

Vurdering i skolen.

Utvikling av kompetanse og fellesskap

Sluttrapport fra prosjektet Forskning på individuell vurdering i skolen (FIVIS)

Lise Vikan Sandvik og Trond Buland (red.)

Trondheim 10. desember 2014

NTNU Program for lærerutdanning
i samarbeid med SINTEF Teknologi og samfunn



NTNU Program for lærerutdanning

Postadresse: NTNU Program for lærerutdanning, 7491 Trondheim

Besøksadresse: Jonsvannsveien 82, Trondheim

Web: <http://www.ntnu.no/plu>

ISBN: 978-82-7923-069-4

Forord

Med dette foreligger sluttrapporten fra prosjektet «Forskning på individuell vurdering i skolen» (heretter FIVIS), utført ved NTNU Program for lærerutdanning i samarbeid med SINTEF Teknologi og samfunn. Prosjektet har i sin helhet vært finansiert av Utdanningsdirektoratet.

I arbeidet med denne rapporten har følgende forskere vært involvert:

- **Trond Buland**, forsker ved NTNU Skole og utdanningsforskning, NTNU Program for lærerutdanning
- **Gunnar Engvik**, førstelektor ved NTNU Program for lærerutdanning
- **Henning Fjørtoft**, førsteamanuensis i norsk fagdidaktikk, NTNU Program for lærerutdanning
- **Inger Langseth**, førstelektor, NTNU Program for lærerutdanning
- **Lise Vikan Sandvik**, førsteamanuensis i fremmedspråk fagdidaktikk, NTNU Program for lærerutdanning
- **Siri Mordal**, forsker, SINTEF Teknologi og samfunn

Buland har vært prosjektleder. I tillegg har forsker Thomas Dahl vært faglig rådgiver for prosjektet. Buland og Sandvik har fungert som redaktører for denne rapporten, der de involverte forskerne har skrevet ulike kapitler. Forfatter(e) er angitt i hvert kapittel. Thomas Dahl har lest og kommentert rapporten. Siri Mordal fra SINTEF har også lest og kommentert deler av arbeidet.

Data til denne rapporten er innhentet gjennom hele prosjektperioden. I siste fase har vi supplert våre tidligere data med intervjuer i til sammen fem skoler. Vi benytter derfor anledningen til å takke alle de skolene som gjennom hele prosjektet har åpnet sine klasserom for oss og brukte av sin tid til å stille opp i intervjuer, samt alle lærere som deltok i vår breddeundersøkelse. Vi har uten unntak blitt svært godt mottatt, og vi håper at resultatet også vil være av interesse og nytte for dere.

Vi vil dessuten takke våre eksterne kvalitetssikrere, Ole Johan Eikeland og Christian Lundahl for verdifulle kommentarer til vårt arbeid. Dessuten vil vi takke våre oppdragsgivere i Utdanningsdirektoratet for god dialog og nyttige innspill gjennom hele prosjektperioden.

Trondheim, 10.12 2014

Innhold

Forord	3
Sammendrag	9
Summary	15
1. Om FIVIS-prosjektets formål, problemstillinger, data og metode.....	21
1.1 Problemstillinger	21
1.2 Om rapporten	24
1.3 Hovedfunn fra delrapport en og 2	26
1.3.1 Delrapport 1: Vurdering i skolen. Intensjoner og forståelser.....	26
1.3.2 Delrapport 2: Vurdering i skolen. Operasjonaliseringer og praksiser.....	28
2. Metode og datakilder.....	31
2.1 Kvalitativ metode	31
2.1.1 Observasjoner.....	32
2.1.2 Intervjuer	33
2.1.3 Dokumenter.....	34
2.1.4 Analyser av det kvalitative datamaterialet	35
2.1.5 Er casene representative?	36
2.1.6 Er resultatene valide?	37
2.2 Kvantitativ breddesurvey	38
2.2.1 Datagrunnlag og utvalg	38
2.2.2 Breddesurvey.....	38
2.2.3 Utvikling av indikatorer	39
2.2.4 Oppsummering	42
3. Hva innebærer validitet i vurdering?.....	43
3.1 Individuell vurdering i skolen	44
3.2 Validitet i klasserommet.....	48
3.3 Validitet i lokalt læreplanarbeid	53
3.4 Validitet i overgang mellom trinn	55
3.5 Reliabilitet i klasserommet og på nasjonalt nivå.....	56
4. Hvordan er vurderingspraksisene og -kulturene i norske skoler?	59
4.1 Praksisfortelling en – Julie som lærende lærer.....	59
4.1.1 Kommentar til praksisfortellingen	60
4.2 Praksisfortelling to – mange praksiser og forståelser under samme tak	61
4.2.1 Kommentar til praksisfortellingen	63
4.3 Praksisfortelling tre – fra mål og kriterier til praksisfellesskap	63
4.3.1 Kommentar til praksisfortellingen	65
4.4. Fire forskjellige vurderingskulturer	66
4.5 Helhet og sammenheng i utviklingen av vurderingskulturer	70
4.5.1 Noen forbehold.....	72
5. Det avhenger mye av læreren – elevers opplevelse av vurdering og læring.....	73
5.1 Overordnet inntrykk – god vurdering er viktig	74
5.2 Ønske om elevaktiv undervisning	74
5.3 Læring i forskjellige kontekster – overganger mellom skoleslag	76
5.4 Hvem er den gode læreren – relasjonskompetanse og fagkunnskap.....	78
5.5 Kriterier og tilbakemeldinger på eget arbeid.....	80
5.6 Vurdering av eller for læring?	82
5.7 Tilpasset undervisning.....	84
5.8 Sammenfattet – underveisvurdering og elevers opplevelse av læring.....	85
6. Ledelse av læreres læring	87

6.1	Praksisfortellinger om barneskole som har deltatt i Vurdering for læring.....	87
6.2	Praksisfortelling fra 1–10 skole som ikke har vært med i Vurdering for læring.....	89
6.3	Analyse av praksisfortellinger om ledelse av skoler.....	90
7.	Kvantitative analyser – Hva skaper god vurderingspraksis?	93
7.1	Analysen og diskusjon	93
7.2	Resultater.....	94
7.2.1	Bakgrunnsvariabler som forklaringsfaktorer	94
7.2.2	Type lærerutdanning og vurderingspraksis	95
7.2.3	Skoleslag og vurderingspraksis	96
7.2.4	Fordypning i fag og vurderingspraksis.....	96
7.3.	Hvilke faktorer forklarer god vurderingspraksis?	97
7.3.1	Analysen av samsvar og kommentarer	97
7.3.2	Analysen av signifikante faktorer	99
7.4	Oppsummering av resultater	104
7.5	Drøfting – God vurderingspraksis – et sammensatt bilde	105
7.6	God vurderingspraksis – kompetanse og synlige fellesskaper i skolen	107
8.	Hvordan utvikle vurderingskulturer?	111
8.1	Kulturhistorisk aktivitetsteori.....	112
8.1.1	Vurderingskultur som aktivitetssystem	112
8.2	Lærernes vurderingspraksiser er mangfoldige	116
8.3	Skolenes vurderingskulturer er varierte	118
8.4	Vurdering for læring som skolebasert utvikling og forankret i fag.....	119
8.5	Tydelige mål og gode tilbakemeldinger viktig for elevens læring.....	120
8.6	Utvikling av vurderingskulturer	123
8.7	Validitet som terskelbegrep.....	125
8.7.1	Validitetsskjeder i undervisning og læring	126
8.7.2	Validitet i utviklingen av praksisfellesskap i skolen.....	127
8.7.3	Validitet som terskelbegrep i utviklingen av vurderingskompetanse	128
9.	Prinsipper for utvikling av vurderingskulturer	133
10.	Litteratur.....	137
	Vedlegg 1: Intervjuguide.....	145
	Vedlegg 2: Observasjonsguide.....	147
	Vedlegg 3: Faktoranalyse med faktorladninger	150

Figurer

Figur 1. Aspekter ved formativ vurdering (Wiliam & Thompson, 2007).	47
Figur 2. Fire ulike vurderingskulturer (modifisert etter O'Donovan mfl., 2006).....	67
Figur 3. Aktivitetssystemet (Engstrøm, 1999)	113
Figur 4. Aktivitetssystemet i utvikling av vurderingskulturer.....	116

Tabeller

Tabell 1: Klasseromsobservasjoner fordelt på klasse og trinn	32
Tabell 2: Intervjuer	34
Tabell 3: Indikatorer og underliggende variabler	40
Tabell 4: Indikatorenes reliabilitet og signifikansnivå på test etter SEM-analyse ¹	42
Tabell 5: Gjennomsnittsscore på indikatoren "Vurderingspraksis" fordelt på "Type lærerutdanning"	95
Tabell 6: Gjennomsnittsscore på indikatoren "Vurderingspraksis" fordelt på "Fordypning i fag"	96
Tabell 7: Gjennomsnittsscore på indikatoren "Vurderingspraksis" fordelt på "Skoleslag"	96
Tabell 8: Korrelasjonstabell for alle indikatorer som er utviklet	98
Tabell 9: Lineær regresjonsanalyse av "Vurderingspraksis" etter signifikante forklaringsfaktorer og fag.....	100

Sammendrag

Dette er sluttrapporten fra prosjektet «Forskning på individuell vurdering i skolen» (FIVIS). Rapporten presenterer resultatene av analysene i prosjektets tredje fase, i tillegg til den sammenfattende analysen av resultater fra hele prosjektet FIVIS, som har løpt fra 2011 til 2014.

Hensikten med prosjektet har vært å gi et bilde av hvordan vurderingspraksis i norske skoler utvikler seg i et samspill mellom sentralt gitte føringer og teorier og lokale fortolkninger og forståelser. Det er i det komplekse samspillet mellom sentrale intensjoner og lokale forståelser og fortolkninger i skolene at vurderingspraksis blir konstruert. Vi har søkt å følge fenomenet vurdering i skolen fra sentrale intensjoner og forståelser via lokal operasjonalisering til lokal praksis.

FIVIS har hatt tre overordnede problemstillinger:

Hvordan kommer skoleeieres, skolelederes og læreres vurderingskompetanse til uttrykk gjennom fortolkninger, forståelse og praksiser?

Hvilke vurderingspraksiser finnes i norske klasserom?

Hvilken betydning har vurderingspraksisene for elevenes opplevelse av læringsprosesser og læringsutbytte?

Metodisk har vi belyst problemstillingene ved hjelp av data innsamlet med både kvantitative og kvalitative metoder.

Resultatene fra FIVIS viser et mangfold i vurderingspraksis, både på kollektivt og på individuelt plan i norske skoler. Det er variasjoner i praksis mellom skoler og skoleslag, men også internt i skoler. Resultatene viser også at det er stor interesse for vurdering i norske skoler i dag.

Lærerens faglige og fagdidaktiske kompetanse har betydning for hvordan læreren forstår og operasjonaliserer arbeidet med vurdering. Lærere med svak faglig og fagdidaktisk kompetanse synes å ha utfordringer med å forstå og operasjonalisere vurdering for læring på en tilfredsstillende måte.

Skolens organisering i team vs. fagteam har betydning for om det etableres tolkningsfellesskap for vurdering for læring i fag. Skoleledelsens og skoleeiers forståelser av vurdering for læring og operasjonaliseringer av arbeidet for å styrke skolens vurderingspraksis er en annen faktor som påvirker lærernes holdninger og praksiser til vurdering. 1–10-skoler har større utfordringer knyttet til etablering av tolkningsfellesskaper, altså rom for felles tolkning, diskusjon og utvikling av en kollektiv vurderingspraksis, enn andre skoler. Også videregående skoler, blant annet på grunn av større grad av individuelt

orienterte lærere enn andre skoleslag, har i mindre grad etablert tolkningsfellesskap omkring vurdering.

Både breddeundersøkelsen, intervjuene og observasjonene viser at det er store variasjoner i skolenes arbeid med vurdering og bevissthet rundt vurdering for læring. FIVIS viser at mange skoler har kommet godt i gang med å etablere tydelige mål for elevenes læring og å involvere elevene i eget vurderingsarbeid. Særlig viser det seg at de som har deltatt i satsingen «Vurdering for læring», i større grad arbeider systematisk med vurdering for læring på skolenivå enn de som ikke har deltatt. Likevel er det en klar tendens i vårt materiale at mange av skolene ikke har lyktes med å etablere en vurderingskultur, altså en kollektivt forankret forståelse av vurdering for læring og vurderingspraksis for hele skolen, blant hele personalet. Målet om å sette elevens læring i fokus i vurderingsarbeidet har ikke nådd alle lærere og alle skoler.

Langs flere dimensjoner er det signifikante forskjeller mellom skolene og lærerne når det gjelder arbeid med god vurderingspraksis. Lærere med fagutdanning på universitet/høgskole og praktisk-pedagogisk utdanning scorer signifikant bedre på indikatoren *individuell vurderingspraksis* enn lærere med allmennlærerutdanning. Lærere med fra ett til tre års utdanning i faget scorer bedre på indikatoren *individuell vurderingspraksis* enn lærere med bare et halvt års utdanning i faget. Det er også et klart skille mellom vurderingskulturer og -praksiser i grunnskolens barnetrinn og ungdomsskolen / videregående skole.

Vi finner også et stort mangfold av vurderingspraksiser i fagene engelsk, norsk, kroppsøving og matematikk. Disse varierende praksisene bygger på forskjellige forståelser av intensjonene med individuell vurdering i skolen og på de ulike fagenes formål i skolen, læreplanmål og egenart. Vi ser også at de forskjellige fagene har utfordringer med hensyn til validitet i vurderingen. Ved flere skoler ser vi eksempel på at man blander den kompetansen som skal prøves, og den representasjonsformen som kunnskapen skal prøves gjennom. Den kvantitative undersøkelsen viser at i alle fagene er lærernes vurderingskompetanse den indikatoren som har sterkest forklaringskraft på hva som bidrar til god vurderingspraksis. Samtidig ser vi at det på skolenivå er utfordringer med hensyn til fagkompetanse. Det kan synes som det er behov for å arbeide mer med kompetanseheving på fagnivå, slik at lærernes vurderingskompetanse i fagene utvikles videre.

FIVIS-studien viser at i norske myndigheters satsing på vurdering for læring synes allmenpedagogiske verktøy å ha blitt vektlagt i større grad enn disiplinfaglige og fagdidaktiske. Dette kan ha ført til at vurderingskompetansen til den enkelte lærer ikke har blitt tilstrekkelig forankret i skolefagenes undervisningstradisjoner. Verktøyene som har blitt vektlagt, har også vært rettet mot enkeltlærere og ikke mot kollektivt orienterte utviklingsprosesser.

Resultatene fra FIVIS viser at skoleledelsen spiller en avgjørende rolle for utviklingen av skolens vurderingskultur og vurderingspraksis. Måten skolelederen legger til rette for utviklingsarbeid i skolen på, har betydning for hvilke spenninger og motsetninger som oppstår i skolekulturen når endringer og utvikling skal foregå. Lederens rolle i arbeidet vil bidra til enten å hemme eller å fremme utviklingsarbeidet.

Utviklingen av god vurderingspraksis må skje både på individuelt nivå og på skolenivå. På individnivå er de viktigste dimensjonene for god vurderingspraksis bruk av forskjellige arbeidsmåter blant elevene, forskjellige vurderingsformer, utvikling av et fagspråk om vurdering samt lærerens kompetanse til å se sammenheng mellom kompetansemål i læreplaner, undervisningsformer og vurdering. For å få til den individuelle utviklingen må det legges til rette for faglig utviklingsarbeid på skolenivå, slik at det etableres tolkningsfellesskaper omkring vurdering som bidrar til lærernes faglige utvikling. I dette utviklingsarbeidet oppstår vurderingskulturer som fremmer elevenes læring. Denne formen for utviklingsarbeid, hvor lærerne i fellesskap arbeider med å styrke sin vurderingspraksis, åpner opp for meningsskaping og deltakelse i praksisfellesskap, den skaper interaksjon, dialog og refleksjoner over læring.

FIVIS-studien viser at elever på alle trinn har relativt høy bevissthet om vurdering, selv om fokuset åpenbart blir sterkere jo høyere årstrinn det er snakk om. Elever på alle trinn opplever vurdering som viktig, også for egen læring. Elever på alle årstrinn viser en betydelig bevissthet rundt mål. Mål står i ukeplaner, og lærere skriver dem opp på tavla og presenterer dem også muntlig i timene. Samtidig er det elever i ungdomsskole og videregående skole som har et klart fokus på den summative vurderingen. Disse elevene er i mindre grad opptatt av hvordan vurdering kan bidra til egen faglig utvikling.

Elever har fortalt om hvordan man har fått informasjon om og vært med på å diskutere vurderingskriterier, altså i betydningen hva som skiller et godt svar fra et dårlig, hva som må være med for å få en viss karakter. Disse elevene har en bevissthet om arbeidet med mål og kriterier, noe som skyldes at det er etablert praksiser i mange klasserom hvor læringsintensjonene blir kommunisert til elevene, både via planer og i klasseromsdialogene.

Gode tilbakemeldinger handler om å gi elevene tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonene deres. Dessuten handler gode tilbakemeldinger om at elevene får informasjon om hva som skal til for å utvikle seg videre i faget. Vurderingsforskriften understreker at denne underveisvurderingen skal «gis løpende og systematisk og kan være både munnleg og skriftleg». FIVIS-prosjektet viser at det er store variasjoner på enkeltskoler når det gjelder å etterkomme forskriftens krav. Elevene uttrykker

at noen lærere er konstruktive og gir grundige pekere på måter de kan forbedre seg på, andre er korte og temmelig intetsigende i sine kommentarer.

Elevgruppene gir et sammensatt bilde av hva det vil si å ha god vurderingskompetanse. «Den gode læreren» er en som er faglig flink, som kan faget sitt, og som også kan lære bort, altså har evne til å formidle faget på en måte som ulike elever kan ha utbytte av. I dette ligger det også å kunne gi gode, læringsfremmende tilbakemeldinger på elevenes prestasjoner. For å kunne gjøre dette framheves ofte betydningen av det relasjonelle, at læreren må se hver enkelt elev, skjønne den enkelte elev og kunne kommunisere med alle på deres egne premisser.

Et sentralt tema for mange av elevene i studien handler om reliabilitet i vurdering. Elevene forventer at samme lærer skal vurdere ulike elever likt. Funn i vår studie kan tyde på at det er sprikende praksiser på dette området.

Egenvurdering bidrar til refleksjoner over egen læring. Det synes som om elevene i FIVIS har svært varierte erfaringer knyttet til vurdering for læring. Graden av elevaktivitet i vurderingsarbeidet varierer, men er jevnt over relativt lav på alle trinn. Likevel kan det synes som barnetrinnet har kommet lenger i arbeidet enn de andre skoleslagene. Det er verdt å merke seg at elevene mener at god vurderingspraksis med tydelige mål og kriterier, der de selv er aktive i egen og andres læringsprosess, er viktig for deres læring, motivasjon og mestring.

FIVIS-prosjektet viser at lærere i liten grad benytter tilbakemeldinger fra elevene og resultat fra vurderinger i videre arbeid med egen undervisning. Vi ser med andre ord et klart potensial for å videreutvikle vurdering for læring som et ledd i utviklingen av tilpasset undervisning.

Lærere trekker slutninger om elevenes læring og tar avgjørelser om elevenes utvikling og om videre undervisningsforløp på grunnlag av disse slutningene. Resultatene fra FIVIS viser at en bevissthet om validitet og validitetskjeder er et godt begrep for å bedømme kvaliteten i arbeidet med vurdering, enten det er fortolkningen av læreplanmålene, det lokale arbeid med læreplaner, dagligdagse samtaler om læring mellom lærere og elever, bruk av standardiserte tester eller rutiner for karaktersetning i skolen. Denne bevisstheten omkring sammenhengen mellom læreplanmål, lokalt læreplanarbeid, arbeidsmåter i faget, tilbakemeldinger til elevene og vurdering i faget er varierende i de ulike skolene. Det kan synes som enkelte skoler operasjonaliserer validitet i vurdering som en enkelthendelse, der det anvendes generelle sjekklister eller kontrollverktøy. Validitet i vurdering bør sees som et gjennomgripende aspekt ved alle vurderingsprosesser, fra klasserom til nasjonalt nivå, som må sikres gjennom bevissthet om og rutiner for kontinuerlig og kritisk evaluering. En mer formell teoretisk forståelse av

validitet og reliabilitet i vurdering kan bidra til profesjonalisering av aktørene i dette arbeidet. Validitet representerer en terskel inn i en ny forståelse av kunnskapsfeltet, og dersom denne terskelen krysses, vil kunnskapen om vurdering kunne endres og integreres med andre former for kunnskap om undervisning og læring.

Forskningsprosjektets funn antyder et spenningsfelt mellom forvaltningen av skolepolitiske styringsmekanismer på den ene siden og det praktiske arbeidet i hvert enkelt klasserom på den andre. Mens styring av skolevirksomheten har oppmerksomheten rettet mot karakterstatistikker og mulige klagesaker, er spørsmålet om validitet i vurdering delegert ned til lærernivå. Dersom forvaltningsapparatet legger overdreven vekt på reliabilitet i karaktersetting, kan resultatet bli byråkratiserende dokumentasjonsprosesser som ikke bidrar til utviklingen av gode validitetskjeder og en læringsstøttende vurderingskultur i skolen. I vårt datamateriale ser vi noen steder tendenser til dette.

Summary

This is the final report from the FIVIS project «Research on individual assessment in schools» (FIVIS). The report presents the results of the analyses in the third phase of the project in addition to the synoptic analysis of results from the whole project, which ran from 2011 to 2014.

The purpose of the project has been to provide a picture of how assessment practices in the Norwegian schools develop in an interaction between centrally set policies and theories and local interpretations and understandings.

Assessment practices are constructed in a complex interplay between governmental intentions and local understandings and interpretations in the schools. We have followed the phenomenon of assessment in schools from central intentions to local understandings and local practice.

The FIVIS project has had three main research questions:

1. How do school owners, school leaders and teachers express their assessment competency through their interpretations, understandings and practices?
2. What assessment practices exist in Norwegian classrooms?
3. What is the impact of assessment practices on students' experience of learning processes and learning outcomes?

We have tried to answer the research questions using data collected through both quantitative and qualitative methods.

The FIVIS project shows the existence of diversity in assessment practices, both on a collective and on an individual level. There are variations in practices not only between schools and school levels but also within schools. The FIVIS project also shows that there is a great interest in assessment in Norwegian schools today.

The teachers' education and didactic competency seem to matter as to how teachers understand and operationalize their assessment practice. Teachers with weak educational and didactic competency seem to meet greater challenges when trying to understand and operationalize assessment for learning in a satisfactory manner.

The way schools organise their work in teams, either across subjects or along subject lines, seems to have an effect upon the establishment of a community of practice about assessment for learning in the school subjects. Another factor that affects teachers' attitudes and practices is how the school management and school owners understand and organize the local work with assessment for learning. 1–10 schools appear to have greater challenges when trying to establish a community of practice than other schools. This involves having a collective room for discussions, interpretations and the development of a

collective practice. Secondary schools, partly due to a higher number of individually oriented teachers than other school levels, have also to a lesser extent established communities of practise.

Data from interviews and observations, as well as the survey, show that there are large variations in the schools' assessment cultures. The FIVIS project shows that many schools have made good progress in establishing clear goals for students' learning and in involving students in their own assessment. Interestingly, it is clear that those who have participated in the government initiative «Assessment for Learning» have, to a larger extent, developed a culture of assessment at their school. Yet, there is a clear tendency in our data that many of the schools have not succeeded in establishing a community of practice regarding assessment that involves the entire school, including the entire staff.

On several dimensions, there are significant differences between schools and teachers when it comes to assessment practices. We find that teachers trained for the lower levels in the school system score significantly differently (lower) on good assessment practices than teachers who are educated and trained at university/college level. Teachers with one – three years of education in the subject they teach report that they have a better assessment practice than teachers with six months of education in the subject. We also see a clear distinction between assessment cultures and practices in primary school, lower secondary and upper secondary school.

We find a great diversity in assessment practices in the subjects English, Norwegian, physical education and mathematics. These diverse practises are based upon different understandings of the intentions behind individual assessment as well as the curricular goals and objectives and the individual nature of the subjects in the schools. We can also see that different subjects meet different challenges when it comes to validity in assessment. In several schools we have observed that the competency that is to be tested is confused with the assessment form that is used. The quantitative study shows that the teachers' expertise in assessment is the factor with the highest degree of explanatory force when it comes to explaining good assessment practices. Also, we observe a challenge with reference to subject knowledge among teachers in the schools. There seems to be a need for more emphasis on subject-knowledge development in order further to develop teachers' subject-related assessment competency.

The FIVIS project shows that in the Norwegian authorities' focus on assessment for learning, general pedagogy has been emphasized to a greater extent than subject-related and didactical approaches to assessment. This may have resulted in an individual assessment competency that is not adequately

anchored in the teacher's traditional teaching of school subjects. The assessment tools that have been emphasized also seem to be aimed at individual teachers and not at a collectively oriented developmental process. The FIVIS project has shown how school leaders play a crucial role for the development of the school's assessment culture. School leaders' facilitation of school development has bearing on the tensions and contradictions that arise in the school culture when change and development take place. A leader's role in the work will contribute either to inhibiting or promoting development.

The development of good assessment practice must take place both at the individual and at the school level. When it comes to the individual teacher, the most prominent assessment features are the use of working methods, assessment forms, development of a terminology for assessment and the teacher's ability to see the connection between competence aims in the curriculum and assessment. For the sake of the individual teacher's development, the school management must facilitate professional development at the school level, thereby establishing communities of practice that work with assessment in a way that contributes to the individual teacher's professional development. In this context, assessment cultures that promote student learning may grow. This form of competency development opens up for meaning making and participation in communities of practice; it creates interaction, dialogue and reflection on learning.

Our study shows that students at all levels have a relatively high awareness of assessment, although the focus is obviously stronger the higher the level in question. Pupils at all levels in the school system rate assessment as important, even for their own learning. Many students in all age groups recalled a significant awareness of goals. Goals are visible in a variety of ways: in the weekly plans, when teachers write goals on the blackboard and when they discuss goals orally in class. Yet, we see that some students in lower secondary school and upper secondary school have a clear focus on summative assessment, and consequently, they are less concerned about, and they understand less about, how assessment can contribute to their own educational development.

Students talk about how they have been informed of and have helped to discuss assessment criteria, i.e., in the sense of what distinguishes a good answer from a bad one, what must be included in order to get a certain grade. This shows awareness among the students of working with goals and criteria. It may seem as if there are established practices in many classrooms where learning intentions are communicated to students, both through planning and classroom dialogues.

Good feedback is about giving students feedback that tells them about the quality of their work or their performance. At the same time, good feedback informs students about what is necessary for further development of their skills in the subject.

National regulations stress that formative assessment should be «continuous and systematic and can be both oral and written». FIVIS shows that there are large variations in this area. Students express that some teachers are constructive and provide in-depth pointers on ways they can improve themselves, others are short and rather futile in their comments. Students provide a composite picture of what it means to have good assessment competency. «The good teacher» is one who is academically competent, knows his or her subject and who can also teach. In other words: a teacher who has the ability to teach the subject in a way which different students can benefit from. This also includes providing good «feed forward» on student performance. In order to do this, the importance of the relation between the teacher and the students is often highlighted. The teacher must see each student, understand the individual learner and be able to communicate with everyone on his or her own terms.

A central theme for many of the students in our study is the reliability of the assessment they receive. They expect that the same teacher will assess individual students even-handedly. Findings in our study suggest that there are divergent practices in this area. Self-assessment contributes to reflection on own learning. When it comes to self-assessment, it appears that primary schools have progressed further in their work than the other school levels. It seems that the students in the FIVIS project have very varied experiences of assessment for learning. The degree of student activity in assessment activities varies but is generally relatively low. The lack of a community of practice about assessment within and across schools is a work in progress that takes time to internalize. It is nevertheless worth noticing that students believe that good assessment practices with clear objectives and criteria, where they are active in their own and others' learning processes, are important for their learning, motivation and mastery.

The FIVIS project shows that teachers make little use of feedback from students and assessment results in adjusting their own teaching. We see, in other words, a clear potential in developing assessment for learning as part of the development of customised training.

Teachers draw conclusions about student learning and make decisions based on these conclusions. The FIVIS project shows that an awareness of validity and validity chains is a good concept for judging the quality of the work on assessment, whether in the interpretation of the learning objectives, the local

work with the curricula, everyday conversations about learning between teachers and students or the use of standardized tests or procedures for grading in school. This awareness occurs to varying degrees in the different schools. It seems that some schools operationalize validity in assessment as a singular event, using general checklists and other similar tools.

Ensuring validity in assessment from the classroom to the national level is not a uniform procedure where teachers apply the general checklists or control tools. Validity is rather a pervasive aspect of all assessment processes that must be ensured through awareness and procedures for continuous and critical evaluation. A more formal theoretical understanding of the validity and reliability of assessment may contribute to the professionalization of educators in this field.

Our research findings suggest a tension between the management of the national policy on the one hand and the practical work in each classroom on the other. While school management focuses on statistics and grades, and formal complaints about grades, the question of validity in assessment is delegated down to the teacher level. If the administrative apparatus overemphasises reliability in grading, the result may be bureaucratic documentation processes that do not contribute to the development of a culture of good validity chains and supportive assessment for learning in schools. In our data, we see tendencies towards this in some places.

1. Om FIVIS-prosjektets formål, problemstillinger, data og metode

Lise Vikan Sandvik, Gunnar Engvik, Trond Buland

Med dette foreligger den tredje og siste rapporten fra prosjektet «Forskning på individuell vurdering i skolen» (FIVIS). I FIVIS-studien har vi undersøkt hvordan læreres vurderingskompetanse kommer til uttrykk gjennom fortolkninger og forståelser, og hvordan dette manifesterer seg i praksiser i norske klasserom. Alt dette har vi undersøkt for å se hvilken betydning vurderingskulturer har for elevenes læring. Vi har undersøkt skoler som har deltatt i den nasjonale satsingen Vurdering for læring, og skoler som ikke har deltatt i satsingen. Vår forskning viser ikke noen klar tendens til at skoler som har deltatt i utviklingsprosjektet Vurdering for læring har en lærerstab der alle lærerne har utviklet den samme forståelsen av hva god vurderingspraksis betyr. Likevel viser studien vår at ved de skolene som har kommet godt i gang med utviklingsarbeidet, har skoleledelsens involvering i utviklingsprosjektet hatt betydning for arbeidet. Vi har observert mange ulike praksiser innad i skolene, og også mange ulike praksiser hos én og samme lærer. Det synes ikke som det er klare, etablerte forståelser av hvordan vurdering for læring kan bidra til økt motivasjon og mestring hos elevene.

Rapporten bygger på kvantitative og kvalitative data (spørreundersøkelser, intervjuer, observasjoner, styringsdokumenter og plandokumenter). Underveis er det publisert to rapporter i prosjektet (Sandvik, L.V., Engvik, G., Fjørtoft, H., Langseth, I.D., Aaslid, B.E., Mordal, S. & Buland, T., 2012; Sandvik & Buland (red.), 2013). Funnene fra disse rapportene inngår som en del av denne sluttrapporten, sammen med nye analyser av breddeundersøkelsen og intervjuer. Funnene settes videre inn i et utviklingsperspektiv for å belyse hvordan den nasjonale satsingen Vurdering for læring er realisert i norsk skole, og for å kunne legge noen føringer på hvordan utviklingsarbeid av denne type kan videreføres og videreutvikles.

1.1 Problemstillinger

Prosjektet FIVIS har løpt fra 2011 til høsten 2014. Utdanningsdirektoratet var i dette prosjektet interessert i å få et bilde av «hvordan vurdering stimulerer til læring og hva som kjennetegner ulike klasseromspraksiser når vurdering brukes som redskap for læring, spesielt knyttet til fire prinsipper for underveisvurdering som er nedfelt i forskrift til Opplæringsloven, kapittel 3. I tillegg skal halvårsvurdering og forholdet mellom underveis- og sluttvurdering belyses» (Utdanningsdirektoratet, 2011a). For å kunne undersøke dette har vi

tatt utgangspunkt i Utdanningsdirektoratets føringer (Utdanningsdirektoratet, 2010, 2011a, 2011b) for hva som kjennetegner en læringsfremmende vurderingspraksis.

Hensikten med prosjektet har vært å gi et bilde av hvordan vurderingspraksis i norske skoler utvikler seg i et samspill mellom sentralt gitte føringer og teorier og lokale fortolkninger og forståelser. Det er i det komplekse samspillet mellom det vi har kalt intensjoner og lokale forståelser og fortolkninger på skoler at vurderingspraksis blir konstruert. I dette har vi søkt å følge fenomenet vurdering i skolen fra sentrale intensjoner og forståelser via lokal operasjonalisering til lokal praksis. Det er denne beskrivelsen vi sammenfatter og syntetiserer i denne rapporten.

Med utgangspunkt i dette oppdraget ble det formulert tre overordnede problemstillinger for prosjektet FIVIS:

1. Hvordan kommer skoleeieres, skolelederens og læreres vurderingskompetanse til uttrykk gjennom fortolkninger, forståelse og praksiser?

Vurdering handler om hvordan ulike aktører forstår hva som ligger i de ulike politiske føringene, intensjonene, og hvordan de snakker om vurdering og egen vurderingspraksis. Hva vil det si å ha vurderingskompetanse? Hva legger de ulike aktørene i *vurderingskvalitet*?

2. Hvilke vurderingspraksiser finnes i norske klasserom?

Vurderingspraksiser handler her om bruk av ulike didaktiske verktøy og metoder, omfanget av elevvurderingen samt likheter og forskjeller i vurderingspraksiser. Har de nasjonale føringene skapt grobunn for innovative undervisningsmetoder? Blir eksisterende praksiser videreført under nye begreper, eller har gode praksiser faktisk forsvunnet?

3. Hvilken betydning har vurderingspraksisene for elevenes opplevelse av læringsprosesser og læringsutbytte?

Læringsprosesser handler her om hvordan elevene drar nytte av skolens vurderingskultur og vurderingspraksiser i sin læringsprosess. Hvordan kan vurdering stimulere til læring? Hvordan opplever elevene skolens vurderingspraksis, og hvilken betydning har den for deres læring?

Denne rapporten søker å sammenfatte og besvare alle de tre problemstillingene.

Vurderingskompetanse og skolens vurderingskulturer belyses gjennom en kvantitativ analyse av en breddeundersøkelse som ble gjennomført i prosjektets første fase. Hvilke faktorer er det som påvirker lærernes vurderingspraksis og

vurderingskompetanse? Vi undersøker i tillegg om det i fagene engelsk, norsk, kroppsøving og matematikk er ulike faktorer som har betydning for lærernes vurderingspraksis og for utvikling av deres individuelle vurderingskompetanse og skolenes samlede vurderingskompetanse. I de tidligere delrapportene har vi referert til O'Donovan mfl. (2006) sin modell for utvikling av vurderingskulturer. Vi ser denne modellen som et mulig kart over utviklingen i norske skolers vurderingskulturer av flere årsaker. For det første bygger den på de samme faktorene som fremheves i forskrift til opplæringslova (Utdanningsdirektoratet, 2011a). For det andre støttes den av de data som vi har samlet inn i løpet av prosjektet. Og for det tredje kan den tjene som en teoretisk modell som kan gi retning til skoleutviklingsprosesser.

Vurderingspraksiser undersøkes videre gjennom kvalitative analyser av klasseromspraksiser, intervjuer og plandokumenter. Validitetsbegrepet, som er et nøkkelbegrep i FIVIS-prosjektet, danner rammen for våre analyser. I FIVIS-prosjektet forstås *validitet i vurdering* som et dynamisk begrep og anses som relevant for hele den komplekse prosessen fra læreren leser kompetansemålet, til elevens produkter og prestasjoner blir vurdert av læreren. Vi har anvendt validitetsbegrepet for å undersøke både isolerte enkelttilfeller av vurderingspraksiser og mer overordnede sammenhenger i skolen (bl.a. gjennom begrepet *validitetstjeder*). Hvordan kan vi forstå ulike vurderingspraksiser hos enkeltlærere? Hva skal til for å utvikle kollektivt orienterte praksiser? Vurdering sees som en serie slutninger som har konsekvenser for den praktiske profesjonsutøvelsen til lærerne. Vi ser også at systemer knyttet til lagring av data, dokumentasjon i forbindelse med klagesaker og lignende må innbefattes i en overordnet forståelse av validitetsbegrepet. Standardiserte vurderingspraksiser som nasjonale prøver, eksamenskarakterer og lignende tillegges stor vekt i den offentlige debatten om skolen. Balansen mellom lærernes arbeid med formativ vurdering som fremmer elevens læring, og mer *accountability*-orienterte diskusjoner kan derfor forårsake spenninger og motsetninger som både hemmer og fremmer elevens læring. Hickey mfl. (2006) påpeker at dikotomien mellom interne, formative vurderingspraksiser og eksterne, summative praksiser tilslører en dypere forståelse av de ulike vurderingspraksisenes reelle funksjon. Denne forståelsen er imidlertid nødvendig for å sikre systemisk validitet. For å få kunnskap om validiteten i et utdannings- og vurderingssystem må vi derfor også studere de underliggende forståelsene og antakelsene om læring som ligger bak de ulike praksisene, sier Hickey mfl. De er også spesielt opptatt av de ulike representasjonsformene som læringssynet kommer til uttrykk i, for eksempel i form av flervalgsprøver, skriveoppgaver eller muntlige presentasjoner. En analyse av det underliggende læringssynet som fremkommer i det kvalitative materialet, vil kunne belyse validitetsproblematikken på en måte som vil bidra

til å kunne forstå og anvende *validitet* som et terskelbegrep i arbeidet med vurdering.

Elevenes læringsprosesser og læringsutbytte er underliggende all vurderingspraksis. *Validitet* er ikke bare en rent teknisk term som beskriver kvaliteten på lærernes målinger av elevenes læringsutbytte, men kan sies å være en indikator på den overordnede sammenhengen mellom lærernes fortolkninger av læreplanen, planleggingspraksiser og vurderingspraksiser. I sluttrapporten vil vi fremholde validitetsbegrepet som analytisk verktøy for å se på vurderingspraksiser på tvers av fag og trinn. Et sentralt spørsmål vil være hvilken betydning validitet og god vurderingspraksis har for elevenes læring. Det er mulig å anlegge et utenfraperspektiv på validitet, der vi som forskere ser inn i skolesystemet for å avdekke hvorvidt de ulike leddene i fortolkningen og operasjonaliseringen av læreplanen har lav eller høy validitet. Men det er også mulig å se på lærernes og de andre aktørenes forståelse av validitet innenfra, for å avdekke i hvilken grad lærerne tilpasser undervisningen etter elevenes vurderingsresultater. Hvilken innstilling til læring finnes i klasserom der lærere kommuniserer eksplisitte mål uten at elevene selv er delaktig i konstruksjonen av dem? Og hvilket syn på læring er representert hos lærere som involverer elevene i vurderingen?

Slike refleksive prosesser kan sies å sikre validitet i vurderingen fordi undervisningen i seg selv fungerer som en kontinuerlig validering av lærernes og elevenes forståelse av læringsprosessene og -utbyttet. I tidligere delrapporter har vi dessuten sett at lærere som samarbeider om implementeringen av standardiserte tester (f.eks. lesetester på barnetrinnet), i større grad utvikler et felles profesjonelt språk for å snakke om elevenes læring innenfor dette feltet. Dette er et eksempel på det vi kan kalle positiv uintendert backwash – vurderingspraksisen er initiert av eksterne myndigheter og ikke av lærerne selv, og den har en positiv konsekvens fordi den fremmer læring for dem som administrerer den.

Svarene på de tre overordnede problemstillingene omkring vurderingskompetanse, vurderingspraksiser og læringsprosesser sammenfattes avslutningsvis i en analyse av vurderingskulturer som aktivitetssystem. Hva hemmer og fremmer utviklingen av gode vurderingspraksiser i skolekulturen? Hvilke elementer ved individuelle og kollektive utviklingsprosesser i skolen kan bidra til elevenes læring, motivasjon og mestring?

1.2 Om rapporten

I dette kapitlet har vi så langt redegjort for forskningsprosjektets formål og problemstillinger. Vi har også plassert denne sluttrapporten inn i prosjektets

ramme. I kapittel 1.3. oppsummeres hovedfunn fra delrapport 1 og 2. I kapittel 2 presenterer vi datakildene og metoden som ligger til grunn for våre analyser. I de videre kapitlene danner praksisfortellinger utgangspunktet for de påfølgende drøftingene. Gjennom disse praksisfortellingene gis en kontekst for å kunne forstå hva som kan forklare de mange vurderingspraksiser og -kulturer vi har kunnet observere gjennom vår forskning. Dessuten kan de danne grunnlag for en forståelse av hva som kan bidra til utvikling av gode praksiser.

For å synliggjøre sammenhengen mellom empiriske funn fra praksis og mer overordnede teoretiske perspektiver har vi valgt å presentere relevant teori i hvert kapittel, direkte knyttet til de empiriske analysene. Dermed framstår rapporten som en kinesisk eske, med ulike fortellinger og forskjellige teoretiske innganger, som belyser de samme overordnede problemstillingene som FIVIS tar utgangspunkt i. På den måten forsterkes også inntrykket av at god vurderingspraksis er knyttet til mange nivåer i skolen – fra enkelteleven, undervisningen i klasserommet, lærerens fagkompetanse via foreldrenes rolle, skolens utviklingsarbeid, ledelsen ved skolen, skoleeiers påvirkning, til nasjonale styringsdokumenter. Denne kompleksiteten søker vi å synliggjøre gjennom rapportens struktur.

I kapittel 3 drøfter vi med utgangspunkt i én praksisfortelling sentrale problemstillinger knyttet til arbeidet med god vurderingspraksis, før dette settes inn en teoretisk kontekst knyttet til validitet og reliabilitet i undervisning.

Kapittel 4 presenterer flere praksisfortellinger fra undervisningssituasjoner. Disse kan sees som fortattede fortellinger og illustrerer mangfoldet i vurderingspraksiser og -kulturer som er observert og analysert i dette forskningsprosjektet. Etter hver fortelling gis det en refleksjon knyttet til fortellingen. Etter disse praksisfortellingene presenteres en modell for å forstå ulike praksiser og kulturer. Slik settes lærernes og skolenes ulike tilnærminger til vurdering inn i et teoretisk rammeverk som kan bidra til å forstå arbeidet med vurdering for læring både som et individuelt anliggende og som et organisatorisk anliggende knyttet til utviklingen av tolkningsfellesskaper i skolene.

Kapittel 5 setter elevenes fortellinger i fokus og presenterer deres opplevelse av vurdering og læring.

Kapittel 6 presenterer to praksisfortellinger om ledelse og læring i skolen. Deretter analyseres innholdet i fortellingene i lys av aktuell teori om læring i organisasjonen skole og ledelse av læreres læring.

I kapittel 7 presenteres funn fra den kvantitative breddeundersøkelse, der hovedspørsmålet er hva som kjennetegner god vurderingspraksis.

Hvert av kapitlene 3–6 står som selvstendige bidrag inn mot en forståelse av vurderingskompetanse, vurderingspraksiser og læringsprosesser i norsk skole. For hvert kapittel er det angitt ansvarlige forfattere. Likevel fremstår rapporten som en helhet, der de sentrale funnene fra de ulike kapitlene drøftes i et utviklingsorientert syn på læring i kapittel 8. I kapittel 9 gis et sammendrag av rapportens hovedfunn med implikasjoner for videre utviklingsarbeid og forskning på feltet vurdering.

1.3 Hovedfunn fra delrapport 1 og 2

I dette kapitlet presenteres en kortfattet sammenfatning av hovedfunnene i delrapport 1 og 2 i FIVIS-prosjektet.

1.3.1 Delrapport 1: Vurdering i skolen. Intensjoner og forståelser

I delrapport 1 var hovedproblemstillingen:

Hvordan kommer skoleeieres, skolelederes og læreres vurderingskompetanse til uttrykk gjennom fortolkninger, forståelser og praksiser?

Rapporten presenterte hovedfunn fra første del av prosjektet, blant annet analyser av sentrale føringer for arbeidet, basert på overordnede styringsdokumenter, og et utvalg av lokale læreplaner. I de sentrale føringene ligger de intensjonene som finnes for vurdering i skolen fra nasjonalt hold, og som er forankret i forskning omkring vurdering. Det viktigste formålet med disse analysene var å gi en forståelse av hvilke politiske føringer, prosesser og vedtak som har bidratt til å utvikle dagens rammer rundt individuell vurdering i skolen.

Gjennom kvalitative fokusgruppeintervjuer med lærere, skoleledere og skoleeiere og kvantitative spørreundersøkelser av de samme gruppene søkte vi å få et innblikk i ulike aktørers forståelser av og begrepsbruk knyttet til vurdering i skolen. Med *forståelser* menes de ulike aktørenes tolkninger av intensjonene.

Resultatene fra prosjektets første del viste at lærerne er overveiende positive til de nasjonale føringene for individuell vurdering. Over halvparten av lærerne sa seg enige, og 20 prosent sa seg faktisk helt enige i at de har god kjennskap til vurderingsforskriften. De som er med i den nasjonale satsingen VFL, var noe mer enige i at forskriftens føringer for undervisningsvurdering har bidratt til å endre deres praksis. Et interessant funn er at mens bare halvparten av lærerne

rapporterte at de hadde inngående kjennskap til forskriften, rapporterte de fleste om vurderingspraksiser som ligger svært nær opp til intensjonene.

Våre data viste videre at fagkompetanse er en viktig faktor i lærernes vurderingskompetanse. Særlig mente lærere med høy fagkompetanse dette. Lærere i barneskolen framhevet i mindre grad dette som viktig for egen vurderingskompetanse. Behovet for fagdisiplinær kompetanse kom også sterkt til uttrykk i fokusgruppeintervjuene. Det oppleves som nødvendig å ha faglig trygghet for å gjøre forventningene i faget eksplisitte og tydelige for elevene.

Et flertall av våre informanter i første fase av prosjektet mente at lokale læreplaner har klare, konkrete mål. Et flertall av lærerne uttrykte at tydelige mål og kriterier i forkant av undervisningen oppleves som motiverende for elevene. Et gjennomgående trekk var at barneskolelærerne har kommet lengst i dette arbeidet, og at vurdering forstås som et arbeid med å bygge opp læringsmål for relativt korte perioder.

Mange lærere er mer opptatt av å utvikle vurderingsverktøy som fungerer i det daglige, enn av å følge opp dette i læringsprosesser over lengre tidsspenn. Klasser med mange elever og liten tid til planlegging pekes på som begrensninger i dette arbeidet. Flertallet i vår undersøkelse var enige i at undervisvurdering styrker muligheten for å tilpasse opplæring til den enkelte elev. Kartleggingsprøver benyttes i høy grad, men mange av lærerne uttrykte frustrasjon når det gjaldt å benytte resultatene fra slike prøver til rangering av elever. Et flertall av lærerne oppga også at de benytter resultater fra elevvurderingen til å videreutvikle egen vurderingspraksis.

Krav om slutt- og halvårsvurdering ser ikke ut til å hindre god undervisvurdering, og lærerne rapporterte at de kjenner til hva som skal inngå i denne. Grunnskolenes barnetrinn har en historie på over 40 år med å vurdere motivasjon, arbeidsinnsats, arbeidsprosess og elevens fremgang ut fra forutsetninger og elevens personlige utvikling, og dette ser ut til å ha gitt et solid grunnlag.

I videregående opplæring har det vært en like lang tradisjon for vurdering med bruk av karakterer og et målrelatert vurderingssystem. Vår analyse antyder at det er ulike vurderingskulturer på tvers av forvaltningsnivå, skoleslag og innad i skolene. Når det gjelder undervisvurdering, viser vår studie få signifikante forskjeller mellom fagene. Et unntak er faget kroppsøving, der lærerne i noe mindre grad enn i de andre fagene opplevde at undervisvurdering er en integrert del av deres undervisning, og at de i noe mindre grad setter tydelige mål for elevenes læring. Kroppsøvingslærerne etterlyste også sterkere nasjonale føringer for vurderingen.

I fokusgruppeintervjuene kom det frem tydelige forskjeller mellom lærernes forståelse av vurderingens rolle i undervisningen. Matematikkfagets utfordringer synes å ligge i lærernes forståelser av hva matematisk kompetanse innebærer, og hvordan man kan vurdere den. For kroppsøvingfagets del synes spenningene å ligge i diskursen omkring prestasjoner og innsats, mens de for språkfagenes del i høy grad handler om sprikende forståelser av hva som utgjør gode bevis for læring, og hva som kan regnes som gyldige slutninger om elevenes kompetanse.

På skoleeiersiden ser det ut som om hovedfokus i engasjementet rundt vurdering ligger på reliabilitet. Skoleeiere ser primært ut til å være opptatt av at skolene utvikler en felles vurderingspraksis, og at alle elever får en rettferdig vurderingspraksis. Utvikling av et godt fortolkningsfellesskap som kan fremme validitet i alle ledd i undervisning og vurdering, er i høy grad delegert ned til fagnivå i skolen. Det viktige fortolkningsfellesskapet rundt vurdering omfatter altså i liten grad skoleeier. Skoleeier spiller dermed i mindre grad en aktiv rolle i utviklingen av vurderingspraksiser med en sterk forståelse av det faglige innholdet.

Å kunne delta i en utvidet diskusjon om vurdering er et viktig element i vurderingskompetanse. Vår studie viser at dette er under utvikling, men at det her er relativt store forskjeller mellom skoler. Et flertall i vårt utvalg oppgir at vurdering er et tema lærerne relativt ofte tar opp i lærerkollegiet, og at lærerne samarbeider om å skape gode vurderingspraksiser. Vi finner at det er relativt store forskjeller mellom ulike skoleslag i spørsmålet om hvorvidt vurdering er tema i diskusjoner/samtaler lærere imellom. Å skape rom for å diskutere vurdering i fag kan synes svært viktig for å sikre en mer valid vurderingspraksis i fagene på skolenivå.

1.3.2 Delrapport 2: Vurdering i skolen. Operasjonaliseringer og praksiser

I delrapport 2 var fokuset i all hovedsak rettet mot vurderingspraksiser i skolen og elevenes opplevelse av læringsprosesser i fagene norsk, engelsk, matematikk og kroppsøving. Det empiriske materialet besto her av klasseromsobservasjoner og intervjuer av elever, lærere, foreldre og skoleledere i åtte skoler i 2012 og 2013. I tillegg samlet vi inn et utvalg av skolens planer – årsplaner, halvårsplaner og periodeplaner – i fagene norsk, engelsk, matematikk og kroppsøving.

Funnene i denne andre fasen av studien var i all hovedsak basert på et kvalitativt materiale. Det betyr at vi søkte etter mønstre av vurderingspraksis med utgangspunkt i de faktiske observasjonene vi gjorde i klasserommet, og i studier av skolens plandokumenter. Gjennom intervjuene understøttes disse

mønstrene i deltakernes opplevelse av vurdering i skolen. Samlet belyser det innsamlede kvalitative materialet arbeidet med vurdering i norske klasserom, og studien belyste også hvilken betydning dette har for elevenes læringsprosess i fagene. Sentrale perspektiver her er tolkningsfellesskaper, validitet i vurdering, læringsprosesser og ledelse av utviklingsarbeid i vurdering.

I delrapport 2 fokuserte vi fortsatt på arbeidet med vurdering i fagene norsk, engelsk, matematikk og kroppsøving for å få en bedre oppfatning av forholdet mellom skoleledernes og lærernes forståelser og praksiser. Dessuten var elevperspektivet sentralt i denne analysen. Vurdering er til for å fremme læring, motivasjon og mestring. Derfor blir elevstemmene, og også foreldrestemmene, viktige for å få mer kunnskap om betydningen av dette arbeidet for læringsprosessene deres.

Når det gjelder utvikling av nye praksiser eller videreutvikling av praksiser i en skole, er også skolelederens rolle av stor betydning. Vi så også i fase 2 av vår studie at skolelederen er sentral i arbeidet med å etablere et tolkningsfellesskap. Dette fellesskapet er nødvendig for å skape en vurderingskultur som bygger opp under en god vurderingspraksis.

Særlig i 1–10-skoler understreker ledere at det er en utfordring å samle kollegiet om en kollektiv forståelse av en slik praksis. I skoler som har arbeidet med Vurdering for læring, gir ledelsen uttrykk for en mer presis og oppdatert forståelse av hva som legges i sentrale begreper i satsingen, enn i skoler som ikke har vært med i Vurdering for læring. Ledere som har vært med i den nasjonale satsingen, legger i større grad vekt på å utvikle forståelse i praksisfellesskap på skolene og har arbeidet med å forankre satsingen på flere nivå i skolen.

Generelt ser det ut som om ungdomsskolene og 1–10-skolene har flere utfordringer knyttet til å utvikle en felles forståelse av hva som er god vurderingspraksis, enn rene barneskoler og videregående skoler. Skoler som har vært med i Vurdering for læring har også kommet lenger enn skoler som ikke har deltatt i satsingen.

Hovedinntrykket av vurderingspraksiser i de ulike fagene etter fase 2 av prosjektet var det store mangfoldet av praksiser vi finner. Disse varierende praksisene bygger på forskjellige forståelser av intensjonene med individuell vurdering i skolen. Vi ser at mange skoler er på vei mot å utvikle vurderingskulturer. Ved flere observerte vi en bevegelse mot en større grad av arbeid i praksisfellesskap, der læringsmål formidles gjennom deltakelse og dialog i kunnskapsutvekslinger i faglige kontekster. De «tradisjonelle» vurderingspraksisene fant vi først og fremst hos lærere i skoler der det så ut til å være svak kultur for å drive utviklingsarbeid på skolenivå.

Studien vår viste at de forskjellige fagene har utfordringer med hensyn til validitet i vurderingen. Ved flere skoler så vi eksempler på at man blander den kompetansen som skal prøves, og den representasjonsformen som kunnskapen skal prøves gjennom.

Med hensyn til vurderingens pålitelighet så vi i flere skoler en utfordring. Dette gjelder spesielt praksis knyttet til beskrivelser av kjennetegn for grad av måloppnåelse. Analysene av lokale læreplaner viser for eksempel ofte relativt svak sammenheng mellom skolens årsplaner og de enkelte trinns og klassers ukeplaner og periodeplaner.

Vår analyse viser at lærerne i høy grad mener at tydelige mål og kriterier i forkant av undervisningen oppleves som motiverende for elevene. Dette støttes også av elevene og foreldrene. Foreldre legger også vekt på at god vurdering er et godt hjelpemiddel i deres eget arbeid med å støtte egne barns læring.

Vårt inntrykk er ellers at foreldre/foresatte ofte tenker på vurdering som vurdering av læring mer enn vurdering for læring. Det er det summative aspektet ved vurderingene som dominerer i foreldrenes forståelse. Den vurdering som skjer i undervisningen, med løpende tilbakemeldinger på elevs arbeid, tolkes av mange foreldre som noe annet, som undervisning.

Vår studie viser at lærernes vurderingspraksis speiler ulike kulturer i skolene og på tvers av trinn og skoleslag. Lærerne rapporterer at det er betydelige forskjeller i vurderingspraksis mellom lærere, noe våre observasjoner understøtter. Det er også tydelig i analysen i delrapport 2 at fagkompetanse er en viktig faktor i lærerens vurderingskompetanse. Lærere med høy fagkompetanse er mer opptatt av at utdanning og egne erfaringer som lærer er viktige for egen utvikling, mens lærere med lavere fagutdanning løfter fram etter- og videreutdanning som sentrale elementer i utviklingen av egen vurderingskompetanse.

Også på skolenivå er det utfordringer knyttet til fagkompetanse. I barneskolen og på ungdomstrinnet observerte vi lærere som underviser i fag de ikke har faglig fordypning i. Dette synes i noen tilfeller å føre til tilfeldige praksiser. Lærerne med liten eller ingen fagkompetanse synes å ha lavere læreplanforståelse enn lærere med høy fagkompetanse, og de evner i mindre grad å iverksette intensjonene som ligger i de ulike nasjonale føringene. Dette kan føre til en instrumentell tilnærming til vurdering. Det virker som om det er et behov blant lærerne for å utvikle flere verktøy som kan fortolke og forstå hva som ligger i de faglige lærings- og vurderingsbegrepene. Dette så vi både gjennom arbeidet med lokale læreplaner og gjennom våre observasjoner og intervjuer.

2. Metode og datakilder

Trond Buland og Gunnar Engvik

I prosjektet FIVIS har vi søkt å belyse problemstillingene med et rikt sett av datakilder. Vårt overordnede prinsipp har vært metodetriangulering, vi har altså i størst mulig grad ønsket å belyse alle problemstillinger ved hjelp av data innsamlet på ulike måter. Prosjektets viktigste datakilder har vært:

- en kartleggende breddesurvey som ble gjennomført blant et utvalg lærere vinteren og våren 2012
- intervjuer med skoleeiere, skoleledere, lærere, foreldre og elever i utvalgte skoler
- observasjon av undervisning i fagene engelsk, matematikk, norsk og kroppsøving i utvalgte skoler
- dokumenter i form av læreplaner, forskrifter, lokale læreplaner og undervisningsplaner

Vi vil i det følgende presentere metode og gjennomføring av det kvalitative og det kvantitative datamaterialet.

2.1 Kvalitativ metode

En betydelig del av datagrunnlaget for våre analyser er framkommet gjennom kvalitative metoder. Analysen i sluttrapporten bygger på observasjoner av undervisning i fagene norsk, engelsk, matematikk og kroppsøving ved åtte skoler gjort i forbindelse med delrapport 2 (Sandvik & Buland, 2012). I tillegg intervjuet vi skoleledelsen, lærere, elever og foreldre/foresatte ved disse skolene. Vi samlet også inn lokale læreplaner, ukeplaner, lokale vurderingskriterier og ulike planer som var utarbeidet i forbindelse med de timene som ble observert. Det sentrale fokuset i observasjonene og intervjuene har vært på vurderingspraksiser.

I valg av caseskoler lå følgende kriterier til grunn:

- Deltakelse / ikke deltakelse i satsingen Vurdering for læring.
- Alle skoleslagene skulle være representert (barneskoler, 1–10-skoler, ungdomsskoler, videregående skoler).
- Beliggenhet både i urbane og rurale strøk.

I kvalitative undersøkelser er hovedmålet å få dypere innsikt i ulike elementer sett fra informantenes side. I motsetning til i kvantitative studier er målet således ikke å skaffe til veie statistiske eller generaliserbare data.

2.1.1 Observasjoner

På alle caseskoler planla vi å foreta en-to observasjoner av undervisningsøkter i de aktuelle fagene på følgende trinn: 4. trinn, 7. trinn, 10. trinn og Vg1. Dette ble i hovedsak gjennomført, men i noen få tilfeller / ved noen få skoler gjorde timeplanmessige forhold det umulig å gjennomføre observasjoner i alle fag.

Tabell 1 viser en oversikt over observasjonene.

Tabell 1: Klasseromsobservasjoner fordelt på klasse og trinn

Fag	Trinn	Antall	Total
Engelsk	4. trinn	5	17
	7. trinn	3	
	10. trinn	5	
	Vg1	4	
Norsk	4. trinn	3	13
	7. trinn	4	
	10. trinn	3	
	Vg1	3	
Matematikk	4. trinn	4	18
	7. trinn	4	
	10. trinn	4	
	Vg1	2	
Kroppsøving	4. trinn	4	12
	7. trinn	3	
	10. trinn	3	
	Vg1	2	
			50

Våre observasjoner kan kategoriseres som passiv deltakende observasjon. Med *deltakende* mener vi at vi var til stede i klasserommet og både elever og lærere var fullstendig klar over vår tilstedeværelse. Elevene var også klar over hvem vi var, hvor vi kom fra, og hvorfor vi var der. Elevene hadde muligheten til å komme og snakke med oss etter undervisning, men ikke i løpet av timen. Med *passiv observasjon* menes det at det ikke skjedde noen form for styring, kommunikasjon eller avbrytelser fra forskernes side i løpet av observasjonsperioden. Slik sett har vi søkt å påvirke observasjonsfenomenet i minst mulig grad. Likevel er det klart at all deltakende observasjon kan påvirke

både aktørene og fenomenet man skal observere. På samme måte vil observatøren tolke sine observasjoner inn i sin forståelsesramme.

Å få tilgang til feltet og å få observere undervisning handler om å etablere gode relasjoner til forskningsdeltakerne. Det er ikke selvsagt at man som forsker blir akseptert, eller at deltakerne vil samarbeide med en. *Diskresjon, sensitivitet og sunn fornuft* er overordnede begreper i en slik kontekst (Hitchcock, 1995). Som forskerteam brukte vi tid på å finne gode måter å komme i kontakt med skolene og lærerne på, og vi diskuterte også hvor følsomt det kan være for læreren å bli observert. Vi gjorde det derfor til en rutine å ha gode samtaler med læreren før og etter observasjonene. I disse samtalen fikk vi også oppklart eventuelle misforståelser og fikk diskutert den observerte undervisningstimen i en større kontekst.

For å styre våre egne observasjoner til en viss grad og for å få et sammenligningsgrunnlag mellom de ulike caseskolene ble det utarbeidet et observasjonsskjema (se vedlegg 2). Dette skjemaet ga føringer i forhold til hva man skulle se etter, og hva som var mest viktig å rette fokuset på. Observasjonsskjemaet finnes som vedlegg til denne rapporten. Samtidig er det klart at forskerne måtte benytte dette skjemaet med fleksibilitet, avhengig av den observerte timens innhold. Hensikten med skjemaet var å fungere som en guide og en huskeliste for observatørene og å sikre at vi hadde et nokså likt fokus under observasjonen i de ulike fagene.

Observasjonene ble i all hovedsak gjennomført av forskere med faglig og fagdidaktisk kompetanse i det aktuelle faget. Et viktig argument i vår studie er nettopp betydningen av fagkompetanse og fagdidaktisk kompetanse i arbeidet med vurdering. For å kunne forstå og fortolke de observerte vurderingspraksisene og kulturene i fagene norsk, engelsk, matematikk og kroppsøving mener vi at også forskeren på feltet bør inneha denne kompetansen. I noen tilfeller observerte forskerne timer alene, noen ganger var vi to og to sammen. Disse parobservasjonene var også med på å sikre validiteten i observasjonene. Analysene av observasjonene danner grunnlag for analyser i sluttrapporten, og noen observasjoner er reanalysert i sluttrapporten.

2.1.2 Intervjuer

Vi benyttet fokusgruppeintervju i alle intervjuene i studien. Intervjuene var kvalitative, semistrukturerte intervjuer, der hensikten var å belyse bestemte problemstillinger og sentrale elementer knyttet til vurdering i skolen. Et fokusgruppeintervju er en godt planlagt gruppediskusjon, designet for å få tak i synspunkter på gitte tema (Wibeck, 2000). Ved hjelp av vår intervjuguide ledet vi intervjuene gjennom de tema som var nødvendige i forhold til prosjektets

målsettinger. Intervjuenes semistrukturerte form ga også rom for å avvike fra guiden og følge interessante problemstillinger der de dukket opp. En slik metode gjør det mulig for forskeren å følge informantens fortelling og utdype temaer som kommer opp, også der forskerne i utgangspunktet ikke hadde inkludert dem i sine intervjuguider.

Alle intervjuene ble gjennomført som gruppeintervjuer. De ble tatt opp elektronisk og transkribert. Lærere i norsk og engelsk ble noen ganger intervjuet sammen, mens lærere i matematikk og kroppsøving ble intervjuet i egne grupper. I noen tilfeller hadde lærere problemer med å møte til avtalt gruppeintervju og ble derfor intervjuet individuelt. Ved de fleste skolebesøkene ble det også gjennomført kortere intervjuer med lærere umiddelbart etter observasjon. Dette dreide seg om individuelle intervjuer/samtaler.

Vi intervjuet elever og foreldre fra 4., 7. og 10. trinn. Foreldreintervju ble gjennomført samlet på tvers av flere trinn, med unntak av ved én skole. Ved noen få skoler ble det også gjort kortere intervjuer med elever i forbindelse med observert time.

Intervjuene med faglærere ble i all hovedsak foretatt av fagdidaktikere i forskergruppen med de relevante fagene. Intervjuer med skoleledelsen ble stort sett gjennomført av forskere med erfaring fra denne type intervju.

Foreldreintervju og intervju med elever ble gjort av alle i forskergruppen. Vi var stort sett to forskere til stede ved alle intervjuer. Intervjuene fra de to delrapportene og funn knyttet til analyser av dem er integrert i sluttrapportens analyser. Enkelte intervjuer er også reanalysert i sluttrapporten. Tabell 2 viser en oversikt over intervjuene.

Tabell 2: Intervjuer

Kategori	Antall intervjuer	Antall respondenter
Skoleledere	8	18
Lærere	59	94
Foreldre	9	35
Elever	17	63
Totalt	93	210

2.1.3 Dokumenter

Ulike dokumenter, som læreplaner i fag, forskrifter, lokale læreplaner, undervisningsplaner fra hver enkelt lærer, skjema med vurderingskriterier og

ulike vurderingsoppgaver, kan gi et godt bilde av konteksten for vurderingspraksisene.

I en analyse av arbeidsplaner i delstudie 1 samlet forskergruppen inn lokale planer fra skolene som deltok i den første intervjuunden (Sandvik mfl., 2012). Planene ble lastet ned fra Internett fra barneskolene, hvor de lå tilgjengelige på skolenes hjemmesider, og gjennom e-postkontakt med de videregående skolene, der planene ligger lukket på læringsplattformer (LMS). Vi henvendte oss også direkte til dem som hadde deltatt på intervjuene, via e-post. Det viste seg vanskelig å innhente læreplaner fra de enkelte skolene og fagene. I kroppsøving fikk vi ikke inn en eneste fagspesifikk plan. Faget stod bare nevnt i timeplanene på ukeplanene. Det kom heller ikke inn planer fra norskfaget på videregående skole. Det viste seg også at ungdomsskolene ikke fulgte opp oppgaven med å sende inn lokale planer. I alt omfattet undersøkelsen 14 ukeplaner. På barnetrinnet er ukeplanene utarbeidet slik at hvert fag har sin plass på planen. I denne analysen regnes hvert fag som en plan. Bare to årsplaner kom inn, en fra engelsk i videregående skole og en årsplan fra barnetrinnet som omfattet alle fagene.

I delstudie 2 (Sandvik & Buland, 2013) var situasjonen annerledes. Under besøkene samlet vi systematisk inn lokale læreplaner, undervisningsplaner og ulike skjemaer lærerne brukte i de forskjellige fagene. Dette dannede utgangspunkt for analysene i delrapport 2. Disse dokumentanalysene brukes som grunnlag for videre analyser i sluttrapporten.

2.1.4 Analyser av det kvalitative datamaterialet

Arbeidet med analysen av det empiriske materialet i denne delstudien har foregått i en konstant dialog med, og i videreutvikling av, de teoriene og modellene vi presenterte i delrapport 1 og 2.

Vi har i fellesskap utviklet noen teoretiske rammeverk som danner utgangspunktet for våre analyser, og som i korthet er presentert i dette kapitlet. Noen temaer og perspektiver er mer sentrale innenfor enkelte fagområder. Underveis i analyseprosessen varierte det derfor hvor stor vekt enkelte perspektiver og temaer fikk. Analysene i denne studien vil også bære preg av denne forskjelligheten.

Selv om vi har prøvd ut ulike teoretiske perspektiver i våre analyser, har også de kvalitative analysene vært preget av en induktiv tilnærming. Vi startet analysen av materialet med en holistisk tilnærming der hovedinntrykket ble oppsummert, hvilke data som støttet dette, og om det eventuelt var utsagn og observasjoner som tilsa noe annet. I de videre analysene ble materialet

analysert og kategorisert i NVivo 9 (International, 2011). Utgangspunktet for kategoriene her var forskningsspørsmålene som ble stilt i første og andre delrapport, som også var utgangspunktet for intervjuguidene og observasjonsguidene. Analyser og presentasjoner av datamaterialet hviler på en konstant komparativ metodisk tilnærming (Gudmundsdottir, 2001; Postholm, 2005), der vi har vært opptatt av å identifisere mønstre og se sammenhenger og motsetninger i materialet.

2.1.5 Er casene representative?

Det er viktig å påpeke at *skolene* vi har besøkt, ikke er valgt på bakgrunn av resultater på nasjonale prøver, tester eller lignende, men på grunnlag av at de passer til utvalgskriteriene. Likevel har vi ikke gjort et tilfeldig, men et strategisk valg av skoler. Vi har ønsket å besøke skoler som vi, basert på forhåndskunnskap, hadde grunn til å tro kunne belyse ulike deler av våre problemstillinger på en konstruktiv måte.

Deltakelsen i denne delen av prosjektet var frivillig. Vi sendte forespørsler til ulike skoler som passet i forhold til våre kriterier, for så å få svar på om de kunne tenke seg å være med eller ikke. Flere skoler vi henvendte oss til, valgte å ikke stille opp som caseskole. Likevel lyktes vi i å få med skoler som samlet sett tilfredsstillte alle kriteriene.

Den enkelte informant er svært avgjørende for resultatene. Utvalget av skoler kan være skjevt. Slik sett er det viktig å reflektere over hvilke informanter som er med i studien, og hva som eventuelt kan særprege disse. I første omgang kan man si at selv det å bestemme seg for å være en del av datainnsamlingen som en caseskole kan ha betydning. Dette er skoler som har sagt ja til å ha besøk av forskere i to dager. Skolene har måttet være behjelpelige med utsendelse av informasjonsskriv til både ansatte, elever og foresatte, organisert tid og sted for ulike intervjuer og observasjoner samt selv stilt opp til intervju. Med andre ord har det vært krevende og medført en del organiseringsarbeid å delta som caseskole i vår studie.

Foreldreintervju ble normalt forsøkt gjennomført etter normal arbeidstid. Disse intervjuene ble altså lagt til perioden 16–20 og fant sted på den enkelte caseskole. Oppmøtet var litt varierende, noe som erfaringsmessig er helt normalt for slike fellesmøter innenfor en skolerelatert studie. De foresatte fikk informasjon om foreldremøtet ca. en uke i forveien. Man kan anta at de foresatte som valgte å møte opp, også er de som er mest engasjert i tematikken, men vårt inntrykk er at vi nådde et bredt utvalg av foreldre/foresatte.

2.1.6 Er resultatene valide?

Validitet i forskning sier noe om hvorvidt de benyttede metoder har bidratt til å svare på forskningsspørsmålene, om man har målt det man har hatt til hensikt å måle.

I all kvalitativ datainnsamling vil forskeren være det viktigste datainnsamlings- og analyseverktøyet. Med dette menes det at *forskerens* oppfattelse, forforståelse og evne til å se sammenhenger og meninger i høy grad vil prege resultatet av studien. Derfor har det vært viktig for oss å arbeide grundig med både observasjonsskjema og intervjuguider for å være best mulig forberedt før datainnsamlingen.

Ved observasjon i et klasserom bør man være klar over at observatørens tilstedeværelse kan påvirke det faktiske fenomenet som skal observeres. Med dette menes at både elever og lærere kan være påvirket av vår tilstedeværelse i de timene vi gjennomførte observasjonene, samt at timene kan være planlagt noe annerledes enn det som er normalen, siden læreren på forhånd visste at vi skulle komme. Dette forholdet må derfor tas som et slags bakteppe inn i analysene.

Vi har gjennomført gruppeintervju med *elever*. Elevene og deres foresatte hadde muligheten til å reservere seg mot deltakelsen både i observasjonen og i intervjuene. Observasjoner og intervjuer har bare foregått der hverken elevene eller deres foresatte motsatte seg dette. Ved skolene var det lærerne som valgte hvilke elever vi kunne snakke med. På forhånd hadde vi kommet med et ønske om å intervju grupper på 5–6 elever fra hvert av de valgte trinnene. Dermed er det klart at utvelgelsen kan ha vært subjektiv og muligens favoriserte elever som læreren visste ville ytre seg positivt om undervisningen. Vi har imidlertid ikke inntrykk av at dette har vært tilfellet. Elevene har gitt uttrykk for svært ulike meninger og holdninger, og både positive og mer kritiske røster har vært rikelig representert.

I kvalitative undersøkelser er som kjent hovedmålet ikke å skaffe til veie statistiske eller generaliserbare data, men å få dypere innsikt i ulike elementer sett fra informantenes side. I all kvalitativ datainnsamling kan man si at forskeren er det viktigste datainnsamlings- og analyseverktøyet. Med dette menes det at vår egen oppfattelse, forforståelse og evne til å se sammenhenger og meninger i høy grad vil prege resultatet av studien. Derfor har det vært viktig for oss å arbeide grundig med både observasjonsskjema og intervjuguider for å være best mulig forberedt før datainnsamlingen. At vi som regel har søkt å være to forskere til stede ved intervjuer og observasjoner, har også bidratt til den løpende kvalitetskontroll, noe som har bidratt til studiens validitet.

2.2 Kvantitativ breddesurvey

2.2.1 Datagrunnlag og utvalg

Datagrunnlaget bygger på en kartleggende breddesurvey som ble gjennomført blant et utvalg lærere vinteren og våren 2012. Hovedfunnene i surveyen er presentert i FIVIS' delrapport 1 (Sandvik mfl., 2012), og hensikten med disse analysene var å få en oversikt over «[...] utbredte rutiner og metoder knyttet til vurdering og hvordan lærerne oppfatter disse» (Sandvik mfl., 2012, s. 19).

I samarbeid med oppdragsgiver, Utdanningsdirektoratet, ble det besluttet å trekke et utvalg skoler fra det utvalget som deltar i direktoratets «spøringer» som sendes ut en gang i halvåret. Utvalget til skolespørningen er representativt. Med tanke på geografi (fylker/landsdel), skolestørrelse, skoletype og kommunestørrelse ble 50 prosent av grunnskoleutvalget og 70 prosent av utvalget på videregående nivå tilfeldig trukket ut til å delta. 530 grunnskoler og 100 videregående skoler fikk i uke 12 i 2012 tilsendt undersøkelsen, og en sentral bakgrunnsfaktor for undersøkelsen var om skolene deltok i den nasjonale satsingen «Vurdering for læring» eller ikke.

Fra de 530 grunnskolene svarte 500 lærere. Det utgjør 24 prosent av populasjonen. Fra de 100 videregående skolene svarte 290 lærere. Svarandelen for disse skolene er 45 prosent.

Svarandelen er ikke tilfredsstillende med tanke på representativitet på lærernivå. Vi har derfor undersøkt om det er noen skjevfordelinger i utvalget med hensyn til fordeling på fag og kjønn. Med bakgrunn i årsrammer for fagene og gjennomsnittlig leseplikt i grunnskolen og i videregående skole finner vi at datasettet er representativt med tanke på antall lærere som underviser i fagene engelsk, kroppsøving, matematikk og norsk i grunnutdanningen.

Siden vi ikke finner noen signifikante avvik i utvalget med hensyn til fordelinger på fag og kjønn, betrakter vi det som representativt for norsk skole også på lærernivå.

2.2.2 Breddesurvey

Breddesurveyen fungerer som et bidrag til å kunne si noe mer om lærernes intensjoner og tolkninger knyttet til individuell vurdering i skolen. Tematisk er undersøkelsen koblet til lærernes holdninger til individuell vurdering i skolen,

deres og skolens vurderingspraksiser og vurderingskompetanse og kjennskap til nasjonale føringer for underveis- og sluttvurdering. Undersøkelsen søker i tillegg å favne lærernes oppfatninger om elevens læring og elevens involvering i å vurdere egen læring og utvikling. Dette danner et viktig bakteppe for utviklingen av spørsmålsbatteriene i breddesurveyen, som ble utviklet i samarbeid mellom hele forskergruppen.

Arbeidet tok utgangspunkt i tidligere forskning rundt vurdering i skolen og skjedde i tett dialog med oppdragsgiver, Utdanningsdirektoratet. På denne måten sikret vi at surveyen var både faglig oppdatert og i tråd med eksisterende forskning samt bestående rammer og strukturelle forhold. Vi vurderer at spørsmålsbatterier og variabler er relevante for praksisfeltets utfordringer, og de favner de aktuelle og viktige dimensjonene vi ønsker å belyse.

Surveyen var utformet med 11 spørsmålsbatterier og 2 enkeltstående spørsmål, til sammen 99 variabler. Analyser av datafilene påviste at 3 variabler hadde over 50 prosent «Vet ikke», og disse ble tatt ut av analysegrunnlaget. De øvrige analysene ble foretatt med grunnlag i 96 variabler.

I tillegg hadde surveyen 32 spørsmål som inneholdt bakgrunnsvariabler. Svaralternativene i de fleste variablene var: hvor enig/uenig respondenten er i en påstand fra 1–4, hvor 1 = helt uenig og 4 er helt enig. I forhold til i hvor høy grad respondenten vurderer at påstanden stemmer, eller i hvor høy grad et fenomen/praksis og lignende er utbredt, var svaralternativene fra 1–3, hvor 1= i liten grad og 3= i høy grad. Respondentene fikk muligheten til å svare «Vet ikke» på de fleste påstandene. De to enkeltstående spørsmålene hadde svaralternativene «Nei», «Ja», «Vet ikke».

2.2.3 Utvikling av indikatorer

De enkelte spørsmålsbatteriene i surveyen er basert på en teoretisk forståelse av bestemte dimensjoner. Batteriene er imidlertid ikke validert. For å teste dimensjonaliteten til batteriene gjennomførte vi eksplorerende faktoranalyse for alle spørsmålsbatteriene samlet. Resultatet fra faktoranalysen (se vedlegg 3) er i høy grad samsvarende med det teoretiske utgangspunktet for utviklingen av surveyen og spørsmålsbatteriene. I tabellen nedenfor har vi angitt grupperingen av enkeltspørsmål ut fra faktoranalysen. Vi har også navngitt indikatorene med navn som kan sies å representere den dimensjonaliteten de respektive grupperinger av indikatorer kan sies å representere.

Tabell 3: Indikatorer og underliggende variabler

Indikator	Underliggende variabler
Vurderingspraksis	Jeg benytter undervisvurdering for å fremme elevenes videre læring
	Jeg benytter undervisvurdering for å kartlegge elevenes fremgang i fag
	Elevene mine får løpende tilbakemeldinger på arbeidet sitt
	Undervisvurdering er en integrert del av min undervisning
	Min vurderingspraksis når det gjelder halvårsvurdering, er god
	Min vurderingspraksis når det gjelder sluttvurdering, er god
	Mine elever får tilbakemeldinger undervis som gir dem mulighet til å forbedre arbeidet før endelig innlevering
Undervisvurdering	Å arbeide med undervisvurdering er en viktig del av læreryrket
	Undervisvurdering bør vise hva elevene kan
	Jeg synes undervisvurdering er et nyttig verktøy for meg som lærer
	Undervisvurdering bør støtte elevenes videre læring
	Undervisvurdering styrker lærerens muligheter til å tilpasse opplæringen til den enkelte elev
	Jeg synes det er interessant å jobbe med vurdering
	Det er viktig å involvere elevene i vurderingsarbeidet
Deltakerbaner	Ekstern veiledning, skolevurdering o.l.
	Lærerstuderenter som er i praksis
	Tolærersystemet
	Konferanser
	Resultater av internasjonale tester
	Faglitteratur, forskningsrapporter, fagtidsskrifter o.l.
Skolens vurderingskultur	I hvilken grad diskuteres det på skolen hva lærerne legger vekt på i vurderingen i fag?
	I hvilken grad diskuteres det på skolen på hvilken måte lærernes vurderingspraksis kan hjelpe elevene å nå målene?
	Vi har ofte samtaler om vurdering i lærerkollegiet
	På vår skole bruker vi vurderingsresultater systematisk for å utvikle vår vurderingspraksis
	Andre lærere ved egen skole
Elevinvolvering	I hvilken grad mener du at din vurdering bidrar til at elevene deltar i samtaler om egen læring og faglige utfordringer?
	I hvilken grad mener du at din vurdering bidrar til at elevene reflekterer over egen læring?
	I hvilken grad mener du at din vurdering bidrar til å involvere elevene i egen læring?
	I hvilken grad mener du at din vurdering bidrar til at elevene deltar i arbeidet med å utvikle mål for læringen i faget?
	I hvilken grad mener du at din vurdering bidrar til å øke elevenes kompetanse i fag?
	I hvilken grad mener du at din vurdering bidrar til at elevene deltar i arbeidet med å utvikle vurderingskriterier?
Lærerens tilbakemeldinger	I hvilken grad mener du at din vurdering bidrar til at elevene vet hva som er forventet av dem?
	I hvilken grad mener du at din vurdering bidrar til at elevene får tilbakemeldinger som forteller dem om kvaliteten på resultatet av arbeidet?
	I hvilken grad mener du at din vurdering bidrar til at elevene får råd om hvordan de skal forbedre seg?
	I hvilken grad mener du at din vurdering bidrar til at elevene forstår hva de skal lære i faget?

	I hvilken grad mener du at din vurdering bidrar til at elevene får tilbakemeldinger på den faglige forståelsen?
	I hvilken grad mener du at din vurdering bidrar til at elevene får tilbakemeldinger som forteller dem om kvaliteten på gjennomføringen av arbeidet?
Egen vurderingskompetanse	Jeg har nok kompetanse til å drive med sluttvurdering
	Jeg føler meg trygg på arbeid med vurdering
	Jeg har nok kompetanse til å drive med undervisvurdering
Skoleledelse	Skolens ledelse arbeider aktivt for kompetanseutvikling knyttet til vurdering
	Skolens ledelse bidrar til å utvikle en felles og god vurderingspraksis
	Skoleeier arbeider aktivt for kompetanseutvikling knyttet til vurdering
Krevende	Jeg synes det er krevende å arbeide med sluttvurdering
	Jeg synes det er krevende å arbeide med undervisvurdering
Planer	I mitt undervisningsfag er det en klar sammenheng mellom læringsmålene for faget og kjennetegn på måloppnåelse
	I mitt undervisningsfag er det en god sammenheng mellom læringsmålene for faget og elevens arbeidsoppgaver
Hindringer	Kravet om halvårsvurdering hindrer god undervisvurdering
	Kravet om sluttvurdering hindrer god undervisvurdering
Forskrift	Forskriftens føringer for undervisvurdering har ført til endringer i min vurderingspraksis
	Vurderingsforskriften er viktig for mitt daglige arbeid som lærer
Pådrivere	Er skoleleder en pådriver for å videreutvikle vurderingspraksis?
	Er skoleeier (kommunen/fylkeskommunen) en pådriver for å videreutvikle vurderingspraksis?
Elevenes læring	I hvilken grad mener du at tilbakemeldingene elevene får, bidrar til at de lærer bedre?
	I hvilken grad mener du at tilbakemeldingene elevene får, bidrar til at de når kompetansemålene?

De 14 indikatorene har til sammen 54 variabler, som forklarer 51,6 prosent av den totale variansen i materialet.

Vi har i den videre analysen holdt oss til den inndelingen av spørsmålene som faktoranalysen gir grunnlag for. Denne inndelingen er testet både gjennom en konfirmerende faktoranalyse (SEM-analyse) og med måling av indre konsistens (med Cronbachs alfa). Den konfirmerende faktoranalysen viser at modellen (indikatoren som et samlemål på de underliggende variablene) ikke avviker signifikant fra de empiriske data for noen av modellene.¹ Cronbachs alfa viser også for de fleste indikatorene gode verdier og for alle tilfredsstillende. Cronbachs alfa og signifikansnivå på testing av modellgenererte data (ved hjelp av en SEM-analyse) og empiriske data er gjengitt i tabellen nedenfor.

¹ SEM-metodikken, hvor SEM er en forkortelse for *Structural Equation Model*, er en probabilistisk analysemetodikk, hvor data generert fra en teoretisk modell (i vårt tilfelle indikatoren) kan testes mot de empiriske data. Testingen er en enkel khi-kvadrat-test av modellgenererte data (forventede) mot empiriske (observerte). Se for eksempel arbeidene til Rex Kline for en god introduksjon til SEM-analyser (Kline, 2011).

I tabell 4 presenteres verdier for den indre konsistensen for den avhengige samlevvariabelen «Vurderingspraksis» og for hver indikator i tabell 3.

Tabell 4: Indikatorenes reliabilitet og signifikansnivå på test etter SEM-analyse¹

Indikator	Antall variabler	Gyldige N ¹	Cronbachs alfa	Signifikansnivå på testmodell vs. empiriske data
Vurderingspraksis	7	633	0,690	0,000
Egen vurderingskompetanse	3	676	0,775	
Lærerens tilbakemelding	6	605	0,836	0,001
Skolens vurderingskultur	5	621	0,722	0,000
Krevende arbeid	2	674	0,811	
Elevinvolvering	6	606	0,781	0,000
Elevens læring	2	620	0,825	
Underveisvurdering	7	619	0,778	0,000
Skoleledelse	3	616	0,843	
Planer	2	650	0,729	
Forskrift	2	633	0,728	
Hindringer	2	640	0,843	
Deltakerbaner	6	612	0,806	0,001
Pådrivere	2	616	0,625	

¹ Antall svar på alle spørsmål (gyldige N) som inngår i indikatorene, og som Cronbachs alfa er beregnet ut fra. Merk at signifikanstesting lar seg gjennomføre bare hvor antall frihetsgrader i modellen er større enn frihetsgradene i det empiriske materialet. Indikatorer basert på færre enn fire items har vi dermed ikke kunnet teste.

Med utgangspunkt i disse variablene er det gjennomført både bivariante analyser (korrelasjon) og multivariate analyser (regresjon). I de multivariate analysene har vi behandlet indikatoren «Vurderingspraksis» som en avhengig variabel.

2.2.4 Oppsummering

Det kvantitative datamaterialet er representativt for Skole-Norge på både skole- og lærernivå med hensyn til de fire utvalgte fagene. Datamaterialet gir grunnlag for utvikling av 14 reliable indikatorer som på forskjellige måter sier noe om vurdering i skolen.

3.Hva innebærer validitet i vurdering?

Lise Vikan Sandvik

I dette kapitlet danner praksisfortellingen om lærer Anne rammen for kapitlet. I lys av Annes undervisning på ungdomstrinnet i fagene engelsk og matematikk rettes fokus mot sentrale problemstillinger knyttet til arbeidet med god vurderingspraksis, før dette settes inn i en teoretisk kontekst knyttet til validitet og reliabilitet i undervisning.

Anne er ansatt som lærer på en ungdomsskole som er med i satsingen Vurdering for læring, der hun i mange år har undervist i fagene engelsk og matematikk. Den første observasjonen av Anne og hennes undervisning er i en engelsktime på 10. trinn. Hun starter med å gi beskjeder om et essay som skal skrives to dager etterpå, og læreren deler ut oppgaveformuleringen til elevene i timen, slik at de kan forberede seg på temaet hjemme. De snakker sammen om hvor de kan hente informasjon til oppgaven, og læreren sier at all informasjon de kan finne, kan brukes i forberedelsene. Deretter følger en undervisningssekvens der læreren gjennomgår det grammatiske fenomenet passiv på tavla. Denne undervisningen har ingenting å gjøre med skriveoppgaven som elevene har fått utdelt, eller med arbeidet i timen som kommer etterpå. Hun avslutter timen med en lytteoppgave som de snakker litt om før timen er over. Lytteoppgaven har ingenting med tema for skriveoppgaven å gjøre. I løpet av timen er det stort sett læreren som prater engelsk. Elevene mottar informasjon om skriveoppgaven, de skriver ned regler for bruk av passiv i skriveboka si, og de lytter til teksten som leses opp. De snakker om lytteoppgaven på norsk.

Etter denne observasjonen stiller jeg meg noen spørsmål: Presenteres elevene for undervisningsmål? Det finnes ikke spor av noen i de undervisningsplanene som er knyttet til undervisningsøkten. Hun kommuniserer heller ingen klare mål for undervisningen i intervjuet knyttet til undervisningen som gjennomføres umiddelbart etter timens slutt. Finnes det vurderingskriterier knyttet til det skriftlige arbeidet? Det finnes ingen kriterier for skrivearbeidet og ingen føringer for hvordan disse forberedelsene skal foregå. I denne engelsktimen er det vanskelig å se sammenheng mellom læringsmålene i Kunnskapsløftet, den undervisningen som gis, og den vurderingen som skal foregå i form av en skriveoppgave. Ingen mål for verken selve undervisningstimen eller for skriveoppgaven presenteres eller drøftes. Grunnleggende ferdigheter er ikke tydelig integrert i arbeidsformene i klasserommet i den forstand at læreren legger opp til arbeidsoppgaver i timen som er knyttet opp til elevenes bruk av ferdigheten skrive, som de skal testes

i to dager senere. Ingen av begrepene som er knyttet til utvikling av skrivekompetanse i engelskfaget, er tydelig i arbeidsformene i denne timen. Det skapes ingen kontekst der det oppstår behov for målbevisst arbeid fra elevenes side med denne ferdigheten, som en del av arbeidet mot vurderingssituasjonen elevene skal ha.

Litt senere samme dag observeres Annes matematikkundervisning. I starten av timen kommuniseres hva de skal arbeide med, og hva de skal ha lært etter timen. De arbeider med konstruksjon, og hele timen dreier seg om problemløsning og dialogisk undervisning. Elevene definerer sentrale begreper gjennom oppdagende læring, og læreren tar i bruk ulike strategier for å få elevene til å forstå begreper og anvende kunnskapen. Læreren får for eksempel et par elever til å forklare hvordan de løste oppgaven, og bruker tid på en dialog omkring ulike strategier for å løse oppgaven, der flere elever forklarer sine fremgangsmåter. Det er elevene som oppsummerer hva de har arbeidet med og lært, når timen avsluttes.

Undervisningstimen kjennetegnes av dialog mellom lærer og elever, mange sekvenser der elevene utfordres til problemløsning både alene og i fellesskap. De utfordres til å formulere regler for det de har lært, og til å reflektere over egen læringsprosess. Denne matematikktimen har en tydelig kobling mellom læreplanmål, konkrete mål for timen, arbeidsformer og vurdering. Muntlighet er et gjennomgående trekk ved hele timen, og elevene presenterer, forklarer, drøfter, stiller spørsmål, gjør seg egne refleksjoner og samhandler i løpet av undervisningstimen. Dette er sentrale begreper for det å kunne uttrykke seg muntlig og skriftlig i matematikkfaget og er integrert i læreplanen for matematikk.

3.1 Individuell vurdering i skolen

Innen vurderingsforskningen er det vanlig å skille mellom *formativ* og *summativ* vurdering. Formativ vurdering har som mål å fremme elevens individuelle læring, mens den summative vurderingen måler og kontrollerer at de definerte læringsmålene er nådd (Black mfl., 2003). FIVIS-studien har spesielt fokus på hvordan vurdering stimulerer til læring, og undersøker hva som kjennetegner ulik klasseromspraksis når vurdering brukes som redskap for læring, spesielt knyttet til de fire prinsippene for underveisvurdering som er nedfelt i forskrift til opplæringslova, kapittel 3. Begrepene *formativ* og *summativ vurdering* brukes i skolesammenheng til å illustrere de ulike aspektene ved elevvurdering. Mange ulike fortolkninger av sammenhengen mellom de to begrepene har utviklet seg. Harlen (2004) peker på at det

viktigste for å forstå skillet mellom de to begrepene er erkjennelsen av at vurderingens formål er forskjellig:

The term «summative assessment» refers to an assessment with a particular purpose – that of providing a record of a student’s overall achievement in a specific area of learning at a certain time. It is the purpose that distinguishes it from assessment described as formative, diagnostic, or evaluative, rather than any particular method of gathering information about students’ performance (Harlen, 2004, s. 1).

Det kan synes enkelt å skille mellom disse to formålene, men for lærere i praksis kan det være svært vanskelig. FIVIS-studien viser i sine analyser at disse utfordringene er reelle. Særlig viser det seg at arbeidet med å gi gode tilbakemeldinger til elevene er en utfordring. I studier av effekten av tilbakemeldinger viser forskningen til Black og Wiliam (1998, 2009) at formativ kriteriebasert vurdering fører til bedre læring på alle utdanningsnivåer. Formativ kriteriebasert vurdering betyr å vurdere elevprestasjoner opp mot forhåndsbestemte kriterier og læringsmål. Black mfl. (2003) bruker kriteriebegrepet når de skriver om å praktisere vurdering for læring, og det brukes om både tilbakemelding fra lærer til elev, medelevvurdering og elevens egenvurdering. I en slik forståelse representerer kriterier et rammeverk som læreren bruker for å utvikle egnede kommentarer som gir informasjon til elevene om hva de har oppnådd, og hva de trenger å forbedre. I tillegg fungerer de som et rammeverk for å vurdere eget arbeid og for å strukturere eller planlegge kommende arbeidsoppgaver. Dette bidrar til en mer transparent læringsprosess for både elever og lærer. Black, Harrison, Lee, Marshall og Wiliam (2003) framhever også motivasjonsaspektet ved en slik kriteriebasert vurdering. Elevene utvikler med andre ord vurderingskompetanse ved at de lærer å anvende kriterier på sitt eget og andres arbeid.

Denne felles forståelsen mellom elever og lærer av grunnlaget for vurderingen kalles ofte tolkningsfellesskap. Kriteriebasert, målorientert vurdering som skaper et tolkningsfellesskap, og som gir transparente læringsprosesser, er sentrale perspektiver i denne studien. En slik forståelse av vurdering bringer inn betydningen av meningsfulle vurderingsoppgaver som oppleves som viktige for elevene (Wiggins, 1998). Wiggins og McTighe (2005) framhever at formativ vurdering kan sees som et system for læring, der autentiske læringssituasjoner og det han kaller *baklengs planlegging*, er sentrale elementer i et slikt system. Med baklengs planlegging forstås en annerledes tankegang enn den tradisjonelle undervisningspraksisen, der læreren planlegger hva et undervisningsopplegg skal inneholde av aktiviteter, og lager prøver i etterkant av undervisningsopplegget. Baklengs planlegging handler om å snu dette på hodet, der målene for undervisningen identifiseres først. Deretter avgjøres hvordan elevene skal dokumentere læringsutbyttet, og hvordan dette skal

vrderes, før selve undervisningen planlegges (Wiggins & McTighe, 2005). Et annet viktig prinsipp med en slik tankegang er å eksemplifisere for elevene hva som kan betraktes som høy måloppnåelse innen de kompetansemålene som skal nås. En måte å gjøre det på er å vise fram eksempler på *best practice* og diskutere disse med elevene (Rust, O'Donovan & Price, 2006).

Særlig har Sadlers forskning (1989, 1998) hatt innflytelse på dette forskningsfeltet. Han har identifisert tre nødvendige betingelser for at elever skal ha nytte av tilbakemeldinger. Han argumenterer for at eleven må vite (1) hva som kjennetegner god praksis (dvs. å ha forståelse av læringsmålene), (2) hvordan hans prestasjon her og nå står i forhold til målene, og (3) hva han skal gjøre for å minske avstanden mellom de ferdighetene han innehar, og de målene han skal nå. For at eleven skal være i stand til å arbeide på denne måten, må han, ifølge Sadler, evne å vurdere seg selv. Å få til en slik læringsfremmende undervisning krever en lærer som har profesjonsfaglig kompetanse til å vite hvordan hun skal fremme de metakognitive ferdighetene som egenvurdering krever. Hattie og Timperley (2007) har gjennomført en metastudie om tilbakemeldinger og hvilken type tilbakemeldinger som er mest effektiv. De peker på at den som skal gi tilbakemelding, må svare på tre svært sentrale spørsmål, som er svært like Sadlers betingelser for god tilbakemelding: Hvor skal jeg? Hvor er jeg? Hvordan skal jeg gå videre? De bruker begrepene *feed up*, *feed back* og *feed forward* for å beskrive den gode tilbakemeldingsstrategien. Dette kalles i dagens norske vurderingsdebatt ofte framovermeldinger.

Formativ vurdering har fått en dynamisk tilnærming kalt *vurdering for læring*, som er en interaktiv, læringsfokusert tilnærming til vurdering. Prinsippet vurdering for læring handler om å få elevene til å bli aktivt deltakende i egen læringsprosess, til å forstå læringsmålene og til å gjennomføre egenvurdering og vurdering av andres ferdigheter. For lærernes vedkommende handler det om å oppmuntre elevene til å delta i vurderingsarbeidet, om å lære dem å forstå læringsmålene, om å gi tilbakemeldinger som er læringsfremmende, og å planlegge undervisningen ut fra læringsmålene (Black & Wiliam, 1998, 2005, 2006, 2009; O'Donovan, Rust, Price & Carrol, 2006).

Dermed kan det sies å være tre nøkkelprosesser i all undervisning, som også er kjernen i god vurderingspraksis: å vite hvor den lærende er i sin læring, å lage mål for læringen og å finne ut hva som skal til for å nå målene (Sadler, 1989; Wiliam & Thomson, 2007). Tradisjonelt sett har læreren hatt ansvaret for disse prosessene, men i et vurderingsteoretisk perspektiv er det viktig å tydeliggjøre hvilken rolle elevene selv har i sin egen læringsprosess, og hvilken betydning de har for hverandres læring i fellesskapet (Sadler, 1998). Slik kan man si at læreren er ansvarlig for å designe og implementere et effektivt læringsmiljø, mens den lærende har ansvar for læringen i dette miljøet (Black & Wiliam,

2009). Ved å koble de tre prosessene sammen med de ulike aktørene i klasserommet kan formativ vurdering sees som et konsept bestående av fem strategier, som vist i figur 1.

	Målsetting	Planlegging	Diagnostisering
Lærer	1. Utarbeide tydelige mål og vurderingskriterier	2. Utvikle effektive klasseromsdiskusjoner og gode vurderingsoppgaver for dokumentasjon av elevenes ferdigheter	3. Gi tilbakemelding som fremmer læring
Medelever	Tolkningsfellesskap for læringsmål og vurderingskriterier	4. Aktivere elevene som ressurser for hverandres læring	
Elev	Forstå læringsmål og vurderingskriterier	5 Aktivere elevenes evne til å lære å lære	

Figur 1. Aspekter ved formativ vurdering (William & Thompson, 2007).

Det er ulike klasseromsaktiviteter og undervisningsprinsipper som kan tas i bruk for å implementere disse fem strategiene. For de to første punktene vil det å lage gode undervisningsopplegg som motiverer eleven, det å stille gode spørsmål og det å ha god forståelse av elevenes læringsutbytte være viktige elementer for å lykkes med den individuelle vurderingen. En slik tilnærming og forståelse av vurdering understreker at den er «far more than just a summative collection of information about learners' achievements: it is also a vital part of the learning process and impact on the course of pupils' learning» (Harlen & Winter, 2004, s. 391). Strategiene i Williams og Thomsons oversikt er utgangspunktet for FIVIS-studiens undersøkelser av vurderingspraksiser i norske klasserom, slik vi også kjenner dem igjen i de fire prinsippene for underveivurdering som er nedfelt i forskrift til opplæringslova, kapittel 3.

Med denne forståelsen av individuell vurdering som bakteppe vil neste delkapittel drøfte begrepet *validitet i vurdering* – eller hvordan man skaffer gode beviser for elevens læring, hvordan disse bevisene fortolkes, for i neste omgang å kunne ta gode beslutninger om tilpassingen av det videre læringsforløpet.

3.2 Validitet i klasserommet

Annes undervisning og vurderingspraksis, som ble presentert innledningsvis i dette kapitlet, danner utgangspunkt for å forstå validitetsbegrepet, som er et sentralt begrep i FIVIS-forskningen. Drøftingen av Annes undervisning tas så videre ved å se på funn knyttet til validitet i læreplanarbeid og i overgangen mellom trinnene.

Hvordan kan det ha seg at en lærer har så ulike praksiser i arbeidet med vurdering og de grunnleggende ferdighetene muntlighet og skriving i sin undervisning? Anne kan fortelle at det er matematikk som er «hennes fag». Det er der hun har den faglige fordypningen, og det er der hun har tatt videreutdanning. Dette har hatt stor betydning for hennes forståelse av arbeidet med grunnleggende ferdigheter i faget og for hennes utvikling av god vurderingspraksis. Engelskfaget er noe hun ikke føler hun behersker i like høy grad, noe som fører til at hun i høy grad støtter seg til læreboka og de oppgavene som presenteres der. Hun har heller ikke stor kunnskap om læreplanen i engelsk og forståelse av de ulike læreplanmålene. Dette innebærer at undervisningen som foregår, ikke er like «gjennomtenkt», som hun sier i intervjuet. Dette viser at lærerens kunnskaper og forståelser i fag og for fagenes egenart har betydning for hvilket læringssyn læreren utvikler, noe som igjen påvirker lærerens undervisningspraksis.

Undersøkelser fra matematikkfaget viser mange praksiser der nettopp den dialogiske instruksjonen er det bærende prinsippet (Sandvik & Buland, 2013; Svenkerud, Klette og Hertzberg, 2012). Dette er samtaler som utdyper og skaper innsikt i det faglige innholdet. I matematikkundervisningen som er observert i FIVIS (Sandvik & Buland, 2013), er det tydelig at lærere som lykkes i å gjøre helklassesamtalen til en engasjerende faglig dialog der læreren stiller utforskende spørsmål, anerkjenner elevenes svar, gjentar elevenes svar for å forsterke og der svar nesten alltid får oppfølgingsspørsmål, også har elever i klasserommet som opplever undervisningen som motiverende. Slik undervisning orienterer seg mot enkeltelevens forståelse og mot det mer dialogiske, samhandlende perspektivet. Disse lærerne ser ut til å kunne etterspørre valide bevis for mestring og forståelse parallelt med elevens læringsprosess.

Når det gjelder muntlighet i engelskfaget, viser undersøkelser mange utfordringer knyttet til validitet i undervisningen (Sandvik & Buland, 2013). Læringsprosessene for barnetrinnet består i høy grad i at elevene lytter til en tekst i læreboka, leser teksten i kor, svarer på noen lukkede spørsmål til teksten og oversetter og skriver glosene til stykket i gloseboka. I klasseromsdialogene på ungdomstrinnet får elevene tilbakemeldinger som omfatter bekreftelser eller korrigerende av svar de avgir, og av høytlesning i klassen. Det ble i liten grad

observert tilbakemeldinger som utfordrer elevenes refleksjon.

Undervisningspraksiser som beskrevet i Annes engelsktime kan knyttes til validitetsbrister i opplæringen. Lærernes kompetanse når det gjelder metoder for utvikling av muntlighet, kan beskrives som mangelfull i flere tilfeller. Det ligger en utfordring i å kunne etablere praksiser i klasserommet der dialog og samhandling med andre kan bidra til utvikling av og læring i muntlige ferdigheter i større grad.

Dette gjelder også for skriveopplæringen. Å utvikle skriveferdigheter i alle fag er en forutsetning for livslang læring og for å kunne delta aktivt i samfunnslivet på en kritisk og reflektert måte. I forskning omkring grunnleggende ferdigheter og skriveopplæring kan det synes som den systematiske opplæringen avtar etter at den første skriveopplæringen er avsluttet – altså når elevene behersker de såkalte basisferdighetene (Aasen, Møller, Rye, Ottesen, Prøitz, Hertzberg, 2012; Sandvik & Buland, 2013). Det virker som om lærere på barnetrinnet i større grad har innarbeidet en læringsfremmende vurderingspraksis når det gjelder grunnleggende ferdigheter, enn hva som er tilfellet på ungdomstrinnet og i videregående skole. Eksemplet med Annes undervisning i engelsk tyder også på det. Det er manglende kulturer for eksplisitt arbeid med planlegging og strukturering av skrivearbeidet og med revidering og bearbeiding av tekster. Det synes også som det er et lite fokus på mottakerorientering i skrivearbeidet, der elevene lærer å tilpasse tekstene sine til mottaker, innhold og formål.

Vurdering har flere hensikter: å dokumentere læring, å forbedre undervisning og å fremme læring. Læreren skal i en vurderingssituasjon samle bevis for elevens læring og fortolke disse, for igjen å skape en undervisning som videreutvikler elevenes kompetanse. For en lærer i engelsk handler det om å skape situasjoner i undervisningen der elevene får utviklet sin kompetanse i for eksempel skriving. Denne kompetansen skal så uttrykkes i en eller annen relevant skriveoppgave og vurderes av læreren. Denne rekken av fortolkninger og operasjonaliseringer fra læreplanmål til vurdering kan sees som en lang kjede av forsøk på å fortolke og skape mening av læreplanmål, kriterier, lærebøker, arbeidsoppgaver, elevens prestasjoner og konsekvenser av disse.

Diskusjoner om validitet kan sies å være det viktigste momentet i enhver vurderingssituasjon, enten det er en nasjonal prøve eller en kort, uformell test i klasserommet. Hvis en lærer mangler gode beviser for elevens læring, eller hvis læreren mangler de nødvendige forutsetningene for å fortolke bevisene, vil det være vanskelig å ta gode beslutninger om tilpassingen av det videre læringsforløpet. Med andre ord: En lærer som ikke kan samle inn god informasjon, fortolke den og tilrettelegge for elevens fortsatte utvikling, risikerer at læringsmålene ikke blir realisert på best mulig måte for den enkelte elev. Annes arbeid i engelskundervisningen kan tyde på at sammenhengen

mellom læreplanmål, arbeidsoppgaver, vurderingsoppgaver og elevers prestasjoner kan utvikles for å gjøre undervisningen i faget mer valid.

Læreplanmålene i fagene er et resultat av en komplisert prosess og vil alltid være et kompromiss mellom forskjellige faglige ståsteder. Lærerne fortolker disse målene basert på sin utdanning og sine erfaringer og omsetter dem i en undervisning bestående av læremidler, aktiviteter, mål og vurderingsformer. Validitet i undervisning og vurdering handler derfor ikke bare om hvorvidt prøveresultatene er gyldige. Det handler også om den lange kjeden av fortolkninger og mulige misforståelser som kan oppstå når kulepunktene i Kunnskapsløftet skal omsettes til lærerens klasseromspraksis og elevenes læring. Det at lærerne ikke har kunnskap nok om bakgrunnen for grunnleggende ferdigheter og en forståelse av hvordan det kan operasjonaliseres i klasserommet, kan være noen av årsakene til at grunnleggende ferdigheter ikke har blitt integrert i alle fagene i tråd med intensjonene i Kunnskapsløftet. Ved å la validitet i undervisning og vurdering være et bærende prinsipp for alt arbeid med grunnleggende ferdigheter vil det også kunne føre til en bedre sammenheng mellom læreplanens mål og hvordan læreren underviser i faget.

Til tross for at validitet regnes som et helt avgjørende element i vurdering, og dermed også undervisning, er det likevel ikke avklart i teoretisk forstand hvorvidt det skal tas hensyn til faktorer som ligger utenfor selve vurderingssituasjonen. Knyttet til eksempelet med lærer Anne er det selve klasseromskonteksten som drøftes, og hvilke vurderingspraksiser knyttet til grunnleggende ferdigheter som kommer til uttrykk. Nettopp fordi det er denne konteksten som er hovedanliggendet, og når denne konteksten kan beskrives med et bredt spekter av vurderingsoppgaver og -situasjoner, er en fortolkende tilnærming til validitet en god måte å forstå begrepet på (Bonner, 2013). Studier med en slik tilnærming (Kane, 2006; Moss, 2003) fremhever hvordan validitetsargumentet hos lærere i klasserommet er forankret i en kvalitativ, selvreflekterende prosess som igjen er forankret i lærerens læringssyn. Dette læringssynet er påvirket av lærerens utdanning og erfaringer, skolens undervisnings- og vurderingskultur og klasseromkulturen. Hvis det er en svak tilpasning mellom undervisningen og vurderingen, kan det ha negativ betydning for elevens motivasjon og læringsmiljøet i klasserommet.

Å undervise etter prinsippet om en læringsfremmende og valid vurderingspraksis betyr at vurderingspraksiser blir tettere integrert i lærernes daglige undervisningspraksis. En slik praksis betyr at vurdering i større grad blir å anse som et prosessuelt fenomen og i mindre grad som en avgjørende enkelthendelse i skolen. Derfor er det viktig å ha en kontekstuell forankret forståelse av samspillet mellom læreplanmål, faglig innhold, lærerens verdissyn og de tausede praksiser som gjerne er en del av faglige tradisjoner i skolen. Med

en slik forståelse av sammenhengen mellom undervisning og vurdering vil det også være lettere for lærere å se hvordan grunnleggende ferdigheter kan integreres i fagene, med en progresjon på fagenes premisser.

I FIVIS-prosjektet forstås *validitet i vurdering* som et dynamisk begrep og anses som relevant for hele den komplekse prosessen fra læreren leser kompetansemålet, til elevens produkter og prestasjoner blir vurdert av læreren. Hovedargumentet for dette er at vi ikke har analysert kvaliteten i enkeltstående tester og prøvesituasjoner, men heller undersøkt hvordan vurderingspraksiser i norsk skole preges av både lærernes læreplanforståelse, tidligere praksiser, fagdidaktiske syn og andre faktorer som ikke direkte berører konkrete vurderingssituasjoner. Vi er interessert i de kontekstuelle og reaktive spørsmålene om aktørenes operasjonalisering og anvendelse av vurderingsresultater. Lærere gjør fortolkninger, trekker slutninger, utformer oppgaver og forsøker å skape en meningsfylt forståelse av elevens kompetanse. Samtidig vil også eleven være en aktiv fortolker av lærerens operasjonalisering slik de kommer til syne i undervisningen, og vedkommende vil forsøke å skape mening i faget slik det blir formidlet, kommunisert og representert gjennom undervisning, læremidler og andre læringssituasjoner.

Dermed kan validitet og verdier i vurdering sees som to aspekter av samme sak, og de sosiale og etiske sidene av vurderingen er like viktige som de mer tekniske aspektene av selve vurderingssituasjonen (Bonnier, 2013; Messick, 1995). Til tross for de metodiske utfordringene som ligger i et slikt perspektiv, har vi forfulgt et slikt syn på vurdering i vårt prosjekt. Det er to hovedårsaker til dette. For det første kan den nasjonale satsingen på vurdering for læring bety at vurderingspraksiser blir tettere integrert i lærernes daglige undervisningspraksis og det blir vanskeligere enn før å skille vurderingssituasjoner fra læringssituasjoner. En slik praksis vil bety at vurdering i større grad blir å anse som et prosessuelt fenomen og i mindre grad som en avgjørende enkelthendelse i skolen. Da er det viktig å ha en kontekstuell forankret forståelse av samspillet mellom læreplanmål, faglig innhold, lærerens verdisyn og de tause praksiser som gjerne er en del av faglige tradisjoner i skolen.

For det andre er vurderingspraksiser sterkt knyttet til styringsmekanismer i skolen. Det betyr at vurdering ikke utelukkende dreier seg om undervisnings- og læringsprosesser i klasserommet, men at det også er et av de områdene der balansen mellom profesjonell autonomi og ekstern styring vil variere. En av intensjonene med Kunnskapsløftet som læreplanreform var å tydeliggjøre ansvaret som ligger på de forskjellige nivåene og aktørene i skolen. Ifølge rapporten *Kunnskapsløftet som styringsreform – et løft eller et løfte?* (Aasen mfl., 2012) skulle Kunnskapsløftet som styringsreform øke desentraliseringen

og gi større handlingsrom og frihet til skoleeier og skoleledere til å drive egen opplæringspraksis. Dette skulle være et ledd i ansvarliggjøring av profesjonen. Rapporten peker imidlertid på at den statlige styringen i økende grad har blitt skjerpet, og at dette også viser seg i satsingen på individuell vurdering. Elevvurdering skal ikke lenger oppleves som et individuelt ansvar, men som et kollektivt ansvar på flere nivå i skoleforvaltningen (Aasen mfl., 2012). Balansen mellom individuelle læreres vurderingspraksiser, utviklingen av organisasjonsforankrede praksiser og utdanningsmyndighetene er altså endret, og det er et viktig argument for å forstå validitet i vurdering som del av et større nettverk av fortolkninger og forståelser.

En god måte å visualisere dette komplekse nettverket av fortolkninger på er å anse validitet som en lang kjede av forsøk på å fortolke og skape mening av læreplanmål, tause fagfellesskap, elevers prestasjoner og så videre. I vurderingslitteraturen har dette blitt referert til som en validitetskjede (Crooks, Kane & Cohen, 1996). Etter vårt syn er dette et fruktbart begrep som kan belyse det komplekse forholdet mellom læreplanens iboende paradokser, skoleeieres forståelse av læreplanreformen, skoleledernes rolleforståelse, lærernes operasjonaliseringer og elevenes handlinger. Vi vil også umiddelbart tilføye at dette ikke er snakk om et perfekt system som gradvis forringes jo lenger man beveger seg bort fra læreplanmålene. Fordi læreplanmål fortolkes inn i allerede eksisterende faglige strukturer, praksiser og forståelser, kan det like gjerne hende at validitetskjeden tilføres noe av verdi i de ulike leddene. For å kunne belyse dette nærmere er det imidlertid nødvendig å forstå vurdering i fagspesifikke kontekster.

I norsk sammenheng er det faglige perspektivet på vurdering viktig fordi læreplanstrukturen i Kunnskapsløftet til dels er svært annerledes enn tidligere planer. Vekten på kompetansemål og observerbare endringer i elevers kunnskaper, ferdigheter og forståelser krever at lærerne behersker læreplananalyse og operasjonalisering i større grad enn tidligere læreplaner. Denne kompetansen krever lærere med fagkompetanse som forstår vurdering ut fra fagenes premisser. Dermed blir også det fagdidaktiske perspektivet sentralt i et vurderingsperspektiv. Dette understrekes i blant annet studien til Dansk Clearinghouse, som i 2008 foretok en systematisk gjennomgang av empirisk forskning på lærerkompetanse og elevenes læring (Nordenbo, Larsen, Tiftikçi, Wendt & Østergaard, 2008). Også fagdidaktisk forskning understreker lærerens fagkompetanse og fagdidaktiske kompetanse som sentralt i arbeidet med vurdering (Tellefsen & Martinussen, 2011; Sandvik, 2011). Eksempelet med Anne som ble presentert innledningsvis, kan anskueliggjøre dette. Det kan synes som om hun ikke har god nok fagdidaktisk kompetanse til å legge til rette for elevenes utvikling av skrivekompetanse i engelsk. Dette fører til tilfeldige skriveoppgaver og en skriveopplæring i engelsk som ikke er fortolket

inn i fagets strukturer. Slik oppstår det brister i validitetskjeden – mellom læreplanmål, arbeidsoppgaver, aktiviteter og vurdering av elevenes prestasjoner.

3.3 Validitet i lokalt læreplanarbeid

Annes klasseromsdialog i matematikkundervisningen er et eksempel på undervisning som motiverer og utfordrer og er orientert mot enkeltelevens forståelse. Dette er et eksempel på god praksis i det daglige vurderingsarbeidet. Det er det kontinuerlige, daglige vurderingsarbeidet som totalt sett har størst betydning for elevenes læring, fordi tilbakemeldingene elevene får, bør komme mens den lærende er i en prosess (Black & Wiliam, 1998; Hartberg, Dobson & Gran, 2012, s. 35). Likevel viser studier at dette ikke er like enkelt å gjennomføre i praksis, fordi det foregår læring i ulike tidsspenn (Black & Wiliam, 2005; Dysthe, 2008; Engh mfl., 2012). Arbeid med vurdering i ulike tidsspenn realiseres gjennom lærerens undervisning og gjennom lærerens og skolens planarbeid.

Analysen i FIVIS' delrapport 1 (Sandvik mfl., 2012) antyder at lærerne ikke har forholdt seg til langsiktige mål i arbeidet med årsplaner, men at de har kommet relativt langt med arbeidet med å operasjonalisere mål på kort sikt. Vi har skilt mellom tre ulike perspektiv i planleggingen: et kort perspektiv, som strekker seg over noen uker, et mellomlangt perspektiv, som strekker seg fra et halvt til et helt skoleår, og et langt perspektiv, som strekker seg til sluttvurdering og studiespesialisering.

Det kan synes som om planarbeidet i det korte tidsspennet er forholdsvis tilfeldig, med et fravær av mønster og felles begrepsbruk. Våre analyser i FIVIS' delrapport 2 (Sandvik & Buland, 2013) kan tyde på at det utvikles kulturer og subkulturer for planarbeidet ved skolene. Vi har også funnet eksempel på at det ikke utvikles arbeidsplaner hos enkeltlærere. I skoler som har deltatt i den nasjonale satsingen på vurdering for læring, er planene utarbeidet i et tolkningsfellesskap på skolen eller i kommunen, og de gjelder for et skoleår eller for et mellomlangt til et langt tidsspenn. Våre observasjoner understøtter oppfatningen av at for eksempel kroppsøvlingslærerens ukeplan og plan for hver time blir tilpasset den lokale konteksten og rammefaktorer som man i liten grad kan forutse når årsplanen blir laget. For skoler som ikke har vært med i satsingen på vurdering for læring, er det i mindre grad samarbeid på skolen om planer i faget.

Det synes å være et misforhold mellom hva som forventes av lærerne gjennom forskrifter og plandokumenter, og hvordan lærerne fortolker og operasjonaliserer dette i arbeidsplaner og undervisning. Analyser av

intervjuene og resultater fra breddeundersøkelsen indikerer at lærerne har en forståelse av hvilken betydning klare mål og kriterier har for elevenes læring, og at de bruker mye tid på dette arbeidet, både i forkant av undervisning og i selve undervisningen. De involverer også elevene i dette arbeidet til en viss grad. Likevel kommer dette arbeidet ikke alltid fram gjennom analyser av lokale læreplaner. Det kan synes som om arbeidet med vurdering er implisitt kunnskap som noen lærere oppfatter som lite relevant å kommunisere til andre i praksisfellesskapet gjennom arbeidsplaner. Utfordringen i skolen ligger i å løfte alle planer opp på et nivå som kan gi elevene like forutsetninger for å utvikle sin kompetanse og sin evne til selvorganisert læring. Dette vil kunne løfte arbeidet med vurdering i det lange tidsspennet, som handler om å sørge for at helhetlige og langsiktige mål ikke glemmes, og at elevene øves i å utvikle kvalitetsbevissthet (Hartberg, Dobson & Gran, 2012).

Generelt viser funn at skoler som har deltatt i satsingen Vurdering for læring, sitter med flere års erfaringer med gode prosesser i et slikt arbeid. Dette arbeidet har ført til grunnleggende endringer i vurderingspraksiser og har slik bidratt til å styrke undervisningen og vurderingspraksisene i faget. Disse skolene har gjennomarbeidet planverket og brukt mye tid på å utarbeide lokale planer og kriterier. Gjennom slike prosesser har lærerne bidratt til å sikre en lang rekke med kjeder av validitet fra læreplan til undervisningsplanlegging.

Det ser ut til at skoler som har satset på vurdering for læring, har en sterkere kultur for operasjonalisering av kompetansemål, bevissthet rundt de ulike delkompetansene som elevene skal tilegne seg, og i større grad er i stand til å bruke vurderingspraksiser i ulike tidsspenn for å følge opp elevenes læringsprosesser. Dette kommer tydeligst til uttrykk på barnetrinnet. Det er imidlertid vanskelig å si om dette betyr at satsingen har hatt effekt, eller om årsaken er at skolen allerede fra før har hatt en kultur for felles utviklingsarbeid, lokalt læreplanarbeid, felles refleksjon og praksisdeling.

Forskere fremhever validitet og verdier i vurdering som to sider av samme sak (Bonnier, 2013; Messick, 1995). De sosiale og etiske sidene ved vurdering er like viktige som de mer tekniske aspektene ved selve vurderingssituasjonen. Dette synes som relevante perspektiver også i denne studien. Den nasjonale satsingen på Vurdering for læring betyr at vurderingspraksiser blir tettere integrert i lærernes daglige undervisningspraksis, og det blir vanskeligere enn før å skille vurderingssituasjoner fra læringssituasjoner.

FIVIS-prosjektet identifiserer et mangfold av praksiser på både individuelt og kollektivt nivå. Nettopp verdiperspektivet, det prosessuelle og sosiale aspektet ved vurdering, synes i mange skoler å ha kommet for kort. Det fremstår som en utfordring å utvikle praksiser som gjenspeiler et syn på læring og vurdering

som læringsfremmende og inkluderende, der tilpasset undervisning bidrar til økt motivasjon og læring.

Balansen mellom individuelle læreres vurderingspraksiser, utviklingen av organisasjonsforankrede praksiser og utdanningsmyndighetene er altså endret, og det er et viktig argument for å forstå validitet i vurdering som del av et større nettverk av fortolkninger og forståelser, der *kontekst* og *verdier* vil være sentrale begreper for å validere vurderingspraksiser på alle nivå. Denne tilknytningen til styringsmekanismer fremstår også tydelig i denne studien. Særlig er lærere og ledere i videregående skole opptatt av de organisasjonsforankrede praksisene som handler om å dokumentere vurdering. Forståelsen av at vurdering i høy grad handler om å ha dokumenterte vurderingssamtaler og å ha nok vurderingsgrunnlag, kan skygge for en forståelse av vurdering som en læringsfremmende prosess som foregår i det daglige, i dialogen med hver enkelt.

3.4 Validitet i overgang mellom trinn

Anne underviser på en 1–10-skole. Vi har undersøkt praksiser i de ulike skoleslagene, og de skoleslagene som synes å ha flest utfordringer knyttet til vurdering, er 1–10-skolene og ungdomstrinnet. Funn viser at skoleledere ved 1–10-skoler synes det er en stor utfordring å samle kollegiet om en kollektiv forståelse av hva som er god vurderingspraksis, nødvendig vurderingskompetanse, og hva som kjennetegner en vurderingskultur bygd på et tolkningsfellesskap.

For ungdomstrinnet sin del spriker vurderingspraksisene i høy grad. Forskjellen her synes å være knyttet til tre faktorer: (1) Lærere har ikke blitt innført i vurdering, (2) lærere har liten kompetanse i faget, (3) ledelsen har ikke vært pådriver for utviklingsprosessen.

Særlig synes engelskfaget å stå i et spenningsforhold mellom pragmatiske løsninger for å få undervisningskabalene til å gå opp og faglig ansvar. I flere tilfeller møter vi lærere på 10. trinn som ikke har undervisningskompetanse til å undervise i engelsk, men som likevel blir satt til denne oppgaven fordi de har et annet språkfag i sin fagkrets. Å kunne drive læreplananalyse og operasjonalisering av denne krever lærere med fagkompetanse som forstår vurdering ut fra fagenes premisser (Nordenbo mfl., 2008). Både analyser av lokalt læreplanarbeid og undervisningspraksiser viser at variasjonene er store på dette området.

3.5 Reliabilitet i klasserommet og på nasjonalt nivå

I FIVIS-studien har vi vært opptatt av å undersøke vurderingsspraksiser på klasseromsnivå. Samtidig vet vi at reliabilitetsproblematikk er viktig på et nasjonalt nivå for både elever, lærere, skoleledere og ikke minst skoleeiere. Utvikling av et godt fortolkningsfellesskap som kan fremme validitet i alle ledd av undervisning og vurdering, er viktig.

Reliabilitetsproblematikk er kjent for skoleeiere gjennom blant annet klagesaker og graden av samsvar mellom standpunkt- og eksamenskarakterer. Validitetsproblematikk er imidlertid delegert ned til skole-, nettverks- og individnivå. Når en skoleeier ytrer håp om «større standardisering» i vurderingen, kan dette først og fremst forstås som et ønske om bedre rutiner og praksiser for å sikre vurderersamsvar. Samtidig skal vurderingsinformasjon og karakterstatistikk tjene flere funksjoner: Sammenligninger av resultater brukes i policyavgjørelser mellom skoleeiere og politikere, og skoleeierne hevder også at lærerne opplever det som «ganske nyttig» å bli informert om utviklingen i karakterresultater over tid. Slik informasjon kan neppe anvendes til å forbedre læringen til elevene, men skoleeierne hevder altså likevel at slik informasjon er et grunnlag for videre beslutninger om skolen.

Reliabilitet kan forstås gjennom begrepene *intervurderingsreliabilitet*, at kjennetegn på måloppnåelse ved fastsettelse av standpunkt- og eksamenskarakter er i overensstemmelse med andre lærere, skoler, kommuner eller fylkeskommuner som vurderer elever i det samme faget, og *intravurderingsreliabilitet*, som handler om lærerens egen stabilitet i vurderinger av elever i fag over tid (Erickson, 2012).

Backmann og Sivesind (2012) peker på et paradoks i Kunnskapsløftet: På den ene siden formuleres og implementeres mål, på den andre siden venter kontroll og ansvarliggjøring. Det vil si at et målprogram krever at målene følges opp og evalueres. Samtidig som målene må formuleres for å kunne evalueres, skal de være tilstrekkelig vage til at lokale nivå kan få et eierforhold til målene gjennom ulik tolkning og ulik konkretisering. Dette kan oppleves som et paradoks hvis systemet forstås som grunnlagt på mål som skal forene uforenlige krav; de skal være diffuse/generelle og evaluerbare på samme tid (Backmann & Sivesind, 2012, s. 251).

Spørsmålet om reliabilitet belyser dermed noen utfordrende sider av kvaliteten på skolens arbeid med lokale læreplaner og elevenes opplevelse av en rettferdig vurdering. Disse utfordringene skal løses på lokalt nivå. Spørsmålet er om oppgaven er overkommelig når systemet innebærer en styring gjennom ansvarliggjøring uten tydelige mål og vurderingsresultatene skal vise kvaliteten på prosessene. Våre analyser i delrapport 1 og 2 antyder en utfordring knyttet

til denne problematikken: Arbeidet lærerne gjør med planer i denne studien, viser ikke en felles forståelse av betydningen av å synliggjøre vurderingsprosessene gjennom blant annet arbeid med lokale læreplaner. Det er få spor etter disse beslutningsprosessene – fra skolens årsplaner til de enkelte trinnenes og klassenes ukeplaner og periodeplaner.

Vurderingsformene i klasserommet varierer fra den ene ytterligheten til den andre, fra enkle gløseprøver og utfyllingsoppgaver via kapitellprøver og temaprøver til halvårsprøver og årsprøver. Noen av prøvene har læreren utformet og gjennomfører, mens i andre tilfeller vurderer elevene seg selv eller sine medelever. Denne store spennvidden av prøvesituasjoner både krever og tillater varierende konseptuelle og metodologiske tilnærminger til reliabilitet. Vurderingen vil påvirke undervisningens form og innhold. McMillan (2003) understreker dette forholdet og sier at all klasseromsaktivitet handler om å legge til rette for situasjoner der læreren kan samle bevis for elevenes læring, fortolke denne informasjonen og ta beslutninger om undervisningen ut fra dette. I et reliabilitetsperspektiv vil det så handle om å vurdere kvaliteten på metoden læreren har brukt for å samle inn informasjon, og å vurdere kvaliteten på informasjonen som er samlet inn. Det handler også om å vurdere om det er nok informasjon av tilstrekkelig kvalitet for å ta beslutninger om undervisning (Smith, 2003). Denne store spennvidden innen prøvesituasjoner krever og tillater med andre ord varierende konseptuelle og metodologiske tilnærminger til reliabilitet (Parkes, 2013). Parkes hevder videre at med det dynamiske og utviklende læringssynet som preger vurderingsdiskursen, handler ikke reliabilitet bare om informasjon som samles inn for å ta beslutninger, men om hele beslutningsprosessen (Parkes, 2013). Slik utfordres det tradisjonelle reliabilitetsbegrepet til å omhandle elevkunnskap og kjennskap til konteksten vurderingen blir gjort i. På den måten blir prosessen både interaktiv og samarbeidende (Moss, 1994, 2004).

Disse ulike forståelsene viser kompleksiteten i debatten omkring reliabilitet i læringsfremmende vurdering. Men hva vet vi så om læreres kunnskap om og forståelser av reliabilitet? Forskning viser at lærere har liten kunnskap om fundamentale konsepter for måling av kompetanse (Campbell & Evans, 2000). Delrapport 1 og 2 viste også at de ulike aktørene forsto reliabilitet i tradisjonell forstand som måling og vurderersansvar. Dette kan ha konsekvenser for hvordan lærerne samler bevis for elevenes læring, og hvordan de fortolker disse bevisene. Vurdering for læring «Vurdering for læring» som konsept utvider begrepsrammene, og dette krever også nye kunnskaper, ferdigheter og forståelser på ulike nivå i skolen for at vurderingen skal være reliabel.

Hvordan kan vi så sikre kvalitet i vurdering dersom lærere har svært liten forståelse av fundamentale begreper innen måling og testing? Parkes (2013)

foreslår følgende strategier for å forbedre reliabiliteten i læreres vurderingspraksiser:

1. eliminere risikoen for menneskelige feil med standardiserte tester;
2. gi detaljerte instruksjoner til oppgaver med åpne svar;
3. etablere beste praksis på utvikling av oppgaver;
4. veilede lærernes vurderingspraksis gjennom retningslinjer eller rubrikker for skåring av oppgaver;
5. øke lærernes forståelse av elevenes læring, slik at deres fortolkninger av elevens svar blir mer konsistente, for eksempel ved at lærerne involveres i utviklingen av vurderingsoppgaver og kriterier. Gjennom dette arbeidet kan lærerne utvikle en dypere forståelse av hva elevene kan tenkes å svare;
6. la lærerne danne praksisfellesskap der de kan få en felles forståelse av hva de forventer av elevene.

Disse kan bidra til å redusere ulikheter og fremme mer rettferdig (i betydningen reliabel) vurdering i klasserommet. Parkes påpeker likevel at det kan være uheldig å stille strenge krav om reliabilitet til vurdering som foregår i klasserommet. Lærere observerer og trekker slutninger om elevenes læring til svært mange formål og over ulike tidsrom. I en slik fortolkende prosess kan det være mer hensiktsmessig å etterspørre gyldigheten i lærerens tolkninger enn påliteligheten (Parkes, 2013).

Et nasjonalt system for vurdering bør etterstrebe rettferdig behandling av alle elever. I et slikt system har reliabilitet en naturlig plass. Når det gjelder vurdering i klasserommet, kan det imidlertid være mer problematisk å stille strenge krav om reliabilitet fordi dette kan gå på bekostning av validiteten i vurderingen. Vi velger derfor å konkludere med at det er behov for mer kunnskap som reliabilitet som fenomen, men at denne kunnskapen alltid må ses i sammenheng med konteksten og formålet for vurderingen, særlig i klasserommet.

Annes ulike praksiser i fagene matematikk og engelsk har dannet rammen om dette kapitlet, som har løftet fram validitetsbegrepet i FIVIS-forskningen. Vi vil videre undersøke hva som gjør at noen skoler lykkes i arbeidet med Vurdering for læring, der vi spesielt ser på balansen mellom individuelle læreres vurderingspraksiser, utviklingen av organisasjonsforankrede praksiser og utdanningsmyndighetenes krav. Dette vil være viktig for å forstå validitet i vurdering som del av et større nettverk av fortolkninger og forståelser, der *kontekst* og *verdier* vil være sentrale begreper for å validere vurderingspraksiser på alle nivå.

4. Hvordan er vurderingspraksisene og -kulturene i norske skoler?

Henning Fjørtoft, Inger Langseth og Lise Vikan Sandvik

Hvilke vurderingspraksiser finnes i norske klasserom? Det er en av tre overordnede problemstillinger i FIVIS. I dette kapitlet presenterer vi et utvalg praksisfortellinger som illustrerer mangfoldet i vurderingspraksiser og -kulturer som er observert og analysert i FIVIS-prosjektet, for å gi noen perspektiver inn mot denne problemstillingen. Fortellingene er konstruert ut fra klasseromsobservasjoner, dokumentanalyser og intervjuer. Vi kommenterer hver fortelling kort. Lærernes og skolenes tilnærminger til vurdering settes deretter inn i et teoretisk rammeverk som kan bidra til å forstå vurdering i både individuelt og organisatorisk perspektiv.

4.1 Praksisfortelling 1 – Julie som lærende lærer

Engelsklæreren Julie arbeider ved en 1–10-skole som har vært med i satsingen Vurdering for læring, og har etter noen år som lærer på ungdomstrinnet utviklet en vurderingspraksis preget av sterk elevmedvirkning og tydelige læringsmål.

Klokka er 09.00 presis. Julie samler klassen og hilser smilende «good morning, everyone». Hun starter med å si at hun er spent på hvordan skriveleksa har gått, og hun spør om noen av elevene vil presentere teksten sin for resten av klassen. Flere elever rekker opp hånda, og tre av elevene leser opp tekstene sine. Det er tydelige kriterier for skriveleksa, og disse blir gjentatt før klassen gir tilbakemeldinger til den enkelte elev: Er *påstandene* underbygget med argumenter? Er *argumentene* overbevisende for tilhørerne? Har eleven en tydelig *uttale*? Bidrar *kroppsspråket* til å formidle innholdet? Og så videre. Elevene bruker begrepene i sine tilbakemeldinger. Så leverer alle elevene inn sine tekster til læreren, som vil gi en vurdering med kommentar.

Elevene diskuterer innhold og struktur i tekstene sine, og samtalen struktureres ved hjelp av faglige begreper. Julie utfordrer elevenes evne til refleksjon. Hun ber dem begrunne argumentene sine og fortelle hvordan de gikk frem for å skrive teksten. Hun skiller en argumenterende, personlig tekst fra en informativ tekst, hvor elevene skal hente inn informasjon og skrive sin forståelse av teksten i en ny tekst. «Argumenterende tekster krever at du har meninger og argumenterer for meningene dine», sier Julie. Julie gir tydelige tilbakemeldinger når elevene gir gale svar eller er upresise. Hun bruker «ja» eller «nei» for å bekrefte eller avkrefte elevutsagn. Hennes tilbakemeldinger

preges av faglighet. Hun stiller klare forventninger til elevenes oppførsel, og det slås ned på avvik. En elevkommentar om at timen ikke har begynt (klokka på veggen er for sen), resulterer i at hun tar ned klokka fra veggen. «Slik at det ikke blir et distraksjonsmoment», kommenterer hun. Elevene stiller spørsmål, og Julie svarer på relevante spørsmål, men lar ikke elevene slippe til med irrelevante spørsmål i klasseromsdialogen om argumenterende tekster.

Deretter gjennomgås leseleksa til neste gang. Elevene lytter til tre forskjellige tekster på tre nivå i læreboka. Elevene får i hjemmelekse å arbeide med glosene og oversette teksten til neste gang. Da får de tilbakemelding på arbeidet fra læreren, slik at de kan forbedre seg til lese- og oversettelsesprøven neste fredag. Dersom de ikke uttaler, leser og oversetter «godt nok» på onsdag, må elevene skrive en logg hvor de beskriver hvordan de har arbeidet med lesing og ordinnlæring, og loggen skal underskrives av foreldrene før den leveres til læreren. Dette beskrives som lite ønskelig fra elevenes side.

4.1.1 Kommentar til praksisfortellingen

Utviklingen av nye vurderingspraksiser er ikke noe som bare kan implementeres i skolen – det er også snakk om en individuell læringsprosess der den enkelte lærer må tilegne seg nye kunnskaper gjennom utprøving og refleksjon. Disse nye kunnskapene må deretter settes i sammenheng med allerede eksisterende undervisningspraksiser, slik at læreren kan integrere bedre vurdering i sin yrkesutøvelse.

Julie er selv i en læringsprosess som lærer. Hun ønsker en bekreftelse på at hun er på rett vei i arbeidet med vurdering for læring, samtidig som hun gjerne vil diskutere områder hun mener er problematiske. Julie tar i bruk vurdering for læring for å tilpasse opplæringen til den konteksten hun underviser i. Hun har tydelige rammer for faglig og sosial samhandling, og hun følger opp arbeidet hos den enkelte elev gjennom tilbakemeldinger som legger til rette for å fremme læring i et samarbeid med hjemmet. Elevene skriver egne tekster om egne tema og presenterer tekstene sine. Så får de tilbakemelding på fremføringen fra lærer og medelever og på oppbygningen av avsnitt og språket fra læreren. Hun mener at vurdering for læring ikke er noe nytt. «Det jeg gjør nå, er å tenke over hva jeg vil med timene, hvilke begrep som er sentrale, og hvordan jeg kan være tydelig for elevene». Hun legger til at de bruker et metaspråk om tekster, de snakker om at teksten må virke planlagt og «ikke bare ramle ut av hodet og ned på arket».

4.2 Praksisfortelling 2 – mange praksiser og forståelser under samme tak

Borgheim er en videregående skole i en større norsk by. De fleste elevene som søker seg hit, har gode eller svært gode karakterer. I undervisningen i de ulike fagene varierer både elevdeltakelsen og synligheten i forventningene. Ved skolen foregår vurdering på ulikt vis mellom fagene, og lærerne har også ulik forståelse av hva vurdering kan være.

Norsklærerne ved Borgheim samarbeider ikke om planlegging av undervisning eller vurdering. De begrunner dette med at de ønsker å utnytte sine sterke sider. Det foreligger ikke en årsplan for faget ved skolen, og mål og kriterier blir heller ikke lagt ut på læringsplattformen. Ifølge lærerne selv mangler de opplæring i teknologien. Lærerne sier imidlertid at de reflekterer over undervisningen sammen. I en norsktime vi observerer, møter elevene synlige vurderingskriterier for muntlige oppgaver i norskfaget. Elevene arbeider i grupper, og læreren forventer at alle deltar. Læreren går noen ganger rundt og forsikrer seg om at alle forstår, og gir veiledning der det trengs. Det er læreren som skal vurdere elevenes prestasjoner når arbeidet er ferdig. Elevene arbeider også med skriveoppgaver, men elevene får ikke informasjon om hvordan teksten skal vurderes av læreren. I en annen time innleder norsklæreren med å be elevene snakke om sine assosiasjoner til en novelletittel. Læreren har selv laget oppgaver til teksten, og disse skal elevene arbeide med individuelt eller i par. Samtalen i klasserommet berører strukturer i teksten: personkarakteristikk, handling, tematikk og retrospeksjon. Bare 4–5 elever deltar. Læreren stiller generelle spørsmål om rasisme, uten at klassen blir engasjert. Når læreren spør om elevene synes teksten er vanskelig, sier de at de kanskje kan komme tilbake til denne teksten senere. Timen avsluttes med en påminnelse om leksene til neste uke. I intervjuer sier elevene at de har opplevd hverandrevurdering i norskfaget som lite seriøst når det foregår uten kriterier. En av elevene sier at kvaliteten på tilbakemeldinger er viktigere enn å ha mål.

I engelskfaget deltar elevene selv i formuleringen av vurderingskriterier. Læreren oppfordrer dem til å lese planene for undervisningen og spørre dersom noe er uklart. Læringsmiljøet fremstår som trygt og tillitsfullt, og samtalen med klassen i plenum foregår uten håndsopprekking. Undervisningen er preget av høy deltakelse og engasjement i klassen, og når elever ytrer seg, kan andre elever følge opp dette i plenum. Det er en eksplisitt forventning fra læreren at elevene legger bort laptopene sine i slike økter. Læreren stiller utfordrende spørsmål som krever at elevene reflekterer over sin begrepsforståelse. De vektlegger også å kommentere innholdet i elevenes muntlige ytringer og overser språklige feil i muntlige kommunikasjonssituasjoner. Læreren gir uttrykk for at valget av muntlige aktiviteter har som formål å fremme tillit og

relasjoner i klassen, og at den positive responsen til elevene skal fremme dette. Medelevvurdering skjer først og fremst spontant mellom elever i grupper. Det skjer ingen organisert form for egenvurdering i timene. I en av timene vi observerer, tilbyr læreren elevene hjelp når de arbeider. Noen elever ønsker ikke dette. Da ber læreren elevene reflektere over hvorfor hun ikke ønsker dette. Én lærer starter en time med å introdusere mål, men disse blir ikke nevnt på slutten av timen. Elevene beskriver «karakter og kommentar» som en typisk tilbakemeldingsform og nevner slurvfeil, for korte tekster, småfeil og «forklar bedre» som typiske kommentarer fra læreren. Kontakt med foreldre er et satsingsområde på skolen, sier læreren: De inviteres til samtale og får tilgang til læringsplattformen. Læreren bruker små tester for å motivere og for å få informasjon om elevenes kunnskapsnivå, slik at undervisningen kan tilpasses til det. Undervisningen planlegges for en måned av gangen. Til muntlige presentasjoner bruker læreren *two stars and a wish* som tilbakemeldingsform.

I matematikkfaget på Vg1 har elevene nettopp hatt prøve. Læreren ber elevene vurdere seg selv: «Skriv på den karakteren dere synes dere fortjener», sier han. De må også løse oppgavene de ikke har fått til, og returnere prøven til læreren. Læreren ønsker innspill fra elevene om hvordan de oppfattet prøven. Det er uvisst hva læreren skal bruke innspillene til. Den endelige karaktersettingen skjer i samråd med lærerne i andre klasser, forteller læreren til oss etter timen. Det er lite snakk om faglig læring eller utvikling i disse aktivitetene. I dialog med enkeltelever i klasserommet rettes oppmerksomheten mer mot faglig forståelse og læringsprosesser, men også her ser vi at læreren er mest interessert i hvorvidt svarene er feil eller rett. Elever som gir uttrykk for at de mestrer de enklere oppgavene, får lov til å hoppe over dem og gå løs på litt vanskeligere deler. «Det er ikke nødvendig å øve på ting du kan», sier læreren. Av og til tar læreren kontakt med elever som ikke rekker opp hånda for å få hjelp mens de løser oppgaver. Læreren fremstår som interessert i elevene, og både humor og dialog er en del av klasseledelsen. Det differensieres først og fremst for de flinke elevene som selv sier fra om at oppgavene er for enkle. Vi observerer flere samtalsituasjoner der læreren ikke sjekker om elevene har forstått de faglige premissene som aktiviteten bygger på.

I kroppsøving på Vg1 bruker læreren tid i starten av timen til å forklare formålet, innholdet og organiseringen av aktiviteter. Vi er på besøk i en time som har utholdenhetstrening med løping av intervall som tema. Elevene må holde seg løpende orientert om fagets kompetansemål på læringsplattformen, og i denne timen skal de være «personlige trenere» for hverandre. Læreren leder, støtter og oppmuntrer elevene kontinuerlig gjennom timen og leder også de personlige trenernes aktivitet. Timen er preget av hyppig og positiv dialog mellom elevene. Mot slutten av timen må elevene oppsummere hva de har gjennomført. Noen elever har med dokumentasjon fra helsepersonell

(lege/kiropraktor) og ønsker tilrettelagte opplegg. De må lage en plan for sine aktiviteter, som de sender til læreren på læringsplattformen, og må selv dokumentere hvordan deres aktiviteter relaterer seg til utvikling av utholdenhet. Elever som kommer til timen med egenmeldinger om helse, må delta som observatører i kroppsøvingsaktivitetene. Timen preges av både formelle treningsaktiviteter og mer uformell lek, men alle aktivitetene er knyttet til det overordnede målet (utvikling av utholdenhet).

4.2.1 Kommentar til praksisfortellingen

Borgheim er et eksempel på en skole der et mangfold av praksiser og forståelser kommer til uttrykk. Det er for eksempel relativt store forskjeller i graden av eksplisering av mål og elevaktivitet. Forskjellene kan henge sammen med manglende fortolkningsfellesskap mellom lærerne. Reliabilitet i karaktersetting omtales mer detaljert enn hva slags læringsprosesser eller -utbytte faget rommer. Elevene involveres i samtaler med læreren om hvilke karakterer de sikter mot i faget. Matematikklærerne har et system med hyppige prøver der deloppgavene tildeles poengsummer, og de bruker en finmasket karakterskala med pluss, minus og skråstrek. Dette regner de som vurderingen i faget. Faglige tilbakemeldinger i timene er ikke en del av deres vurderingsbegrep – «det kaller jeg tilbakemelding», sier en matematikklærer: «Jeg kaller ikke dét *vurdering*.» Ulikhetene som vi ser i og mellom språkfagene, indikerer at lærerne i liten grad samarbeider om undervisningen, og at lærerne derfor ikke har felles forståelse av hva vurdering er eller kan være i ulike fag og læringsprosesser.

4.3 Praksisfortelling 3 – fra mål og kriterier til praksisfellesskap

Langnes barneskole er en 1–7-skole i en større norsk by. Skolen har jobbet mye med vurderingskriterier og respons i sitt utviklingsarbeid. Det lokale arbeidet med læreplanen har blitt opplevd som både nyttig og vanskelig. Lærerne ønsket å gjøre det tydelig for elevene hva de skulle lære, og hvordan arbeidet skulle vurderes. Dette ble opplevd som viktig, men tidkrevende av lærerne. Den eksplisitte vektleggingen av forventninger til elevenes læring har imidlertid blitt erstattet av en sterkere felles forståelse:

Nå ligger det mye mer i ryggmargen til oss og til elevene. På førstetrinnet har de det faktisk klart for seg, det at vi snakker om forventning og hvilke krav er det her, og sånne ting. Det synes jeg er en ganske riktig måte å bevege seg videre på. Det er noe som ligger i klasserommet, ikke bare sånn som vi må løfte opp og skrive opp.

Gjennom arbeidet med eksplisitte forventninger har det blitt bygd opp en sterkere felles forståelse av læringsmålene i skolefelleskapet. Behovet for å skrive ned målene har blitt redusert, og isteden er forventningene noe som «ligger i klasserommet» og praksisfelleskapet. Ifølge lærerne har elevene en forståelse av hva kvalitet er, og den kommer «mer naturlig» – de har den «under huden». Selv om ikke alle er med hele tiden, deltar overraskende mange også på de laveste trinnene. «Du skulle hørt de samtalene de har med medelevene sine om hvor de står fast», sier en lærer til oss. Endringen mot en mer kollektiv vurderingskultur har også ført til at det har blitt lettere for lærere å overføre praksiser mellom fag.

Allerede på fjerde trinn er elevene sterkt involvert i vurderingen av sin egen læring. Dette fører til en sterk bevissthet om egne læringsprosesser, sier lærerne. Den sterke elevinvolvingen er en relativt ny utvikling ved skolen. Lærerne er overrasket over hvordan dette har utviklet seg ved skolen – også for de aller yngste elevene. De beskriver formen på elevenes egen- og medelevvurderinger som «utilslørt, usensurert, direkte» og kommenterer at «ungene tåler hverandres kommentarer». Selv på første trinn blir elevene engasjert i parvurdering, og lærerne tilrettelegger for at elevene skal utvikle en forståelse av sin egen arbeidsinnsats og hva de arbeider med. «Når du er seks år, klarer du det helt fint», sier en av lærerne.

Lærerne opplevde tidligere vurderingsarbeidet som krevende, men sier at endringer i skolekulturen og lærerrollen ved skolen har gjort dem tryggere, slik at de tør slippe til elevene. Gjennom fast ukentlig erfaringsutveksling på fellestid og tette koblinger mellom refleksjon og utprøving har lærerne utviklet en kultur for samarbeid. De presenterer for hverandre og reflekterer over egen praksis. «Det er ingen vits å ha et foredrag og tre uker med litt arbeid», sier de. Lærerne kommenterer at det er klare fellestrekk mellom måtene elevene og lærerne lærer på.

Vurderingskompetanse handler om å ha et repertoar av praksiser, at lærerne kan kartlegge hvor elevene er, og hvor de skal, at de kan bruke kunnskap og erfaring for å se fremover i tid, forstå elevenes ståsted, ha god dialog med elevene og ha trygghet i egen praksis. De ser også fagkunnskaper, klasseledelse, planlegging og organisering på teamet som en grunnleggende del av vurderingskompetanse. Lærerne mener at rektorer bør ha denne samme kompetansen, selv om en skoleleder selvfølgelig ikke kan alle fagene like godt. De nevner at assistentene som arbeider på trinnet, ikke alltid har forståelsen som trengs i slike situasjoner, og at lærerne derfor må veilede deres praksis. Dette kan oppleves som utfordrende for begge parter.

Lærerne på Langnes skole forsøker å balansere ønsket om høye forventninger med godt skole-hjem-samarbeid. Dersom elever av ulike årsaker ikke gjør

leksene sine, får de med seg et skjema hjem der foreldrene må signere ukentlig. Dette har bidratt til å strukturere arbeidet og ansvarliggjøre foreldre og elever, sier lærerne. Mestring for alle elever er et gjennomgående trekk ved lærernes tenkning om undervisningen – til tross for at elevene opplever slike læringsprosesser som anstrengende. Lærerne sier at de foretrekker å «senke hastigheten i aktivitetene» fremfor å haste gjennom undervisningsprosessene. De vil heller «bli ferdig og få kvalitet» enn at elevene lager «et nestenprodukt til 30–40 prosent»: «At alle kommer i mål, det er det viktige.» Gjennom god planlegging blir ulike skrivestrategier forsterket i ulike fag på samme trinn. For eksempel må lærerne på femte trinn integrere skriving av sammendrag i både naturfag, norsk og matematikk. Kunnskapen om slike praksiser har kommet som følge av at enkelte lærere har tatt videreutdanning. Skolens ledelse og plangruppe har sørget for at de andre lærerne har fått utfordringer innen samme tema. Teamene har fått oppgaver som de må gjennomføre i praksis og dele med kolleger. Slik blir kompetansen fra den individuelle videreutdanningen anvendt i kollegiet som helhet. Lærerne kommenterer at dette krever mot til å holde på målsettingene over tid og unngå at det blir for mange baller i luften samtidig. Skolen har også samarbeidet med andre skoler i nettverk for å lære mer om vurdering for læring.

Langnes kombinerer løpende dialog om elevens utvikling med utstrakt bruk av medelevvurdering i par og mer formelle summative stoppesteder (fast elevsamtale en gang per halvår). Den faste elevsamtalen beskrives som en arena som det ryddes tid til, og der elevene virkelig får tid til å reflektere over hvordan det går. I tillegg til dette kommer samtaler med foreldre. De har tidligere forsøkt å bruke skolens læringsplattform for å sette mål og gi tilbakemeldinger, men noen lærere har opplevd at teknologien skaper avstand, og at den personlige kontakten ansikt til ansikt er en «viktig del av kommunikasjonen». En forutsetning for elevmedvirkning i vurderingen er at lærerne forstår den som «ferskvare».

4.3.1 Kommentar til praksisfortellingen

Langnes er et godt eksempel på at arbeid med eksplisitte mål og kriterier over tid fører til et sterkere fortolkningsfellesskap. Både lærere og elever har nå en forståelse av forventningen til læring. Lærerne gir dessuten uttrykk for et syn på læring som vektlegger forståelse, mestring og motivasjon for alle elever, og de ser vurdering som en måte å underbygge slike prosesser på. Dette er også et godt eksempel på en skole som har lyktes i å involvere elevene sterkt i vurderingsprosessene, og vi ser mange eksempler på egen- og medelevvurdering i aktiv bruk. Selv om dette er relativt nytt ved skolen, uttrykker lærerne tiltro til at elevene mestrer det selv fra første klasse.

Elevinvolvingen har imidlertid betydd at arbeidet med synliggjøring av mål og kriterier har blitt mindre viktig. Operasjonaliseringen av læringsmål skjer nå gjennom deltakelse i elevaktive praksiser og ikke gjennom ulike målformuleringer.

Arbeidet med vurdering har over tid ført til at lærerne ved denne skolen ser mange sammenhenger mellom vurdering og mange andre aspekter av sin yrkesutøvelse. Deres forståelse av læring og læringsprosesser har ført til at de tør prioritere få og komplekse oppgaver for elevene. Mestring av disse oppgavene er det primære, ikke å haste gjennom læreplan eller pensumbøker. Lærernes vurderingskompetanse kommer altså til uttrykk gjennom de praksisene de leder, og gjennom skoleledelsens organisering av lærernes læringsprosesser. De anvender et bredt repertoar av både standardiserte og lokalt utformede vurderingsformer og bruker resultatene for å tilpasse undervisningen til elevenes læringsutbytte.

4.4. Fire forskjellige vurderingskulturer

De ulike vurderingskulturene i praksisfortellingene illustrerer graden av elevmedvirkning og tydelige forventninger i arbeidet med vurdering. I de forrige delrapportene har vi påpekt at lærere anvender læringsmål over relativt korte tidsperioder, og at det er mer utfordrende å operasjonalisere kompetansemål over lengre perspektiver (Sandvik mfl., 2012). Vi har også sett tendenser til utvikling av parallelle kulturer på noen skoler: Mens noen lærere har gode planer for undervisningen, kan andre lærere la være å planlegge (Sandvik & Buland, 2013). Det kan være mange årsaker til disse ulikhetene, som vi antyder i kapittel 3.3, «Validitet i lokalt læreplanarbeid»: for høye forventninger til lærernes evne til å tolke og anvende læreplanmål, en opplevelse blant lærerne av at eksplisitte mål er irrelevante eller unødvendige, manglende kapasitet til å lede utviklingsarbeid i skolen, verdimesse eller etiske syn og så videre. Vi vil likevel fremheve *tolkning* som et sentralt element i slike prosesser fordi lærernes tolkninger av intensjonene i læreplanen vil legge føringer på hvordan læringsmålene formuleres, og på graden av elevinvolving i vurderingen.

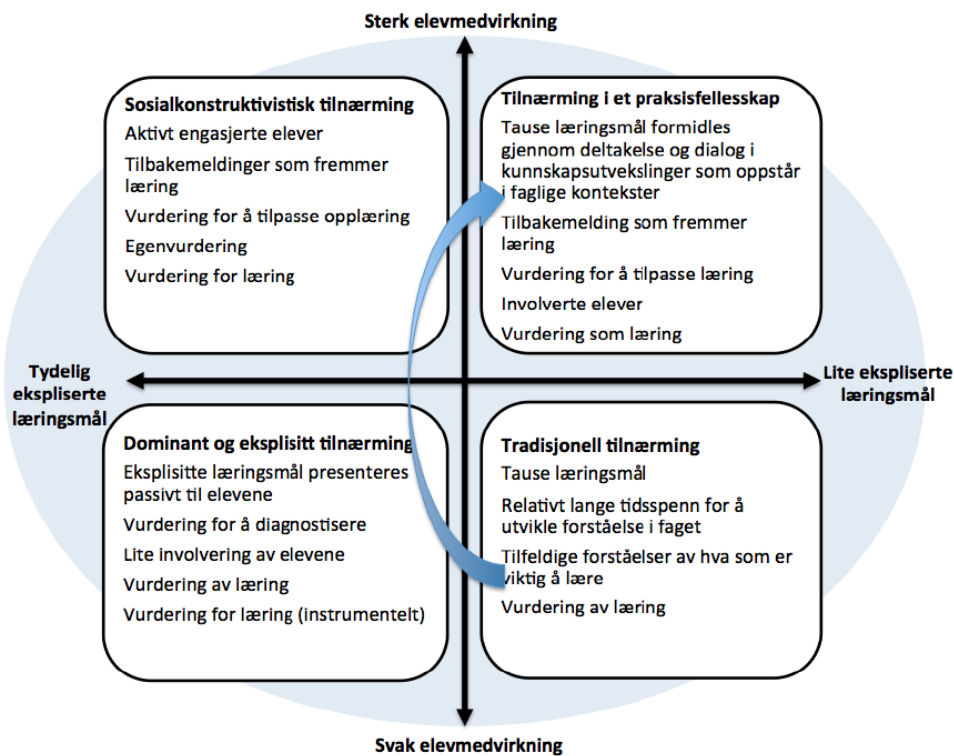
Slik vi ser det, er praksisfortelling 2 et eksempel på at lærerne i noen fag validerer sine vurderingspraksiser gjennom å etterspørre om slutningene holder. Kroppsøvingsfaget fremstår for eksempel som gjennomsyret av en forståelse av både utholdenhetstrening og aktiv deltakelse som overordnede mål. Det samme kan sies å kjennetegne undervisningen vi observerer i engelskfaget. I andre fag er dette mer tilfeldig gjennomført. Norskfaget er preget av lav deltakelse, og dermed får læreren liten tilgang til elevenes

forståelse av den. Elevenes syn på undervisningsaktivitetene og læringsprosessene etterspørres i mindre grad, og dermed blir det også vanskelig å se hvorvidt lærerens slutninger stemmer med elevenes egne erfaringer og opplevelser i læringsprosessen.

I forskrift til opplæringslova, kapittel 3, presiseres elevenes rettigheter i ulike vurderingssituasjoner i skolen. Det «skal vere kjent for eleven, lærlingen og lære kandidaten kva som er måla for opplæringa og kva som blir vektlagt i vurderinga av hennar eller hans kompetanse» (forskrift til opplæringslova § 3-1). De skal også «delta aktivt i vurderinga av eige arbeid, eigen kompetanse og eiga fagleg utvikling» (§ 3-3). Med andre ord fremhever forskriften særlig to faktorer:

- 1) eksplisitte mål eller forventninger for elevenes læring
- 2) aktiv deltakelse fra hver enkelt elev i vurderingen av læringsprosesser og læringsutbytte

Lovteksten krever ikke eksplisitt at elever skal vurdere hverandres læring (hverandrevurdering), men dette kan likevel forstås som en del av de vurderingspraksiser som kan anvendes for å realisere kravene i forskriften. De to faktorene som forskriften fremhever (eksplisitte forventninger og aktiv elevdeltakelse), kan settes inn i et firefeltsdiagram, som vist i figur 2.



Figur 2. Fire ulike vurderingskulturer (modifisert etter O'Donovan mfl., 2006).

Vi ser denne modellen som et mulig kart over utviklingen i norske skolars vurderingskulturer av flere årsaker. For det første bygger den på de samme faktorene som fremheves i forskrift til opplæringslova. For det andre støttes den av de data som vi har samlet inn i løpet av prosjektet. Og for det tredje kan den tjene som en teoretisk modell som kan gi retning til skoleutviklingsprosesser. De fire vurderingskulturene kan forklares slik:

1. Tradisjonell modell

Ved skoler som følger en tradisjonell modell, er målene for elevenes læring sjelden synlige. De nevnes ikke i timene, og sammenhengene mellom læringsaktivitetene og læringsmålene fremheves ikke som en del av praksis i klasserommet. Læring forstås som en prosess over relativt lange tidsspenn, og elevenes innsikt i sentralt faglig innhold skjer tilfeldig mer enn planlagt. Praksisfortelling 2 kan fungere som et eksempel på en skole og en kultur der læringsmålene for elevene er lite synlige.

2. Dominant og eksplisitt logikk

Noen skoler legger mye vekt på å formidle læringsmålene til elevene, enten i form av ordrette kompetansemål fra læreplanen eller i form av operasjonaliserte læringsmål som er tilpasset læringsaktiviteter eller elevenes utviklingstrinn. Elevene presenteres for målene, men deltar ikke selv i formuleringen av dem eller i utformingen av oppgaver og aktiviteter. Operasjonaliseringen av målene kan derfor være begrenset.

Både praksisfortelling 1 og 2 har elementer av dominant og eksplisitt logikk. Engelsklærerens undervisning i den videregående skolen er et eksempel på slike praksiser. Læreren arbeider godt med mål og kriterier, men involverer elevene i liten grad i vurderingsarbeidet. Vurderingsarbeidet er blant elevene i høy grad forstått som summativ vurdering. Denne tilnærmingen er typisk for skoler som er i startgropen for arbeidet med vurdering for læring, men som ikke har klart å utvikle en forståelse blant kollegiet av hvilket læringssyn en læringsfremmende vurdering representerer, og hvilke implikasjoner det har for pedagogisk praksis.

3. Sosialkonstruktivistisk tilnærming

Skoler som har en utpreget sosialkonstruktivistisk tilnærming til vurdering, vektlegger at elevene skal være aktivt engasjerte og deltakende i utformingen av læringsmål, vurderingskriterier, oppgaver og lignende. Lærerne tilrettelegger for formelle prosesser som involverer elevene i å gjøre forventningene til læringen eksplisitt.

Praksisfortelling 1 viser en lærer som er på vei mot en sosialkonstruktivistisk tilnærming. Vi møter en lærer som er i ferd med å utvikle en praksis der hun tar

i bruk vurdering for læring for å tilpasse opplæringen til den konteksten hun underviser i. Hun er en tydelig leder med tydelige rammer for faglig og sosial samhandling, og hun følger opp arbeidet hos den enkelte elev gjennom tilbakemeldinger som legger til rette for å fremme læring i et samarbeid med hjemmet. Hennes klasseledelse og de forventningene hun signaliserer og følger opp, virker også motiverende på elevene.

4. *Praksisfellesskap*

Praksisfortelling 3 er et eksempel på en skole som kan plasseres i *praksisfellesskap* i modellen. Selv om lærerne på denne skolen tar i bruk en lang rekke praksiser, er det ikke praksisene i seg selv som konstituerer de verdifulle aspektene av vurderingskulturen på skolen. Wiliam (2011) har påpekt at enkelte kan ha en tendens til å se vurdering for læring som *verktøy eller instrument*. Dette er en feilslutning, for den samme prøven kan anvendes til både formative og summative hensikter. Istedenfor å forstå kartlegging eller egenvurdering som konkrete verktøy bør vi heller forstå dem som del av *prosesser* der informasjon fra ulike vurderingssituasjoner anvendes med den hensikt å støtte elevenes læring og utvikling. I dette arbeidet inngår gode instrumenter som en nødvendig komponent, men de må brukes på en gjennomtenkt måte, sier Wiliam. Slike prosesser kan finne sted over ulike tidsspenn, fra korte prosesser av 1–2 dagers varighet til lengre sykluser over to år. Prosessene deler imidlertid ett kjennetegn: (1) Bevis for elevens læring blir samlet inn og fortolket av lærere og elever for å forbedre undervisnings- og lærings situasjoner, og (2) de avgjørelsene som tas, blir kvalitativt bedre eller sterkere fundert gjennom fortolkningen av disse bevisene (Wiliam, 2011, s. 38 ff.). Lærerne i vårt tilfelle uttrykker klart at deres egen forståelse av vurdering har utviklet seg gjennom felles læringsprosesser i organisasjonen. Utprøving og erfaringsdeling sammen med mer formell videreutdanning av enkeltlærere har styrket den felles kompetansen til å forankre vurderingen i elevenes læringsprosesser. Det er dermed nære paralleller mellom lærernes kollektive læringsprosesser i skoleutviklings sammenheng og den prosessuelle forståelsen av vurdering som undervisningen bygger på i klasserommet. Vurderingen har blitt *embedded*: integrert i andre undervisnings- og læringsprosesser.

Skoler som vektlegger å dyrke frem praksisfellesskap i vurdering, har gjerne arbeidet med eksplisitte mål i det lokale utviklingsarbeidet tidligere. De har kommet frem til at en overdreven vekt på synlige læringsmål ikke nødvendigvis fremmer elevenes læring, og prioriterer derfor å sikre at elevene forstår hva de skal lære, gjennom mer uformelle utvekslinger, som kontinuerlig dialog og høy elevdeltakelse i ulike læringsaktiviteter. O'Donovan mfl. (2006) argumenterer for å nedtone betydningen av læringsmål til fordel for økt elevdeltakelse. I norsk sammenheng kan dette virke overraskende siden forskriften til opplæringsloven krever at målene er kjent for eleven. De påpeker

imidlertid at det i praksis er vanskelig å formulere læringsmål og vurderingskriterier som tilfredsstillende krav til både faglig presisjon og praktisk nytteverdi. For eksempel kan det være vanskelig for elever å gjøre bruk av slike verktøy fordi de ikke forstår dem, eller fordi de mangler kunnskaper om den konteksten som de skal anvendes i. Når O'Donovan mfl. påpeker at dette er en utfordring for studenter på universitetsnivå, er det ikke tvil om at elever i grunnskole og videregående opplæring vil oppleve lignende vansker. De argumenterer for at læringssituasjoner bør kombinere eksplisitte forventninger med overføring av taus kunnskap gjennom sosialiseringprosesser der elevene får erfaring og tilegner seg kunnskaper gjennom observasjon, imitasjon og øvelse. O'Donovan mfl. understreker også at de fire kulturene ikke er gjensidig ekskluderende, og at de kan sies å ta opp i seg praksiser og tenkning fra de andre stadiene. Det gjelder spesielt de to siste kulturene (*sosialkonstruktivistisk tilnærming* og *praksisfellesskap*).

4.5 Helhet og sammenheng i utviklingen av vurderingskulturer

Lærere som kan plassere vurdering i en større sammenheng, har en fordel når de skal skape forståelse også utenfor skolens vegger. Fullan påpeker dette når han sier at evnen til å delta i større diskusjoner om testing og vurdering kan regnes som en komponent i vurderingskompetanse: «[A]ssessment literacy is a powerful coherence-maker» (Fullan, 2001, s. 117). Med andre ord: Når elever og lærere har høy vurderingskompetanse, er de bedre i stand til å skape meningsfulle sammenhenger mellom de aktiviteter og kontekster som skjer innenfor rammene av skolens virksomhet. Slik meningsskapende prosesser bør i seg selv forstås som sentrale for elevenes opplevelse og forståelse av sammenhengene mellom undervisning, vurdering og læring i skolen. Mål og kriterier kan være viktige verktøy for både elever og lærere i disse prosessene, men med sterk vektlegging av elevdeltakelse og felles forståelse kan de større sammenhengene mellom vurderingen og andre aspekter av elevenes læringsprosesser bli like viktige. Vurdering er altså ikke bare innsamling av data om elevenes utbytte eller et grunnlag for å endre lærernes undervisning. Disse aspektene er nødvendige, men ikke tilstrekkelige vilkår for dannelsen av et praksisfellesskap. Vurderingskulturen ved skolen er også preget av at lærerne og elevene *verdsetter* bestemte former for læring: aktive elever som bidrar i hverandres læringsprosesser, og som kan reflektere over sine egne, lærere som lærer sammen og utvikler nye praksiser kontinuerlig, en skoleledelse som tilrettelegger for og forventer at lærerne skal samarbeide om vurderingen, og et tett samarbeid mellom skole og hjem om elevenes utvikling.

Både formative og summative vurderingssituasjoner kan spille en rolle når lærere skal etterprøve sin forståelse av elevenes læring. En *teach to the test-*

holdning blir gjerne kritisert i mediene som en snevert orientert form for undervisning. Popham (2003) påpeker imidlertid på at tester i noen tilfeller er *for lite* anvendt i skolen. Lærere kan bruke tester for å vurdere hvorvidt elever faktisk har tilstrekkelig progresjon i faget. Denne informasjonen kan i neste omgang brukes for å avgjøre hvorvidt vi kan gå raskere frem, eller for å se om vi må undervise noe på nytt. En slik praksis betinger imidlertid at læreren sørger for førvurdering i en eller annen form, og at informasjonen som samles inn underveis, fungerer som konkrete eksempler på hvordan læreplanmålene kan nås. Popham innrømmer imidlertid at en slik bruk av tester eller prøver avhenger av tolkningen av læreplanmålene. Lærere forstår læreplanmål på ulike måter, og dersom det er store avvik mellom fortolkningene, vil også elevenes forståelse påvirkes av dette. Popham kaller dette *teach toward your interpretation* og påpeker dermed at vi ikke egentlig underviser mot målene direkte, men mot fortolkningen av dem (Popham, 2003, s. 7). Popham påpeker videre at en lærer kan trekke gyldige slutninger om elevenes læring uten at dette alltid fører til briljant undervisning. Vi har altså ikke noen garanti for at utvikling av et sterkere fortolkningsfellesskap automatisk fører til mer læring. Men *ugyldige* slutninger om elevenes læring vil nesten alltid føre til feilslåtte avgjørelser, sier han. Det er derfor viktig at lærere arbeider godt med å utvikle representative måter å vurdere elevenes kompetanse på, og at lærere samarbeider om å sikre validitet i slike praksiser, spesielt for større og viktige arbeider i skolen (Popham, 2003, s. 48 ff.). Validitet i klasserommet er altså ikke bare et spørsmål om den enkelte lærers fortolkninger – vi kan like gjerne snakke om validitetskjeder som et felles ansvar for lærerne på en skole.

Et felles ansvar for validitet i vurdering kan innebære å

- granske læremidler og vurderingssituasjoner for å undersøke hvilken fortolkning av læreplanmålene som ligger bak, samt drøfte hvorvidt denne fortolkningen skiller seg betydelig fra lærernes egen forståelse av læreplanen;
- koble læreplanens kompetansemål til et sett med konkrete eksempler på bevis for kompetanse;
- bruke kolleger som kritiske venner for å diskutere hvilke slutninger som kan trekkes om en elevs læring ut fra en bestemt vurderingssituasjon;
- diskutere gyldigheten av vurderingen med foreldre/foresatte. Læreren kan for eksempel spørre om bildet av elevens læring som fremkommer i vurderingen, stemmer med deres inntrykk;
- diskutere lærerens slutninger om elevens kompetanse med elevene selv og spørre om lærerens inntrykk stemmer med deres egen forståelse av læringsprosessene og -utbyttet (Popham, 2003, s. 48 ff.).

Som listen antyder, innebærer ikke formativ bruk av vurdering utelukkende at læreren samler inn data for å fastslå hvor elevene er. Frey og Fisher påpeker at elevenes evne til å vurdere seg selv må fremmes av de daglige vurderingsteknikkene som læreren bruker (Frey & Fisher, 2011, s. 35). Når elevene står i sentrum for vurderingspraksisene i skolen, kan det også oppstå en tettere sammenheng mellom læringsprosessene som finner sted, og vurderingspraksisene som følger elevenes faglige utvikling.

4.5.1 Noen forbehold

Vi må presisere at vi ikke ser på firefelttdiagrammet som en garanti for å lykkes med utviklingsarbeid innen vurdering, og at det slett ikke er sikkert at synlige læringsmål er den eneste riktige vei å gå for å utvikle undervisningen. Som alle teoretiske modeller har den sine begrensninger. Vi må også presisere at prosedyrer for å sikre både reliabilitet og sterke validitetskjeder er et underliggende premiss for kvalitet i vurdering. Det gjelder alle de fire vurderingskulturene som er skissert her. Dersom operasjonaliseringen av kompetansemålene i læreplanen ikke er gjort på en hensiktsmessig og innsiktsfull måte, hjelper det neppe å gjøre læringen mer eksplisitt. Å arbeide med læreplanenes krav til læring innebærer et omfattende fortolkningsarbeid for den enkelte lærer. For å realisere intensjonene om utviklingen av grunnleggende ferdigheter i alle fag kreves det videre fortolkningsarbeid både innenfor og på tvers av tradisjonelle faggrenser. Kvalitet i det enkelte læringsmål er altså like viktig som graden av synlighet eller tydelighet, enten det handler om samfunnsfaglige kunnskaper eller skriving som grunnleggende ferdighet.

Vi ser likevel i datamaterialet til FIVIS-prosjektet at mange skoler har beveget seg i lignende baner, fra arbeid med synlige mål via økt vekt på elevdeltakelse og (i noen få tilfeller) til en fjerde fase der målene tones ned til fordel for videreutviklingen av et praksisfellesskap.

5. Det avhenger mye av læreren – elevers opplevelse av vurdering og læring

Trond Buland

I dette kapitlet gir vi et bilde av elevenes opplevelse av skolens vurderingspraksis, slik det har kommet til syne gjennom FIVIS. Elevenes opplevelse av hvordan vurdering påvirker deres læringer, er den sentrale problemstillingen i denne sluttrapporten og er også belyst gjennomgående i denne rapporten. Hensikten med dette kapitlet er å la elevenes egne fortellinger stå i sentrum. Vi ønsker altså her å gjøre elevene til de sentrale informantene om deres opplevelse av vurdering i skolen, for på den måten å få økt innsikt i hvordan elever på ulike nivå selv tenker rundt og opplever vurdering av og for læring. Et element i dette blir å se på elevenes opplevelse av hvordan vurderingspraksis forandrer seg ved overganger mellom trinn og skoleslag.

Dette kapitlet er basert på kvalitative data, primært elevintervjuer og observasjon, men også intervjuer med lærere og foresatte har vært viktig for å forstå forholdet mellom vurdering og elevers læring.²

Elevenes læring og utvikling er sentralt i vurdering for læring. Samtidig er elever og foreldre de som står lengst fra det teoretiske grunnlaget for god vurderingspraksis. Wiliam & Thomson (2007) har i sin forskning kommet fram til konsept bestående av fem strategier ved læringsfremmende vurderingspraksis som er viktig for elevenes læring. Disse aspektene ble drøftet i kapittel 3.1 og danner i dette kapitlet rammen for å få en større forståelse av elevenes opplevelse av vurdering og læring. Disse fem strategiene er:

1. å utarbeide tydelige mål og vurderingskriterier
2. å utvikle effektive klasseromsdiskusjoner og gode vurderingsoppgaver for dokumentasjon av elevenes ferdigheter
3. å gi tilbakemelding som fremmer læring
4. å aktivere elevene som ressurser for hverandres læring
5. å aktivere elevens evne til å lære å lære

² Vi planla i utgangspunktet å koble data fra vår breddeundersøkelse med data om elevenes læringsresultater (karakter/eksamen etc.). I dialog med oppdragsgiver valgte vi å ikke gjøre dette, da man ikke forventet å finne signifikante resultater på det nivå vi kunne gjennomføre våre analyser på. Data fra vår breddestudie viser heller ingen signifikante funn på sammenheng mellom læringsresultat og vurderingspraksis.

Disse strategiene vil være et utgangspunkt når vi lar elevenes stemme komme fram.

5.1 Overordnet inntrykk – god vurdering er viktig

Våre kvalitative data viser stor enighet i alle informantgrupper om at god vurderingspraksis bidrar til elevers læring. Elever er i intervjuer opptatt av at tilbakemeldinger fra lærer og medelever er et viktig element for å lære, at lærere er bevisst på hvordan tilbakemelding kan bidra til læring. Vi finner at foreldre deler denne vurderingen.

Klasseromsobservasjoner i FIVIS-studien viser mange eksempler på hvordan god, konstruktiv tilbakemelding setter i gang og forsterker læringsprosesser. I dialogen mellom elev og lærer skapes et rom for veiledninger og vurderinger. Denne dialogen i det korte tidsrommet på et konkret arbeid har vi analysert som læringsfremmende prosesser.

Elevene forteller i intervjuene om stor diskrepans mellom lærernes bruk av dialogen i læringsrommet for å få innsikt i elevenes kompetanse og hvilke tanker eleven har om egen læring. Samtidig ser vi også at dialogen mellom lærer og elev også påvirker læreres praksis. I første FIVIS-rapport (Sandvik mfl., 2012, s. 93) er det en tabell som viser hvilke kilder som er viktige for utviklingen av lærerens vurderingspraksis. Lærerne i studien rapporterer at «andre lærere ved egen skole» og «tilbakemeldinger fra elevene» er de kildene som er aller viktigst. Resultatet fra en regresjonsanalyse forteller at variabelen «Tilbakemelding fra elevene» har en signifikant betydning i forklaringen av god vurderingspraksis. Selv om forklaringskraften er liten, er elevens direkte og indirekte tilbakemelding fra elever til lærer interessant. Vi støtter oss her til Hatties metastudie (2009), som hevder at «[...] feedback is most powerful when it is from the student to the teacher» (Hattie, 2009, s. 173). Samtidig ser vi i våre intervjuer og i observasjoner at lærere i liten grad bruker resultatene av sine elevvurderinger til å utforme videre undervisning.

5.2 Ønske om elevaktiv undervisning

Med bakgrunn i mange elevintervjuer er hovedinntrykket at mange elever, uavhengig av trinn, etterlyser variert opplæring. Dette gjelder elever i alle skoleslag. På spørsmål om hva som er de morsomste fagene i skolen, trekker mange fram fag med praktisk islett: mat og helse, kunst og håndverk, kroppsøving. Flere elever nevner også prosjekt- og gruppearbeid som gode arbeidsformer, og deres begrunnelser er at de får muligheter til å arbeide på

ulike måter og de unngår å bli sittende ved pulten i undervisningstiden: «Det blir litt kjedelig å sitte så lenge på stolen og hele tiden bare høre læreren snakke, ikke å få gjøre noe selv, liksom.» Våre klasseromsobservasjoner viser også at det er en betydelig andel lærerstyrt undervisning i skolen.

Dette tilsvarer funn i andre prosjekter, og det er særlig elever i ungdomsskolen som beskriver skolehverdagen som ensformig og kjedelig, og de etterlyser mer varierte undervisningsformer og læringsarenaer og mer elevaktiv undervisning (se f.eks. Kaspersen mfl., 2012). Marthe Blikstad-Balas peker i sin doktoravhandling på at elever bruker for mye tid på ikke-faglig nettbruk i skolen (se f.eks. Bedre Skole 12/2012), og at de allikevel presterer godt på tester fordi de leser hjemme før prøver i skolefag.

Elevaktive undervisningsformer kan bidra til å skape interesse for læring, eller en situasjonsinteresse, som senere kan føre til utvikling av indre motivasjon (Hidi & Renninger, 2006). Læreren oppgave er å legge til rette for at elevenes læring foregår i en inkluderende læringskontekst. Elevene etterlyser kontekster der de får utfordret seg selv, bruke sine kunnskaper og ferdigheter kreativt i nye kontekster. En viktig forutsetning er gode relasjoner mellom elever og lærer, men også faglig gode omgivelser som utfordrer og motiverer. Når denne interessen er skapt, er det ifølge Hidi og Renninger (2006) et godt utgangspunkt for å utvikle den individuelle interessen for det faglige innholdet. God vurderingspraksis i fagene kan være et viktig bidrag for å skape den individuelle interessen for fagets innhold og egen læringsprosess. Vi finner det er en god sammenheng mellom elevers og foreldres forståelser av og erfaring med god vurderingspraksis og intensjonene i satsingen på vurdering for læring.

Sammenfattende kan vi si at i skolene som FIVIS-studien har undersøkt, er det forskjeller knyttet til hvordan lærerne arbeider med tilbakemeldinger til elevene. Elevene uttrykker at gode tilbakemeldinger er viktig for deres læring. Det synes som om elevene har en tydelig oppfatning om viktigheten av god vurderingspraksis for deres utvikling og læring, og at de fem strategiene for læringsfremmende vurdering, slik Wiliam og Thomson (2007) uttrykker det, er kjente strategier for elevene. Samtidig sier FIVIS-studien at elevene savner varierte og utfordrende arbeidsmetoder i fagene. Det kan synes som en satsing på vurdering for læring som ikke tar inn over seg behovet for variasjon i elevaktive undervisningsformer, kan stå i fare for å bli for instrumentell og ikke ha den ønskede effekten på økt motivasjon og læring.

5.3 Læring i forskjellige kontekster – overganger mellom skoleslag

I vår studie har vi blant annet intervjuet elever på 5. og 8. trinn. De erfarer at overgangen fra «småskoletrinnet» til «mellomtrinnet» og mellom barnetrinnet og ungdomstrinnet er stor. Overgangene innebærer på mange måter klare brudd i opplæringsløpet, og overgangene preges av møter med nye lærere og kamerater, nye lokaler, nye fag og ikke minst nye vurderingspraksiser. Overganger kan også være «skummelt». En elev på 7. trinn beskrev den kommende overgangen slik:

Det var vel egentlig det som var det skumleste med å gå fra barneskolen til ungdomsskolen og, at man er redd for å komme i en ny klasse hvor man ikke kjenner hverandre, og er redde for å dumme seg ut, liksom.

Møtet med en sterk summativ vurderingstradisjon på ungdomstrinnet nevnes av elevene som den største forskjellen i skolehverdagen. Men vi finner også eksempler på at elever opplever overgangen fra småskoletrinnet til mellomtrinnet som stor, og at denne overgangen innebærer tydeligere forventninger om læring og vurdering fra både lærere og elever. Karakterer er også det de fleste elever forbinder med vurdering. Vi registrerer at elevene ser ut til å være mindre opptatt av de nasjonale prøvene enn deres lærere.

Det er relativt vanlig i vårt utvalg at elever forteller om lærere som har forsøkt å forberede dem på overgangen til ungdomsskolen og den karakterbaserte vurderingen som vil skje der. Noen «trener» på overganger ved å gi uoffisielle karakterer skrevet i parentes. Andre elever forteller at lærere har søkt å forberede overgangen gjennom å gi skriftlige vurderinger, som de så muntlig «oversetter» til tallkarakterer i tilbakemeldingen. En gruppe elever på 7. trinn fortalte for eksempel følgende:

De har ikke lov til å gi karakterer, så tar de og skriver sånn der meget bra, supert, dårlig, bra pluss, bra minus.

Ja, bra, da er du 1–2, meget bra, da er du 3–4, og supert, da er du 5 eller 6.

Ja, en 3-er, da er du slapp. Da har du ikke gjort noe.

Men en 4 og 5, da gjør du det du skal, og kan tema, mens en 6-er, da har du liksom gått i dybden og forstår det ordentlig.

For eksempel ikke å ha skrevet bare det som står i boka, skrevet mer ut av det som står på Wikipedia, og har googlet det og lest andre bøker.

Elevene i vårt utvalg har forventninger om og opplever større krav i ungdomsskolen enn i barneskolen, og tilsvarende større faglige krav i videregående skole enn i ungdomsskolen. Mange elever i vårt utvalg gir uttrykk for at de ser fram til forventninger om faglige krav, men at det også ligger et snev av bekymring knyttet til overgangene. De fleste elevene forteller også om mer skolearbeid som følge av overgangene fra 7. til 8. trinn og fra 10. trinn til Vg1.

Også i intervjuene med foreldre møtte vi noen som fortalte om store overganger mellom skoleslagene, barnetrinn til ungdomstrinn og til videregående skole. Særlig er nok overgangen til ungdomstrinnet stor. Noen foreldre til elever ved 1–10-skoler, der barnetrinn og ungdomstrinn er i samme lokale, fortalte at overgangen følt som «å gå inn i en annen verden». Dette kom til uttrykk på mange måter, og ikke minst med hensyn til vurderingspraksis. Noen har etterlyst mer smidige, planlagte overganger. En foresatt til en elev med spesielle behov hadde opplevd denne overgangen som svært problematisk, selv om de to skoleslagene befant seg i samme bygning:

Altså, når du da går fra for eksempel barneskolen, da, og over på ungdomsskolen, så skal du egentlig ikke kunne oppleve det at da går du fra det å ha femteklassenivå på stort sett det meste, og så kommer du dit, og så «ja men her er det åttende klasse, så det må du ta». Det skal de egentlig ikke... For en vanlig elev er det for så vidt greit, han har jo sine papirer med, han òg, men de det er litt ekstra med, der er det ikke litt papirer med. Og hva er poenget da med halvårsrapporter i hytt og gevær, for det er jo en veldig bra vurdering i utgangspunktet. Når du da havner på en ny skole, så kan det nesten virke som at de på en måte skal lære eleven å kjenne på nytt.

For noen foreldre er denne overgangen positiv. Én foresatt fortalte følgende om sitt barns overgang til ungdomstrinnet:

Når hun gikk på barneskolen, opplevde vi en veldig fokusering, evaluering, i forhold til det som hun var svak på. Var veldig fokus på det hun var svak på, liksom. At hun måtte ha ekstra på det, istedenfor på en måte en stimulering på det som hun var sterk på. Og det, for hun så virket det som, hun ble på en måte utbrent. Jeg hadde tenkt at jeg skulle si at overgangen når hun kom til ungdomsskolen her, var veldig stor, for her ble det på en måte mye mer opp til henne. Så hun blomstret veldig opp når hun kom hit, og fikk på en måte en annen type oppfølging da. Hun følte ikke det presset på den måte, men ble på en måte mer litt opp til seg selv. Og hun vokste veldig på det og fikk mye bedre resultater etter det.

På godt og vondt synes det riktig å si at elever, og foresatte, opplever at det sammenhengende opplæringsløpet er preget av mange til dels store overganger, eller sporskifter, der mye i læringsmiljøet endrer seg brått. Konteksten for elevers utvikling er også basert på et sammenfall av alder og intellektuell modning, hvor alle elever anses å ha en lik progresjon (Tessmer & Richmer, 1997). Noen etterlyser også mer planlagte sporvekslinger. Vi erfarer gjennom samtaler med elever og foreldre at skolens og lærernes vurderingspraksiser endrer seg gjennom grunnskolen. Lærerne på småskoletrinnet har mindre vekt på vurdering for læring i fag, og de vektlegger en mer helhetlig vurderingstradisjon. Samtidig erfarer vi at lærere på barnetrinnet har en bevisst holdning til elevinvolvering og bruk av tydelige mål for undervisningen. Basert på utsagn fra elever på 4. og 7. trinn og foreldre til elever på disse trinnene registrerer vi vurderingspraksiser som våre informanter mener involverer elevene (og foreldre) i større grad, og at det i lokale planer og i klasserommet er tydelig hva som er mål for elevens læring. På samme tid understreker vi at elevinvolvering og tydelig mål for læringen ikke er like mye utbredt i skolene, og at det er store forskjeller innad i skolene.

5.4 Hvem er den gode læreren – relasjonskompetanse og fagkunnskap

Elevenes beskrivelse av hva som er en bra skole og en bra lærer, er relativt entydige. En god lærer er en som kan faget hun eller han underviser i (Hattie, 2012), og samtidig er god på formidling og varierte læringsprosesser (Dale, Engelsen & Karseth, 2011). Fagkompetanse er ikke nok, en god lærer må kunne lære bort, slik to elever i ungdomsskolen beskrev:

Og jeg synes at det er kanskje litt dumt, fordi vi har hatt noen lærere som har hatt veldig mye kompetanse, eller vært veldig flink i forskjellige fag, men de har ikke greid og lært det fra seg ordentlig nok. Sånn at vi har i åttende masse i matte og sånt, for da syntes jeg ikke at jeg greide å lære noe særlig fra læreren. Hun var veldig flink, men jeg forsto det ikke.

Når man er lærer, kan man være veldig god i det faget han holder på med, men han trenger ikke å være noe god til å lære bort. Og det er jo generelt på ungdomsskole òg. Du kan være en flink lærer, men hvis du ikke er flink med ungdom, så kommer du ingen vei fordi at det er en nødvendig egenskap på en ungdomsskole, hvor det går ungdommer som er midt i en utviklingsperiode.

Den gode læreren må dessuten mestrer det relasjonelle (Gardner, 2008). En god lærer er den som har tid til den enkelte, og som elevene opplever ser akkurat dem. En ungdomsskoleelev beskrev en god skole slik:

Ja, jeg og synes at det er viktig at det er trivsel, og at alle blir sett av lærerne, fordi at jeg tror ikke det er sånn på alle skolene, og det er kanskje ikke alltid sånn her heller, da. Men jeg synes at det er viktig at en god lærer greier å oppmuntre elevene og se alle sammen. Det tror jeg er viktig.

En annen elev ved samme skole tilføyde:

Jeg synes det er veldig viktig å være trygg på hverandre i klasserommet, sånn at det er ikke farlig, liksom, å rekke opp hånden og svare feil. Og det er jo delvis læreren sin oppgave å sørge for. Men det er jo så klart forskjell, da, på hvor stor innsats lærerne legger i det.

Engasjement fra lærerne er også viktig, sammen med det som vel best kan beskrives som god klasseledelse. To elever i ungdomsskolen beskrev det slik:

Han må være engasjert, han må være interessert i det han står og prater om. Hvis en lærer kommer inn og du merker at han er negativ og egentlig ikke har lyst til å være der, så er det veldig lett for elevene å dette av og ikke gidde å følge med. Det merker jeg hvert fall for meg selv, det er ikke noe artig å sitte og høre på at en lærer står og prater monotont om et tema.

Og at han er engasjert og vil at vi skal prestere det beste, for at de lærerne legger alt i det og vil at vi skal lykkes og gjør det, det blir mye mer motiverende. Og sånn som Ane sa, at vi blir trygge. At vi ikke er redde for å tabbe oss ut eller noe sånt. Det tror jeg er viktig.

En god lærer må se den enkelte, hvilke behov og begrensninger den enkelte har (Postholm, 2012). Dette opplever elever ofte også virker inn på vurderingene. For mange elever er det en forutsetning for god og rettferdig vurdering at læreren kjenner dem og forstår dem. En elev på 7. trinn beskrev det blant annet slik:

Men de gode lærerne, de må også være sånn at de følger med, den vet hvem den og den personen er, hva det gjelder, og så at de vet hvordan nivå det er på, og være litt streng. De kan ikke være altfor snille.

Viktig i denne elevens opplevelse av en god lærer er som vi ser, evnen til å vurdere den enkeltes faglige nivå og å balansere det som oppleves som snill, med nødvendig klasseledelse og disiplin.

Mange elever gir også uttrykk for at de ikke blir like godt sett av lærerne på høyere trinn, at det relasjonelle hadde en annen karakter på lavere trinn. Elever kan være tvisynte i forhold til dette. For noen er det greit at lærerne har et mer avgrenset faglig fokus, men samtidig gir nok noen uttrykk for at de savner de tettere relasjonene til lærere som de hadde på lavere trinn, som det framgår i denne samtalen i et gruppeintervju fra videregående skole:

Den største overgangen som jeg har merket, er at du har ikke samme forhold til lærerne som du hadde før. Lærerne på ungdomsskolen var mer personlige, og du fikk lyst til å snakke med dem.

Du kunne være kompis og sånt.

Ja, kompis. Jeg er jo kompis med lærerne mine på ungdomsskolen, de var dritkule, kan jeg si. Men her er det liksom mer restriktivt, du er ikke helt på godsiden til lærerne.

De er ganske nøytrale til deg, egentlig.

Mangel på motivasjon er noe som også elever i vårt materiale forteller om, som for eksempel denne eleven ved en videregående skole:

Ja, jeg tenker at jeg vil videre, da, og jeg vet selv at hvis jeg skal studere videre, så vet jeg at jeg må få opp snittet mitt, men jeg finner bare ikke motivasjon til det. For jeg er sånn at jeg må på en måte bli pisket litt rundt, om jeg skal si det sånn, da, og på en måte bli kjeftet litt på, bli håndtert litt sånn ... Og ikke ha det så fritt, da.

Generelt etterlyser elevene mer motiverende og utfordrende arbeidsoppgaver. Dette kan forstås som et ønske om mer elevaktivitet, der elevenes situasjonsinteresse skal vekkes. Dette har betydning for utvikling av en mer individuell interesse for faget og for læring generelt (Hidi & Renninger, 2006). Dette forskningsprosjektet undersøker hvordan vurdering kan bidra til økt læring, motivasjon og mestring hos elevene i skolen. Hovedinntrykket er at de elevene vi intervjuet, hadde presise beskrivelser av den gode lærer som tar opp i seg sentrale kompetanseområder som kjennetegner lærerprofesjonen (se Nordenbo mfl., 2008).

5.5 Kriterier og tilbakemeldinger på eget arbeid

Vår studie viser at elever på alle trinn har relativt høy bevissthet om vurdering, selv om fokuset åpenbart blir sterkere desto høyere årstrinn eleven befinner seg på. Elever på alle trinn opplever vurdering som viktig, også for egen læring. De ønsker gode, tydelige tilbakemeldinger på hva de er gode til, hva som kan bli

bedre, og hvordan. Lærerne beskriver også tilbakemeldinger som viktige i læringsprosessene, men som en lærer uttrykte det: Jeg kan ikke gi for mange oppgaver, for veiledningen er for tidkrevende: «ikke hele stiler, det syns jeg tar for lang tid, rett og slett». Elevene viser også til at «noen lærere gir flere oppgaver enn andre».

Elever på fjerde trinn svarte slik på spørsmål om hvordan de fikk tilbakemeldinger fra lærer:

De sier det på en måte til oss, og så kommer de bare og sier bra.

Og så skriver de kanskje nederst på siden av boken «bra jobba, fortsett å øve» eller noe sånt.

De sier bare, når de skriver at vi skal øve mer, så vet vi hva vi skal øve mer på, fordi de skriver det i boka vi holder på med de der tingene i.

Det viser seg at elevene er opptatt av tilbakemeldinger, og at disse skal være knyttet til noe konkret. Innholdsløs ros og kritikk oppleves ikke læringsfremmende. De søker med andre ord tilbakemeldinger som viser at de gjør framskritt (Hattie, 2007), men de rapporterer at de i varierende grad får læringsfremmende tilbakemeldinger. En elev på Vg1 i videregående skole fortalte for eksempel om denne vurderingspraksisen i kroppsøving:

Læreren sa et visst antall armhevinger, tror det var 45, ga karakteren 6, det samme med andre øvelser, som sit-ups og benkpress, tror det var 95 i benkpress, som ga 6, og så hadde hun en skala der antall ble til karakter. Tror hun sa at 20 armhevinger var bestått. Og så springer vi en runde, der fortalte læreren at hun hadde gått gjennom runden og ut fra det vurdert hvor raskt vi burde løpe. Så satte hun karakterer ut fra tiden vi fikk. Tror en tid på litt under 4 minutter ga karakteren 6. Noen andre kriterier fikk vi ikke høre om. Jeg tror ikke innsats teller noe, i alle fall.

Den samme eleven var frustrert over dette og opplevde ikke at det ga noen tilbakemeldinger han hadde nytte av. Det ble åpenbart ikke opplevd som læringsfremmende av denne eleven. Han så det også i kontrast til andre lærere i kroppsøving ved samme skole, som visstnok hadde samtaler med hver enkelt der de sammen kom fram til vurderingskriterier. Også i ungdomsskolen fortalte den samme eleven om en kroppsøvingslærer som opererte med et sett av kriterier, og der innsats ble sagt å skulle telle.

Flere elever i vårt materiale har vært opptatt av at det er viktig å kjenne kriteriene de blir vurdert etter. De ønsker å vite hva de har å strekke seg mot, hva som forventes, og der det er aktuelt, vil de vite hvilken karakter det svarer til. Dette handler om vurderingens reliabilitet. Elever er opptatt av rettferdig,

lik vurdering, de vil ha samme karakter for samme arbeid. Vi har tidligere (Sandvik & Buland, 2013) fortalt hvordan både elever og foreldre kan bli svært engasjert hvis det oppleves som at ulike lærere vurderer samme arbeid ulikt.

Gjennom intervjuer og i observasjon har vi sett hvordan vurderingskriterier på ulike måter blir presentert for elevene. Dette kan være ført opp som læringsmål i ukeplaner, det kan settes opp på tavle i klasserommet og så videre. I mindre grad har vi sett eksempler på at elever involveres i dialog rundt vurderingskriteriene eller selv er med på utforming av kriterier, men dette skjer også.

5.6 Vurdering av eller for læring?

I delrapport 2 (Sandvik & Buland, 2013) har vi vist at mange foreldre/foresatte forstår fenomenet og begrepet vurdering relativt entydig som vurdering av læring. Vurdering for læring, altså den løpende vurderingen og tilbakemeldingen som skjer i en læresituasjon, forstås av mange som læring, noe annet enn vurdering. Dette ser vi nok også hos noen elever. De assosiasjoner som først skapes av ordet *vurdering*, handler om karakterer og om vurdering *av* læring.

Noen opplever helt klart at vurdering handler om å vurdere hvor mye de kan, og at det i liten grad handler om læring. En gruppe elever på 4. trinn ble spurt om hvorfor de fikk vurderinger, og svarte slik:

Jeg tror det er for å sjekke hvor mye man kan, jeg.

Jeg tror ikke det er for å lære.

Bare sjekke at du husker det, liksom. En siste test før vi hopper over til neste kapittel, kanskje.

Felles for mange av elevene vi har intervjuet, er at dette oppleves som lite positivt. Det man ønsker, er vurderinger og tilbakemeldinger som gir dem noe å arbeide videre med, for å bli bedre i faget. En elev i videregående skole forklarte:

Han engelsklæreren er veldig flink til å gi en positiv tilbakemelding og så en tilbakemelding på hvor vi kan forbedre oss, da. Så jeg synes det er ganske bra, for da kan jeg forbedre meg på ...

En annen elev i samme gruppe utdypet ønsker om å få konkrete tilbakemeldinger slik:

Det står med en karakter, og så står det kommentar, om eventuelt, ja, jeg fikk i hvert fall at det var noen verbformer som jeg kunne jobbe litt mer med, og at jeg måtte passe på at jeg skrev riktig, og sånn her. Så det er jo bare å øve seg litt mer på.

Vi observerer altså betydelig grad av variasjon med hensyn til hvor konkrete og læringsfremmende tilbakemeldingene blir opplevd. Denne variasjonen gjelder ikke bare på skolenivå, men også på trinn- og lærernivå innad i skolene.

Å ikke få gode nok begrunnelser på gitt vurdering/karakter er noe som flere i vårt utvalg fortalte om. Dette oppleves som negativt, som for eksempel i denne dialogen i en ungdomsskole:

NTNU: Hvor ofte får dere karakterer som dere syns er helt feil? Som dere syns ikke stemmer?

Sjeldent, syns jeg.

Ja, det er sjeldent at det er karakteren som er feil. Det er ikke det at jeg er misfornøyd med karakteren, men jeg kan være misfornøyd med undervisningen som har ligget i forkant.

Eller vurderingen av det. Når de forklarer oss hvorfor vi fikk det. At ikke på en måte de greier å forklare oss nok. Jeg hadde forstått det hvis de hadde en god forklaring på det, at jeg skjønnte, ok, da forstår jeg. Men når de ikke greier å forklare det, da er det litt sånn, da skjønner du ikke alltid.

Du får en tilbakemelding, så står det ingenting som du bør gjøre bedre, men du får fortsatt ikke 6, liksom.

Da er det liksom, hvorfor fikk jeg ikke det.

Det står ingenting, det står bare veldig bra bla bla bla. Og kjempebra bla bla bla, og så står det karakteren 5, og så ok, hva var det som var i veien med teksten, hvorfor fikk jeg ikke 6?

Ja, og så spør man, og så får man ikke svar, sånn har det vært litt i niende og sånn.

I delrapport 2 (Sandvik mfl., 2012, s. 67 ff.) er samtalestrategier i matematikkrommet beskrevet. Observasjonene understreker at mange av lærerne vektlegger dialog og spørsmål-svar-sekvenser for å få informasjon om elevenes forståelse. Slike aktiviteter kan synliggjøre tenkningen i faget og gi læreren kontinuerlig informasjon om hvor mange av elevene som forstår det faglige innholdet. Lærerens vurderingsarbeid omfatter også «å finne bevis for læring» (Black & Wiliam 1998; Sadler 1998). I samtalene ber

matematikk læreren elevene konkretisere utsagn, aktivere førkunnskap, forsterke ordforråd ved å nevne synonymer sammen eller ved å koble hverdagslige ord til matematiske fagbegreper. Det ble også observert at lærer bruker egen vurdering for å kartlegge elevens forståelse underveis i timen samt stiller metakognitive spørsmål etter læringsaktiviteter. Tilbakemelding blir dermed også en toveis prosess.

5.7 Tilpasset undervisning

Mange elever opplever hverdager i skolen der lærerens tilpasning av undervisningen er svak. Undervisningen følger gjerne kapitlene i læreboka. Vi observerte undervisning der læreren i liten grad anstrengte seg for å tilpasse læringen til elevenes kompetansenivå og inkludering av hendelser som preger deres liv og det samfunnet de er en del av. For å kunne utvikle et bredt repertoar av læringsstrategier anbefales det at elevene får muligheter til å lære i meningsfulle kontekster (Hattie, 2007). De rapporterer om mye høytlesing i hel klasse, der tilpasningen ligger i hvor mye tekst hver enkelt må lese høyt, og om arbeidsplaner som deles inn i tre læringsløp, og som baseres på prinsippet «mer av det samme».

Tilpasset undervisning oppleves av mange elever dessuten som et spørsmål også om det relasjonelle. For at læreren virkelig skal kunne drive tilpasset opplæring, er det viktig at han eller hun kjenner den enkelte elevens forutsetninger. En elev på videregående beskrev det slik, med referanse til overgangen fra ungdomsskole til videregående:

På forrige skolen min, da var vi så liten klasse at du kjente liksom læreren din som din venn, egentlig, mens nå på (skolens navn) så er det litt sånn stort fokus på faget.

Å kunne vurdere eget og medelevers arbeid er et viktig aspekt av elevenes engasjement for økt læringsutbytte (Hattie, 2007). Elevene rapporterer om ulike praksiser på dette feltet. Samtidig fremhever de betydningen av egen vurdering og medelevvurdering for både læring og motivasjon.

Det synes som om elevene i denne studien har svært varierte erfaringer knyttet til vurdering for læring og tilpasset opplæring. Dette gjenspeiler andre funn i denne rapporten. Manglende tolkningsfellesskap omkring vurdering innad i og på tvers av skoler viser et arbeid som er i gang, men som tar tid å internalisere. Det som likevel er verdt å merke seg, er at elevene mener at god vurderingspraksis med tydelige mål og kriterier, der de selv er aktive i egen og andres læringsprosess, er viktig for deres læring, motivasjon og mestring.

Et typisk funn i våre elevintervjuer er at elevene opplever at det er store forskjeller mellom fag og mellom lærere med hensyn til hvordan vurdering praktiseres. Vi har tidligere snakket om at mange skoler mangler et fungerende tolkningsfellesskap rundt vurdering, og at dette gir mange ulike praksiser innenfor samme skole. Dette er noe som elevene observerer, og noe som ofte diskuteres i elevgruppen og mellom elever og foreldre. For mange elever og foreldre blir dette et spørsmål om reliabilitet. Man opplever at ulike lærere vurderer det samme arbeidet ulikt, og dette blir oppfattet som urettferdig.

Vi ser også at mange elever, og foresatte, opplever et skille mellom vurdering av læring og læring. I den kjeden av oversettelser av sentrale begreper og signaler som karakteriserer feltet vurdering, kan vi si at mange elever og foreldre har utviklet en lokal forståelse, eller en lokal teori, der vurdering er forstått som vurdering av læring, og primært ved bruk av karakterer, mens vurdering for læring forstås som noe annet, som «læring» eller undervisning, noe kvalitativt annet enn «vurdering».

Sentrale aktører i feltet forstår altså begreper ulikt. Disse ulikhetene i forståelse vil selvsagt også påvirke den kontinuerlige diskursen rundt vurdering som er viktig for å utvikle feltet (Meyer & Land, 2005).

Det er naturlig å oppfatte vurdering for læring som et nytt paradigme for vurdering, i betydningen et nytt, felles sett av skrevne og uskrevne teorier og metoder som samler profesjonsutøvere innen feltet for vurdering og læring omkring en bestemt praksis (se f.eks. Kuhn, 1962). Vi ser at dette nye paradigmet, som også involverer begrepsbruk og begrepsforståelse, i ulik grad har fått gjennomslag blant skolens ulike aktører: lærere, skoleledere og skoleeiere. Særlig kommer alternative paradigmer til uttrykk i mange elevers og foreldres opplevelser av og fortolkning av fenomenet vurdering.

5.8 Sammenfattet – underveisvurdering og elevers opplevelse av læring

Selv om vi i våre analyser ikke har funnet kvantitative mål på sammenheng mellom underveisvurdering og elevers læring, viser våre intervjuer og observasjoner klart at elever opplever at gode tilbakemeldinger og underveisvurderinger bidrar positivt til læring. Gode tilbakemeldinger er essensielt i både elevers og foresattes opplevelse av hva som gir læring. På mange måter kan man si at underveisvurdering *er* læring, mens vurdering forstås som sluttvurdering.

Samtidig opplever elevene at det er store forskjeller på vurderingspraksis. Dette gjelder selvsagt mellom skoleslag, men også mellom lærere ved samme skole.

Elever er seg ofte svært bevisst slike forskjeller, og aller mest når de opplever at dette gir utslag i ulike vurderinger av samme arbeid. Elever, og foresatte, er primært opptatt av vurderingens reliabilitet. Vurdering skal være rettfærdig, i den forstand at man får samme vurdering av ulike lærere. Samtidig er det altså en klar bevissthet om at undervisvurdering og gode tilbakemeldinger er viktig. For at undervisvurdering skal oppleves som læringsfremmende, viser våre studier at noen betingelser må være til stede:

- Elev og lærer må ha et omforent forhold til vurderingskriterier. Kriteriene for hva som er et godt resultat, må være kjent.
- Vurdering og tilbakemelding må være knyttet til et konkret arbeid, og tiden mellom arbeidet og tilbakemeldingen bør være kort.
- Tilbakemelding må være konkret og peke på hva som er bra, og hva som kan gjøres bedre.
- Tilbakemelding må peke framover og gi eleven klare meldinger om hva som skal til for å levere et bedre resultat. En «bra» eller en tallkarakter gir ikke alltid nødvendig informasjon som kan være av nytte i videre læring.
- Tilbakemeldingen bør være personlig, den enkelte elev bør føle at han eller hun har blitt sett og behandlet individuelt.

Dette samsvarer som vi ser, godt med Wiliam og Thomson (2007) og deres fem strategier for læringsfremmende vurderingspraksis. Det elevene særlig opplever som viktig, er imidlertid de læringsfremmende tilbakemeldingene og de gode dialogene. Som vi har pekt på, er aktiviteter som medelevvurdering relativt lite utbredt i vårt empiriske materiale. Elevene opplever dermed i mindre grad å være ressurser for hverandres læring. Her er det etter vår mening et potensial for videre utvikling av praksis.

6. Ledelse av læreres læring

Gunnar Engvik

I dette kapitlet presenterer vi først to praksisfortellinger. Deretter er det analyser av innholdet i fortellingene i lys av aktuell teori om læring i organisasjonen skole og hvordan lede læreres læring. Både nasjonal og internasjonal forskning viser at læring som skjer på skolen i samarbeid med andre lærere og med en skoleledelse som støtter opp om en sosial læring, er den beste måten for lærere å utvikle sin egen undervisning på, og at dette fører til at elevene lærer enda mer (Postholm, 2012).

6.1 Praksisfortellinger om barneskole som har deltatt i Vurdering for læring

Skolelederen har over flere år arbeidet systematisk med vurdering for læring. Leder har masterutdanning i skoleledelse, og elevmedvirkning i skolen og i undervisningen har vært sentralt i lederens skolekarriere. Skolen deltok i prosjektet «Bedre vurderingspraksis» (2007–2009), og etter den tid har skolen i tette samarbeid med kommunen og en UH-institusjon viet mye tid til vurdering. I første fase var oppmerksomheten på lokalt arbeid med læreplaner, der faglige og sosiale læringsmål og kjennetegn på måloppnåelse ble utformet. Målet med dette arbeidet var å skape felles språk om læring (framfor «gjøring») som kommuniserte med elever og foreldre. Dette var et spennende arbeid som krevde mye tid og mange prosesser i kollegiet.

Arbeidet med halvårsvurderingene skapte imidlertid motstand i kollegiet. Leder ordsetter dette slik: «Jeg har tapt noen slag, men vant krigen!» Skoleleder løfter fram at poenget med skolens vurderingsarbeid var at det skulle fremme læring, og da måtte skolens lærere komme fram til et språk som kommuniserte det på en god måte. Skoleleder nevner at kommunale planer ga klare føringer for at skolene skulle involvere foreldrene i vurderingsarbeid. Involveringen skulle være dialogbasert. Foreldrene i FAU har gitt meget gode tilbakemeldinger på skolens arbeid med vurdering for læring.

Vurdering for læring er godt forankret i hele organisasjonen, fra skoleeiernivå til skole- og teamnivå. Forankringen er «kjempeviktig». Skoleeier støtter aktivt opp om skolens prioriteringer og at utviklingsarbeidet vurdering for læring «styres i front». Sammen med skoleeier uttrykker skoleleder klare forventninger til skolens lærere om hva som er gode vurderingspraksiser.

Oppfølging er et annet stikkord som skoleleder bruker i beskrivelsen av sin ledelse. Å gjenta budskapet på strategiske arenaer i skolen er usedvanlig viktig. Dette innebærer å være tett på, samtidig som lærerne får muligheter (tid) til å bruke sin profesjon til å vurdere seg selv. En viktig del av oppfølgingen er å skape sammenhenger mellom grunnlaget for vurdering for læring, planer fra kommunalt hold og skolens arbeid med vurdering for læring. Disse sammenhengene må kommuniseres og ha en plass i foreldremøter, i elevrådet, i samtalene lærerne har med elevene.

Skoleleder har bidratt til engasjementet i læringsprosessene gjennom et tiltak han benevner som «Elevkonferansen». I løpet av «Elevkonferansen» snakker rektor med alle elevene på skolen om deres læring og utvikling. De prosessene som er knyttet til konferansen, er på mange måter en vurdering av validitetskjedene til lærerne. I tillegg driver skolens leder en problembasert ledelse, og han foretrekker å lede problembasert framfor å ha fokus på detaljer. Ved å spørre lærerne om hvorfor de gjør ting slik de gjør, vekkes også tanker om hvilke dilemma lærerne står overfor i vurdering for læring hos lærerne. Skoleleder mener skolen i dag har en kollektiv forståelse og praksis i vurdering.

Skolen og kommunen har i flere år benyttet eksterne krefter til satsingen. Nettverk har vært bedre med tanke på implementeringen av vurdering for læring enn kurs med eksterne krefter, sier skolelederen. Gjennom kurs og samarbeid i nettverk har skolens leder sørget for at lærerne samhandler i et profesjonelt fellesskap, men denne samhandlingen har avtatt de siste årene. Skoleleder forteller at lærerne etter hvert har blitt dyktige til å bruke andre læringsbaner, som nettressurser og lignende.

Dagens skolelederrolle er mye knyttet til økonomi og administrasjon, og det blir for liten ledertid til å bruke egen pedagogisk kompetanse. Skoleleder savner også noen å diskutere og reflektere sammen med.

Skoleleder forteller at skolens deltakelse i den nasjonale satsingen på vurdering for læring har gitt skoleleder erfaringer med et omfattende skoleutviklingsprosjekt. Skoleleder har kommet tettere på skolens kjerneområde, som er elevens læring og utvikling, men også lærernes undervisning. Skoleleder erfarer hvor viktig det er at lederne evner å utvikle og kommunisere en klar visjon, at de støtter lærernes profesjonelle utvikling, og at skoleledelsen ser betydningen av ekstern kompetanse og støttende nettverk.

6.2 Praksisfortelling fra 1–10-skole som ikke har vært med i Vurdering for læring

Skoleledelsen understreker at det er en stor utfordring å samle kollegiet om en kollektiv forståelse av hva som er god vurderingspraksis og nødvendig vurderingskompetanse, og hva som kjennetegner en god vurderingskultur. Ledelsen nevner at de har fått en del kritikk for innhold av «kursdager» og utviklingsprosjektet fordi «det her gjelder da ikke oss!». Med bakgrunn i en 40-årig tradisjon med forskjellige forståelser av hva som skal inngå i lærernes og skolens vurderingspraksiser, erfarer skolens ledelse at det er vanskelig «å treffe alle lærere» når én gruppe lærere har ansvaret for barnetrinnet og innlæring av lese- og skrivekompetanse, mens andre er opptatt av avsluttende prøver, karakterer og eksamen på 10. trinn.

Ledelsen er kjent med at elevene ved skolen opplever at det er «stor overgang» fra 4. til 5. trinn og fra 7. til 8. trinn. Det er stor forskjell på faglige krav til elevene når de flyttes opp fra 7. til 8. trinn, mener ledelsen, og de finner tilsvarende – om enn i mindre grad – endringer i forventninger og krav fra 4. til 5. trinn. En i ledelsen mener at «det er veldig viktig når vi har en 1–10-skole, at barnetrinnet og også ungdomstrinnet har en viss diskusjon om det rent faglige, slik at det kanskje ikke blir noen overraskelser for eleven, altså at vi dyktiggjør elevene til å gå videre i skolen». Ledelsen hevder at skolekulturen er preget av at lærerne på mellomtrinnet synes lærerne på småskoletrinnet har jobbet for dårlig, og at ungdomsskolelærerne mener lærerne på mellomtrinnet burde gjort en bedre jobb.

Ledelsen har satt av noe tid til kollegialt arbeid av «faglig karakter», og det fortelles at skolens vurderingspraksiser har endret seg de siste årene. Ledelsen har bidratt til å legge føringer på et tydeligere system rundt vurdering for læring. Spesielt fremheves forberedelser til foreldresamtaler og til samtaler med elevene. Lederutfordringene har vært knyttet til halvårsvurderinger for 1.–7. trinn. Skal disse være skriftlige eller muntlige? Ledelsen mener barnetrinnet har de største utfordringene knyttet til halvårsvurderinger. Hvilke skjema skal vi bruke, hva slags forarbeid skal vi gjøre med tanke på foreldresamtalene?

Ledelsen forteller at lærerne i kommunen har vært gjennom en «vanvittig prosess» hva gjelder lokalt arbeid med læreplaner. De planene som ble utarbeidet, er ikke i bruk i dag. Skolen støtter seg til læreplanverket i Kunnskapsløftet, og ifølge ledelsen er det er en svak sammenheng mellom lokale planer og lærerens vurderingsarbeid.

Ledelsen bruker mest «ungdomsskolen» og noen ganger «ungdomstrinnet» når de omtaler 8.–10. trinn. Ledelsens beskrivelser av lærernes vurderingspraksiser på ungdomstrinnet er i høy grad koblet til summativ vurdering, i form av

standpunkt og eksamen. Én leder hevder: «Vi måler jo alt i karakterer på ungdomstrinnet, det handler om å få en karakter.»

Ledelsen forteller at skolen har «litt tradisjoner til å hoppe av» prosjekter og tiltak. Resultatet av dette og at ikke alle tiltak har vært like godt forankret på skolen, har ført til en «backlash» i flere sammenhenger. Noen ganger har slike erfaringer «hemmet og litt ødelagt» grunnlaget for skoleutvikling. Ledelsen ser på dette som en svakhet ved skolekulturen, som har ført til at resultatene ikke har blitt så bra. Skolelederne forteller at enkeltlærere «tar seg fram på egen hånd» på nettet og søker aktuelle nettsider og andre kilder for å få impulser.

6.3 Analyse av praksisfortellinger om ledelse av skoler

I denne delen vil vi se nærmere på hvordan skoleledernes praksiser og forståelser kommer til uttrykk.

Fortellingene i det foregående delkapitlet presenterer ulike praksiser. Vi ser det kanskje tydeligst i forankringen av arbeidet med vurdering for læring. Mens den ene rektoren forteller om ulike faser i forankringen av utviklingsarbeidet, erkjenner lederteamet i den andre praksisfortellingen at skolen har tradisjon for «å hoppe av utviklingsprosjekter». Rektoren i 1–7-skolen kan vise til et strategisk arbeid som innbefatter forankring til politisk nivå (skoleeier), i tillitsapparatet, i lærerkollegiet og til elever og foreldre. Lederteamet i den andre skolen gir klart uttrykk for at deres skole har en tradisjon og kultur som er lite kollektivt orientert. Lederteamet vedgår at utviklingsprosjekter i for liten grad har vært forankret i skolen som organisasjon, men at enkeltlærere gis muligheter til «å dra på kurs».

Praksisfortellingene speiler på mange måter rektors og lederteamets strategier og arbeidsformer. Et mangeårig samarbeid med et kompetansemiljø har vært viktig for rektor og hennes utvikling og forståelse av rollen, men også for læringen mellom lærere. Strategiske tiltak som favner lærernes profesjonsutvikling, har etter rektors erfaringer bidratt til læring i organisasjonen.

Et annet trekk som kan løftes fram, er erfaringer fra arbeidet med lokale planer. Lederteamet i 1–10-skolen erfarer at det store aldersspranget, sammen med ulike vurderingstradisjoner på barnetrinnet og på ungdomstrinnet, gjør det vanskelig å samle hele kollegiet om en kollektiv forståelse av vurdering for læring. I fortellingen fra denne skolen får vi et klart inntrykk av at arbeidet med lokale planer ble «en vanvittig prosess». Rektoren i 1–7-skolen vedgår også at dette arbeidet kan skape spenninger innad i kollegiet. «Jeg tapte noen slag, men vant krigen», forteller rektoren. Selv med en uttalt motstand fra deler av

lærergruppen, ble lokale planer med vekt på læring og mindre på «gjøring» utviklet. Hovedmålet med planene var at de skulle kommunisere innad i skolen og utad til eiere, foreldre og offentligheten.

Forståelsen av elevinvolvering er til stede i praksisfortellingen fra 1–10-skolen. Planlegging, gjennomføring og vurderinger av samtaler med foreldre og elever har vært lederteamets strategi. Noe av bakgrunnen for dette utviklingstiltaket var at overgangene fra 4. til 5. trinn og fra 7. til 8. trinn av elevene ble betegnet som store. Forståelsen av elevinvolvering er svært tydelig i fortellingen fra 1–7-skolen. Rektors strategi er å la elevenes stemmer bidra til en skolebasert evaluering og utvikling.

Samlet gir praksisfortellingene et innblikk i skoleledelse rundt vurdering for læring, som peker mot at skoleledere «utgjør en forskjell». Ser en nærmere på forskning utført i skoler over hele verden fra 1980-tallet og fram til 2005, hevder Hallinger (2005) at den bekrefter hva «aktører i skolen» og foreldre alltid har kjent til: Skoleledelse utgjør en forskjell i skolenes evne til utvikling (Hallinger, 2005). Dette har ført til økt fokus på å identifisere hvordan skoleledere «utgjør en forskjell».

Robinson, Lloyd og Rowe (2008) fant at jo mer skolelederne fokuserte på relasjoner og oppgavene knyttet til skolens kjernevirksomhet, som er undervisning og læring, jo større ble skoleleders innflytelse på elevens resultater. Det argumenteres også med at i skoler hvor lærerne rapporterer at skoleledelsen er opptatt av og bruker tid på kjernevirksomheten undervisning og læring, har elevene bedre resultater enn i tilsvarende skoler der lærerne rapporterer om mindre oppmerksomhet på kjernevirksomheten fra skoleledelsens side (Robinson, Lloyd & Rowe, 2008).

Rektor ved 1–7-skolen leder skolen ved å delta i lærernes læring og utvikling. Vi kan forklare rektorens form for ledelse ut fra teori om *instructional leadership*, som har fem lederdimensjoner: å lede gjennom å fremme og delta i lærernes læring og utvikling, å etablere mål og uttrykke forventninger, å planlegge, å evaluere undervisning og arbeidet med læreplaner samt strategisk ressursbruk og å utvikle et støttende læringsmiljø (Engvik, 2012).

Robinson (2009, 2010) peker på andre sider av *instructional leadership* som er relevante i den norske skolekonteksten og for norske skoleleder. Gitt at en skoleleder velger å lede sin skole ved å fremme og samtidig delta i lærernes læring og deres profesjonelle utvikling, vil dette kreve en skoleleder som i første rekke har faglig kompetanse som gjør at hun/han kan diskutere lærernes oppgaver innen fag og fagdidaktikk, og en skoleleder med nødvendig innsikt i og forståelse av de utfordringer lærerne møter i sin hverdag (Robinson, Hohepa & Lloyd, 2009; Robinson, 2012).

I praksisfortellingen fra 1–10-skolen understreker lederteamet at det er en stor utfordring å samle kollegiet om en kollektiv forståelse av hva som er god vurderingspraksis og nødvendig vurderingskompetanse, og hva som kjennetegner en god vurderingskultur. Denne vurderingen av praksiser og forståelser i lærerkollegiet kan forstås ut fra norske forskningsbidrag som peker på at satsingen på vurdering for læring hviler på en detaljert og styrende forskrift og svært åpne læreplaner, som skal fortolkes og bygges ut lokalt (Sivesind & Backmann, 2013). Denne koblingen mellom et lovverk med forskrift for vurdering og læreplaner som skal utformes lokalt i grunnskole og videregående skole, skaper utfordringer også på skoleledernivå.

Her møter vi igjen utfordringer og problemstillinger knyttet til skoleleders kompetanse og kapasitet.

I den tette oppfølgingen av lærernes undervisning, som er til stede i praksisfortellingen i 1–7-skolen, og som ligger tett opptil teorien *instructional leadership*, er det en underliggende forståelse av at de fleste lærerne utvikler sin undervisning gjennom nøye utformet støtte og hjelp. Dette synet på lærernes læring og utvikling står i sterk kontrast til tradisjonen med en autonom lærerrolle i vårt land, som er kartlagt i flere internasjonale studier (se Vibe, Aamodt & Carlsten, 2009). Praksisfortellingen fra 1–7-skolen viser at rektor har en høyere grad i skoleledelse. Dette medfører at skolen har en godt utviklet læringskultur basert på samhandling mellom skoleleder og lærere og mellom lærere. Samhandling mellom aktørene i skolen blir slik kjernen i effektiv skoleledelse, og det vil kreve at skolelederen har en bred kompetanse, på alt fra klasseromsobservasjoner til datainnsamlinger, og kunnskap om lærernes læring og utvikling og gode relasjoner og kommunikasjonsevner.

7. Kvantitative analyser – hva skaper god vurderingspraksis?

Gunnar Engvik

I denne rapporten ser vi nærmere på hva som bidrar til god vurderingspraksis i norsk grunntidning. En survey med et utvalg lærere som underviser i fagene engelsk, norsk, kroppsøving og matematikk i grunnskolen og i videregående skole, er grunnlaget for analysene. Det er gjort rede for datagrunnlag, utvalg og instrumentene i surveyen i kapittel 2.2. Faktoranalyser, analyser av reliabilitet og utvikling av indikatorer i kapittel 2.2 danner utgangspunktet for analyser og diskusjoner i dette kapitlet. Statistikkprogrammet IBM SPSS Statistics 20 er brukt som programvare for analysene.

7.1 Analyser og diskusjon

Vi er særlig opptatt av å kartlegge og analysere faktorer som påvirker lærernes vurderingspraksis. Vi vil i tillegg undersøke om det i fagene engelsk, norsk, kroppsøving og matematikk er ulike faktorer som har betydning for lærernes vurderingspraksis.

Enkeltvariabler som inngår i indikatoren «Vurderingspraksis», er: «Jeg benytter undervisvurdering for å fremme elevenes videre læring», «Jeg benytter undervisvurdering for å kartlegge elevenes fremgang i fag», «Elevene mine får løpende tilbakemeldinger på arbeidet sitt», «Undervisvurdering er en integrert del av min undervisning», «Min vurderingspraksis når det gjelder halvårsvurdering, er god», «Min vurderingspraksis når det gjelder sluttvurdering, er god» og «Mine elever får tilbakemeldinger undervis som gir dem mulighet til å forbedre arbeidet før endelig innlevering». Svaralternativene er fra 1–3, hvor 1= i liten grad, 2 = i noen grad og 3= i høy grad. Respondentene fikk muligheten til å svare «Vet ikke» på disse påstandene (se kap.2.2.3). Følgende forskningsspørsmål har dannet grunnlaget for analysene:

1. Hvilke faktorer forklarer god vurderingspraksis
2. Hvilke faktorer forklarer god vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk?

Forskningsspørsmålene er undersøkt gjennom deltakernes selvrappotering i en kvantitativ survey vinteren 2012. Det er metodisk utfordrende å måle læreres vurderingspraksis. Eksempler på problemstillinger som forskningsgruppen stilte seg, og som ligger til grunn for utformingen av spørsmålsbatteriene i

surveyen, er: I hvilken grad er vurderingen målorientert og kriteriebasert? Hvordan anvendes klasseromsdialoger som vurderingsverktøy? I hvilken grad får elevene tilbakemelding på arbeidet de gjør, underveis i prosessen? I hvilken grad brukes egenvurdering som vurderingsverktøy? I hvilken grad brukes medelevvurdering som vurderingsverktøy? Forskjeller/likheter i vurderingspraksis gjennom utdanningsløpet. I hvilken grad anvendes digital teknologi (LMS, programvare) som del av vurderingen?

Metodiske utfordringer knytter seg også til den lave svarprosenten fra lærere i grunnskolen (24 prosent). Måletidspunktet kan være én forklaring på at tre av fire grunnskolelærere i utvalget ikke svarte på undersøkelsen. Selv med lav svarprosent er vår vurdering at undersøkelsen blant lærere i grunnskolen har god nøyaktighet og gir forholdsvis stor reliabilitet. Ut fra at lærernes svar på påstander om vurderingspraksis kan knyttes til måletidspunktet, til måten dette måles på, og til hvilke kontekstuelle sammenhenger lærerne var i da de svarte på surveyen, bør konklusjonene fra analysene tolkes med varsomhet.

7.2 Resultater

Utgangspunktet for denne delen er å analysere faktorer som samlet forklarer god vurderingspraksis i fagene, samt analysere faktorer som forklarer god vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk. Først presenteres statistikk hvor vi har testet om bakgrunnsvariabler kan forklare variansen i indikatoren «Vurderingspraksis». Her benyttes analyser av varians (ANOVA). Så følger analyser av bakgrunnsvariabler som forklaringsfaktorer, og deretter presenteres bivariat statistikk i form av korrelasjonsanalyser og til sist multivariat statistikk, regresjonsanalyser.

Indikatoren «Vurderingspraksis» og de øvrige indikatorene som er utviklet for hele materialet, er benyttet i multiple regresjonsanalyser og analyser av korrelasjon. Når det gjelder datagrunnlaget og utvalg i surveyen og utvikling av indikatorer, viser vi til kap. 2.2 *Kvantitativ breddesurvey*.

7.2.1 Bakgrunnsvariabler som forklaringsfaktorer

En enveis variansanalyse (ANOVA) viser at det er signifikante forskjeller for 3 av 32 bakgrunnsvariabler mellom de ulike kategoriene i bakgrunnsvariablene når det gjelder score på variabelen «Vurderingspraksis». Disse er: «Skoleslag» ($p = 0,000$), «Fordypning i fag» ($p = 0,000$) og «Type lærerutdanning» ($p = 0,000$). I de følgende avsnitt viser vi forskjellen i skåre på variabelen «Vurderingspraksis» mellom de forskjellige kategoriene i disse variablene.

I framstillingen av de målte skårverdier i tabellene 5, 6 og 7, er indikatoren «Vurderingspraksis» normert til en tretrinnskala, som går fra 1 «i liten grad», 2 «i noen grad», 3 «i stor grad».

7.2.2 Type lærerutdanning og vurderingspraksis

Bakgrunnsvariabelen «Type lærerutdanning» har sju (7) svaralternativer, men i denne oversikten tar vi med de to utdanningene som er mest representert i surveyen. Hensikten med analysen er å vise om det er en signifikant forskjell i gjennomsnittsscorene for indikatoren «Vurderingspraksis», mellom lærer med allmennlærerutdanning og lærere med fagutdanning fra universitet/høgskole og praktisk-pedagogisk utdanning (PPU), og om forskjellene er signifikante.

Tabell 5: Gjennomsnittsscore på indikatoren "Vurderingspraksis"³ fordelt på "Type lærerutdanning"

Type lærerutdanning	N	Gjennomsnittsscore på «Vurderingspraksis» med 99 % konfidensintervall i parentes
Allmennlærer	248	2,19 (2,15-2,24)
Fagutdanning fra UH med PPU	310	2,46 (2,42-2,51)
Total	558	2,34 (2,30-2,38)

Gjennomsnittsscoren for både «Allmennlærere» og lærere med «Fagutdanning fra UH med PPU» på 2,34, ligger en del over midtpunktet i skalaen (2) for gjennomsnittsscore på indikatoren «Vurderingspraksis». For «Allmennlærer» er gjennomsnittsscoren 2,19 nærmere midtpunktet sammenlignet med lærere som har «Fagutdanning fra UH med PPU», som har gjennomsnittsscoren 2,46. Ut fra dette plasseres gjennomsnittsscoren for lærere med «Fagutdanning fra UH med PPU» mellom svaralternativene 2 «i noen grad» og 3 «i stor grad», mens gjennomsnittsscoren for «Allmennlærere» ligger noe over svaralternativet 2 «i noen grad».

Konklusjonen er at bakgrunnsvariabelen «Type lærerutdanning» bidrar til en signifikant forskjell i lærernes vurderingspraksis. Tabell 5 viser at gjennomsnittsscoren på «Vurderingspraksis» for «Allmennlærer» og «Fagutdanning fra UH med PPU» er forskjellig. Målt på 99 % konfidensintervall er det en signifikant forskjell mellom skårene for disse to gruppene som inngår i bakgrunnsvariabelen «Type lærerutdanning».

³ Skala for «Vurderingspraksis» er: 1 = «i liten grad», 2 = «i noen grad», 3 = «i stor grad».

7.2.3 Fordypning i fag og vurderingspraksis

Det er en signifikant sammenheng mellom indikatoren «Vurderingspraksis» og bakgrunnsvariabelen «Fordypning i fag».

Tabell 6: Gjennomsnittsscore på indikatoren "Vurderingspraksis"⁴ fordelt på "Fordypning i fag"

Fordypning i fag	N	Gjennomsnittsscore på «Vurderingspraksis» med 99 % konfidensintervall i parentes
10 vekttall / 30 studiepoeng	131	2,20 (2,095-2,284)
20 vekttall / 60 studiepoeng	226	2,35 (2,286-2,420)
60 vekttall / 180 studiepoeng	127	2,51 (2,430-2,601)

Bakgrunnsvariabelen «Fordypning i fag» bidrar til en signifikant forskjell i lærernes vurderingspraksis. Dette er vår konklusjon. Ved å sammenholde skårene for tre av seks svaralternativer for faglig fordypning (10 vekttall / 30, 20 vekttall / 60 studiepoeng, og 60 vekttall / 180 studiepoeng) er det på 99 prosent konfidensintervall ikke overlappende skårer mellom disse tre svaralternativer i surveyen (se tabell 6). Vår konklusjon er at lærere med fra ett (1) til tre (3) års fordypning i faget skårer signifikant høyere på vurderingspraksis enn lærere som har inntil ½ års faglig fordypning.

7.2.4 Skoleslag og vurderingspraksis

Bakgrunnsvariabelen «Skoleslag» har en signifikant betydning i analysene av indikatoren «Vurderingspraksis». Den tredje analysen av gjennomsnittsscore på indikatoren «Vurderingspraksis» har bakgrunnsvariabelen «Skoleslag» med fire skoleslag som utgangspunkt (se tabell 7).

Tabell 7: Gjennomsnittsscore på indikatoren "Vurderingspraksis"⁵ fordelt på "Skoleslag"

Skoleslag	N	Gjennomsnittsscore på «Vurderingspraksis» med 99 % konfidensintervall i parentes
Barnetrinn (1.-7. trinn)	167	2,13 (2,06-2,21)
Kombinert b/u. skole	67	2,25 (2,12-2,37)
Ungdomsskole	160	2,40 (2,32-2,48)
Videregående skole	238	2,47 (2,41-2,54)
Total	623	2,34 (2,30-2,38)

Vi konkluderer med at bakgrunnsvariabelen «Skoleslag» bidrar til en signifikant forskjell i lærernes vurderingspraksis. Ved å sammenholde gjennomsnittsscore på «Vurderingspraksis» til lærere på ungdomstrinn (2,40)

⁴ Skala for «Vurderingspraksis» er: 1 = «i liten grad», 2 = «i noen grad», 3 = «i stor grad».

⁵ Skala for «Vurderingspraksis» er: 1 = «i liten grad», 2 = «i noen grad», 3 = «i stor grad».

og i videregående skole (2,47), med lærere som arbeider på barnetrinnet (gjennomsnittsscoren er 2,13), kan vi ut fra analyser på 99 % konfidensintervall si at lærernes gjennomsnittsscore på «Vurderingspraksis» er signifikant forskjellig mellom «Barnetrinn» og «Ungdomsskoler»/»Videregående skole». Forskjellen er størst mellom lærere på barnetrinn og videregående skole. Vi finner i tillegg at det er betydelige forskjeller i skårene på indikatoren «Vurderingspraksis» mellom lærere på barnetrinnet og lærere på ungdomsskoler. Forskjellen er signifikant på 99 % konfidensintervall-nivå (se tabell 7).

Funnene i tabellene 5, 6 og 7 viser noen interessante sammenhenger. Lærere med allmennlærerutdanning og som arbeider i barnetrinnet, har mindre faglig fordypning enn lærere i videregående skole. Det viser våre analyser. Vi ser at type lærerutdanning og grad av fordypning i fag har signifikant betydning for lærernes selvrapporing av god vurderingspraksis. Ut fra dette er de signifikante forskjellene i vurderingspraksis også knyttet til skoleslag ettersom tre av fire lærere i barnetrinnet og i 1–10-skoler er allmennlærere og rundt 80 prosent i videregående skole har fagutdanning fra universitet og høgskole og praktisk-pedagogisk utdanning. Dette er ett av hovedfunnene i våre analyser.

7.3. Hvilke faktorer forklarer god vurderingspraksis?

7.3.1 Analyser av samsvar og kommentarer

Med bakgrunn i utviklingen av indikatoren «Vurderingspraksis» og 13 andre indikatorer for denne studien, er det interessant å analysere samvariasjonen mellom «Vurderingspraksis» og de øvrige indikatorene og mellom disse. Til dette er det benyttet en bivariat analyse, se tabell 8.

Tabell 8: Korrelasjonstabell for alle indikatorer som er utviklet

Indikatorer	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Vurderingspraksis	-	,423**	,361**	,286**	,278**	,230**	,212**	,204**	,170**	,110**	,103*	,061	,038	-,031
2. Egen vurderingskompetanse		-	,151**	,254**	,546**	,053	-,051	,093*	,231**	,076	,036	,245**	-,044	,050
3. Lærerens tilbakemeldinger			-	,153**	-,011	,486**	,347**	,232**	,101*	,193**	,090*	-,065	,109**	-,067
4. Skolens vurderingskultur				-	,185**	,210**	,095*	,111**	,544**	,181**	,108**	,155**	,153**	,159**
5. Krevende arbeid					-	-,071	-,120**	,084*	,213**	,043	,087*	,413**	-,022	,069
6. Elevinvolvering						-	,342**	,175**	,081*	,165**	,154**	-,045	,179**	-,095*
7. Elevens læring							-	,214**	,065	,140**	,027	-,036	,064	-,023
8. Undervisvurdering								-	,028	,269**	,243**	-,063	,078	,031
9. Skoleledelse									-	,107**	,011	,228**	,165**	,369**
10. Planer										-	,092*	,051	,038	,036
11. Forskrift											-	-,104*	-,021	-,120**
12. Hindringer												-	,086*	,132**
13. Deltakerbaner													-	,162**
14. Pådrivere														-

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Indikatoren «Vurderingspraksis» har en signifikant korrelasjon til 9 av 13 indikatorer: «Egen vurderingskompetanse», «Lærerens tilbakemeldinger», «Skolens vurderingskultur», «Krevende arbeid», «Elewinvolvering», «Elevens læring», «Undervisvurdering», «Skoleledelse» og «Planer» (se tabell 9). Det er en svak signifikant sammenheng mellom «Vurderingspraksis» og indikatoren «Forskrift», men ingen pålitelig korrelasjon mellom indikatoren «Vurderingspraksis» og indikatorene «Hindringer», «Deltakerbaner» og «Pådrivere».

Flytter vi blikket til den andre indikatoren i tabell 8, «Egen vurderingskompetanse», har indikatoren en sterk sammenheng med «Vurderingspraksis», men også til fem (5) andre indikatorer: «Lærerens tilbakemeldinger», «Skolens vurderingskultur», «Krevende arbeid», «Skoleledelse», og «Hindringer». Korrelasjonstabellen over viser at indikatorene «Egen vurderingskompetanse» og «Krevende arbeid» har en meget sterk korrelasjon ($,546^{**}$).

Indikator nr. 3 «Lærerens tilbakemeldinger» har en signifikant sammenheng med «Elewinvolvering» ($,486^{**}$) og «Elevens læring» ($,347^{**}$), men også til indikatorene «Vurderingspraksis», «Egen vurderingskompetanse», «Skolens vurderingskultur», «Undervisvurdering», «Planer» og «Deltakerbaner».

Indikatoren «Skolens vurderingskultur» knytter seg signifikant til 12 av 13 indikatorer i korrelasjonstabellen (se tabell 8): «Vurderingspraksis», «Egen vurderingskompetanse», «Lærerens tilbakemeldinger», «Skolens vurderingskultur», «Krevende arbeid», «Elewinvolvering», «Undervisvurdering», «Skoleledelse», «Planer», «Forskrift», «Hindringer» og

«Deltakerbaner». Det er en svak signifikans mellom indikatorene «Skolens vurderingskultur» og «Elevens læring».

Vi har så langt vist at indikatorene «Egen vurderingskompetanse», «Lærerens tilbakemeldinger» og «Skolens vurderingskultur» til sammen korrelerer signifikant med alle studerte variabler i matrisen. Ut fra dette ble alle variabler i matrisen tatt inn i den multiple regresjonsanalysen.

Den signifikante korrelasjon vi finner mellom indikatorene «Skolens vurderingskultur» og «Skoleledelse» ($,544^{**}$), kan forklares ut fra forskning som peker på at skolens ledelse (rektor) er ansvarlig for å respondere på eksterne krav, i dette tilfellet god vurderingspraksis, samtidig som ledelse må ha en høy grad av indre ansvarlighet (Elmore, 2008). Rektorer som utøver indre ansvarlighet, bidrar til å utvikle et profesjonelt læringsfellesskap på skolen, i vår sammenheng skolens vurderingskultur. Studier viser at det kreves en aktiv og støttende ledelse for å legge til rette for profesjonelle læringsfellesskaper der det er fokus på læring på alle nivå (Emstad, 2014; Robinson, Hohepa & Lloyd, 2009).

7.3.2 Analyser av signifikante faktorer

I denne delen har vi gjort regresjonsanalyser av indikatorer med den avhengige variabelen «Vurderingspraksis», de 13 indikatorene i tabell 8 og to bakgrunnsvariabler («Type lærerutdanning» og «Fordypning i fag»). De 13 indikatorene som inngår i regresjonsanalysen, er framstilt med bakgrunn i en eksplorerende faktoranalyse (se vedlegg 3). Indikatorene er utviklet på grunnlag av analyser av indre konsistens (Cronbachs alfa).

Regresjonsmodell

For å se hva som påvirker indikatoren «Vurderingspraksis», har vi gjennomført en lineær regresjonsanalyse. Vi har utviklet én regresjonsmodell (se tabell 9) som er brukt på fem forskjellige datasett. Modellen er utviklet med bakgrunn i 13 uavhengige indikatorer i tabell 8, samt to bakgrunnsvariabler («Type lærerutdanning» og «Fordypning i fag»). I datasett 1 er det svar fra 520 lærere. Datasett 2 representerer svar fra 112 lærere som underviser i engelsk, datasett 3 representerer svar fra 197 norsklærere, mens datasett 4 har svar fra 65 respondenter som underviser kroppsøving. I datasett 5 er det svar fra 137 matematikklærere. Hensikten er å bidra til å svare på forskningsspørsmålene i dette kapitlet. Modellen representerer en regresjonsanalyse som omfatter fagene i studien (engelsk, norsk, kroppsøving og matematikk).

Tabell 9: Lineær regresjonsanalyse av "Vurderingspraksis"⁶ etter signifikante forklaringsfaktorer og fag

Indikatorer	Alle fag		Engelsk		Norsk		Kroppsøving		Matematikk	
	B-verdi	Beta	B-verdi	Beta	B-verdi	Beta	B-verdi	Beta	B-verdi	Beta
Egen vurderingskompetanse ⁷	,056***	,353***	,041***	,276***	,059***	,365***	,085***	,469***	,052***	,338***
Lærerens tilbakemeldinger ⁸	,233***	,234***	,203***	,223***	,156**	,155**	,258**	,241**	,370***	,367***
Type lærerutdanning ⁹	,134***	,179***	,209***	,312***	,099**	,135**	-	-	,146***	,195***
Elevens læring ¹⁰	,118***	,155***	-	-	,127***	,164***	,201**	,242**	,138***	,194***
Fordypning i fag ¹¹	,021**	,078**	-	-	-	-	-	-	-	-
Skolens vurderingskultur ¹²	-	-	,092**	,165**	-	-	-	-	-	-
Forklart varians (Justert R ²)	,367		,403		,295		,359		,477	
Gyldig N	520		112		197		65		137	

Signifikansnivå: *** = P < 0,01; ** = < 0,05; * = < 0,1.

Regresjonsanalysen av hele datasettet (alle fag) viser at fem av seks indikatorer har en signifikant betydning for variansen til den avhengige variabelen «Vurderingspraksis». Andelen forklart varians (Justert R² er 0,367) viser at de uavhengige indikatorene forklarer til sammen 36,7 prosent av den totale variansen i «Vurderingspraksis». Forklaringsprosenten på 36,7 er god (se tabell 9).

Den høyeste forklaringskraften finner vi for indikatorene «Egen vurderingspraksis» og «Lærerens tilbakemeldinger». På mange måter er indikatorene knyttet til lærerens vurderingskompetanse, som konkretiseres i lærerens vurderingspraksis. Indikatorene representerer viktige sider ved god vurderingspraksis, som kompetanse og trygghet i arbeidet med vurdering og hvordan lærerens vurdering bidrar til at eleven forstår hva hun/han skal lære, og om læreren gir råd om hvordan eleven skal forbedre seg.

Når vi ser indikatorene «Egen vurderingspraksis», «Lærerens tilbakemeldinger» og «Elevers læring» under ett, bidrar disse til å forklare hoveddelen av god vurderingspraksis. Samtidig representerer disse indikatorene underliggende faktorer som kan videreutvikles i et utviklingsprosjekt som vurdering for læring.

Et annet interessant funn i analysen for alle fag, er at bakgrunnsvariablene «Type lærerutdanning» og «Fordypning i fag» har en signifikant betydning i forklaringen av god vurderingspraksis. Sammenlignet med de indikatorene som

⁶ Skala for «Vurderingspraksis» er: 1 = «i liten grad», 2 = «i noen grad», 3 = «i stor grad».

⁷ Skala for «Egen vurderingskompetanse»: 1= helt uenig, 2= uenig, 3= enig, 4 = helt enig.

⁸ Skala for «Lærerens tilbakemeldinger»: 1= i liten grad, 2= i noen grad, 3 = i stor grad.

⁹ Skala for «Type lærerutdanning»: 1= Allmennlærerutdanning», 2= Fagutdanning ved UH med PPU.

¹⁰ Skala for «Elevers læring»: 1= helt uenig, 2= uenig, 3= enig, 4 = helt enig.

¹¹ Skala for «Fordypning i fag»: 1= ingen, 2= mindre enn 30 stp., 3= minst 30 stp., 4= minst 60 stp., 5= minst 120 stp., 6= minst 180 stp.

¹² Skala for «Skolens vurderingskultur»: 1= helt uenig, 2= uenig, 3= enig, 4 = helt enig.

er nevnt i avsnittet over, er det på mange måter vanskeligere å påvirke lærernes utdanning og faglige fordypning innenfor rammene av prosjektet Vurdering for læring.

Samlet sett kan de fem indikatorene som belyser god vurderingspraksis i tabell 9 knyttes til samlebegrepet *kompetanse*, med indikatorene «Egen vurderingskompetanse», «Fordypning i fag» og «Type lærerutdanning», og til en dialogbasert undervisning representert ved indikatorene «Lærerens tilbakemeldinger» og «Elevens læring».

Det første forskningsspørsmålet i dette kapitlet har oppmerksomheten på faktorer som forklarer god vurderingspraksis. Vi oppsummerer første del med følgende konklusjoner:

- God vurderingspraksis kan forklares med bakgrunn i fem (5) indikatorer. Disse er «Egen vurderingskompetanse», «Lærerens tilbakemeldinger», «Type lærerutdanning», «Elevens læring», «Fordypning i fag».
- «Type lærerutdanning» bidrar signifikant til å forklare variansen i «Vurderingspraksis». Sammenlignet med lærere som har fagutdanning på universitet/høgskole og praktisk-pedagogisk utdanning, svarer lærere med allmennlærerutdanning signifikant forskjellig (lavere) på «Vurderingspraksis».
- Lærernes «Fordypning i fag» er en signifikant indikator i analysene av variansen i «Vurderingspraksis». Vi finner at lærere med fra ett til tre års utdanning i faget oppgir en bedre vurderingspraksis enn lærere med et halvt års utdanning i faget.

I engelsk er det spesielt indikatorene «Egen vurderingskompetanse», «Lærerens tilbakemeldinger» og «Skolens vurderingskultur» som bidrar mest til å forklare god vurderingspraksis. Ser vi nærmere på bakgrunnsvariabelen «Type lærerutdanning», har denne indikatoren størst betydning for faget engelsk, sammenlignet med de øvrige fagene i tabell 9. Det er den indikatoren som har sterkest effekt for engelsk (høyest standardisert beta), og betaværdien (B) er høyere for engelsk enn for de andre fagene. Fagutdanning i engelsk fra universitet/høgskole med PPU har en betydning for «Vurderingspraksis» i faget. Når «Type lærerutdanning» har en mer framtrædende rolle i å forklare vurderingspraksis i engelsk enn i de øvrige fagene, kan det forklares ut fra to funn i FIVI-delrapport 2 (Sandvik og Buland, 2013). Med bakgrunn i observasjoner og samtaler med lærere som underviser i faget engelsk, blir det i rapporten løftet fram at det er både en faglig og en metodisk utfordring å

utvikle lokale læringsmål, arbeidsoppgaver og vurderingsformer ut fra kompetansemål i læreplanen. Det andre er mangelen på formell kompetanse i engelsk blant lærere som underviser i faget. Denne kombinasjonen av utfordringer og ulik kompetanse til lærere som underviser i faget, kan forklare at indikatoren «Skolens vurderingskultur» har betydning for «Vurderingspraksis» i engelsk.

Konklusjonen er indikatorene «Egen vurderingskompetanse», «Lærerens tilbakemeldinger» og «Skolens vurderingskultur», bidrar signifikant til å forklare god vurderingspraksis i engelsk. «Skolens vurderingskultur» har bare en signifikant betydning for engelsk. Når det gjelder «Lærerens tilbakemeldinger» og «Egen vurderingskompetanse», finner vi større effekt i noen av de andre fagene. Når det gjelder «Lærerens tilbakemeldinger», er beta-verdien (B) for engelsk 0,20. For norsk er den 0,16, mens den for kroppsøving er hele 0,37.

Indikatorene «Egen vurderingskompetanse», «Lærerens tilbakemeldinger» og «Skolens vurderingskultur» representerer underliggende faktorer som kan videreutvikles i et utviklingsprosjekt som vurdering for læring. Det er vanskeligere å påvirke lærernes utdanning og formelle kompetanse i faget innenfor rammene av prosjektet Vurdering for læring.

Dersom vi ser på språkfagene norsk og engelsk under ett, er vurderingspraksis i norsk mer rettet mot indikatoren «Elevens læring». Dette kan tyde på at lærere i norsk i større grad enn lærere i engelsk, knytter forståelsen av «Vurderingspraksis» til at eleven når kompetansemålene og at deres dialog med elevene bidrar til at de lærer bedre. Vi kan se norsklærernes vektlegging av en mer elevsentrert vurderingspraksis ut fra funn i FIVIS-delrapport 2. Det hevdes i rapporten at norsklærere ser operasjonalisering og eksplisering av kompetansemål som er potensielt hjelpemiddel i undervisning og vurdering (se Sandvik og Buland, 2013). Det er viktig å merke seg at forklaringskraften til de fire indikatorene i norsk er på knapt 30 prosent, noe som innebærer at en del av variansen ikke er forklart av indikatorene. «Type lærerutdanning» har ikke samme betydning i faget norsk sammenlignet med engelsk (se tabell 9).

Det interessante fra analysene i kroppsøving er at tre indikatorer svarer for 36 prosent av variansen i «Vurderingspraksis». Indikatoren «Egen vurderingskompetanse» bidrar med størst forklaringsprosent. Det betyr at forståelsen av god vurderingspraksis i kroppsøving er påvirket av lærernes svar på utsagnene: «Jeg har nok kompetanse til å drive med sluttkompetanse», «Jeg føler meg trygg på arbeidet med vurdering» og «Jeg har nok kompetanse til å drive med underveisvurdering». Indikatorene «Elevens læring» og «Lærerens tilbakemelding» forklarer den øvrige delen av variansen. Dette angir en elevsentrert vurderingspraksis i kroppsøving, noe som også løftes fram i

FIVIS-delrapport 2. Observasjoner fra undervisning i faget synliggjør en rekke sekvenser med tilbakemeldinger fra lærer til elev som hadde til hensikt å utvikle elevens kompetanse (se Sandvik og Buland, 2013: 197).

Indikatorene «Type lærerutdanning», «Skolens vurderingskultur» og «Fordypning i fag» har ingen signifikant betydning for vurderingspraksis i kroppsøving. Det spesielle med analysene for kroppsøving er at lærerens egen kompetanse og tilbakemelding og elevenes læring har betydning for vurderingspraksis i faget. Vi finner at tilknytningen til skolens vurderingskultur og faglig fordypning har ingen signifikant rolle i å forklare vurderingspraksis i kroppsøving.

Til sammen fire (4) indikatorer bidrar signifikant til å forklare god vurderingspraksis i matematikk. Forklaringsfaktorene fordeles på kompetanse («Egen vurderingskompetanse») på dialogbasert undervisning («Lærerens tilbakemeldinger») som understøtter «Elevens læring». Dette funnet kan sees i lys av internasjonal matematikdidaktisk forskning der betydningen av hyppige formative vurdering blir ansett som en sentral del av elevens læringsprosess. I FIVIS-delrapport 2 oppsummeres intervjuer med lærere i matematikk og observasjon av undervisning i faget slik: «I [...] matematikkfaget synes to vurderingspraksiser å være fremtredende: skriftlige prøver og samtaler. Dialogen mellom lærer og elev nevnes som en viktig kolde til informasjon om elevens kompetanse, og som en viktig arena for å kunne gi ulike former og «feedback» underveis i læringen.» (Sandvik og Buland, 2013: 81). «Type lærerutdanning» har en signifikant en sentral plass i å forklare variansen i den avhengige variabelen «Vurderingspraksis». Vi ser at «Type lærerutdanning», som også kan knyttes til grad av faglig fordypning, har signifikant betydning for lærernes selvrapporing av god vurderingspraksis i matematikk.

Med utgangspunkt i kapitlets andre forskningsspørsmål, «Hvilke faktorer forklarer god vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk?», presenterer vi følgende konklusjoner:

- Indikatorene «Egen vurderingskompetanse» og «Lærernes tilbakemeldinger» er felles og har betydning for vurderingspraksis i alle fagene.
- Indikatorene «Type lærerutdanning» og «Elevens læring» har en signifikant rolle i tre av fire fag. «Type lærerutdanning» har størst betydning for vurderingspraksis i engelsk, mens indikatoren «Elevens læring» har en sentral plass i engelsk og matematikk. «Elevens læring» har litt svakere betydning i faget kroppsøving.
- God vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk påvirkes av noen ulike forhold.

- Kroppsøving skiller seg mest ut fra de øvrige fagene i denne studien.

7.4 Oppsummering av resultater

I dette kapitlet er det gjort rede for resultater fra de kvantitative analysene av data fra breddesurveyen, som ble gjennomført vinteren 2012. Hensikten er å undersøke samsvar mellom den avhengige variabelen «Vurderingspraksis», 13 uavhengige indikatorer, som er utviklet med bakgrunn i en eksplorerende faktoranalyse, og to bakgrunnsvariabler. Regresjonsanalysen av hele datasettet (alle fag) viser at fem av seks indikatorer forklarer deler av variansen i «Vurderingspraksis», og som bidrar til å gi svar på kapitlets første problemstilling, «Hvilke faktorer forklarer god vurderingspraksis?». På basis av en multippel regresjonsanalyse er det identifisert tre (3) indikatorer og to (2) bakgrunnsvariabler som forklarer nær 36,7 prosent av variansen. Dette er tilfredsstillende.

For å svare på kapitlets andre problemstilling, «Hvilke faktorer forklarer god vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk?», er det gjennomført en regresjonsanalyse der datasett 1 er splittet i fag. Disse fire datasettene har to felles indikatorer («Egen vurderingskompetanse» og «Lærerens tilbakemeldinger»), noe som antyder at vurderingspraksisene i fagene som inngår i surveyen, har ulike underliggende forklaringsfaktorer. Forklaringskraften i datasettene for fagene engelsk, kroppsøving og matematikk er på 30–48 prosent, og dette er tilfredsstillende. I datasettet for faget norsk er forklaringskraften på 29,5 prosent, som tyder på at hovedtyngden av variasjonen i vurderingspraksis har sammenheng med andre faktorer.

Det er utfordrende å måle læreres vurderingspraksis, og det knytter seg noen metodiske problemstillinger (lav svarprosent) til datagrunnlaget. Vi understreker at konklusjonene fra analysene bør tolkes med varsomhet. Selv med disse forbeholdene mener vi at våre analyser peker på noen interessante sammenhenger:

- Analyser av god vurderingspraksis danner et bilde som er koblet til ulike underliggende forklaringsfaktorer.
- God vurderingspraksis har sterk sammenheng med lærerens egen kompetanse i vurdering og lærerens tilbakemelding, type lærerutdanning, elevens læring og lærerens faglige fordypning.
- God vurderingspraksis knyttes også til hvordan lærere bruker tilbakemeldinger for å understøtte elevens læring.

- God vurderingspraksis i engelsk, norsk, kroppsøving og matematikk påvirkes av flere forhold. God vurderingspraksis i kroppsøving har færrest forklaringsfaktorer.

Disse konklusjonene drøftes mer inngående under, og kapitlet avrundes med å løfte fram noen implikasjoner for å utvikle god vurderingspraksis i skolen.

7.5 Drøfting – god vurderingspraksis – et sammensatt bilde

Lærernes vurderingspraksis utvikles i en kontekst med styringsmekanismer og forvaltningspraksiser i skolen (Eggen, 2009). Målstyring, rammestyring og resultatstyring er kjennetegn ved styringsfilosofien i Kunnskapsløftet, og arbeid med de lokale planene krevde en desentralisering. En særlig utfordring har grunnskolen eiere, ledere og lærere ettersom kompetansemålene for dette skoleslaget er utformet på 4. (noen på 2.), 7. og 10. trinn. Skolen, skoleledere og lærere har et tydelig ansvar for å konkretisere de nasjonale målene og iverksette tiltak som kan bidra til å realisere forventningene. Det gis imidlertid få anvisninger om innhold i det fagstoffet som skal formidles, og hvilke metoder som skal benyttes for å nå målene.

Denne surveyen favner lærere i fire fag som underviser i grunnskolen og i videregående skole. Dette er et utdanningsløp som strekker seg fra seksåringens første dag i grunnskolen til avgangseleven i videregående skole. Med Kunnskapsløftet ble det fra 2006 en rekke endringer i skolens innhold, struktur og organisering fra første trinn i grunnskolen til siste trinn i videregående opplæring. En av de strukturelle endringene var at fagene engelsk, norsk, kroppsøving og matematikk (og andre fag) fikk en felles læreplan for grunn- og videregående skole. Dessuten fikk endringene i forskrift til opplæringslova, kapittel 3, «Individuell vurdering», fra 2009 og rundskrivnet til forskriften i 2010 stor betydning for lærernes vurderingspraksis.

Noe av det sammensatte bildet av god vurderingspraksis som vi ser, kan ha bakgrunn i at grunnskolelærernes vurderingspraksis utvikles i en kontekst med styringsmekanismer og forvaltningspraksiser som før Kunnskapsløftet var forskjellig fra forvaltning og styring i videregående skole.

Når vi ser analysene av indikatorene «Type lærerutdanning», «Fordypning i fag» og «Egen vurderingskompetanse» under ett, trer det fram noen signifikante sammenhenger som peker mot et skille mellom ulike forståelser av god vurderingspraksis i norsk grunnutdanning. På den ene siden finner vi lærere i videregående skole og i ungdomsskolen (8.–10. trinn), hvor det har vært tradisjon for vurdering med bruk av karakterer og vurderingen i utgangspunktet skulle foretas i forhold til målene og innholdet i fagene.

Overgangen til et målrelatert vurderingssystem skjedde i 1968 i forbindelse med vurdering i videregående skole. Samlet sett har vi flere tiårs vurderingspraksis i henholdsvis grunnskolens barnetrinn, med vekt på elevens forutsetninger, arbeidsmåter og arbeidsinnsats, og i grunnskolens ungdomstrinn og i videregående opplæring, der vurderingen har utgangspunkt i målene for og innholdet i fagene, har ført til en utvikling av vurderingspraksiser i grunnutdanningen som skiller seg fra hverandre. Et funn i FIVIS-delrapport 1 understreker dette. På spørsmål om lærerne på skolen bruker vurderingsresultater systematisk for å utvikle vurderingspraksisen, rapporterer lærere i videregående skole om mer systematisk arbeid med vurderingsresultater enn de øvrige skoleslagene.

En av konklusjonene i analysene er at god vurderingspraksis i engelsk, norsk, kroppsøving og matematikk påvirkes av flere, delvis ulike forhold. Denne konklusjonen kan vi belyse ved å peke på fagenes formål i skolen, læreplanmål og egenart. Mens formålet med faget kroppsøving er bevegelse, er formålet med matematikk å medvirke til å utvikle den matematiske kompetansen som samfunnet og den enkelte trenger. Vi hevder i FIVIS' delrapport 2 at det er en klar sammenheng mellom lærerens arbeidsmåter og praktiseringen av vurdering for læring i klassen. Lærernes metodiske valg må sees i sammenheng med det handlingsrommet som tolkes ut fra analyser av læreplanverket og utformes i lokale planer for fag. Fagene i skolen bæres også oppe av faktorer utenfor skolen, og dette kan ha betydning for forståelsen av god vurderingspraksis. Ut fra dette vil en vektleggingen av det som skjer utenfor skolen – i den lokale konteksten – ha større innvirkning på vurderingspraksis i faget kroppsøving enn i faget matematikk.

I lærerens vurderingspraksis vil prestasjoner og observasjoner variere etter fag, aldersgrupper, oppgave- og situasjonstype. Lærerne må slik sett beherske et bredt spekter av fortolkningsmetoder for å trekke slutninger om elevenes læring. Observasjon av elever i kroppsøving og, i klasserommet krever en annen tilnærming enn resultater av prøver, som nasjonale prøver.

Et annet perspektiv som kan løftes fram, er å forstå det sammensatte bildet av vurderingspraksis i lys av validitet, som er et sentralt begrep i hele FIVIS-prosjektet (se kap.3, «Hva innebærer validitet i vurdering?»). God vurderingspraksis handler blant annet om å samle bevis for læring og å tolke og anvende denne informasjonen i beslutninger om undervisning og læring. En slik forståelse av vurdering har betydning for hvordan det sentrale begrepet *validitet* defineres og anvendes. Et annet viktig aspekt ved validitet er at vurdering foregår i en kontekst for å måle kompetanse og resultatene brukes til å fortolke, forstå og planlegge tiltak. Dette er sentrale elementer i god vurderingspraksis.

Sett ut fra vår forståelse av validitet i vurdering vil god vurderingspraksis hvile på en rekke kjeder av validitet, som er innbyrdes avhengige av hverandre. Resultatene fra den multiple regresjonsanalysen i tabell 9 peker nettopp på dette: God vurderingspraksis påvirkes av lærerens kompetanse i vurdering, lærerens tilbakemeldinger, skoleslag, skolens vurderingskultur, lærerens faglige fordypning og type lærerutdanning og at lærerens tilbakemeldinger understøtter elevens læring. I FIVIS-prosjektet forstås *validitet i vurdering* som et dynamisk begrep, og slik sett kan det sammensatte bildet av god vurderingspraksis være et uttrykk for at lærerne er aktivt søkende for å utvikle sin kompetanse i vurdering.

I forlengelsen av dette kan gode kjeder av validitet i lærerens vurderingspraksis (se kap.3) forsterke evnen til å forstå sammenhengen mellom læringsprosesser og elevens læring og mellom rammevilkårene for utdanningen og forståelsen av hva som er god vurderingspraksis. Som et grunnlag for en slik positiv tolkning av funnene i de kvantitative analysene vil lærerens vurderingskompetanse, skolens vurderingskultur og et synlig praksisfellesskap i skolen være avgjørende for å utvikle god vurderingspraksis.

7.6 God vurderingspraksis – kompetanse og synlige fellesskaper i skolen

Vi har konkludert med at god vurderingspraksis støtter seg til flere forhold. Sammenhengen mellom indikatorene «Skolens vurderingskultur», «Type lærerutdanning», og «Egen vurderingskompetanse» er presentert. I tillegg har indikatorene «Egen vurderingskompetanse» og «Fordypning i fag» en sterk forklaringskraft i analysene av vurderingspraksis. Av de underliggende faktorene som forklarer god vurderingspraksis, er det på kort sikt vanskelig å oppnå endringer av bakgrunnsvariablene «Type lærerutdanning» og «Fordypning i fag» i skolen. Derimot er det lettere å påvirke den signifikante forklaringsfaktoren «Skolens vurderingskultur», som igjen kan utvikle lærerens egen vurderingskompetanse.

Utviklingen av god vurderingspraksis kan forstås som et individuelt anliggende, der det mest fremtredende er bruk av arbeidsmåter (metoder), vurderingsformer, utvikling av et fagspråk om vurdering samt lærerens kompetanse til å se sammenheng mellom kompetansemål i læreplaner og vurdering. Dette individuelle perspektivet på god vurderingspraksis i skolen underbygges av forskere som peker på at et optimalt læringsutbytte er avhengig av at læreren er bevisstgjort den pedagogiske plattformen som den statlige satsingen Vurdering for læring hviler på, og er i stand til å analysere sin egen

praksis og kontinuerlig videreutvikle den i formativ retning (Steinsholt & Dobson, 2011).

I FIVIS-prosjektet er oppmerksomheten også på utvikling av god vurderingspraksis i et tolkningsfellesskap. Et funn som understreker dette, er at indikatoren «Skolens vurderingskultur» knytter seg signifikant og sterkt til alle de øvrige indikatorene i en bivariat analyse, men sterkest til indikatoren «Skoleledelse» (se tabell 8). God vurderingspraksis kan påvirkes av et synlig praksisfellesskap i skolen. Læring skjer på ulikt vis, både formelt og uformelt, og forutsatt at refleksjon over egen undervisning og observasjon av og refleksjon over andres undervisning foregår i samarbeid med kolleger, kan det skje en utvikling i skolen. Læring vil kunne finne sted i planlagte refleksjonsmøter lærere imellom eller i «uformelle» samtaler med andre kolleger, eller i samtaler med foreldre.

Vi finner at vurderingspraksis i kombinerte skoler (1.–10. trinn) ligger nær barnetrinnets forståelser av god vurderingspraksis (se tabell 7). Skoleledelsen er en sentral faktor i tilrettelegging av lærernes profesjonsutvikling ved den enkelte skole. I opplæringsloven § 10-8, «Kompetanseutvikling», har skoleeier ansvaret for å ha riktig og nødvendig kompetanse i virksomheten. Skoleeier skal også ha et system som gir undervisningspersonale, skoleleder og andre arbeidsgrupper i skolen nødvendig kompetanseutvikling. Hensikten er å fornye og utvide den faglige og pedagogiske kunnskapen og å holde seg orientert om og være på høyde med utviklingen i skolen og i samfunnet (Kunnskapsdepartementet, 2010).

Med bakgrunn i våre analyser kan vi ikke trekke noen vidtrekkende slutninger om hvorvidt en vurderingspraksis som vektlegger utvikling av forståelser gjennom forhandlinger og dialoger mellom lærer og elever, er et gjennomgående trekk i norsk skole. Men vi ser at kompetanseutvikling i skolen som favner variablene i indikatorene «Egen vurderingskompetanse», «Lærerens tilbakemeldinger» og «Elevens læring», åpner opp for meningsskapning og deltakelse i praksisfellesskap, interaksjon, dialog og refleksjon om læring. Læring gjennom aktive og sosiale prosesser, samarbeid i klasserommet, der læreren er tettere på elevenes læring, preger praksiser i et sosiokulturelt syn på vurdering for læring. En utvikling av egen vurderingskompetanse som vektlegger tilbakemeldinger, som fremmer læring gjennom blant annen egenvurdering, og der tause læringsmål formidles gjennom deltakelse og dialog i kunnskapsutvekslinger i faglige kontekster, understreker forståelsen av god vurderingspraksis.

Det som best fører til god vurderingspraksis på skolenivå, er at skoleeier, skolen og dens ledelse har kompetanse og kapasitet til å kunne støtte opp om profesjonell læring og å involvere lærerne i samarbeidende aktiviteter. Sett opp

mot vår forståelse av god vurderingspraksis kan styringsmekanismer og forvaltningspraksiser i 1–10-skolene, som det er rundt 750 av i dette landet, aktualisere spørsmålet om hvorvidt et desentralisert og lokalt arbeid med læreplaner har bidratt til å utvikle ulike forståelser av hva som er god vurderingspraksis i dette skoleslaget.

Fagenes egenart og tradisjoner for vurdering, i form av eksamen i et omfattende og komplekst eksamineringssystem til nasjonale og lokale læreplaner, gjør at fag som engelsk, norsk og matematikk har en annen posisjon enn kroppsøving. En disiplinering av vurderingsarbeidet innenfor et faglig tolkningsfellesskap, som arbeid med eksamensoppgaver og sensoroppdrag, kan bidra til en større åpenhet og forståelse av hva som er en god vurderingspraksis. Dette er en type faglig fellesskap som faget kroppsøving til en viss grad mangler.

Fagenes posisjon i opplæringen varierer, og dette må ses i lys av den samfunnsmessige og politiske verdsetting av fagene. Faget kroppsøving blir betraktet som supplement til andre fag (Annerstedt, 2007), og et skille mellom «kjernefag» som engelsk, norsk og matematikk og «støttefag» som kroppsøving kan ha betydning for forståelsen av god vurderingspraksis. I fag der det finnes tolkningsfellesskap, åpnes det også i større grad for åpenhet og dialog om grunnlag for karaktersetting lærere imellom, men også mellom lærere og elever og deres foresatte. I fag uten eksamen, som kroppsøving, vil individuelt skjønn og kompetanse i større grad kunne danne grunnlaget for lærerens forståelse av hva som er god vurderingskompetanse. Lærernes vurdering kan i et slikt perspektiv gjenspeile oppfatninger om fagenes ulike roller innenfor en samlet helhet i opplæringen.

Avslutningsvis vil vi henlede noe oppmerksomhet på vurderingspraksis i kroppsøving, som i høy grad påvirkes av «Egen vurderingskompetanse» og «Elevinvolvering». Dette kan tolkes som et uttrykk for av skolens vurderingskultur, skoleslag og faglig fordypning har en underordnet betydning for god vurderingspraksis. Den svake forankringen som kroppsøvingslærerne har til skolens vurderingskultur og til skoleledernivået i våre analyser, kan forklares ut fra behovet og ønsket om å søke tolkningsfellesskap utenfor skolen. Et teoretisk perspektiv på dette er deltakerbaner, som innebærer bevegelse i rom, sted og praksisfellesskaper. Lærere deltar i ulike praksiser som har forskjellig betydning for læreren. Relevante erfaringer og kunnskap fra ulike deltakerbaner trekkes inn og blir en del av lærerens yrkeskompetanse og yrkesutøvelse. Deltakerbaner kan være praksisfellesskaper både innenfor og utenfor skolen hvor læringsprosessene fordeles over tid og sted. Ved de fleste skoler er det få lærere i kroppsøving, og kroppsøvingslærere har lang tradisjon for å finne tolkningsfellesskap med kolleger i andre skoler. Vi har over

presentert begrepet *validitet*, eller *kjeder av validitet*, som sentralt i FIVIS-prosjektet. Hvis vi ser kroppsøvlingslærernes søken etter synlige praksisfellesskaper utenfor egen skole i lys av validitet, kan funnene i studien forklares med at lærere i kroppsøving, på linje med lærere i de tre andre fagene i studien, er aktive i å utvikle en god vurderingspraksis.

Vi setter spørsmålstegn ved om denne tolkningen gir en helhetlig og gunstig utvikling av læreres vurderingspraksis. Analysene av god vurderingspraksis forteller at det er flere forhold som bidrar til god vurderingspraksis i fagene engelsk, norsk, kroppsøving og matematikk. Den svake tilknytning til skolens vurderingskultur og skolens ledelse som vi finner for fagene norsk og kroppsøving, åpner for spørsmål rundt styringsmekanismer og forvaltningspraksiser av vurdering i skolen.

Implikasjoner fra gjennomgangen av hvilke forhold som bidrar mest til god vurderingspraksis i fag, er:

- utrede vurderingspraksiser i 1–10-skolene
- undersøke hvordan lærernes vurderingspraksis utvikles i lys av styringsmekanismer og forvaltningspraksiser i skolen

8. Hvordan utvikle vurderingskulturer?

Lise Vikan Sandvik og Henning Fjørtoft

Denne rapporten har gitt svar på prosjektets overordnede problemstillinger.

- Hvordan kommer skoleeieres, skolelederes og læreres vurderingskompetanse til uttrykk gjennom fortolkninger, forståelse og praksiser?
- Hvilke vurderingspraksiser finnes i norske klasserom?
- Hvilken betydning har vurderingspraksisene for elevenes opplevelse av læringsprosesser og læringsutbytte?

I kapittel 3 beskrives én lærers undervisnings- og vurderingspraksiser i fagene matematikk og engelsk. Som vi så, var det til dels store variasjoner i denne lærerens praksis. Mens hun i det ene faget har et klart mål med undervisningen, involverer elevene i vurderingsarbeidet og har en klar sammenheng mellom mål, arbeidsoppgaver og vurderingsformer, er dette ikke til stede i hennes andre fag. FIVIS-forskningen viser at det er flere slike forskjeller innad i skolene som har deltatt i dette prosjektet, både på lærernivå og på ledernivå. Dette kan forstås som en indikasjon på kompleksiteten i pedagogisk endrings- og utviklingsarbeid: Forandringer i skolens tenkning og praksis på vurderingsfeltet kan gripe inn i den enkelte lærers grunnleggende syn på undervisning og læring. Men den har også vist at det er noen lærere, noen skoler og noen skoleledere som lykkes i arbeidet med å utvikle skolens vurderingskultur.

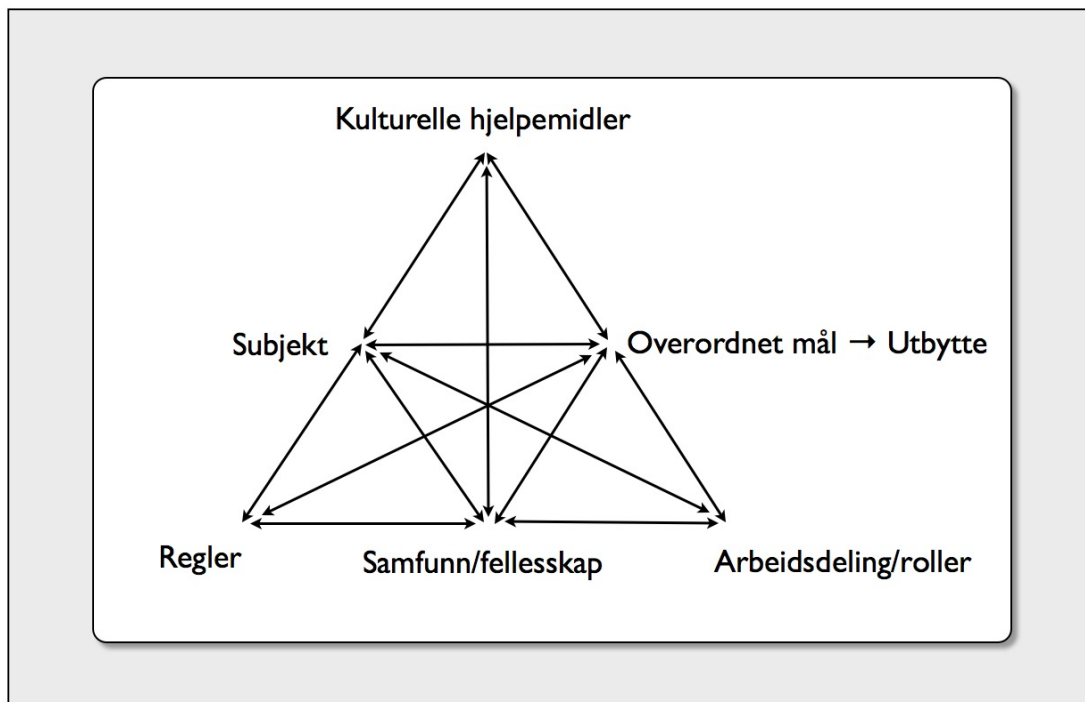
I dette kapitlet sammenfatter vi våre hovedfunn og forsøker å forklare dem ut fra et utviklingsorientert perspektiv. I FIVIS-prosjektet forstås vurdering som en serie fortolkninger og som en dialogisk prosess mellom de involverte deltakerne. Vi undersøker om det kan identifiseres noen forskjeller mellom skoler som har deltatt i Vurdering for læring, og skoler som ikke har deltatt, når det gjelder arbeidet med kvalitet i vurderingen. For å forstå og drøfte våre funn vil vi bruke kulturhistorisk aktivitetsteori. Vi vil derfor innledningsvis redegjøre for denne teorien som et verktøy for å forstå hva som kan hemme og fremme utvikling av vurderingskulturer, før vi går inn på ulike deler av vurderingskulturen ved skolene. Avslutningsvis drøfter vi hvordan kunnskap om validitet i vurdering er en teoretisk terskel som må krysses for å utvikle enkeltlærers vurderingskompetanse og skolens praksis- og fortolkningsfellesskap i arbeidet med vurdering.

8.1 Kulturhistorisk aktivitetsteori

FIVIS-prosjektet er teoretisk forankret i sosiokulturell læringsteori og bygger på en oppfatning om at kunnskap og forståelse blir skapt i sosial interaksjon. Vi forstår vurdering som situert sosial praksis, noe som betyr at det som regnes som legitim kunnskap og praksis i et klasserom, er preget av både tradisjoner og den institusjonelle diskursen omkring vurdering (Pryor & Crossouard, 2008). En slik sosiokulturell tilnærming til vurdering kan identifisere spenninger og konflikter i den sosiale praksisen der vurderingen foregår. Dette har også vært bakteppet for den vurderingsteoretiske diskusjonen.

8.1.1 Vurderingskultur som aktivitetssystem

Skolen og klasserommet er det stedet der aktivitet og læring foregår gjennom bruk av språket. Kulturhistorisk aktivitetsteori (KHAT) er utviklet av Leontjev på grunnlag av Vygotskijs utviklingsteori (Wertsch, 1981) og kan brukes som et teoretisk rammeverk for å analysere ulike betingelser for læring (Engeström, 1999). Essensen i teorien er ifølge Engeström (1999) å ta tak i en situasjon og endre den for å skape noe kvalitativt nytt. I FIVIS undersøker vi hvordan Vurdering for læring bidrar til kunnskapskonstruksjon i aktivitetssystemet. KHAT er utviklet til det som kalles aktivitetssystem (Engeström, 1999), og kan framstilles grafisk slik figur 3 nedenfor viser.



Figur 3. Aktivitetssystemet (Engström, 1999)

Vi ser på hver enkelt skole i satsingen som et aktivitetssystem. Kulturelle hjelpemidler står sentralt i aktivitetsteorien. Det samme gjelder forholdet mellom individene, der aktivitetsteorien framhever betydningen av menneskelige relasjoner som forutsetning for aktivitet og utvikling. All aktivitet er mediert, og ifølge teorien har hver aktivitet i kulturen en kobling til de historisk og kulturelt formede artefaktene (Wertsch, 1981). Denne modellen har vist seg å være et produktivt rammeverk for å kartlegge og transformere kompleksiteten i sosiale praksiser (Lantolf & Thorne, 2006).

Satsingen Vurdering for læring har dreid seg om å uttrykke eksplisitte mål for og forventninger til læringen og involvere elevene i vurdering av eget arbeid. Med Engeström kan vi si at som handlende «subjekt» tar mennesker i bruk «kulturelle hjelpemidler» for å nå sine «mål», som i denne sammenheng defineres som overordnede mål. Dette kan være en skoleleder som inviterer ressurspersoner inn i arbeidet med vurdering for læring for å utvikle praksis. Det handlende subjektet «act on the object» (mål), og motivasjonen knyttes derfor til det overordnede målet. Faktoren «utbytte» står for sluttresultatet av handlingene som utføres. I den nasjonale satsingen Vurdering for læring har de fire prinsippene om god vurderingspraksis vært et mål. Elevene skal forstå hva de skal lære, og hva som er forventet av dem, de skal få tilbakemeldinger som forteller dem om kvaliteten på arbeidet og prestasjonene, de skal få råd om hvordan de kan forbedre seg, og de skal være involvert i eget læringsarbeid ved blant annet å vurdere eget arbeid og utvikling. I tillegg skal vurderingen brukes til å tilpasse undervisning for hver enkelt elev. Det overordnede målet for

satsingen, og det vi i denne sammenhengen kan kalle utbyttet, har vært at skoleeier, skoler og lærebedrifter/opplæringskontorer skal videreutvikle en vurderingskultur og en vurderingspraksis som har læring som mål.

Retningslinjene for satsingen er presentert gjennom Kunnskapsløftet og forskrift til opplæringslova. Slik vil konteksten for handlingene representeres med faktorene «regler», «samfunn/fellesskap» og «arbeidsdeling/roller», som danner de tre nederste trekantene i aktivitetssystemet. Faktoren «regler» innebærer retningslinjer for handlinger (Engeström, 1987; Cole, 1996; Engeström & Miettinen, 1999). «Fellesskap/samfunn» står for alle personer i et fellesskap som deler samme målsetting. Når lærere og lærerteam utgjør det handlende subjektet, kan fellesskap/samfunn i skolesammenheng være skoleledere og skoleeier. «Arbeidsdeling/roller» innebærer at arbeidet eller de målrettede handlingene er fordelt mellom de ulike menneskene som tilhører fellesskapet i skolen. Faktoren «arbeidsdeling/roller» gjør det også mulig å skille mellom kollektiv aktivitet og individuell handling (Engeström, 1987; Cole, 1996; Engeström & Miettinen, 1999). I et lærerteam kan den enkelte lærer ha ulike roller, på samme vis som skoleledere som en del av fellesskap/samfunn har andre roller enn lærerne på skolen. Faktorene som ligger på aktivitetssystemets grunnlinje, kan legge premisser for eller også mulige begrensninger for hvilke handlinger som kan utføres.

Det er også forbindelseslinjer på tvers av de ulike trekantene. Disse linjene er synliggjort i figur 3 ovenfor. Linjen mellom det «handlende subjektet» og «arbeidsdeling/roller» representerer en arbeidsdeling mellom ulike personer som er en del av det handlende subjektet. Det betyr at ulike personer kan utføre forskjellige oppgaver for å nå et felles mål. I systemet er det også en forbindelseslinje mellom «regler» og «mål». Dette innebærer at reglene gir retning for hvilke mål som skal nås, og hvordan disse skal oppnås. Mellom faktorene «fellesskap/samfunn» og «redskaper/artefakter» er det også en forbindelseslinje. Mennesker skaper hjelpemidler samtidig som disse hjelpemidlene påvirker menneskenes handlinger. I satsingen Vurdering for læring kan ledere og lærere på en skole utvikle nye rutiner og bestemmelser, og de kan utvikle nye samarbeidsmønstre og undervisningsopplegg for å gjøre læringsmålene mer eksplisitte og for å involvere elevene i egen læring. Disse rutinene, bestemmelsene og oppleggene kan skrives ned og slik fungere som kulturelle hjelpemidler for videre praksis.

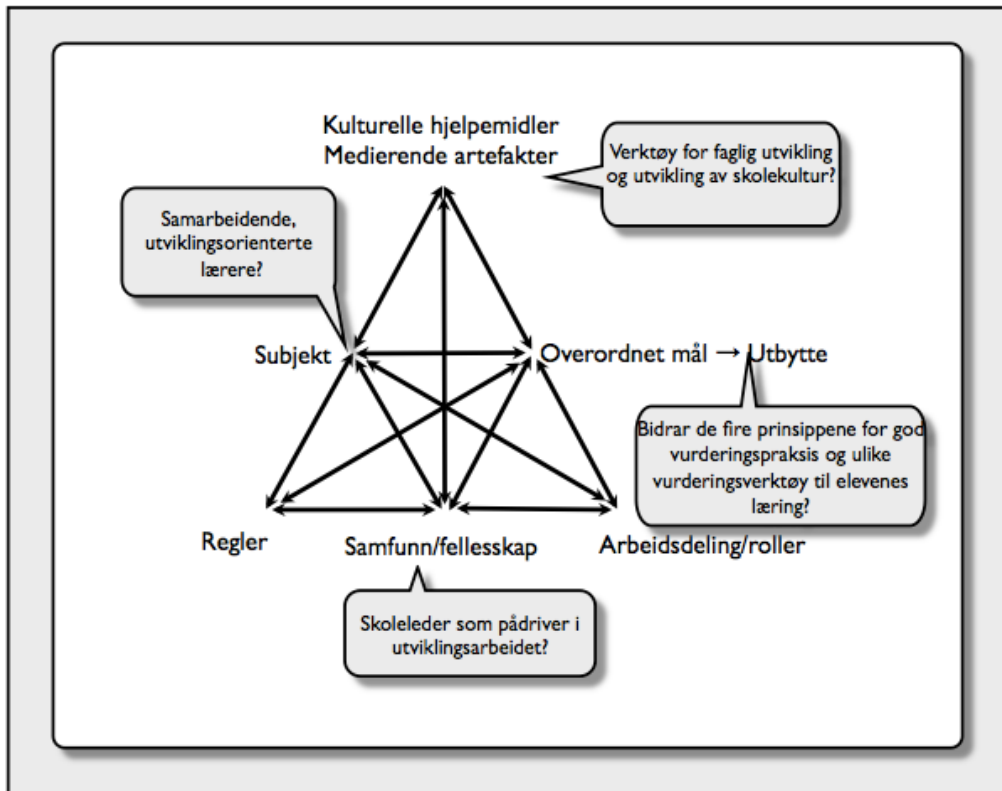
Aktivitetssystemet viser at det er en nær sammenheng mellom det handlende subjektet, som kan være et individ eller en gruppe med mennesker, og konteksten. Konteksten er ikke redusert til noe som bare omgir handlingene, men den er vevd inn i handlingene som pågår (Cole, 1996). I disse handlingene tar menneskene bevisst i bruk ulike redskaper eller artefakter for å nå sine mål.

I en undervisningssituasjon kan læreren som det handlende subjektet bruke ulike vurderingsverktøy for å fremme elevenes læring.

Spenninger eller konflikter mellom de ulike faktorene i aktivitetssystemet utgjør utgangspunkt for endring og utvikling (Engeström & Miettinen, 1999). Vurdering for læring er et eksempel på hvordan det forsøkes skapt nye måter å utføre handlinger på sammen med de lokale deltakerne. I disse tilfellene må én eller flere faktorer endres. Dette innebærer at de andre faktorene i det dynamiske systemet også endres i forhold til de justeringer som iverksettes. Dermed blir forhåpentlig positive endringsprosesser satt i gang. Ved å fokusere på spenninger og konflikter i den eksisterende praksisen undersøkes hva som skjer, for videre å komme med råd for nye og endrende handlinger i arbeidet med Vurdering for læring.

Klasseromsaktiviteter er som regel målorienterte. Gjennomføringen av disse aktivitetene handler om interaksjon mellom ulike deltakere og mellom deltakerne og de ulike artefaktene som tas i bruk for å gjennomføre aktiviteten. I et klasserom vil læreren ha stor innflytelse på hvordan denne interaksjonen materialiserer seg, og i FIVIS er det særlig lærernes vurderingskompetanse og betydningen av denne i elevenes læringsprosess som kommer til syne. I aktiviteter som utføres, kan man se handlingsmønstre som kommer av kulturen man er oppdratt i, eller kulturen man handler i. Å se på målorienterte aktiviteter i forskningen kan si noe om hvilke motiver som ligger bak de ulike handlingene (Wells, 1999). Det er komplekst å forstå hvordan kunnskap konstrueres i en konstant dialog mellom individ og fellesskap. Vi vil i følgende kapitler undersøke hvordan arbeidet med Vurdering for læring påvirker de ulike delene av aktivitetssystemet. Det kan oppstå spenninger og konflikter i de praksisene som eksisterer og er i ferd med å endre seg. Vi starter med subjektet og lærerperspektivet, for så videre å drøfte hvordan samfunn/fellesskap – skoleledere og skoleeiere – påvirker og påvirkes av de nye måtene å utføre undervisning og vurdering på. Avslutningsvis er det målet for praksisene som drøftes, det som omhandler elevenes opplevelse av læring, motivasjon og mestring knyttet til arbeidet med Vurdering for læring. Drøftingene sentreres derfor om fire kjernesporsmål:

1. Hvordan samarbeider lærerne i arbeidet med vurdering for læring?
2. Hvordan leder skolelederen arbeidet med vurdering for læring på egen skole?
3. Hvordan fungerer vurdering for læring som hjelpemiddel for faglig utvikling og utviklingen av skolens vurderingskultur?
4. Hvordan bidrar vurdering for læring til elevenes læring?



Figur 4. Aktivitetssystemet i utvikling av vurderingskulturer.

8.2 Lærernes vurderingspraksiser er mangfoldige

I FIVIS-studien har vi undersøkt hvordan læreres vurderingskompetanse kommer til uttrykk gjennom fortolkninger og forståelser, og hvordan dette manifesterer seg i praksiser i norske klasserom. Alt dette har vi undersøkt for å se hvilken betydning vurderingskulturer har for elevenes læring, som kan sees som den objektorienterte aktiviteten i aktivitetssystemet vi har undersøkt. Vår forskning viser ikke noen klar tendens til at skoler som har deltatt i utviklingsprosjektet Vurdering for læring har en lærerstab der alle lærerne har utviklet den samme forståelsen av hva god vurderingspraksis betyr. Vi har observert mange ulike praksiser innad i skolene, og også mange ulike praksiser hos én og samme lærer. Dette skaper spenninger og motsetninger i aktivitetssystemet, der det ikke er klare, etablerte forståelser av hvordan vurdering for læring kan bidra til økt motivasjon og mestring hos elevene.

Hva påvirker så lærerne, og hva er det som gjør at det er så sprikende praksiser? Lærernes ulikheter kan forstås og forklares ut fra flere aspekter. Vi vil peke på de aspektene vi mener er de mest sentrale, basert på våre funn i FIVIS-studien.

1. Lærers faglige og fagdidaktiske kompetanse synes å ha betydning for hvordan han forstår og operasjonaliserer arbeidet med vurdering.

Vi har gjennom flere eksempler sett hvordan læreres praksiser henger sammen med deres forståelse av vurdering, samt hvordan de operasjonaliserer sine forståelser i undervisningen. Lærere med svak faglig og fagdidaktisk kompetanse synes å ha utfordringer med å forstå og operasjonalisere vurdering for læring på en tilfredsstillende måte. Samtidig ser vi at faglig sterke lærere har en implisitt forståelse av hva god vurdering betyr, og har praksiser som i høyeste grad bidrar til å fremme elevenes læring og motivasjon i fagene, selv om de selv ikke har en klar forståelse av at den undervisningen de bedriver, kan kjennetegnes av de fire prinsippene for god vurderingspraksis.

Det å ha god faglig og fagdidaktisk kompetanse handler også om å ha endrings- og utviklingskompetanse som lærer. Studier (Postholm mfl., 2013) viser at målrettet og systematisk observasjon, refleksjon og erfaringsdeling er viktige aktiviteter som bringer kompetanseutvikling hos lærere fremover. Lærere som lar kompetanse drive eget utviklingsarbeid, vil ha større sjanse for å lykkes med implementering av vurdering for læring enn andre.

2. Skolens organisering i team vs. fagteam synes å ha betydning for hvordan det etableres tolkningsfellesskap omkring vurdering for læring.

FIVIS-rapport 2 viste at skolenes organisering av lærere har betydning for utvikling av tolkningsfellesskap om vurdering. På ungdomstrinnet er organiseringen i trinnteam mest vanlig. Dette gjør at størsteparten av tiden går med til praktisk organisering av undervisning og aktiviteter, og lærere på ungdomstrinnet uttrykker ønske om mer faglig samarbeid for å utvikle seg faglig. På videregående skole synes det som organisering i fagteam er den etablerte samarbeidsformen, og disse lærerne opplever arbeidet i fagteamene som svært nyttig for egen utvikling. Her ser vi imidlertid at det i mindre grad samarbeides om undervisningen. På barnetrinnet underviser lærerne samme klasse i mange fag. Der er trinnteam den etablerte samarbeidsformen, noe lærerne opplever som både praktisk og nyttig ut fra den konteksten de underviser i.

3. Skoleledelsens og skoleeiers forståelser og operasjonaliseringer av arbeidet med vurdering for læring påvirker lærernes holdninger og praksiser.

1–10-skoler har større utfordringer knyttet til etablering av tolkningsfellesskaper enn andre skoler. Også videregående skoler synes i større grad å ha individuelt orienterte lærere enn andre skoleslag. Vi har i tidligere rapporter pekt på at videregående skoler har lange tradisjoner i summativ vurdering. Disse praksisene og forståelsene henger igjen og påvirker lærernes handlinger i høy grad. I neste kapittel vil nettopp spenningene og

motsetningene på skoleleder- og skoleeiernivå løftes fram for å få en større forståelse av hva som bidrar til utvikling av vurderingskulturer.

8.3 Skolenes vurderingskulturer er varierte

Både breddeundersøkelsen, intervjuene og observasjonene vi har gjort med skoleleder- og skoleeiernivået, viser at det er store variasjoner i skolenes vurderingskulturer. I skolene som har deltatt i satsingen, gis det inntrykk av at utviklingen har kommet betydelig lenger, men også blant disse er det variasjoner. Ifølge Timperley, Wilson, Barrar og Fung (2007) er manglende forståelse av hensikten med utviklingsarbeidet et problematisk aspekt ved skoleutvikling. Vi skal se nærmere på ulike aspekter som kan være til hinder for utvikling av gode vurderingskulturer fra et ledelsesperspektiv.

I kapittel 5 refererte vi til O'Donovan mfl. (2006) sin modell for utvikling av vurderingskulturer. Denne modellen brukes som et mulig kart over utviklingen i norske skolars vurderingskulturer. Når vi forsøker å plassere skolene vi har undersøkt, i de fire vurderingskulturene som beskrives i denne modellen, viser det seg at de fleste skolene havner i kategorien «dominant og eksplisitt logikk», med et eksempel i den tradisjonelle modellen og med noen på vei mot en sosialkonstruktivistisk tilnærming. Likevel må vi igjen understreke at det er store variasjoner innad i en skole, slik at det vil være utfordrende å kategorisere den ene skolen som det ene eller det andre.

Det kan være flere årsaker til at variasjonene mellom skolene er så tilsynelatende store. Vi vil peke på to forklaringsgrunner som peker seg ut som sentrale for å utvikle vurderingskulturen på skolenivå.

1. Prosjektet må være forankret i kollegiet.

Ifølge Robinson må lærerne ha et *felles rammeverk* for undervisningen (Robinson, 2011). Det er et lederansvar å sørge for at en slik felles forståelse og samhandlingspraksis utvikles. I FIVIS-forskningen har vi registrert at skoleledelsen ved den enkelte skole ivaretar dette ansvaret på svært ulike måter.

2. Ledelsen må ha forståelse og verktøy for å drive utviklingsprosjekter.

Det er rektor og skoleledelsen som skal lede arbeidet på egen skole slik at det skjer en kunnskapsbasert utvikling av lærernes kompetanse i den hensikt å utvikle vurderingspraksiser som bedrer elevenes læring. Robinsons metastudie (2011) over hvilken form for ledelse som gir sterkest effekt på elevenes læring,

viser at sjansene for å få engasjerte elever med gode resultater er størst når følgende kjennetegner en skole:

- Forskjellige lærere forsterker det samme budskapet.
- Lærerne anvender samme begrepsapparat for å kommunisere budskapet.
- De vet hvordan de skal skape forbindelse til det som har vært tatt opp tidligere.
- De forbedrer sin undervisning basert på felles evaluering (Robinson, 2011, s. 85).

Robinson har utviklet fem kjennetegn på lærende skoleledelse: å etablere mål og kommunisere forventninger, strategisk ressursbruk, å sikre kvalitet i undervisningen, å lede læreres læring og utvikling samt å sikre et velordnet og trygt miljø. Vi ser at på de skolene der man lykkes med utvikling av vurderingskulturen, har lederne klart å etablere en kultur på skolen for at læring og utvikling er et bærende kjennetegn ved skolekulturen.

8.4 Vurdering for læring som skolebasert utvikling og forankret i fag

Satsingen Vurdering for læring er forankret i forskrift til opplæringslova (Utdanningsdirektoratet, 2010) og i læreplan for Kunnskapsløftet. Rammene og reglene for satsingen er med andre ord klare. Disse rammene skal forstås av de ulike handlende subjektene i aktivitetssystemet og manifesteres og etableres som et grunnlag for arbeidet i fellesskapet på skolen. Et verktøy for å forstå rammene og reglene er nettopp satsingen Vurdering for læring. Den er de kulturelle hjelpemidlene i aktivitetssystemet. De kulturelle hjelpemidlene er ment å bidra til utvikling av skolens vurderingskultur, slik at elevene lærer.

Vår studie viser at mange av skolene har kommet godt i gang med å etablere tydelige mål for elevenes læring og å involvere elevene i eget vurderingsarbeid. Særlig viser det seg at de som har deltatt i satsingen, har utviklet en kultur for dette på sin skole. Dette er mest synlig på barnetrinnet. Det kan synes som man på de skolene som har lyktes, har klart å utnytte de spenningsene som oppstår mellom de ulike delene i et utviklingsarbeid, til noe konstruktivt som kan føre til varig endring og grunnlag for videre utviklingsarbeid.

Likevel er det en klar tendens i vårt materiale at mange av skolene ikke har lyktes med å etablere en vurderingskultur på hele skolen, blant hele personalet, der elevenes læring står sentralt, selv om Vurdering for læring har vært et

satsingsområde. Det er flere aspekter i våre funn som kan bidra til å forstå disse motsetningene.

1. De kulturelle hjelpemidlene som anvendes på skolene, kan være av generell karakter og mangle forankring i skolefagene.

Vi har tidligere i dette kapitlet pekt på elementer i satsingen som gjør at enkeltlæreren ikke har forstått eller klart å implementere satsingen i eget arbeid med vurdering. Store deler av satsingen har handlet om å konkretisere mål og bruke ulike skjema og rubrikker i vurderingsarbeidet. Dette er i utgangspunktet en fornuftig inngang til arbeid med vurdering for læring. Til tross for en slik inngang kan det synes som om mange lærere operasjonaliserer dette på en instrumentalistisk måte uten å reflektere over hvordan god vurderingspraksis kan operasjonaliseres på *fagets* premisser. Og så lenge læreren ikke har faglig og fagdidaktisk kompetanse nok til å omsette og fortolke intensjonene til gode praksiser i de fagene han underviser i, vil inngangen til enkeltlæreren fort bli for instrumentalistisk.

2. Den overordnede målsettingen for satsingen er rettet mot enkeltlærerens arbeid.

Et annet viktig punkt i denne sammenhengen er satsingens fokus på enkeltlærerens arbeid med god vurderingspraksis. Det viser seg at verktøyene knyttet til Vurdering for læring i størst grad har blitt anvendt til individuell læring. I hvilken grad verktøyene har blitt anvendt til både individuell og organisatorisk læring, avhenger av skolelederens ledelse av disse prosessene. Skoleledere har varierende kompetanse innenfor ledelse av læreres læring, og deres kunnskaper, ferdigheter og forståelse påvirker det endelige resultatet av utviklingsarbeidet i høy grad. I skoler der lærere og ledelse uttrykker et felles begrepsapparat, har det i høy grad blitt internalisert gjennom systematisk arbeid med satsingen Vurdering for læring.

8.5 Tydelige mål og gode tilbakemeldinger viktige for elevens læring

God vurderingspraksis i fagene kan være et viktig bidrag for å skape den individuelle interessen for fagets innhold og egen læringsprosess. Vurdering for læring handler dessuten om å legge til rette for at elevenes læring foregår i en inkluderende læringskontekst. Det er dette som er det overordnede målet for den nasjonale satsingen Vurdering for læring. I en inkluderende læringskontekst handler det om gode relasjoner mellom elever og lærer, men også om faglig gode omgivelser som utfordrer og motiverer. Når denne

interessen er skapt, er det ifølge Hidi og Renninger (2006) et godt utgangspunkt for å utvikle den individuelle interessen for det faglige innholdet.

Vi vil peke på sentrale funn i FIVIS-studien knyttet til satsingens overordnede målsetting, som også kan bidra til å gi en retning i videre arbeid med vurdering for læring.

1. Bevissthet rundt mål og kriterier er viktig for elevenes læring.

Vår studie viser at elever på alle trinn har relativt høy bevissthet om vurdering, selv om fokuset åpenbart blir sterkere jo høyere årstrinn det er snakk om. Elever på alle trinn opplever vurdering som viktig, også for egen læring. Mange elever vi har snakket med, på alle årstrinn, har fortalt om en betydelig bevissthet rundt mål, at det står i ukeplaner, at lærere skriver det opp på tavla, og også presenterer det muntlig i timene. Likevel ser vi at noen elever i ungdomsskole og videregående skole har et klart fokus på den summative vurderingen og dermed også i mindre grad er opptatt av og forstår hvordan vurdering kan bidra til egen faglig utvikling.

Flere har fortalt om hvordan man har fått informasjon om og vært med på å diskutere vurderingskriterier, altså i betydningen hva som skiller et godt svar fra et dårlig, hva som må være med for å få en viss karakter. Det vitner om en bevissthet hos elevene om arbeidet med mål og kriterier. Det kan synes som om det er etablert praksiser i mange klasserom der nettopp læringsintensjonene blir kommunisert til elevene, både via planer og i klasseromsdialogene.

Delrapport 1 viste at barnetrinnet har kommet langt i arbeidet med mål og kriterier. Også i arbeidet med egenvurdering har de kommet lenger enn videregående skole. Lærerne bruker ulike skjemaer for dette arbeidet, og elevene tar disse i bruk.

FIVIS-studien viser at det i mange klasserom tas i bruk vurderingsverktøy som fremmer refleksjon omkring egen læring. Samtidig viser denne studien at vurderingsverktøy fort kan bli rent tekniske strategier som kommer fra læreren. I stedet bør læringsstrategier betraktes som fleksible muligheter for å nå et mål. Med en forståelse av at mening skapes i interaksjonen, bør det å lære å lære derfor være en gjensidig refleksjon som gjøres i fellesskap med de ulike aktørene i klasserommet.

2. Læreren må ha fagkompetanse og relasjonskompetanse.

Elevgruppene gir et sammensatt bilde av hva det vil si å ha god vurderingskompetanse. «Den gode læreren» er selvsagt en som er faglig flink, som kan faget sitt, og som også kan lære bort, altså har evne til å formidle faget på en måte som ulike elever kan ha utbytte av. I dette ligger det også å kunne gi

gode, framoverrettede tilbakemeldinger på elevenes prestasjoner. For å kunne gjøre dette framheves ofte betydningen av det relasjonelle, at læreren må se hver enkelt elev, skjønne den enkelte elev og kunne kommunisere med alle på deres egne premisser.

Et sentralt tema i mange av elevintervjuene vi har gjort i FIVIS-studien, handler om reliabilitet i vurdering. Elevene forventer at samme lærer skal vurdere ulike elever likt. Funn i vår studie kan tyde på at det er sprikende praksiser på dette området.

3. Gode tilbakemeldinger er viktig for elevens læring.

Gode tilbakemeldinger handler om å gi elevene tilbakemeldinger som forteller dem om kvaliteten på arbeidet eller prestasjonene deres. Forskriften understreker at denne undervisvurderingen skal «gis løpende og systematisk og kan vere både munnleg og skriftleg» (Utdanningsdirektoratet, 2010). FIVIS-studien viser at det er store variasjoner på dette området. Elevene uttrykker at noen lærere er konstruktive og gir grundige pekere på måter de kan forbedre seg på, andre er korte og temmelig intetsigende i sine kommentarer.

4. Egenvurdering bidrar til refleksjon om egen læring.

Når det gjelder egenvurdering, kan det synes som barnetrinnet har kommet lenger i arbeidet enn de andre skoleslagene. Helt ned på fjerde trinn har vi snakket med elever som har fortalt at de til en viss grad driver egenvurdering, for eksempel i form av at de selv må krysse av for om de kan en ting godt eller ikke, om de har nådd målene for uka helt, delvis eller ikke i det hele tatt. Mange gir uttrykk for at egenvurdering er greit, og at de stort sett vet hvor godt de har gjort det på en prøve eller oppgave. Noen forteller også at da de gjorde en konkret oppgave om dette, fungerte det som en bevisstgjøring; elevgruppen ble mer bevisst på vurdering og kriterier, og ikke minst fikk de et mer bevisst forhold til vurdering av sitt eget arbeid.

Dette eksemplet understreker vurdering som en del av undervisningen og som et læringsfremmende verktøy. Dermed er det elevenes læring som står i sentrum, noe som innebærer ikke bare en ny lærerrolle, men også en ny elevrolle. Eleven blir en aktiv, ansvarliggjort deltaker i egen læringsprosess, og egenvurdering fungerer som et verktøy for å fremme en mer aktiv elevrolle (Birenbaum, 1996; Black & Wiliam, 2005, 2006; Sadler, 1989). Dette kan også bidra til å utvikle elevenes individuelle interesse for læring, som motivasjonsforskerne Hidi og Renninger (2006) peker på.

5. Medelevvurdering er viktig for økt bevissthet om egen læring.

I de fire prinsippene som ligger til grunn for den nasjonale satsingen, sies det mye om tilbakemeldinger, lærerens rolle og elevens rolle i egen læringsprosess. Den sier ingenting om medelevvurdering, det til tross for at vurderingsforskere hevder at både egenvurdering og medelevvurdering er svært effektive læringsfremmende verktøy (Black & Wiliam, 2009; Sadler, 1989, 1998). Medelevvurdering er også noe Thomson & Wiliam (2007) løfter fram som et av de fem kategoriene som kjennetegner god vurderingspraksis, som vi har drøftet nærmere i kapittel 3.1.

Forskning viser at tilbakemeldinger i elevpar er svært effektivt, og at elevene i større grad griper og utnytter veiledningen de får av medelever, enn den de får av eksperten (Lantolf, 2000; van Lier, 2000; Sandvik, 2011). Flere forskere i det sosiokulturelle forskningsfeltet har vært opptatt av dette fenomenet og har forsket på hvordan elever er i stand til å gi hverandre effektiv støtte gjennom varierte strategier som er sensitive mot den nærmeste utviklingssonen. En viktig implikasjon av dette arbeidet er at læring kan oppstå i fravær av en ekspert. I stedet konstruerer elevene i samhandling ekspertise gjennom samtalen fordi de deler et mål om å finne svar på en felles utfordring (Lantolf, 2000). Dette er et interessant perspektiv også i dette forskningsprosjektet, der medelevvurdering er en strategi som vi har observert i mange klasserom.

Det synes som om elevene vi har intervjuet i FIVIS- studien, har svært varierte erfaringer knyttet til vurdering for læring. Dette gjenspeiler andre funn i studien. Manglende tolkningsfellesskap omkring vurdering innad i og på tvers av skoler viser et arbeid som er i gang, men som tar tid å internalisere. Det som likevel er verdt å merke seg, er at elevene mener at god vurderingspraksis med tydelige mål og kriterier, der de selv er aktive i egen og andres læringsprosess, er viktige for deres læring, motivasjon og mestring.

8.6 Utvikling av vurderingskulturer

Gjennom de foregående delkapitlene har vi pekt på ulike aspekter som kan fremme og hemme utvikling av vurderingskulturer. Vi har sett på hvordan lærerens fagkompetanse har betydning for hvordan både individuelt og kollektivt utviklingsarbeid oppfattes hos den enkelte læreren og blant elevene. Den kvantitative undersøkelsen av hva som fremmer god vurderingspraksis, fremhever fordypning i fag som en vesentlig faktor for å lykkes med arbeidet. Elevenes oppfatning om den faglig sterke læreren som viktig for deres faglige utvikling utvides også til å omfatte læreren som har evne til å bygge gode relasjoner til sine elever.

Vi har videre sett hvordan lederen spiller en avgjørende rolle for utviklingen av skolens vurderingskultur. I den kvantitative undersøkelsen i kapittel 7 ser vi en sterk korrelasjon mellom skolens vurderingspraksis og skoleledelse. Dette indikerer at måten skolelederen legger til rette for utviklingsarbeid i skolen på, har betydning for hvilke spenninger og motsetninger som oppstår i skolekulturen når endringer og utvikling skal foregå. Lederens rolle i arbeidet vil bidra til enten å hemme eller å fremme utviklingsarbeidet.

Det tredje punktet vi har drøftet, er hvordan de kulturelle hjelpemidlene, eller sagt med andre ord hvordan arbeid med vurdering for læring, bidrar til faglig utvikling og utvikling av vurderingskulturer. Vi har pekt på at et større fokus på utvikling av kollektive kulturer kan bidra til at satsinger som Vurdering for læring vil føre til endring og utvikling i enda større grad enn det som til nå er resultatet. Særlig viser funn i FIVIS-studien at 1–10-skoler har utfordringer knyttet til å etablere tolkningsfellesskaper omkring vurdering. Dessuten kan det synes som de ulike organiseringene i de forskjellige skoleslagene – trinnteam, fagteam – utfordrer ledelsen i å utvikle gode kulturer for utviklingsarbeid i skolene. De store forskjellene kommer til uttrykk i det lokale læreplanarbeidet på de forskjellige skolene, der vi ser store variasjoner både internt mellom trinn og mellom lærerne på det enkelte trinnet.

Det siste spørsmålet vi har drøftet når det gjelder utvikling av vurderingskulturer, er elevenes oppfatning av mestring og motivasjon knyttet til vurdering for læring. Satsingen Vurdering for læring har som mål at elevene skal forstå hva de skal lære, og hva som er forventet av dem, de skal få tilbakemeldinger som forteller dem om kvaliteten på arbeidet og prestasjonene, de skal få råd om hvordan de kan forbedre seg, og de skal være involvert i eget læringsarbeid ved blant annet å vurdere eget arbeid og utvikling. I tillegg skal vurderingen brukes til å tilpasse undervisning for hver enkelt elev. Det overordnede målet for satsingen, og det vi i denne sammenhengen kan kalle utbyttet, har vært at skoleeier, skoler og lærebedrifter/opplæringskontorer skal videreutvikle en vurderingskultur og en vurderingspraksis som har læring som mål. Vi har sett på hvordan ulike deler av aktivitetssystemet kan bidra til å fremme eller hemme elevenes utvikling. Vi tror det er viktig å huske elevperspektivet i denne, og andre, nasjonale satsinger som handler om skoleutvikling. Elevene som vi har intervjuet i FIVIS-prosjektet, er opptatt av en rettfærdig vurdering, at elever blir behandlet på likt grunnlag. Det sier mye om hvilke forventninger til de har til lærernes vurderingskompetanse og skolens vurderingskultur.

Utviklingen av god vurderingspraksis kan forstås på to måter. For det første som et individuelt anliggende hos enkeltlæreren, der det mest fremtredende er bruk av arbeidsmåter, vurderingsformer, utvikling av et fagspråk om vurdering samt lærerens kompetanse til å se sammenheng mellom kompetansemål i

læreplaner og vurdering. I tillegg skal læreren kunne analysere sin egen praksis og kontinuerlig videreutvikle den i formativ retning. For det andre er den et kollektivt anliggende hos skolen som organisasjon. For å få til den individuelle utviklingen må det legges til rette for faglig utviklingsarbeid på skolenivå, slik at det etableres tolkningsfellesskaper omkring vurdering som bidrar til lærernes faglige utvikling. I dette utviklingsarbeidet oppstår vurderingskulturer som fremmer elevenes læring.

Denne formen for kompetanseutvikling åpner opp for meningsskaping og deltakelse i praksisfellesskap, den skaper interaksjon, dialog og refleksjoner over læring. Validitetsbegrepet i FIVIS-prosjektet handler om å skape læring gjennom aktive og sosiale prosesser, gjennom samarbeid i klasserommet, der det er sammenheng mellom mål, arbeidsmåter og vurderingsformer, og der læreren er tettere på elevenes læringsprosess. Gjennom denne forståelsen av validitet vil også vurderingen oppfattes som mer rettferdig og læringsfremmende for eleven.

8.7 Validitet som terskelbegrep

Til grunn for FIVIS-prosjektet ligger elevens læring, mestring og motivasjon som det overordnede målet for skolens virksomhet. I arbeidet for å nå dette målet er bevissthet om eksplisitte forventninger i form av mål og kriterier en viktig faktor (William & Thomson, 2007). FIVIS-studien viser at både fagkompetanse og relasjonskompetanse er viktig for å ha en konstruktiv dialog med elevene om deres læring og utvikling. Videre vet vi at tilbakemeldinger, egenvurdering og medelevvurdering er viktig for elevenes læring (William & Thomson, 2007). FIVIS-studien viser at det er store variasjoner i praksis på dette området i skolen.

Som del av den politiske styringen av skolen har det blitt innført en mer systematisk oppfølging av elevenes læring. Det norske kvalitetsvurderingssystemet (NKVS) har som mål å samle tilgjengelige data for å analysere samsvaret mellom mål og resultat i opplæringen. Forskningen på det norske kvalitetsvurderingssystemet har påpekt at læreplanmålene i Kunnskapsløftet er åpne og generelle, og at de derfor åpner for et betydelig lokalt tolkningsrom (Høst, 2012). Når myndighetene i tillegg stiller krav om tilpasset opplæring, blir ifølge Høst systematisk utvikling av en felles forståelse av målene en viktig forutsetning. Som OECD har påpekt, er det likheter mellom vurdering for læring i klasserommet, evaluering på skolenivå og større datainnsamling på skolepolitisk nivå: Det handler om å samle inn informasjon som kan anvendes for å utforme strategier som kan utvikle forståelse, og som

kan gi grunnlag for videre handling (CERI, 2008). Vurdering på individnivå i skolen bør altså forstås som en del av det større nasjonale systemet.

I norske myndigheters satsing på vurdering for læring synes allmennpedagogiske verktøy å ha blitt vektlagt i større grad enn disiplinfaglige og fagdidaktiske (Sandvik & Buland, 2013), og dette kan ha ført til at vurderingskompetansen til den enkelte lærer ikke har blitt tilstrekkelig forankret i skolefagenes undervisningstradisjoner. Verktøyene som har blitt vektlagt, synes også å være rettet mot enkeltlærere og ikke mot kollektivt orienterte utviklingsprosesser i skolen.

Et flertall av informantene i FIVIS-prosjektet sier at de lokale læreplanene har konkrete mål, og at arbeidet med vurderingskriterier i klasserommet virker motiverende for elevene. Det kan forstås som en indikasjon på at lærerne klarer å håndtere det lokale tolkningsrommet. FIVIS-prosjektet identifiserer likevel visse skillelinjer i norsk sammenheng: Lærere i barneskolen er kommet lengst i arbeidet med å operasjonalisere mål og kriterier. Lærere på ungdomstrinnet og i videregående opplæring sier at disiplinfaglig kompetanse er et viktig element i deres vurderingskompetanse (Sandvik mfl., 2012).

8.7.1 Validitetskjeder i undervisning og læring

Lærere trekker slutninger om elevenes læring og tar avgjørelser på grunnlag av disse slutningene. I FIVIS-prosjektet har vi sett at en bevissthet om *validitetskjeder* kan være et nyttig begrep for å bedømme kvaliteten i arbeidet med vurdering, enten det er fortolkningen av læreplanmålene, det lokale arbeid med læreplaner, dagligdagse samtaler om læring mellom lærere og elever, bruk av standardiserte tester eller rutiner for karaktersetting i skolen. Dette skjer i varierende grad på de ulike skolene (Bonnier, 2013; Kane, 2006; Moss, 2003). Vi understreker at å sikre validitet i vurdering fra klasserom til nasjonalt nivå imidlertid ikke er en enkelthendelse der vi anvender generelle sjekklister eller kontrollverktøy. Validitet kan snarere forstås som et gjennomgripende aspekt ved alle vurderingsprosesser. En mer formell teoretisk forståelse av validitet og reliabilitet i vurdering kan bidra til profesjonalisering av aktørene i dette arbeidet.

Implementering av kompetansemål i læreplanen og vurdering for læring er et utgangspunkt for tilpasset opplæring. Dette krever gode validitetskjeder, som igjen krever kontinuerlige fortolkningsprosesser i skolen. Uten en mer formell forståelse av validitet i vurdering vil det være vanskelig å ha en god forståelse av sammenhengene mellom læringsmål, undervisning, vurderingspraksiser, læringsprosesser og læringsutbytte – noe vår studie viser. Konsekvensen kan bli sprikende operasjonaliseringer og rent erfaringsbasert arbeid uten

tilstrekkelig forankring i grunnleggende vurderingsteoretiske spørsmål. En annen konsekvens kan bli overdreven vektlegging av reliabilitetsspørsmål hos både skoleeiere, ledere og lærere. Det gjelder særlig på trinnene der lærerne setter karakterer. Våre forskningsfunn antyder et spenningsfelt mellom forvaltningen av skolepolitiske styringsmekanismer på den ene siden og det praktiske arbeidet i hvert enkelt klasserom på den andre. Mens styringen av skolevirksomheten har oppmerksomheten rettet mot karakterstatistikker og mulige klagesaker, er spørsmålet om validitet i vurdering delegert ned til lærernivå. Dersom forvaltningsapparatet legger overdreven vekt på reliabilitet i karaktersetting, kan resultatet bli byråkratiserende dokumentasjonsprosesser som ikke bidrar til utviklingen av gode validitetskjeder og en læringsstøttende vurderingskultur i skolen. I vårt datamateriale ser vi noen steder tendenser til dette.

8.7.2 Validitet i utviklingen av praksisfellesskap i skolen

Vi har i kapittel 3 antydnet at vurdering for læring ikke bør forstås som et sett med verktøy eller instrumenter, men heller som del av større prosesser. Én årsak til dette er at samme instrument kan anvendes til både formative og summative formål. En annen er at validering avhenger av en forståelse av konteksten for og formålet med vurderingen. Lærere forsøker å samle inn og fortolke informasjon om elevenes læringsprosesser og -utbytte for å ta avgjørelser om den videre undervisningen. Slike prosesser kan foregå over ulike tidsspenn, fra noen dager til flere år. Samlet fungerer de som en meningsskapende virksomhet i skolen, og over tid kan lærere og hele skoler lykkes i å skape praksisfellesskap der elevene aktivt deltar i vurderingen av sin egen kompetanse. Kapittel 4 forteller om en barneskole som i høy grad har lyktes med dette, og som har skapt et arbeidsfellesskap der både lærere og elever bidrar i og reflekterer over læringsprosesser gjennom ulike vurderingspraksiser. Dette leder frem til det siste teoretiske aspektet av validitet som vi ønsker å fremheve.

Utviklingen av vurderingskompetanse og vurderingspraksiser på de ulike skolene i FIVIS-studien er i varierende grad preget av eksplisitte forventninger og har varierende grad av elevinvolvering. Det ser imidlertid ut til at mange skoler innleder sitt utviklingsarbeid med å gjøre forventningene eksplisitte *før* de involverer elevene i vurderingen. Dersom vi legger teoriene til O'Donovan mfl. (2006) til grunn, er dette et nødvendig steg for å skape en felles forståelse blant lærerne. Formuleringen av lokale læringsmål og vurderingskriterier fremtvinger diskusjoner om og refleksjoner over hva de faglige forventningene består i, og hva som kan regnes som bevis for læring. Dette arbeidet når imidlertid en praktisk grense når det ikke lenger er mulig for læreren å være

mer eksplisitt eller bruke flere ord for å beskrive forventningene. Elevene har på sin side ulike forutsetninger for å forstå formuleringene som ligger bak de eksplisitte læringsforventningene. Ønsket om å gjøre faglige mål forståelige for elevene leder derfor til en kultur der elevene involveres i større grad, og der forståelsen av de faglige forventningene deles gjennom en form for akademisk sosialisering der elevene inkluderes i et praksisfellesskap. Vi understreker imidlertid at dette rammeverket for utvikling av vurderingskulturer ikke er en lineær prosess, og at én og samme skole kan ha trekk som minner om flere av disse fasene eller kulturene, samtidig. Det er til og med mulig for én og samme lærer å ha svært ulike vurderingspraksiser i forskjellige fag (jf. Sandvik, 2011). På 1–10-skoler, der arbeidet i mange fag og trinn skal koordineres, ser det ut til å være særlige utfordringer med å skape felles forståelse i arbeidet med vurdering.

Vi har også forsøkt å forstå vurdering som et sett med kulturelle hjelpemiddel som anvendes i arbeidet med å støtte elevenes læring i skolen (Engeström, 1999). Vurderingspraksiser inngår i et komplekst sett med relasjoner der ulike aktører handler sammen for å bidra til elevenes læring og utvikling. Å forstå vurdering som et kulturelt hjelpemiddel er imidlertid ikke det samme som å forstå vurdering som et verktøy eller instrument. Både standardiserte prøver, egenutviklede vurderingspraksiser og eksamensformer kan forstås som kulturelle hjelpemidler i skolen som aktivitetssystem. De inngår som del av relasjoner i skolens virksomhet og opptrer i både individuelle og mer kollektivt orienterte prosesser. Dette systemet er dynamisk og preget av spenninger og motsetninger. FIVIS-prosjektet har for eksempel avdekket spenninger mellom ulike fag, mellom ulike trinn på 1–10-skoler og mellom skoleeieres og skolelederes syn på hva vurdering er. De ulike aktørenes vurderingskompetanse kommer til uttrykk på ulike måter i dette systemet. Lærere med svak faglig og fagdidaktisk kompetanse synes for eksempel å ha større utfordringer med operasjonalisering av fagenes kompetansemål. Organiseringen av driften på den enkelte skole ser ut til å påvirke skolens evne til å etablere fortolkningsfellesskap. De forståelser og operasjonaliseringer som skoleledere og -iere er ansvarlige for, påvirker trolig lærernes holdninger og praksiser. Vurdering kan derfor ikke forstås som et spørsmål om juridisk definerte plikter og rettigheter alene. Det er snarere et kulturelt hjelpemiddel som inngår i et sett med relasjoner til allerede eksisterende prosesser i skolen som aktivitetssystem.

8.7.3 Validitet som terskelbegrep i utviklingen av vurderingskompetanse

Vi forstår alle former for vurdering som en serie fortolkninger og som en dialogisk prosess mellom de involverte aktørene – fra elever og lærere via

skoleledere og -eiere til nasjonale myndigheter. Arbeidet med kvalitet i vurdering krever imidlertid ikke bare en forståelse av validitet. Like viktig er kunnskaper om *validering* av vurderingen. Ett eksempel på en slik valideringsprosess er lærere som samarbeider om å utvikle prøver i videregående opplæring. Gjennom samarbeidsprosessen etablerer de en felles forståelse av hva som kan regnes som en god oppgave og et godt svar, og hvordan disse svarene kan tolkes som indikasjoner på at eleven forstår eller mestrer. De etablerer også en forståelse av hva som kan være tvetydige indikasjoner, og hva som eventuelt ikke kan regnes som en gyldig slutning. Et annet eksempel er når lærere utvikler vurderingskriterier for en oppgave sammen med elevene på ungdomsskolen. Ved å involvere elevene i beskrivelsen av ulike dimensjoner ved den ferdige oppgaven kan lærere etablere et fortolkningsfellesskap i klassen som bidrar til at elevene forstår hensikten og hva det vil si å mestre oppgaven. Et tredje eksempel er lærere på barnetrinnet som samarbeider om leseundervisningen. De henger opp plakater med lesestrategier i alle klasserom, diskuterer tekstutvalg til undervisningen og gjennomfører kartleggingsprøver i fellesskap. Gjennom disse aktivitetene etablerer de et praksisfellesskap som definerer hva god leseundervisning er, hva som skal regnes som indikasjoner på god lesekompetanse, og hva slags ressurser som kan støtte elevenes leseutvikling.

Det teoretiske grunnlaget for dette henter vi fra nyere validitetsteori, som forstår validitet som et kontekstuel og pragmatisk fenomen (Markus & Borsboom, 2013). Mens man tidligere har ment at validitet er et universelt fenomen, har innflytelsen fra filosofi og språk teori gjort validitet til et mer situert og dynamisk begrep (Bonnier, 2013; Kane, 2006; Moss, 2003). I dag er det derfor mer vanlig å forstå *validering* som en prosess der man utvikler et resonnement for å støtte en bestemt tolkning med et bestemt formål. Dersom formålet med en vurderingssituasjon endrer seg, må valideringen skje på nytt. Det finnes altså ikke en generell standard for validitet i et slikt syn, for nye resonnementer må utvikles dersom formen eller formålet for en vurderingspraksis endres (Markus & Borsboom, 2013). Nasjonale prøver er et godt eksempel på en slik vurderingspraksis som har flere formål. Eventuelle tolkninger av prøveresultatene må derfor valideres på ulike måter: Dersom formålet er å identifisere enkeltelevers behov for mer støtte til å tolke tekster, må vi utvikle et valideringsresonnement som understøtter prøvens gyldighet til dette formålet. Dersom formålet imidlertid er å sammenligne skoler for å se hvor elevene utvikler seg raskest som lesere, må vi utvikle et helt nytt valideringsresonnement.

Vi foreslår derfor at det legges større vekt på validitet som et *terskelbegrep* i utviklingen av vurderingskompetanse i skolen. Med «terskelbegrep» menes her et teoretisk ord eller uttrykk som fungerer som en åpning inn i nye og tidligere

utilgjengelige kunnskapsfelt eller tenkemåter (Meyer & Land, 2006). Et terskelbegrep er en måte å forstå, tolke eller betrakte noe på, og uten tilgang til denne forståelsen vil det ikke være mulig å utvikle seg innenfor fagfeltet. Når et terskelbegrep blir forstått, kan synet på et fagfelt, et kunnskapsområde eller et helt verdenssyn endres. Dette kan skje plutselig eller over lengre tidsrom, men vil ofte innebære vanskelige læringsprosesser og en grad av transformasjon. Forståelse av et terskelbegrep kan innebære helt ny innsikt i hvordan man tenker eller handler innenfor en disiplin, eller hvordan konkrete situasjoner og fenomener forstås og tolkes. Terskelbegrep kan derfor være både *transformative* (de endrer forståelsen av et fenomen eller et felt betydelig), *irreversible* (de innebærer en omstrukturering av forståelsen som neppe lar seg glemme eller «avlære»), *integrerende* (de gjør det mulig å se sammenhenger som tidligere har vært skjult), *avgrenset* (de forholder seg eksplisitt til andre begreper eller kunnskapsområder) og *problematisk* (fordi de representerer en type kunnskap som er taus, fremmed, språklig kompleks eller lignende; Meyer & Land, 2006).

Når vi definerer validitet i vurdering som et terskelbegrep, ønsker vi å fremheve at en forståelse av validitet synes å være både en grense og en terskel i utviklingen av vurderingskompetanse. Vi har sett at læreres erfaring i klasserommet er en sentral kilde til utvikling av vurderingskompetanse i Norge. Det indikerer at refleksjon over egne erfaringer i klasserommet bidrar til lærernes forståelse av vurdering. Denne forståelsen vil imidlertid nå en øvre grense hos den enkelte lærer: Uten en teoretisk fundert forståelse av validitetsbegrepet vil det være vanskelig å videreutvikle innsikten i vurdering hos den enkelte lærer, og kunnskapen forblir på et intuitivt, erfaringsbasert plan. Men validitet representerer også en terskel inn i en ny forståelse av kunnskapsfeltet, og dersom denne terskelen krysses, vil kunnskapen om vurdering kunne endres og integreres i andre former for kunnskap. Det vil gjøre det enklere å diskutere aspekter av vurdering avgrenset fra andre områder av yrkesutøvelsen, men vil også kunne gjøre noen diskusjoner *mer* problematiske når tidligere forklaringer blir forkastet for et mer teoribasert resonnement. Vi mener imidlertid at dette samlet vil være en gevinst fordi en forståelse av validitet vil innebære en mer profesjonalisert forståelse av de intensjoner, forståelser og operasjonaliseringer som kommer til uttrykk gjennom ulike vurderingspraksiser i skolen.

Forståelsen av validitet er en teoretisk terskel som i første omgang må krysses av den enkelte lærer. Det medfører at skoleledere må kjenne til de potensielt transformative og problematiske aspektene av utviklingen av den enkelte skoles praksis- og fortolkningsfellesskap i arbeidet med vurdering. Det innebærer også at skoleeiere, som er ansvarlige for innsamling av karakterstatistikk, behandling av klagesaker og kompetanseutvikling i skolen,

har en forståelse av validitetsbegrepet. En bevissthet om fortolkning og validering av konklusjoner og oppfølgingstiltak vil imidlertid også bety at ny og potensielt «fremmed» kunnskap kan gjøre styrings- og forvaltningsrutiner mer problematiske enn før. Årsaken til dette er at validitetsteoretiske problemstillinger representerer en ny type forståelse som skiller seg betraktelig fra den mer individ- og rettighetsorienterte språkdrakten i opplæringsloven. Funnene som vi har gjort i FIVIS-studien, antyder imidlertid at dette vil være en viktig terskel å krysse for å videreutvikle vurderingskulturen i det norske skolesystemet.

9. Prinsipper for utvikling av vurderingskulturer

Lise Vikan Sandvik, Henning Fjørtoft, Gunnar Engvik, Inger Langseth og Trond Buland

Forståelsen av validitet er en teoretisk terskel som i første omgang må krysses av den enkelte lærer. Det medfører at skoleledere må kjenne til de potensielt transformativ og problematiske aspektene av utviklingen av den enkelte skoles praksis- og fortolkningsfellesskap i arbeidet med vurdering. Det innebærer også at skoleeiere, som er ansvarlige for innsamling av karakterstatistikk, behandling av klagesaker og kompetanseutvikling i skolen, har en forståelse av validitetsbegrepet. I dette kapitlet oppsummerer vi derfor noen implikasjoner for praksis, forskning og videre satsing på vurdering i skolen. Vi har valgt å formulere dette som en serie prinsipper.

Med utgangspunkt i våre funn ser vi følgende faktorer som viktige for utviklingen av sterke validitetskjeder i skolen:

- **Læreplanforståelse er et sentralt element i utviklingen av vurderingskompetanse.** Lærernes forståelse av skolefagenes formål, hovedområder og kompetansemål slik de er definert i læreplanen, legger føringer på det videre arbeidet med undervisning, læring og vurdering. Læreplanforståelsen må dessuten ses i sammenheng med lærernes pedagogiske kunnskaper om hvordan elevenes kompetanse utvikles innenfor hovedområdene i et lengre tidsperspektiv. For eksempel vil det være store forskjeller mellom utvikling av lese- og skrivekompetanse i norskfaget og elevenes evne til å praktisere *fair play* i kroppsøving gjennom de ulike skoletrinnene. Et slikt utviklingsperspektiv kan forstås som en del av den fagdidaktiske kompetanse som kreves for å arbeide med undervisning innenfor de enkelte skolefagene.
- **Lærernes disiplinfaglige kompetanse ser ut til å spille en fremtredende rolle, særlig i ungdomsskole og videregående opplæring.** Lærere i de høyere trinnene arbeider i parallelle fortolkningsfellesskap på én og samme tid: Størsteparten av tiden tilbringes i klasserommet i en disiplinfaglig kontekst, og samarbeid om undervisning skjer primært innenfor denne arenaen. Deltakelse i ulike disiplinfaglige sammenhenger styrker fortolkningsfellesskapet. Når lover og forskrifter endres, må de disiplinfaglige fortolkningsfellesskapene mobiliseres for at de juridiske endringene skal gi mening. God sammenheng mellom nasjonale eller skolebaserte satsinger på den ene siden og samarbeid i fagseksjoner og -nettverk på

den andre kan gjøre det enklere å forankre forståelsen av ny praksis i skolen.

- **Lærernes og skoleledernes vurderingskompetanse spiller en rolle når skolen skal sikre sammenheng mellom læreplanen, vurdering for læring og vurdering av læring.** Både lærere og skoleledere må ha forståelse av vurdering og læreplan når skolens vurderingskultur skal utvikles. En slik systemisk forståelse er nødvendig for å sikre at lokalt arbeid med læreplaner er et felles prosjekt. Dersom en slik forståelse ikke er til stede hos alle berørte aktører, kan arbeidet med læreplan og vurdering bli mer tilfeldig eller usammenhengende.
- **Dialogen mellom lærer og elev er preget av tydelige forventninger og tilbakemeldinger og har synlige resultat i klasserommet.** Samspillet mellom eksplisitte læringsmål, dialogiske arbeidsformer og synlige bevis for læring ser ut til å føre til konstruktive læringsprosesser og en opplevelse av verdsettelse og meningsfylte erfaringer. Når elever opplever å få delta i og kunne påvirke vurderingsprosesser på en meningsfylt måte, fremmer det gode læringsprosesser. Imidlertid kan tilbakemeldinger oppleves som monologiske, lite læringsfremmende eller uten særlig betydning for læring eller karakterresultater. Noen klasserom kan dessuten være preget av aktiviteter der elevenes læring ikke kommer til syne. I slike sammenhenger kan en ensidig vektlegging av tilbakemeldinger, rangering eller læringsmål oppleves monotont og lite transparent av elevene, selv der intensjonene er de beste fra lærernes side. Samspillet mellom lærer og elev og en kontinuerlig dialog om konkrete læringssituasjoner og prosesser lar seg derfor ikke beskrive som standardiserte metoder.
- **Forståelse av intensjonene i den skolepolitiske satsingen Vurdering for læring øker mulighetene til å realisere dem.** Når lærere, skoleledere og skoleeiere har god forståelse av de bakenforliggende ideene i skolepolitiske satsinger, blir det enklere for lokale myndigheter og skoler å realisere dem på en god måte. Dersom denne forståelsen er mangelfull, vil teorier, metoder og praksiser tolkes inn i allerede eksisterende forståelsesrammer. For eksempel kan en satsing på vurdering *for* læring oppleves som overflødig og unødvendig dersom en gruppe lærere eller skoleledere allerede har investert mye tid i vurdering *av* læring, gjennom kvalitetssikring av standpunkt karakterer, sensur og eksamensgjennomføring.
- **Lærerens evne til å bruke vurdering som utgangspunkt for å utvikle undervisningen øker kvaliteten på elevenes læring.** Lærere som involverer elevene i å utvikle vurderingskriterier og se på modeller for gode prestasjoner, og som synliggjør forventninger til læringsprosesser og -utbytte, integrerer nye vurderingspraksiser i allerede eksisterende undervisningsopplegg. De forbedrer kontinuerlig

sine egne rutiner og videreutvikler rammene for elevenes læring. Gjennom samarbeid med andre lærere treffer de beslutninger om den videre retningen som undervisningen skal gå i, slik at elevenes individuelle potensial kan realiseres.

- **En felles forståelse av validitet i vurdering og de ulike formålene med vurdering styrker skolens samlede vurderingskultur.** Egne erfaringer ser ut til å være et viktig bidrag til utviklingen av den enkelte lærers vurderingskompetanse. En rent erfaringsbasert tilnærming kan imidlertid gjøre det vanskelig å sette seg inn i innsikter fra forskning. Teoretiske perspektiver på vurdering og faglig funderte begreper som *validitet* kan gjøre det lettere å forstå mer komplekse sider av profesjonsutøvelsen og arbeidet i skolen som organisasjon.
- **Lærere behøver fagdidaktiske kunnskaper om relevante undervisningspraksiser som legger til rette for at alle elevene deltar, forstår og mestrer.** De ulike vurderingspraksisene som den enkelte elev deltar i, kan ikke ses isolert fra den øvrige undervisningssituasjonen i de ulike skolefagene. Lærere som skal realisere prinsipper for god vurdering, må også ha en forståelse av god undervisning i det enkelte fag.

Punktene over fremhever samspillet mellom organisasjonens samlede kompetanse og den enkelte lærers ekspertise innenfor vurdering. Som Helen Timperley har påpekt, er en felles utvikling av kollegiets felles begrepsapparat og forståelse viktig for å kunne utdype kunnskap og ferdigheter hos lærerne. Deres personlige teorier om undervisning og læring utfordres innenfor felles begrepsrammeverk, og mulighetene for å lære mer blir systematisk utformet som deltakelse i profesjonelle læringsfellesskap (Timperley, 2011). For å lykkes med kollektive utviklingsprosesser på skolene er det derfor viktig at

- utviklingen av vurderingskulturen på den enkelte skolen forankres i et felles rammeverk
- utviklingsarbeidet ledes av skoleledelsen
- lærere og skoleledelse har felles begrepsapparat
- lærere og skoleledelse kan skape forbindelser til allerede eksisterende kunnskap om undervisning, vurdering og læring
- lærerne kan utvikle undervisningen basert på felles utviklingsprosesser med etablerte kriterier for evaluering

Flere av disse momentene ser ut til å være gjensidig forsterkende, for eksempel når skoleledere tilrettelegger for skolebaserte utviklingsprosesser der alle lærere deltar, samtidig som lærerne deltar i andre former for kunnskapsutvekslinger. En bevissthet om fortolkning og validering av konklusjoner og oppfølgingstiltak vil bety at ny og potensielt «fremmed» kunnskap kan gjøre styrings- og forvaltningsrutiner mer problematiske enn før.

Årsaken til dette er at validitetsteoretiske problemstillinger representerer en ny type forståelse som skiller seg betraktelig fra den mer individ- og rettighetsorienterte språkdrakten i opplæringsloven. Funnene som vi har gjort i FIVIS, antyder imidlertid at dette vil være en viktig terskel å krysse for å videreutvikle vurderingskulturen i det norske skolesystemet.

10. Litteratur

- Aasen, P., Møller, J., Rye, E.M., Ottesen, E., Prøitz, T.S. & Hertzberg, F. (2012). *Kunnskapsløftet som styringsreform – et løft eller et løfte? Forvaltningsnivåenes og institusjonenes rolle i implementeringen av reformen*. Oslo: NIFU, UiO.
- Annersted, C. (2007). *Att (lära sig) vara lärare i idrott och hälsa*. Göteborg: Multicare.
- Birenbaum, M. (1996). Assessment 2000. Towards a Pluralistic Approach to Assessment. I M. Birenbaum & F.J.R.C. Dochy (red.), *Alternatives in Assessment for Achievement, Learning Processes and Prior Knowledge* (s. 3–29). Boston, MA: Kluwer.
- Black, P. Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning. Putting It into Practice*. Buckingham: Open University Press.
- Black, P. & Wiliam, D. (1998). *Inside the Black Box. Raising Standards through Classroom Assessment*. London: King's College London.
- Black, P. & Wiliam, D. (2005). Changing Teaching through Formative Assessment. Research and Practice. I CERI (red.), *Formative Assessment. Improving Learning in Secondary Classrooms* (s. 223–240). Paris: OECD.
- Black, P. & Wiliam, D. (2006). Assessment for Learning in the Classroom. I J. Gardner (red.), *Assessment and Learning. Theory, Practice and Policy* (s. 9–25). London: Sage Publications.
- Black, P. & Wiliam, D. (2009). Developing the Theory of Formative Assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Bonnier, S.M. (2013). Validity in Classroom Assessment. Purposes, Properties and Principles. I J.H. McMillan (red.), *SAGE Handbook of Research on Classroom Assessment* (s. 87–106). Los Angeles, London, New Dehli, Singapore, Washinton DC: SAGE Publications.
- Campbell, C. & Evans, J.A. (2000). Investigation of Preservice Teachers' Classroom Assessment Practices during Student Teaching. *The Journal of Educational Research*, 93(6), 350–355.

- CERI (2008). *Assessment for Learning. Formative Assessment*. OECD.
www.oecd.org/site/educeri21st/40600481.pdf
- Cole, M. (1996). *Cultural Psychology. A Once and Future Discipline*.
Cambridge, MA: The Belknap Press of Harvard University Press.
- Crooks, T.J., Kane, M.T. & Cohen, A.S. (1996). Threats to the Validity of Assessments. *Assessment in Education*, 3(3), 265–286.
- Dale, L.E, Engelsen, B.U. & Karseth, B. (2011). *Kunnskapsløftets intensjoner, forutsetninger og operasjonalisering. En analyse av en læreplanreform. Sluttrapport*. Oslo: Universitetet i Oslo.
- Dysthe, O. (2008). Klasseromsvurdering og læring. I *Bedre Skole*, 4, 16–23.
- Eggen, A.B. (2009). Vurdering, evaluering og læreplaner som kunnskapsområde. I S. Dobson, A.B. Eggen & K. Smith (red.), *Vurdering, prinsipper og praksis. Nye perspektiver på elev- og læringsvurdering* (s. 75–93). Oslo: Gyldendal Akademisk.
- Elmore, R.F. (2008). Leadership as the Practice of Improvement. I B. Pont, D. Nusche & D. Hopkins (red.), *Improving School Leadership* (s. 37–67). Paris: OECD.
- Emstad, A.B. (2014). *Skolevurdering – et verktøy for ledelse, utvikling og læring i skolen*. Bergen: Fagbokforlaget.
- Engh, R., Dobson, S., Sandvik, L.V., Gamlem, S.M., Engvik, G., Hartberg, E. & Tellefsen, H.K. (2012). *Teoretisk bakgrunnsdokument for arbeid med implementering av vurdering for læring på ungdomstrinnet*. Hentet fra http://www.udir.no/Upload/Ungdomstrinnet/Rammeverk/Ungdomstrinnet/Bakgrunnsdokument_vurdering_for_laring_vedlegg_4.pdf
- Engeström, Y. (1987). *Learning by Expanding*. Helsinki: Orienta-Konsultit Oy.
- Engeström, Y. (1999). Activity Theory and Individual Social Transformation. I Y. Engeström, R. Miettinen & R.-L. Punamäki (red.), *Perspectives on Activity Theory*. Cambridge: Cambridge University Press.
- Engeström, Y. & Miettinen, R. (1999). Introduction. I Y. Engeström, R. Miettinen & R.-L. Punamäki (red.), *Perspectives on Activity Theory* (s. 1–16). Cambridge, MA: Cambridge University Press.
- Engvik, G. (2012). Instructional Leadership. I M.B. Postholm (red.), *Læreres læring og ledelse av profesjonsutvikling* (s. 51–66). Trondheim: Tapir Akademisk forlag.

- Erickson, G. (2012). Forelesning 29. februar 2012 for Nätverk för Språk-, läs- og skrivutvecklare: Bedömning för lärande och likvärdighet. Göteborg: Institutionen för pedagogik och specialpedagogik, Göteborgs universitet.
- Frey, N. & Fisher, D. (2011). *The Formative Assessment Action Plan. Practical Steps to More Successful Teaching and Learning*. Alexandria: ASCD.
- Fullan, M. (2001). *Leading in a Culture of Change*. San Francisco: Jossey Bass.
- Gardner, H. (2008). Five Minds for the Future. Paper presented at the Ecolint Meeting, Genova.
- Gudmundsdottir, S. (2001). Narrative Research in School Practice. I V. Richardson (red.), *Fourth Handbook for Research on Teaching* (s. 226–240). New York: Macmillan.
- Hallinger, P. (2005). Instructional Leadership and the School Principal. A Passing Fancy that Refuses to Fade Away. *Leadership and Policy in Schools*, 4(3), 221–239.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77.
- Hattie, J. (2012). *Visible Learning for Teachers. Maximizing Impact on Learning*. London: Routledge.
- Harlen, W. (2004). A Systematic Review of the Evidence of Reliability and Validity of Assessment by Teachers Used for Summative Purposes. I *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Harlen, W. & Winter, J. (2004). The Development of Assessment for Learning. Learning from the Case of Science and Mathematics. *Language Testing*, 21(3), 390–408.
- Hartberg, E., Dobson, S. & Gran, L. (2012). *Feedback i skolen*. Oslo: Gyldendal Akademisk.
- Hidi, S. & Renninger, A. (2006). A Four-phase Model of Interest Development. *Educational Psychologist*, 41, 111–127.
- Hitchcock, G. & Hughes, D. (1995). *Research and the Teacher. A Qualitative Introduction to School-Based Research*. New York: Routledge.
- Høst, H. (2012). *Kunnskapsgrunnlag og faglige perspektiver for en studie av kvalitet i fag- og yrkesopplæringen*. Rapport 1. Forskning på kvalitet i

- fag- og yrkesopplæringen. Oslo: NIFU.
www.nifu.no/publications/943391/.
- International, Q. (2011). NVivo 9.
- Kane, M.T. (2006). Validation. I R.L. Brennan (red.), *Educational Measurement* (s. 17–64). Westport, CT: Praeger.
- Kaspersen, S.L., Bungum, B., Buland, T., Slettbakk, R. & Ose, S.O. (2012). *Holdninger til fravær og nærvær i skole og arbeidsliv*. Trondheim: SINTEF Teknologi og samfunn, avd. Helse.
- Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Kunnskapsdepartementet. (2010). *LOV 1998–07–17 nr 61: Lov om grunnskolen og den vidaregåande opplæringa (opplæringslova)*. (<http://www.lovdatab.no/all/hl-19980717-061.html>). Oslo: Lovdata.
- Lantolf, J.P. (2000). Second Language Learning as a Mediated Process. *Language Teaching*, 33, 79–96.
- Lantolf, J.P. & Thorne, S.L. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford, New York: Oxford University Press.
- Markus, K.A. & Borsboom, D. (2013). *Frontiers of Test Validity Theory. Measurement, Causation, and Meaning*. New York: Routledge.
- Meyer, J.H.F. & Land, R. (2005). Threshold Concepts and Troublesome Knowledge (2): Epistemological Considerations and a Conceptual Framework for Teaching and Learning. *Higher Education: The International Journal of Higher Education and Educational Planning*, 49(3), 373–388.
- Meyer, J.H.F. & Land, R. (2006). Threshold Concepts and Troublesome Knowledge. An Introduction. I J.H.F. Meyer & R. Land (red.), *Overcoming Barriers to Student Understanding. Threshold Concepts and Troublesome Knowledge*. London: Routledge.
- McMillan, J.H. (2008). *Assessment Essentials for Standard-based Education*. Thousand Oaks, California: Corwin Press.
- Messick, S. (1995). Validity of Psychological Assessment. Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning. *American Psychologist*, 50(9), 741–749.

- Moss, P.A. (1994). Can There Be Validity without Reliability? *Educational Researcher*, 23(2), 5–12.
- Moss, P.A. (2003). Reconceptualizing Validity for Classroom Assessment. *Educational Measurement: Issues and Practice*, 22(4), 13–25.
- Nordenbo, S.E., Larsen, M.S., Tiftikçi, N., Wendt, R.E. & Østergaard, S. (2008). *Lærerkompetanser og elevers læring i barnehage og skole. Et systematisk review utført for Kunnskapsdepartementet, Oslo*. Universitetet i Aarhus: Dansk Clearinghouse for Uddannelsesforskning.
- O'Donovan, B., Rust, C., Price, M. & Carroll, J. (2006). «Staying the distance». The Unfolding Story of Discovery and Development Through Long-term Collaborative Research into Assessment. *Brookes eJournal of Learning and Teaching*, 1(4).
- Parkes, J. (2013). Reliability in Classroom Assessment. I J.H. McMillan (red.), *SAGE Handbook of Research on Classroom Assessment* (s. 107–124). Los Angeles, London, New Dehli, Singapore. Washinton DC: SAGE Publications.
- Popham, W.J. (2003). *Teach Better, Test Better. The Instructional Role of Assessment*. Alexandria: ASCD.
- Postholm, M.B. (2005). *Kvalitativ metode. En innføring med fokus på fenomenologi, etnografi og kasusstudier*. Oslo: Universitetsforlaget.
- Postholm, M.B. (2012). *Læreres læring og ledelse av profesjonsutvikling*. Trondheim: Tapir akademisk forlag.
- Postholm, M.B. (2012). Innledning: Tekster i kontekst. I M.B. Postholm (red.), *Læreres læring og ledelse av profesjonsutvikling* (s. 9–17). Trondheim: Tapir akademisk forlag.
- Postholm, M.B., Dahl, T., Engvik, G., Fjørtoft, H., Irgens, E.J., Sandvik, L.V. & Wæge, K. (2013). *En gavepakke til ungdomstrinnet? En undersøkelse av piloten for den nasjonale satsingen på skolebasert kompetanseutvikling*. Oslo: Akademika forlag.
- Pryor, J. & Crossouard, B. (2008). A Socio-cultural Theorisation of Formative Assessment. *Oxford Review of Education*, 34(1), 1–20.
- Robinson, V., M. Hohepa og & C. Lloyd. (2009). *School Leadership and Student Outcomes. Identifying What Works and Why (Best Evidence Synthesis Iteration)*. Auckland: University of Auckland.

- Robinson, V., C. Lloyd & K. Rowe. (2008). The Impact of Leadership on Student Outcomes. An Analysis of Differential Effects of Leadership Type. *Educational Administration Quarterly*, 44(5), 636–674.
- Robinson, V. (2011). *Student-centered Leadership*. San Francisco: Jossey Bass.
- Robinson, V.M.J. (2010). From Instructional Leadership to Leadership Capabilities. Empirical Findings and Methodological Challenges. *Leadership and Policy in Schools*, 9(1), 1–26.
- Sadler, R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18, 119–144.
- Sadler, R. (1998). Formative Assessment. Revisiting the Territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77–84.
- Sandvik, L.V. (2011). *Via mål til mening. En studie av skriving og vurderingskultur i grunnskolens tyskundervisning*. Ph.d.-avhandling. Trondheim: NTNU.
- Sandvik, L.V., Engvik, G., Fjørtoft, H., Langseth, I.D., Aaslid, B.E., Mordal, S. & Buland, T. (2012). *Vurdering i skolen – intensjoner og forståelser. Delrapport 1 fra prosjektet «Forskning på individuell vurdering i skolen» (FIVIS)*. Trondheim: NTNU Program for lærerutdanning og SINTEF.
- Sandvik, L.V. & Buland T. (red.) (2013). *Vurdering i skolen. Operasjonaliseringer og praksiser. Delrapport 2 fra prosjektet «Forskning på individuell vurdering i skolen» (FIVIS)*, Trondheim: NTNU Program for lærerutdanning.
- Sivesind, K. & Bachmann, K.E. (2013). Læreplaner – fra dokumenter til dokumentasjon. I R.J. Krumsvik & R. Säljö (red.), *Praktisk-pedagogisk utdanning – en antologi* (s. 303–342). Bergen: Fagbokforlaget.
- Slemmen, T. (2010). *Vurdering for læring i klasserommet* (2. utgave). Oslo: Gyldendal Akademisk.
- Smith, J.K. (2003). Reconsidering Reliability in Classroom Assessment and Grading. *Educational Measurement: Issues and Practice*, 22(4), 26–33.
- Steinsholt, K. & Dobson, S. (2011). *Dannelse – introduksjon til et ullent pedagogisk landskap*. Trondheim: Tapir akademisk forlag.
- Stokke, K. mfl. (2008). *Evaluering av vurdering for læring. Underveisrapport fra følgeforskningen «Evaluering av modeller for kjennetegn på måloppnåelse i fag»*. Rapport. Oslo: Universitetet i Oslo.

- Svenkerud, S., Klette, K. & Hertzberg, F. (2012). Opplæring i muntlige ferdigheter. *Nordic Studies in Education*, 1.
- Tellefsen, H.K. & Martinussen, G. (2011). Vurdering som en del av undervisning og læring. I T.L. Hoel, T.M. Guldal, C.F. Dons, S. Sagberg, T. Solhaug & K. Wæge (red.), *FoU i praksis 2010. Rapport fra konferanse om praksisrettet FoU i lærerutdanning. Trondheim, 10. og 11. mai 2010* (s. 341–352). Trondheim: Tapir akademisk forlag.
- Timperley, H., Wilson, A., Barrar, H. & Fung, I. (2007). *Teacher Professional Learning and Development. Best Evidence Synthesis Iteration*. Wellington, New Zealand: Ministry of Education.
http://www.educationcounts.govt.nz/_data/assets/pdf_file/0017/16901/TPLandDBESentire.pdf Lastet ned 13. oktober 2014.
- Timperley, H. (2011). *Realizing the Power of Professional Learning*. Berkshire: McGraw Hill / Open University Press.
- Tessmer, M. & Richey, R.C. (1997). The Role of Context in Learning and Instructional design. *ET&D*, 45(2).
- Utdanningsdirektoratet (2010). Rundskriv Udir – 1 – 2010 – Individuell vurdering i grunnskolen og videregående opplæring etter forskrift til opplæringsloven kapittel 3. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2011a). *Elevundersøkelsen 2011*. Oslo: Utdanningsdirektoratet.
- Utdanningsdirektoratet (2011b). *Nasjonal satsing på vurdering for læring (2010–2014)*. Oslo: Utdanningsdirektoratet.
- van Lier, L. (2000). From Input to Affordance. Social-interactive Learning from an Ecological Perspective. I J.P. Lantolf (red.), *Sociocultural Theory and Second Language Learning* (s. 247–262). Oxford: Oxford University Press.
- Wells, G. (1999). *Dialogic Inquiry. Towards a Sociocultural Practice and Theory of Education*. Cambridge: Cambridge University Press.
- Wertsch, J.V. (1981). The Concept of Activity in Soviet Psychology. An Introduction. I J.V. Wertsch (red.), *The Concept of Activity in Soviet Psychology* (s. 3–36). Armonk: M.E. Sharp.
- Wibeck, V. (2000). *Fokusgruppen. Om fokuserande gruppintervju som undersökningsmetod*. Lund: Studentlitteratur.

Vibe, N., Aamodt, P.O. & Carlsten, T.C. (2009). *Å være ungdomsskolelærer i Norge. Resultater fra OECDs internasjonale studie av undervisning og læring (TALIS)*. Oslo: NIFU STEP.

Wiliam, D. (2011): *Embedded Formative Assessment*. Bloomington: Solution Tree.

Wiliam, D. & Thomson, M. (2007). Integrating Assessment with Instruction. What Will It Take to Make It Work? I C.A. Dwyer (red.), *The Future of Assessment. Shaping Teaching and Learning*. Mahwah, NJ: Erlbaum.

Vedlegg 1: Intervjuguide

Begrepsforståelse

Hva legger du i begrepet vurdering?

Individuell vurdering/elevvurdering?

Underveisvurdering/sluttvurdering?

Vurderingspraksis

Beskriv din egen vurderingspraksis

Hva gjør du når du vurderer?

Hva legger du vekt på?

Ha er/blir avgjørende i din vurdering?

Hvordan gir du tilbakemeldinger?

Når skjer vurdering?

Underveisvurdering/sluttvurdering?

Hvordan involverer du elevene i vurderingen av eget arbeid?

Hvordan bruker du teknologi i vurderingsarbeidet (her tenker vi på LMS (Fronter/It's learning), SkoleArena eller andre verktøy)?

Er vurdering noe som diskuteres mellom lærere?

Har skolen en «kollektiv» vurderingspraksis eller mange individuelle praksiser?

Hvordan tenker skolen rundt utvikling av vurderingspraksis? Finnes det en felles strategi for dette?

Vurderingskompetanse

Hva er vurderingskompetanse?

Hva slags kompetanse er nødvendig for å kunne vurdere?

Har du endret vurderingspraksis i løpet av din tid som lærer?

Hva vil du si har vært sentralt i formingen av din praksis?

Hvordan har du «lært» vurdering? / Hvorfor vurderer du slik du gjør?

Læringsprosesser / elevenes læring

Hvordan bruker du dine vurderingsresultater videre i undervisning?

Hvordan påvirker vurderingspraksisen din måte å jobbe på som lærer / i klassen?

Vurdering og læring: Hvordan henger vurdering og undervisning sammen?

Sammenheng mellom din vurderingspraksis og din undervisning

Hvordan påvirker vurderingen elevenes læring?

Hvordan formidler du vurderingsresultater til elever og foreldre?

Hvordan ser du på forholdet mellom underveisvurdering og sluttvurdering?

metodisk?

formål?

Operasjonalisering

Hvor godt kjenner du vurderingsforskriften?

Hvor god/viktig er vurderingsforskriften i forhold til det arbeid du gjør med vurdering?

Har endring i forskrift ført til endringer i din praksis?

Hvordan bruker du læreplanens kompetansemål i din vurdering?

Lokal læreplan og vurdering: Hvordan henger dette sammen?

Vedlegg 2: Observasjonsguide

Observasjon

foretatt av:

Skole _____ Klasse

Fag _____

Dato

Observasjonsperiode: innledende , midtveis eller avsluttende i en arbeidsperiode på _____ timer

Innledende (introduksjon av et tema/arbeidsperiode)

Læringsmål	
Vurderingskriterier	
Vurderingsformer (produkt)	
Faglig kvalitet	
Forventninger	
Engasjement (motivasjon, mestring, relasjon elev/lærer)	
Klasseromsdialogen	
Elevens mulighet til påvirkning, aktiv deltakelse	
Klasseledelse	
Teknologibruk	

Midtveis (arbeid underveis i økta/opplegget)

<p>Arbeidsoppgaver</p> <ul style="list-style-type: none">- fra lærebøker- teknologistøttede- egendefinerte <p>Grad av tilpasset opplæring</p> <p>Elevens mulighet til aktiv deltakelse / påvirkning</p> <p>Organisering (metoder)</p> <ul style="list-style-type: none">- samarbeid- individuelt arbeid- klassediskusjon <p>Bruk av refleksjon</p> <ul style="list-style-type: none">- instrumentelt- læringsstrategier <p>Uformelle tilbakemeldinger</p> <ul style="list-style-type: none">- fra lærer- medelever- egenvurdering <p>Klasseledelse</p>	
---	--

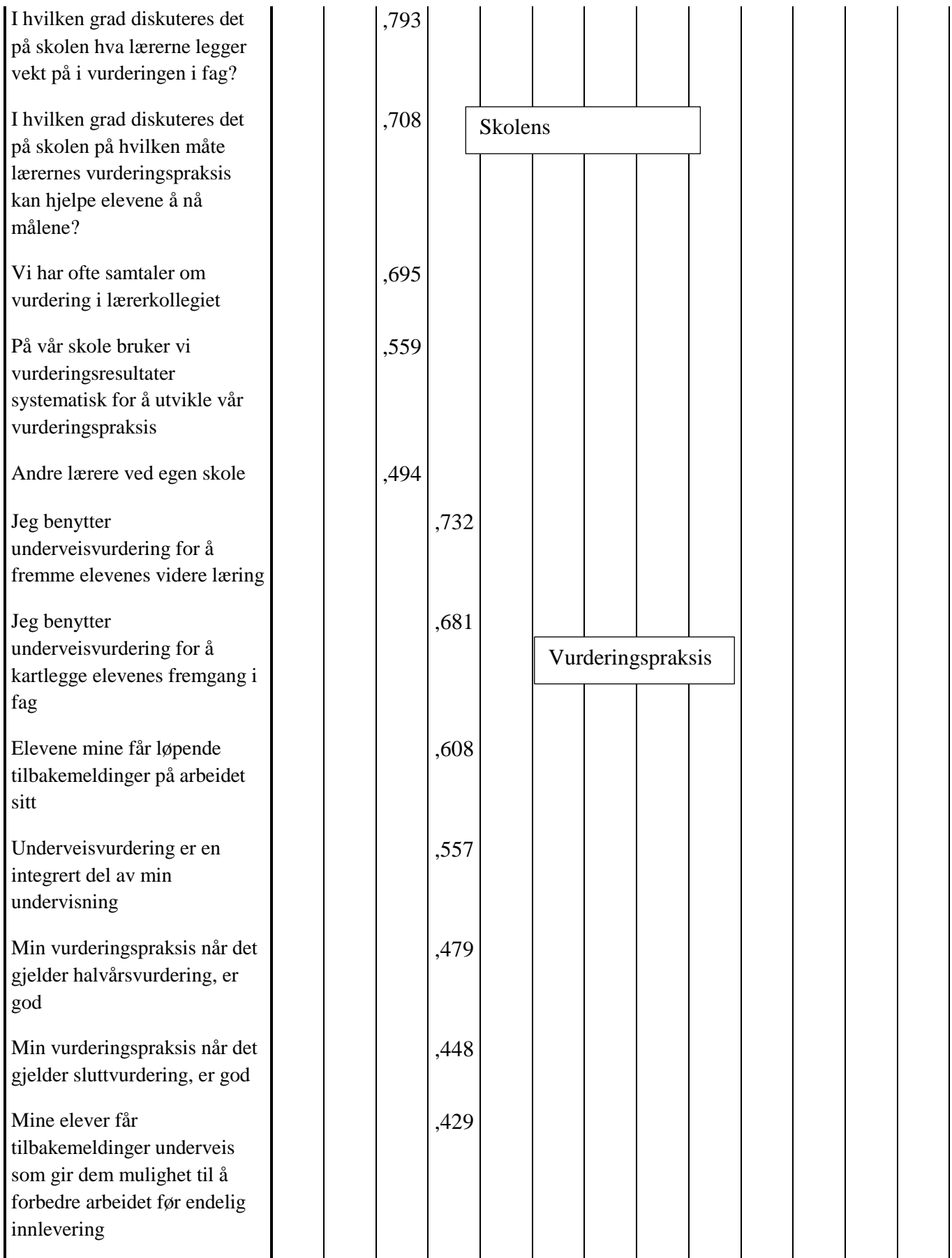
Avsluttende (gjennomføring av vurdering og tilbakemeldinger)

<p>Vurderingsform (produkt)</p> <ul style="list-style-type: none">- kjennetegn på måloppnåelse- vurderingskriterier <p>Tilbakemeldinger (metode)</p> <ul style="list-style-type: none">- klasseromsdialog- individuell vurdering- egenvurdering- medelevvurdering- informasjonsflyt hjem-skole <p>Læringsprosess:</p> <p>Elevers oppfattelse av tilbakemeldinger</p> <p>Elevers anvendelse av tilbakemeldinger</p> <ul style="list-style-type: none">- videre arbeid <p>Lærerens anvendelse av vurderingsresultat</p> <ul style="list-style-type: none">- tilpasse undervisningen <p>Lærerens bruk av evalueringer (logg etc.)</p> <p>Klasseledelse</p>	
--	--

Vedlegg 3: Faktoranalyse med faktorladninger

Rotated Component Matrix^a

	Component													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Underveisvurderingen bør støtte elevenes videre læring	,792													
Jeg synes at underveisvurdering er et nyttig verktøy for meg som lærer	,750		Underveisvurderin											
Underveisvurdering styrker lærerens muligheter til å tilpasse opplæringen til den enkelte elev	,684													
Min vurderingspraksis er viktig for elevenes læring og faglige utvikling	,680													
Underveisvurderingen bør vise hva elevene kan	,678													
Å arbeide med underveisvurdering er en viktig del av læreryrket	,542													
Det er viktig å involvere elevene i vurderingsarbeidet	,457													
Ekstern veiledning, skolevurdering o.l.	,846													
Lærerstudenter som er i praksis	,783		Deltakerbaner											
Tolærersystemet	,704													
Konferanser	,599													
Resultater av internasjonale tester	,526													
Faglitteratur, forskningsrapporter, fagtidsskrifter o.l.	,498													



får tilbakemeldinger på den faglige forståelsen?

I hvilken grad mener du at din vurdering bidrar til at elevene får tilbakemeldinger som forteller dem om kvaliteten på gjennomføringen av arbeidet?

Jeg har nok kompetanse til å drive med sluttvurdering

Jeg føler meg trygg på arbeid med vurdering

Jeg har nok kompetanse til å drive med underveisvurdering

Skolens ledelse arbeider aktivt for kompetanseutvikling knyttet til vurdering

Skolens ledelse bidrar til å utvikle en felles og god vurderingspraksis

Skoleeier arbeider aktivt for kompetanseutvikling knyttet til vurdering

Jeg synes det er krevende å arbeide med sluttvurdering

Jeg synes det er krevende å arbeide med underveisvurdering

I mitt undervisningsfag er det en klar sammenheng mellom læringsmålene for faget og kjennetegn på måloppnåelse

I mitt undervisningsfag er det en god sammenheng mellom læringsmålene for faget og elevens arbeidsoppgaver

Kravet om halvårsvurdering hindrer god underveisvurdering

,497

,797

,739

,729

Egen

,815

Skoleledelse

,744

,623

,876

Krevende

,872

,825

Planer

,777

,890

Hindringe

