

Evaluering av vurdering for læring

**Underveisrapport fra følgeforskningen
”Evaluering av modeller for kjennetegn på måloppnåelse i fag”**

Institutt for lærerutdanning og skoleutvikling (ILS)

September 2008

Kjersti Hage Stokke (ILS)

Inger Throndsen (ILS)

Svein Lie (ILS)

Erling Lars Dale (PFI)

Forord

Denne rapporten tar utgangspunkt i følgeforskningen *Evaluering av modeller for kjennetegn på måloppnåelse i fag*. Forskergruppen har fulgt prosjektet og skolene som deltar gjennom fem måneder og har fått innspill fra lærere gjennom spørreundersøkelse og intervjuer. Spørreundersøkelsen ble gjennomført tidlig i prosjektperioden (februar/mars 2008), mens intervjuene ble gjennomført i april. Vi ønsker i denne rapporten å legge vekt på lærerstemmene inn i prosjektet. For å få et reelt bilde av situasjonen ute i skolen, er det etter vår mening avgjørende å gå ut dit arbeidet med kjennetegn foregår, nemlig blant lærerne. De har bidratt med å dele erfaringer knyttet til sitt arbeid med kjennetegn med oss, verdifull informasjon som er svært viktig for det videre arbeidet med kjennetegn på måloppnåelse og elevvurdering generelt. Vi vil takke alle lærerne som har bidratt med informasjon og data til undersøkelsen. Dataene er analysert i lys av den historiske og utdanningspolitiske bakgrunnen for prosjektet, og i lys av følgeforskningens teoretiske forankring.

Vi har hatt jevnlig kontakt med vår oppdragsgiver, Utdanningsdirektoratet v/Grethe Hovland, og prosjektgruppen for *Bedre vurderingspraksis* (Vivi Bjelke, Hedda B. Huse og Marthe Amundsen) i forbindelse med følgeforskningen. Vi takker alle for et godt og konstruktivt samarbeid. Når det gjelder anbefalingene i rapporten, står de for forskergruppens regning. Vi håper denne underveisrapporten vil bidra positivt til det videre arbeidet med prosjektet.

Oslo, september 2008
Forfatterne

Innhold

Forord	2
Sammendrag og kommentarer	6
1 Innledning	10
1.1 Bakgrunn	10
1.2 Utprøving av modeller	14
1.3 Mer rettferdig, faglig relevant og motiverende	15
1.4 Deltakende skoler	16
1.5 Følgeforskningen etter oppdrag fra Utdanningsdirektoratet	17
2 Teoretisk forankring	18
2.1 Målrelatert vurdering for hele grunnopplæringen	18
2.2 Kriteriebasert vurdering	19
2.3 Taksonomi	21
2.4 Vurdering for læring	23
2.5 Kontinuitet – et enhetlig vurderingssystem	24
3 Metode	26
3.1 Design	26
3.2 Deltakelse på samlinger og møter	26
3.3 Undersøkelse 1	27
3.4 Undersøkelse 2	28
3.5 Analysemetoder	29
4 Resultater	31
4.1 Innledning	31
4.2 Lærere	31
4.3 Skolens utgangspunkt	32
4.3.1 Skolens eventuelle plan for elevvurdering før oppstart av prosjektet	32
4.3.2 Skolens vurderingspraksis før oppstart av prosjektet	33

4.3.3	” Ikke bare fint og flott, det må være noe mer enn det”	35
4.3.4	Lærernes syn på vurdering	36
4.3.5	Elevdeltakelse i vurderingen	38
4.3.6	Lærernes forståelse av oppdraget	39
4.4	Hvordan har skolen organisert arbeidet med prosjektet?	40
4.4.1	Skolenes interne organisering	40
4.4.2	Ledelse av prosjektet på skolen	41
4.4.3	Samarbeid mellom skoler	42
4.5	Hvor langt har skolene kommet i prosjektet?	43
4.5.1	Hvor skolene er i prosjektet ved undersøkelses- tidspunktet	43
4.5.2	Progresjonen i pilotskolene i forhold til resten av skolene	44
4.5.3	Sammenhengen mellom ”før og nå”	45
4.6	Skolenes kompetanseutvikling i elevvurdering	45
4.6.1	Viktige faktorer for å utvikle lærernes kompetanse i elevvurdering	45
4.6.2	Kontakt med eksterne fagmiljøer	48
4.6.3	Hvor stort utbytte har skolene hatt av kontakten med eksterne fagmiljøer?	49
4.7	Lærernes tanker om prosjektet	50
4.7.1	Utvikling av kjennetegn lokalt eller nasjonalt?	50
4.7.2	Lærernes erfaringer med bruk av kjennetegn	51
4.7.3	Forholdet mellom læreplan og kjennetegn	52
4.7.4	Hvor fornøyde er lærerne med å delta i prosjektet?	53
4.7.5	Nyttig og lærerikt, men tidkrevende arbeid og uklar organisering	54
4.7.6	Informasjon til lærerne om prosjektet	56
4.7.7	Hva fikk lærerne av informasjon ved prosjektstart	57
4.7.8	Hva slags informasjon hadde vært viktig å få på et tidlig tidspunkt?	58

4.8	Hva kjennetegner lærere og skoler som er positive til prosjektet og har hatt god progresjon?	59
5	Anbefalinger	61
	Litteratur	65
	Vedlegg 1 Programområder/fag Vg1 og Vg2	68

Sammendrag og kommentarer

Utprøving av modeller for kjennetegn på måloppnåelse i fag

Utdanningsdirektoratet har ansvar for gjennomføringen av et prosjekt om elevvurdering hvor ulike modeller for kjennetegn på måloppnåelse i fag skal prøves ut. Utprøvingen går over to skoleår (fra høsten 2007 til våren 2009), og den er et ledd i Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis*. Fire ulike modeller prøves ut. Modellene A (barnetrinn) og D (ungdomstrinn og videregående) innebærer at skolene selv utvikler egne kjennetegn på måloppnåelse i fag, mens modellene B og C (begge på barnetrinnet) går ut på å prøve ut ferdigstilte kjennetegn i fagene norsk, matematikk, samfunnsfag samt mat og helse. På ungdomstrinnet og videregående skole skal kjennetegnene utvikles i tilknytning til karakterskalaen.

Om rapporten

En forskergruppe ved Institutt for lærerutdanning og skoleutvikling og Pedagogisk forskningsinstitutt ved Universitetet i Oslo gjennomfører følgeforskning med utgangspunkt i denne utprøvingen. Dette er første rapport til Utdanningsdirektoratet fra følgeforskningen "Evaluering av modeller for kjennetegn på måloppnåelse i fag". I denne rapporten gjør vi rede for metoder og funn fra første fase av datainnsamlingen. Resultatene er tidligere presentert i Skolerapport 1 (juni 2008).

Denne rapporten er en tilbakemelding om skolenes utprøving av modeller for kjennetegn på måloppnåelse i fag. Tilbakemeldingen er rettet mot de ulike nivåene i prosjektet, det vil si Utdanningsdirektoratet og involverte skoleeiere, skoleledere og lærere. Rapporten bør også kunne være av interesse for Kunnskapsdepartementet i deres arbeid med elevvurdering, samt være av mer allmenn interesse.

Bakgrunn og teoretisk forankring

I kapittel 1 trekkes bakgrunnen for prosjektet fram med et tilbakeblikk på den vurderingspraksis som de siste tiårenes læreplaner og vurderingsforskrifter har lagt opp til. Som det går fram av denne gjennomgangen, representerer Utdanningsdirektoratets prosjekt nye signaler i forhold til tidligere vurderingspraksis. Målrelatert elevvurdering er for første gang gjort gjeldende for hele grunnopplæringen. Læreplantenkningen i Kunnskapsløftet bygger på kompetansemål som beskriver hva elevene skal mestre, og kompetansemålene skal være utgangspunktet for vurderingen av elevenes måloppnåelse i fagene. Avgjørende blir det da å ha *kriterier* for hvordan graden av måloppnåelse skal vurderes. Eller sagt på en annen måte, det blir viktig å lage en form for *taksonomi* som kan si noe om hva som skal gjelde for måloppnåelse på ett eller flere nivåer. I kapittel 2 gjøres det derfor rede for kriteriebasert vurdering, og ulike taksonomimodeller blir presentert. I diskusjonen om taksonomi settes Blooms måltaksonomi opp mot den måten å beskrive kompetanse på som benyttes i PISA. Kjennetegnene som prøves ut i dette prosjektet (og på noen skoler utvikles), skal utgjøre et hjelpemiddel til å kunne vurdere ulik grad av måloppnåelse hos elever. Ambisjonene er at slike kjennetegn skal bidra til en mer rettfærdig vurdering, gi bedre signaler om hvordan elevene kan forbedre seg, og derigjennom bidra til høyere motivasjon for læring for å oppnå et best mulig læringsutbytte. Tilbakemeldinger som fokuserer på det elevene mestrer og som angir hva og hvordan det er nødvendig å arbeide videre for å oppnå høyere grad av måloppnåelse, omtales gjerne som *vurdering for læring*.

Metode

Datainnhentingene har bestått av en kvantitativ og en kvalitativ del. Dette gjøres det nærmere rede for i kapittel 3. Den kvantitative delen besto i en spørreskjemaundersøkelse til lærere ved alle prosjektskolene. Den ble gjennomført i februar/mars 2008. Kvalitative data har vi fått gjennom åtte gruppeintervjuer med lærere fra fire deltakerskoler, gjennomført i april 2008.

Resultater

Funn fra spørreundersøkelse og intervju blir presentert i kapittel 4. Hovedfunnene er følgende:

Et viktig og nyttig initiativ

Et entydig budskap er at utprøvingen er positivt mottatt ute på skolene. Svært mange lærere gir uttrykk for stor entusiasme og interesse for arbeidet med kjennetegn på måloppnåelse. Denne entusiasmen er sterk, på tvers av klassetrinn og modell. De fleste lærerne rapporterer at de er glade for at skolen er med i prosjektet, og at de opplever arbeid med kjennetegn som nyttig og lærerikt. Den store interessen for og bevisstgjøringen rundt elevvurdering gjennomfører all tilbakemelding i prosjektet. Engasjementet i arbeidet med elevvurdering som kommer fram i intervjuene, vitner om stor interesse blant lærere til å videreutvikle vurderingskompetansen og vurderingspraksisen i skolen. Når så mange lærere trekker fram at de tror eller opplever at bruk av kjennetegn virker læringsfremmende og motiverende på elevene, betyr det at behovet for et kriteriebasert vurderingsverktøy i skolen er stort.

Elevvurdering i prinsipp og praksis

Lærerne gjør gjennom sine svar i spørreskjemaet rede for noen prinsipielle sider ved god elevvurdering som i betydelig grad stemmer overens med det overordnede formålet med kjennetegnene. Spesielt er det få lærere (omtrent en av sju) som gir sin tilslutning til at det er sammenlikning mellom elever som skal danne grunnlaget for elevvurderingen. Videre er det nesten ingen tilslutning til at det bare er elever som har kompetanse på høyt nivå, som har mestring. Ett viktig spørsmål er det imidlertid delte meninger om, nemlig hvorvidt innsats og aktivitet skal ha innflytelse på den faglige vurderingen av måloppnåelse. Det er ikke tvil om at det "riktige" svaret, ifølge formålet med prosjektet *Bedre vurderingspraksis* og kjennetegnene, er "Nei". Men mange lærere (så mye som tre av fire på barnetrinnet og omtrent halvparten på ungdomstrinnet og i videregående) markerer likevel at innsats og aktivitet bør "telle med" i vurderingen. Dette synet har solid tradisjon i norsk skole, altså også i forbindelse med karaktersetning i ungdomsskolen og i videregående skole. Det ligger en viktig utfordring om diskusjon og avklaring her. Når en kultur skal endres, blir det en særlig viktig oppgave å trekke fram de innebygde motsetningene som må overvinnes hvis man skal komme fram til en enhetlig vurderingskultur. I den sammenhengen er det et viktig funn at de lærerne som er mest fornøyd med å delta i prosjektet, er de som målbærer et prinsipielt syn på elevvurdering som gjenspeiler tenkningen som ligger til grunn for kriteriebasert vurdering.

Bare omtrent en av fire lærere rapporterer at skolen hadde en samlet plan for elevvurdering før oppstart av utprøvingen, men her kan riktignok "plan" forstås på litt forskjellig måte. Oppfølgingsspørsmålene fortalte at det særlig dreide seg om systematisk elevdeltakelse, rutiner for informasjon til elever og foreldre, for utvikling av kriterier osv. Det synes å være et sterkt behov for en styrket og mer systematisk vurderingskultur rundt om på skolene.

Behov for veiledning

Lærerne uttrykker både gjennom spørreundersøkelsen og intervjuene at de har behov for kompetanseutvikling både når det gjelder elevvurdering generelt og spesielt i forhold til arbeidet med kjennetegnene. I forbindelse med intervjuene kom det fram at lærere er usikre på hva som ligger i betegnelsen *kjennetegn*, og det ble registrert at det eksisterer ulike oppfatninger av selve begrepet. Det engasjementet lærerne viser i forhold til prosjektet er imidlertid et godt utgangspunkt for arbeidet med kompetanseutvikling i elevvurdering.

Organisering og samarbeid

Langt de fleste skolene har organisert arbeidet i faggrupper/fagseksjoner når de arbeider med kjennetegn. Mer tid til fagsamarbeid ble sett på som positivt. Elevdeltakelsen i dette arbeidet ser imidlertid ut til å være lav. Bare rundt 10 prosent av lærerne (på alle trinn) rapporterer om slik elevdeltakelse, men dette må riktignok ses i sammenheng med at skolene ikke har kommet langt i arbeidet. Det er særlig skoler som er kommet langt, som rapporterer om slik elevdeltakelse.

Et annet viktig spørsmål om organiseringen av utprøvingen gjelder forholdet til lokalt læreplanarbeid, som skolene og lærerne er forpliktet til å arbeide med. Gjennom å se utprøvingen av kjennetegn som en del av denne forpliktelsen, får man en helhetlig utvikling der elevvurdering blir en del av skolens fagplanarbeid. Omtrent 40 prosent av lærerne krysser av for at arbeidet med kjennetegnene er organisert som lokalt læreplanarbeid.

Omtrent halvparten av lærerne rapporterer at skolen samarbeider med andre skoler. Slikt samarbeid er mest omfattende på barnetrinnet. Utveksling av erfaringer med utvikling og bruk av kjennetegn er viktig i denne sammenhengen. Lærerne rapporterer at samarbeid mellom skoler fører til økt motivasjon og bidrar til kompetanseutvikling innenfor elevvurdering.

Behovet for en tydelig ledelse som sørger for kontinuitet, som koordinerer og setter klare mål og rammer for arbeidet med kjennetegn, går igjen i tilbakemeldingene fra lærerne. Skoleleders involvering er viktig for framdriften i prosjektet, for lærernes motivasjon i arbeidet og for en helhetlig kompetanseutvikling knyttet til elevvurdering.

Informasjon

Lærerne rapporterer både direkte og indirekte at informasjonen de har fått, ikke har gitt dem god nok forståelse av hva utprøvingen av kjennetegn egentlig går ut på. Det er to sider ved dette. Et spørsmål er om Utdanningsdirektoratet gjennom skoleeierne har nådd ut med informasjon med definisjonen av hva kjennetegn er. Et annet spørsmål er om den informasjonen som har nådd ut, er tilskrekkelig for å kunne forstå behovet for en ny vurderingspraksis. Våre dybdeintervjuer gir grunn til å peke på at det synes å være et stort behov for avklaring på vesentlige områder. Lærerne sitter dessuten med mange spørsmål som dreier seg om selve prosjektet, og de er usikre på hva som forventes av dem.

I intervjuene kommer det fram at mange av lærerne forstår utvikling av kjennetegn som utvikling av kriterier til delmål av kompetansemålene i LK06. Mange opplever de ferdigstilte kjennetegnene som uklare og for generelle, og skoler som skal utvikle egne kjennetegn, rapporterer at det er vanskelig å bruke de ferdigstilte eksemplene i informasjonsheftet fra Utdanningsdirektoratet. Dette resulterer i at de famler mye og sliter med å finne *hva* og *hvordan*. Flere av skolene etterlyser en "mal" og flere gode eksempler. Vi tror det er viktig å legge sterkere vekt på prinsippet bak prosjektet om å utvikle et tolkningsfelleskap og gi en nasjonal retning.

Progresjon

Det er stor variasjon i hvor langt skolene er kommet i prosjektet. Sammenligning mellom hovedtrinnene viser at lærerne på ungdomstrinnet i større grad enn de andre har tatt kjennetegnene i praktisk bruk, det vil si i forbindelse med vurderingsarbeidet og i sin kommunikasjon med elevene. Dette er et område som det blir spennende å følge i fortsettelsen av prosjektet. Skoler som har tatt utgangspunkt i allerede eksisterende lokale planer og utviklet kjennetegnene i forhold til disse, rapporterer å ha kommet lenger enn skoler som ikke har valgt en slik strategi.

Anbefalinger

Kapittel 5 peker på forhold som det er viktig å fokusere på i det videre arbeidet. Anbefalingene retter seg mot ulike nivå som har ansvaret for gjennomføringen av prosjektet. Vi oppsummerer anbefalingene slik:

- *Et godt grunnlag for endring*
Ta vare på og forsterke lærernes positive holdning til selve prosjektet og til arbeidet med elevvurdering generelt.
- *Uklarheter knyttet til selve prosjektet*
Klargjøre prosjektets mål og hensikt tydeligere.
- *Usikkerhet og valg av terminologi*
Klargjøre terminologien, dvs. forholdet mellom kriterier, kjennetegn og oppnåelse av kompetansemål.
- *Faglig relevant tilbakemelding*
Gjøre enda klarere hva som forstås med faglig relevante tilbakemeldinger.
- *Gode eksempler*
Utvikle flere gode eksempler på kjennetegn på måloppnåelse på ulike nivå, som kan være en støtte i skolens videre arbeid med prosjektet.
- *Et helhetlig vurderingssystem*
Sørge for at det nasjonalt blir etablert en bedre forbindelse mellom de ulike vurderingssystemene i norsk skole (løpende undervisningsvurderinger, nasjonale prøver med ”nivåer”, muntlig og skriftlig eksamen med karakterer).
- *Kompetanseutvikling*
Følge opp kompetansemiljøene slik at en sikrer at veiledningsarbeidet er i tråd med prosjektets intensjoner.
- *Forskning*
Behov for forskning rettet mot forholdet mellom elevvurdering og elevenes læringsutbytte.

1 INNLEDNING

1.1 Bakgrunn

Utdanningsdirektoratet har fått i oppdrag av Kunnskapsdepartementet å iverksette en rekke tiltak for å bedre vurderingspraksisen i norsk skole. Et av tiltakene er en bredt anlagt utprøving av kjennetegn på måloppnåelse i fag på trinn med kompetansemål. Institutt for lærerutdanning og skoleutvikling ved Universitetet i Oslo fikk i brev av 14.11.2007 fra Utdanningsdirektoratet tildelt ansvaret for følgeforskningen knyttet til prosjektet *Evaluering av modeller for kjennetegn på måloppnåelse i fag*.

Den skolepolitiske bakgrunnen for prosjektet *Utprøving av modeller for kjennetegn på måloppnåelse i fag* står beskrevet i St. meld. nr. 16 (2006-2007) ... og *ingen stod igjen* og St.meld. nr. 31 (2007-2008) *Kvalitet i skolen*.

I St. meld. nr. 16 (2006-2007) ... og *ingen stod igjen* fins et avsnitt om individvurdering (s. 79). Etter departementets oppfatning er det fire hovedutfordringer knyttet til individvurdering. Den første utfordring er at regelverket for individvurdering ikke oppfattes som klart nok. Den andre utfordringen er at både lærerutdanningen og skolen mangler tilstrekkelig kompetanse. Den tredje utfordring er at det er svak vurderingskultur og vurderingspraksis i skolen. Og den fjerde utfordring er at det er forsket lite på individvurdering i Norge.

Når det gjelder den første utfordringen, sies det at departementet vil arbeide med å forberede bestemmelser i regelverket om vurdering slik at de kan tas i bruk fra skoleåret 2007-2008. Det sies at bestemmelsene skal legge til rette for mer likeverdig og rettferdig vurdering og at det skal bli lagt til rette for individvurdering som skal bidra til motivasjon og læring. I forlengelsen av omtalen av hva departementet vil gjøre, kommer også informasjon om det som ble til prosjektet "Bedre vurderingspraksis". Departementet sier at det "fra høsten vil bli iverksatt en bredt anlagt utprøving av ulike virkemidler for mer faglig relevant, likeverdig og rettferdig vurdering" (s. 79). Det vil blant annet omfatte utprøving av kjennetegn på måloppnåelse i fag.

Ved å gå nærmere inn i bakgrunnen for prosjektet "Bedre vurderingspraksis", kan en lettere forstå hva en ønsker å oppnå med utprøvingen av modeller som omfatter kjennetegn på måloppnåelse. Bakgrunnen fins i en sammenheng av tre faktorer; 1) svak vurderingskultur og vurderingspraksis i skolen, 2) utilstrekkelig kompetanse i lærerutdanning og i skolen, og 3) mangelfull relevant tilbakemelding. Dette reiser viktige spørsmål: Hva er faglig relevant tilbakemelding? Dersom en ikke kan svare på det, blir det tilsvarende uklart hva en har utilstrekkelig kompetanse på. Og dersom en ikke skjønner, men aner at det er mangler i den eksisterende vurderingspraksis, men ikke får tak i målene og hensikten med forbedringen, vil trolig enkelte oppleve frustrasjoner overfor uklarhet. Det er derfor gode grunner til å gi seg tid til å forstå mer av den situasjonen som foreligger før utprøvingen av modellene med kjennetegn på måloppnåelse.

I St.meld. nr. 31 (2007-2008) *Kvalitet i skolen* fins et avsnitt om tilbakemeldinger. Her heter det, som i den stortingsmeldingen som er henvist til ovenfor, at tilbakemeldinger som fremmer læring og mestring, ser generelt ut til å ha vært en mangelvare i grunnopplæringen. Og følgende utsagn er avgjørende for å forstå hensikten med prosjektet "Bedre vurderingspraksis": "Reform 97 avdekket en kombinasjon av bruk av allmenn ros og fravær av eksplisitte og klare faglige standarder som et klart trekk i arbeidet på barnetrinnet" (s. 30). Her uttales premisen for prosjektet "Bedre vurderingspraksis"; nemlig tilbakemelding til

elevene ut fra ”eksplisitte og klare faglige standarder”. Det første skrittet er å skjønne hva som ikke er ”eksplisitte og klare faglige standarder”. Her heter det i stortingsmeldingen at i noen klasserom er tilbakemeldingen rettet mot å ”støtte elevene emosjonelt og motivere dem”. Den generelle typen av tilbakemeldinger er av typen ”flott” og ”du er flink”, men uten konkretisering av hva elevene er flinke til, eller hva de må jobbe mer med. I andre klasserom er tilbakemeldingen mer av ”faglige kommentarer” som bidrar til at elevene stiller seg selv spørsmålet ”hvordan henger dette sammen med det jeg har lært før?”. Og dermed er vi ved den sammenhengen prosjektet ”Bedre vurderingspraksis” tas opp i. I stortingsmeldingen heter det videre at departementet har satt i gang et arbeid for bedre vurderingspraksis i skolen. Og som premiss for prosjektet er forskrifter fra høsten 2007 om vurdering, hvor det blant annet er innført et klart skille mellom ”faglig tilbakemeldinger og andre tilbakemeldinger” (s. 30). St. meld. nr. 31 (2007-2008) fremhever at i prosjektet ”Bedre vurderingspraksis” arbeider en med ”vurdering og målrettet oppfølging av elevens læringsutvikling og læringsutbytte”. Prosjektet rammes inn i tanker om økt ”læreplanforståelse” ved at skoler blir mer oppmerksomme på ”sammenhengen mellom kompetansemål, læringsaktiviteter og oppgaver som gis til elevene” (s. 31), og at det for mange skoler er en ”utfordring å gjøre vurdering til en integrert del av elevens læringsprosesser” (s. 30).

Vurderingspraksisen på barnetrinnet

Å forandre en kultur er vanskelig. Det gjelder også å forandre en vurderingspraksis i skolen og særlig på barnetrinnet i retning av mer ”faglig relevant tilbakemelding”. Helt fra Mønsterplanen av 1974 har ”arbeidsinnsats” og ”læreprosessen i seg selv” vært del av elevvurderingen (Mønsterplanen 1974, s.59 – s.60). Også elevens ”personlige utvikling” som omfatter mer enn det som omtales som ”framgang i de enkelte fag”, har vært med i vurderingen. Og når lærere på barnetrinnet har vurdert elevens ”framgang i de enkelte fag”, har det vært elevvurdering ut fra ”elevens forutsetninger for fremgang”. Disse retningslinjene har mer eller mindre vært praktisert i over 30 år. I Mønsterplanen av 1987 omtales og knyttes ”den løpende, uformelle vurderingsprosessen” på barnetrinnet til elevvurdering uten bruk av karakterer. Retningslinjene er at overfor elever med ”svake forutsetninger for arbeidet”, må læreren være spesielt oppmerksomme på behovene for ”varsomhet, oppmuntring, råd og veiledning” (s. 76). Det hører sympatisk ut, men funksjonen er nettopp at en har fått ”allmenn ros”, ”flott”, ”du er flink” og fravær av ”eksplisitte og klare faglige standarder”. I retningslinjene om den løpende, uformelle vurderingsprosessen blir lærerens tilbaketrakkethet en del av veilederrollen i og med at det heter at elevene må få ”søke råd og veiledning når de selv føler behov for det” (s. 75). Også på midten av 1990-tallet i St. meld. nr. 47 (1995-1996) omtales elevinnsats som del av det som skal inngå i elevvurderingen uten bruk av karakterer, og som har vært enerådende på barnetrinnet. Vurderingen av innsats sies her å være avhengig av ”evner og anlegg” og av om eleven har ”tillit til egne evner og eigenmotivasjon”. L-97 bygger i hovedsak på tankegangen om at elevens motivasjon og innsats skal stå i fokus når det gjelder elevvurdering på barnetrinnet. Tilbakemeldingene til elevene skulle omfatte ”fremgangen til elevene” og den skulle omhandle ”arbeidsprosessen”. Denne tradisjonen med å vurdere motivasjon, arbeidsinnsats, arbeidsprosess, elevens fremgang ut fra forutsetninger og elevens personlige utvikling, har altså en historie på over 40 år. Prosjektet ”Bedre vurderingspraksis” innebærer ikke en forbedring av disse retningslinjene. Prosjektet innebærer et *brudd* med denne tradisjonen på barnetrinnet når det gjelder elevvurdering sammenholdt med intensjonen om ”faglig relevant tilbakemelding”.

Den videregående opplæringen og ungdomstrinnet har en vurderingstradisjon med bruk av karakterer. Den er basert på at en vurderer elevene med utgangspunkt i målene og innholdet i

fagene. Spørsmålet blir om den vurderingspraksisen som praktiseres med bruk av karakterer, kan gjelde som grunnlag for relevant faglig tilbakemelding i hele grunnopplæringen, inklusive barnetrinnet.

Vurderingspraksis med bruk av karakterer

En utdanningspolitisk premiss for prosjektet "Bedre vurderingspraksis" blir omtalt i St. meld. nr. 31 (2007- 2008) *Kvalitet i skolen* i avsnittet om mål for kvalitet i opplæringen: "Alle elever og lærlinger skal inkluderes og oppleve mestring" (s. 11). Et avsnitt i stortingsmeldingen bærer overskriften "Inkludering og mestring" (s. 22). Et mål for kvalitet i skolen er undervisning som fremmer "mestring og fremgang blant elevene. Både elever som sliter i fagene, og faglig sterke elever skal få utfordringer". Det heter at elevene skal få oppleve faglig fremgang og mestring og at det forutsetter at elevene møtes med "høye og realistiske forventninger til hva de kan få til". Og lærerne må kunne formidle til elevene hva de trenger å jobbe mer med for å bli bedre (s. 22).

En avgjørende bakgrunn for prosjektet "Bedre vurderingspraksis" er at den elevvurderingspraksisen med bruk av karakterer som hittil har vært gjeldende, også har vesentlige mangler. Å "inkludere" alle og at alle opplever "mestring", samt at alle får realistiske utfordringer har ikke vært hensikten med elevvurdering med bruk av karakterer. Bruk av karakterer har vært *normbasert*. Normbasert eller grupperelatert vurdering brukes når det foretas sammenligninger elevene imellom og hvor kvaliteten på den enkelte elevs resultater ses i lys av de andre elevenes prestasjoner. Manglene i en normbasert vurdering med bruk av karakterer trer fram dersom en ser nærmere på modellen som ble brukt tidligere på folkeskolen. I Normalplanen av 1939 var det et eget hovedavsnitt om *karakterer og karaktergivning*. De karaktergradene som skulle brukes var: Særs godt (Sg), Meget godt (Mg), Godt (G), Nokså godt (Ng) og Lite godt (Lg). Utgangspunktet var, det som er godt. Og Godt gis når eleven er "middels flink". Kjernepunktet var med andre ord det som var "middels", og graderingen ga seg ut fra det som var "avgjort over middels" (Mg) og "under middels dyktig, men følger noenlunde med i faget" (Ng). En elev som "i særlig grad" var under eller over middels fikk enten Lg eller Sg. Lg skulle benyttes om eleven i "særlig grad står tilbake" og Sg når "særskilt utmerking" var kjennetegnet av "en ytterst sjelden dyktighet".

En slik gradering med utgangspunkt i "middels dyktighet" passet med karaktergivning ut fra en tenkning om det som ble kalt normalfordeling. Konsekvensen av en slik vurderingsform er at karakterene til den enkelte elev avgjøres på bakgrunn av hvordan de andre elevene presterer. Karakterene ble brukt riktig hvis det ved et større antall karakter, det vil si over flere tusen karakterer, var slik at en fikk følgende prosentfordeling: 4 % Sg, 24 % Mg, 44 % G, 24 % Ng og 4 % Lg. Systemet hadde ikke som intensjon at "alle elever og lærlinger skal inkluderes og oppleve mestring". Dersom en ser på måten karakterskalaen ble brukt, regnet en med at 24 % av elevene fulgte "så noenlunde med i faget", mens 4 % "i særlig grad står tilbake". Eller sagt på en annen måte: Til sammen 28 % av elevene "står tilbake" og av dem står 4 % "i særlig grad" tilbake. Karaktersystemet medførte således at en kunne tale om "mange elever med manglende faglig dyktighet".

Normalplanen av 1939 påpekte også at det hadde vært en tendens til å gi "urimelig gode karakterer". Konsekvensen var at karakterene hadde mistet den betydningen som det opprinnelig var meningen at de skulle ha. Dessuten hadde karakterene også blitt brukt "meget forskjellig" av lærerne. Læreplanen ga derfor overlærere, skoleinspektører og lærere råd om å

”gjøre alt hva de kan for å hindre utgliding her”. I denne forbindelse skulle ”normalfordelingen” brukes som et hjelpemiddel for at elevvurderingen skulle være rettferdig. Dette førte til at karaktergivning basert på normalfordeling i en del sammenhenger ble anvendt på klasse- eller skolenivå, noe som ikke var den egentlige intensjonen. Denne vurderingsmåten kan sies å ha som intensjon å være rettferdig i utvelgelsen av de elevene som er bedre enn andre. Men den kan ikke fungere som faglig relevant tilbakemelding for elevens måloppnåelse og for inkludering og mestring. Å se bakover til denne modellen for å finne målestokk for hvordan en skal utvikle bedre vurderingspraksis i skolen i dag, er derfor ikke tilrådelig.

I Norge har vi i dag et *målbasert vurderingssystem* (jfr. Utdanningsdirektoratets heftet ”Vurdering”, s. 13). Det innebærer at elevene vurderes i forhold til målformuleringer i læreplanen. Overgang til et målrelatert vurderingssystem skjedde i 1968 i forbindelse med vurdering i videregående skole. Her ble ”innsikt og forståelse” brukt som kriteriet på måloppnåelse i det enkelte fag. Innsikt og forståelse kan være et kjennetegn på måloppnåelse i et fag. Kriteriet om ”innsikt og forståelse” i faget gir grunnlag for faglig relevant tilbakemelding. Det er også mulig å gi en rettferdig tilbakemelding om karakterbeskrivelsene er basert på et nasjonalt fortolkningsfellesskap. Det forutsetter at kjennetegnene på måloppnåelsen forstås på samme måte av lærerne uansett hvor i landet eleven tar utdanningen, og at kjennetegnene på måloppnåelse anvendes tilsvarende likt.

Det målbaserte vurderingssystemet, utviklet fra 1968 og gjaldt til og med opplæringsforskriftene i 2001, er imidlertid ikke fullt ut formålstjenlige som modell for ”Bedre vurderingspraksis”. Også dette målbaserte vurderingssystemet, som gjaldt fra 1968 og frem til innføringen av Kunnskapsløftet, var normert med utgangspunkt i over og under middels. Vurdering er *målrelatert* i den forstand at målene for vurderingen er *innsikt og forståelse*. Men vurderingen er også *grupperelatert* i og med at elevene blir sammenlignet med hverandre i forhold til hvem som er over og under middels. I Opplæringslova med forskrifter (2001) ble det i forbindelse med vurdering i grunnskolen slått fast at når en setter karakterer, skal en legge vekt på *i hvilken grad* eleven har oppnådd kompetanse i faget (§ 3-6). Det skal benyttes tallkarakterer på en skala fra 1–6. Ifølge forskriften innebærer karakteren 6 at eleven har oppnådd ”særlig høy kompetanse”, mens karakteren 1 innebærer at eleven har oppnådd ”liten kompetanse i faget”. Det slås også her fast at karakterer skal gi uttrykk for *i hvilken grad* elevene har nådd de målene de blir prøvd i (§ 4-8). Forskriften for videregående opplæring er mer utdypende. I forskriften blir vurdering med karakterer delt inn på følgende måte: *over middels* måloppnåing i faget, *middels* måloppnåing i faget og *under middels* måloppnåing i faget (§ 4-12). *Over middels* måloppnåing i faget knyttes til karakterene 6 og 5, der det skilles mellom svært høy (6) og høy (5) måloppnåelse i faget. *Middels* måloppnåing i faget knyttes til karakterene 4 og 3, der det skilles mellom noe over middels (4) og noe under middels (3) måloppnåelse i faget. Elever som viser *svært svak* innsikt og forståelse av lærestoffet får karakteren 1, mens elever som viser *mangelfull* innsikt og forståelse av lærestoffet får karakteren 2.

Tankegangen er følgende: Elever som viser *svært høy grad* av innsikt og forståelse av lærestoffet fikk karakteren 6. Elever som viser *høy grad* av innsikt og forståelse av lærestoffet fikk karakteren 5. Dette var karakterene for mestring og måloppnåelse. Selv om tilbakemeldingen derfor kunne være faglig relevant, var vurderingssystemet basert på at bare de som hadde *over middels* måloppnåelse, kunne oppleve mestring av fagets mål om ”innsikt og forståelse”. Om vi går inn i sensorveiledninger, kan en se at mestring i hovedsak er på karakteren 5 og at en allerede på karakteren 4 begynner å vurdere ut fra en tankegang om at

en besvarelse er ”mindre fullstendig i svar”, ”noe uklar og upresis”, at oppgaven er ”lite original” eller ”lite selvstendig”. Språket er ”noe enkelt” og kjennetegnet av ”mindre variasjon” i ordvalg og syntaks. Oppgavebesvarelser som får 4, har måloppnåelse som kjennetegnes av å ha en del ”kommunikasjonsnedbrytende feil”. Tenkemåten innebærer altså at kjennetegn på måloppnåelse og følgelig mestring fins i beskrivelsen av karakter 5, og resten står i en gradering fra måloppnåelse om forståelse og innsikt i faget. *Under middels* måloppnåing i faget knyttes til karakterene 2 og 1, der det skilles mellom mangelfull og usikker måloppnåelse (2), samt svak og utilstrekkelig måloppnåelse i faget (1). Det vil selvsagt være en absurd tanke dersom elevene skulle lære å mestre ”svak” forståelse og innsikt i faget, lære å skrive faglige tekster med ”uklar struktur, springende og lite logisk sammenheng” og med ”magert innhold” og ”lite relevant budskap”.

Problemstillingen som prosjektet *Bedre vurderingspraksis* tar utgangspunkt i, er derfor følgende: Hvordan skal en kunne gi faglig relevant tilbakemelding som er rettferdig, og som samtidig ivaretar intensjonen om at alle elever og lærlinger skal inkluderes og oppleve mestring?

1.2 Utprøving av modeller

Utdanningsdirektoratets prosjekt, hvor kjennetegn på måloppnåelse i fag skal prøves ut, retter seg mot hele grunnopplæringen. På barne- og ungdomstrinnet skal kjennetegn på måloppnåelse utvikles og/eller prøves ut i fagene norsk, matematikk, samfunnsfag samt mat og helse på 2., 4., 7. og 10. trinn. I videregående skole deltar Vg1 og Vg2 i prosjektet med fagene norsk og matematikk, i tillegg til andre fag (se vedlegg 1). Kjennetegnene skal være spesifikke for fag og trinn. Ifølge Utdanningsdirektoratet (jfr. heftet ”Vurdering” s. 5) er fellestrekk ved kjennetegn på måloppnåelse at de skal være:

- beskrivelser av hva elevene faktisk mestrer
- beskrivelser som kan skape tolkningsfellesskap
- beskrivelser som utgangspunkt for lokalt arbeid med vurdering
- beskrivelser som knytter seg til flere kompetansemål av gangen

Følgende fire modeller for felles kjennetegn på måloppnåelse prøves ut:

Modell A: Skolene *utvikler selv og prøver ut* kjennetegn på høy og lav måloppnåelse i norsk, matematikk, samfunnsfag og mat og helse på årstrinn med kompetansemål på barnetrinnet

Modell B: Skolene *prøver ut* ferdigstilte eksempler på kjennetegn på høy måloppnåelse i norsk, matematikk, samfunnsfag, mat og helse på årstrinn med kompetansemål på barnetrinnet

Modell C: Skolene *prøver ut* ferdigstilte eksempler på kjennetegn på høy og lav måloppnåelse i norsk, matematikk, samfunnsfag, mat og helse på årstrinn med kompetansemål på barnetrinnet

Modell D: Skolene *utvikler selv og prøver ut* kjennetegn på måloppnåelse i ulike fag knyttet til karakterskalaen. Kjennetegnene utvikles på årstrinn med karakter og med utgangspunkt i de generelle karakterbeskrivelsene i forskriften. På ungdomstrinnet og i videregående skole skal det utvikles kjennetegn på måloppnåelse på 10. årstrinn, Vg1 og Vg2. På 10. årstrinn skal det utvikles kjennetegn knyttet til karakterskalaen i norsk, matematikk, samfunnsfag og mat og helse. I videregående opplæring skal det utvikles kjennetegn på måloppnåelse knyttet til karakterskalaen i fellesfagene norsk og matematikk, i

tillegg til fag i studieforbereidende og yrkesfaglige utdanningsprogram. Det skal utvikles kjennetegn felles for karakterene 6 og 5, felles for karakterene 4 og 3 og for karakteren 2.

De ferdigstilte eksemplene på kjennetegn i modell B og C er utarbeidet av Utdanningsdirektoratet i samarbeid med faggrupper. Kompetansemålene i læreplanene er utgangspunktet for utformingen av de nasjonale kjennetegnene på måloppnåelse i fag.

I oppdragsbrev nr. 06 2007 fra Kunnskapsdepartementet til Utdanningsdirektoratet går det fram at målet med utprøvingen er å se om kjennetegn på måloppnåelse i fag på trinn med kompetansemål kan gi en felles nasjonal retning for vurderingsarbeidet og en mer faglig relevant og rettferdig vurdering. Lik vurdering av samme prestasjon er et viktig mål, og tanken er at kjennetegnene skal bidra til et tolkningsfellesskap rundt vurderingen av elevers arbeid. Dessuten anses mer faglig presise tilbakemeldinger om hva elevene mestrer som læringsfremmende, noe som vil føre til økt motivasjon for læringsarbeidet hos elevene. Det er også et mål at læreren i sin vurderingspraksis har et klarere skille mellom vurdering av elevenes faglig kompetanse og deres innsats, interesse osv. Tydeligere tilbakemeldinger til både elever og foresatte vil kunne bidra til økt innsikt i enkeltelevens faglige kompetanse og læringsprosess og til bevisstgjøring blant lærere om betydningen av vurdering i læringsarbeidet. Et annet viktig mål med prosjektet er økt vurderingskompetanse i skolen.

Utdanningsdirektoratet vil gjennom prosjektet sørge for at strukturert kompetansebygging i vurdering og samarbeid mellom lærerutdanning, kommuner, fylkeskommuner og skoler blir en del av utprøvingen.

I utredninger og evalueringer blir det påpekt at systematisk vurdering i liten grad brukes som utgangspunkt for læring i den norske skolen (St.meld. nr. 16, 2006-2007). Individvurdering er imidlertid et sentralt virkemiddel med tanke på å fremme læring og utvikling hos elever og lærlinger for å legge til rette for tilpasset opplæring. Departementets mål er at et helhetlig system for individvurdering skal foreligge fra skolestart høsten 2009 (Oppdragsbrev nr. 06 2007 fra KD).

1.3 Mer rettferdig, faglig relevant og motiverende

I St.meld. nr. 16 (2006-2007) presiseres viktigheten av en mer rettferdig vurdering og at vurdering skal "... bidra til motivasjon og læring" (s. 79). Videre påpekes det at skolens vurderingspraksis skal videreutvikles for å sikre at "... alle elever skal få faglige, realistiske og nødvendige tilbakemeldinger og slik øke læringsutbyttet" (s. 97).

Behov for økt grad av rettferdighet i vurderingssystemet kan illustreres ved en undersøkelse av Hægeland m.fl. (2005). De konkluderer med at lærere i ungdomsskolen gir standpunkt karakterer på en relativ måte som har sammenheng med skolens elevsammensetning. Dersom en elev har mange skoleflinke medelever, blir eleven tildelt en lavere standpunkt vurdering enn hva eleven oppnår som avgangsprøveresultat. Omvendt har elever fått tildelt en relativt høy standpunkt vurdering dersom de har få andre medelever med høyt karakternivå til avgangsprøven. Deres analyser av eksamens- og standpunkt karakterer bygger på data fra alle elever i avgangskullene i 10. klasse våren 2002 og 2003. Analysene viser også at gjennomsnittskarakterene ved skriftlige eksamener er lavere enn standpunkt, mens de for muntlige eksamener er høyere. De konkluderer med at dette enten kan skyldes at

det vurderes feil eller at vurderingene kan være ulike fordi det fokuseres på ulike deler av fagkompetansen. Videre viser undersøkelsen til Hægeland og Kirkebøen (2006) at standpunktkarakterene er svakt økende i norsk og engelsk. Et annet viktig moment er at karakterer er den enkeltfaktoren som entydig viser seg å ha sterkest statistisk effekt på sannsynligheten for om elever blir i videregående skole, eller om de avbryter opplæringen (Utdanningsdirektoratet 2006). Innfridd førsteønske ved opptak til videregående skole, som kan ses som et indirekte mål på elevens karakterer, øker dessuten sannsynligheten for å bli i opplæringen (Grøgaard 1997, Markussen og Sandberg 2004, Støren mfl. 2007). Alle disse eksemplene viser hvor viktig det er at elevvurderingen blir rettferdig.

En viktig målsetting for Kunnskapsløftet er å øke læringsutbyttet for alle elever og lærlinger. Et virkemiddel i denne sammenhengen er faglig relevante tilbakemeldinger. Faglige og relevante tilbakemeldinger fra lærer anses som en viktig del av elevens læringsprosess. Resultatene fra Elevundersøkelsen 2007 viser at ikke alle elever i norsk skole har nok kunnskap om hva som skal til for å bli bedre i faget (Danielsen mfl. 2007). Nesten 60 prosent av elevene svarer at læreren 1-3 ganger per måned eller sjeldnere forteller dem hva de skal gjøre for å bli bedre i fagene. Dessuten svarer 19 prosent av elevene at læreren i noen fag eller i svært få fag ikke har snakket med dem om hva som kreves for å oppnå de ulike karakterene, mens 23 prosent av elevene svarer at i noen fag eller i svært få fag vet de ikke hva som kreves for å oppnå de ulike kompetansemålene. Elevundersøkelsen viste også at hele 44 prosent av elevene på ungdomstrinnet og i videregående opplæring svarer at læreren i noen fag eller i svært få fag ikke spør hvordan de selv vurderer sitt eget arbeid. Disse tallene indikerer at på dette området har skolen et betydelig forbedringspotensial.

Som det er redegjort for i bakgrunnen for prosjektet, har det i Norge helt fram til nå vært tradisjon for å benytte et vurderingssystem som beskriver mangel på mestring for alle elever som ikke får beste karakter. I all opplæring er det imidlertid et viktig prinsipp å vektlegge hva eleven mestrer i stedet for at en benytter betegnelser som ”mangelfull”, ”usikker” eller ”utilstrekkelig” i beskrivelsen av elevens kompetanse. Vurdering som en del av elevenes læringsprosess, er således nært forbundet med spørsmål om motivasjon. En mer motiverende og læringsfremmende vurderingspraksis innebærer at tilbakemeldinger gis i form av tydelige beskrivelser om hva elevene mestrer og hva de bør jobbe videre med for å bedre sin kompetanse ytterligere.

1.4 Deltakende skoler

I alt 77 skoler deltar i den nasjonale utprøvingen av kjennetegn på måloppnåelse i fag. Prosjektskolene kommer fra alle landets fylker. De er delt inn i regioner (Sør, Øst, Vest, Midt og Nord), og hver region har sin kontaktperson i Utdanningsdirektoratet. Det var opprinnelig 78 prosjektskoler, men en av skolene trakk seg fra prosjektet i februar 2008. Av de 77 skolene er 59 grunnskoler. Både barnetrinn, ungdomstrinn og to voksenopplæringsinstitusjoner inngår i utvalget. I tillegg kommer 18 videregående skoler. Sju av de 77 skolene deltok dessuten i et pilotprosjekt i perioden 20. august–9. november 2007. Hensikten med piloteringen var ifølge Utdanningsdirektoratet å få kunnskap om hva som ville være hensiktsmessige rammer og prosesser i hovedutprøvingen ved projektskolene (for sluttrapport fra pilotfasen, se Engh 2007). For pilotkolene er premissene for deltakelse i utprøvingen dermed litt forskjellig fra de andre skolene. De har arbeidet med kjennetegnene i lengre tid, og de har fått mer veiledning. I noen tilfeller er disse skolene derfor behandlet som en egen gruppe i vår undersøkelse.

Tabell 1.1: De 77 skolene fordelt etter modell og trinn

Modell A	Modell B	Modell C	Modell A+D	Modell C+D	Modell C+D	Modell D	Modell D
Barne-trinnet	Barne-trinnet	Barne-trinnet	Barne- og ungdoms-trinnet	Barne- og ungdoms-trinnet	Voksen-opp-læring	Ungdoms-trinnet	Videre-gående
17	18	6	2	10	2	5	18

Tabell 1.1 viser hvordan skolene fordeler seg på de ulike modellene. Til sammen 12 av skolene er kombinerte barne- og ungdomsskoler. Som det går fram av tabellen, prøver disse skolene og de to voksenopplæringsinstitusjonene ut to ulike modeller; én modell knyttet til undervisningen på barnetrinnet og én for ungdomstrinnet. Som tidligere nevnt, kommer prosjektskolene fra alle landets fylker. Tabell 1.2 viser hvordan prosjektskolene fordeler seg etter modeller og regioner.

Tabell 1.2: Fordeling av skoler etter modeller og regioner (Summen overstiger antall skoler, da noen skoler deltar med to modeller, se tabell 1.1)

Regioner	Modell A	Modell B	Modell C	Ungdomstrinnet Modell D	Videregående Modell D
Sør	3	3	1	2	3
Øst	4	5	6	6	6
Vest	3	4	4	3	3
Midt	3	3	4	4	3
Nord	6	3	3	4	3
Totalt	19	18	17	19	18

Av tabellen går det fram at antall skoler som prøver ut henholdsvis modell A, B, C og D, varierer noe fra region til region. Dersom en ser på det totale antallet prosjektskoler innenfor de ulike modellene, er imidlertid fordelingen noenlunde lik.

1.5 Følgeforskning etter oppdrag fra Utdanningsdirektoratet

En forskergruppe ved ILS (Institutt for lærerutdanning og skoleutvikling) og PFI (Pedagogisk forskningsinstitutt) ved Universitetet i Oslo gjennomfører følgeforskning på oppdrag av Utdanningsdirektoratet med utgangspunkt i denne utprøvingen. Dette er første delrapport til Utdanningsdirektoratet fra denne forskningen. Her har vi gjort rede for metoder og funn fra første fase av datainnsamlingen i perioden februar-april 2008. Datainnhentingene har bestått i en kvantitativ og en kvalitativ del. Den kvantitative delen av undersøkelsen har bestått av en spørreskjemaundersøkelse til lærere ved alle skolene som deltar i utprøvingen. Kvalitative data har vi fått gjennom åtte gruppeintervjuer med involverte lærere fra fire deltakerskoler. Resultatene fra disse undersøkelsene ble også presentert i Skolerapport 1, som Utdanningsdirektoratet mottok i juni 2008.

Formålet med følgeforskningen er å gi en vurdering av hvordan de ulike modellene og bruk av felles referanser for kjennetegn på måloppnåelse for fag har fungert for elevene, og hvordan lærerne har opplevd at deres vurderingsarbeid har fungert som følge av utprøvingen. Forskergruppen vil gi tilbakemelding til skolene og oppdragsgiver underveis i prosjektperioden. Formålet med forskningen er også å sikre dokumentert kunnskap om erfaringer med denne type elevvurdering for å kunne gi et godt grunnlag for beslutninger på nasjonalt nivå.

2 TEORETISK FORANKRING

I forrige kapittel ble den utdanningspolitiske og historiske bakgrunnen for prosjektet "Bedre vurderingspraksis" presentert. I dette kapitlet vil vi presentere den teoretiske forankringen som ligger til grunn for følgeforskningens datainnsamling og analyser.

2.1 Målrelatert vurdering for hele grunnopplæringen

Den historiske bakgrunnen for prosjektet viser at vi har hatt to ulike tradisjoner for elevvurdering; vurdering uten bruk av karakterer på barnetrinnet, og med bruk av karakterer fra ungdomstrinnet. Når det gjelder elevenes opplevelse av skolens vurderingspraksis, har overgangen fra barnetrinnet til ungdomstrinnet representert en brå og vanskelig overgang for mange elever. Dette har vært særlig problematisk for de "svake" elevene og gått ut over deres motivasjon for videre læring.

Med innføringen av L-97 ble det signalisert at vurdering på barnetrinnet både skulle ha en ny funksjon og få økt vekt. I læreplanverket står det at "Hovedformålet med elevvurdering er å fremme læring og utvikling", i tillegg til at "... elevene skal være aktive deltakere i arbeidet med vurdering og øve opp evne til å vurdere sitt eget arbeid" (KUF 1996, s. 79). På barnetrinnet fortsatte allikevel nedtoningen av målrelatert vurdering. Målrelatert vurdering har i hovedsak vært knyttet til vurdering med bruk av karakterer fra ungdomsskolen. Vurderingen på barnetrinnet skulle fremdeles fokusere på elevenes framgang med utgangspunkt i elevenes forutsetninger. Dette var først og fremst av hensyn til de "svake" elevene. I for eksempel veiledningsheftet til L-97 uttrykkes det også en viss reservasjon overfor målrelatert vurdering: "Somme elever kan oppleve vurderinga som eit uttrykk for at dei kjem til kort i forhold til felles mål og kriterium" (KUF 1998). Det ble også advart mot hyppig bruk av ensidig vurdering på grunnlag av felles kriterier overfor elever som strever med å nå fellesmålene.

I Læreplanverket for Kunnskapsløftet (2006) videreføres imidlertid en målrelatert elevvurdering, men nå for hele grunnopplæringen. Læreplantenkningen i Kunnskapsløftet bygger på kompetansemål som beskriver hva elevene skal mestre etter endt opplæring innen de ulike fagene etter 2., 4., 7., og 10. trinn i grunnskolen, samt etter hvert av årene i videregående. Kompetansemålene slik de står beskrevet i læreplanene skal være utgangspunktet for vurderingen av elevenes måloppnåelse i fagene.

Når en bruker uttrykk som "hvordan eleven står i forhold til kompetansemålene", tenker en ofte på "graden" av elevens måloppnåelse. En tenker seg i hvilken grad eleven har nådd et gitt kompetansemål. Her er det imidlertid viktig å være oppmerksom på at det er to måter å tenke "grad av måloppnåelse" på. Den ene måten er knyttet til skillet mellom det eleven kan og det eleven ikke kan. Når elevene ikke kan, tenker en på hvor langt eleven er fra å nå et kompetansemål. Da graderer en elevene i relasjon til et gitt kjennetegn på kompetansemålet. Den andre måten å tenke gradering på er at mange elever kan oppnå det samme kompetansemålet, men de kan nå kompetansemålet med ulik kompetanse. Graderingen blir helt ulik i disse to tilfellene. I det ene tilfellet får en gradering av elevene i forhold til et gitt kompetansemål som noen når og andre ikke. Det vil si at noen elever mestrer kompetansemålet, mens andre gjør det ikke. I det andre tilfellet kan elever nå samme kompetansemål, men med ulikt mestringsnivå. Spørsmålet her blir med hvilken grad av kompetanse målet er nådd med.

Ifølge Dale (2008) er det viktig å forstå disse to måtene å tenke gradering på, og han hevder det er behov for å bruke begge disse måtene å gradere på i den faglige tilbakemeldingen til elevene. Mange av målene i læreplanen kan tolkes entydig. Lærerne (og elevene) kan lett fastslå om målene er nådd for eksempel når det gjelder å ”nemne dager, måneder og enkle klokkeslett”, som det blir uttrykt i kompetansemål i matematikk innenfor kompetanseområdet måling etter 2. årstrinn. I slike tilfeller passer det å si at noen elever når målet og andre gjør det ikke, altså den første formen for gradering. Men ikke alle målformuleringer kan tolkes like entydig. I naturfag innenfor hovedområdet ”Mangfold i naturen” er for eksempel et av kompetansemålene etter 4. årstrinn at eleven skal kunne ”samle og systematisere informasjon og beskrive noen utdødde dyrearter og dyregrupper og hvordan de levde”. Et annet eksempel hentet fra naturfag på samme trinn innen hovedområdet ”Teknologi og design” er at eleven skal kunne ”beskrive konstruksjoner og samtale om hvorfor noen er mer stabile og tåler større belastninger enn andre”. Elevenes beskrivelser kan i slike tilfeller spenne fra relativt enkle utsagn til avanserte beskrivelser og argumentasjon hvor eleven forklarer sammenhenger i lys av matematisk kunnskap om målinger og geometri. På denne måten kan *ulike* ferdighets- og kunnskapsnivåer knyttes til *en* formulering av kompetansenivå. Kompetansemålet kan med andre ord nås fra ulike kompetansenivåer eller i ulik grad.

Dale (2008) argumenterer også for at ikke alle kompetansemålene skal vurderes ut fra ulike ferdighets- og kunnskapsnivåer. Det er ikke meningen at hver enkelt lærer skal utvikle ulike kompetansebeskrivelser på hvert enkelt kompetansemål. Et kompetansemål i faget kunst og håndverk etter andre 2. årstrinn innenfor hovedområdet ”Design” er følgende: ”Eleven skal kunne lage enkle gjenstander og former i papir og tekstil gjennom å rive, klippe, lime, tvinne og flette”. Et kompetansemål etter 2. årstrinn innenfor hovedområdet ”Arkitektur” er følgende: ”Eleven skal kunne lage enkle modeller av hus i naturmaterialer”. Til disse to formuleringene er det ikke naturlig å utvikle vurderingsbeskrivelser på ulike kompetansenivåer.

2.2 Kriteriebasert vurdering

Det å tenke elevvurdering ut fra spørsmålet om grad av måloppnåelse, står sentralt i Utdanningsdirektoratets prosjekt ”Bedre vurderingspraksis”. Lysne (1999) hevder at det er et svært omfattende og faglig krevende arbeid å utlede fra målformuleringer et relevant og tilstrekkelig sett av vurderingskriterier som kan angi grad av måloppnåelse innen de ulike fagene og på de ulike trinnene. Dette vil vi utdype nærmere, både i dette avsnittet og i neste avsnitt om taksonomi. Synliggjøring og utvikling av klare kriterier er ifølge Black mfl. (1998) en forutsetning for at elevene skal kunne forstå hva som skal læres og hvorfor. En vurderingsform hvor kvaliteten på elevens kompetanse utelukkende bestemmes på grunnlag av vedkommendes grad av måloppnåelse, omtales vanligvis som *kriteriebasert vurdering*. En nødvendig forutsetning som ligger til grunn for praktiseringen av kriteriebasert vurdering er at det utvikles presise kjennetegn som beskriver mestring på flere kompetansenivåer innen de forskjellige fagene og på de ulike trinnene i opplæringen.

Robert Glaser var den som først introduserte kriteriebasert vurdering som vurderingsform. Han var særlig opptatt av forholdet mellom kriteriebasert og normbasert vurdering. Når det gjelder forholdet mellom disse to vurderingsformene, hevder han at den prinsipielle forskjellen ligger i den standarden eller målestokken som benyttes som referanse i forbindelse med vurderingen; ”*Criterion-referenced* measures depend on an absolute standard of quality while *norm-referenced* measures depend on a relative standard” (Glaser og Klaus 1962, s.

421). Glaser påpeker at vurdering av kompetanse eller prestasjoner på bakgrunn av kriterier kun kan foretas dersom kriteriene eller standardene som vurderingen skal foretas i forhold til, er presist definert *på forhånd*. Ifølge Glaser kan en tenke seg at en elevs kompetanse eller prestasjoner innen et fagområde vil befinne seg langs et kontinuum som strekker seg fra nærmest totalt fravær av kompetanse til fullstendig perfekt kompetanse. For å kunne vurdere elevens kompetanse, er det en forutsetning at en på forhånd har definert hva som kjennetegner kompetanse på de ulike nivåene langs dette kontinuumet. Han hevder derfor at det må utarbeides presise beskrivelser for hva som er minimumskravet til prestasjoner på hvert kompetansenivå, og elevene må mestre dette før de begynner å arbeide mot neste kompetansenivå. En slik vurderingsform gir ifølge Glaser god informasjon om hva eleven faktisk mestrer, i motsetning til normbasert vurdering. Normbasert vurdering gir kun informasjon om hvordan eleven presterer i forhold til resten av gruppen og sier ingen ting om den enkelte elevs faktiske kompetanse.

Ifølge Glaser (1972a, 1972b) vil en kriteriebasert vurderingspraksis også kunne ha positiv effekt på undervisningen. Dette forklarer han med at tydelige beskrivelser av hva som forventes på de ulike kompetansenivåene, vil bidra til å klargjøre for læreren hva som er viktig å vektlegge i undervisningen. Læreren vil med andre ord bli mer bevisst på sin egen undervisning. Han hevder også at en slik vurderingsform krever at læreren utarbeider kriteriebaserte prøver, det vil si prøver som er tilpasset de ulike kompetansenivåene og som har til hensikt å sjekke om elevene virkelig mestrer det som forventes på de ulike nivåene. Gjennom kriteriebaserte prøver vil lærer få direkte informasjon om et kompetansenivå er nådd eller ikke. I tillegg vil det gi tilbakemelding til lærer om undervisningen har vært vellykket, eller om det er grunn til å endre eller justere den for at elevene skal lære bedre.

Når det gjelder vurderingsforskningen i Norge, har Lysne (1999) vært opptatt av hvordan vi kan få et vurderingssystem hvor karakterskalaen forstås og brukes likt blant lærere. Han stiller seg skeptisk til detaljerte målpresiseringer når det gjelder beskrivelser av konkret innhold, slik vi finner i L-97. Han argumenterer for at det utvikles kriteriesett. Det betyr at det utarbeides klare og faglige vurderingskriterier for hvert karakternivå innen de ulike fagene. Dette er særlig viktig i forbindelse med vurdering hvor målene i læreplanen er beskrevet i generelle vendinger, noe som kan gi svært stort rom for ulike tolkninger og subjektivt skjønn i karaktervurderingen. Lysne hevder at jo lavere presisjonsnivå målbeskrivelsene har, desto mer presise må vurderingskriteriene være for at det skal være mulig gi en rettfærdig kvalitetsbedømming av elevenes prestasjoner. Samtidig hevder han at det er svært krevende å utarbeide vurderingskriterier fordi målene i læreplanen som regel er formulert i så generelle vendinger at de er vanskelige å konkretisere. Ifølge Lysne vil imidlertid målrelatert vurdering som ikke gjøres i henhold til et adekvat kriteriesett, i praksis som regel bli normbasert. Det har vist seg at det er lettere for lærere å foreta vurdering ved å sammenligne en elevs prestasjoner med andre elevs prestasjoner. I pedagogisk sammenheng oppstår denne praksisen som regel når målbeskrivelsene er lite presise og det ikke er utarbeidet klare kriterier for de ulike kompetansenivåene eller karaktergradene. Han hevder derfor at i vurderingssammenheng må generelle målbeskrivelser representeres ved et sett av felles kriterier som lærerne så langt det er mulig må bruke. Videre påpeker han at vurdering i forhold til målkriterier forutsetter at oppgavenivået på prøver til en hver tid må være avstemt i forhold til de kriteriene som elevene skal vurderes i forhold til. Lysne hevder at sentrale skolemyndigheter bør ha ansvaret for at arbeidet med målbaserte vurderingskriterier blir utført, men at det må gjøres av fagfolk som har pedagogisk kunnskap og erfaring. Slikt kriteriearbeid anser han som særlig viktig i forbindelse med avsluttende standpunktvurdering og eksamen.

2.3 Taksonomi

I prosjektet ”Bedre vurderingspraksis” arbeides det etter modeller som beskriver høy og lav måloppnåelse på barnetrinnet. Når det gjelder modellen som er rettet mot karakterer, skal grad av måloppnåelse knyttes til karakterskalaen. Det å tenke i forhold til høy og lav måloppnåelse og i forhold til karakterskalaen, krever ikke bare forståelse av kriteriene, men også forståelse av det som kalles taksonomi. Taksonomi er et slags klassifiseringssystem. Analyser av ulike taksonomisystemer er derfor en viktig diskusjon når det gjelder kriteriebasert vurdering.

Benjamin Bloom (1956) var forskningsleder og mest kjent for sitt arbeid med å klassifisere og operasjonalisere pedagogiske mål, også kalt Blooms måltaksonomi. Behovet for et slikt klassifiseringssystem ble begrunnet med at målene danner grunnlaget for utformingen av læreplaner og undervisningsprogrammer, tester og prøver. En forventet dessuten at måltaksonomi ville lette kommunikasjonen innad i skolen og utad overfor blant annet foreldre. Formålet med klassifiseringsarbeidet var å definere nærmere de mål og delmål som undervisningen skulle sikte mot. Bloom har i sin måltaksonomi ensidig lagt vekt på kognitive læringsmål i skolen, og han klassifiserte kognitiv mestring som det mest sentrale læringsområdet. De enkelte målene ble organisert hierarkisk. Det innebar at han rangerte dem i seks nivåer etter hvor avanserte de var:

1. Kunnskap
2. Redegjørelse
3. Anvendelse
4. Analyse
5. Syntese
6. Evaluering

Den hierarkiske organiseringen av de kognitive læringsmålene i Blooms taksonomimodell innebærer at hvert nivå (f.eks. anvendelse) forutsetter at eleven mestrer det som kreves på lavere nivå (dvs. kunnskap og redegjørelse). Bloom anser *kunnskap* som det mest sentrale i undervisningen. Han forbinder kunnskap med det å huske spesifikke fakta. I læringssituasjonen forventes det at eleven lagrer ulike typer informasjon i sin bevissthet. Dette kan omfatte informasjon om for eksempel spesifikke hendelser eller størrelsen på et bestemt fenomen. På dette første nivået inngår også måter som eleven bruker for å organisere slike isolerte informasjonsbiter i hukommelsen. I tillegg forventes det at eleven kan hente fram den lagrede informasjonen senere, for eksempel i forbindelse med prøver. Det andre nivået, *redegjørelse*, innebærer at eleven skal kunne beskrive og i noen grad utdype med egne ord det en har lært, mens det tredje nivået, *anvendelse*, blant annet innebærer at eleven kan ta i bruk lærte prinsipper og regler i løsning av ulike typer problemer. Det fjerde nivået, *analyse*, representerer dyktighet i å dele opp eller bryte ned helheter i elementer og være i stand til å forstå sammenhengen mellom dem i den totale konteksten, mens det femte, *syntese*, referer til å kunne sette sammen deler eller elementer slik at de danner en helhet. *Evaluering* anses som det mest avanserte nivået i modellen og henspiller på dyktighet i å foreta ulike former for vurderinger. Blooms taksonomimodell er senere blitt revidert av Anderson og Krathwohl (2001), men prinsippet er det samme. I revisjonen har en identifisert nivå 1 til hukommelse, nivå 2 til forståelse og beholdt anvendelse og analyse som nivå 3 og 4, men endret evaluering til nivå 5 og omdøpt syntese til kreativitet som nivå 6. Dessuten er det i den reviderte modellen utviklet fire typer kunnskap: faktisk kunnskap, begrepskunnskap,

prosedyrekunnskap og metakognitiv kunnskap. De tre første typene av kunnskap inngikk også i den første utgaven av modellen.

Blooms taksonomi omfatter ikke bare det kognitive området, som er beskrevet over i seks ulike nivåer. En viktig side ved hans taksonomi er at det skilles mellom det kognitive, affektive (holdning) og psykomotoriske området (ferdigheter). Dette er også en tenkning vi finner igjen i våre læreplaner når det gjelder skillet mellom kunnskap, ferdigheter og holdninger. I forbindelse med prosjektet "Bedre vurderingspraksis" er faget Mat og helse med. Her er det nødvendig å overskride den tradisjonelle taksonomitenkningen ved at ferdigheten til å lage mat kan forbindes med det kognitive aspektet. Holdning er også en del av faget.

En kan rette innvendinger ikke bare mot skillet mellom det kognitive, holdninger og ferdigheter, men også mot Blooms taksonomitenkning innenfor det kognitive. En mulig følge av hans taksonomitekning er for eksempel at "å beskrive..." og "gjøre rede for..." alltid uttrykker lav måloppnåelse i et fag, i forhold til "å forklare ..." og "vurdere..." som er uttrykk for høy måloppnåelse. Det kan imidlertid tenkes situasjoner hvor det å beskrive eller gjøre rede for kan være uttrykk for "høy måloppnåelse", som for eksempel når eleven skal beskrive kompliserte situasjoner eller problemstillinger. En elev kan ha "lav måloppnåelse" ved å analysere forholdet mellom to fenomener når situasjonen er velkjent og enkel. En må derfor ikke se isolert på verbene som beskriver måloppnåelse. En må relatere beskrivelsen til kompleksiteten i den aktuelle situasjonen. Det er derfor grunn til å stille seg kritisk til en slik taksonomitenkning som Bloom står for.

Et alternativ er å se nærmere på taksonomisystemet som PISA bygger på. I PISA benyttes formuleringer om hva som kreves for å mestre hver av oppgavene innen henholdsvis lesing, matematikk og naturfag, og det gis en generell beskrivelse av kjennetegn ved kompetansen til elever på de ulike nivåene (Lie mfl. 2001, Kjærnsli mfl. 2004, 2007). Her finner vi en annen måte å vurdere på og beskrive ulike kompetanse- eller ferdighetsnivåer på. Dersom vi tar for oss lesing, er det å "finne informasjon" ett av kompetansemålene. Innen dette kompetanseområdet er det utviklet fem ferdighetsnivåer som eleven vurderes i forhold til. Beskrivelsene på de ulike nivåene er utformet slik at kriteriene for vurderingen går fra det enkle til det avanserte. Tilsvarende er gjort for de andre kompetanseområdene innen lesing ("tolke, forstå" og "vurdere, reflektere"). I matematikk foretas vurderingen av elevenes kompetanse i forhold til seks ulike ferdighetsnivåer (fra enkel til avansert kompetanse) knyttet til kompetanseområdene "forandring og sammenheng", "rom og form", "tall og mål" og "usikkerhet". Tilsvarende er gjort for naturfag. Overført til elevvurdering i skolen betyr en slik tankegang at det er mulig å ha måloppnåelse selv på et lavt kompetansenivå.

Beskrivelsene av de ulike ferdighetsnivåene i PISA er utviklet i en kombinasjon av teori og empiri (Lie mfl. 2001). I for eksempel lesing er det som nevnt utviklet fem kompetansenivåer som elevenes besvarelse skal vurderes i forhold til, og hvert av spørsmålene i testen er definert inn i ett av disse fem ferdighetsnivåene. Nivåene er definert etter vanskegrad, der nivå 1 regnes som det letteste og nivå 5 som det mest kompliserte. Når oppgavens nivå skal defineres, er det avgjørende hvilke lese- og tankeoperasjoner som må foretas for å svare riktig. Noen spørsmål krever kun at en henter ut enkel informasjon fra teksten, mens andre spørsmål krever at leseren foretar kompliserte slutninger eller tolkninger på bakgrunn av informasjon som både kan være eksplisitt og implisitt formulert. Ut fra en teoretisk vurdering av hva som synes å definere vanskegraden for en oppgave, blir oppgaven foreløpig plassert på et bestemt ferdighetsnivå. Til syvende og sist er det imidlertid hvor godt elevene svarer, med

andre ord elevenes vurdering av vanskegraden, som er avgjørende for nivåplasseringen. Noen oppgaver kan i praksis vise seg å være vanskeligere enn antatt ut fra den teoretiske analysen. Dette krever en ny gjennomgang av hva hver oppgave krever, og vil kunne medføre at definisjonene av nivåene må endres noe slik at oppgavene passer inn på det nivået der de ut fra løsningsfrekvensen hører hjemme. Nasjonale prøver med sine beskrivelser av mestringsnivåer er inspirert av PISAs måte å beskrive kompetanse- eller ferdighetsnivåer på.

2.4 Vurdering for læring

Overfor er beskrevet overgangen fra normrelatert til målrelatert vurdering. Den målrelaterte er knyttet til kriterier. Kriteriene kan igjen stå innenfor ulike former for taksonomi. Dette er vesentlige momenter i prosjektet "Bedre vurderingspraksis" hvor lærerne skal forholde seg til elevenes mestring på ulike ferdighetsnivå. Når læreren bruker kriteriene, er intensjonen å forbedre elevenes læring. Å bruke kriterier i tilbakemeldingene til elevene, kan forbindes med det etter hvert kjente uttrykket *vurdering for læring*.

Faglige vurderinger og tilbakemeldinger som har til hensikt å forbedre elevenes kompetanse, blir gjerne omtalt som *formativ* vurdering (Engelsen 2006). Den formative vurderingen skal bidra til å lede, korrigere og styrke læreprosessen gjennom løpende tilbakemeldinger til eleven om faglig framgang og mestring i forhold til mål og de kriteriene som er satt for måloppnåelse. Slike løpende tilbakemeldinger skal være en gjennomgående del av undervisningen og blir i Forskrift til opplæringslova (2006) omtalt som *underveisvurderinger*. Underveisvurderinger skal bidra til å fremme elevenes læring, utvikle deres kompetanse og gi grunnlag for tilpasset opplæring (jfr. § 3-3, § 3-4, § 4-4, § 4-5). Dette innebærer at elevene underveis i opplæringen skal få tydelige tilbakemeldinger på hvor de står i forhold til kompetansemålene. Ifølge Haugstveit (2005) har skolen en utfordring i å utforme en pedagogikk der formativ vurdering inngår tydelig i opplegg og læringsprosesser. Lærerne gir uttrykk for at ansvaret for dette i for stor grad overlates til enkeltlæreren.

Haugstveit (2005) fant i sin undersøkelse at det blant lærere er en økende forståelse for vurdering som et viktig element i elevenes læring. Formativ vurdering, som vektlegger vurderingen underveis i prosessen for å oppnå et best mulig læringsutbytte, ses ofte i sammenheng med "vurdering for læring". I internasjonal forskning er det særlig Black og kolleger som har studert formativ vurdering i betydningen vurdering for læring (assessment for learning). Black mfl (1998) gikk igjennom mer enn 250 forskningsartikler som viste at formativ vurdering hadde positiv effekt på elevenes faglige prestasjoner, og at en gjennom å forbedre den formative vurderingspraksisen kunne øke elevenes læringsutbytte. Forbedring forutsetter imidlertid at vurderingskriteriene er tydelige for både elever og lærer. Gjennomgangen viste også at det var helt avgjørende for elevenes utbytte av den faglige tilbakemeldingen at læreren også ga elevene konkret hjelp til å forstå *hvordan* de skal kunne forbedre seg. Kjennskap til kriterier for måloppnåelse er med andre ord ikke tilstrekkelig i seg selv. Eller sagt på en annen måte; elevene bør gjennom *framovermeldinger* både få kjennskap til hvor langt de er kommet i forhold til de ulike læringsmålene, hva som gjenstår og hvordan de bør arbeide for å oppnå høyere grad av måloppnåelse. På denne måten griper den formative vurderingen inn i selve læringsprosessen. Samtidig blir vurderingen et viktig redskap i det videre arbeidet med eleven. Black mfl (1998) konkluderer med at formative vurderingsmetoder er særlig nyttig overfor elever med svake faglige prestasjoner.

Sadlers (1998) syn på elevvurdering er helt i tråd med en slik oppfatning. Han opererer med tre nødvendige elementer i den faglige tilbakemeldingen til elevene: 1) et mål som skal nås, 2) en beskrivelse av elevens nåværende ståsted og 3) en forståelse av hvordan gapet mellom de to kan lukkes. Alle disse tre momentene må være forstått før en kan sette i gang handling med tanke på å forbedre læring. En viktig forutsetning for at elevene skal kunne forstå og lykkes med dette, er ifølge Black mfl (2006) at de trenes i *egenvurdering*, der eleven lærer å bruke de oppsatte kriteriene for måloppnåelse. De hevder at egenvurdering er en krevende prosess fordi det krever at elevene har et metaperspektiv på egen læringsaktivitet og læringsutbytte. Mange lærere hevder imidlertid at de mangler kunnskap om hvordan de kan fremme de metakognitive ferdighetene som egenvurdering krever. Haugstveit (2005) hevder at en i norsk skole har fokusert på verktøy for å måle resultater, mens en mangler verktøy for ”vurdering for læring”.

Når det gjelder *summativ* vurdering (sluttvurdering), gir den en beskrivelse av elevens kompetansenivå på et bestemt tidspunkt (Engelsen 2006). Den summative vurderingens primære formål er ifølge Lysne (1999) å gi en vurdering av læringsresultater på ulike tidspunkt eller stadier i et læreprogram og er et mål på ”hvor mye” eleven har lært til nå, enten innen en bestemt fagmodul eller i forhold til hva eleven har oppnådd totalt i faget. Betegnelsen sluttvurdering blir ofte assosiert med uttrykket ”vurdering av læring”. Fordi ”vurdering av læring” og ”vurdering for læring” i utgangspunktet tjener ulike hensikter, opereres det i mange sammenhenger med et skille mellom disse to vurderingsformene. Black m.fl. (2006) argumenterer imidlertid for at også summative prøver kan bidra positivt til elevenes læring ved at de brukes formativt i undervisningen. Dette forutsetter at eleven forstår at prøver og prøveresultater kan brukes til å forbedre egen læring, noe Black m.fl. hevder best lar seg fremme gjennom elevinvolvering. I den forbindelse foreslår de følgende (Black m.fl. 2006, s. 56):

- Students should be engaged in a reflective review of the work they have done to enable them to plan their revision effectively.
- Students should be encouraged to set questions and mark answers to help them, both to understand the assessment process and to focus further efforts for improvement.
- Students should be encouraged through peer- and self-assessment to apply criteria to help them understand how their work might be improved.

Hovedpoenget med disse forslagene er at summative prøver kan bidra positivt til elevenes læring.

2.5 Kontinuitet – et enhetlig vurderingssystem

I forbindelse med prosjektet ”Bedre vurderingspraksis” skal tre modeller for kjennetegn på måloppnåelse i fag med kompetansemål prøves ut på barnetrinnet. I tillegg skal en modell på kjennetegn på måloppnåelse i fag knyttet til karakterskalaen prøves ut på ungdomstrinnet og i videregående skole. Spørsmålet er om modellene fungerer som et godt verktøy for å skape sammenheng mellom vurdering underveis på de ulike trinnene og til slutt i opplæringsforløpet. Til nå har det imidlertid som påpekt ovenfor, eksistert ulike vurderingstradisjoner i norsk skole. Vurdering uten bruk av karakterer har etablert seg som tradisjon på barnetrinnet og vurdering med bruk av karakterer har vært tradisjonen på ungdomstrinnet og i videregående opplæring. Prosjektet legger opp til at det er en relasjon med faglig relevante tilbakemeldinger gjennom hele grunnopplæringen, og for første gang i historien har vi gjennomløpende læreplaner for hele grunnopplæringen. I ungdomsskolen og

videregående opplæring kreves det en felles begrepsbruk, beskrivelse og forståelse av karakternivåene slik at en mer enhetlig oppfatning av hva karakterene uttrykker kan utvikles. Klare kjennetegn på måloppnåelse i fagene kan gi en mer rettferdig og faglig relevant vurdering av elevens måloppnåelse. Kriteriebasert vurdering representerer en måte å tenke vurdering av fagkompetanse på som også er relevant å bruke på barnetrinnet. Gjennom å bruke en kriteriebasert vurdering kan det utvikles en kontinuitet i den faglige vurderingen i grunnopplæringen, selv om barnetrinnet ikke har vurdering med bruk av karakterer.

Gjennom erfaring med gradering av måloppnåelse i fag på barnetrinnet, vil elevene gradvis kunne lære at en prestasjon kan vurderes ut fra ulike kompetansekriterier. Elevens kjennskap til den faglige vurderingen utvikles gradvis. Det skjer ved at kriteriene blir gjentatt, og at eleven forstår regelmessigheten i bruken av vurderingskriteriene. En innføring av målrelatert vurdering ved hjelp av kriterier i barneskolen kan med andre ord føre til at overgangen til ungdomsskolen ikke blir så vanskelig for mange elever. Selv om elever på barnetrinnet ikke får karakterer, kan vurderingen følge de samme prinsipper. Ved at vurderingskriteriene som er i bruk på barnetrinnet videreføres på ungdomstrinnet og i videregående opplæring, men nå med karakterer, vil det altså bli en kontinuitet i skolens vurderingsarbeid. Elevene vil på den måten ikke oppleve noe brudd i vurderingen i overgangen fra barnetrinnet til ungdomstrinnet og senere videregående opplæring. Å utvikle felles referanser som kommer i regelmessig bruk, øker dessuten muligheten for en rettferdig vurderingspraksis.

3 METODE

3.1 Design

I vår følgeforskning om utprøvingen benyttes både en kvantitativ og en kvalitativ tilnærming for innhenting av data. Datainnhentingene består av til sammen fire undersøkelser som gjennomføres i løpet av to perioder; første gang i februar–april 2008 og andre gang i oktober–november 2008. Ved begge anledninger benyttes både spørreskjema og intervju, men i noe ulikt omfang. Spørreundersøkelser gir god oversikt over problemstillinger som en ønsker å belyse, og svar kan hentes fra brede grupper av respondenter. Ved intervjuer kan man gå i dybden og få større innsikt i de temaene som tas opp, men bare et fåtall respondenter kan nå på denne måten. Tabell 3.1 viser hvordan undersøkelsen er organisert i forbindelse med datainnhentingene.

Tabell 3.1: Tidspunkt, instrument og informanter i de ulike undersøkelsene

	Tidspunkt	Instrument	Informanter
Undersøkelse 1	februar–mars 2008	spørreskjema	- alle lærerne som underviser i bestemte fag og på bestemte trinn ved alle prosjektskolene
Undersøkelse 2	april 2008	intervju i fokusgrupper	- lærerne ved et utvalg av prosjektskolene
Undersøkelse 3	oktober 2008	intervju i fokusgrupper	- lærerne ved et utvalg av prosjektskolene - enkelte elever fra bestemte trinn ved prosjektskolene - ledelsen ved skolen - representant for skoleeier
Undersøkelse 4	november 2008	spørreskjema	- alle lærerne som deltar i prosjektet ved den enkelte skole - et utvalg av elever ved prosjektskolene - ledelsen ved alle prosjektskolene - representanter for skoleeierne - ressurspersoner fra UH-sektoren/ andre kompetansemiljøer

Følgeforskningen er godkjent av NSD (Norsk samfunnsvitenskapelig datatjeneste).

3.2 Deltakelse på samlinger og møter

I forbindelse med prosjektstart arrangerte Utdanningsdirektoratet i månedsskiftet november/desember 2007 regionale oppstartseminar for prosjektskolene. Representanter fra vår forskergruppe var til stede ved samlingene for region Øst og Sør. De var også representert på samling for skoleeiere 25. januar og 14. mai 2008. Leder av forskergruppen var også til stede på et møte for referansegruppen for Nasjonalt kvalitetsvurderingssystem for å gi en kort orientering om følgeforskningen. Denne underveisrapporten er for øvrig blitt presentert og drøftet i et møte med Felles referansegruppe for forskning i Utdanningsdirektoratet 29. august 2008. Representanter fra forskergruppen var også til stede på fellessamlingen for alle prosjektskolene 4. september 2008.

3.3 Undersøkelse 1

Som det går fram av tabell 3.1, ble det i første datainnhentingsperiode aller først gjennomført en spørreundersøkelse. Spørreundersøkelsen var rettet mot *alle lærerne* ved prosjektskolene som underviser i de aktuelle fagene og på de aktuelle trinnene. Når det gjelder barne- og ungdomstrinnet, besto utvalget i den første spørreundersøkelsen av lærere som underviser i fagene norsk, matematikk, samfunnsfag samt mat og helse på 2., 4., 7., og 10. trinn. I videregående skole deltok lærere som underviser i norsk og matematikk samt andre fag i studieforbereende og yrkesfaglige utdanningsprogrammer på Vg1 og Vg2.

Spørreskjemaene ble sendt skolene 15. februar med beskjed om å returnere dem så fort som mulig og senest innen 7. mars. I forkant av spørreundersøkelsen (i januar) ble spørreskjemaet prøvd ut blant lærerne ved prosjektets pilotskoler. Det endelige spørreskjemaet, som ble diskutert med oppdragsgiver før utsending, inneholdt spørsmål som ga opplysninger om (her i rekkefølge som i spørreskjemaet):

- den enkelte lærer (kjønn, antall års erfaring, fag, trinn, utdanning), *spm 1-7*
- lærernes syn på hva som er god elevvurdering, *spm 8*
- den enkelte skoles praksis i elevvurdering før prosjektstart, *spm 9 -10*
- skolens situasjon i prosjektet så langt, *spm 11-12*
- skolens kompetanseutvikling i elevvurdering, *spm 13-18*
- elevdeltakelse i elevvurdering i fag, *spm 19*
- elevenes motivasjon for læring (ikke omtalt i denne rapporten), *spm 20*
- lærernes holdninger til kjennetegnene og prosjektet som helhet, *spm 21-22*
- forholdet mellom læreplan og kjennetegn, *spm 23*
- bruk av kjennetegn i egen vurderingspraksis, *spm 24*
- forholdet mellom kjennetegn og karakterer, *spm 25*

Lærerspørreskjemaet besto av i alt 25 ”spørsmål” eller grupper av spørsmål i samme format og innen samme tema. I oversikten ovenfor er det henvist til spesifikke deler av spørreskjemaet. To spørsmål var åpne, mens resten av spørsmålene var lukkede og skulle besvares ved avkryssing. To av de siste gruppene av spørsmål (*spm 23-24*, som gjaldt forholdet mellom læreplan og kjennetegn, og bruk av kjennetegn i egen vurderingspraksis) skulle bare besvares dersom lærerne var i gang med utprøving av kjennetegnene i eget arbeid. Det siste spørsmålet (*spm 25*) gjaldt forholdet mellom kjennetegn og karakterer og skulle bare besvares av lærere på ungdomstrinnet og i videregående skole. Spørreskjemaet er i sin helhet tilgjengelig som vedlagt pdf-fil og på www.skolenettet.no/vurdering.

I forkant av spørreundersøkelsen ble skolene bedt om å rapportere til oss hvor mange lærere ved skolen som deltar i prosjektet. Det ble presisert at det bare var lærere som underviser i de aktuelle fagene på de aktuelle trinnene, som skulle delta. Skolenes opplysninger om dette skulle bl.a. benyttes som grunnlag for å beregne svarprosenten. Ifølge skolenes egen rapportering skulle til sammen 1010 lærere delta i den første spørreundersøkelsen. Det knytter seg imidlertid stor usikkerhet til dette tallet, da det virket som om flere av skolene ikke selv var klar over hvilke lærere som egentlig deltok i prosjektet. Det er grunn til å tro at det tallet enkelte skoler oppga kan være for høyt, særlig ved at noen skoler har tatt med lærere som ikke underviser på de aktuelle trinnene, men som likevel har vært engasjert i utvikling av kjennetegn. Det knytter seg derfor stor usikkerhet til svarprosenten på lærernivå i undersøkelsen (se 4.2).

Videre er det ulikt antall skoler på de ulike modellene i den enkelte region, og disse skolene er nokså forskjellige med hensyn til størrelse og beliggenhet. Skolene i hver region er samlet på noen få kommuner. Alt dette får konsekvenser for vår strategi for analyse og rapportering av data.

3.4 Undersøkelse 2

For å kunne gå i dybden og få større innsikt i de temaene som prosjektet var rettet mot, ble det gjennomført intervju med lærere ved enkelte av prosjektskolene. Intervjuundersøkelsen (undersøkelse 2) ble gjennomført i april 2008 og omfattet intervju med lærere på fire skoler. I valg av skoler ble det sørget for en viss geografisk spredning og at alle modellene som skulle prøves ut, var representert. Skolene kom fra tre ulike regioner. Av grunnskolene var en av pilotskolene med i utvalget, det var også en kombinert barne-/ungdomsskole som prøvde ut to modeller. Når det gjelder videregående, ble en skole med både studieforberevende og yrkesfaglige utdanningsprogram valgt ut.

Den kvalitative undersøkelsen var organisert som intervju med lærere i grupper. Gruppeintervjuer anses som særlig godt egnet for å få svar på deltakernes erfaringer og deres nærmere utdypelse av dem. Når diskusjon i grupper er strukturert omkring et relativt begrenset tema, benyttes gjerne betegnelsen ”fokusgrupper” (Stewart mfl 2007). Fokusgruppeintervjuer kjennetegnes dessuten ved at deltakerne har anledning til mer fritt å diskutere bestemte temaer.

Det ble gjennomført intervjuer med i alt åtte lærergrupper bestående av til sammen 49 lærere. Fokusgruppene besto av lærerne ved den enkelte skole som deltok i prosjektet med de aktuelle fagene og på de aktuelle trinnene, og som i tillegg hadde besvart spørreskjemaet. Når det gjelder modell A og B, ble det gjennomført to gruppeintervjuer på hver skole med lærere fra henholdsvis småskole- og mellomtrinn. På den kombinerte barne-/ungdomsskolen skolen som prøver ut både modell C og D, ble det dannet fokusgrupper med lærere fra henholdsvis barne- og ungdomstrinn. I videregående skole var ikke fokusgruppene organisert verken etter trinn eller fag; dvs. at lærerne i samme gruppe underviste både på Vg1 og Vg2 og innen både studieforberevende og yrkesfaglige utdanningsprogram. Gruppene varierte noe i størrelse og besto av fra fem til åtte personer. Tabell 3.2 gir en oversikt over de gjennomførte intervjuene.

Tabell 3.2: Oversikt over gjennomførte gruppeintervjuer (fokusgrupper).

Modell	Antall lærere	Antall fokusgrupper				
		Småskole-trinnet	Mellom-trinnet	Ungdoms-trinnet	Videregående, studieforberevende fag	Videregående, yrkesfag
Modell A	11	1	1			
Modell B	16	1	1			
Modell C	5	1				
Modell D – ungdomstrinnet	6			1		
Modell D – videregående	11				1	1

To personer fra forskergruppa tilbrakte én dag på hver skole i forbindelse med intervjuene. Den ene hadde ansvaret for å lede selve intervjuet, mens den andre hadde i oppgave å notere

det som kom fram i løpet av samtalen. Notatene ble renskrevet umiddelbart etterpå. Hvert gruppeintervju varte i 1 ½ time. Intervjuguiden besto av følgende hovedpunkter:

- Innledende informasjon om oss og vår rolle i følgeforskningen
- Skolens vurderingspraksis før prosjektstart
- Informasjon de har mottatt om prosjektet
- Hvor langt skolen er kommet i prosjektet
- Organisering av arbeidet
- Utvikling av kjennetegn
- Kontakt med eksterne kompetansemiljøer
- Erfaringer med bruken av kjennetegn
- ”Nyttig” og ”lærerikt”
- Anbefalinger

Intervjuguiden inneholdt også mulige oppfølgingsspørsmål innenfor de områdene som en ønsket mer informasjon om.

I løpet av intervjuene deltok lærerne aktivt og bidro både med egne erfaringer og refleksjoner, samt i diskusjoner som oppsto i gruppen. De var svært engasjerte, og vi opplevde en åpen og tillitsfull atmosfære under intervjuene. I denne sammenhengen var det nyttig at vi som intervjuere hadde god kjennskap til praksisfeltet gjennom mange års praksis i skolen. I etterkant av intervjuene ble intervjudataene bearbeidet og analysert og vurdert i lys av resultatene og funn i Undersøkelse 1.

3.5 Analysemetoder

Undersøkelse 1 var en kvantitativ undersøkelse med nesten bare avkryssing av ferdig oppstilte kategorier (se spørreskjemaet i vedlegg 2). De få åpne spørsmålene ble i sin helhet kodet av to personer i samarbeid etter på forhånd oppsatte kriterier/koder. Disse var i hovedsak utviklet på bakgrunn av svarene til lærerne som deltok i piloteringen av spørreskjemaet.

De statistiske analysene ble foretatt med statistikkprogrammet SPSS. I hovedsak dreide det seg om enkel deskriptiv statistikk, og de fleste av resultatene er gitt i tabellform med sammenlikninger mellom hovedtrinn (barnetrinn, ungdomstrinn og videregående) og for barnetrinnets vedkommende også modell (A, B eller C). For at tabellene ikke skal bli for omfattende, har vi valgt å oppgi de fleste resultatene i form av prosent av antall svar uten hver gang å oppgi det absolutte antall.

Når det gjelder forskjeller mellom svarene på tvers av trinn og modeller, har vi forsøkt å løse et dilemma på en enkel måte. Utvalget av skoler representerer på ingen måte et tilfeldig utvalg, og generalisering ut over de deltakende skoler og lærere er i prinsippet ikke mulig. På den annen side er det liten grunn til å kommentere forskjeller hvis de ikke peker seg tydelig ut. For dette siste formålet har vi angitt med en stjerne (*) i tabellene der forskjellene er ”signifikante” i betydningen at en slik forskjell ikke med rimelig sannsynlighet (<5 %) kunne oppstå ved en ren tilfeldighet HVIS lærerne hadde vært tilfeldig trukket fra en og samme populasjon.

Resultatene fra spørreundersøkelsen blir som nevnt ovenfor gjennomgående illustrert som prosenter på lærernivå. Antall deltakere på spørreundersøkelsen fra hver av de 72 deltaker-skolene i vår datafil varierer mye (fra 3 til 24 lærere), så vi har i rapporten valgt å ikke illustrere resultatene på skolenivå. Dette valget er også tatt på bakgrunn av videre analyser av datamaterialet. Når vi går inn og undersøker svarene på spørsmål som dreier seg om forhold på skolen som helhet, finner vi forskjeller i svarene fra lærerne innad på mange av skolene. Opplevelsen av hvor skolen befinner seg i forhold til ulike sider ved prosjektet, er altså ikke entydig blant lærere på samme skole. Dette kan illustreres ved svarene på spørsmålet om skolen har en plan for vurdering før oppstart av prosjektet (*spm 9*). Her svarer lærerne likt innad bare ved omtrent 40 % av skolene. Det er også bare rundt eller under 40 prosent enighet innad blant svarene på spørsmålene om skolen samarbeider med andre skoler (*spm 14*), om skolen har knyttet til seg eksterne fagmiljøer i sin kompetansebygging i elevvurdering (*spm 16*), og hvordan skolen har organisert arbeidet i prosjektet (*spm 12*).

I tillegg til motstridende svar, er det på mange av skolene stor usikkerhet knyttet til disse spørsmålene. For eksempel på spørsmålet om skolen har knyttet til seg eksterne kompetansemiljøer (*spm 16*), svarer hele 30 prosent av lærerne "Vet ikke". Total enighet (om enten "Ja" eller "Nei") blant lærerne på en og samme skole i svaret på dette spørsmålet finner vi bare på 11 av 72 skoler. (Her kan imidlertid grunnen være en viss usikkerhet om hva som kan inkluderes i begrepet "eksterne fagmiljøer".)

Når det gjelder data fra gruppeintervjuene, har vi bestrebet oss på å finne fram til representative uttalelser som dekker typiske holdninger blant lærerne. Slike sitater er tatt med i rapporten for å illustrere og utdype funn i spørreundersøkelsen eller for å illustrere signaler som i stor grad går igjen i fokusgruppene og derfor er representative for de lærerne som er intervjuet. Disse uttalelsene er ikke tilfeldig utvalgte fra vårt materiale. Vi kan likevel ikke si noe pålitelig om hvor vanlig slike holdninger er for hele prosjektet, men i de fleste tilfellene bidrar uttalelsene til å forstå bedre hva som ligger bak svarene på de strukturerte spørsmålene i spørreskjemaet. Slik sett har vi erfart at de to metodene, spørreundersøkelsen og gruppeintervjuene, har utfyllt hverandre på en god måte.

4 RESULTATER

4.1 Innledning

Når vi videre i rapporten refererer til *prosjektet*, mener vi arbeidet med *Utprøving av kjennetegn på måloppnåelse i fag* ute på skolene. Utprøvingen er som tidligere forklart en del av Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis*.

Rapporten inneholder mange sitater fra lærere, både fra de åpne spørsmålene i spørreskjemaet og fra intervjuene. Etter hvert sitat benytter vi bokstavene b, u og v, som referer til henholdsvis barnetrinn, ungdomstrinn og videregående skole. Ved henvisning til bestemte spørsmål i spørreskjemaet brukes som i kapittel 3 betegnelsen *spm* etterfulgt av nummeret på spørsmålet i skjemaet (for eksempel *spm 15*).

I presentasjonene av resultater i tabellform har vi noen ganger valgt å ta med bare de spørsmålene der det er forskjeller av betydning mellom trinn eller modeller, oppsiktsvekkende store likheter, eller der svarene i betydelig grad avviker fra gjeldende lovverk og forskrifter. Det betyr at vi i en del av tabellene utelater enkelte spørsmål der vi mener at svarene ikke er av nevneverdig betydning.

4.2 Lærere

Spørreskjema ble sendt ut 15.02.08. Da svarfristen gikk ut 7. mars, hadde 29 skoler returnert lærerspørreskjemaene (lærerne hadde ca. tre uker på seg til å besvare spørreskjemaet). Dagen etter mottok vi spørreskjema fra ytterligere 21 skoler (det vil si at i alt 50 av 77 skoler hadde besvart). For å få inn spørreskjemaene fra resten av skolene, ble det gjennomført to ringerunder. Den første purringen ble foretatt 10. mars, mens siste ringerunde foregikk 14. mars (på dette tidspunktet manglet fortsatt ti skoler). Det ble satt strek for datainnhenting tre uker etter at fristen for innsending var gått ut (28. mars). Da manglet fortsatt tre skoler.

I utgangspunktet skulle alle lærere som deltok i prosjektet på de aktuelle trinnene og i de aktuelle fagene svare på spørreskjemaet (se 3.3). Tar en utgangspunkt i skolenes egne tall over hvor mange lærere som deltar i prosjektet, gir dette en svarprosent på noe over 70 %. Det er imidlertid grunn til å tro at svarprosenten er betydelig høyere. Dette har sammenheng med at det var stor usikkerhet på skolene da de skulle rapportere inn hvor mange lærere som deltar, gitt kriteriene om at det er lærere som deltar i prosjektet i de aktuelle fagene, som skulle besvare skjemaet. På henvendelser til enkelte av skolene er det brakt på det rene at antall lærere på skolen som faktisk deltar i prosjektet, er ned mot halvparten av det antallet skolen i utgangspunktet hadde oppgitt og fått tilsendt hefter til. Forklaringen på dette er trolig at noen skoler har involvert flere lærere i prosjektet enn de som underviser i aktuelle fag og trinn.

Tabell 4.1: Antall lærerrespondenter i undersøkelse 1 fordelt etter trinn og modell.

Barnetrinnet modell A	Barnetrinnet modell B	Barnetrinnet modell C	Ungdoms-trinnet modell D	Videregående skole modell D
175	138	95	95	183

I alt inneholder datafila svar fra 686 lærere ved 72 skoler. Én av skolene er ikke lagt inn i datamaterialet pga for sen tilbakesending. To av skolene har ikke besvart spørreskjemaet, mens data fra to skoler ikke er tatt med på grunn av andre årsaker. Tabell 4.1 viser hvordan disse lærerne fordelte seg på henholdsvis trinn og modell. Tabell 4.2 viser fordelingen av lærerne på fag som inngår i prosjektet. Det viser seg å være en overvekt av lærere som opplyser at de underviser i fagene norsk og matematikk. På barnetrinnet er det ikke uvanlig at enkelte lærere underviser i opptil tre av fagene som inngår i prosjektet. Dette gjelder særlig på småskoletrinnet. Når det gjelder andre fag enn norsk og matematikk i videregående skole, er det stor spredning, og vi har ikke gjort noe forsøk på å analysere forskjeller mellom lærerne i disse fagene.

Tabell 4.2: Antall lærere som deltok i undersøkelse 1 i ulike fag (mange lærere underviser i flere av fagene).

Matematikk	Norsk	Samfunnsfag	Mat og helse	Andre fag
363	355	179	103	89

På spørsmålet om hvilken stilling lærerne har, oppgir 12 prosent *lærer*, 72 prosent *adjunkt* eller *adjunkt med tillegg*, og 16 prosent *lektor* eller *lektor med tillegg*. På spørsmålet om hvor lærerne *hovedsakelig* har sin pedagogiske grunnutdanning fra, svarer 24 prosent *universitetet*. Fordelt etter trinn var denne prosentandelen 4, 32 og 62 for henholdsvis barnetrinn, ungdomstrinn og videregående. Videre var i overkant av 70 prosent av deltakerne på spørreundersøkelsen kvinner, og kvinneandelen på de tre trinnene er henholdsvis 79, 64 og 59 prosent. Gjennomsnittlig antall års erfaring i skolen er 14,2 år med liten forskjell mellom trinnene. 20 prosent av lærerne har mindre enn fem års erfaring, og 10 prosent av dem har 30 års skoleerfaring eller mer.

Lærernes stilling, kjønn, antall års erfaring i skolen eller hvor de hovedsakelig har sin pedagogiske grunnutdanning fra, ser ikke ut til å ha særlig betydning for hvordan de svarer på spørreskjemaet, så lenge vi sammenlikner lærere på samme hovedtrinn i skolen. Vi kommer derfor gjennomgående ikke til å diskutere resultater i forhold til kjønn, undervisningserfaring eller utdanningsbakgrunn i denne rapporten.

4.3 Skolenes utgangspunkt

4.3.1 Skolens eventuelle plan for elevvurdering før oppstart av prosjektet

Vi antok at en viktig faktor for prosessen med utvikling og utprøving av kjennetegn i arbeidet med elevvurdering, ville være hva slags evalueringskultur som preget skolene i utgangspunktet før oppstart av prosjektet. En del skoler har hatt fokus på elevvurdering gjennom lengre tid og har utviklet sitt eget system for vurdering, mens på andre skoler har det vært lite fokus på dette. Et system for elevvurdering kan bl.a. innebære at skolen har egne vurderingsverktøy og en plan for bruken av det. Spørsmål 9 og 10 i spørreskjemaet dreier seg om dette.

Spørsmålet *Hadde skolen fra før en plan for elevvurdering?* (spm 9), ble besvart av 97 prosent av alle som deltok i spørreundersøkelsen. Av de som besvarte spørsmålet, svarte kun 27 prosent *ja*, mens 48 og 25 prosent av lærerne svarte henholdsvis *nei* og *vet ikke*.

Dette bildet støttes av resultatene fra intervjuene. Lærerne ble bedt om beskrive skolens vurderingspraksis før de startet opp med prosjektet, både med hensyn til organisering og

innhold. Svært få av lærerne ga uttrykk for at de kjente til at skolen hadde en plan for arbeidet med elevvurdering, og ved kun én av skolene rapporterte lærerne om et systematisk utviklingsarbeid på bakgrunn av en helhetlig plan for hele skolen på dette området. Alle skolene rapporterte derimot i større eller mindre grad om felles tiltak som er innarbeidet i lærernes evalueringspraksis, men ikke nødvendigvis nedfelt i skolens plandokumenter;

- ”Vi har hatt en del rutiner, uten at det har vært formalisert.” (b)

Dette utsagnet er karakteristisk for uttalelser fra lærere i de fleste av fokusintervjuene.

Skolenes formaliserte rutiner i form av planer rapporteres hovedsakelig å være knyttet opp til offisielt lovverk og retningslinjer. Planer utover dette rapporteres for det meste å være samarbeid på fagseksjon og trinnnivå, eller som resultat av individuelt arbeid knyttet opp mot et spesifikt fag. ”Skolens kultur” rapporteres å være styrende i denne sammenheng;

- ”Styrt av tradisjonen på skolen, ...forskjell mellom trinn med hensyn til klarhet.” (v)

I flere av intervjuene kommer det klart fram et ønske om en felles helhetlig plan for arbeidet med elevvurderingen;

- ”Savner en mer helhetlig plan som går på fagsamarbeid...” (u)

4.3.2 Skolens vurderingspraksis før oppstart av prosjektet

Et viktig premiss for arbeidet på skolene synes å være skolens system for elevvurdering før de gikk inn i prosjektet. Tabell 4.3 gir en oversikt over dette (fra *spm 10*). Da det er såpass få av lærerne som rapporterer at skolen har en slik plan (se 4.3.1), knytter det seg relativt stor usikkerhet til prosenttallene i tabellen, særlig når svarene fordeles på hovedtrinn. Til tross for dette mener vi at lærernes svar på dette spørsmålet gir et visst bilde av rådende vurderingspraksis ute på skolene før prosjektet startet opp, og dermed også et grunnlag for det videre arbeidet med kjennetegn.

Tabell 4.3: Skolenes vurderingspraksis før oppstart av prosjektet. Prosentandel av de 181 lærerne (27 % av alle) som sier at det allerede var en slik plan (*spm 10*).

Planen omfatter:	Barnetrinn	Ungdomstrinn	Videregående	Alle
a) Elevdeltakelse i vurderingsarbeidet	85	85	76	82
b) Rutiner for informasjon til foresatte om skolens system for elevvurdering *	82	53	34	67
c) Faste rutiner for informasjon til foresatte om elevens måloppnåelse *	76	85	34	67
d) En systematisk plan for innhold, hyppighet og form når det gjelder tilbakemelding til elevene	58	55	71	61
e) Egne modeller eller systemer for å gjenkjenne og beskrive elevenes faglige nivå	61	56	58	60
f) Utvikling av kriterier knyttet til kompetansemål *	48	84	60	55
g) Bruk av kriterier i faglige tilbakemeldinger til elevene *	46	90	74	57

* Forskjeller mellom trinnene er ”signifikante” (se 3.5)

Som det framkommer av tabell 4.3, finner man at faste rutiner for informasjon til foresatte om skolens vurderingssystem er nedfelt i planene i ulik grad på skolene som deltar i undersøkelsen. Som forventet og av naturlige grunner, ligger videregående skole lavt her. Forøvrig er forskjellen mellom barnetrinnet og ungdomstrinnet verd å merke seg på bakgrunn av felles forskrift (§3-2) til opplæringsloven. Bildet er derimot et helt annet når det dreier seg om faste rutiner for informasjon til foresatte om elevens måloppnåelse. Her bekrefter hele 85 prosent av lærerne på ungdomstrinnet at de har nedfelt slike rutiner i sitt planverk.

På spørsmålet om skolene hadde utviklet kriterier knyttet til kompetansemål før oppstart av prosjektet, fant vi store forskjeller mellom hovedtrinnene (*spm 10f*). På ungdomstrinnet svarte 84 prosent av lærerne bekreftende på dette spørsmålet, mens tallene for barnetrinnet og videregående skole er henholdsvis 48 og 60 prosent. Dette viser at lærere på ungdomstrinnet som deltar i prosjektet har et annet utgangspunkt for arbeidet med kriterier enn de andre lærerne. Dette er interessant å se opp mot hvor langt hovedtrinnene rapporterer at de har kommet i prosjektet (se 4.5). Vi fant en betydelig forskjell mellom trinnene med hensyn til progresjon i prosjektet. Ungdomstrinnet synes å ha kommet klart lengst med å ta kjennetegnene i praktisk bruk. Når det gjelder elevdeltakelse i vurderingsarbeidet, ser vi nærmere til dette senere i rapporten (4.3.5).

Intervjuene viste at systematikken i vurderingsarbeidet var ulik fra skole til skole. Opplevelsen av å ha en samlet felles plan på skolen varierte ikke bare mellom skoler, men også innad på den enkelte skole. Som tidligere nevnt, var lærerne i fravær av en felles samlet plan for elevvurdering, samkjørte på noen faste ”tradisjonelle” elementer og innarbeidede rutiner på skolen. Faste elementer som ble rapportert fra alle skolene, var utviklingssamtalen (tidligere betegnet som konferansetimen), samtale med eleven før konferanse med foresatte på grunnskolen, og tilbakemelding til eleven før karaktersetting hver termin på videregående. Ellers ble tilbakemeldingen til elevene hovedsakelig gitt som skriftlige kommentarer på skriftlige arbeider eller prøver.

Gjennom intervjuene kom det fram at skolene hadde innført ulike varianter av vurderingsskjemaer, der elevene og lærerne på forskjellige måter vurderte sin egen kompetanse opp mot mål. Målene var delmål av læreplanenes kompetansemål, enkelte også med konkrete kriterier koblet til delmålene. Det ble rapportert om svak, men stadig økende grad av elevinvolvering. Noen av skolene presenterte egenutviklede vurderingsskjemaer, mens andre skoler hadde ”kjøpt” ferdig utviklede verktøy, som mållark, moduler etc.

Når det gjelder utvikling og bruken av kriterier knyttet til kompetansemål og bruk av disse i tilbakemelding til elevene, svarer omlag dobbelt så mange av lærerne på ungdomstrinnet som på barnetrinnet at planen omfatter dette. Her finner vi spesielt stor forskjell mellom barnetrinn og ungdomstrinn. På ungdomstrinnet og i videregående er dette knyttet til karaktergivningen i faget, mens på barnetrinnet er samarbeidet konsentrert om teamets diskusjon rundt mål på elevens arbeidsplaner. Som tidligere påpekt, har det ikke vært vanlig å vektlegge målrelatert vurdering på barnetrinnet. Dette kan illustreres med følgende utsagn fra barnetrinns lærerne:

- *”Kravet til høy måloppnåelse har vært fraværende tidligere.”* (b)

Vurderingen har i også i liten grad dreid seg om veien framover sammenlignet med det eleven har fått til. Dette kommer fram i følgende lærerutsagn:

- *”Framovervurderingen er vanskelig, spesielt på 2. trinn.”* (b)

Når det gjelder lærerens tilbakemelding på elevens kompetanse, rapporterer mange av lærerne derfor om stor usikkerhet knyttet til om eleven mester den kompetansen som målet uttrykker. De gir følgelig uttrykk for å savne et verktøy som kan måle elevens kompetanse;

- *”Setter kryss ved målet, men er usikker på om de kan det.”* (b)

Imidlertid har det å presentere mål for elevenes arbeid blitt rapportert å være godt innarbeidet på alle skolene som ble intervjuet. Skolene ga uttrykk for å ha fokus på målene både ved starten av året og underveis i elevenes ulike planer.

På alle skolene samarbeidet lærerne i elevvurderingsarbeidet. Hyppighet og form på samarbeidet varierte fra skole til skole, men vektleggingen av samarbeidet økte med økende klassetrinn. Det ble gitt uttrykk for ett tett samarbeid på ungdomstrinnet og i videregående skole, mens på småskoletrinnet hadde man ikke samme kultur for dette;

- *"Kunne vært flinkere til å snakke oss lærere imellom."* (b)
- *"Vi mangler en felles basis for hva vi legger i elevvurdering."* (b)

Der samarbeidet er tett, foregår dette hovedsakelig innen samme fag, og sentralt i samarbeidet er diskusjoner om hvilke krav som skal stilles til elevenes kompetanse;

- *"Vi diskuterer hvor lista skal legges."* (v)
- *"Kjører samme prøver og blir enige om hvor lista skal ligge, så vurderinger er ikke tilfeldig."* (u)

4.3.3 "Ikke bare fint og flott, det må være noe mer enn det!" (b)

Hvilke opplevelser har lærerne av endringer når de sammenlikner tidligere vurderingspraksis og arbeidet med kjennetegn? Gjennom de åpne spørsmålene (*spm 13* og *spm 22*) i spørreskjemaet, samt intervjuene, rapporterer lærerne om refleksjoner knyttet til "før" og "etter" oppstart av prosjektet. Lærernes tilbakemeldinger i de åpne spørsmålene har vi gjengitt flere steder i rapporten.

Bevisstheten rundt den faglige tilbakemeldingen til elevene basert på kompetansemålene i faget i motsetning til elevenes arbeidsprosess, er tydelig til stede i alle fokusgruppene;

- *"Fokus fra det sosiale til det faglige."* (b)

I *det sosiale*, definerte de også elevens innsats og aktivitet i timene inn. Dette er spesielt tydelig på barnetrinnet, noe som ikke er overraskende ut fra tradisjonen med bruk av kriterier i karaktersetningen på ungdomstrinnet og i videregående skole. Bevisstheten rundt denne endringen av fokus gjenspeiler seg i tilbakemeldingene fra lærere på alle trinn, men det oppleves spesielt problematisk i de praktiske fagene i videregående;

- *"Innsatsbiten faller bort i vurderingen. Det er vanskelig."* (v)

Kulturen for å vurdere innsats har vært sterk på alle trinn i skolen. Hva betyr denne sterke kulturen for innføring av kjennetegn? Tilbakemeldingene viser tydelig at denne utfordringen er satt på dagsorden i større grad og har blitt et sentralt punkt i den pedagogiske debatten mellom lærere ute på skolene (se også 4.3.4).

Ellers er tilbakemeldingene fra lærerne i både spørreskjema og intervju tydelige på at de har fått økt bevissthet når det gjelder alle sidene ved elevvurdering. Denne bevisstgjøringen blir entydig framhevet som meget positiv og verdifull;

- *"Vi er blitt mer bevisst gjennom prosjektet."* (b)

Til tross for at mange av skolene har erfaring med bruk av målrelatert vurdering, oppleves arbeidet med innføring og bruk av kriterier i form av kjennetegn for mange som en god og naturlig videreutvikling av elevvurderingen;

- *"..viderefører målbasert vurdering. Nytt er innføring av kriterier."* (b)

Arbeidet har satt i gang prosesser rundt på skolene, prosesser som får ringvirkninger utover de fagene som omfattes av prosjektet;

- *"Tenkningen smitter over på andre fag."* (u)

Bevisstgjøringen gjelder ikke bare vurderingsprosessen, men lærerne rapporterer i like stor grad om økt bevisstgjøring rundt den læringsfremmende effekten tilbakemeldingene til elevene skal ha;

- *"Måten vi tenker på og måten vi snakker med elevene på."* (b).

Denne endringen av fokus for tilbakemelding til elevene gjelder ikke bare skoler som tidligere har manglet en kultur for systematisk elevvurdering knyttet opp mot læreplanens kompetansemål. Skoler som fra før har en solid plattform mht målrelatert vurdering, rapporterer at den økte bevisstheten nettopp ligger i begrunnelsen overfor elevene;

- ”Ikke så ulik tidligere praksis, men nå må det begrunnes.” (b)

Den økte bevisstheten rundt elevvurdering og de positive signalene fra lærerne om effekten av refleksjonene og endret praksis som dette har medført, gjenspeiles i lærernes kommentarer i spørreskjemaet (*spm 22*). Arbeidet med kjennetegn oppleves som nyttig og lærerikt for en stor majoritet av lærerne som deltar i prosjektet (jfr 4.7.5).

4.3.4 Lærernes syn på vurdering

I forbindelse med spørreundersøkelsen ønsket vi (ved hjelp av *spm 8*) å få innblikk i lærernes prinsipielle syn på vurdering, uavhengig av den aktuelle situasjonen. Flere av utsagnene som lærerne skulle ta stilling til, dreide seg om forhold ved elevvurdering som står sentralt i prosjektet og også ligger til grunn for tenkningen bak kriteriebasert vurdering. Tabell 4.4 viser hvordan lærerne svarte på noen av disse utsagnene.

Tabell 4.4: Lærernes svar på utsagn om hva som kjennetegner god elevvurdering (fordeling i prosent for hvert utsagn). Svar på noen av utsagnene i *spm 8*.

	Enig	Nøytral	Uenig
a) Sammenlikning mellom elever bør ligge til grunn for vurdering av den enkelte elev.	14	24	62
b) Elevenes innsats og aktivitet i klasserommet bør telle med som en viktig del av vurderingen av eleven.	66	17	17
d) Et godt vurderingssystem skiller klart mellom elevenes læringsprosess og deres faglige kompetanse.	54	37	9
f) Det er viktig å få tydelig fram det eleven ikke mestrer i den faglige tilbakemeldingen.	59	22	19
h) Det er bare elever som har kompetanse på høyt nivå i et fag, som har mestring i faget.	2	10	89
k) Høy måloppnåelse i et fag kjennetegnes ved at eleven viser stor interesse for faget.	16	29	55
m) Elevvurdering bør inngå som en viktig del av undervisningen i faget.	70	25	5
n) Å hente informasjon fra en tekst representerer lav måloppnåelse, mens høy måloppnåelse alltid forutsetter drøfting.	34	37	30

Det viste seg at 6 av 10 lærere er uenige i den første påstanden (*spm 8a*) om at sammenlikning mellom elever bør ligge til grunn for vurderingen av den enkelte elev. Vekt på sammenlikning mellom elever representerer et normbasert vurderingssystem, noe som ikke skal være regelen i skolen (jfr. 2.1). Karakteren eller vurderingen av en elevs prestasjoner skal nettopp ikke foretas på bakgrunn av hva andre elever presterer. Når det gjelder vurdering med karakter, skal den i henhold til regelverket være målrelatert. Også vurdering uten karakter skal uttrykke grad av måloppnåelse. Når det gjelder lærernes svar på dette, finner vi liten forskjell mellom barnetrinn, ungdomstrinn og videregående skole.

I underkant av 70 prosent av lærerne mener at elevens innsats og aktivitet bør inngå som en viktig del i vurderingen av eleven (*spm 8b*), mens litt over halvparten var uenige i påstanden om at høy måloppnåelse i et fag kjennetegnes ved at eleven viser stor interesse for faget (*spm 8k*). Det er et klart krav fra departementets side om at faglige prestasjoner på den ene siden og

aktivitet og innsats på den andre, holdes fra hverandre i forbindelse med vurdering (St.meld. nr. 30 2003-2004). Det er med andre ord utelukkende kompetansen i fagene og elevenes måloppnåelse som skal ligge til grunn for den faglige vurderingen av eleven. På denne måten blir det også klarere for elevene hvilke kriterier som er utgangspunktet for den faglige vurderingen. På dette området finner vi tydelig forskjell mellom barnetrinns lærere på den ene siden og lærer på ungdomstrinnet og i videregående skole på den andre. Dette samsvarer med den historiske bakgrunnen som ble presentert i kapittel 1. Lærerne på barnetrinnet vektlegger i mye større grad innsats i elevvurderingen. Hele 75 prosent av lærerne på barnetrinnet var enige i dette, mens 48 og 55 prosent av lærerne på henholdsvis ungdomstrinnet og videregående skole var av samme oppfatning. Et interessant poeng i denne sammenhengen er imidlertid at lærere på barnetrinnet i forbindelse med intervjuene, gir uttrykk for at de gjennom arbeid med prosjektet er i ferd med å endre holdning og praksis i forhold til dette. De gir uttrykk for at de er i ferd med å få en mer bevisst holdning til hva som skal inngå i elevvurderingen (jfr. 4.3.3). Også på videregående skole rapporterer lærerne i forbindelse med intervjuene om utfordringene ved at innsats ikke skal inngå i elevvurderingen. Spesielt på de yrkesfaglige studieretningene gir lærerne uttrykk for at dette er et problem. Lærernes svar på disse utsagnene må ses i sammenheng med at mange av dem har arbeidet innenfor en læreplantradisjon hvor innsats og interesse skal "premieres";

- *"Innsatsbiten faller bort i vurderingen. Det er vanskelig."* (v).
- *"Dette bryter tradisjonen fra tidligere praksis."* (v)

Kun halvparten av lærerne er enige i at en i elevvurderingen bør skille mellom elevenes læringsprosess og deres faglige kompetanse. Det at hele 37 % av lærerne har valgt ikke å ta stilling til dette utsagnet, kan være en indikasjon på stor usikkerhet blant lærerne på dette punktet eller at de har problemer med å forstå hva som menes med utsagnet. Når det gjelder dette utsagnet, finner vi ingen forskjeller mellom barnetrinnet, ungdomstrinnet og videregående skolen. I forbindelse med intervjuene ga lærerne uttrykk for at det er vanskelig å skille mellom elevens kompetanse i faget og arbeidet de legger ned for å nå læringsmålet;

- *"Det er vanskelig å skille fag og prosess."* (b)

I alt seks av ti lærere mener det er viktig å få fram det eleven ikke mestrer i den faglige tilbakemeldingen (*spm 8f*). Dette er først og fremst relevant i situasjoner hvor det er viktig å formidle til elevene hva de bør arbeide mer med. På dette området ble det ikke registrert forskjeller mellom de tre hovedtrinnene, og flere av lærerne ga i intervjuet uttrykk for at det i den faglige tilbakemeldingen også burde inngå hva elevene ikke mestrer.

- *"Det er vanskelig å ikke ha fokus på og si det de ikke mestrer."* (v)

Når det gjelder utformingen av kjennetegnene, er det imidlertid viktig å være klar over at de i prinsippet skal fokusere på det som elevene mestrer (Utdanningsdirektoratet 2007). Kjennetegn på måloppnåelse i fag skal med andre ord beskrive hva elevene faktisk kan og ikke det de ikke kan.

Bortimot ni av ti lærere var imidlertid uenige i påstanden om at det bare er elever som har kompetanse på høyt nivå, som har mestring i faget (*spm 8h*). Det vil med andre ord si at på dette området har de fleste lærerne en oppfatning som er i tråd med tenkningen som prosjektet bygger på.

Når det gjelder utsagnet om at elevvurderingen bør inngå som en viktig del av undervisningen (*spm 8m*), viser det seg å være større tilslutning blant lærere på barnetrinnet enn på ungdomstrinnet og aller minst blant lærere på videregående skole. Ser man de tre trinnene under ett, er nesten sju av ti lærere enig i dette utsagnet.

Ett av utsagnene var rettet mot taksonomi (*spm 8n*). Utsagnet som ble benyttet, er i tråd med Blooms taksonomitenkning; det å drøfte vil alltid være på et mer avansert nivå enn det å hente ut informasjon (jfr. 2.2). Det går fram av tabellen at lærernes svar fordeler seg relativt jevnt på de tre svaralternativene, noe som kan være en indikasjon på lærernes usikkerhet når det gjelder å formulere beskrivelser av kjennetegn på måloppnåelse på ulike nivå. Dette kom også fram i intervjuene. Det var i første rekke lærerne på ungdomstrinnet og i videregående som var opptatt av dette. De hadde hatt store diskusjoner i faggruppene om hvilke betegnelser som best beskriver hva som kjennetegner ulike kompetansenivåer.

- *"Det er vanskelig å bli enige om begrepene. Er det å kunne vurdere mer avansert enn det å kunne beskrive?"* (v)
- *"Formuleringene er vanskelige."* (u)
- *"Vi må finne en mal som det ikke var retningslinjer for (med mal mente lærerne uttrykk som beskriver det elevene kan og som er satt i system, dvs. de må utvikle sin egen taksonomi). Hadde det vært mulig å få noen retningslinjer her? Hvorfor må vi finne opp kruttet på nytt?"* (v)

Dette illustrerer tydelig de utfordringene som lærerne står overfor i arbeidet med å utvikle kjennetegn på ulike kompetansenivåer.

Disse resultatene kan være en indikasjon på at det blant lærere er en viss motstand mot de prinsipper som ligger til grunn for kriteriebasert vurdering og den tenkningen som prosjektet bygger på. Det at såpass mange som 1/4 av lærerne (i noen tilfeller flere) har krysset av på "nøytral" på flere av utsagnene, kan være et signal om usikkerhet eller manglende bevissthet hos lærerne når det gjelder disse spørsmålene, framfor direkte uenighet. I intervjuene var lærerne entydige i ønsket om økt vurderingskompetanse (jfr. 4.6.2).

I forbindelse med de statistiske analysene ble det registrert signifikant sammenheng mellom et syn på elevvurdering som er i tråd med tenkningen som prosjektet bygger på, og lærernes tilfredshet med å delta i prosjektet. Eller sagt på en annen måte; de lærerne som er fornøyde med å delta i prosjektet, har også en holdning til elevvurdering som gjenspeiler tenkningen som ligger til grunn for kriteriebasert vurdering. Dette er også prosjektets utgangspunkt.

4.3.5 Elevdeltakelse i vurderingen

Både i Læreplanverket og i Forskrift til opplæringsloven påpekes det at elevene skal trekkes aktivt med i den faglige vurderingen. Av tabell 4.5 går det fram hvor stor andel av lærerne som var "Enig" eller "Svært enig" på noen sentrale utsagn om elevdeltakelse (*spm 19*).

Det er naturlig nok stor enighet om at elevene spiller en hovedrolle i elevsamtalen. Om lag halvparten av lærerne er enig i at elevene kan vurdere sin egen kompetanse, at elevene selv har forslag til hvordan de skal arbeide for å nå sine mål, og at de jevnlig har samtaler med elevene om deres måloppnåelse.

Tabell 4.5: Prosent av lærerne som er ”Enig” eller ”Svært enig” i noen sentrale utsagn om elevdeltakelse i vurderingen (spm 19).

	Barne- trinn	Ungdoms- trinn	Videre- gående	Alle
a) Elevene spiller en hovedrolle i elevsamtalen.	97	99	97	97
b) De fleste elevene er i stand til å vurdere sin egen kompetanse knytte til kompetansemålene i faget.*	50	49	65	53
c) De fleste elevene har forslag til hva de skal jobbe med for å nå læringsmålene sine.	50	46	53	50
e) Jeg involverer elevene i alle sider av vurderingen i faget.*	30	27	46	34
f) Jeg har jevnlig samtaler med den enkelte elev om måloppnåelsen i faget.*	53	53	75	59
h) De fleste elevene deltar aktivt med å sette opp egne mål i faget.*	20	4	13	16
i) Jeg forteller elevene hva de skal gjøre for å bli bedre i faget.	97	100	97	97
j) Jeg forteller elevene hva de mestrer i faget.	98	97	92	96

* Forskjeller mellom trinnene er ”signifikante” (se 3.5)

For flere av utsagnene i tabell 4.5 er det nokså små forskjeller mellom hovedtrinnene. I og for seg er dette litt overraskende, idet vi ville forvente at elevene ble betydelig mer involvert i evalueringen på høyere trinn i skolen. Derimot finner vi en stor forskjell mellom trinnene når det gjelder om lærerne har jevnlig samtaler med den enkelte elev om måloppnåelse i faget. Her rapporterer $\frac{3}{4}$ av lærerne på videregående skole at de er enige i utsagnet, mens omtrent halvparten av lærerne på de to andre hovedtrinnene er enige. Med tanke på den aktuelle debatten om læreplanmålene er det interessant å se at såpass få elever deltar aktivt med å sette opp egne mål i fagene (spm 19h). Det er først og fremst lærerne på ungdomstrinnet som rapporterer om en praksis der dette foregår i liten grad. Når det gjelder de to siste utsagnene (spm 19i og 19j), representerer de kjernen i det som forstås med framovermeldinger (jfr 2.3). Her rapporterer lærerne på alle trinn at de forteller elevene hva de mestrer og hva de skal gjøre for å bli bedre i faget. Lærerne ble imidlertid ikke spurt om hyppigheten av slike tilbakemeldinger.

4.3.6 Lærernes forståelse av oppdraget

I spørreskjemaet var det ingen spørsmål som var direkte rettet mot lærernes forståelse av oppdraget. På bakgrunn av sluttrapporten fra piloteringen (Engh 2007) og gjennom signaler som ble fanget opp i spørreskjemaet, anså vi dette som et viktig moment som burde belyses nærmere i forbindelse med intervjuene. Lærernes usikkerhet når det gjelder prosjektet som helhet og deres rolle i prosjektet, blir beskrevet senere i rapporten (jfr. 4.7.6 og 4.7.7). I denne sammenheng ønsker vi å sette fokus på lærernes forståelse av sentrale begreper i prosjektet. Her tenker vi først og fremst på hva lærerne legger i betegnelsen *kjennetegn*, og hvilke konsekvenser denne forståelse har for deres videre arbeid med prosjektet og for prosjektet som helhet.

Ifølge Utdanningsdirektoratets veiledningshefte er det ingen fasit på hvordan felles kjennetegn i et fag skal utformes, men det fremheves som et viktig prinsipp at kjennetegnene

skal dekke de samlede kompetansemålene i faget. Kjennetegnene skal være beskrivelser som knytter sammen flere kompetansemål av gangen. Det står videre at felles kjennetegn ikke er ment å skulle kunne brukes direkte inn mot vurdering av elevenes arbeid, men gi felles referanser på hva som kjennetegner ulik måloppnåelse i fag.

Den forståelsen av kjennetegn som vi har registrert på skolene, er i varierende grad i tråd med Utdanningsdirektoratets definisjon og forståelse. I forbindelse med intervjuene kom det fram at lærerne ser på de ferdigstilte kjennetegnene som ulne, uklare og lite hensiktsmessige med tanke på å skulle anvende dem i vurderingsarbeidet. Lærerne gir uttrykk for at de ønsker seg kjennetegn som er så klare at de kan brukes i konkrete situasjoner, som for eksempel i kommunikasjonen med elevene. I flere av fokusgruppene kom det fram at kjennetegnene som blir utviklet, i realiteten er kriterier på delmål av kompetansemålene i LK06.

Det er viktig å presisere at dette funnet bygger på intervju med et relativt begrenset antall fokusgrupper. Vi har derfor ikke belegg for å hevde at en slik forståelse også skulle gjelde ved de skolene som ikke var med i intervjurunden. En må imidlertid ta høyde for at dette kan være tilfellet.

4.4 Hvordan har skolene organisert arbeidet med prosjektet?

4.4.1 Skolenes interne organisering

Lærerne ble spurt om hvordan skolen organiserer arbeidet med prosjektet (*spm 12*). Tabell 4.6 viser hvordan ulike lærergrupper beskriver hvordan dette arbeidet er organisert på skolen. Som det går fram av tabellen, har de fleste skolene valgt å organisere arbeidet i faggrupper.

Tabell 4.6: Prosentandel av lærerne som svarer "Ja" for hvert trinn og modell (*spm 12*).

	Barnetrinn				Ungdoms-	Videre-
	A	B	C	Alle	trinn	gående
					D	D
a) Organisert i faggrupper *	89	43	80	71	95	78
b) Organisert på det enkelte trinn *	41	63	44	49	28	42
c) Organisert med elevdeltakelse *	8	14	9	10	7	8
d) Organisert som del av lokalt læreplanarbeid *	35	29	60	39	51	33

* Forskjeller mellom trinnene og modellene er "signifikante" (se 3.5)

Denne tendensen ble bekreftet i intervjuene, og de fleste lærerne så på dette som en hensiktsmessig organisering, da det ga dem anledning til å jobbe i forhold til eget (eller egne) fag;

- "Arbeidet er organisert på stortrinn på grunn av arbeidsmengden, for eksempel matte 1. - 4.trinn og 5. - 7.trinn." (b)

På noen skoler var faggruppene også organisert i forhold til det enkelte trinnet. Flere av lærerne kommenterte at som konsekvens av denne organiseringen, kan forståelsen av kjennetegn og utprøvingen bli ulik mellom fagene innad på skolen. På en av skolene rapporterte lærerne at fagseksjonene hadde laget hver sin plan, noe som resulterte i;

- "... ulike planer for de ulike fagseksjonene" (v)

I et åpent spørsmål i spørreskjemaet ble lærerne bedt om å beskrive hvordan de opplever å delta i prosjektet (*spm 22*). Det viser seg at en stor andel av kommentarene går på organiseringen av arbeidet. Mange av lærerne opplever organiseringen av prosjektet som

frustrerende (jfr. 4.7.5 hvor det gjøres nærmere rede for dette). Det er likevel viktig å merke seg at ingen her kommenterte organisering i faggrupper, hvilket tyder på at det ikke er fagorganiseringen i seg selv som oppleves som frustrerende. Lærernes kommentarer i intervjuene bekrefter hovedsakelig også en slik tendens. Ser vi dette i sammenheng med uttalelser i intervjuene, er det grunn til å anta at det hovedsakelig er organiseringen på skolenivå lærerne sikter til.

Et trekk ved organiseringen av prosjektet lokalt er at elevene i så liten grad trekkes inn. Dette gjelder særlig på de høyeste trinnene;

- ”*Elevene kan gradvis bli med på det og reflektere over sine egne prestasjoner.*” (v)

Her er det viktig å påpeke at skolene hovedsakelig er i oppstartfasen, men graden av elevinvolvering er et område som det blir interessant å følge opp videre. Det er naturlig å se dette resultatet opp mot den vurderingskulturen som vi har hatt i Norge. Fravær av elevdeltakelse har vært betydelig, noe som bekreftes gjennom intervjuene.

I våre analyser for å avdekke faktorer som er viktige for *framdriften* av prosjektet (*spm 11c*), noe vi har definert som et suksesskriterium for prosjektet, fant vi *elevdeltakelse* sentralt (jfr 4.8). Faktoren *at arbeidet er organisert med elevdeltakelse* (*spm 12c*), ser ut til å være viktig i forhold til om lærerne har tatt kjennetegnene i bruk (*spm 11c*). Sett i lys av dette er lærernes rapportering av elevdeltakelsen meget interessant.

Det kan synes å være noe uavklart når det gjelder forholdet mellom utprøvingen og det lokale læreplanarbeidet ved skolene, da ca en tredel svarer at dette spørsmålet foreløpig er uklart. Derimot er det interessant at 40 prosent svarer at *arbeidet er organisert som en del av det lokale læreplanarbeid* (*spm 12d*). Dette er en av faktorene som i våre data henger sammen med suksesskriteriet *motivasjon*, slik det kommer til uttrykk i hvor fornøyd lærerne er med at skolen deltar i prosjektet (*spm 21d*) (jfr. 4.8). Vi finner en betydelig forskjell både mellom trinn og modeller på dette utsagnet, se tabell 4.6.

Gjennom intervjuene rapporterte et betydelig antall av lærere på barnetrinnet at de opplevde et økende fagfokus i forbindelse med arbeidet med kjennetegn. Spesielt blant lærere på småskoletrinnet var dette en hyppig kommentar;

- ”*Må bli mer fagsentrert for å lykkes med dette.*” (b)

Ser vi dette i lys av tradisjonen for samarbeid på trinnivå på barnetrinnet, og spesielt på småskoletrinnet, kan dette representere en interessant dreining av fokus på de lavere trinnene. Til tross for lærerne jevnt over en ga uttrykk for en positiv holdning til dreiningen over på fag, kom det tydelig fram i intervjuene en stor skepsis knyttet til mengde arbeid som dette vil kunne medføre;

- ”*Kan bli veldig mye ... kommer til å drukne på grunn av mange fag.*” (b)

4.4.2 Ledelse av prosjektet på skolen

Det kommer fram av forrige avsnitt at skolene har organisert prosjektet på ulike måter. Mange har dannet en prosjektgruppe ledet av en prosjektleder med et spesielt mandat i forhold til initiering og gjennomføring av prosjektet. I intervjuene uttrykkes sterkt behovet for å ha en prosjektledelse som informerer, koordinerer, strukturerer og setter klare rammer og mål for prosjektet;

- ”*Vi hadde trengt noen konkrete mål å jobbe mot.*” (b)

I alle intervjuene ble det framholdt nødvendigheten av å ha en plan for arbeidet og med prosjektet på skolen og en oppfølging med milepæler og evaluering underveis;

- *”Vi har ingen plan på skolen for framdriften av prosjektet.”* (u)

I flere av fokusgruppene trekkes det fram behovet for en ledelse som sørger for kontinuitet i arbeidet med prosjektet. Arbeidet med utvikling og utprøving av kjennetegn stykkes opp, med de konsekvensene at tidsbruket øker, effektiviteten blir mindre og kvaliteten på arbeidet oppleves som dårligere;

- *”Kommet bare som små drypp, som splitter opp hverdagen, får ikke sammenhengen og får ikke gjort det ordentlig.”* (v)
- *”Hvorfor ikke smi mens jernet er varmt?”* (b)

Der man ikke har en ledelse som ivaretar disse funksjonene, kommer etterspørselen etter ekstern hjelp for å fylle disse funksjonene;

- *”Vi kunne trenge hjelp med hensyn til framdrift og organisering.”* (b)

På skoler der prosjektlederfunksjonen er delegert til en lærer, rapporteres det om usikkerhet knyttet til rolleoppfatning og lærerens funksjon inn i personalet;

- *”Lærere i prosjektgruppa er usikre på sin rolle – må også motivere frustrerte lærere. Mye faller på én lærer.”* (b)

På en av skolene som ble intervjuet, ble prosjektet ledet av rektor. Lærerne rapporterte om planer både på region og skolenivå. Dette var planer med klare mål og milepæler. På denne skolen opplever lærerne prosjektet som en videreføring av utviklingsarbeidet med målbasert vurdering, der innføring av kriterier i form av kjennetegn kommer inn som et nytt element. Prosjektet er en del av skolens lokale læreplanarbeid. I motsetning til de andre skolene beskriver lærerne her en aktiv satsing på vurdering også på kommunalt og regionalt nivå. De rapporterer om et pedagogisk senter som driver veiledning og legger til rette for kompetanseheving innenfor satsingsområdene for alle skolene i kommunen. Lærerne opplever samsvar mellom kommunal og regional satsing og skoleleders prioriteringer og føringer.

4.4.3 Samarbeid mellom skoler

På spørsmål om skolen samarbeider med andre skoler i forbindelse med prosjektet (spm 14), svarte lærerne på følgende måte (tabell 4.7):

Tabell 4.7: Prosentandel av lærerne som sier at skolen samarbeider med andre skoler i forbindelse med prosjektet *

Barnetrinn				Ungdomstrinn	Videregående
A	B	C	Alle	D	D
54	86	47	63	55	43

* Forskjeller mellom trinnene er ”signifikante” (se 3.5)

Ser man alle skolene under ett, rapporterer litt over halvparten av lærerne i spørreundersøkelsen at skolen har innledet samarbeid med andre skoler. Tilbakemeldingene viser noe mer samarbeid på barnetrinnet (og spesielt mellom skolene med modell B) enn på ungdomstrinnet og i videregående skole.

De som rapporterer om samarbeid med andre skoler, ble også spurt om hvordan dette samarbeidet var organisert. Her kom ”samarbeidsmøter i skolens regi” og ”samarbeidsmøter i skoleeiers regi” noenlunde likt ut. Diskusjoner og erfaringsutvekslinger på nettet forekom i

noe mindre grad. Bare 30 prosent av skoler som samarbeider, rapporterer om at dette foregår på nettet.

I spørreskjemaet ble lærerne bedt om å nevne to viktige faktorer for at skolen skal lykkes i arbeidet med kjennetegn (*spm 13*). Det viser seg at 20 prosent av de som svarer, nevner samarbeid innad og mellom skoler som viktige faktorer i denne sammenhengen. Dette bekreftes også i intervjuene. Samarbeid mellom skoler ble framhevet som en viktig faktor i utviklingen og det videre arbeidet med kjennetegnene. Motivasjon og arbeidsfordeling ble spesielt framhevet i denne sammenhengen. Der det foregår et samarbeid mellom skoler, ble dette løftet fram som en viktig faktor. På samme måte rapporteres fravær av samarbeid som et savn på de skolene som ikke har etablert et slikt type nettverk;

- ”Det hadde vært nyttig med innspill fra andre, fra andre skoler og fra andre grupper.” (v)

Til tross for positiv holdning til et slikt nettverk, kan samarbeidet for enkelte skoler også representere en utfordring. Det kommer tydelig fram på en av skolene som ble intervjuet;

- ”De jobber helt annerledes enn oss (de andre skolene) mange skoler gir mange individuelle tolkninger.” (b)

Totalt sett finner vi en klar tendens blant lærerne til å ønske et samarbeid med andre skoler der utveksling av erfaringer med utvikling og bruk av kjennetegn står sentralt. Dette er viktig å merke seg på alle nivåer innenfor skolesektoren i den videre planleggingen av arbeidet med kjennetegn.

4.5 Hvor langt har skolene kommet i prosjektet?

4.5.1 Hvor skolene er i prosjektet ved undersøkelsestidspunktet

Vi presenterer i tabell 4.8 informasjon om hvor langt skolene var kommet i prosjektet ved tidspunktet for undersøkelsen (februar/mars 2008). Siden pilotkolene startet prosjektet tidligere, er disse skolene holdt utenfor i denne sammenstillingen. Det går fram av tabellen at de fleste skolene har lagt planer for prosjektet, og at de har startet arbeidet med å utvikle kjennetegn. I intervjuene kom det fram at tre av de fire skolene opplevde å ha en plan for det videre arbeidet med prosjektet og bruken av kjennetegn. Når det gjaldt innholdet i planen, varierte det mellom de tre skolene; Fra én av skolene der planen besto av møtedatoer, til én annen skole der man hadde en plan for skolen og hele regionen med utprøving av to kjennetegn for alle fag med oppfølgende tidfestet evaluering.

Når det gjelder spørsmålet om hvorvidt de har prøvd ut kjennetegnene i vurderingsarbeidet (*spm 11c*), og om de har tatt dem i bruk i kommunikasjonen med elever og foresatte (*spm 11e*), er det langt færre som svarer bekræftende. Dette betyr at prosjektet så langt synes å ha hatt nokså begrenset betydning for hvordan elevvurderingen faktisk foregikk på prosjektskolene ved dette tidspunktet. Med tanke på implementering av de prinsippene dette dreier seg om, er det svært viktig å ha en forståelse av at det dreier seg om en endring i vurderingskultur som nødvendigvis må ta tid, særlig sett i lys av at det skal innføres i en ellers travel hverdag (jfr. kapittel 1 om behovet for brudd med tidligere vurderingspraksis).

Tabell 4.8: Hvor langt skolene (utenom pilotskolene) er kommet i prosjektet. Prosent av lærerne som svarer bekreftende på ulike utsagn (spm 11).

	Barnetrinn				Ungdoms-	Videre-
	A	B	C	Alle	trinn	gående
a) Skolen har lagt planer for hvordan arbeidet organiseres. *M	88	78	84	84	D	D
b) Skolen har startet arbeidet med å utvikle kjennetegn. *M	89	66	85	81	87	80
c) Jeg (lærer) har tatt i bruk kjennetegnene i vurderingsarbeidet. *	17	36	29	25	45	39
d) Jeg (lærer) har tatt i bruk kjennetegnene i kommunikasjonen om vurdering med elevene. *	15	25	23	20	40	35
e) Jeg (lærer) har tatt i bruk kjennetegnene i kommunikasjonen om vurdering sammen med elevene og foresatte. *T	7	8	9	8	21	2

* Forskjeller mellom både trinn og modeller er "signifikante" (se 3.5).

*M "Signifikante" forskjeller bare mellom modeller, ikke mellom trinn

*T "Signifikante" forskjeller bare mellom trinn, ikke mellom modeller

Som det framgår av tabell 4.8, er det er en betydelig forskjell mellom trinnene og mellom modellene hva gjelder progresjon. Ungdomstrinnet synes å ha kommet klart lengst med å ta kjennetegnene i praktisk bruk, noe som er litt overraskende, sett i lys av at disse skolene også har måttet utvikle kjennetegnene selv. Når det derimot gjelder barnetrinnet, er det naturlig at de som må lage kjennetegn selv (modell A), er kommet kortest. I intervjuene fant vi en klar forskjell mellom trinnene på barnetrinnet. Lærerne på mellomtrinnet rapporterte i intervjuene å ha en klarere forståelse av kjennetegn enn lærerne på småskoletrinnet. Blant lærere på 1. til 4. klassetrinn fant vi større grad av usikkerhet med hensyn til utvikling og bruk;

- "Vi er jo litt usikre på det vi har laget, om det er bra nok. Skulle gjerne hatt noen som så på det før vi prøvde det ut". (b)

Også på videregående skole ga flere av lærerne uttrykk for at de ventet med å ta i bruk kjennetegn med elevene i undervisningen. Begrunnelsen der gikk imidlertid ikke på usikkerhet knytte til utviklingen, men derimot behovet for mer tid til utvikling av kjennetegn;

- "Vi bruker denne våren til å utvikle kjennetegn og så prøver vi dem ut til høsten." (v)

4.5.2 Progresjonen i pilotskolene i forhold til resten av skolene

Når det gjelder pilotskolene, har også disse lærerne svart på spørsmålene om progresjon. I tabell 4.9 har vi sammenliknet prosenttallene for pilotskolene med de andre skolene, og disse tallene er tankevekkende. For ungdomstrinnets vedkommende framgår det at som ventet er pilotskolene kommet betydelig lenger enn de andre skolene. Kjennetegnene er i stor grad tatt i bruk i den praktiske vurderingen. Men for barnetrinnet framstår pilotskolene med en svakere framdrift. De to første spørsmålene kan kanskje ha vært noe forvirrende for pilotskolene, idet de har holdt på med prosjektet en god stund, og lærerne kan tenke at de er forbi "planleggingsstadiet". Mens rundt halvparten av disse lærerne på barnetrinnet rapporterer at kjennetegnene er tatt i bruk i vurderingsarbeidet, sier nesten alle på ungdomstrinnet at dette har skjedd. Det er vanskelig å se at dette henger sammen med at kjennetegnene måtte utvikles først, da bare én av pilotskolene på barnetrinnet har modell A. Det blir interessant å følge opp

dette spørsmålet i neste datainnsamling for å se hvordan forskjellen mellom barnetrinnet og ungdomstrinnet endrer seg på sikt.

Tabell 4.9: Hvor langt pilotskolene er kommet i prosjektet sammenliknet med de andre skolene. Prosent av lærerne som svarer bekreftende på ulike utsagn (spm 11).

	Pilotskoler		Ikke pilotskoler	
	Barne-trinn	Ungd.-trinn	Barne-trinn	Ungd.-trinn
a) Skolen har lagt planer for hvordan arbeidet organiseres.	69	81	84	83
b) Skolen har startet arbeidet med å utvikle kjennetegn.	63	100	81	87
c) Jeg (lærer) har tatt i bruk kjennetegnene i vurderingsarbeidet.	53	94	25	45
d) Jeg (lærer) har tatt i bruk kjennetegnene i kommunikasjonen om vurdering med elevene.	39	81	20	40
e) Jeg (lærer) har tatt i bruk kjennetegnene i kommunikasjonen om vurdering sammen med elevene og foresatte.	19	50	8	21

4.5.3 Sammenhengen mellom ”før og nå”

Det kan virke som om de lærerne som klart ser ”linken” mellom tidligere vurderingspraksis og arbeid med kjennetegn, er mer positive og har kommet lenger i prosjektet enn lærere som ikke trekker disse linjene. Evnen til å koble ”nytt og gammelt”, se sammenhengen og overføre kompetanse på målrelatert vurdering, ser ut til å være et suksesskriterium i forhold til framdrift;

- ”Vi har fått noe som gjør det vi hadde bedre.” (b)

Denne refleksjonen fant vi på skoler som fra før har en vurderingskultur som var kjennetegnet ved systematikk og felles forståelse. Ved skoler der vi ikke fant denne holdningen, rapporterte lærerne om en vurderingskultur som var preget av en større grad av individualisering.

Lærere som tok utgangspunkt i eksisterende lokale planer og utviklet kjennetegnene i forhold til disse, rapporterte at dette var en god og hensiktsmessig måte å utvikle og implementere kjennetegn i elevvurderingen. Dette er i tråd med Utdanningsdirektoratets retningslinjer for prosjektet, der det står at deltakerskolene kan fortsette å bruke egne lokale planer og framgangsmåter i vurderingsarbeidet. Som forventet var det også blant disse lærerne vi fant forslag til hvordan bruken av kjennetegn inngår som en naturlig del av undervisningen. Vi fant store forskjeller mellom hovedtrinnene på dette området. Mens lærere på videregående skole, ungdomstrinnet og til dels mellomtrinnet i hovedsak hadde klare tanker om koblingen mellom bruk av kjennetegn og den daglige undervisningen, ”slet” flere av småskoletrinnets lærere med å finne ”rom og plass” for dette.

4.6 Skolens kompetanseutvikling i elevvurdering

4.6.1 Viktige faktorer for å utvikle lærernes kompetanse i elevvurdering

Et viktig mål for Utdanningsdirektoratets prosjekt *Bedre vurderingspraksis* er å utvikle lærernes kompetanse i elevvurdering, slik at vurderingen blir rettferdig, mer faglig relevant og læringsfremmende. I spørreskjemaet ble lærerne i et åpent spørsmål bedt om å nevne to viktige faktorer for at skolen skal lykkes med dette (spm 13). Spørsmålet ble besvart av 602 lærere og svarene ble kodet i forhold til syv kategorier av faktorer.

Tabell 4.10: Faktorer som er nevnt som viktige for å lykkes i å utvikle læreres kompetanse, gitt i prosent av de som svarte på dette spørsmålet. (Summen blir over 100 % fordi hver lærer kunne nevne inntil to faktorer.)

Faktor	Prosent av respondentene
Alt som har med tidsaspektet å gjøre	60
Felles for hele skolen/felles verktøy	33
Behov for kompetanseutvikling, informasjon og ekstern hjelp	20
Forankring i ledelsen	11
Samarbeid internt på skolen og mellom skoler	19
Annet	47

Tid

Som det framgår av tabellen, opplever lærerne *tid* som den viktigste faktoren for å lykkes med å utvikle lærernes vurderingskompetanse. Mange gir uttrykk for at arbeidet med kjennetegn tar uforholdmessig mye tid og kommer på toppen av alt annet;

- ”Det er ikke satt av nok tid til arbeidet. Det kommer på toppen av alt som skal gjøres.” (b)

For mange av lærerne oppleves utviklingen og utprøvingen av kjennetegn som et merarbeid, og ikke som en del av undervisningen;

- ”Meir papirarbeid. Mindre tid til å planleggje undervisningen.” (b)

Denne uttalelsen er også interessant sett i lys av prinsippet *vurdering for læring*, og synet på vurdering som en viktig del av elevenes læringsprosess.

Til tross for at svært mange rapporterte bruken av tid som en kritisk faktor i prosjektet, setter mange av lærerne denne sentrale betenkningen inn i en konstruktiv sammenheng;

- ”Svært positivt med et slikt prosjekt. Det er svært nyttig og viktig med en felles vurderingspraksis! Like fullt bør det settes av mye tid til et slikt prosjekt. Når det blir en ekstrem belastning i det daglige arbeidet så virker det mot sin hensikt.” (v)

Denne kommentaren er helt i tråd med lærernes tilbakemeldinger når det gjelder prosjektet generelt; de opplever arbeidet som nyttig og lærerikt, men tidkrevende. Dette kommer vi nærmere tilbake til i 4.7.5.

Disse funnene er helt i tråd med tilbakemeldinger i intervjuene. Ikke uventet trakk mange fram tidspresset i en aktiv skolehverdag som en stor utfordring, og for mange en kilde til frustrasjon. På spørsmålet om hva som nå skulle til for at å komme videre, svarer mange av lærerne;

- ”Tid!” (b)
- ”Mer sammenhengende tid.” (u)
- ”Tid til samlinger som er planlagt.” (v)

Felles for hele skolen

En annen viktig faktor for at skolen skal lykkes med å utvikle lærernes kompetanse i tråd med prosjektets mål, er et felles prosjekt der alle på skolen er med. Stikkord her er samarbeid, kompetanse- og erfaringsdeling og hensiktsmessig organisering av skolens ressurser. En tredel av de som svarte på spørsmål 13 rapporterer at dette er en viktig faktor.

Kompetanseutvikling

Behovet for kompetanseutvikling er også et viktig stikkord. Lærerne rapporterer behov for ekstern hjelp og informasjon for både å komme i gang med utvikling av kjennetegn og videre med utprøvingen.

Lærerne gir i spørreskjemaet uttrykk for at arbeidet med kjennetegn er krevende, men viktig;

- *”Vanskelig, men viktig.”* (b)

Også i intervjuene uttrykte lærerne behov for kompetanseutvikling knyttet til elevvurdering. På spørsmål om det var noe de savnet, kom flere av fokusgruppene inn på behovet for ekstern hjelp på dette området;

- *”Savner en ekstern kontaktperson som vi kan stille spørsmål til”* (v)
- *”Etterlyste en veileder. Skoleeier lovet å knytte kontakter, uten at det har skjedd så mye!”* (u)

Samarbeid internt på skolen og mellom skoler

Samarbeid både innad på skolen og mellom skoler er en annen faktor som vektlegges av mange. Dette er omtalt i 4.4.1 og 4.4.3.

Forankring i skoleledelsen

Behov for at skoleledelsen og spesielt rektor, er involvert i prosjektet framheves i alle intervjuene som en helt sentral faktor for å få en god utvikling i arbeidet med kjennetegn (se 4.4.2). På de skolene som rapporterer om gode prosesser og positive erfaringer, har rektor en solid forankring inn i prosjektledelsen. Skolene preges av tydelig ledelse, med klare forventninger og bestillinger knyttet til utvikling og utprøving av kjennetegn.

På skoler der lærerne ikke opplever at rektor engasjerer seg i prosjektet, oppleves dette som svært demotiverende for arbeidet;

- *”Få i ledelsen er med på møtene, det gir dårlige signaler.”* (u)

Skoleledelsens tilstedeværelse og engasjement i prosjektet øker fokus og gir prosjektet integritet. Gjennom bred skoleforskning i de senere årene er det grundig dokumentert at rektor har en helt sentral rolle for at utvikling på skolen skal finne sted. Da er det oppsiktsvekkende at lærerne ved tre av de fire skolene som ble intervjuet, i dette prosjektet opplever utydelig eller fravær av skoleleder her representert ved rektor. Ledelsen av prosjektet er ofte delegert, i beste fall til en inspektør. I mangel av nærværende leder, rapporterer lærerne om stor frustrasjon knyttet til organiseringen, mangel på kompetanseutvikling, noe som resulterer i mye famling, stor usikkerhet og liten motivasjon for arbeidet;

- *”Kun en inspektør engasjerer seg.....rektor burde vært inspirator, men deltar ikke...”* (b)

Ved en av skolene er lærerne veldig tydelige på rektors positive bidrag inn i prosjektet i form av tydelig ledelse, klare rammer og uttalte bestillinger. Denne skolelederen blir beskrevet ved;

- *”Dyktig og aktiv ledelse av skolen.”* (b)

Lærerne ved denne skolen rapporterer om en leder som legger til rette for systematisk jobbing med målrelatert vurdering. Lærerne reagerer med;

- *”Stor lojalitet overfor pålegg og andre innspill fra ledelsen.”* (b)

De opplever en kontinuitet i arbeidet med vurdering, både med hensyn til videreføring av tidligere satsing og med utvikling av kjennetegn.

4.6.2 Kontakt med eksterne fagmiljøer

På spørsmål om skolene har hatt kontakt med fagmiljøer i forbindelse med prosjektet, svarer lærerne på følgende måte:

Tabell 4.11: Prosentandel av lærerne som sier at skolen har knyttet til seg eksternt fagmiljø. Pilot skolene er her holdt utenfor.*

Barnetrinn				Ungdomstrinn	Videregående
A	B	C	Alle	D	D
44	45	34	42	31	22

* Forskjeller mellom trinnene og modellene er ”signifikante” (se 3.5)

Tabell 4.11 viser hvor stor andel av lærerne som rapporterer at skolen har kontakt med eksternt fagmiljø i forbindelse med prosjektet. Sett i lys av at skoleeierne gjennom avtalen med Utdanningsdirektoratet forpliktet seg til å etablere samarbeid med eksterne i forbindelse med kompetanseutvikling for lærerne på egen skole, oppfatter vi disse tallene som uventet lave. Det er for øvrig verdt å merke seg at de skolene som skal utvikle kjennetegn på egen hånd og som man i utgangspunktet ville tro hadde klart mest behov for veiledning i forbindelse med arbeidet (dvs. modell A og D), samlet sett i mindre grad enn de andre skolene har hatt kontakt med eksterne fagmiljøer. Det er verd å merke seg at det er en tydelig forskjell mellom trinnene; jo høyere trinn, desto mindre er ekstern kompetanse trukket inn. Dette er vel å merke lærernes individuelle oppfatning av situasjonen, og det har heller ikke vært noe krav fra Utdanningsdirektoratets side om at en slik kontakt skal etableres helt fra starten av prosjektet.

I forbindelse med intervjuene kom det fram at skolene også definerte kontakt med ansatte i Utdanningsdirektoratet som ”kontakt med eksternt fagmiljø”. Da dette spørsmålet i første rekke var ment å skulle fange opp skolenes kontakt med pedagogiske fagmiljøer (f.eks. innen universitets- og høgskolesektoren), er derfor tallene i tabellen noe misvisende. Men de generelt lave tallene i tabellen kommentert ovenfor blir i lys av dette enda mer påfallende.

I spørreskjemaet var det også spørsmål om *hva* eksterne fagmiljøer har bidratt med så langt. Det viser seg at det i første rekke dreier seg om ”kontakt med skolens ledelse”, ”kurs om elevvurdering” og ”deltakelse ved oppstart av prosjektet”. Som tidligere nevnt, viser intervjuene at lærerne også oppfatter ansatte i Utdanningsdirektoratet som en del av et eksternt fagmiljø i forbindelse med prosjektet. Her inngår blant annet de regionale oppstartsamlingene som Utdanningsdirektoratet arrangerte for representanter fra skolene. Dataene er derfor noe usikre med hensyn til skolenes kontakt med pedagogiske fagmiljøer. Betydelig færre lærere svarte ja på spørsmålet om noen fra eksternt fagmiljø har bidratt med ”veiledning av lærergrupper/enkeltlærere”.

Da lærerne ble bedt om å nevne to faktorer som var viktige for å lykkes med prosjektet (*spm 13*), oppga over 20 prosent av de som har besvart dette spørsmålet, at de har behov for å bedre sin kompetanse i elevvurdering. Dette inntrykket ble forsterket i forbindelse med fokusgruppeintervjuene. På skoler som ennå ikke hadde etablert kontakt med en ekstern kompetanseperson, signaliserte lærerne et sterkt behov for kontakt med en ekstern veileder med kompetanse på elevvurdering, og som også hadde god kjenneskap til prosjektet. Ved en av intervju skolene kom det fram at skolen hadde meldt i fra om dette behovet til skoleeier, men det hadde ikke skjedd noe i saken ennå.

- ”Vi har etterlyst en veileder overfor skoleeier. Skoleeier har lovt å knytte kontakt, uten at det har skjedd noe.” (u)
- ”Vi savner ekstern veiledning. En som kunne si om vi har tenkt riktig. Vi ønsker tilbakemelding på det vi har gjort.” (u)
- ”Vi ønsker en ekstern kontaktperson som vi kan stille spørsmål til.” (v)

Bildet her er imidlertid noe mer nyansert, for dette behovet ble ikke like sterkt uttrykt blant alle lærerne i videregående som ble intervjuet.

- ”Jeg savner det ikke. Det kan lett bli altfor teoretisk hvis vedkommende ikke kjenner faget.” (v)
- ”Vi kan mest om faget selv, kan klare det selv bare vi får tid til å utvikle kjennetegn.” (v)

Slik vi ser det, kan det synes som om det er hensiktsmessig for skolene å knytte kontakt med eksterne kompetansemiljøer på et tidlig stadium av prosjektet. Dette er verd å merke seg for skolene når de nå står foran et nytt skoleår med videreføring av prosjektet.

4.6.3 Hvor stort utbytte har skolene hatt av kontakten med eksterne fagmiljøer?

Lærerne ble også spurt om hvor fornøyd de er med den eksterne veiledningen. Kun 37 % av lærerne besvarte dette spørsmålet, og fordelingen av svar er gitt i tabell 4.12 i form av prosentandel av de som rapporterte slik kontakt. Pilotskolene har fått mye spesiell oppfølging og er derfor ikke representative. De er derfor behandlet særskilt i denne sammenhengen.

Det framgår av tabell 4.12 at lærerne i hovedsak er fornøyd med den veiledningen de har fått fra eksternt fagmiljø, selv om det er noe variasjon mellom modeller og trinn. Misnøyen synes å være større på videregående skole enn på de lavere trinnene, men det er viktig å huske på at det er få på videregående som har fått slik veiledning i det hele tatt. Og når det gjelder barnetrinnet, synes mange av lærerne med modell A tydeligvis at veiledningen ikke har vært god nok. Dette må selvsagt ses i lys av det krevende arbeidet med å utvikle kjennetegnene selv.

Tabell 4.12: Hvor fornøyde lærerne er med bidraget fra eksterne fagmiljøer. Svarene er gitt som prosent av de som har rapportert slik kontakt. Resultater er gitt både for ikke-pilotskoler og for pilotskoler.

	Trinn og modell	Svært fornøyd	Fornøyd	Misfornøyd	Svært misfornøyd
Ikke pilotskoler (*M)	Alle	3	73	22	3
	Barnetrinn, modell A	2	61	30	7
	Barnetrinn, modell B	3	91	6	0
	Barnetrinn, modell C	5	70	20	5
	Barnetrinn, alle	3	74	19	4
	Ungdomstrinn	5	75	20	0
	Videregående	0	68	32	0
Pilotskoler	Barnetrinn, alle	14	76	10	0
	Ungdomstrinn	7	67	27	0

*M Forskjeller mellom modellene (men ikke mellom trinnene) er ”signifikante” (se 3.5)

Tabellen viser også data for pilotskolene spesielt. Her er det få skoler, og prosenttallene (regnet av de 57 lærerne som har svart) må derfor vurderes med forsiktighet. På barnetrinnet framstår likevel pilotskolene generelt med stor entusiasme for den veiledningen de har fått.

Entusiasmen på ungdomstrinnet framstår derimot som vesentlig lavere, faktisk lavere enn skolene som ikke var pilotskoler. Ingen videregående skoler var pilotskoler.

4.7 Lærernes tanker om prosjektet

4.7.1 Utvikling av kjennetegn lokalt eller nasjonalt?

Lærerne ble også bedt om å ta stilling til noen utsagn som dreide seg om utvikling av kjennetegn (*spm 21*). Tabell 4.13 viser lærernes svar på disse utsagnene.

Tabell 4.13: Lærernes syn på utvikling av kjennetegn (angitt som prosent som er enig i utsagnet).

	Barnetrinn				Ungdoms-	Videre-
	A	B	C	Alle	trinn	gående
a) Jeg mener det er viktig at skolene selv utvikler kjennetegn på måloppnåelse i fag. *	37	36	30	35	D 45	D 49
b) Jeg mener kjennetegnene bør utvikles nasjonalt.	86	82	88	85	82	82
c) Jeg mener skolene bør kunne velge om de vil utvikle kjennetegn lokalt eller bruke nasjonalt fastsatte kjennetegn.	29	39	30	33	40	35
<i>Skal kun besvares av lærere som prøver ut modell A og D: ¹</i> e) Jeg synes det går greit å utvikle kjennetegn på måloppnåelse på tvers av kompetanser i fag.	57	-	-	-	61	51

¹ Dette spørsmålet ble besvart av til sammen 58 % av lærerne.

* Forskjeller mellom trinn og modeller er ”signifikante” (se 3.5)

Det går tydelig fram at et stort flertall av lærerne ønsker nasjonalt utviklede kjennetegn. Dette inntrykket ble ytterligere forsterket i forbindelse med intervjuene. På dette området var fokusgruppene helt samstemte.

- ”De bør være nasjonale. Det bør være noe felles for hele landet, det blir mest rettferdig.” (u)
- ”Kjennetegnene bør være felles nasjonale. Det sikrer at det blir likt på alle skoler. Skal vi gjøre jobben?” (b)
- ”Dette burde vært utviklet av de som utarbeidet læreplanen.” (v)
- ”Jeg ønsker nasjonale, pga tiden. Hvorfor skal alle gjøre denne jobben? Men de ferdige kjennetegnene er ulne og vage.” (b)
- ”Utdanningsdirektoratet burde ha engasjert noen dyktige fagfolk til å gjøre dette.” (b)
- ”Dette er en for stor jobb å legge på lærerne.” (b)
- ”Det er en fare for at det blir halvgjort, det kommer på toppen av alt det andre vi skal gjøre.” (b)
- ”Hvorfor er ikke dette laget sentralt? Hvorfor må vi gjøre dette?” (u)
- ”Det sikrer at det blir likt i hele landet. Det sikrer ved inntak til andre skoler. Skolen skal ikke bety noe for vurderingen.” (v)
- ”Læreplanen er utydelig, og nå må vi gjøre jobben selv.” (v)

I spørreskjemaet var det imidlertid ikke alle som oppfatter at de to første utsagnene (a og b) står i motsetning til hverandre. Beregninger viser at så mange som 25 prosent av lærerne er enig i begge utsagnene. Dette kan altså innebære at kjennetegnene bør utvikles og bestemmes nasjonalt, men at de allikevel ”utvikles” videre lokalt før de kan komme i praktisk bruk. Vi oppfatter dette som et viktig poeng.

Det at godt over 80 prosent av lærerne ønsker nasjonalt utviklede kjennetegn, kan foruten de argumentene som kommer fram i sitatene foran, ha sammenheng med lærernes usikkerhet når det gjelder taksonomi (jfr. 2.2 og 4.3.4). Ett av utsagnene i spørreskjemaet (*spm 21e*) var kun rettet mot de skolene som skulle utvikle kjennetegn på egen hånd (dvs. modell A og D). Som det går fram av tabell 4.13, svarte litt over halvparten av lærerne på A- og D-skolene at ”det går greit å utvikle kjennetegn”. Det betyr på den andre siden at nesten halvparten av lærerne sa seg uenig i dette utsagnet. Dette kan ha sammenheng med lærernes usikkerhet når det gjelder å formulere beskrivelser av kjennetegn på måloppnåelse på ulike nivå.

4.7.2 Lærernes erfaringer med bruk av kjennetegn

Tabell 4.14 viser lærernes svar på noen utsagn knyttet til hvordan de opplever å bruke kjennetegnene i sin vurderingspraksis (*spm 24 A-C* for ulike fag). Utsagnene var stilt knyttet til bestemte fag, og svarene for de fire mest sentrale fagene i prosjektet er vist i tabellen. (Her er det viktig å huske på at spørsmålene bare skulle besvares av lærere som hadde kommet i gang med å bruke kjennetegnene i praksis, og særlig i samfunnsfag og i mat og helse gjaldt dette få lærere).

Tabell 4.14: Prosent av lærerne som er ”Enig” eller ”Svært enig” i noen utsagn om bruk av kjennetegnene i vurderingen (Svar i prosent av de som har tatt kjennetegnene i bruk og svarte på spørsmålene).

	Mate- matikk N=145	Norsk N=165	Samf. fag N=38	Mat og helse N=24
a) Kjennetegnene klargjør for meg hva elevene må gjøre for å forbedre seg faglig.	77	83	84	83
b) Kjennetegnene bidrar til en mer rettferdig vurdering.	78	89	92	83
c) Kjennetegnene bidrar til en mer faglig relevant tilbakemelding til elevene.	84	90	97	92
d) Kjennetegnene er et nyttig verktøy i kommunikasjonen mellom lærer og elev i forbindelse med vurdering.	83	90	97	92
e) Ved bruk av kjennetegn er det lettere enn tidligere å vurdere elevene kun ut fra faglige kriterier.	69	82	83	79
f) Ved bruk av kjennetegn er skillet mellom vurdering av prosess og vurdering av elevens kompetanse tydelig for lærere og elever.	72	81	87	88
g) Kjennetegnene har resultert i større åpenhet rundt vurdering av elevene.	68	76	89	79
h) Kjennetegnene er et nyttig verktøy i kommunikasjonen med de foresatte i forbindelse med vurdering av eleven.	81	82	91	92
i) Kjennetegnene gjør det lettere for meg å skille mellom prosess og kompetanse i tilbakemeldinger til eleven.	66	80	82	83

Tabell 4.14 viser at lærerne gjennomgående vurderer kjennetegnene som positive og nyttige bidrag i elevvurderingen. Det er en tendens til at det gjennomgående er kjennetegnene i matematikk som i minst grad er beskrevet i positive termer.

Lærerne på ungdomstrinnet og i videregående skole fikk også spørsmål om forholdet mellom kjennetegnene og karakterer (spm 25). Omtrent halvparten av lærerne på disse trinnene svarte på spørsmålene, trolig fordi de ikke hadde kommet langt nok i prosjektarbeidet. Men så godt som alle som svarte, var enige i at kjennetegnene:

- klargjør hva som ligger i de ulike karakterene (spm 25a)
- fører til en vurdering med karakterer på bakgrunn av kriterier (spm 25b)
- bidrar til å konkretisere karakternivåene (spm 25c)
- gjør det mer meningsfullt å kommunisere med elevene om karakterene (spm 25f)

Derimot er det kanskje uventet at så ”mange” som 20 prosent av lærerne er enig i utsagnet ”Sammenlikning mellom elever er viktig i min elevvurdering” (spm 25e). Som vi også kommenterer i 4.4.4 representerer vektlegging av sammenlikning mellom elever på mange måter det motsatte av det prosjektet står for.

4.7.3 Forholdet mellom læreplan og kjennetegn

Tabell 4.15 viser lærernes svar på noen spørsmål om forholdet mellom kjennetegn og kompetansemål (spm 23A-C). Det er viktig å merke seg at de som har besvart dette spørsmålet, utgjør en mindre del av alle lærerne som var med i spørreundersøkelsen, da dette var på et tidlig tidspunkt i prosjektet for skolene. Bare lærere som har kommet i gang med bruk av kjennetegnene, skulle svare på dette. Til tross for lavt antall besvarelser på spørsmålet, kan det med tanke på det videre arbeidet med prosjektet være interessant å ta med resultatene. Det er gjennomgående bred tilslutning om alle disse utsagnene bortsett fra det siste.

Tabell 4.15 Prosent av lærerne som er ”Enig” eller ”Svært enig” i noen utsagn om forholdet mellom kjennetegnene og kompetansemålene i læreplanen (spm 23). Svar oppgitt i prosent av de som har tatt kjennetegnene i bruk i vedkommende fag og svarte på spørsmålene

	Mate- matikk N=145	Norsk N=165	Samf. fag N=38	Mat og helse N=24
a) Det er samsvar mellom kompetansemålene og kjennetegnene i fagene.	85	88	91	95
b) Kjennetegnene gir uttrykk for kvaliteten på elevenes kompetanse.	84	88	91	90
c) Kjennetegnene måler elevenes kompetanse på tvers av kompetansemålene.	72	73	80	80
d) Jeg har god kjennskap til kompetansemålene i LK06.	86	87	98	80
e) Jeg har god kjennskap til kjennetegnene i de fagene jeg underviser i.	78	81	79	85
f) Arbeidet med kjennetegnene bidrar til å klargjøre kompetansemålene i fagene for meg som lærer.	84	91	91	85
g) Kjennetegnene bidrar til å klargjøre kompetansemålene for elevene.	74	79	78	85
h) Kjennetegnene erstatter kompetansemålene i læreplanen.	20	21	27	11

Resultatene viser at disse lærerne i stor grad er enige i at det er samsvar mellom kompetansemålene og kjennetegnene, at kjennetegnene gir uttrykk for kvaliteten på elevenes kompetanse og at arbeidet med kjennetegn bidrar til å klargjøre kompetansemålene for dem som lærere. Dette er i samsvar med viktige prinsipper som ligger til grunn for prosjektet fra Utdanningsdirektoratets side.

Hvis vi sammenlikner fagene, ser vi likevel at disse utsagnene får noe lavere tilslutning når det gjelder matematikk enn for de andre fagene. Denne forskjellen ble også bekreftet i intervjuene. I flere av fokusgruppene ble matematikk trukket fram som det faget der lærerne opplevde det vanskelig å se sammenhengen mellom de ferdigstilte kjennetegnene og kompetansemålene i faget. Dette gjenspeiles også i lærernes noe lavere tilslutning til spørsmålet om hvor enige de er i at kjennetegnene måler elevenes kompetanse på tvers av kompetansemålene (*spm 23c*).

Når det gjelder det siste utsagnet om at kjennetegnene erstatter kompetansemålene (*spm 23h*), ser vi at såpass mange som en av fem lærere gir sin tilslutning. Dette står i strid med intensjonen om hvordan kjennetegnene skal fungere. I veiledningen til prosjektet beskrives forskjellen mellom kjennetegn og kompetansemål på følgende måte:

- ”Mens kompetansemålene sier noe om hva som skal nåes, beskriver kjennetegnene hvordan eleven mesterer kompetansemålene” (Utdanningsdirektoratet 2007, s. 5)

Det er viktig at man skiller klart mellom disse to begrepene. Kompetansemålene beskriver hva eleven skal mestre etter endt opplæring på ulike trinn, mens kjennetegnene beskriver hva eleven faktisk mesterer. Kjennetegnene skal ikke erstatte kompetansemålene, men være en støtte i vurderingen av måloppnåelse.

I intervjuene rapporterte mange av lærerne at de opplevde skillet mellom kompetansemål og kjennetegn som vanskelig når det kommer til utvikling og anvendelsen av kjennetegnene i fagene. De beskriver en forståelse av kjennetegn som en utvikling av kriterier til delmål av kompetansemålene i LK06. Mange opplever de ferdigstilte kjennetegnene som vide, uklare og ikke hensiktsmessige til å anvende i vurdering av elevenes måloppnåelse. Skolene som skal prøve ut ferdigstilte kjennetegn, omarbeider og lager nye utgaver av kjennetegnene før de prøves ut. De av skolene som skal utvikle egne kjennetegn, rapporterer at de har liten nytte av eksemplene i informasjonsheftet fra Utdanningsdirektoratet. Dette resulterer i at de famler mye og sliter med å finne *hva* og *hvordan*. Flere av skolene etterlyser en ”mal” og flere gode eksempler.

Disse resultatene viser at å avklare forholdet mellom kjennetegn og kompetansemålene i fagplanene representerer en stor utfordring, som det er viktig at skolene er klar over i det videre arbeidet med prosjektet.

4.7.4 Hvor fornøyd er lærerne med å delta i prosjektet?

Det framgår av tabell 4.16 at den store majoriteten av lærerne opplever deltakelsen i prosjektet som positivt. Forskjellen mellom trinn og mellom modeller er ubetydelig.

Tabell 4.16: Prosent av lærere som er positive til at skolen er med i prosjektet, fordelt på trinn og modell (pilotskolene er holdt utenfor)(Spm 21d)

Barnetrinn				Ungdomstrinn	Videregående
A	B	C	Alle	D	D
83	82	73	81	81	87

Lærernes positive holdning gjenspeiles i følgende sitater hentet fra spørreskjemaet (spm 22):

- *"Interessant og engasjerende. Har bidratt til at jeg som lærer har gitt elevene mer utstrakt underveisvurdering og en mer målrettet tilbakemelding. Dette igjen har resultert i tilfredse, motiverte elever."* (b)
- *"Jeg opplever at dette er et interessant, spennende og svært nyttig utviklingsarbeid. Jeg har lært mye om vurdering og måloppnåelse, og jeg opplever at denne kompetansen styrker min profesjonalitet som lærer."* (b)

Lærernes positive holdning til prosjektet blir også bekreftet gjennom mange andre sitater i denne rapporten (jfr 4.7.5).

4.7.5 Nyttig og lærerikt, men tidkrevende arbeid og uklar organisering

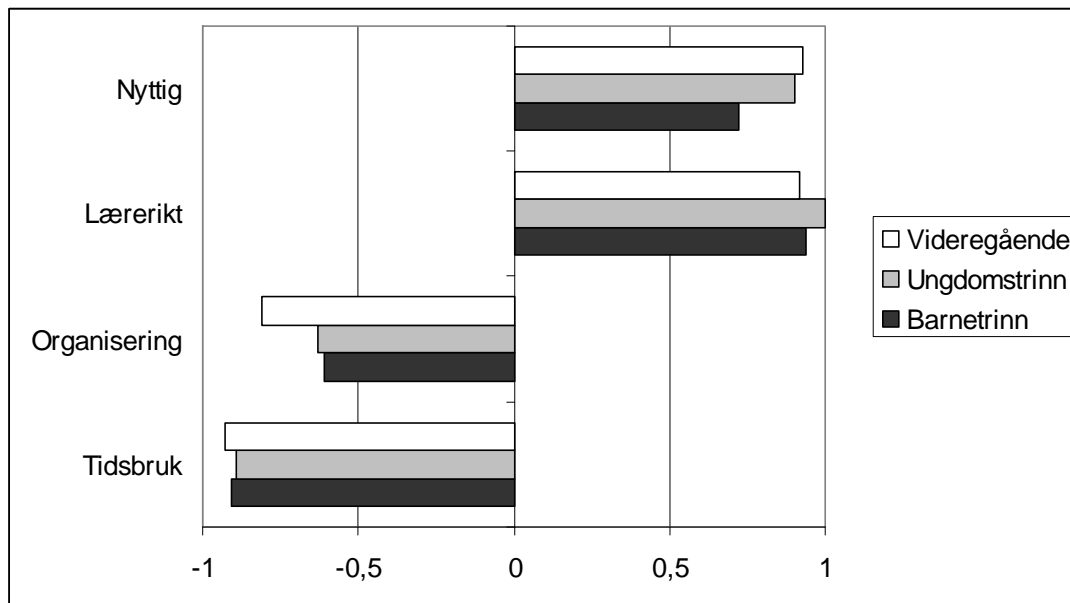
Lærerne fikk også anledning til å beskrive med egne ord hvordan de opplever at det er å delta i prosjektet *Bedre vurderingspraksis*. Dette var utformet som et åpent spørsmål (spm 22), og lærernes svar ble kodet ut fra om de inneholdt utsagn om følgende:

- Nyttig
- Lærerikt
- Organisering
- Tidsbruk

Disse kategoriene var definert på forhånd på bakgrunn av lærernes svar på dette spørsmålet i forbindelse med utprøvingen av spørreskjemaet. Det viste seg at de fungerte godt som analysekategorier. Alle utsagnene ble kodet med 1 for positive utsagn, 0 for nøytrale og -1 for negative. Det var derfor lett å beregne gjennomsnitt for hvert klassetrinn, og dette er vist grafisk på figur 4.1.

Som det går fram av figuren, var lærerne særlig positive til prosjektet når det gjelder opplevelsen av at det lærerikt å delta. Det at så mange av lærerne har trukket fram nettopp dette, underbygger antakelsen om behovet lærerne har for kompetanseutvikling innen elevvurdering. Dessuten var det mange lærere som påpekte at de opplever arbeidet med kjennetegnene som nyttig i vurderingsarbeidet sitt.

Figur 4.1: "Gjennomsnittlig" klassifisering av svar etter en skala fra -1 for bare negative utsagn og til 1 for bare positive utsagn. Gjennomsnittsverdier er vist for hvert skoletrinn. (Spm 22)



Nedenfor gjengis noen utsagn fra spørreskjemaet som ytterligere kan illustrere lærernes holdning til prosjektet:

- "Jeg opplever det som et incitament både til å utvikle egen vurderingspraksis og til å sikre høyere grad av rettferdighet i vurderingen." (u)
- "Hvorfor har ikke dette skjedd før er for meg en gåte, og for mange et spørsmål." (u)
- "Jeg lærer mye av dette og får gått mye i dybden og tenkt grundig gjennom min vurderingspraksis. Det er bra." (v)
- "Interessant, relevant for det faglige arbeidet, utfordrer oss på å være konkrete, gir oss et bedre verktøy, mer rettferdig vurdering for elevene, jeg blir bedre til å gi konkrete og utviklingsfremmende kommentarer." (u)
- "Jeg får god fokus på egen praksis. Dette stimulerer til refleksjon og utvikling for meg. – Fra dag en! Jeg liker utfordringer – dette er et spennende prosjekt, svært nyttig og nødvendig. Jeg har lenge ventet på noe likt. Det var på høy tid." (b)

Figur 4.1 viser at lærerne også trakk fram høy tidsbruk og mangelfull organisering som negative forhold ved prosjektet. Lærerne har først og fremst fokusert på tidsfaktoren. Det er ikke fullt så mange som har ytret seg negativt om organiseringen. Følgende utsagn hentet fra spørreskjemaet viser noe av denne misnøyen:

- "Feil tidspunkt for vår skole – deltar allerede i et toårig prosjekt – det er nok. Pga pkt 1 (dvs. punktet over) blir det for lite tid til å utvikle prosjektet. Tilgjengelig informasjon er mangelfull og uklar. Lav motivasjon for å arbeide med prosjektet pga pkt 1-3" (de tre punktene over) (b)
- "Prosjektet oppleves som uklart og lite strukturert. Alt for mye tid går med til å finne ut hva vi egentlig skal gjøre. Vi jobber etter modell B og ingen har klart for seg hva som er de ferdige kjennetegnene." (b)
- "Fikk beskjed om at dette var et ferdig opplegg, men det viser seg at vi må gjøre alt selv. Det burde vært et ferdig opplegg som vi kunne bruke direkte, det har gått med mye tid til et arbeid mange oppfatter som unyttig eller lite matnyttig." (u)

- *”Tidkrevende og kommer i tillegg til allerede bestemte oppgaver. For uklare kriterier knyttet til utprøving av modellene. Kriteriene i modellen er beskrevet, men vi får signaler om at vi kan gjøre en del som vi vil – tilpasse til egne behov. Dette skaper forvirring. Vi har brukt for mye tid på Kunnskapsløftet fra før. Kompetansemålene er altfor lite konkrete, og det er et møysommelig arbeid å tilpasse dette til innhold som gir mening.” (b)*

Det er ikke slik at lærerkommentarene utelukkende er enten bare positive eller negative. Flere av dem inneholdt utsagn med både positivt og negativt fortegn, noe sitatene under viser.

- *”Interessant, men det tar lang tid å komme inn i det. Nyttig i forhold til tilbakemelding til elever.” (b)*
- *”Vi må få bedre tid, slik at vi kan arbeide sammenhengende med dette. Ikke bare få 1-2 timer av og til. Ellers er det spennende å delta.” (u)*
- *”Vanskelig, men viktig ...” (b)*

Den tendensen som kom fram i analysen av dette spørsmålet i spørreskjemaet, ble bekreftet i forbindelse med intervjuene på skolene. Intervjuskolene var imidlertid forskjellige ved at de i varierende grad la vekt på henholdsvis de positive eller de negative aspektene ved å delta i prosjektet. Noen var gjennomgående mer positive til skolens deltakelse enn andre.

4.7.6 Informasjon til lærerne om prosjektet

Når vi i dette kapittelet går inn og refererer til lærernes utsagn knyttet til informasjon, er det viktig å presisere at dette er lærernes egne opplevelser av ulike aspekter ved informasjonen. Skolene har mottatt informasjon fra ulike hold, og det er skoleeier som etter avtale med Utdanningsdirektoratet har ansvaret for gjennomføringen av prosjektet ved sine skoler. I dette ligger også ansvaret for at nødvendig informasjon kommer ut til alle ledd i skolesystemet. Når vi velger å legge vekt på lærernes egne opplevelser, er det fordi vi mener at dette er av stor viktighet for det videre arbeidet med prosjektet ute i skolen. Vi ønsker å løfte fram lærerstemmene som tydelig framholder nødvendigheten av god og nødvendig informasjon for å gjøre et best mulig arbeid med kjennetegnene.

I spørreskjemaet er det ingen spørsmål som er direkte rettet mot hvordan informasjonen om prosjektet til lærerne har vært. I de åpne spørsmålene i spørreskjemaet var det allikevel så mange av lærerne som trakk fram og hadde fokus på nettopp dette, at vi vurderte dette som et viktig område å se nærmere på i forbindelse med intervjuene.

I det første åpne spørsmålet (*spm 13*) ble respondentene bedt om å nevne to viktige faktorer for at skolen skulle lykkes i arbeidet med å utvikle lærernes vurderingskompetanse. Svarene ble klassifisert etter kategorier som ble utarbeidet på bakgrunn av svarene på dette spørsmålet i utprøvingen av spørreskjemaet. En av analysekategoriene dreier seg om informasjon/veiledning. Av de som svarte på dette spørsmålet, trakk over 20 % av lærerne fram dette som en viktig faktor.

I det andre åpne spørsmålet i spørreskjemaet (*spm 22*) ble lærerne bedt om å beskrive hvordan de opplevde det å delta i prosjektet. En av analysekategoriene som ble benyttet, dreide seg om organiseringen. Flere av utsagnene som omhandlet organiseringen av prosjektet, var koblet opp mot informasjon om prosjektet. Mest fremtredende var kommentarene som gikk på at lærerne syntes de hadde fått for lite informasjon om prosjektene, samt uklare instruksjoner og målsettinger.

- *”Det kjem uklare meldingar. Ingen som veit kva vi skal gjere. Må få klarare instruksar frå sentralt hold”.* (b) (Her ble det ikke presisert hva som ble lagt i *sentralt hold*.)
- *”Altfor mye uklart”.* (v)

På bakgrunn av tilbakemeldingene på disse spørsmålene, ble dette emnet tatt opp i forbindelse med intervjuene. Fokusgruppene ble bedt om å fortelle hva de hadde fått av informasjon og hvordan denne var blitt formidlet til dem. De ble også bedt om å si noe om hva de nå i etterkant savnet av informasjon ved oppstart og hva slags informasjon som ville ha vært nyttig for dem å få, noe som vi anser som viktig med tanke på det videre arbeidet med prosjektet.

4.7.7 Hva fikk lærerne av informasjon før prosjektet startet?

I forbindelse med intervjuene kom det fram at lærerne både savnet informasjon knyttet til selve organiseringen og gjennomføringen av prosjektet, formålet med det og hva som ble forventet av dem. De etterlyste også informasjon som kunne gitt dem en bedre forståelse av den faglige siden ved prosjektet (dvs. arbeidet knyttet til kjennetegnene). Vi har mye datamateriale fra fokusgruppeintervjuene som illustrer dette.

”Vi fikk bare beskjed; dette skal vi være med på”

I alle fokusgruppene ble informasjon knyttet til den enkelte skoles deltakelse i prosjektet tatt opp av lærerne. Det viste seg at prosessen fram og begrunnelsen for skolens deltakelse i prosjektet var uklar for alle, noe som gjenspeiler seg i mange av kommentarene.

- *”Vi fikk bare beskjed; dette skal vi være med på. Vi ble ikke tatt med på råd. Det kunne vi godt ha blitt.”* (b)
- *”Det kom som lyn fra klar himmel i forbindelse med LK06, etter at LK06 var innført”*(v)
- *”Vi har ikke noe eierforhold til dette.”* (v)

I utgangspunktet valgte fylkesmannen ut skoleeiere som skulle delta. Det videre ansvaret for gjennomføringen ligger da på skoleeier og skoleledere. Kjennskap til denne prosessen kom ikke fram i intervjuene.

”Det gikk utrolig mange timer før vi skjønnte hva det gikk ut på.”

Også mangelen på informasjon om selve prosjektet før og ved oppstart ble sterkt framhevet i alle fokusgruppene. Dette dreide seg både om informasjon om selve prosjektet (organisering, formål), selve prosessen med å utvikle kjennetegn og usikkerheten som lærerne opplevde i forhold til sentrale begreper som kjennetegn, høy og lav måloppnåelse etc.

- *”Vi fikk veldig lite informasjon, visste ikke hva det gikk ut på.”* (v)

”Ulike forutsetninger pga ulik mengde informasjon”

I forbindelse med intervjuene hadde lærerne også synspunkter på at lærerne på samme skole hadde fått ulik mengde og type informasjon. På skoler hvor lærernes første møte med prosjektet var i forbindelse med de regionale oppstartsamlingene, ble det vurdert som uheldig at bare noen få av lærerne på skolen fikk delta på dette. De lærerne som deltok på oppstartsamlingen, opplevde på sin side det som vanskelig å skulle formidle til resten av personalet hva som lå i prosjektet da de selv syntes at de hadde mangelfull kjennskap til det. Lærerne rapporterte også om høy tidsbruk, stor usikkerhet og mye famling knyttet til ”å få tak på prosjektet og kjennetegnene”.

- *”Uheldig at så få var med på oppstartsamlingen.”* (u)

Lærernes usikkerhet med hensyn til *"Er vi på riktig vei?"* kom også til uttrykk i fokusgruppene. Denne usikkerheten var både knyttet til manglende kjennskap til målet med prosjektet og deres egen rolle i det.

- *"Hva gjør vi nå?"* (v)
- *"Vi vet ikke hva som forventes."* (b)

Motivasjonen øker med økt informasjon

Ikke uventet ble det uttrykt en klar sammenheng mellom lærernes motivasjon for å delta i prosjektet og mengden informasjon som de oppfattet som nyttig. Fokusgruppene rapporterte at dette spesielt gjaldt informasjon som førte til bedre forståelse av begrepet kjennetegn og arbeidet knyttet til utvikling og utprøving av disse.

- *"Først var det et ork – det ble bedre etter hvert."* (b)

4.7.8 Hva slags informasjon hadde vært viktig å få på et tidlig tidspunkt?

På spørsmål om hva slags informasjon lærerne nå i etterkant ser ville vært særlig nyttig for dem å få på et tidlig tidspunkt i prosjektet, ble det i gruppene fokusert på to områder. Det ene gikk på selve den faglige siden ved prosjektet (dvs. forhold knyttet til kjennetegn og utviklingen av dem), mens den andre dreide seg om kjennskap til prosjektet som helhet. Det er viktig å peke på at dette er lærernes opplevelser av den informasjonen som har vært tilgjengelig, både når det gjelder muntlig og skriftlig informasjon.

Kjennetegnenes egenart, utforming og modeller

I fokusgruppene var lærerne helt entydige når det gjaldt behovet for informasjon som gikk direkte på utformingen av kjennetegn. Dessuten uttrykte de stor grad av usikkerhet når det gjaldt kvaliteten på eget arbeid.

- *"Vi hadde trengt veiledning i forhold til hvordan vi helt konkret skulle utforme kjennetegn, og hadde behov for flere eksempler."* (v)

Når det gjelder informasjon knyttet til modeller og utvikling/bruk av kjennetegn, hadde lærerne ved en skole som skulle ta utgangspunkt i ferdigstilte kjennetegn, fått en forståelse av at de oppstilte eksemplene skulle kunne brukes direkte inn i vurderingsarbeidet uten videre bearbeiding.

- *"Vi trodde vi fikk ferdig utviklede kjennetegn som skulle prøves ut, men vi må jo tolke alle og konkretisere dem."* (b)

Som vi tidligere har vært inne på, er det ikke ment at kjennetegnene skal brukes direkte inn mot vurdering av elevenes arbeid, men være utgangspunkt for lærernes bearbeiding i samarbeid med elevene (jfr Utdanningsdirektoratets veiledningshefte *Vurdering*).

Ved en annen skole som også skulle prøve ut ferdigstilte eksempler på kjennetegn, hadde lærerne ut fra den informasjonen som var gitt, oppfattet at det var opp til dem å bestemme om de skulle ta utgangspunkt i de ferdigstilte kjennetegnene eller ikke.

- *"Skal vi følge modellen, eller skal vi det ikke? Vi forsto at det var opp til oss å bestemme"*. (b)

På grunn av det lærerne betegnet som "svært forvirrende signaler", hadde de selv tatt avgjørelsen om å utvikle kjennetegn uten å ta utgangspunkt i de ferdigstilte kjennetegnene.

Informasjon om prosjektets mål og forventninger

Ønsket om bedre kjennskap til prosjektet som helhet, kom også fram i alle intervjuene.

Flere av lærerne framhevet tidlig og god informasjon som helt vesentlig, både for å få kjennskap til hensikten med prosjektet og hele prosessen videre, forståelse for deres rolle i det og hva som forventes av dem. Dette var dessuten avgjørende for at lærerne skulle kunne føle et eierforhold til prosjektet og for deres motivasjon til å engasjere seg i arbeidet med kjennetegn. Behovet for å se sammenhengen mellom alle nivåer i prosjektet ble også poengtert.

- "Vi ønsker informasjon om helheten nasjonalt. Hvor er vi i forhold til hele prosjektet?" (b)
- "Vi ser ikke sammenhengen." (v)

Til tross for at lærerne var relativt samstemte når det gjaldt opplevd mangel på informasjon, fant vi store forskjeller i hvordan skolene og grupper innenfor enkelte skoler hadde løst denne utfordringen. Her er det grunn til å anta at skolens kultur har avgjørende betydning. Når vi vet at all forståelse er kontekst- og kulturavhengig, så blir det viktig at man på alle nivåer kvalitetssikrer en felles forståelse av sentrale begreper og føringer i prosjektet.

Vi som forskergruppe tar ikke her stilling til årsakssammenheng, men ønsker å løfte fram og peke på disse sentrale problemstillingene ute i skolene. Det er grunn til å anta at her er det mange faktorer som spiller inn, noe som tilsier at det blir desto viktigere at man tar disse signalene alvorlig og har fokus på informasjon videre i prosjektet. Det er viktig at man sikrer god nok informasjon om prosjektet i alle ledd i systemet i det videre arbeidet. Det er et felles ansvar å legge til rette for at informasjonen når fram og at den er i henhold til informasjonen fra Utdanningsdirektoratet.

4.8 Hva kjennetegner lærere og skoler som er positive til prosjektet og har hatt god progresjon?

Vi har studert nærmere hvilke suksessfaktorer som kan knyttes til prosjektet: Hva kjennetegner skoler og lærere med positiv motivasjon til prosjektet og god framdrift i prosjektarbeidet. Dette har vi studert ved å bruke verktøyet multippel regresjon. Uten å gå for mye i detalj vil vi bare peke på at vi med en slik metode kan finne fram til hvor mye av variansen (et mål for spredningen) i den avhengige variabelen (her: suksesskriteriene) som kan "forklares" (ikke nødvendigvis i kausal forstand) av de uavhengige variablene ("forklaringsvariablene") til sammen, og hvilke variabler som bidrar sterkt til dette.

Våre to suksesskriterier er

1. Motivasjon for prosjektet, slik dette kommer til uttrykk i hvor godt fornøyd lærerne er med at skolen deltar i prosjektet (*spm 21d*)
2. Framdriften i prosjektet slik dette kommer til uttrykk ved at kjennetegnene er tatt i bruk i vurderingsarbeidet (*spm 11c*)

Det første suksesskriteriet (motivasjon) påvirkes positivt først og fremst av disse faktorene:

- At skolen har knyttet til seg et eksternt fagmiljø (*spm 16*)
- At lærerne mener at det er viktig at skolen selv utvikler kjennetegn (her inkluderes trolig betydningen "eventuelt utvikler videre") (*spm 21a*)

- At lærerne er uenig i at det bare er kompetanse på høyt nivå som kan sies å ha mestring i faget (*spm 8h*)
- At arbeidet med prosjektet er organisert som en del av det lokale læreplanarbeidet.

Angående kriterium nr 2 ovenfor (god framdrift i prosjektet), så er det disse faktorene som framtrer som spesielt viktige:

- At arbeidet med prosjektet er organisert med elevdeltakelse (*spm 12c*)
- At lærerne er uenige i at elevenes innsats og aktivitet bør telle med som en viktig del av den faglige vurderingen (*spm 8b*)
- Økende framdrift med høyere skoletrinn
- At lærerne er positivt motivert for å delta i prosjektet (*spm 21d*)
- At skolen allerede hadde en plan for elevvurdering (*spm 9*)

For begge disse to suksesskriteriene gjelder at rundt 20 % av variansen kan ”forklares” av de faktorene som er trukket fram. Vi minner igjen om hvor nødvendig det er å være varsom med å tolke sammenhengene som kausale. Men likevel får vi noen signaler om hva det kan være viktig å være oppmerksomme på ved videre planlegging av prosjektet.

5 Anbefalinger

Vi avslutter rapporten med å peke på forhold som det er viktig å fokusere på i det videre arbeidet med prosjektet. Rådene bygger på analyse av data og faglige drøftinger. Analysene og drøftingene er sett i relasjon til bakgrunnen for prosjektet og den teoretiske forankringen som følgeforskningen har. Anbefalingene retter seg mot de ulike nivåene som har ansvaret for prosjektets gjennomføring, det vil si Utdanningsdirektoratet, skoleeiere og skoleledere.

Et godt grunnlag for endring

Dataene viser at lærerne har en svært positiv holdning til prosjektet. De aller fleste opplever arbeidet med kjennetegn som både nyttig og lærerikt. Lærerne rapporterer om økt bevissthet når det gjelder elevvurdering. Prosjektet har også ført til økt fokus på fag og fagsamarbeid, noe de ser på som særlig positivt. Flere lærere ga uttrykk for at de gjennom arbeidet med prosjektet er i ferd med å endre holdning og praksis i forhold til elevvurdering. Dette gjaldt særlig lærere på barnetrinnet. I det videre arbeidet anses det derfor som viktig at de ansvarlige på alle nivå i prosjektet bidrar til å opprettholde og videreutvikle lærernes positive erfaringer.

Foruten positivt engasjement fra deltakerne i prosjektet, viser dataene at det er andre sider ved prosjektet som også bør vies oppmerksomhet dersom en skal lykkes i det videre arbeidet. I det følgende vil vi peke på noen sentrale utfordringer.

Uklarheter knyttet til selve prosjektet

I forbindelse med intervjuene kom det fram at lærerne har mange spørsmål knyttet til prosjektet. Uklarhetene dreier seg både om mer overordnede forhold ved prosjektet, samt deres egen rolle i det. For det første er det uklart for mange hva som er prosjektets konkrete formål. Et tegn på det er at det knytter seg en viss usikkerhet til hva skolene egentlig er blitt bedt om å gjøre, og lærerne er usikre på hva som forventes av dem. De lurer også på "hva som skjer" når prosjektperioden er over.

Vi anser det som viktig at aktører på alle nivå tar disse signalene på alvor og gir dem oppmerksomhet, slik at det skapes bedre klarhet rundt prosjektets mål og hensikt. Dette vil bidra til økt motivasjon for det videre arbeidet. God informasjonsflyt er helt sentralt. En må imidlertid være oppmerksom på at dette er en utprøving, og det er i utprøvingens natur at ikke alle spørsmål vedrørende prosjektet lar seg besvare nå. I slike sammenhenger er det viktig at Utdanningsdirektoratet også formidler dette.

Usikkerhet og valg av terminologi

Dataene viser at det råder usikkerhet når det gjelder selve begrepet *kjennetegn*. Vi registrerte at det eksisterer ulike oppfatninger av selve termen, og lærere gir uttrykk for at det er vanskelig å få en klar forståelse av hva som ligger i betegnelsen.

Kjennetegn på måloppnåelse står helt sentralt i prosjektet. I den internasjonale forskningslitteraturen benyttes ofte betegnelsen *kriterier* i forbindelse med målrelatert vurdering. Kriterier brukes når en har en målestokk som utgangspunkt for vurderingen av et fenomen. I samtalene med lærerne kom det fram at mange av dem allerede benytter og er fortrolig med denne betegnelsen. Når det gjelder begrepet *kjennetegn*, brukes det gjerne når en skal beskrive det som er karakteristisk eller typisk for et fenomen. Internasjonalt brukes ofte betegnelsen standardisering når kriterier får felles beskrivelser (altså *kjennetegn*). Hovedsaken

er å ha kriterier når en vurderer om kompetansemålet er nådd. Når en har kriterier, kan kjennetegn brukes som beskrivelse på hva som er karakteristisk eller typisk for det eleven mestrer på dette gitte nivået. Vi anbefaler at denne terminologien (dvs. kriterier, kjennetegn og oppnåelse av kompetansemål) klargjøres i den videre gjennomføringen av prosjektet.

Faglig relevant tilbakemelding

I skolen har det vært vanlig å skille mellom kunnskap, holdning og ferdigheter (jfr. omtalen av taksonomi i kapittel 2). Kunnskap har en knyttet til det kognitive området, mens ferdigheter gjerne er knyttet til praktiske fag, altså til det å kunne gjøre noe. Holdning er knyttet til det affektive området, og her inngår verdier, interesser og elevens motivasjon og arbeidsinnsats. Når det gjelder spørsmålet om hva som inngår i *faglig relevant tilbakemelding*, er det viktig å være klar over at det ikke er elevens interesser, motivasjon og arbeidsinnsats som skal vurderes. Våre data vedrørende lærernes syn på vurdering tyder på at en her står overfor utfordringer. Det er for eksempel hele 66 prosent av lærerne som er enige i at innsats og aktivitet bør telle med i den faglige vurderingen, mens bare 17 prosent er uenig i dette standpunktet. Utsagn som at ”kravet til høy måloppnåelse har vært fraværende tidligere” og at en i stedet har lagt vekt på interesse og aktivitet i vurderingen illustrerer dette. Lærernes svar tyder på at det i vurderingen kan være vanskelig å skille mellom faglig kompetanse og læringsprosess. Faglig relevante tilbakemeldinger som samsvarer med den tenkningen som prosjektet ”Bedre vurderingspraksis” bygger på, innebærer et brudd med en tradisjon som vektlegger innsats, aktivitet og interesse i den faglige vurderingen. Vi anbefaler derfor at de ulike nivåene i prosjektet i det videre arbeidet bidrar til å gjøre enda klarere hva som forstås med *faglig relevant tilbakemelding*. Prosjektet kan også ses i sammenheng med den nye forskriften til opplæringsloven. I forskriften skilles det mellom faglig relevant tilbakemelding og annen tilbakemelding. Mye tyder på at prosjektet har satt i gang en god prosess for å utvikle en tradisjon for faglig relevant tilbakemelding som er i samsvar med forskriften.

Dataene viser også at det eksisterer oppfatninger vedrørende elevvurdering som Utdanningsdirektoratet, skoleeiere, skoleledere og eksterne kompetansmiljøer kan forsterke og bygge videre på. For eksempel er nærmere 90 prosent av lærerne uenige i påstanden om at det bare er elever som har kompetanse på høyt nivå, som har mestring i faget. Lærerne uttrykker dessuten at de gjennom arbeidet med prosjektet er i ferd med å endre praksis i forhold til elevvurdering, og mange opplever arbeidet med kriterier i form av kjennetegn som en god og naturlig videreutvikling av egen vurderingspraksis.

Gode eksempler

Lærere gir også uttrykk for at arbeidet med kjennetegn oppleves som krevende, og at dette er et utviklingsarbeid som tar mye tid. De ser det imidlertid som svært nyttig og viktig med tanke på en felles vurderingspraksis. Når det gjelder de ferdigstilte eksemplene på kjennetegn i veiledningsheftet til Utdanningsdirektoratet, oppfattes disse som ulne og generelle (skal prøves ut av modell B- og C-skolene). Enkelte lærere påpeker at de gir rom for mange tolkningsmuligheter, og de etterlyser derfor flere gode eksempler. Utdanningsdirektoratet har helt fra prosjektet startet opp oppfordret skoler og skoleeiere til å sende inn sine forslag til ”gode kjennetegn”. Disse arbeidene legges ut på nettet og er ment å skulle være utgangspunkt for refleksjon på den enkelte skole om hva gode nasjonale kjennetegn kan være og en støtte i skolens eget arbeid med å utvikle en felles vurderingspraksis. Vi mener at Utdanningsdirektoratet i større grad må kvalitetssikre disse bidragene for å unngå at det legges ut materiale på nettet som åpenbart er i strid med prosjektets hovedideer.

Utdanningsdirektoratet bør også formidle erfaringer fra prosjektet til skoler som ikke deltar i prosjektet, men som ønsker å utvikle sin kompetanse i elevvurdering.

Et helhetlig vurderingssystem

Lærere i grunnskolingen har ulike vurderingskriterier som de må forholde seg til. Vi tenker her spesielt på karakterskalaen (1-6) slik den skal brukes både for termin-, standpunkt- og eksamenskarakterer. Videre har nasjonale prøver etablert et system som er inspirert av internasjonale undersøkelser (TIMSS, PIRLS og PISA), med henholdsvis tre (5. trinn) og fem (8. trinn) kompetansenivåer, der kompetansen på hvert nivå er beskrevet i form av eksempler på hva elever på dette nivået typisk mestrer. Det må anses som et problem dersom kjennetegn for måloppnåelse i prosjektet står i motsetning til kriterier og kompetansebeskrivelser som benyttes i nasjonale prøver. Faren er at det kan oppstå unødvendig forvirring dersom en har ulike vurderingssystemer når det gjelder vurdering av elevens måloppnåelse. Vi anser det derfor som viktig at det nasjonalt blir etablert en klar forbindelse mellom de ulike vurderingssystemene.

Kompetanseutvikling

Ifølge St. meld. nr. 16 (2006-2007) er det en svak vurderingskultur og –praksis i skolen, og det hevdes at "... både lærerutdanningen og skolen mangler tilstrekkelig kompetanse" (s. 79). I kapittel 1 trakk vi fram bakgrunnen for at vi har en svak vurderingskultur og –praksis i skolen. En lærerutdanning som er i samsvar med den praksis som er beskrevet der, vil mangle den kompetansen som dette prosjektet krever. I en slik lærerutdanning vil det dessuten kunne eksistere en utbredt skepsis til målrelatert vurdering i forhold til oppsatte faglige krav på barnetrinnet. Dette tilsier at det ikke er en selvfølge at skolene har tilgang til kompetente veiledere i forbindelse med prosjektet. Dataene viser at behovet for kompetanseutvikling i elevvurdering er stort, og at relativt få skoler har etablert samarbeid med eksterne kompetansemiljøer.

I prosjektet legges det vekt på elevmedvirkning. Her er det imidlertid viktig å fastholde at prosjektet "Bedre vurderingspraksis" ikke bygger på en tankegang at det er eleven selv som fastsetter kriterier. Perspektivet er at eleven skal lære å vurdere sine prestasjoner i lys av gitte kriterier. Dette prosjektet er innrettet mot å gi faglige tilbakemeldinger på kompetansemål som er nasjonalt fastsatt i læreplaner for fag. Vurderingskriterier og kjennetegn på måloppnåelse kan utvikles på forskjellige nivåer i utdanningssystemet, for eksempel på en enkelt skole, altså i en lokal kontekst. Kriteriene og de utviklede kjennetegnene på måloppnåelse er imidlertid ikke bundet til en lokal kontekst. De er derimot forankret til nasjonalt fastsatte kompetansemål. Det er derfor viktig å fastholde at også de lokalt utviklede kjennetegnene på måloppnåelse må ha læreplaner for fag som sin referanse.

Både ut fra de utdanningspolitiske markeringene som det er referert til, og data som viser behovet for kompetanseutvikling når det gjelder elevvurdering, anbefaler vi at særlig Utdanningsdirektoratet følger opp kompetansemiljøene slik at veiledningen er i tråd med prosjektets intensjoner.

Forskning

I St. meld. 31 (2007-2008, s. 67-68) fins et avsnitt om styrking av utdanningsforskningen. Her skisseres et nytt og langsiktig forskningsprogram med tre tematiske prioriteringer, hvorav vurdering, vurderingsformer og kvalitetsutvikling er en av dem.

Vår følgeforskning tar opp spørsmål knyttet til evaluering av vurderingspraksis, men berører ikke sammenhengen mellom en eventuell forbedring av faglig relevant tilbakemelding og økt læringsutbytte hos elevene. Det er behov for mer kunnskap om læringsfremmende vurdering, og forholdet mellom elevvurdering og læringsresultat bør særlig vektlegges i fremtidig forskning. I denne sammenhengen vil vi også fremheve behovet for forskning rettet mot vurdering *av* læring (resultatvurdering) som kan fungere som læringsfremmende vurdering, altså som vurdering *for* læring. Nå som elevvurdering i hele grunnopplæringen kan ses under ett, er det behov for å se nærmere på hvordan elevens kompetanseoppnåelse etter de ulike trinnene i opplæringen utgjør en forutsetning for videre læring. Det er særlig grunn til å fremheve forskning rettet mot sammenhengen mellom sluttvurdering på grunnskolen som forutsetning for læring i videregående opplæring.

Litteratur:

Anderson, L.W. og Krathwohl, D. (2001). *A Taxonomy for learning, teaching, and assessing*. NY: Longman.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (1998). *Working inside the black box – Assessment for learning in the classroom*. London: School of Education, King's college.

Black, P., Harrison, C., Lee, C., Marshall B. & Wiliam, D. (2006). *Assessment for learning. Putting it into practice*. Berkshire: Open University Press.

Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1. Cognitive domain* (ed.). New York: McKay.

Dale, E.L. (2008). *Felleskolen – skolefaglig læring for alle*. Oslo: Cappelen Akademisk Forlag.

Dale, E.L. og Wærness J.I. (2006). *Vurdering og læring i en elevaktiv skole*. Oslo: Universitetsforlaget.

Danielsen, I.J., Skaar, K. og Skaalvik, E.M. (2007). *De viktige få – Analyse av Elevundersøkelsen 2007*. Oxford Research/NTNU.

Engelsen, B.U. (2006). *Kan læring planlegges? Arbeid med læreplaner – hva, hvordan og hvorfor*. Oslo: Gyldendal Akademisk.

Engh, R. (2007). *Sluttrapport for piloteringen med utprøvingen av nasjonale kjennetegn på måloppnåelse i fire fag i grunnskolen*. Høyskolen i Vestfold, Avdeling for lærerutdanning.

Forskrift til opplæringslova (2006). Kunnskapsdepartementet.

Glaser, R. (1972a). Instructional technology and the measurement of learning outcomes: Some questions. I: E.J. Popham (ed): *Criterion-referenced measurement* (5-14). Englewood Cliffs, NJ: Educational Technology Publications.

Glaser, R. (1972b). A criterion. Referenced test. I: E.J. Popham (ed): *Criterion-referenced measurement* (41-51). Englewood Cliffs, NJ: Educational Technology Publications.

Glaser, R. og Klaus, D.J. (1962). Proficiency measurement: Assessing human performance. I: R.M. Gagné (ed.): *Psychological principles in system development* (419-474). NY: Holt, Rinehart and Winston.

Grøgaard, J.B. (1997). En historie som har fått vasket seg? Om oppfølgingstjenestens målgruppe – rekruttering og tiltak første skoleår. I: B. Lødding og K. Tornes (red.): *Idealer og paradokser. Aspekter ved gjennomføring av reform 94*. Oslo: Tano Aschehoug.

Haugstveit, T.B. (2005). Vurdering som profesjonsfaglig kompetanse – Læreres refleksjoner over egen vurderingspraksis på 5., 6. og 7. trinn. *Norsk pedagogisk tidsskrift*, 6, 417-430.

Hægeland, T. og Kirkebøen, L.J. (2006). *Skoleresultater 2006. En kartlegging av karakterer fra grunn- og videregående skoler i Norge*. Statistisk sentralbyrå.

Hægeland, T., Kirkebøen, L.J., Raaum, O. og Salvanes, K.J. (2005). Familiebakgrunn, skoleressurser og avgangskarakterer i norsk grunnskole. *Utdanning 2005 – deltakelse og kompetanse*. Statistisk sentralbyrå.

KD (2006). Forskrift til opplæringslova. Kunnskapsdepartementet.

KD (2006). *Læreplanverket for Kunnskapsløftet. Midlertidig utgave*. Kunnskapsdepartementet.

KD (2007). St.meld. nr. 16 (2006-2007). *...og ingen sto igjen. Tidlig innsats for livslang læring*. Kunnskapsdepartementet.

KD (2008). St. meld. nr. 31 (2007-2008). *Kvalitet i skolen*. Kunnskapsdepartementet.

Kirke- og undervisningsdepartementet (1957). *Normalplanen for byfolkeskolen* (3. opplag), Oslo: Aschehoug.

Kirke- og undervisningsdepartementet (1974). *Mønsterplan for grunnskolen*. Oslo: Aschehoug.

Kirke- og undervisningsdepartementet (1987). *Mønsterplanen for grunnskolen*. Oslo: Aschehoug.

Kjærnsli, M., Lie, S., Olsen, R.V., og Roe, A. (2007). *Tid for tunge løft. Norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006*. Oslo: Universitetsforlaget.

Kjærnsli, M., Lie, S., Olsen, R.V., Roe, A. og Turmo, A. (2004). *Rett spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003*. Oslo: Universitetsforlaget.

KUF (1996). *Læreplanverket for den 10-årige grunnskolen*. Oslo: Nasjonalt læremiddelsenter.

KUF (1996). St. meld. nr. 47 (1995-1996). *Om elevvurdering, skolebasert vurdering og nasjonalt vurderingssystem*. Oslo: Kirke-, utdannings- og forskningsdepartementet.

KUF (1998). *Elevvurdering. Rettleiing til Læreplanverket for den 10-årige grunnskolen (L97)*. Kirke-, utdannings- og forskningsdepartementet.

Lie, S., Kjærnsli, M., Roe, A. og Turmo, A. (2001). *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv*. Acta Didactica, 4/2001. Institutt for lærerutdannings og skoleutvikling, Universitetet i Oslo.

Lysne, A. (1999): *Karakterer og kompetanse. Stridstema i norsk skolehistorie* (bind 1). Oslo: Anders Lysne, AVA forlag.

Markussen, E. og Sandberg, N. (2004). *Bortvalg og prestasjoner. Om 9798 ungdommer på Østlandet, deres vei gjennom, ut av, eller ut og inn av videregående opplæring, og om deres prestasjoner et år etter avsluttet grunnskole*. NIFU skriftserie 4/2004, Oslo.

Opplæringslova med forskrifter (2001). Oslo: PEDLEX Norsk Skoleinformasjon.

Sadler, R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5 (1), 77-84.

Stewart, D.W., Shamdasani, P.N. og Rook, D.W. (2007). *Focus groups. Theory and practice*. CA: Sage Publications.

Støren, L.A., Helland, H. og Grøgaard, J. (2007). *Og hvem stod igjen...? Sluttrapport fra prosjektet gjennomstrømning i videregående opplæring blant elever som startet i videregående opplæring i årene 1999-2001*. Rapport 14/2007 NIFU STEP, Oslo.

UFD (2004). St.meld. nr. 30 (2003-2004). *Kultur for læring*. Utdannings- og forskningsdepartementet.

Utdanningsdirektoratet (2006). *Utdanningsspeilet. Analyse av grunnskolen og videregående opplæring i Norge*.

Utdanningsdirektoratet (2007). *Vurdering – Et felles løft for bedre vurderingspraksis – en veiledning*.

Vedlegg 1

Oversikt over antall videregående skoler i utprøvingen som deltar med hvilke utdanningsprogram og programområder/fag på vg1 og vg2.

	Utdanningsprogram eller fag	Antall skoler	Programområde vg2 (gjelder for enkelte av skolene)
Yrkesfaglige utdanningsprogram	Norsk Y	4	
	Matematikk P (Y)	4	
	EL	2	Elenergi, Automatisering
	SS	3	Salg og service
	DH	1	Frisør
	BA	1	
	HS	8	Helsearb.fag, Hudpleie, Barne- og ungdomsarb.fag, Helseservice
	TP	4	Produksjons- og industrideknikk
	MK	1	
	RM	1	Kokk- og servitørfag
Studieforberedende utdanningsprogram	Norsk ST	14	
	Matematikk T (ST)	10	
	Matematikk P (ST)	7	
	Tysk	1	
	Engelsk	1	
	Historie	1	
	Samfunnsfag	1	
	Rettslære	1	
	Markedsføring og ledelse	1	
	Idrettsfag	2	
	Musikk, dans og drama	1	Drama