3,57 (1,74)	3,03 (1,74)	3,91 (1,8)
Fremmedspråk	3,13 (1,84)	3,87 (1,87)	2,7 (1,61)	3,52 (1,65)
Helse og oppvekst	2,91 (1,44)	3,47 (1,63)	3,19 (1,57)	3,13 (1,66)
KRLE/religion	3,5 (1,55)	3,75 (1,73)	3,63 (1,26)	4,06 (1,91)
Kroppsøving	3,67 (1,46)	3,61 (1,97)	2,11 (1,37)	4,5 (1,5)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	4 (0)
Matematikk	3,71 (1,32)	3,59 (1,52)	3,14 (1,44)	4,1 (1,45)
Musikk	3,13 (1,13)	2,63 (1,69)	2,25 (1,04)	3,25 (1,28)
Naturfag	3,63 (1,42)	4,11 (1,68)	3,49 (1,5)	4,03 (1,58)
Norsk	3,44 (1,51)	3,77 (1,69)	3,11 (1,48)	3,87 (1,63)
Samfunnsfag	3,21 (1,54)	3,48 (1,56)	2,61 (1,43)	3,42 (1,66)
Annet	3,64 (1,45)	3,58 (1,69)	3,17 (1,51)	3,7 (1,56)
<i>Totalt</i>				

Tabell 9Svar på påstanden: «Jeg ville ha brukt teksten om *underveisvurdering* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
Alle deltakere	3,44 (1,45)	3,67 (1,64)	3,13 (1,5)	3,81 (1,59)

Merknad. Relevant = «hvis teksten var mer relevant»; tid = «hvis jeg hadde mer tid»; kunnskap = «hvis jeg hadde mer kunnskap»; konkret = «hvis teksten var mer konkret».

Tabell 10Svar på påstanden: «Jeg ville ha brukt teksten om *standpunktavurdering* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	3,26 (1,42)	3,35 (1,69)	3,12 (1,57)	3,62 (1,6)
Fremmedspråk	2,91 (0,54)	3,27 (1,1)	3,18 (0,75)	3,55 (1,21)
KRLE/religion	2,67 (1,53)	3 (1)	2,67 (1,53)	3 (1)
Kroppsøving	3,22 (1,73)	3,39 (1,69)	3 (1,68)	3,72 (2,02)
Kunst og håndverk	3,25 (1,73)	3 (1,55)	3,06 (1,77)	3,44 (1,82)
Mat og helse	3,88 (1,36)	4,25 (0,89)	3,5 (1,2)	3,88 (1,73)
Matematikk	3,58 (1,3)	4,14 (1,41)	3,48 (1,33)	4,14 (1,46)
Musikk	3,43 (1,56)	3,62 (1,67)	3,27 (1,52)	3,7 (1,56)
Naturfag	2,9 (1,58)	2,9 (1,45)	3,05 (1,53)	3,19 (1,57)
Norsk	2,87 (1,51)	3,23 (1,8)	2,76 (1,3)	3,16 (1,71)
Samfunnsfag	3,17 (1,11)	3,92 (1,38)	3,42 (1,38)	3,83 (1,53)
Annet	3,57 (1,9)	3,14 (1,77)	3,14 (1,07)	4,14 (1,77)
<i>VGO</i>				
Engelsk	3,31 (1,43)	3,31 (1,71)	2,97 (1,64)	3,74 (1,69)
Fremmedspråk	3,13 (1,77)	3,52 (1,75)	3,17 (1,8)	3,74 (1,86)
Helse og oppvekst	2,88 (1,45)	3,25 (1,74)	2,97 (1,49)	3,09 (1,63)
KRLE/religion	3,44 (1,36)	3,63 (1,63)	3,38 (1,31)	4 (1,83)
Kroppsøving	3,06 (1,59)	3 (1,75)	2,5 (1,47)	4,22 (1,59)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	4 (0)
Matematikk	3,33 (1,51)	3,15 (1,64)	2,86 (1,43)	3,95 (1,69)
Musikk	2,5 (0,93)	2,63 (1,41)	2,38 (1,19)	3,5 (1,2)
Naturfag	3,2 (1,43)	3,71 (1,74)	3,03 (1,4)	3,8 (1,64)
Norsk	3,41 (1,54)	3,64 (1,71)	3,14 (1,53)	3,76 (1,64)
Samfunnsfag	2,88 (1,75)	3,24 (1,6)	2,27 (1,35)	2,94 (1,64)
Annet	3,41 (1,53)	3,33 (1,77)	3 (1,5)	3,45 (1,59)
<i>Totalt</i>				
Alle deltakere	3,25 (1,51)	3,41 (1,67)	3,01 (1,48)	3,62 (1,65)

Merknad. Relevant = «hvis teksten var mer relevant»; tid = «hvis jeg hadde mer tid»; kunnskap = «hvis jeg hadde mer kunnskap»; konkret = «hvis teksten var mer konkret».

Tabell 11Svar på påstanden: «Jeg ville ha brukt teksten om *kjennetegn på måloppnåelse* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	3,21 (1,47)	3,35 (1,7)	2,91 (1,56)	3,68 (1,7)

Tabell 11Svar på påstanden: «Jeg ville ha brukt teksten om *kjennetegn på måloppnåelse* mer hvis...»

	Relevant	Tid	Kunnskap	Konkret
	M (SD)	M (SD)	M (SD)	M (SD)
Fremmedspråk	2,73 (0,9)	3,27 (1,42)	3 (1,1)	3,55 (1,51)
KRLE/religion	3 (2)	2,67 (1,53)	3,33 (2,52)	3,33 (2,52)
Kroppsøving	3,44 (1,46)	3,67 (1,75)	3,06 (1,59)	4 (1,71)
Kunst og håndverk	3,5 (1,63)	3,06 (1,34)	3,13 (1,5)	4 (1,59)
Mat og helse	2,75 (1,58)	3,75 (1,58)	2,75 (1,16)	3,5 (1,93)
Matematikk	3,86 (1,32)	4,22 (1,31)	3,36 (1,45)	4,5 (1,42)
Musikk	3,43 (1,54)	3,41 (1,72)	3,24 (1,46)	3,81 (1,54)
Naturfag	3,19 (1,6)	3,57 (1,6)	3,48 (1,47)	4,05 (1,72)
Norsk	2,66 (1,33)	3,32 (1,82)	2,55 (1,31)	3,15 (1,71)
Samfunnsfag	3,42 (1,78)	4 (1,54)	3 (1,54)	4,58 (1,16)
Annet	3,57 (1,62)	3,29 (1,6)	3,14 (1,21)	4 (1,73)
<i>VGO</i>				
Engelsk	3,54 (1,65)	3,26 (1,82)	3 (1,7)	3,89 (1,66)
Fremmedspråk	3,13 (1,87)	3,39 (1,85)	3 (1,71)	3,52 (1,86)
Helse og oppvekst	3,16 (1,59)	3,72 (1,75)	3,19 (1,64)	3,31 (1,77)
KRLE/religion	3,25 (1,73)	4 (1,46)	3,63 (1,02)	4,06 (1,81)
Kroppsøving	3,67 (1,53)	3,17 (1,79)	2,39 (1,29)	4,28 (1,56)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	4 (0)	5 (0)	5 (0)	4 (0)
Matematikk	3,66 (1,61)	3,33 (1,69)	3,04 (1,64)	4,1 (1,69)
Musikk	3,38 (2)	3,25 (2,25)	3,13 (2,17)	3,5 (1,77)
Naturfag	3,37 (1,77)	3,46 (1,92)	2,94 (1,68)	3,69 (1,88)
Norsk	3,42 (1,55)	3,66 (1,66)	3,16 (1,52)	3,65 (1,66)
Samfunnsfag	3,18 (1,69)	3,48 (1,7)	2,45 (1,48)	3,48 (1,72)
Annet	3,52 (1,62)	3,42 (1,85)	3,18 (1,7)	3,7 (1,69)
<i>Totalt</i>				
Alle deltakere	3,38 (1,58)	3,5 (1,72)	3,05 (1,56)	3,77 (1,69)

Merknad. Relevant = «hvis teksten var mer relevant»; tid = «hvis jeg hadde mer tid»; kunnskap = «hvis jeg hadde mer kunnskap»; konkret = «hvis teksten var mer konkret».

Skolens arbeid med standpunktvurdering

Denne delen av undersøkelsen dreide seg om skoleleders forventninger og veiledning med tanke på de fem sentrale tekstene, samt samarbeid med kollegaer om disse tekstene. I tabell 12 gjengis gjennomsnittsverdier relatert til påstanden «skoleledelsen har gjort det tydelig at jeg skal bruke disse tekstene for å kvalitetssikre min standpunktvurdering». Teksten som ser ut til å ha blitt mest tematisert er kompetansemålene for fagene ($M = 4,94$; $SD = 1,29$). Avstanden til minst tematiserte tekst, den om standpunktvurdering, var cirka ett skalatrinn ($M = 4,04$; $SD = 1,58$). Generelt ser det ut til at lærerne har oppfattet det slik at skoleledelsen forventer bruk av tekstene.

Tabell 12

Svar på påstanden: «Skoleledelsen har gjort det tydelig at jeg skal bruke disse tekstene for å kvalitetssikre min standpunktvrdering...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	4,41 (1,35)	5,18 (1,11)	4,41 (1,33)	4,68 (1,47)	4,62 (1,23)
Fremmedspråk	4,91 (0,83)	5,55 (0,69)	4,82 (1,17)	5,18 (0,87)	4,91 (0,83)
KRLE/religion	4,67 (1,53)	5,33 (0,58)	4,67 (2,31)	5 (1,73)	5 (1)
Kroppsøving	3,83 (1,65)	5 (1,33)	4,28 (1,49)	3,89 (1,49)	3,67 (1,46)
Kunst og håndverk	4,88 (1,63)	5,06 (1,34)	4,56 (1,31)	4,63 (1,36)	4,69 (1,35)
Mat og helse	4,5 (1,31)	4,75 (1,39)	3,63 (1,69)	4,25 (1,16)	4,88 (1,13)
Matematikk	4,2 (1,6)	4,86 (1,26)	3,98 (1,49)	4,36 (1,47)	4,3 (1,47)
Musikk	4,41 (1,36)	5,19 (1,02)	4,05 (1,61)	4,32 (1,68)	4,38 (1,48)
Naturfag	4,14 (1,46)	4,76 (1,51)	4,1 (1,14)	4,29 (1,23)	3,62 (1,47)
Norsk	4,05 (1,65)	5,08 (1,27)	3,92 (1,67)	4,23 (1,63)	4,37 (1,62)
Samfunnsfag	4,5 (0,8)	5,17 (1,03)	4 (1,71)	4,25 (1,42)	4,42 (1,08)
Annet	4,86 (1,46)	4,71 (1,5)	4,29 (1,5)	4,14 (1,68)	4,14 (2,04)
<i>VGO</i>					
Engelsk	4 (1,75)	4,69 (1,68)	3,8 (1,75)	4,06 (1,86)	3,63 (1,96)
Fremmedspråk	4,35 (1,53)	5,04 (1,4)	4,78 (1,31)	4,78 (1,35)	4,26 (1,74)
Helse og oppvekst	3,94 (1,74)	5,19 (1,09)	4,34 (1,41)	4,56 (1,46)	3,69 (1,69)
KRLE/religion	3,75 (1,34)	4,94 (0,85)	3,75 (1,29)	4,19 (1,05)	4,19 (0,98)
Kroppsøving	4 (1,53)	5,17 (1,2)	3,72 (1,45)	4,39 (1,46)	3,94 (1,51)
Kunst og håndverk	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)
Mat og helse	5 (0)	6 (0)	5 (0)	6 (0)	6 (0)
Matematikk	4,1 (1,42)	4,86 (1,32)	4,4 (1,3)	4,58 (1,3)	3,85 (1,48)
Musikk	4,13 (1,64)	4,88 (0,99)	3,75 (1,91)	4 (1,6)	4 (2,07)
Naturfag	3,89 (1,66)	4,89 (1,32)	4,43 (1,52)	4,57 (1,46)	3,6 (1,83)
Norsk	3,85 (1,57)	4,71 (1,57)	4,3 (1,59)	4,27 (1,71)	3,97 (1,73)
Samfunnsfag	4,15 (1,5)	4,88 (1,49)	4,24 (1,3)	4,48 (1,46)	3,97 (1,69)
Annet	4,22 (1,38)	4,94 (1,12)	4,33 (1,32)	4,3 (1,37)	4,05 (1,51)
<i>Totalt</i>					
Alle deltakere	4,14 (1,51)	4,94 (1,29)	4,22 (1,46)	4,38 (1,48)	4,07 (1,58)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

Tabell 13 presenterer resultater relatert til påstanden om at skoleledelsen hadde gitt veiledning i bruk av tekstene. Med få unntak var lærerne uenige i denne påstanden. Mest uenige var de når det gjaldt teksten om standpunktvrdering ($M = 3,28$; $SD = 1,64$), og minst uenige når det gjaldt kompetansemålene ($M = 3,62$; $SD = 1,67$).

Tabell 13

Svar på påstanden: «Skoleledelsen har gitt meg veiledning i hvordan jeg kan bruke disse tekstene for å kvalitetssikre min standpunktvrdering...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	3,82 (1,31)	4,21 (1,41)	3,85 (1,58)	3,91 (1,64)	3,76 (1,65)
Fremmedspråk	4,55 (1,21)	5 (1,26)	4,36 (1,29)	4,82 (1,17)	4,45 (1,21)
KRLE/religion	3,67 (1,15)	3,67 (1,15)	3,33 (0,58)	3,67 (1,15)	3,33 (0,58)
Kroppsøving	2,89 (1,53)	3,11 (1,81)	2,89 (1,53)	2,94 (1,51)	2,67 (1,46)
Kunst og håndverk	4,13 (1,54)	4,31 (1,62)	3,75 (1,65)	4 (1,71)	3,94 (1,48)
Mat og helse	3,5 (2,07)	3,63 (2,2)	3,13 (1,81)	3,13 (1,73)	3,38 (1,92)
Matematikk	3,44 (1,57)	3,54 (1,59)	3,3 (1,46)	3,42 (1,54)	3,48 (1,53)
Musikk	3,7 (1,84)	3,86 (1,84)	3,59 (1,8)	3,59 (1,79)	3,57 (1,79)
Naturfag	2,86 (1,49)	2,9 (1,48)	2,95 (1,5)	3,1 (1,58)	2,71 (1,38)
Norsk	2,89 (1,52)	3,4 (1,74)	3,18 (1,66)	3,23 (1,72)	3,24 (1,67)
Samfunnsfag	3,67 (1,23)	3,92 (1,16)	3,83 (1,47)	3,75 (1,76)	3,58 (1,56)
Annet	4,29 (1,5)	4,14 (1,46)	4 (1,29)	3,86 (1,35)	3,86 (1,68)
<i>VGO</i>					
Engelsk	3,09 (1,62)	3,37 (1,8)	3,23 (1,8)	3,31 (1,68)	2,91 (1,79)
Fremmedspråk	3,48 (1,62)	3,48 (1,62)	3,65 (1,64)	3,48 (1,53)	3,17 (1,61)
Helse og oppvekst	3,59 (1,6)	4,13 (1,6)	3,75 (1,63)	4,03 (1,64)	3,22 (1,58)
KRLE/religion	2,75 (1,13)	3,31 (1,35)	3,13 (1,2)	3,06 (1,12)	3,06 (1,12)
Kroppsøving	3,56 (1,92)	3,44 (1,79)	3,39 (1,61)	3,67 (1,75)	3,22 (1,77)
Kunst og håndverk	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)
Mat og helse	6 (0)	6 (0)	6 (0)	6 (0)	6 (0)
Matematikk	3,13 (1,39)	3,45 (1,51)	3,3 (1,44)	3,33 (1,39)	3,04 (1,38)
Musikk	3,13 (1,81)	3,13 (1,81)	3,13 (1,81)	3,13 (1,81)	3,13 (1,81)
Naturfag	3,23 (1,55)	3,46 (1,62)	3,6 (1,56)	3,71 (1,51)	3,06 (1,66)
Norsk	2,98 (1,63)	3,32 (1,77)	3,3 (1,75)	3,25 (1,77)	3,13 (1,74)
Samfunnsfag	3,76 (1,77)	3,97 (1,81)	3,55 (1,73)	3,73 (1,77)	3,67 (1,88)
Annet	3,42 (1,61)	3,8 (1,63)	3,67 (1,55)	3,64 (1,61)	3,34 (1,66)
<i>Totalt</i>					
Alle deltakere	3,33 (1,6)	3,62 (1,67)	3,46 (1,6)	3,5 (1,63)	3,28 (1,64)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

I tabell 14 gjengis svar på påstander om bruk av tekster i diskusjoner med kollegaer. To av tekstene ser ikke ut til å ha blitt omfattende brukt (fagets kjerneelementer: $M = 3,88$; $SD = 1,44$; teksten om underveisvurdering: $M = 3,88$; $SD = 1,4$), men den typiske læreren var enig i at kompetansemålene ble brukt til dette formålet ($M = 4,86$; $SD = 1,27$). Denne trenden gikk igjen på tvers av fag og trinn.

Tabell 14

Svar på påstanden: «Jeg bruker disse tekstene for å diskutere og kvalitetssikre standpunktvurdering med mine kollegaer...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	4,41 (1,26)	5,09 (1)	4,35 (1,1)	4,56 (1,13)	4,79 (1,15)
Fremmedspråk	4 (1,26)	4,55 (1,29)	3,91 (1,3)	4,09 (1,22)	4,45 (1,44)
KRLE/religion	5 (1)	5,33 (0,58)	5 (1)	5,33 (0,58)	5 (1)
Kroppsøving	3,33 (1,57)	4,89 (1,13)	3,78 (1,4)	3,67 (1,08)	3,83 (1,25)
Kunst og håndverk	3,69 (1,78)	5 (1,41)	4,13 (1,59)	4,25 (1,61)	4,69 (1,45)
Mat og helse	4,63 (1,19)	4,38 (1,69)	3,63 (1,77)	3,75 (1,28)	4,5 (1,69)
Matematikk	4,08 (1,23)	4,78 (1,28)	3,5 (1,27)	3,88 (1,41)	4,1 (1,37)
Musikk	4,19 (1,47)	4,81 (1,58)	3,68 (1,51)	3,86 (1,7)	4,08 (1,72)
Naturfag	3,67 (1,46)	4,67 (1,39)	3,62 (1,32)	3,95 (1,56)	4 (1,34)
Norsk	3,94 (1,3)	4,98 (1)	4,1 (1,24)	4,23 (1,26)	4,81 (1,14)
Samfunnsfag	4,5 (1)	4,75 (0,97)	4 (1,21)	3,92 (1,38)	4,25 (1,36)
Annet	5 (1,15)	5 (1)	4,14 (1,46)	4 (1,63)	4,43 (1,9)
<i>VGO</i>					
Engelsk	3,97 (1,48)	4,89 (1,35)	4,03 (1,6)	4,17 (1,36)	4,17 (1,76)
Fremmedspråk	3,39 (1,47)	4,39 (1,44)	3,65 (1,56)	3,96 (1,4)	4,22 (1,41)
Helse og oppvekst	4,31 (1,38)	5,34 (0,94)	4,28 (1,53)	4,53 (1,44)	4 (1,7)
KRLE/religion	3,5 (1,55)	4,94 (1,44)	3,31 (1,25)	3,81 (1,42)	4,06 (1,57)
Kroppsøving	4,56 (1,2)	5,28 (1,18)	3,89 (1,57)	4,22 (1,63)	4,17 (1,92)
Kunst og håndverk	5 (0)	5 (0)	5 (0)	5 (0)	5 (0)
Mat og helse	6 (0)	6 (0)	6 (0)	6 (0)	6 (0)
Matematikk	3,58 (1,41)	4,93 (1,16)	3,65 (1,3)	4,04 (1,28)	3,68 (1,48)
Musikk	4,25 (0,89)	5 (0,93)	3,63 (1,51)	4,13 (1,25)	3,75 (1,98)
Naturfag	3,4 (1,4)	4,49 (1,48)	3,71 (1,49)	3,94 (1,41)	3,66 (1,7)
Norsk	3,74 (1,37)	4,85 (1,33)	4 (1,34)	4,17 (1,36)	4,24 (1,44)
Samfunnsfag	3,45 (1,5)	4,64 (1,39)	3,73 (1,4)	4 (1,52)	4,12 (1,47)
Annet	3,89 (1,56)	4,85 (1,34)	3,93 (1,47)	4,05 (1,49)	3,86 (1,7)
<i>Totalt</i>					
Alle deltakere	3,88 (1,44)	4,86 (1,27)	3,88 (1,4)	4,09 (1,4)	4,13 (1,54)

Merknad. FK = fagets kjerneelementer; KP = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

Sett under ett var det ikke noen av tekstene som så ut til å bli mye brukt i kommunikasjon med foresatte. I tabell 15 kan vi se at gjennomsnittsverdien for alle tekster var under fire. Et unntak kan vi se for kompetansemålene på ungdomstrinnet, hvor lærere i flere fag meldte at teksten ble brukt i samtale med foresatte. Høyest verdi hadde kroppsøving ($M = 4,83$; $SD = 1,15$).

Tabell 15

Svar på påstanden: «Jeg bruker disse tekstene for å diskutere standpunktvurdering med foresatte...»

	FK	KP	UV	SP	KM
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	3,32 (1,43)	4,18 (1,73)	3,41 (1,42)	3,65 (1,5)	3,97 (1,55)
Fremmedspråk	2,82 (1,08)	3,82 (1,72)	2,91 (1,04)	3,09 (1,3)	3,36 (1,43)
KRLE/religion	4,33 (1,15)	4,33 (1,15)	4 (1)	4 (1)	2,33 (1,15)
Kroppsøving	3 (1,53)	4,83 (1,15)	3,5 (1,2)	3,39 (1,24)	3,83 (1,47)
Kunst og håndverk	3,25 (1,73)	3,75 (1,77)	2,88 (1,54)	2,81 (1,52)	3,38 (1,78)
Mat og helse	3,5 (1,31)	4,63 (1,06)	3,63 (1,3)	3,25 (1,16)	4,63 (1,3)
Matematikk	3,44 (1,34)	4,08 (1,52)	3,1 (1,3)	3,52 (1,5)	3,7 (1,42)
Musikk	3,35 (1,58)	4,14 (1,75)	3,14 (1,42)	3,32 (1,63)	3,59 (1,66)
Naturfag	3,14 (1,8)	4,05 (1,69)	3,1 (1,41)	3,19 (1,36)	3,05 (1,36)
Norsk	2,89 (1,57)	4,27 (1,6)	3,19 (1,61)	3,27 (1,67)	3,77 (1,74)
Samfunnsfag	3,17 (1,64)	4,17 (1,11)	3,58 (1,62)	3,5 (1,62)	3,58 (1,88)
Annet	3,14 (2,12)	3,71 (1,98)	2,57 (1,62)	3,14 (1,68)	3,57 (1,99)
<i>VGO</i>					
Engelsk	2,51 (1,54)	3,11 (1,88)	2,8 (1,68)	3,09 (1,8)	2,8 (1,76)
Fremmedspråk	2,3 (1,61)	2,87 (2,07)	2,48 (1,75)	2,65 (1,85)	2,87 (2,07)
Helse og oppvekst	2,75 (1,61)	3,63 (1,9)	3,03 (1,62)	3,13 (1,74)	2,91 (1,75)
KRLE/religion	1,81 (1,28)	2,69 (1,92)	1,69 (1,14)	2,13 (1,45)	1,94 (1,44)
Kroppsøving	3,06 (1,89)	3,78 (2,29)	2,56 (1,72)	3 (2,03)	2,56 (1,72)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)
Mat og helse	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)
Matematikk	2,5 (1,65)	3,14 (1,89)	2,71 (1,68)	2,75 (1,74)	2,64 (1,69)
Musikk	2,88 (1,89)	3 (2,07)	2,63 (1,69)	2,75 (1,91)	2,38 (1,77)
Naturfag	2,37 (1,57)	3,09 (2,01)	2,43 (1,75)	2,83 (1,89)	2,66 (1,89)
Norsk	2,31 (1,35)	3,14 (1,8)	2,83 (1,62)	2,86 (1,73)	2,8 (1,71)
Samfunnsfag	2,33 (1,61)	2,82 (2,02)	2,58 (1,79)	2,67 (1,87)	2,45 (1,66)
Annet	2,86 (1,7)	3,48 (1,94)	2,95 (1,76)	3,02 (1,78)	2,92 (1,82)
<i>Totalt</i>					
Alle deltakere	2,78 (1,6)	3,56 (1,87)	2,9 (1,61)	3,04 (1,7)	3,06 (1,76)

Merknad. FK = fagets kjerneelementer; KM = kompetansemålene; UV = underveisvurdering; SP = standpunkt, KM = kjennetegn på måloppnåelse.

Karakterskalaen

I del 3 fokuserte spørreundersøkelsen på læreres oppfatning av karakterskalaen og læreres bruk av flere skalasteg ved hjelp av markeringer (pluss, minus, skråstrek). I tabell 16 ser vi svar på påstanden om at avstanden mellom karaktertrinnene er lik. Det var i det nærmeste unison uenighet om dette, noe som indikerer at læreren i denne undersøkelsen var av oppfatningen at noen karaktertrinn er «bredere» eller «smalere» enn andre. Sagt på en annen måte kan det se ut

til at lærerne mente at enkelte karaktertrinn rommet større variasjon i elevkompetanse enn andre. I tabell 17 ser vi svar på påstanden om flere skalasteg. Et maksimumbruk av markeringer med pluss og minus innebærer at skalaen går fra seks trinn (1, 2, ... 6) til 18 trinn (1-, 1, 1+, 2-, 2, 2+, ... 6+).¹ Vi vet ikke hvordan lærerne brukte markeringer, men vi kan se i tabellen at lærerne var enige i at bruk av flere skalatrinn var en forekommende praksis. Størst enighet med påstanden hadde musikk lærere på videregående ($M = 5,88$; $SD = 0,35$). Minst enige var engelsklærere på ungdomstrinn ($M = 3,97$; $SD = 1,66$).

Tabell 16

Svar på påstanden: «Avstand mellom karaktertrinnene er lik»

Fag	U.trinn	VGO
	M (SD)	M (SD)
Engelsk	3,79 (1,39)	3,14 (1,65)
Fremmedspråk	3,18 (1,33)	3,43 (1,38)
Helse og oppvekst	-	3,72 (1,51)
KRLE/religion	3,39 (1,09)	3,25 (1,65)
Kroppsøving	3,63 (1,45)	3,61 (1,29)
Kunst og håndverk	3,5 (1,07)	6 (0)
Mat og helse	3,04 (1,43)	6 (0)
Matematikk	3,19 (1,37)	3,43 (1,55)
Musikk	3,33 (1,49)	3,13 (1,55)
Naturfag	3,06 (1,44)	2,89 (1,6)
Norsk	3,39 (1,09)	3,62 (1,49)
Samfunnsfag	3,33 (1,5)	3,27 (1,48)
Annet	2,71 (1,38)	3,65 (1,52)
<i>Totalt</i>	M (SD)	
Alle deltakere	3,4 (1,49)	

Merknad. U.trinn = ungdomstrinn; VGO = videregående opplæring.

Tabell 17

Svar på påstanden: «Jeg bruker karakterer med markeringer (f.eks. '+', '-')»

Fag	U.trinn	VGO
	M (SD)	M (SD)
Engelsk	3,97 (1,66)	4,69 (1,64)
Fremmedspråk	4,27 (1,79)	5,48 (0,9)
Helse og oppvekst	-	4,53 (1,67)
KRLE/religion	5,67 (0,58)	4,63 (2,19)
Kroppsøving	4 (2,06)	4,06 (2,07)

¹ Ifølge intervjuene deler noen lærere skalaen opp i enda flere kategorier ved å bruke skråstreks-karakterer, slik at det blir 1-, 1, 1+, 1/2, 2/1, 2-, 2 ... 6+.

Tabell 17

Svar på påstanden: «Jeg bruker karakterer med markeringer (f.eks. '+', '-')»

Fag	U.trinn	VGO
	M (SD)	M (SD)
Kunst og håndverk	4,94 (1,39)	6 (0)
Mat og helse	4,25 (1,67)	1 (0)
Matematikk	4,62 (1,59)	5,06 (1,16)
Musikk	4,84 (1,48)	5,88 (0,35)
Naturfag	4,43 (1,83)	4,4 (1,82)
Norsk	4,98 (1,3)	5,06 (1,41)
Samfunnsfag	4,42 (1,93)	4,79 (1,69)
Annet	4,14 (1,95)	4,73 (1,53)
<i>Totalt</i>	M (SD)	
Alle deltakere	4,74 (1,57)	

Merknad. U.trinn = ungdomstrinn; VGO = videregående opplæring.

Arbeid med standpunktvurdering

I den fjerde delen presenterte vi påstander om hva lærerne vektla i forbindelse med fastsettelse av standpunktkarakter. I tabell 18a presenteres svarene relatert til fem påstander: 1) «et gjennomsnitt av hva eleven har vist i løpet av året» (forkortet *Sn* i tabellen), 2) «'store' oppgaver i større grad enn 'små' oppgaver» (*Op*), 3) «resultater fra prøver og arbeid som er gjennomført på skolen i større grad enn resultater på hjemmearbeid og innleveringsoppgaver» (*Sk*), 4) «elevens kompetanse mot slutten av opplæringen i større grad enn hva eleven har vist gjennom året» (*Se*), 5) «oppgaver hvor elever ikke har hatt hjelpemidler tilgjengelig» (*Uh*). Sett under ett var det uenighet i at det ble vektlagt oppgaver uten hjelpemiddel og uenighet i at standpunkt var et gjennomsnitt av elevenes resultater. Det var enighet i at store oppgaver veide tyngre enn små ($M = 4,15$; $SD = 1,14$) og at skoleoppgaver var viktigere enn hjemmearbeid ($M = 4,29$; $SD = 1,36$). Størst enighet var det i påstanden at det viktigste for standpunktkarakteren var elevenes kompetanse mot slutten av opplæringen ($M = 4,82$; $SD = 1,12$).

Tabell 18a

Svar på påstanden: «Når jeg fastsetter standpunktkarakterer vektlegger jeg...»

	Sn	Op	Sk	Se	Uh
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>					
Engelsk	3 (1,21)	4,24 (0,89)	4,68 (1,12)	5 (0,92)	3,88 (1,12)
Fremmedspråk	2,82 (1,6)	3,45 (1,57)	3,91 (1,81)	5,09 (0,7)	3,91 (1,7)
KRLE/religion	2,33 (1,15)	3,33 (0,58)	4,33 (0,58)	4 (1)	2,67 (0,58)
Kroppsøving	3,94 (1,59)	3,83 (1,15)	4 (1,33)	4,22 (1,35)	2,44 (1,42)
Kunst og håndverk	4 (1,21)	4,75 (1,24)	3,75 (1,84)	4,19 (1,17)	2,63 (1,45)
Mat og helse	4 (0,76)	3,75 (0,89)	3,38 (1,3)	4,25 (0,89)	3 (0,76)
Matematikk	3,46 (1,28)	4,26 (1,08)	4,56 (1,15)	4,96 (1,01)	3,44 (1,2)
Musikk	3,62 (1,4)	4,32 (1,2)	4,19 (1,49)	4,7 (1,08)	2,97 (1,32)
Naturfag	3,52 (1,29)	4,33 (1,2)	4,33 (1,32)	4,52 (1,33)	2,95 (1,24)

Tabell 18a

Svar på påstanden: «Når jeg fastsetter standpunktkarakterer vektlegger jeg...»

	Sn	Op	Sk	Se	Uh
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Norsk	3,21 (1,54)	3,95 (1,36)	4,06 (1,44)	4,45 (1,35)	3,21 (1,36)
Samfunnsfag	3,92 (1,16)	4,08 (0,67)	4,33 (0,98)	4,67 (0,78)	3,92 (1,08)
Annet	3 (1,41)	4,71 (0,76)	3,71 (0,76)	4,29 (1,38)	3,14 (1,77)
<i>VGO</i>					
Engelsk	2,74 (1,24)	4,09 (1,07)	4,2 (1,32)	5,11 (0,96)	3,34 (1,39)
Fremmedspråk	2,78 (1,51)	4,39 (1,2)	5,13 (0,81)	5,39 (0,66)	4,04 (1,36)
Helse og oppvekst	2,81 (1,71)	3,56 (1,13)	3,47 (1,46)	4,69 (1,28)	2,59 (1,21)
KRLE/religion	2,88 (1,45)	4,38 (1,09)	4,38 (1,09)	5,06 (0,77)	3,69 (1,08)
Kroppsøving	3,61 (1,29)	4,22 (0,94)	3,94 (1,26)	4,28 (1,27)	3,17 (1,54)
Kunst og håndverk	1 (0)	1 (0)	1 (0)	6 (0)	3 (0)
Mat og helse	1 (0)	4 (0)	4 (0)	5 (0)	3 (0)
Matematikk	2,74 (1,31)	4,13 (1,01)	4,73 (1,14)	4,95 (1,05)	3,38 (1,23)
Musikk	3,63 (1,41)	4,13 (1,13)	4 (1,77)	5,38 (1,06)	2,25 (1,16)
Naturfag	3,09 (1,4)	4,46 (0,85)	4,57 (1,38)	4,29 (1,34)	3,57 (1,22)
Norsk	2,93 (1,42)	4,31 (1,22)	4,64 (1,21)	4,97 (1,07)	3,56 (1,36)
Samfunnsfag	2,91 (1,68)	4,36 (1,17)	4,15 (1,64)	5,06 (0,9)	3 (1,64)
Annet	3,34 (1,5)	3,98 (1,13)	4,04 (1,38)	4,96 (1,07)	3,29 (1,48)
<i>Totalt</i>					
Alle deltakere	3,16 (1,45)	4,15 (1,14)	4,29 (1,36)	4,82 (1,12)	3,31 (1,37)

Merknad. Sn = «et gjennomsnitt av hva eleven har fått til i løpet av året»; Op = «‘store’ oppgaver i større grad enn ‘små’ oppgaver»; Sk = «resultater fra prøver og arbeid som er gjennomført på skolen i større grad enn resultater på hjemmearbeid og innleveringsoppgaver»; Se = «elevens kompetanse mot slutten av opplæringen i større grad enn hva eleven har vist gjennom året»; Uh = «oppgaver hvor elever ikke har hatt hjelpemidler tilgjengelig».

I tabell 18b presenteres resultatene til resterende fire påstander: 6) «muntlig aktivitet og deltakelse i klasserommet i vesentlig grad» (*Ma*), 7) «oppgaver hvor elever har arbeidet individuelt i større grad enn samarbeidsoppgaver» (*In*), 8) «om eleven har utviklet seg fra sitt individuelle utgangspunkt i større grad enn om eleven når målene i læreplanen» (*Ut*), 9) «om eleven har utviklet seg på de områdene der jeg har gitt underveisvurdering» (*Uv*). Den typiske læreren var enig i påstand 6 ($M = 4,15$; $SD = 1,08$) og 9 ($M = 4,12$; $SD = 1,12$) og helte mot enig i påstand 7, som dreide seg om individuelt arbeid kontra gruppearbeid ($M = 3,92$; $SD = 1,19$). Det var stort sett unison uenighet ($M = 3,26$; $SD = 1,33$) i at det var viktigere om eleven har utviklet seg fra sitt individuelle utgangspunkt i større grad enn om eleven når målene i læreplanen. Et unntak som er verdt å merke seg er kroppsøving på videregående ($M = 4,28$; $SD = 0,96$). I dette faget ser det altså ut til at elevens utvikling fra sitt individuelle utgangspunkt spiller en større rolle for standpunktvurderingen enn i andre fag.

Tabell 18b

Svar på påstanden: «Når jeg fastsetter standpunktkarakterer vektlegger jeg...»

	Ma	In	Ut	Uv
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	4,18 (0,9)	4,09 (1,06)	3,47 (1,28)	4,15 (0,99)
Fremmedspråk	4,18 (1,25)	3,64 (1,29)	3 (1)	4 (1,26)
KRLE/religion	4 (1)	3,33 (0,58)	2,33 (1,53)	3,67 (1,53)
Kroppsøving	4,5 (1,2)	2,89 (1,49)	3,83 (1,2)	4,61 (1,04)
Kunst og håndverk	3,63 (0,81)	4 (1,03)	3,19 (1,11)	3,88 (0,96)
Mat og helse	3,63 (1,41)	3 (1,2)	3,38 (1,06)	4,13 (1,13)
Matematikk	3,98 (0,89)	3,98 (0,89)	3,1 (1,3)	3,68 (1,19)
Musikk	4,14 (0,92)	3,49 (1,22)	3,3 (1,33)	4,05 (1,35)
Naturfag	4,38 (1,12)	3,57 (1,21)	2,86 (1,15)	4,14 (1,11)
Norsk	4,34 (0,9)	3,9 (1,2)	2,77 (1,26)	4,02 (1,03)
Samfunnsfag	4,17 (0,94)	3,67 (1,15)	3,25 (1,29)	3,92 (0,67)
Annet	4,71 (0,95)	3,14 (1,35)	3,86 (0,69)	4,29 (0,95)
<i>VGO</i>				
Engelsk	4,4 (1,03)	4 (1,19)	3,57 (1,63)	4,26 (1,22)
Fremmedspråk	4,43 (1,2)	4,22 (0,95)	3,43 (1,27)	4,35 (1,07)
Helse og oppvekst	4,5 (0,95)	3,25 (1,3)	3,47 (1,16)	4,31 (1,15)
KRLE/religion	4,75 (0,77)	3,38 (1,36)	3,19 (1,28)	4,63 (0,89)
Kroppsøving	5,11 (0,96)	3 (1,08)	4,28 (0,96)	4,44 (0,78)
Kunst og håndverk	1 (0)	1 (0)	3 (0)	4 (0)
Mat og helse	3 (0)	3 (0)	2 (0)	3 (0)
Matematikk	3,68 (1,17)	4,3 (0,88)	2,65 (1,16)	3,8 (1,13)
Musikk	3,5 (1,31)	4 (1,41)	2,63 (1,3)	4 (0,76)
Naturfag	3,89 (1,13)	4,06 (1,26)	2,94 (1,37)	3,86 (1,22)
Norsk	4,21 (1,12)	4,21 (1,15)	3,36 (1,21)	4,26 (1,04)
Samfunnsfag	4,27 (0,88)	4,24 (1,35)	3,24 (1,58)	4 (1,48)
Annet	4,07 (1,11)	4,09 (1,14)	3,64 (1,4)	4,3 (1,08)
<i>Totalt</i>				
Alle deltakere	4,15 (1,08)	3,92 (1,19)	3,26 (1,33)	4,12 (1,12)

Merknad. Ma = «muntlig aktivitet og deltakelse i klasserommet i vesentlig grad»; In = «oppgaver hvor elever har arbeidet individuelt i større grad enn samarbeidsoppgaver»; Ut = «om eleven har utviklet seg fra sitt individuelle utgangspunkt i større grad enn om eleven når målene i læreplanen»; Uv = «om eleven har utviklet seg på de områdene der jeg har gitt underveisvurdering».

Samarbeid og kompetanseutvikling om standpunktvurdering

I den femte delen tok deltakerne stilling til påstander om samarbeid med kollegaer samt skoleledelsens innsats for å styrke kompetansen innen standpunktvurdering. Til sammen ble

lærerne presentert for syv påstander (tabell 19a og tabell 19b). De første fire fokuserte på samarbeid: 1) «Jeg og mine kolleger samarbeider om å lage oppgaver i fagene» (forkortet *Lage oppg.* i tabellen), 2) «Jeg og mine kolleger vurderer elevbesvarelser i fellesskap» (*Vurdere*), 3) «Jeg og mine kolleger har møter/arenaer for felles arbeid om vurdering i fagene» (*Arena*), 4) «Samarbeid med andre faglærere er viktig for at min vurdering er pålitelig og gyldig» (*Gyldig*). Resultatene gjengis i tabell 19a, og som det framgår, var lærerne enige i påstandene om samarbeid om å lage oppgaver ($M = 4,34$; $SD = 1,44$), om at det fantes arenaer for felles arbeid ($M = 4,14$; $SD = 1,43$) og at samarbeid var viktig for å styrke pålitelighet og validitet ($M = 4,59$; $SD = 1,23$). Den typiske læreren var ikke enig i at elevbesvarelser ble vurdert i fellesskap ($M = 3,75$; $SD = 1,42$). I tabell 19b kommer det frem at lærerne ikke var enige i påstander om skoleledelsens tilrettelegging. Gjennomsnittet for den første påstanden («Skoleledelsen legger til rette for samarbeid om standpunktvurdering» [*Til rette*]) var 3,57 ($SD = 1,48$). Gjennomsnittet for den andre påstanden (Skoleledelsen jobber systematisk med utvikling av vurderingskompetanse blant lærerne [*Systematisk*]) var 3,62 ($SD = 1,47$) og gjennomsnittet for den tredje påstanden («Skoleledelsen legger til rette for kurs/opplæring i regelverket for vurdering» [*Kurs*]) var 3,35 ($SD = 1,45$).

Tabell 19a

Svar på påstander om samarbeid og kompetanseutvikling om standpunktvurdering

	Lage oppg.	Vurdere	Arena	Gyldig
	M (SD)	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>				
Engelsk	4,68 (1,17)	4,12 (1,27)	3,91 (1,31)	4,85 (0,82)
Fremmedspråk	3,82 (1,66)	3,36 (1,57)	3,82 (1,6)	4 (1,61)
KRLE/religion	4 (1)	2,33 (0,58)	3,33 (1,53)	4,67 (0,58)
Kroppsøving	3,89 (1,49)	3,67 (1,19)	3,56 (1,2)	4,61 (1,46)
Kunst og håndverk	4,75 (1,61)	3,88 (1,59)	3,94 (1,73)	4,31 (1,7)
Mat og helse	4 (1,85)	3,75 (1,91)	3,63 (1,77)	4,25 (1,67)
Matematikk	4,32 (1,36)	3,48 (1,22)	4,02 (1,22)	4,36 (1,21)
Musikk	3,89 (1,68)	3,11 (1,22)	4 (1,35)	4,27 (1,1)
Naturfag	3,43 (1,91)	2,9 (1,48)	3,1 (1,73)	3,67 (1,62)
Norsk	4,5 (1,52)	3,98 (1,29)	3,92 (1,49)	5 (1,02)
Samfunnsfag	4,58 (1,31)	3,67 (1,37)	3,92 (1,16)	4,25 (1,14)
Annet	4,86 (1,46)	4,71 (1,11)	4,71 (0,95)	4,43 (1,4)
<i>VGO</i>				
Engelsk	3,94 (1,16)	3,8 (1,41)	4,17 (1,4)	4,71 (1,13)
Fremmedspråk	4,04 (1,22)	4,13 (0,97)	4,09 (1,24)	4,65 (0,93)
Helse og oppvekst	5,03 (1,12)	4,34 (1,36)	4,88 (1,31)	5,16 (0,88)
KRLE/religion	3,94 (1,73)	3,06 (1,61)	4 (1,71)	4,63 (1,59)
Kroppsøving	4,72 (1,45)	3,94 (1,63)	5,17 (1,04)	4,61 (1,29)
Kunst og håndverk	6 (0)	4 (0)	6 (0)	6 (0)
Mat og helse	6 (0)	6 (0)	6 (0)	6 (0)
Matematikk	4,74 (1,13)	3,61 (1,29)	4,36 (1,34)	4,55 (1,14)
Musikk	5,13 (1,46)	4,88 (1,73)	5,25 (0,89)	5,25 (1,16)

Tabell 19a

Svar på påstander om samarbeid og kompetanseutvikling om standpunktvrdering

	Lage oppg.	Vurdere	Arena	Gyldig
	M (SD)	M (SD)	M (SD)	M (SD)
Naturfag	3,89 (1,57)	3,2 (1,51)	3,63 (1,57)	4,03 (1,52)
Norsk	4,64 (1,35)	4,09 (1,37)	4,58 (1,34)	5,04 (1,08)
Samfunnsfag	4,06 (1,12)	3,33 (1,31)	4,06 (1,32)	4 (1,27)
Annet	4,11 (1,47)	3,78 (1,53)	4,05 (1,46)	4,52 (1,17)
<i>Totalt</i>				
Alle deltakere	4,34 (1,44)	3,75 (1,42)	4,14 (1,43)	4,59 (1,23)

Merknad. Lage oppg. = «Jeg og mine kolleger samarbeider om å lage oppgaver i fagene»; Vurdere = «Jeg og mine kolleger vurderer elevbesvarelser i fellesskap»; Arena = «Jeg og mine kolleger har møter/arenaer for felles arbeid om vurdering i fagene»; Gyldig = «Samarbeid med andre faglærere er viktig for at min vurdering er pålitelig og gyldig»

Tabell 19b

Svar på påstander om samarbeid og kompetanseutvikling om standpunktvrdering

	Til rette	Systematisk	Kurs
	M (SD)	M (SD)	M (SD)
<i>U.trinn</i>			
Engelsk	3,76 (1,33)	4,06 (1,43)	3,82 (1,57)
Fremmedspråk	3,55 (1,37)	4,45 (1,37)	3,73 (1,1)
KRLE/religion	2,33 (0,58)	2,67 (2,08)	2,33 (0,58)
Kroppsøving	3,39 (1,29)	3,22 (1,48)	2,56 (1,1)
Kunst og håndverk	3,56 (1,71)	3,75 (1,48)	3,63 (1,54)
Mat og helse	3,75 (1,49)	3,38 (0,92)	2,5 (1,31)
Matematikk	3,58 (1,33)	3,42 (1,33)	3,24 (1,3)
Musikk	3,57 (1,52)	4,05 (1,61)	3,59 (1,57)
Naturfag	3,05 (1,83)	2,9 (1,89)	2,86 (1,77)
Norsk	3,53 (1,49)	3,44 (1,48)	2,98 (1,38)
Samfunnsfag	3,42 (1,38)	3,33 (1,07)	3,75 (1,14)
Annet	4,14 (1,57)	4,43 (0,79)	4,29 (0,95)
<i>VGO</i>			
Engelsk	3,43 (1,54)	3,69 (1,68)	3,54 (1,52)
Fremmedspråk	3,39 (1,47)	1,24 (4,65)	3,43 (1,53)
Helse og oppvekst	4,13 (1,56)	1,31 (5,16)	3,47 (1,48)
KRLE/religion	3,19 (1,6)	1,71 (4,63)	3,31 (1,58)
Kroppsøving	3,94 (1,47)	1,04 (4,61)	3,61 (1,5)
Kunst og håndverk	6 (0)	0 (6)	4 (0)
Mat og helse	6 (0)	0 (6)	6 (0)
Matematikk	3,76 (1,31)	1,34 (4,55)	3,4 (1,33)
Musikk	4,38 (1,19)	0,89 (5,25)	4,25 (1,28)
Naturfag	2,97 (1,52)	1,57 (4,03)	3,03 (1,44)

Tabell 19b

Svar på påstander om samarbeid og kompetanseutvikling om standpunktvrdering

	Til rette	Systematisk	Kurs
	M (SD)	M (SD)	M (SD)
Norsk	3,58 (1,56)	1,34 (5,04)	3,2 (1,46)
Samfunnsfag	3,36 (1,5)	1,32 (4)	3,45 (1,58)
Annet	3,58 (1,48)	1,46 (4,52)	3,43 (1,45)
<i>Totalt</i>			
Alle deltakere	3,57 (1,48)	3,62 (1,47)	3,35 (1,45)

Merknad. Til rette. = «Skoleledelsen legger til rette for samarbeid om standpunktvrdering»; Systematisk = «Skoleledelsen jobber systematisk med utvikling av vurderingskompetanse blant lærerne»; Kurs = «Skoleledelsen legger til rette for kurs/opplæring i regelverket for vurdering».

SAMMENDRAG

Formålet med denne undersøkelsen var å kartlegge læreres arbeid med standpunktvrdering. Vi hadde følgende forskningsspørsmål: *Hvilke vurderingstekster kjenner lærere til og hvilke vurderingstekster bruker lærere? Hva kjennetegner praksis i standpunktvrdering? Hvordan oppfatter lærere samarbeid og skoleledelsens tilbud og krav?* Resultatene fra undersøkelsen kan sammenfattes i punktene nedenfor:

- Fra del 1 av undersøkelsen kan vi trekke slutningen at den typiske læreren kjenner til og bruker alle de fem tekstene som er inkludert i undersøkelsen. Den mest brukte teksten var kompetansemål for fagene. De minst brukte var tekstene om standpunktvrdering og underveisvrdering. Undersøkelsen lyktes ikke i å klarlegge faktorer som ville ført til økt bruk; den typiske læreren mente at hverken mer tid, større relevans, økt grad av konkretisering eller mer kunnskap ville gjort at tekstene ble mer brukt.
- I del 2 så vi at skoleledere samlet sett så ut til å forvente bruk av de fem sentrale tekstene i vurderingsarbeidet, samtidig som lærerne samlet sett var uenige i at de hadde fått veiledning i hvordan de kan brukes. Lærerne oppga også at de brukte kompetansemål for å diskutere og kvalitetssikre vurdering sammen med kollegaer. Samlet sett ble ikke noen av tekstene mye brukt i kommunikasjon med foresatte.
- Del 3 viste at lærerne ikke oppfattet at trinnene i karakterskalaen har like avstand, men at noen skalasteg rommet «mer kompetanse». Del 3 viste også at det var vanlig å bruke flere skalasteg (med hjelp av markeringer som f.eks. pluss og minus).
- I del 4 kunne vi se at lærerne var enige i at «store oppgaver» veide tyngre enn «små oppgaver» i forbindelse med standpunktvrdering. Det fremkom også at skoleoppgaver var viktigere enn hjemmearbeid, at elevens kunnskap mot slutten av året veide mest, og at muntlig aktivitet og elevens utvikling på områder som har vært tematisert i underveisvrdering var viktig for standpunktvrdering.
- Den femte delen indikerte at lærerne samarbeidet med kollegaer i vurderingsarbeid, men at de ikke opplevde at skoleledelsen la til rette for systematisk arbeid.

Det er viktig å være bevisst på at formålet med denne undersøkelsen *ikke* var å evaluere læreres arbeid opp mot lovtekst eller andre styringsdokumenter. Vi tror likevel at fagpersoner i Utdanningsdirektoratet og i andre stillinger kan finne det interessant å reflektere rundt tallene for gruppen som helhet og for enkeltfag, samt vurdere i hvilken grad de representerer en praksis som er i tråd med intensjonene.

Undersøkelsens begrensninger

Vi vil peke på noen begrensninger som er viktige å legge til grunn for tolking og anvendelse av resultatene. For det første vet vi ikke i hvilken grad deltakerne representerer populasjonen lærere på ungdomstrinn og i videregående, hverken i de tolv fagene vi hadde spesifisert på forhånd eller i de mange fagene som inngår i sekkeposten «andre fag». For det andre bygger resultatene på selvrapportering av praksiser som er lovfestet. Det kan gjøre det lite fristende å presentere sin egen praksis som avvikende fra gjeldende normer. For eksempel vil det kanskje oppleves som uaktuelt å «innrømme» dårlig kjennskap til sentrale styringsdokumenter. Variasjonen i svarene tyder på at ikke alle deltakere har opplevd det slik, men denne kjensgjerningen kan ha bidratt til at mange så seg enige i spørsmål om kjennskap til og bruk av de fem tekstene. En tredje begrensning ligger i spørreskjemaets oppsett, som ikke fanget *hvordan* eller *hvor mye* lærere arbeidet med de fem tekstene, samarbeidet med andre eller tok hensyn til hva eleven kunne på slutten av året.

Avslutning

Den viktigste teksten, slik det fremstår på bakgrunn av denne undersøkelsen, var kompetansemålene i fagene. Dersom resultatene skal brukes som en bakgrunn for kompetanseheving for lærere, kan det være verdt å benytte kompetansemålene som en inngang for arbeid med de andre tekstene, da kompetansemålene ser ut til å fungere godt både som et verktøy for enkeltlæreren og som et middel i kommunikasjon med kollegaer, elever og foresatte.

VEDLEGG 11 RAPPORT FRA SPØRREUNDERSØKELSE 2 OM PRESS OG TILLIT I STANDPUNKTVURDERING

Gustaf B. Skar, Anne Holten Kvistad og Lise Vikan Sandvik

INNLEDNING

Denne rapporten presenterer funn fra en kvantitativ spørreundersøkelse som inngår som en opsjonsdel i prosjektet Forskning på standpunktvurdering, gjennomført på oppdrag for Utdanningsdirektoratet. Undersøkelsen bygger videre på hovedundersøkelsen som er presentert i den øvrige kvantitative rapporteringen (se vedlegg 10), men skiller seg fra denne ved å ha et særskilt fokus på press, tillit, karakterinflasjon og styringsrelasjoner i arbeidet med standpunktvurdering.

Mens hovedrapporten primært kartlegger læreres bruk av vurderingstekstene, praksiser i sluttvurdering og organisering av samarbeid, retter denne delstudien oppmerksomheten mot systemiske spenninger: hvordan lærere opplever forventninger og press fra elever, foresatte og skoleledelse, hvordan de vurderer utviklingen i karakternivåer over tid, og hvordan ledelse og profesjonsfellesskap fungerer som støtte i arbeidet.

Formålet med undersøkelsen var å gi et mer nyansert bilde av læreres handlingsrom i sluttvurdering, og å undersøke hvordan accountability-mekanismer og profesjonell autonomi faktisk erfares i praksis. Undersøkelsen ble gjennomført blant lærere på ungdomstrinnet og i videregående opplæring, på tvers av fag og regioner. Formålet var å belyse læreres opplevelse av press og tillit i karaktersetting, oppfatninger om utvikling i karakternivå over tid og skolens støtte i krevende vurderingssituasjoner

Resultatene fra undersøkelsen kan sammenfattes i punktene nedenfor:

- Et stort flertall av lærerne rapporterer høy grad av tillit fra elever, foresatte og skoleledelse til deres standpunktvurderinger, og et lavt nivå av formelle klager.
- Samtidig opplever mange lærere forventninger om høyere karakterer fra elever, mens press fra foresatte og skoleledelse vurderes som mindre utbredt.
- En majoritet av lærerne oppgir at de opplever at kravene for å bestå (karakter 2) er blitt lavere, og at vurderingspraksis i fagmiljøer har beveget seg i en mer generøs retning, med tydelige forskjeller mellom fag.
- Funnene viser et spenn mellom lav selvrapportert medvirkning til karakterinflasjon og en utbredt oppfatning av at tersklene for bestått og høye karakterer er senket over tid.
- Lærerne beskriver i hovedsak skoleledelsen som støttende i konflikt- og klagesituasjoner, men rapporterer i mindre grad at ledelsen arbeider systematisk og forebyggende med kvalitetssikring av vurderingspraksis.
- Mange lærere forteller at de tilpasser praksis under press, blant annet gjennom økt dokumentasjon, tydeligere begrunnelser og tettere dialog med kollegaer i vanskelige saker.

- Analysene viser at fag og fagklynge har betydning for hvordan press og vurderingsutfordringer oppleves, mens kjønn og erfaring forklarer lite av variasjonen. Dette peker mot at vurderingskulturer i fagene er en sentral forklaringsfaktor.
- Funnene viser også at lærere i stor grad står alene i krevende vurderingssituasjoner, og at støtte fra profesjonsfellesskap varierer betydelig mellom skoler og fag.

METODE

Deltakere

Deltakere var 636 ungdomsskolelærere og 580 lærere fra videregående skole. Kvinner utgjorde 67 % av utvalget og menn 32 % av utvalget (1 % ønsket ikke å oppgi kjønn). Deltakernes gjennomsnittsalder var 46.5 år ($SD = 10.7$) og gjennomsnittslæreren hadde 16.9 års erfaring i skolen ($SD = 10.6$). Lærerne representerte også ulike fag. Ved å analysere deltakernes fritekstsvare kunne vi utlede at følgende fag var representert: HO, Engelsk, Kroppsøving, Mat og helse, Matematikk, Musikk, Naturfag, Norsk og Samfunnsfag. Videre kunne vi utlede at følgende «fagklynger» var representert: Praktisk-estetiske fag, realfag, samfunnsfag, språkfag og yrkesfag. Lærerne representerte skoler i urbane strøk (48 %), forsteder (19 %) og på landsbygd (33 %) og samtlige regioner i Norge. Videre representerte de skoler av varierende størrelse: 421 lærere (35 %) underviste på skoler med 100–300 elever, 450 lærere (37 %) på skoler med 400–600 elever. Resterende 345 lærere (28 %) underviste på skoler med 700 elever eller fler.

Tabell 1 gir oversikt over deltakers kjønn og skolens plassering. Tabell 2 gjengir gjennomsnittsalder og erfaring. Tabell 3a og tabell 3b presenterer fag og fagklynger representert. Tabell 4 viser fordelingen av lærere per region. Som det framkommer, var noen regioner bedre representert enn andre (for eksempel utgjorde lærere i Trøndelag 23 % av utvalget, mens lærere i Vestfold utgjorde om lag 1 %).

Tabell 1
Deltakere

Variabel		Alle		U.tr.		Vgs	
		N	Pst	N	Pst	N	Pst
Kjønn	Kvinne	810	67%	427	67%	383	67%
	Mann	395	32%	208	33%	187	33%
Plassering	By	580	48%	232	36%	348	60%
	Forstad	231	19%	152	24%	79	14%
	Landsbygd	405	33%	252	40%	153	26%

Tabell 2
Deltakernes alder og erfaring

Gruppe	N	Alder		Erfaring	
		M	SD	M	SD
Alle	1216	46.5	10.7	16.9	10.6
U.tr	636	45.4	10.8	16.9	10.2
Vgs	580	47.8	10.5	16.9	11.0

Tabell 3a

Fag

Fag	Alle	U.tr.	Vgs
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
HO	41 (8.6)	–	41 (20.4)
Engelsk	1 (0.2)	–	1 (0.5)
Kroppspøving	39 (8.2)	22 (7.9)	17 (8.5)
Mat og helse	30 (6.3)	30 (10.8)	–
Matematikk	108 (22.6)	58 (20.9)	50 (24.9)
Musikk	11 (2.3)	11 (4)	–
Naturfag	38 (7.9)	21 (7.6)	17 (8.5)
Norsk	180 (37.7)	105 (37.9)	75 (37.3)
Samfunnsfag	30 (6.3)	30 (10.8)	–
Missing	738	359	379

Merknad. Prosent er basert på de deltakere for hvilke fag/fagklynge ble registrert

Tabell 3b

Fagklynger

Fagklynge	Alle	U.tr.	Vgs
	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
Praktisk-estetiske fag	168 (13.9)	119 (18.8)	49 (8.6)
Realfag	258 (21.4)	150 (23.7)	108 (18.9)
Samfunnsfag	179 (14.9)	66 (10.4)	113 (19.8)
Språkfag	484 (40.2)	298 (47.1)	186 (32.5)
Yrkesfag	116 (9.6)	–	116 (20.3)
Missing	11	3	8

Merknad. Prosent er basert på de deltakere for hvilke fag/fagklynge ble registrert

Tabell 4

Regioner representert

Region	Alle		U.tr.		Vgs	
	<i>n</i>	Pst	<i>n</i>	Pst	<i>n</i>	Pst
Agder	51	4%	43	7%	8	1%
Akershus	107	9%	86	14%	21	4%
Buskerud	29	2%	25	4%	4	1%
Finnmark	23	2%	22	3%	1	0%
Innlandet	60	5%	47	7%	13	2%
Møre og Romsdal	68	6%	40	6%	28	5%
Nordland	39	3%	35	6%	4	1%
Oslo	194	16%	75	12%	119	21%
Rogaland	181	15%	44	7%	137	24%
Telemark	5	0%	2	0%	3	1%
Troms	33	3%	10	2%	23	4%
Trøndelag	285	23%	82	13%	203	35%
Vestfold	7	1%	6	1%	1	0%
Vestland	110	9%	96	15%	14	2%
Østfold	24	2%	23	4%	1	0%

Vi har dessverre ikke tilgang til data som kan brukes for å estimere deltakernes representativitet for en større gruppe lærere. Likevel vurderer vi utvalget som så sammensatt at det er trolig at erfaringene til disse deltakerne ikke er eksklusive for kun dette utvalget. Lærerne var i varierende alder, hadde varierende mengde erfaring, representerte by og land, alle regioner og store og små skoler. I tillegg representerte de ulike fag.

Rekruttering

For å rekruttere deltakere sendte prosjektleder følgende epost til postmottak ved samtlige ungdomsskoler og videregående skoler i Norge.

Hei

Jeg er professor ved Skrivesenteret, NTNU. Jeg leder et forskningsprosjekt, Standpunktprosjektet, som er et forskningsprosjekt ved NTNU som gjennomføres på oppdrag av Utdanningsdirektoratet. Undersøkelsen går til alle lærere i videregående og i ungdomsskolen. En sentral målsetting er å undersøke læreres praksiser og prosesser knyttet til standpunktvurdering, inkludert karakterfastsetting, i ulike fag. Det tar maksimalt ti minutter å svare på undersøkelsen. Kunne dere ha sendt ut undersøkelsen til deres lærere? Det hadde vært til stor hjelp for oss.

Her er lenke til spørreundersøkelsen: <https://nettskjema.no/a/510353>

På forhånd takk!

Lise Vikan Sandvik
Prosjektleder for standpunktprosjektet
Professor (Ph.D.)
Skrivesenteret, NTNU
Tel: +4791785002.

Spørreundersøkelsen mottok svar i perioden 10. april 2025–11. november 2025.

Instrument for datainnsamling

Data ble samlet inn gjennom en spørreundersøkelse som besto av 57 items, som oftest formulert som påstander. Alle spørsmål med tilhørende gjennomsnittsverdier presenteres i tabellene nedenfor.

Elleve av itemene var demografiske (for eksempel alder, erfaring med karaktersetting, fag og skolestørrelse). Seks items omhandlet «tillit og klager». Her ble deltakerne bedt om å angi på en seksdelt skala («helt uenig», «ganske uenig», «litt uenig», «litt enig», «ganske enig», «helt enig») i hvilken grad de opplevde at elever, foresatte og skoleledelse hadde tillit til deres standpunktvurdering. De oppga også i hvilken grad de stolte på kollegaers karaktersetting, om de mottok klager, og om de fikk flere klager enn kollegaer.

Karakterpress-delen besto av 18 items. De første seks brukte samme skala som tillitsitems og ba deltakerne angi i hvilken grad de følte press om å sette høyere karakterer enn det de mente det var faglig grunnlag for. Én påstand gjaldt felles praksiser ved egen skole. De tre neste påstandene ba deltakerne oppgi hvor ofte de hadde blitt forsøkt påvirket til å endre en standpunkt-karakter («aldri», «én gang», «2–5 ganger», «6–10 ganger», «11–15 ganger», «mer enn 16 ganger»). De siste ni påstandene ba deltakerne angi grad av enighet om ulike aspekter ved karakterpress, for

eksempel om press er blitt diskutert med skoleledelse eller om press har ført til økt dokumentasjon.

Spørreskjemaet inneholdt også syv påstander om ledelsesstøtte. Her anga deltakerne i hvilken grad de opplevde at skoleledelsen støttet dem i arbeidet med karaktersetting. Et eksempel er følgende påstand: «Skoleledelsen støtter meg i arbeidet med standpunktvurdering.»

Videre inneholdt skjemaet elleve items om karakterinflasjon, som for eksempel: «Jeg setter høyere standpunktkarakterer nå enn for noen år siden uten at elevene har blitt flinkere eller viser høyere faglig kompetanse.» Også her anga deltakerne grad av enighet.

To påstander handlet om samarbeid og kompetanseutvikling («Jeg diskuterer standpunktkarakterer med kolleger før de fastsettes» og «Jeg har tilstrekkelig kompetanse til å sette rettferdige og faglig godt begrunnede standpunktkarakterer»).

Til slutt ble deltakerne bedt om å svare på om de hadde deltatt i kompetanseutvikling knyttet til standpunktvurdering, og om de kunne tenke seg å delta i slik kompetanseutvikling fremover.

Prosedyre

Alle deltakere gjennomførte spørreundersøkelsen på egen maskin (f.eks. datamaskin, telefon) på et selvvalgt tidspunkt. Deltakerne fikk følgende instruks før de svarte på spørsmål:

Forskningsprosjekt om standpunktvurdering

Standpunktprosjektet er et forskningsprosjekt ved NTNU som gjennomføres på oppdrag av Utdanningsdirektoratet. Et av formålene med prosjektet er å undersøke læreres praksiser og prosesser knyttet til standpunktvurdering i ulike fag.

I denne spørreundersøkelsen skal du tenke på ett bestemt fag når du svarer på spørsmålene. Dette faget velger du selv. Vi ønsker at du svarer oppriktig på spørsmålene, slik at vi kan danne oss et mest mulig riktig bilde av praksis rundt standpunktvurdering. Det er ikke riktige eller gale svar på spørsmålene, og du vil ikke kunne identifiseres i etterkant.

Takk på forhånd for at du tar deg tid til å svare på spørsmålene.

Utover dette fikk deltakerne ingen ytterligere instruks enn det som sto i hvert enkelt item. Etter at deltakeren hadde levert sine svar, fikk vedkommende denne meldingen: «Takk for din deltakelse. Ta kontakt med Gustav Skar (e-post) dersom du har kommentarer eller spørsmål.»

Vi mottok to henvendelser: Den første gjaldt svaralternativene i ett av spørsmålene og etterlyste en begrunnelse for disse. Den andre gjaldt frivilligheten i å delta. At kun to av 1216 deltakere (0,2 %) tok kontakt, vurderer vi som et tegn på at spørreundersøkelsen ble oppfattet som forståelig, i den forstand at det var tydelig hva som skulle gjøres.

Analyse

All data ble bearbeidet og analysert med pakkene «Tidyverse» (Wickham et al., 2019) og «emmeans» (Lenth & Piaskowski, 2017) i R (R Core Team, 2025). Vi brukte også basefunksjoner i R. Items med skalaen «helt uenig» til «helt enig» ble rekodet til verdier 1–6 («helt uenig» = 1, «helt

enig» = 6), slik at 1–3 representerte «uenig» og 4–6 representerte «enig». Vi gjør oppmerksom på at en slik gruppering kan skjule variasjon i svarfordelingen. For flere påstander er standardavviket relativt høyt, noe som tyder på at erfaringer og vurderinger varierer mellom lærere og skoler, selv der gjennomsnittet samlet sett peker i én retning. I resultatpresentasjonen kommenteres derfor spredning i svar der dette er særlig relevant. Samtidig er gjennomsnittsskårer og gruppering av svar nyttige for å gi en samlet oversikt over hovedtendenser i materialet og for å kunne sammenligne påstander på tvers av grupper. I dette vedlegget brukes derfor både gjennomsnitt, standardavvik og andeler «enig/uenig» som komplementære mål. Der gjennomsnittet ligger nær midtpunktet eller der standardavviket er høyt, legges det særlig vekt på å beskrive funnene som uttrykk for et mer delt svarbilde, og resultatene tolkes som indikasjoner på variasjon i erfaringer og praksiser snarere enn entydig enighet eller uenighet.

Skjemaet inneholdt ett item med fritekstsvar, nemlig faget læreren underviser i. Spørsmålet lød: «Hvilket fag underviser du i og ønsker å svare for i denne undersøkelsen?» Selv om deltakerne ble bedt om å oppgi ett fag, valgte flere å oppgi flere. Vi bearbeidet derfor svarene på to måter:

1. Vi laget en fagklyngevariabel, og
2. vi laget en variabel som kun inkluderte svar der det var oppgitt ett fag, og der dette faget var blant de som var i fokus for undersøkelsen (Norsk, Matematikk, Naturfag, Helse og oppvekst, Kroppsøving, Musikk [ungdomsskole], Samfunnsfag [ungdomsskole], Mat og helse [ungdomsskole]).

Vi brukte følgende fagklynger: språkfag, samfunnsfag, realfag og praktisk-estetiske fag.

Alle de 57 itemene ble analysert med deskriptiv statistikk. For et utvalg av itemene testet vi også forskjeller mellom fag, kjønn og erfaring. Disse testene ble gjennomført ved hjelp av regresjonsanalyse med basefunksjonen *lm* i R samt *emmeans* fra pakken «emmeans». De aktuelle påstandene og begrunnelsene for å inkludere dem var følgende:

- «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig.» Dette ble vurdert som relevant fordi det kvalitative materialet viser tydelige forskjeller i pressnivå mellom fag.
- «Min skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet.» Itemet ble vurdert som relevant fordi intervjudata viser at fagmiljøer varierer i hvor etablerte og konsistente rutinene er. Enkelte fagmiljøer, som norsk og matematikk, beskriver mer systematiserte praksiser enn andre.
- «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere.» Dette ble vurdert som relevant fordi det kvalitative materialet viser tydelige forskjeller mellom fag. I kroppsøving, mat og helse og helse- og oppvekstfag rapporterer lærere at det ofte er slik at det er nok å være til stede i timene for å berge en ståkarakter, noe som i praksis kan senke terskelen for å bestå. I matematikk og i noen grad norsk beskrives andre vurderingspraksiser.
- «Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer.» Det kvalitative materialet viser at tendenser til karakterinflasjon særlig kommer til uttrykk i enkelte fag. I matematikk rapporterer lærere i mindre grad om endrede normer, mens flere praktisk-estetiske fag (musikk, kunst og håndverk, mat og helse) beskriver en annen vurderingskultur der «generøsitet» forstås annerledes enn i teoretiske fag.

Resultat

Av svarene på påstander relatert til tillit og klager kan vi se at deltakerne var klart enige i at elever, foresatte og skoleleder hadde tillit til deres standpunktvrdering. De var uenige i at de mottok klager og at de oftere mottok klager enn sine kollegaer. Se tabell 5.

Tabell 5

Påstander relatert til tillit og klager

Påstand	Alle		U.tr		Vgs	
	M	SD	M	SD	M	SD
Elever har tillit til min standpunktvrdering	5.41	0.70	5.46	0.64	5.35	0.77
Foresatte har tillit til min standpunktvrdering	5.36	0.71	5.38	0.66	5.34	0.77
Skoleledelsen har tillit til min standpunktvrdering	5.72	0.66	5.82	0.49	5.62	0.80
Jeg stoler på mine kollegers standpunktvrdering	5.35	0.82	5.36	0.75	5.32	0.89
Jeg får klager på min standpunktvrdering	1.57	1.01	1.54	1.02	1.59	1.00
Jeg får oftere klager på min standpunktvrdering enn mine kolleger	1.20	0.62	1.17	0.56	1.24	0.68

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

I tabell 6 gjengis påstander relatert til karakterpress og elevers «faktiske» faglige nivå. Gjennomsnittslæreren helte mot å si seg enig i at det var forventninger fra elever om å sette en høyere standpunktkarakter enn læreren mente er faglig riktig. Gjennomsnittet ligger like under grensen for enighet ($M = 3,77$), samtidig som spredningen i svarene er relativt stor ($SD = 1,45$). Dette tyder på at enkelte lærere opplever tydelige forventninger om høyere karakterer, mens andre i liten grad kjenner igjen dette. Lærerne var uenige i at foresatte, skoleledelse og skoleeier ha uttrykk for forventninger om å sette høyere standpunktkarakterer enn deltakeren mente er faglig riktig. Når det gjelder hvorvidt skolen har en tydelig, felles praksis for å håndtere press, peker svarene på betydelig variasjon. Gjennomsnittet ligger noe under grensen for enighet ($M = 3,20$), men med høy spredning ($SD = 1,50$). Dette kan indikere at noen skoler har etablerte rutiner som oppleves tydelige for lærerne, mens andre i mindre grad har en felles praksis.

Tabell 6

Påstander relatert til karakterpress og faglig nivå

Påstand	Alle		U.Tr		Vgs	
	M	SD	M	SD	M	SD
Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	3.77	1.45	3.66	1.48	3.89	1.40
Jeg opplever forventninger fra foresatte om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	2.82	1.51	3.09	1.50	2.52	1.47
Jeg opplever forventninger fra skoleledelse om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	1.90	1.29	1.53	0.96	2.31	1.48
Jeg opplever forventninger fra skoleeier om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	1.81	1.28	1.42	0.88	2.23	1.50

Jeg opplever at det forventes at standpunktkarakterene jeg setter, skal samsvare med eksamenskarakterene i faget	3.24	1.53	3.25	1.55	3.23	1.52
Mine skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet	3.20	1.50	3.38	1.48	3.00	1.49

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

Tabell 7 presenterer påstander relatert til press i karaktersetting. Tabellen viser at lærerne samlet sett rapporterer høy grad av trygghet i egne vurderinger dersom de opplever press i forbindelse med karaktersetting. Samtidig viser svarene variasjon mellom lærere og skoler i hvordan slike situasjoner håndteres. Mange oppgir at de diskuterer saken med kollegaer og i noen tilfeller skoleledelse, og at de legger vekt på å sikre et grundigere vurderingsgrunnlag. Lærerne rapporterer gjennomgående lav tilslutning til at press i seg selv fører til endrede vurderinger, men spredningen i svarene på flere påstander indikerer at press og håndteringsmåter oppleves ulikt. Lærerne rapporterer i gjennomsnitt at de søker støtte i skolens retningslinjer ved press ($M = 4,14$), men svarene varierer betydelig ($SD = 1,59$). Dette kan peke mot at retningslinjer fungerer som et aktivt støtteverktøy ved noen skoler, mens de i mindre grad brukes eller oppleves relevante ved andre.

Tabell 7

Påstander relatert til ulike aspekter av karakterpress

Påstand	Alle		U.Tr		Vgs	
	M	SD	M	SD	M	SD
<i>Ved eventuelt press i karaktersetting ...</i>						
... føler jeg meg trygg og står ved mine vurderinger	5.36	0.78	5.41	0.71	5.30	0.84
... diskuterer jeg saken med kolleger	5.65	0.70	5.72	0.56	5.57	0.81
diskuterer jeg saken med skoleledelse	4.21	1.58	4.51	1.50	3.90	1.59
... opplever jeg usikkerhet i vurderingsarbeidet	2.81	1.40	2.73	1.36	2.89	1.43
... endrer jeg vurderinger for å unngå ubehag	1.89	1.14	1.87	1.15	1.92	1.14
... dokumenterer jeg mine vurderinger grundigere enn vanlig	4.75	1.26	4.82	1.20	4.67	1.31
... søker jeg støtte i skolens retningslinjer for vurdering	4.14	1.59	4.40	1.51	3.88	1.63
... opplever jeg at skolens ledelse gir tydelig støtte i slike situasjoner	4.78	1.34	5.03	1.17	4.50	1.45
... har jeg opplevd at presset har påvirket min vurderingspraksis over tid	2.73	1.44	2.70	1.44	2.76	1.44

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

Tabellene 8a og 8b gjengir lærernes svar på spørsmål om hvor mange ganger de har blitt forsøkt påvirket å endre standpunktkarakterer av elever, foresatte og skoleledelse. Blant ungdomstrinns lærere rapporterte 69 % å ha blitt forsøkt påvirket av elever én eller flere ganger og 50.9 % av foresatte én eller flere ganger. Lærerne hadde i mindre utstrekning blitt forsøkt påvirket av skoleledere; 11.5 % hadde opplevd det én eller flere ganger. Blant lærere på videregående var tendensene de samme, selv om flere (79.5 %) hadde opplevd seg forsøkt påvirket av elever, og færre (29.3 %) av foresatte. Det var flere lærere på videregående (27.1 %) som hadde opplevd å bli forsøkt påvirket av skoleledelse.

Tabell 8a

Antall ganger deltakere fra *ungdomsskole* er blitt forsøkt påvirket til å endre standpunktkarakterer. Verdiene er prosent.

Påstand	a	b	c	d	e	f
Jeg har blitt forsøkt påvirket av elever til å endre standpunktkarakterer	31	11.8	35.8	10.5	2.2	8.6
Jeg har blitt forsøkt påvirket av foresatte til å endre standpunktkarakterer	49.1	20.4	22.8	5.5	1.1	1.1
Jeg har blitt forsøkt påvirket av skoleleder til å endre standpunktkarakterer	88.5	8.5	2.5	0.5	0	0

Merknad. a = Aldri; b = Én gang; c = 2–5 ganger; d = 6–10 ganger ; e = 11–15 ganger; f = 16+ ganger

Tabell 8b

Antall ganger deltakere fra *videregående skole* er blitt forsøkt påvirket til å endre standpunktkarakterer. Verdiene er prosent.

Påstand	a	b	c	d	e	f
Jeg har blitt forsøkt påvirket av elever til å endre standpunktkarakterer	20.5	13.3	37.1	14.8	5	9.3
Jeg har blitt forsøkt påvirket av foresatte til å endre standpunktkarakterer	70.7	15.9	12.2	1	0.2	0
Jeg har blitt forsøkt påvirket av skoleleder til å endre standpunktkarakterer	72.9	10	13.3	3.1	0.5	0.2

Merknad. a = Aldri; b = Én gang; c = 2–5 ganger; d = 6–10 ganger ; e = 11–15 ganger; f = 16+ ganger

I tabell 9 presenteres gjennomsnitt og standardavvik for påstander relatert til ledelsesstøtte. Lærerne rapporterer samlet sett relativt høy støtte fra skoleledelsen i arbeidet med standpunktvurdering ($M = 4,73$). Samtidig viser standardavviket ($SD = 1,34$) at opplevelsen av støtte varierer betydelig, noe som kan tyde på store forskjeller mellom skoler i hvordan vurderingsarbeidet følges opp og støttes. Lærerne var enige i at ledelsen støttet arbeidet med standpunktvurdering, at det fantes tydelige retningslinjer, at ledelsen støttet den enkelte læreren i forbindelse med uenighet mellom lærer og elev/foresatte om en standpunktkarakter og at skoleledelsen la til rette for faglige diskusjoner om vurderingspraksis samt ga veiledning dersom læreren sto i en vanskelig vurderingssituasjon. Lærerne var uenige i at skoleledelsen fulgte med på vurderinger for å sikre en faglig forsvarlig praksis og at leder ga etter for press uten å støtte læreren dersom det oppsto uenighet mellom lærer og elev/foresatte om en standpunktkarakter.

Tabell 9

Påstander relatert til ledelsesstøtte

Påstand	Alle		U.Tr		Vgs	
	M	SD	M	SD	M	SD
Skoleledelsen støtter meg i arbeidet med standpunktvurdering	4.73	1.34	4.92	1.27	4.52	1.38
Skoleledelsen har tydelige retningslinjer for vurderingsarbeidet ved min skole	4.09	1.41	4.22	1.37	3.96	1.45
Skoleledelsen følger med på mine vurderinger for å sikre en faglig forsvarlig praksis	3.00	1.47	3.06	1.46	2.92	1.48
Dersom det oppstår uenighet mellom lærer og elev/foresatte om en standpunktkarakter, støtter leder lærerens vurdering	4.91	1.05	5.11	0.92	4.69	1.14

Dersom det oppstår uenighet mellom lærer og elev/foresatte om en standpunktkarakter, gir leder etter for press uten å støtte læreren	1.93	1.17	1.81	1.14	2.06	1.19
Skoleledelsen legger til rette for faglige diskusjoner om vurderingspraksis	4.62	1.26	4.68	1.25	4.56	1.27
Skoleledelsen gir meg veiledning dersom jeg står i en vanskelig vurderingssituasjon	4.42	1.32	4.50	1.31	4.34	1.32

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

I påstander relatert til karakterinflasjon var lærerne enige i to påstander («Standpunktkarakterene jeg setter gjenspeiler elevenes faktiske kompetanse»; «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere»). De helte mot å være enige i påstanden «Jeg opplever at kravene for å få de høyeste karakterene (5 og 6) har blitt lavere enn tidligere». Ellers var lærerne i gjennomsnitt uenige i at de bidro til karakterinflasjon.

Tabell 10

Påstander relatert til karakterinflasjon

Påstand	Alle		U.Tr		Vgs	
	M	SD	M	SD	M	SD
Jeg setter høyere standpunktkarakterer nå enn for noen år siden uten at elevene har blitt flinkere/viser høyere faglig kompetanse	2.80	1.57	2.59	1.50	3.03	1.60
Jeg setter høyere standpunktkarakterer nå enn for noen år siden fordi elevene har blitt flinkere/viser høyere faglig kompetanse	2.50	1.38	2.45	1.32	2.56	1.43
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra elever	2.01	1.25	1.87	1.16	2.16	1.33
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra foresatte	1.63	1.00	1.71	1.05	1.54	0.93
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra skoleledelse	1.68	1.11	1.47	0.90	1.91	1.26
Jeg setter høyere standpunktkarakterer nå enn for noen år siden for å sikre bedre samsvar med eksamensresultater	1.64	1.00	1.55	0.92	1.74	1.07
Standpunktkarakterene jeg setter gjenspeiler elevenes faktiske kompetanse	5.39	0.89	5.46	0.80	5.30	0.97
Jeg opplever at kravene for å få de høyeste karakterene (5 og 6) har blitt lavere enn tidligere	3.81	1.63	3.69	1.60	3.94	1.66
Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere	4.02	1.66	3.75	1.62	4.32	1.65
Forventninger fra samfunnet om generøs vurderingspraksis påvirker min vurderingspraksis	2.81	1.59	2.60	1.51	3.04	1.64
Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer	3.44	1.58	3.25	1.57	3.65	1.57

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

Tabell 11 presenterer de siste påstandene fra spørreundersøkelsen. Lærerne var enige i begge påstander, altså at de diskuterte standpunktkarakterer med kollegaer før fastsettelse og at de opplevde å ha tilstrekkelig kompetanse for å sette standpunktkarakterer. I tillegg svarte 82 % av deltakerne (ikke presentert i tabell) at de kunne ha tenkt seg å

delta i kompetanseutvikling om standpunktvurdering. Denne andelen var den samme for lærere på ungdomsskolen som lærere i videregående opplæring.

Tabell 11

Påstander relatert til samarbeid og kompetanseutvikling

Påstand	Alle M	SD	U.Tr M	SD	Vgs M	SD
Jeg diskuterer standpunktkarakterer med kolleger før de fastsettes	4.68	1.28	4.75	1.28	4.61	1.28
Jeg har tilstrekkelig kompetanse til å sette rettfærdige og faglig godt begrunnede standpunktkarakterer	5.53	0.73	5.54	0.72	5.52	0.74

Merknad. U.tr = ungdomsskolen; Vgs = videregående skole; M = gjennomsnitt; SD = standardavvik

Fra svarene på påstandene tegner det seg et bilde av lærere som opplever at elever, lærere og ledere stoler på deres standpunktvurderinger og at de ikke gir etter for press eller setter standpunktkarakterer som ikke gjenspeiler det faglige nivået til elever. Samtidig er det viktig å minne om at tabellene viser gjennomsnitt. Påstanden «Jeg setter høyere standpunktkarakterer nå enn for noen år siden uten at elevene har blitt flinkere/viser høyere faglig kompetanse» er belysende: 76 % av lærerne var uenige i dette, men 24 % var enige. Så mens gjennomsnittslæreren fremstår som trygg og lite orientert mot å bidra til karakterinflasjon, er ikke dette tilfelle for alle lærere. I appendix A presenterer vi hvor mange som har valgt de ulike enighetsalternativene for samtlige items.

Regresjonsanalyser

Vi gjennomførte en rekke regresjonsanalyser for å estimere hvordan fag/fagklynge, kjønn og erfaring hang sammen med fire avhengige variabler. Analysene resulterte i modeller (Mi), en for hver avhengig variabel: M1: «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig»; M2: «Min skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet»; M3: «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere»; M4: «Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer». I analyser med fagklynger ble all data brukt, men i analyser med fag utelot vi fag med færre enn 20 observasjoner (se tabell 3a).

I tabell 12 presenteres estimat fra regresjonsanalyser basert på ungdomstrinns-data og med fagklynge som uavhengig variabel. For lesbarhetens skyld presenterer tabellen kun estimat og p-verdi. Fullstendige tabeller gjengis i appendix B. I tabellen fremkommer at fagklynge, kjønn og erfaring alle var signifikante prediktorer i modell 1 (M1); for eksempel var menn mer enige enn kvinner om forventninger fra elever om å sette en høyere standpunktkarakter enn det de mener er faglig riktig. Videre hadde stigende erfaring negativ effekt, hvilket vil si at jo mer erfaring, desto mindre enig i påstanden. Tabellen viser også at alle fagklynger i sammenligning med referanse kategorien (og etter at kjønn og erfaring holdes konstant) var signifikant negative prediktorer. For modellene M2 – M3 var kun erfaring en signifikant prediktor. Det bør noteres at effektstørrelsen var lav for alle fire modeller, hvilket indikerer at det er en rekke andre faktorer som henger sammen med grad av enighet i de fire påstandene. For M1 gjennomførte vi en post hoc-analyse for å undersøke mellom hvilke fagklynger det var signifikante forskjeller. Analysen

indikerte signifikante forskjeller mellom praktisk-estetiske fag og realfag ($t = 3.3$, $p = .006$), med modellbaserte gjennomsnitt for praktisk-estetiske fag = 4.21 og realfag = 3.63. Forskjellen var også signifikant mellom praktisk-estetiske fag og språkfag ($M_{\text{språkfag}} = 3.61$; $t = 3.9$, $p = .001$). Altså: lærere som representerte praktisk-estetiske fag var signifikant mer enige i påstanden i M1 enn lærere fra realfag og språkfag.

Tabell 12

Parameter	M ₁		M ₂		M ₃		M ₄	
	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>
Intercept	4.51	.000	3.14	.000	4.2	.000	3.52	.000
Realfag	-0.58	.001	-0.08	.671	0.36	.083	-0.03	.898
Samfunnsfag	-0.47	.032	-0.17	.443	0.03	.909	-0.17	.485
Språkfag	-0.6	.000	-0.22	.164	0.01	.973	-0.16	.374
Mann	0.43	.001	-0.08	.523	0.15	.299	0.24	.095
Erfaring	-0.03	.000	0.02	.000	-0.03	.000	-0.01	.028
	M ₁		M ₂		M ₃		M ₄	
	<i>R</i> ²		<i>R</i> ²		<i>R</i> ²		<i>R</i> ²	
Effektstørrelse	.08		.02		.05		.01	

Merknad. Avhengige variabler for M₁–M₄: M₁: «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig»; M₂: «Min skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet»; M₃: «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere»; M₄: «Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer» Fagklynge referensnivå: praktisk-estetiske fag.

Regresjonsanalysen med fag istedenfor fagklynger viste at erfaring var en signifikant negativ prediktor for M₁–M₃ og at fag var en signifikant prediktor for M₁. Ellers var ikke noen av prediktorene signifikante. For M₁ indikerte post hoc-analyser signifikante forskjeller mellom Kroppsøving og Matematikk ($M_{\text{kroppsøving}} = 4.49$; $M_{\text{matematikk}} = 3.19$; $t = 3.7$, $p = .004$), Mat og helse ($M_{\text{mat og helse}} = 4.17$) og matematikk ($t = 3.0$, $p = .03$) og matematikk og naturfag ($M_{\text{naturfag}} = 4.27$; $t = 3.0$, $p = .036$). Disse analysene pekte altså på den samme tendensen som fagklyngeanalysen, men utvalget var mindre og alle fag var ikke representert.

Tabell 13 presenterer estimat fra regresjonsanalyser med vgs-data, som kan leses på samme måte som tabell 12. Erfaring var kun signifikant for M₁. Alle modeller inneholdt én signifikant fagklyngeprediktor. Post hoc-analyser for M₁ viste at gjennomsnittet for yrkesfag ($M_{\text{yrkesfag}} = 3.27$) var signifikant mindre enn gjennomsnittet for realfag ($M_{\text{realfag}} = 4.14$; $t = 4.7$, $p < .0001$), samfunnsfag ($M_{\text{samfunnsfag}} = 4.0$; $t = 3.9$, $p = .001$) og språkfag ($M_{\text{språkfag}} = 4.12$; $t = 5.0$, $p < .0001$). Med andre ord var yrkesfagslærere mindre enige i utsagnet «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig». For M₂ viste post hoc-analyser ikke signifikante forskjeller mellom fagklynger. For M₃ var det signifikante forskjeller mellom yrkesfag ($M_{\text{yrkesfag}} = 3.75$) og realfag ($M_{\text{realfag}} = 4.56$; $t = 3.6$, $p = .004$) og mellom yrkesfag og språkfag ($M_{\text{språkfag}} = 4.71$; $t = 4.7$, $p < .0001$). Yrkesfagslærere var altså mindre enige i påstanden «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere» enn realfags- og språkfagslærere. For M₄ var det signifikante forskjeller mellom yrkesfag ($M_{\text{yrkesfag}} = 3.24$) og språkfag ($M_{\text{språkfag}} = 3.87$; $t = 3.2$, $p = .014$). Yrkesfagslærere var dermed i gjennomsnitt mindre enige enn språkfagslærere i følgende påstand: «Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer».

Tabell 13

Parameter	M ₁		M ₂		M ₃		M ₄	
	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>	<i>b</i>	<i>p</i>
Intercept	4.04	.000	3.1	.000	3.88	.000	3.88	.000
Realfag	0.32	.190	-0.26	.322	0.54	.070	-0.43	.129
Samfunnsfag	0.17	.477	-0.08	.758	0.25	.392	-0.27	.339
Språkfag	0.3	.190	-0.49	.046	0.68	.015	-0.15	.571
Yrkesfag	-0.55	.024	-0.05	.863	-0.28	.359	-0.78	.007
Mann	-0.05	.684	-0.04	.753	0.12	.451	0.14	.347
Erfaring	-0.01	.045	0.01	.163	0.01	.446	0	.562
	M ₁		M ₂		M ₃		M ₄	
	<i>R</i> ²		<i>R</i> ²		<i>R</i> ²		<i>R</i> ²	
Effektstørrelse	.05		.01		.04		.02	

Merknad. Avhengige variabler for M₁–M₄: M₁: «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig»; M₂: «Min skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet»; M₃: «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere»; M₄: «Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer» Fagklynge referansenivå: praktisk-estetiske fag.

Regresjonsanalysen med fag (HO, matematikk, norsk) istedenfor fagklynger viste at fag var en signifikant prediktor for M₁ og M₃. Ellers var ikke noen av prediktorene signifikante. For M₁ indikerte post hoc-analyser signifikante forskjeller mellom HO ($M_{HO} = 3.23$) og matematikk ($M_{matematikk} = 4.18$; $t = 3.0$, $p = .008$) og norsk ($M_{norsk} = 4.19$; $t = 3.6$, $p = .001$). Tendensene fra fagklyngeanalysen var altså den samme, med yrkesfagslærere (HO) mindre enige i påstanden «Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig». For M₃ var forskjellen mellom HO og norsk signifikant ($M_{HO} = 4.12$, $M_{norsk} = 5.05$; $t = 3.0$, $p = .001$), med HO-lærere mindre enige i påstanden «Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere».

Sammenfattende kan det sies at regresjonsanalysene indikerte visse forskjeller mellom fagklynger og enkelte fag. Erfaringene med standpunktvurdering er sannsynligvis altså ikke lik på tvers av fag. Samtidig viste analysene at fag/fagklynge, kjønn og erfaring ikke forklarte mye av variansen, hvilket tyder på at det er andre faktorer som spiller stor rolle.

SAMMENDRAG

Formålet med denne kvantitative delundersøkelsen var å kartlegge hvordan lærere opplever press, tillit og systemiske rammer rundt standpunktvurdering. Undersøkelsen gir et nasjonalt bilde av læreres erfaringer med forventninger fra elever, foresatte og skoleledelse, samt opplevelser av endringer i vurderingspraksis over tid.

Funnene viser et sammensatt bilde av høy profesjonell tillit kombinert med økende uformelt press og varierende grad av systematisk støtte. Et stort flertall av lærerne rapporterer høy grad av tillit til deres standpunktvurderinger fra elever, foresatte og skoleledelse, og et lavt nivå av formelle klager. Samtidig opplever mange lærere tydelige forventninger, særlig fra elever, om høyere karakterer enn det de selv vurderer som faglig begrunnet. Press fra foresatte og skoleledelse vurderes som mindre utbredt, men ikke fraværende.

Et særlig interessant funn i materialet er et spenn i lærernes svar om endringer i vurderingspraksis. På den ene siden rapporterer flertallet av lærerne at de i liten grad opplever seg selv som mer generøse i vurderingen enn tidligere, og de uttrykker høy tillit til eget profesjonelt skjønn. Samtidig oppgir en betydelig andel at kravene for å bestå (karakter 2) og tersklene for de høyere karakterene har blitt lavere over tid. Disse oppfatningene varierer systematisk mellom fag og fagklynger.

Mønsteret kan tolkes som uttrykk for endringer i kollektive vurderingsnormer snarere enn individuell ettergivenhet. Funnene peker mot at vurderingsstandarder i noen fagmiljøer gradvis justeres gjennom lokale praksiser og profesjonskulturer, uten at dette nødvendigvis oppleves som et resultat av direkte press. Dette indikerer en utvikling der karakterstandarder forskyves innenfor profesjonsfellesskapet, og peker på betydningen av styrkede kollektive kalibreringsarenaer dersom målet er større grad av likeverdighet i vurderingspraksis på tvers av fag og skoler.

Lærerne beskriver skoleledelsen som støttende i klage- og konfliktsituasjoner, men rapporterer i mindre grad at ledelsen arbeider systematisk og forebyggende med kvalitetssikring, kalibrering og utvikling av vurderingspraksis. Mange lærere oppgir at de tilpasser egen praksis under press, blant annet gjennom økt dokumentasjon, grundigere begrunnelser og tettere samarbeid med kollegaer i krevende saker.

Samlet sett viser funnene et tydelig spenningsfelt mellom høy grad av profesjonell autonomi og behovet for sterkere kollektive strukturer. Analysene viser at fag og fagklynge forklarer mer av variasjonen i opplevd press og vurderingsutfordringer enn kjønn og erfaring. Dette understreker betydningen av faglige vurderingskulturer framfor individuelle kjennetegn ved lærerne.

Resultatene peker mot et behov for tydeligere rammer for kollektiv kvalitetssikring, sterkere profesjonsfellesskap og mer systematisk støtte til lærere i møte med press og krav til dokumentasjon i standpunktvurderingen.

Undersøkelsens begrensninger

Resultatene må tolkes i lys av flere metodiske begrensninger.

For det første bygger undersøkelsen på selvrapporterte data. Standpunktvurdering er et normregulert og potensielt sensitivt praksisfelt, og det kan ikke utelukkes at enkelte deltakere har hatt insentiver til å fremstille egen praksis mer normativ enn den faktisk er.

For det andre har vi begrenset kunnskap om utvalgets representativitet. Selv om deltakerne representerer ulike fag, regioner og skolestørrelser, kan vi ikke sikkert si at utvalget speiler hele lærerpopulasjonen på ungdomstrinn og videregående opplæring.

For det tredje gir spørreskjemaets lukkede svarformat begrenset innsikt i hvordan lærerne konkret håndterer press og dilemmaer i praksis. Undersøkelsen fanger i liten grad dynamikken i profesjonsfellesskap, lokale forhandlinger og uformelle normer i fagmiljøene. Til sammen innebærer dette at funnene gir et bredt, men ikke fullstendig bilde av praksis og utfordringer knyttet til standpunktvurdering.

Avslutning

Denne kvantitative delundersøkelsen gir et bredt bilde av hvordan lærere opplever rammer, press og støtte knyttet til standpunktvurdering i ungdomsskole og videregående opplæring. Funnene viser at lærere i stor grad opplever tillit til sitt profesjonelle skjønn, samtidig som de står i et komplekst landskap preget av forventninger fra særlig elever, men også i noen grad fra foresatte og ledelse.

Undersøkelsen synliggjør et tydelig spenn mellom høy grad av profesjonell autonomi og behovet for sterkere kollektive strukturer rundt vurderingsarbeidet. Variasjonene i erfaringer ser i større grad ut til å være knyttet til fag og fagkulturer enn til individuelle kjennetegn ved lærerne. Dette understreker behovet for videre arbeid med å utvikle sterkere profesjonsfellesskap og mer systematisk arbeid med kalibrering, felles tolkningsrammer og kollektive kvalitetssikringsprosesser i skolene.

Funnene kan gi et viktig kunnskapsgrunnlag for skoleeiere, skoleledelse og nasjonale myndigheter i arbeidet med å styrke kvalitet, legitimitet og likeverdighet i standpunktvurdering. Særlig peker resultatene mot behovet for tydeligere forventninger til kollektivt vurderingsarbeid og bedre rammer for å håndtere press og dokumentasjonskrav i sluttvurderingen

LITTERATUR

- Lenth, R. V., & Piaskowski, J. (2017). *emmeans: Estimated Marginal Means, aka Least-Squares Means (s. 2.0.0)* [Datasett]. <https://doi.org/10.32614/CRAN.package.emmeans>
- R Core Team. (2025). R: A language and environment for statistical computing [Programvare]. *R Foundation for Statistical Computing*. <https://www.R-project.org/>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>

Appendix A Spørreundersøkelse 2

Appendix A viser spørreskjemaet for spørreundersøkelse 2 med påstander og svarskala.

item	1	2	3	4	5	6	NA
Dersom det oppstår uenighet mellom lærer og elev/foresatte om en standpunktkarakter, gir leder etter for press uten å støtte læreren	589	309	163	88	38	13	16
Dersom det oppstår uenighet mellom lærer og elev/foresatte om en standpunktkarakter, støtter leder lærerens vurdering	10	33	65	229	473	389	17
Elever har tillit til min standpunktvurdering	7	4	6	42	569	587	1
Foresatte har tillit til min standpunktvurdering	7	2	11	54	586	551	5
Forventninger fra samfunnet om generøs vurderingspraksis påvirker min vurderingspraksis	343	215	144	254	115	68	77
Jeg diskuterer standpunktkarakterer med kolleger før de fastsettes	51	45	55	301	398	366	0
Jeg får klager på min standpunktvurdering	787	307	35	48	20	17	2
Jeg får oftere klager på min standpunktvurdering enn mine kolleger	1053	105	30	17	4	3	4
Jeg har tilstrekkelig kompetanse til å sette rettferdige og faglig godt begrunnede standpunktkarakterer	5	8	9	48	392	754	0
Jeg opplever at det forventes at standpunktkarakterene jeg setter, skal samsvare med eksamenskarakterene i faget	236	187	195	320	205	73	0
Jeg opplever at kravene for å bestå (karakter 2) har blitt lavere enn tidligere	112	149	121	261	183	300	90
Jeg opplever at kravene for å få de høyeste karakterene (5 og 6) har blitt lavere enn tidligere	129	187	117	298	200	227	58
Jeg opplever at vurderingspraksisen i mitt fagmiljø har endret seg i retning av mer generøse karakterer	184	182	145	329	174	126	76
Jeg opplever forventninger fra elever om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	111	180	109	422	266	128	0
Jeg opplever forventninger fra foresatte om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	323	267	160	289	127	50	0
Jeg opplever forventninger fra skoleeier om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	769	171	108	108	32	28	0
Jeg opplever forventninger fra skoleledelse om å sette en høyere standpunktkarakter enn jeg mener er faglig riktig	693	224	109	130	34	26	0
Jeg setter høyere standpunktkarakterer nå enn for noen år siden for å sikre bedre samsvar med eksamensresultater	669	212	124	49	13	6	143
Jeg setter høyere standpunktkarakterer nå enn for noen år siden fordi elevene har blitt flinkere/viser høyere faglig kompetanse	325	268	225	148	73	31	146
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra elever	543	242	149	117	36	15	114
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra foresatte	687	199	118	57	9	7	139
Jeg setter høyere standpunktkarakterer nå enn for noen år siden på grunn av forventninger eller press fra skoleledelse	695	213	81	81	19	12	115
Jeg setter høyere standpunktkarakterer nå enn for noen år siden uten at elevene har blitt flinkere/viser høyere faglig kompetanse	312	224	182	214	91	77	116
Jeg stoler på mine kollegers standpunktvurdering	7	4	29	89	478	606	3
Mine skole har en tydelig felles praksis for hvordan man håndterer press fra elever og foresatte i vurderingsarbeidet	201	236	258	231	222	68	0
Skoleledelsen følger med på mine vurderinger for å sikre en faglig forsvarlig praksis	244	262	224	271	158	53	4
Skoleledelsen gir meg veiledning dersom jeg står i en vanskelig vurderingssituasjon	39	82	134	320	346	290	5
Skoleledelsen har tillit til min standpunktvurdering	10	0	8	20	226	950	2
Skoleledelsen har tydelige retningslinjer for vurderingsarbeidet ved min skole	78	106	177	311	344	196	4
Skoleledelsen legger til rette for faglige diskusjoner om vurderingspraksis	24	71	99	314	338	365	5
Skoleledelsen støtter meg i arbeidet med standpunktvurdering	40	68	82	242	341	440	3
Standpunktkarakterene jeg setter gjenspeiler elevenes faktiske kompetanse	9	15	24	70	406	654	38
Ved eventuelt press i karaktersetting...diskuterer jeg saken med kolleger	3	5	5	65	208	807	123
Ved eventuelt press i karaktersetting...diskuterer jeg saken med skoleledelse	92	94	111	247	235	284	153

Ved eventuelt press i karaktersetting...dokumenterer jeg mine vurderinger grundigere enn vanlig	32	51	59	223	371	347	133
Ved eventuelt press i karaktersetting...endrer jeg vurderinger for å unngå ubehag	548	272	135	101	13	14	133
Ved eventuelt press i karaktersetting...føler jeg meg trygg og står ved mine vurderinger	2	7	26	61	453	525	142
Ved eventuelt press i karaktersetting...har jeg opplevd at presset har påvirket min vurderingspraksis over tid	273	246	150	247	81	34	185
Ved eventuelt press i karaktersetting...opplever jeg at skolens ledelse gir tydelig støtte i slike situasjoner	39	36	89	172	286	390	204
Ved eventuelt press i karaktersetting...opplever jeg usikkerhet i vurderingsarbeidet	214	331	142	274	80	40	135
Ved eventuelt press i karaktersetting...søker jeg støtte i skolens retningslinjer for vurdering	113	80	92	243	273	234	181

Merknad. 1 = «helt uenig», 2 = «ganske uenig», 3 = «litt uenig», 4 = «litt enig», 5 = «ganske enig», 6 = «helt enig»); NA = missing.

Appendix B Regresjonsanalyser spørreundersøkelse 2

Appendix B presenterer regresjonsanalysene som ligger til grunn for enkelte av de kvantitative funnene, med modellspesifikasjon og resultattabeller.

Regresjonsanalyse – utrinn fagklynge

term	b	S.E.	t	p	modell
Intercept	4.51	0.17	26.92	.000	M1
Realfag	-0.58	0.18	-3.30	.001	M1
Samfunnsfag	-0.47	0.22	-2.15	.032	M1
Språkfag	-0.60	0.15	-3.92	.000	M1
Mann	0.43	0.13	3.39	.001	M1
Erfaring	-0.03	0.01	-5.52	.000	M1
Intercept	3.14	0.17	18.08	.000	M2
Realfag	-0.08	0.18	-0.42	.671	M2
Samfunnsfag	-0.17	0.22	-0.77	.443	M2
Språkfag	-0.22	0.16	-1.39	.164	M2
Mann	-0.08	0.13	-0.64	.523	M2
Erfaring	0.02	0.01	4.19	.000	M2
Intercept	4.20	0.20	21.17	.000	M3
Realfag	0.36	0.21	1.74	.083	M3
Samfunnsfag	0.03	0.25	0.11	.909	M3
Språkfag	0.01	0.18	0.03	.973	M3
Mann	0.15	0.15	1.04	.299	M3
Erfaring	-0.03	0.01	-5.18	.000	M3
Intercept	3.52	0.20	17.79	.000	M4
Realfag	-0.03	0.20	-0.13	.898	M4
Samfunnsfag	-0.17	0.25	-0.70	.485	M4
Språkfag	-0.16	0.18	-0.89	.374	M4
Mann	0.24	0.15	1.67	.095	M4
Erfaring	-0.01	0.01	-2.20	.028	M4

Regresjonsanalyse – u.trinn fag

term	b	S.E.	t	p	modell
Intercept	3.81	0.21	18.31	.000	M1
Kroppsøving	0.94	0.35	2.71	.007	M1
Mat og helse	0.63	0.29	2.18	.030	M1
Matematikk	-0.36	0.24	-1.52	.130	M1
Naturfag	0.72	0.35	2.08	.039	M1
Samfunnsfag	0.04	0.29	0.15	.882	M1
Mann	0.34	0.21	1.65	.100	M1

Erfaring	-0.02	0.01	-2.88	.004	M1
Intercept	2.65	0.22	12.27	.000	M2
Kroppsøving	0.28	0.36	0.78	.437	M2
Mat og helse	0.32	0.30	1.06	.292	M2
Matematikk	0.25	0.24	1.01	.315	M2
Naturfag	0.20	0.36	0.55	.582	M2
Samfunnsfag	-0.12	0.31	-0.38	.701	M2
Mann	0.00	0.21	0.00	1.000	M2
Erfaring	0.03	0.01	3.36	.001	M2
Intercept	4.19	0.25	16.70	.000	M3
Kroppsøving	0.15	0.41	0.38	.706	M3
Mat og helse	-0.57	0.34	-1.66	.099	M3
Matematikk	0.19	0.28	0.67	.503	M3
Naturfag	0.32	0.41	0.78	.434	M3
Samfunnsfag	0.05	0.34	0.14	.885	M3
Mann	0.10	0.24	0.42	.678	M3
Erfaring	-0.03	0.01	-2.46	.014	M3
Intercept	3.35	0.24	13.73	.000	M4
Kroppsøving	0.37	0.40	0.94	.349	M4
Mat og helse	-0.41	0.33	-1.22	.224	M4
Matematikk	-0.23	0.27	-0.87	.385	M4
Naturfag	-0.04	0.40	-0.09	.927	M4
Samfunnsfag	-0.08	0.34	-0.23	.817	M4
Mann	0.23	0.24	0.97	.334	M4
Erfaring	-0.01	0.01	-0.57	.569	M4

Regresjonsanalyse – vgs fagklynge

term	b	S.E.	t	p	modell
Intercept	4.04	0.24	16.68	.000	M1
Realfag	0.32	0.24	1.31	.190	M1
Samfunnsfag	0.17	0.24	0.71	.477	M1
Språkfag	0.30	0.23	1.31	.190	M1
Yrkesfag	-0.55	0.24	-2.26	.024	M1
Mann	-0.05	0.13	-0.41	.684	M1
Erfaring	-0.01	0.01	-2.01	.045	M1
Intercept	3.10	0.26	11.77	.000	M2
Realfag	-0.26	0.26	-0.99	.322	M2
Samfunnsfag	-0.08	0.26	-0.31	.758	M2
Språkfag	-0.49	0.25	-2.00	.046	M2
Yrkesfag	-0.05	0.27	-0.17	.863	M2
Mann	-0.04	0.14	-0.31	.753	M2
Erfaring	0.01	0.01	1.40	.163	M2

Intercept	3.88	0.30	12.89	.000	M3
Realfag	0.54	0.30	1.82	.070	M3
Samfunnsfag	0.25	0.29	0.86	.392	M3
Språkfag	0.68	0.28	2.45	.015	M3
Yrkesfag	-0.28	0.30	-0.92	.359	M3
Mann	0.12	0.15	0.75	.451	M3
Erfaring	0.01	0.01	0.76	.446	M3
Intercept	3.88	0.29	13.49	.000	M4
Realfag	-0.43	0.28	-1.52	.129	M4
Samfunnsfag	-0.27	0.28	-0.96	.339	M4
Språkfag	-0.15	0.26	-0.57	.571	M4
Yrkesfag	-0.78	0.29	-2.71	.007	M4
Mann	0.14	0.15	0.94	.347	M4
Erfaring	0.00	0.01	0.58	.562	M4

Regresjonsanalyse – vgs fag

term	b	S.E.	t	p	modell
Intercept	3.51	0.25	14.01	.000	M1
Matematikk	0.95	0.32	3.03	.003	M1
Norsk	0.96	0.27	3.61	.000	M1
Mann	0.05	0.29	0.17	.869	M1
Erfaring	-0.02	0.01	-1.84	.067	M1
Intercept	2.61	0.34	7.62	.000	M2
Matematikk	-0.25	0.30	-0.86	.393	M2
Norsk	0.21	0.35	0.60	.548	M2
Mann	0.18	0.32	0.57	.571	M2
Erfaring	0.01	0.01	0.49	.622	M2
Intercept	4.55	0.37	12.33	.000	M3
Matematikk	0.24	0.31	0.77	.442	M3
Norsk	-0.70	0.36	-1.91	.058	M3
Mann	0.51	0.33	1.52	.130	M3
Erfaring	0.00	0.01	0.05	.963	M3
Intercept	3.52	0.36	9.88	.000	M4
Matematikk	0.33	0.30	1.10	.275	M4
Norsk	-0.15	0.36	-0.42	.676	M4
Mann	0.42	0.32	1.29	.199	M4
Erfaring	0.00	0.01	0.14	.891	M4

