

RAPPORT 21/2009

Hva er viktige kvaliteter ved god sensur?

En kartlegging av bruk av ekstern sensor på lavere grad med fokus på kvalitet, økonomi, vurdering og læring

Nicoline Frølich, Vibeke Opheim, Synnøve Brandt og Tine Sophie Prøitz



© NIFU STEP Norsk institutt for studier av innovasjon, forskning og utdanning
Wergelandsveien 7, 0167 Oslo

Rapport 21/2009
ISBN 978-82-7218-628-4
ISSN 1504-1824

For en presentasjon av NIFU STEPs øvrige publikasjoner, se www.nifustep.no



Norsk institutt for studier av innovasjon, forskning og utdanning
Norwegian Institute for Studies in Innovation, Research and Education
Wergelandsveien 7, 0167 Oslo
Tlf. +47 22 59 51 00 • www.nifustep.no

RAPPORT 21/2009

Nicoline Frølich, Vibeke Opheim, Synnøve Brandt og Tine Sophie Prøitz

Hva er viktige kvaliteter ved god sensur?

En kartlegging av bruk av ekstern sensor på lavere grad med fokus på kvalitet, økonomi, vurdering og læring

Forord

Kvalitetsreformen medførte en endring i sensurpraksis i Norge. Denne rapporten kartlegger dagens sensurpraksis på lavere grad og diskuterer problemstillinger knyttet til forholdet mellom sensorordning og kvalitet, økonomi, læring og vurdering.

Prosjektet er gjennomført på oppdrag for Kunnskapsdepartementet. Vi takker informantene som gav av sin tid og stilte til intervju og alle som har deltatt i innsamlingen av kartleggingsdataene. Rapporten er skrevet av Synnøve Brandt, Nicoline Frølich, Vibeke Opheim og Tine Sophie Prøitz med Nicoline Frølich som prosjektleder. Liv Langfeldt, Bjørn Stensaker og Per Olaf Aamodt har bidratt med kommentarer.

Oslo, mai 2009

Per Hetland
Direktør

Agnete Vabø
Forskningsleder

Innhold

Sammendrag	7
1 Innledning	11
1.1 Tema og problemstillinger	11
1.2 Rammeverk for analysen.....	13
1.2.1 Fire grunnleggende hensyn	13
1.3 Fremgangsmåte og datagrunnlag.....	14
1.3.1 Sensorordninger og spørsmål om kvalitet.....	14
1.3.2 Sensorordninger og spørsmål om kostnader	15
1.3.3 Sensorordninger og spørsmål om ressurser til læring og vurdering	15
1.3.4 Kartlegging av sensorordninger	16
1.3.5 Intervjudata	16
1.3.6 Begrensinger ved undersøkelsen.....	17
1.4 Om rapporten.....	18
2 Sensorordning: kvalitet, kostnader, læring og vurdering	19
2.1 Sensurering og sensorordninger	19
2.2 Ekstern sensurering av eksamen.....	20
2.3 Aspekter ved ordninger for vurdering	21
2.3.1 Sensorordninger og kvalitet	22
2.3.2 Sensorordning og kostnader.....	22
2.3.3 Sensur, læring og vurdering	23
3 Kvaliteter ved sensorordninger.....	25
3.1 Kartlegging av sensurpraksiser	25
3.1.1 Resultater fra en bredt anlagt kartleggingsstudie.....	25
3.1.2 Hvilke ordninger brukes mest?	27
3.1.3 Sensorordning en del av kvalitetssikringssystemene?	30
3.2 Vurderinger av kvaliteten på sensuren	33
3.2.1 Metodiske hensyn	33
3.2.2 Praktiske hensyn	35
3.2.3 Formelle hensyn.....	35
3.2.4 Kommunikasjonsmessige hensyn	35
3.3 Konkluderende betraktninger om sensorordning og kvalitet	35
4 Sensurpraksis og kostnader.....	37
4.1 Er ekstern sensur kostnadskrevende?	37
4.2 Hvilke kostnader skal tas hensyn til?	38
4.3 Henger erfaringene med kostnader sammen med valg av sensorordning?	39
4.4 Tidsbruk til eksamensarbeid.....	39
4.5 Konkluderende betraktninger om sensurpraksis og kostnader	40

5	Sensur, læring og vurdering	42
5.1	Sensorordning og ressurser til læring og vurdering	42
5.2	Ekstern sensur og typer vurderingsformer	43
5.2.1	Begrunnelser for sensurpraksis og forholdet mellom læring og vurdering .	43
5.3	Konkluderende betraktninger om forholdet mellom sensur, læring og vurdering	44
6	Konklusjon	46
	Tabell- og figuroversikt	49
	Litteratur	50
	Vedlegg 1 Kartlegging av sensorordninger	51
	Vedlegg 2 Intervjuguide	54
	Vedlegg 3 Oversikt over informanter	56

Sammendrag

Kvalitetsreformen og implikasjoner for sensorordningene

Før Kvalitetsreformen krevde loven (minst) to sensorer, hvorav (minst) en ekstern ved avsluttende eksamen ved universiteter og høyskoler i Norge. I den nye universitets- og høyskoleloven av 2005 ble kravet om ekstern sensor opprettholdt på selvstendige arbeider på høyere grad. For øvrig ble kravet om ekstern sensor endret til et lovfestet krav om ekstern evaluering av vurderingene eller vurderingsformene. Lærestedene fikk med dette større frihet til å utforme ordninger for sensur på lavere grad.

Problemstillinger og datagrunnlaget

Rapporten presenterer resultatene av en kartlegging av sensurpraksis ved 31 offentlige læresteder og diskuterer sensurpraksis i forhold til spørsmål om kvalitet på sensuren, ressursbruk og forholdet mellom læring og vurdering.

Datagrunnlaget består av en spørreundersøkelse om sensurpraksis sendt samtlige universiteter, vitenskapelige høyskoler og statlige høyskoler i mars 2009. Trettien læresteder besvarte undersøkelsen. Samtidig ble det gjennomført 30 kvalitative intervju blant studiedirektører, studieledere og tilsynssensorer ved fem fagområder ved 10 læresteder.

Sensurering og sensorordninger

Eksamenssensur i høyere utdanning betyr å vurdere en eksamensprestasjon, uttrykke dette i en karakter, kommunisere vurderingen til eksamenskandidaten, samt å forholde seg til eksamenskandidatens reaksjon på vurderingen. Å sensurere betegner således en rekke profesjonaliserte handlinger knyttet til å vurdere en kandidats skriftlig, muntlig eller praktiske prestasjon ved eksaminering.

Å sensurere en eksamensprestasjon innebærer å balansere en rekke formelle, kommunikasjonsmessige, metodiske og praktiske hensyn. Det handler om spørsmål om hva som skal sensureres og hvordan, samt hvilke hensyn som må tas i gjennomføringen av sensuren. Spørsmål om hvem som sensurerer er knyttet til skillet mellom ekstern og intern sensor.

I høyere utdanning kan det skilles mellom formative og summative vurderingsformer. Summativ vurdering innebære i hovedsak prøving med sikte på karaktersetting eller sertifisering, mens formålet med formativ vurdering er i hovedsak å gi tilbakemeldinger som kan modifisere arbeidet med læring. Ved eksamen står altså sertifiseringsformålet sterkt. Imidlertid har kravet om andre vurderingsformer preget diskusjonen om og utviklingen av nye studieopplegg i høyere utdanning i Norge. Bruk av eksempelvis mappevurdering, åpenbok/kilde oppgaver, gruppe- og hjemmeoppgaver er metoder som

betraktes som mer læringsorienterte enn tradisjonell eksamen (avsluttende individuell skoleeksamen). Dette er arbeidsmåter som kan brukes i vurderingsarbeidet både som underveis og som avsluttende prøving i studieløpene.

Hovedfunn

På lavere grad i norsk høyere utdanning er det fortsatt betydelig innslag av ekstern sensur. Sammenlignet med situasjonen før Kvalitetsreformen er praksis likevel endret. Variasjonen i sensorpraksis er betydelig og varierer innen for enkeltstudium, institutt eller avdeling, faggruppe, mellom faggrupper og mellom læresteder. Den mest brukte sensorordningen er en kombinasjon av intern sensur med et element av ekstern sensur. Denne kombinasjonen organiseres på ulike måter, også i kombinasjon med tilsynssensor. Tilsynssensor innebærer at ekstern evaluering av vurderingene, vurderingsordningene og sammenhengen med studieopplegget er formalisert i en egen funksjon. Bruk av tilsynssensor ser ut til å ha økt noe sammenlignet med situasjonen i 2005. Intervjumaterialet viser at tilsynssensorfunksjonen har forskjellige betegnelser ved de ulike lærestedene. Lærestedene har etablert ordninger som benevnes programsensor, bedømmersensor, eksamenskomiteer og nivåkontrollør. Hensiktene med disse ordningene er å bidra til kvalitetssikringen av sensuren.

Den sentrale begrunnelsen for endringer av sensurpraksis er økonomisk, og ikke kvalitetsmessig. Det er avgjørende for fagmiljøene å sikre kvaliteten på sensuren og der man har gått bort fra tradisjonell sensur er man opptatt av å sikre kvaliteten på vurderingene gjennom å etablere ulike kvalitetssikringsmekanismer. Enkelte mener at kvaliteten på sensuren har blitt styrket gjennom de nye ordningene og hevder at det ikke ligger en automatisk garanti for kvalitet i den tradisjonelle sensorordningen. Skillet mellom enesensur og det å sensurere i fellesskap (to eller flere personer) synes vel så viktig som skillet mellom ekstern og intern sensur. Enesensur medfører et stort ansvar som vitenskapelig ansatte synes seg svært bevisst. Enkelte fagmiljø går nå bort fra denne sensurpraksisen.

Ved vurdering av økonomiske konsekvenser av sensurpraksis legges hovedsakelig vekt på de direkte kostnadene knyttet til ekstern sensur. Kostnader knyttet til ulike interne ordninger trekkes i mindre grad med i vurderingene.

Det er ikke en entydig sammenheng mellom sensorordning og ressurser til læring eller ressurser til vurdering, dette varierer mellom fagmiljøene. Således kan det ikke konkluderes klart i spørsmålet om ekstern sensur flytter ressurser fra læring til vurdering.

Det mest utfordrende spørsmålet er hva variasjoner i sensurpraksis betyr for kvaliteten på sensuren. Vi har ikke kunnet undersøke sammenhengen mellom karaktersetning og sensorordning på en måte som gir grunnlag for robuste konklusjoner. I fagmiljøene er det imidlertid en utbredt oppfatning at ekstern sensur på best måte sikrer kvaliteten på

sensuren. Det betyr imidlertid ikke at de fagmiljøene som velger andre løsninger ønsker seg tilbake til et system med tradisjonell sensur.

Begrensinger ved undersøkelsen

Data ble samlet inn over en kort tidsperiode, som umuliggjorde en trinnvis sammenligning av informasjon fra de ulike datakildene underveis i prosessen. Det var heller ikke mulig å gjøre en grundig pilotundersøkelse for å vurdere hvordan kategoriseringen i kartleggingen dekker det empiriske mangfoldet.

Gruppen av informanter dekker fem svært ulike fagområder og det er få informanter innenfor hvert område. Materialet ikke gir grunnlag for robuste konklusjoner om forskjeller mellom fagområder og mellom læresteder.

1 Innledning

1.1 Tema og problemstillinger

Kvalitetsreformen ga lærestedene for høyere utdanning større frihet til å avgjøre hvilke sensorordninger som skulle benyttes (Bernt, 2006). Konsekvensene av dette handlingsrommet er foreløpig lite dokumentert. Før reformen var en slik problematikk mindre aktuell. Sensurpraksis var lovfestet: Eksamenskommissjoner bestod av to (eller tre) sensorer, hvorav minst en ekstern (Bernt, 1996): Variasjoner i bruken av sensorordningene er større etter reformen (Solum, 2005).

Vurderinger kan kategoriseres som formative eller summative, sluttvurdering eller underveisvurdering. I denne rapporten benyttes begrepsparet sluttvurdering og underveisvurdering som begge kan ha summative og formative elementer. Summativ vurdering innebærer vurdering av kandidatens arbeider med sikte på karaktersetting eller sertifisering (Black & Wiliam, 2003; Bloom, Hasting, & Madaus, 1971). Formative vurderinger tilslutt eller underveis vil si vurdering med sikte på å gi tilbakemelding som justerer læringen (Black et al., 2003).

Som nå, var norsk høyere utdanning preget av heterogenitet i vurderingsmetodene, også før Kvalitetsreformen ble iverksatt (Vabø, 2000: 7). Praksis varierte mellom profesjons- og disiplinbaserte fag og mellom lavere og høyere grad. Den viktigste formen for vurdering og tilbakemelding til studentene særlig ved lavere gradstudier i humanistiske og samfunnsvitenskapelige fag samt ved jus var tradisjonell eksamen (avsluttende individuell skoleeksamen). I tillegg ble vurderingsformer som hjemmeeksamen og semesteroppgave benyttet (Vabø, 2000: 7).

Et utgangspunkt for endringene av sensorordninger i Kvalitetsreformen var kritikk av det norske systemet for høyere utdanning fremmet av OECD (OECD, 1998). Sett i lys av den økte studenttilstrømningen, ble Norge ble kritisert for i for stor grad å bruke tradisjonell eksamen som eneste vurdering av studentenes arbeid (Dysthe, Raaheim, Lima, & Bygstad, 2006: 48; NOU, 2000: 14: 127). OECD konkluderte med at Norge burde fokusere mer på undervisning og læring (fremfor sluttvurdering) (Dysthe et al., 2006: 48). En problemforståelse bak Kvalitetsreformen var at ”en tradisjonell tilnærming til eksamen forstått som noe som skal bestås, fremfor å tilby god undervisning og et rikt tilfang av læringserfaringer” (NOU, 2000: 14: 127) var rådende.

Sensorordninger er i endring i norsk høyere utdanning og i andre land (Stensaker, Brandt, & Solum, 2008). En grunn synes å være endringer i forståelsen av hvordan kvaliteten på sensuren kan sikres, og endringer i oppfatninger om kvalitetssikring i høyere utdanning mer generelt. Dette kan sees i sammenheng med Bologna-prosessens fokus på standardisering av kvalitetssikring. Utviklingen er knyttet til oppretting av eksterne systemer for kvalitetssikring av nasjonale institusjoner. I mange land har dette i stor grad

foregått på siden av allerede eksisterende og tradisjonsrike ordninger for ekstern kvalitetssikring av eksamen (Stensaker et al., 2008).

Vurdering av eksamensprestasjoner kan betraktes som en av de mest sentrale kvalitetssikringsprosedyrene i høyere utdanning. Det er imidlertid stor variasjon i hvordan dette gjøres i ulike land. Norge, sammen med en del andre land, har systemer der vurdering og karaktergivning involverer eksterne krefter. Dette har også flere av Commonwealth landene, Storbritannia og Danmark. Land som opprettholder bruk av ekstern sensur har ofte lange tradisjoner for dette, og disse ordningene synes fortsatt å ha høy anseelse her (Stensaker et al., 2008). Det finnes nasjonale utdanningssystemer der vurdering og karaktergivning gjøres helt og holdent (internt) ved den enkelte institusjon. Dette er sannsynligvis den mest vanlige ordningen ved institusjoner i USA og i deler av Europa (Stensaker et al., 2008). For land i Norden gjelder dette også Finland og Sverige (Vabø, 2000).

I Norge har lærestedenes sensureringspraksis vært et tilbakevendende tema i etterkant av Kvalitetsreformen. Debatten har dreid seg om spørsmål om rettsikkerhet og likebehandling, og om bekymringer for kvalitetssenkning. Vi skal ikke direkte adressere spørsmålet om rettsikkerhet og likebehandling, men fokusere på spørsmål knyttet til bruken av ekstern sensor og forhold som kvalitet, økonomi og vurdering og læring. I Kunnskapsdepartementets oppdrag for dette prosjektet stilles en rekke spørsmål:

- Vil økt andel tradisjonell ekstern sensur føre til bedre kvalitet på sensuren?
- Hva er de økonomiske konsekvensene ved økt bruk av ekstern sensor?
- Vil økt bruk av ekstern sensur flytte ressurser fra læring til vurdering?
- Vil en gjeninnføring av tradisjonell ekstern sensur føre til begrensninger på hvilke typer vurderingsformer institusjonene vil benytte?

Flere av disse problemstillingene er hypotetiske spørsmål som ikke kan besvares empirisk. Likevel, gjennom å undersøke dagens praksis og hvordan nøkkelaktører i feltet begrunner sine valg, kan vi gi en fremstilling av rådende praksis og oppfatninger knyttet til de tre sentrale spørsmålene vi behandler her. Disse spørsmålene er:

1. Er det en sammenheng mellom sensorordning og kvalitet på sensuren?
2. Er det en sammenheng mellom sensorordninger og kostnader?
3. Er det en sammenheng mellom sensorordning og læring og vurdering?

For å analysen av disse problemstillingene benyttes en kvalitativ og en kvantitativ tilnærming, med følgende hovedelementer:

- En avklaring av begrepene sensur og eksamen
- En kvalitativ beskrivelse av statlige læresteders bruk av ekstern sensor på lavere grad

- Intervjuer med studiedirektører, studieledere og tilsynssensorer på et utvalg læresteder og fag med sikte på å kartlegge deres vurderinger av spørsmålene om sensorordning, kvalitet, kostnader, læring og vurdering.

1.2 Rammeverk for analysen

1.2.1 Fire grunnleggende hensyn

Å etablere gode ordninger for sensur handler om å balansere en rekke hensyn og prioriteringer. Basert på Lauvås et al. (2000) har vi utviklet et rammeverk som gjør det mulig å diskutere sensur i forhold til kvalitet, økonomi, læring og vurdering (se kapittel 2). Dette rammeverket vektlegger fire hovedhensyn: formelle forhold, kommunikasjonsmessige forhold, metodiske aspekter og praktiske forhold knyttet til sensorordninger.

Formelle forhold gjelder hensynet til rettssikkerhet, som betyr at både studenters og bedømmeres juridiske rettigheter ivaretas for å forebygge og minimalisere vilkårlighet og overgrep. Andre formelle aspekter gjelder offentlighet ved eksamen og sensur, som betyr at det sikres innsyn og åpenhet i evalueringen slik at minst mulig foregår bak lukkede dører og er unndratt innsyn.

Kommunikasjonsmessige aspekter ved sensorordninger er knyttet til legitimitet, det vil si at studenter, lærere, administrasjonen, fremtidige arbeidsgivere og samfunnet generelt, aksepterer evalueringen som rimelig og rettferdig. Kommunikasjonsmessige aspekter gjelder også problemstillinger knyttet til dialog om sensurpraksis; hva som er relevant kunnskap og relevante kriterier for vurdering.

Spørsmål om kvalitet og sensurpraksis belyser vi gjennom *metodiske aspekter* ved sensuren som innebærer spørsmål om:

- Standardisering: Evalueringssituasjonen er så lik som mulig for alle studenter (tid, oppgaver, utstyr, osv.).
- Objektivitet: Vurdering av det som skal evalueres (kunnskap, ferdigheter, osv.) og ikke noe annet (personlig vennskap, slektskap, arbeidsrelasjoner, osv.) sikres vanligvis gjennom anonymisering.
- Reliabilitet: Tiltak for å sikre presisjon i evalueringen (med hensyn til evalueringsskala) og mellom vurderingene til ulike personer som evaluerer.
- Ytre (økologisk) validitet: Det som blir evaluert, er relevant i forhold til utdanningens totale formål; den kunnskap og kompetanse som utdanningen skal gi.
- Målrelevans (indre validitet): Evalueringen er designet slik at den dekker det området som skal evalueres, både med hensyn til innholdsområder og nivå. Bedømmerne skal bruke de samme kriterier, tolkninger og forståelse.

Spørsmålet om økonomiske konsekvenser knyttet til ulike sensorordninger diskuterer vi ved å se på *praktiske hensyn* knyttet til ulike sensorordninger (se kapittel 3). Praktiske hensyn knyttet til sensorordninger handler om hvorvidt det er optimal balanse mellom ressursbruk (budsjett, arbeidsmengde, tilgjengelige evaluatorene, tid til å utføre evalueringsarbeidet, osv.) og omfanget på modulen/kurset (vekt tall, innholdets karakter, osv.).

Vår problemstilling handler primært om metodiske og praktiske sider ved sensur. Vi skal til en viss grad også berøre formelle og kommunikasjonsmessige sider ved forskjellige sensorordninger.

1.3 Fremgangsmåte og datagrunnlag

For å svare på problemstillingene benyttes en spørreundersøkelse til lærestedene (vedlegg 1), samt kvalitative intervju med nøkkelinformanter. I det følgende redegjør vi for hvordan datakildene kombineres for å besvare problemstillingene, samt for vurderinger i forhold til hvilke konklusjoner som kan trekkes basert på hvilke data.

1.3.1 Sensorordninger og spørsmål om kvalitet

Den første problemstillingen relaterer seg til sammenhengen mellom måten sensuren gjennomføres på og kvaliteten på sensuren (i.e. metodiske aspekter). Her benyttes flere datakilder. For det første gjennomfører vi en bredt anlagt kartlegging av sensureringspraksis ved alle universiteter, vitenskapelige høyskoler og statlige høyskoler (se kapittel 3). Det kan tenkes at ulike sensorordninger er basert på ulike avveininger mellom de fire aspektene ved sensur som vi diskuterte over. I kartleggingen ser vi både på variasjoner i praksis med sensur og på hvilken praksis som er mest utbredt. Tilslutt, har vi gjennomført intervju med et utvalg tilsynssensorer, studieledere og studiedirektører om deres vurderinger og erfaringer knyttet til forholdet mellom sensorordning og kvaliteten på sensuren.

Vi har ikke funnet at sammenhengen mellom karakterdata og sensureringspraksis på lavere grads studier er mulig å undersøke på en faglig forsvarlig måte. En lang rekke andre forhold, i tillegg til sensurering, vil ha betydning for karaktergivning ved hvert enkelt lærested og innenfor de ulike fagområdene. Et er lærestedenes inntakskrav; læresteder med høye inntakskrav vil kunne forvente å ha studenter med relativt bedre prestasjoner (i form av bedre karaktersnitt) enn læresteder med lavere inntakskrav, forutsatt likt faglig nivå mellom lærestedene. I tillegg til studentenes faglige nivå og andre kjennetegn ved studentene, kan en rekke andre forhold ha betydning for karakternivået. Eksempelvis kan lærestedenes organisering av undervisningen, kvaliteten på denne, samt grad av oppfølging av studentene, ha betydning. Vi har heller ikke funnet grupper med kombinasjoner ulike sensorordning og antall karakterer som er store nok til statistiske undersøkelser.

Kartleggingen gir en oversikt over lærestedenes bruk av ekstern sensor og tilsynssensor innenfor de ulike fagområdene. Vi undersøker også bruken av ekstern sensor sammenlignet med bruken av intern sensor.

I tråd med den dokumenterte nasjonale og internasjonale utviklingen på området er det interessant å se utviklingen av ordninger for sensur i relasjon til utviklingen av lærestedenes kvalitetssikringssystemer. Den gjennomførte kartleggingen omhandler også i hvilken grad de ulike sensorordningene på lavere grad er knyttet opp til lærestedenes kvalitetssikringssystem.

1.3.2 Sensorordninger og spørsmål om kostnader

Vi diskuterer kostnader knyttet til ulike sensorordninger, spesielt sentralt er spørsmålet om de økonomiske konsekvensene ved økt bruk av ekstern sensor (se kapittel 4). For å belyse denne problemstillingen har vi intervjuet studieledere i et utvalg av fag. Det er dem som i utgangspunktet kan ha detaljert kunnskap om de økonomiske konsekvensene av økt bruk av ekstern sensor. I tillegg er tilsynssensorer en sentral kilde til informasjon om det eventuelt er en sammenheng mellom ordninger for sensurering og kostnadsnivå. Datamaterialet relateres til tidligere undersøkelser av vitenskapelige ansattes tidsbruk på eksamensarbeid (Frølich & Waagene, 2008; Smeby, 1993).

1.3.3 Sensorordninger og spørsmål om ressurser til læring og vurdering

Den tredje problemstillingen gjelder sammenhengen mellom sensorordning og læring og vurdering (se kapittel 5). Vil økt bruk av ekstern sensur flytte ressurser fra læring til vurdering? Her kombinerer vi intervjuer med studielederne som kan bidra med faktakunnskap om dagens situasjon, og de kostnader som ulike praksiser innebærer, og intervju med tilsynssensorene som kan vurdere kostnadsstipuleringene. Begge grupper av informanter kan bidra med refleksjoner knyttet til det antatte skillet mellom "læring" og "vurdering" og sensurens rolle i forhold til begge disse.

Til slutt drøfter vi spørsmålet om en gjeninnføring av tradisjonell ekstern sensur vil føre til begrensinger i hvilke vurderingsformer lærestedene vil benytte (se kapittel 5). For å diskutere dette spørsmålet sammenligner vi først nåværende sensurpraksis med situasjonen i 2005 som ble undersøkt av Solum (2005). Solum fant variasjoner i sensurpraksis, lærestedene hadde på det tidspunktet tilpasset seg valgfriheten når det gjaldt sensorordninger og dermed økt antallet vurderingsformer sammenlignet med situasjonen før lovendringen. Spørsmålet er hvilken praksis som i dag er rådende. Deretter bruker vi intervju med studiedirektørene og studieledere for å belyse deres begrunnelser for eventuelle endring av praksis.

1.3.4 Kartlegging av sensorordninger

Vi har gjennomført en kartlegging av bruken av ekstern sensor for alle statlige høyskoler, vitenskapelige høyskoler og universitet. Kartleggingen foregikk på avdelings-/fakultetsnivå og identifiserte:

- hvilke *typer* av sensorordninger som eksisterer på lavere grads studier
- hvilket *omfang* de ulike typene sensorordninger har
- i hvilken grad sensorordningene er *formalisert* i forskrifter og eventuelt koplet til institusjonenes kvalitetssikringssystem

For å ta hensyn til at det høyst sannsynlig eksisterer et mangfold av systemer, utarbeidet vi en klassifisering som lettere identifiserer utviklingstrekk og tendenser når det gjelder sensurpraksis. Klassifiseringen skiller mellom fire hovedtyper av sensurpraksis fordelt på fem hovedgrupper av fag innenfor de ulike typene av læresteder (mer om dette i kapittel 2). I tillegg identifiseres i hvilken grad sensorordningen er koblet til lærestedets kvalitetssikringssystem (se kapittel 3). Dataene i kartleggingen ble samlet inn via en henvendelse til alle studiedirektører og tilsvarende ved universitetene, de vitenskapelige høyskolene og de statlige høyskolene (Vedlegg 1).

1.3.5 Intervjudata

Den andre datakilden er kvalitative intervju med nøkkelaktører: studiedirektører, studieledere og tilsynssensorer. Hensikten med intervjuene var å belyse aktørenes begrunnelser for de ulike sensorordningene og deres vurderinger av problemstillingene knyttet til sammenhengen mellom sensorordning og kvalitet; kostnader og læring og vurdering (intervjuguide vedlagt bak i rapporten). Vi har gjennomført til sammen 30 intervju (10 studiedirektører, 15 studieledere og 5 tilsynssensorer). De ulike gruppene av informanter har noe ulik funksjon i forhold til sensur. Mens studiedirektørene kan gi informasjon om sensurpraksis og vurderinger knyttet til denne for lærestedene som helhet, kan studieledere og tilsynssensorer belyse problemstillingene ut fra sitt faglige ståsted.

Informantene dekker ulike fag og ulike typer læresteder (vedlegg 3). De samme 10 lærestedene som Solum (2005) benyttet ble undersøkt for å kunne belyse sentrale utviklingstendenser om sensurpraksis. Utvalget av læresteder består av følgende læresteder:

- Universitetet i Bergen (UiB)
- Universitetet i Stavanger (UiS)
- Universitetet i Oslo (UiO)
- Norges Handelshøyskole (NHH)
- Høgskolen i Lillehammer (HiL)
- Høgskolen i Bodø (HBO)
- Høgskolen i Sogn og Fjordane (HSF)

- Høgskolen i Oslo (HiO)
- Høgskolen i Nesna (HiNe)
- Høgskolen i Østfold (HiØ)

Blant de ti lærestedene identifiserte vi studieprogram på lavere grad innenfor de samme fem fagområdene som ble benyttet i den nevnte kartleggingen (naturvitenskapelige og teknologiske fag; samfunnsfag; humanistiske fag; juridiske fag og pedagogiske fag). Vi kontaktet til sammen 30 studieledere. For å identifisere tilsynssensorer som kunne inngå i studien, ble studielederne bedt om å formidle kontaktinformasjon om tilsynssensorer i sitt studium. I tillegg sendte vi henvendelse til fagområder som oppgav at tilsynssensor ble benyttet for å komme i kontakt med disse. Det viste seg at langt de fleste fagene vi hadde valgt ut ikke hadde etablert en tilsynssensorordning, derfor ble kun fem tilsynssensorer intervjuet.

1.3.6 Begrensinger ved undersøkelsen

Undersøkelsen må leses som en tentativ strukturering av et svært komplekst felt. Datagrunnlaget består av en spørreundersøkelse blant alle offentlige læresteder i Norge. Undersøkelsen ble sendt til 37 studiedirektører eller tilsvarende, svar kom inn fra 31 læresteder. Svarene fordeler seg nokså jevnt mellom hovedtypene læresteder.

Lærestedene organiserte selv fremgangsmåten for å samle inn svar på spørsmålene. Det betyr at det er varierende kvalitet på svarene og at vi ikke kjenner detaljene i fremgangsmåten som ble benyttet. De fleste besvarte henvendelsen før fristen og uten å rette spørsmål til oss. Inndelingen i faggrupper tilsvarende ikke direkte fakulteter/avdelinger ved samtlige læresteder, dette bidro til noe vanskeligheter i utfyllingen av skjemaet.

Data ble samlet inn over en kort tidsperiode, som umuliggjorde en trinnvis sammenligning av informasjon fra de ulike datakildene underveis i prosessen. Det var heller ikke mulig å gjøre en grundig pilotundersøkelse for å vurdere hvordan kategoriseringen i kartleggingen dekker det empiriske mangfoldet.

Gruppen av informanter dekker fem svært ulike fagområder og det er få informanter innenfor hvert område. Informanten kommer fra 10 av 37 læresteder. Det betyr at intervjumaterialet ikke gir grunnlag for robuste konklusjoner om forskjeller mellom fagområder og mellom læresteder.

1.4 Om rapporten

I neste kapittel ser vi på bakgrunn og problemstillinger knyttet til sensur og sensorordninger. Det omfatter en analytisk gjennomgang av ulike formål og aspekter ved ordninger for vurdering av eksamensresultater. Deretter presenterer vi de empiriske resultatene fra analysene av data fra kartleggingen av bruken av ekstern sensor ved offentlige læresteder i Norge, samt intervjuer av tilsynssensorer, studiedirektører og studieledere. De empiriske analysene er delt inn i tre kapitler. Kapittel 3 omhandler forholdet mellom sensorordninger og oppfatninger om kvalitet. I kapittel 4 diskuteres sammenhengen mellom sensurpraksis og kostnader, mens tema i kapittel 5 er sensur, læring og vurdering. I avslutningskapittelet diskuteres hovedfunnene fra de empiriske kapitlene opp mot de tre sentrale problemstillingene vi behandler i denne rapporten: 1) sensorordning og kvalitet, 2) sensorordninger og kostnader, og 3) sensorordning og læring og vurdering.

2 Sensorordning: kvalitet, kostnader, læring og vurdering

Hensikten med dette kapittelet er å etablere en bredere analytisk ramme for å kunne diskutere spørsmål om kvalitet, økonomi, læring og vurdering i tilknytning til ulike sensorordninger. Vi starter gjennomgangen med å avklare hva sensurering er.

2.1 Sensurering og sensorordninger

Sensurering er en del av de faglig-administrative arbeidsoppgavene som utføres ved lærestedene. Vitenskapelige ansatte utfører en lang rekke arbeidsoppgaver knyttet til forskning og undervisning, eksamensarbeid og sensur er blant disse (Frølich et al., 2008).

Å sensurere betegner en rekke profesjonaliserte handlinger knyttet til å vurdere en kandidats prestasjon ved eksaminering, det være seg skriftlig, muntlig eller ved løsning av en praktisk oppgave. I høyere utdanning innebærer eksamenssensur å vurdere en (eksamens)prestasjon, uttrykke denne vurderingen i en karakter, kommunisere vurderingen til eksamenskandidaten, samt å forholde seg til eksamenskandidatens reaksjon på vurderingen.

Sensorordninger er en samlebetegnelse på hvordan sensur organiseres ved eksamen. Sensorordninger handler om hva som skal sensureres, hvem og hvor mange som skal sensurere, og om hvordan sensuren gjennomføres. Hva som skal sensureres gjelder spørsmål knyttet til hvordan og hvilke hensyn som må tas i gjennomføringen av arbeidet. Hvem som sensurerer viser ofte til skillet mellom ekstern og intern sensor.

Sensurarbeid kan gjøres av en person alene eller av flere sensorer i samråd. Når sensuren gjøres av *en* person er dette med utgangspunkt i dennes formelle kvalifikasjoner og erfaring. Argumenter for bruk av en sensor er at denne gjennom sine formelle og uformelle kvalifikasjoner har integritet og kunnskap nok til å gjøre de riktige vurderingene alene. I tilfeller der sensureringsarbeidet gjøres av flere sensorer i samarbeid er dette med utgangspunkt i både de individuelle og deres felles formelle kvalifikasjoner og erfaring. Argumenter for sensur der flere sensorer inngår er at dette vil kunne gi en kvalitativt bedre sensur i forhold til prinsipper om standardisering, legitimitet, objektivitet og rettsikkerhet enn en enkelt sensor vil alene. Dessuten vil sensurering i samarbeid innebære at vurderingsarbeidet læres, deles, kalibreres og revurderes i kollegiale fagfellesskap.

Når flere fagfeller sensurerer sammen innebærer det en innebygget gjensidig kontroll av vurderingen om hva som er en god og en dårlig prestasjon. Slike systemer for sensur innehar også klare elementer av sosialisering av sensorer opp mot sensorrollen innenfor et fagområde. Bruk av flere sensorer kan gjøres på flere måter, eksempelvis ved at flere institusjonsinterne vitenskapelige ansatte vurderer resultatene av en eksaminering.

Tilsvarende ved bruk av en kombinasjon av interne og eksterne sensorer. Ordninger med ekstern sensur kan organiseres på flere måter. I Norge var det frem til Kvalitetsreformen lovfestet at alle eksamenskomisjoner skulle bestå av (minst) to sensorer, hvorav en ekstern (Bernt, 1996: 138). Etter reformen kan lærestedene selv i større grad velge sensorordning på lavere grad, men fortsatt skal det være et element av ekstern sensur også på lavere grad. Formelt sett skal det være ekstern sensur av vurderingene eller vurderingsformene (Bernt, 2006: 63)¹. Dette kan organiseres på forskjellige måter, blant annet i form av tilsynssensor eller tradisjonell sensur på noen av eksamenene² (nærmere om sensurpraksis i kapittel 3). Tilsynssensorfunksjonen er en alternativ måte å sikre et eksternt blikk på vurderingene og vurderingsformene, og også en vurdering av studieopplegget mer generelt.

2.2 Ekstern sensurering av eksamen

Det er tre sentrale formål med ekstern sensur:

1. Ekstern sensur skal sikre at karakterer som gis ved ulike institusjoner er sammenliknbare i forhold til nasjonale (akademiske) standarder.
2. Studenter skal sikres rettferdig vurdering ved å skape distanse mellom dem som underviser og dem som vurderer prestasjoner.
3. Ekstern sensurering kan anses som et hensiktsmessig virkemiddel for kontroll og kvalitetssikring gjennom det innsyn og transparens systemet åpner for (Stensaker et al., 2008).

De to første formålene er direkte knyttet til sensurarbeidet, mens den tredje er av mer indirekte karakter. Det tredje formålet kan sees i lys av de siste års fokus på ”accountability”.

I høyere utdanning finnes det mange interessenter med ansvar for situasjonen ved den enkelte institusjon, studenter, ansatte, administratorer, styremedlemmer og lovgivere har alle en særlig rolle (Aamodt, Prøitz, Hovdhaugen, & Stensaker, 2007). Studiestedet har en utdanningsmessig, økonomisk, juridisk og moralsk forpliktelse til å være transparent – og å demonstrere offentlig for seg selv og omverdenen at løftene som er gitt til studenter, foreldre, finansører og samfunnet som helhet blir holdt (Hersh, 2004).

Kritikere av systemer for ekstern sensurering hevder at det ikke er mulig å oppnå sammenliknbarhet for karakterer, spesielt i store land (Nilsson & Näslund, 1997). De stiller også spørsmål ved om det finnes en bestemt akademisk standard og mulighetene til å vurdere dette i det hele tatt (Jacobsen & Lauvås, 2001). Andre igjen er bekymret for

¹ Kunnskapsdepartementet presiserte i brev av 10.7.2007 at ”institusjonene har [derfor] ikke anledning til kun å ta et tilfeldig utvalg av kandidater som vurderes og får karakterer fastsatt av ekstern sensor” og videre at ”det avgjørende er at alle studenter på en eksamen eller et emne blir gjort gjenstand for ekstern sensur i samme omfang”.

² Eksempelvis ved muntlig eksamen eller på BA oppgaven

negative konsekvenser ved systemer der institusjonene for høyere utdanning organiserer ekstern sensur selv. Antakelsen er at slike ordninger står i fare for å etablere praksiser der kjennskap og vennskap mellom sensorer ved ulike institusjoner vil føre til for ”snill” sensur (Jacobsen et al., 2001). Systemene er for lukkede og for lite informasjon når utenomverden. Ordninger med ekstern sensur er i for stor grad orientert mot å opprettholde akademiske standarder innenfor disiplinene slik at systemene kan bli konserverende og således lite åpne for innovasjon og kreativitet. Dette er argumenter som ofte går hånd i hånd med argumenter om at ekstern sensur er for kostbart og tidkrevende (Silver, Stennet, & Williams, 1995).

Til tross for at ekstern sensur i hovedsak er utformet for kontroll, så bidrar systemene til spredning av god praksis og styrker institusjonell læring mellom institusjoner for høyere utdanning (Stensaker et al., 2008: 213). En studie fra Storbritannia viser at ideer blir utvekslet, sensorenes egen kompetanse blir videreutviklet og motivasjonen for å være ekstern sensor er høy som følge av dette (Hannan & Silver, 2006).

Vi skal nå se på ulike aspekter ved sensorordninger.

2.3 Aspekter ved ordninger for vurdering

For å kunne diskutere ulike sensorordninger tar vi utgangspunkt i at vurderingsordninger for studenters prestasjoner ved eksamen står i forhold til ulike hensyn. Vi bygger på Lauvås et al. (2000) som diskuterer en lang rekke ulike aspekter ved eksamensordninger når vi argumenterer for at sentrale aspekter ved sensorordninger er formelle, kommunikasjonsmessige, metodiske og praktiske. Lauvås (2000) peker på at det er vanskelig å ivareta alle hensynene like godt og at det gjøres prioriteringer mellom dem når et system etableres. Sensorordninger er forankret i utdanningssystemenes tradisjoner, historie og institusjonelle strukturer. Derfor tas de gjerne for gitt og diskuteres sjelden.

Sensorordninger må ivareta formelle aspekter knyttet til rettssikkerhet og offentlighet. Både studentenes og bedømmernes juridiske rettigheter må sikres for å forebygge og minimalisere vilkårlighet og overgrep. Videre må sensorordningene utformes slik at det sikres innsyn og åpenhet i evalueringen slik at minst mulig foregår bak lukkede dører og er unndratt innsyn. Sensorordninger må også sikre konfidensialitet.

Kommunikasjonsmessige aspekter ved sensorordninger er knyttet til ordningenes legitimitet og hvordan dialog ivaretas. Sensorordningers legitimitet innebærer at studenter, lærere, administrasjonen og potensielle arbeidsgivere aksepterer evalueringen som rimelig og rettferdig. De kommunikasjonsmessige aspektene innebærer også at evalueringen legger opp til dialog om hva som er relevant kunnskap og om kriteriene som evalueringen skjer ut fra.

Metodiske aspekter ved sensorordninger er knyttet til *standardisering*, det vil si at evalueringssituasjonen er så lik som mulig for alle studenter både når det gjelder tiden som er til disposisjon for å løse oppgaven, utforming av oppgaver og tilgjengelig utstyr. Videre må *objektivitet* ivaretas. Kun kunnskap og ferdigheter skal evalueres og ikke andre forhold eksempelvis personlig vennskap, slektskap og arbeidsrelasjoner. Anonymisering brukes ofte for å få til dette. *Reliabilitet* er tredje side ved metodiske hensyn sensorordninger bør vurderes i forhold til. Det settes inn tiltak for å sikre presisjon i evalueringen med hensyn til evalueringsskalaen og mellom vurderingene til de ulike personene som evaluerer. *Validiteten* ved sensorordningen må også vurderes. Dette gjelder den ytre validiteten, at det som blir evaluert er relevant i forhold til utdanningens totale formål det vil si den kunnskap og kompetanse som utdanningen skal gi. Det gjelder også den indre validiteten (målerelevansen), at evalueringen er designet slik at den dekker det området som skal evalueres. Dessuten må bedømmerne bruke de samme kriterier, tolkninger og forståelse.

Praktiske aspekter ved sensorordninger innebærer at ressursbruken knyttet til sensur og omfanget på modulen eller kurset er optimalt. Budsjett knyttet til sensur og undervisning generelt, arbeidsmengden, tilgjengelige evaluatorene samt tid til å utføre evalueringsarbeidet må sees i forhold til antallet studiepoeng som avlegges og type oppgave som evalueres.

2.3.1 Sensorordninger og kvalitet

I forbindelse med utredningen av Kvalitetsreformen (NOU, 2000: 14: 610) ble det hevdet at ”ekstern sensur representerer omtrent den eneste kvalitetssikringen av eksamen”. Det ble også argumentet for at ”den eksterne sensorfunksjonen (burde) kunne gjøres rimeligere og mer målbevisst som kvalitetssikring”. To spørsmål kan sies å ha vært flettet sammen i denne debatten: En klar oppfatning av at lærestedene brukte både for mye penger og tid på sensurering, og et ønske om å etablere bredere kvalitetssikringsordninger, ordninger som ikke kun tok hensyn til kvaliteten på den avsluttende vurderingen, men som også bredere vurderte utdanningstilbudet som møtte studentene ved lærestedene. Som det fremgår av diskusjonen av Lauvås’ aspekter, er det imidlertid ikke entydig hvordan kvaliteten på sensuren skal bestemmes. Design av sensorordninger innebærer å vurdere en lang rekke forhold og balansere delvis motsetningsfylte aspekter. En måte å nærme seg dette kompliserte spørsmålet på er gjennom å diskutere hva kvaliteter ved god sensur er.

2.3.2 Sensorordning og kostnader

I underlaget for Kvalitetsreformen ble det argumentert for at en kombinasjon av flere og mindre moduler i utdanningsprogrammene og flere år med generelle budsjettinnstramninger gjorde det ”tvingende nødvendig å se kritisk på utgiftene til eksamen” (NOU, 2000: 14: 609). Det ble også lagt til grunn at ”en relativt tung utgiftspost var eksterne sensorer” (NOU, 2000: 14: 609).

I spørsmålet om sensorordninger og kostnader ligger det implisitt et spørsmål om merkostnadene ved å bruke ekstern sensur sammenlignet med kostnadene ved å bruke

intern sensor. For fullt ut å ta stilling til dette spørsmålet må vi også ta hensyn til alternativkostnaden ved de ulike sensorordningene: Hva kunne vi alternativt brukt pengene vi bruker til ekstern sensur til? Og hva kunne vi alternativt brukt interne sensorers arbeidstid til? Hva er kostnadene ved en god sensorordning? Hvordan skal en god sensorordning finansieres? Hva må prioriteres bort for å få en god sensorordning?

Informasjon om lærestedenes utgifter er ikke umiddelbart tilgjengelig. I dag rapporteres både drifts- og lønnsutgifter i store kategorier, både på lærestedsnivå, på fakultetsnivå og instituttnivå. Disse regnskapsrapportene gir i seg selv ikke informasjon om utgifter knyttet til forskjellige aktiviteter som i vår sammenheng undervisning, eksamen eller sensur. Offentlig lett tilgjengelig informasjon om utgifter til sensur er således ikke umiddelbart enkelt å fremskaffe. Det betyr at for å vurdere spørsmålet om en eventuell sammenheng mellom sensorordninger og kostnader, er vi henvist til å basere oss på aktørenes egne vurderinger.

I diskusjonen av kostnader og målsettinger knyttet til sensorordninger bør, også diskusjonen utvides til å handle om ”prisen på gode undervisnings- og evalueringsopplegg”. Kostnadene knyttet til sensorordninger bør ikke sees isolert fra kostnadene og kvaliteten knyttet til utdanningstilbudet lærestedene gir.

Den sentrale debatten handler om kostnader og kvalitet ved utdanningstilbud i et langt bredere perspektiv, problemstillingen i vår sammenheng er hvordan vi kan frembringe valid informasjon om kostnader knyttet til ulike sensorordninger, og mer spesifikt merkostnadene ved ekstern sensur.

2.3.3 Sensur, læring og vurdering

I diskusjonen av vurderingsformer kan det skilles mellom formative og summative vurderingsformer. Begge vurderingsformene innebærer at informasjon om ytelse og måloppnåelse relatert til læring samles inn og fortolkes, samt at informasjonen brukes som bakgrunn for handling, det vil si tilbakemelding. Summativ vurdering gjøres vanligvis ved slutten av en innlæringsperiode og har til hensikt å sette en karakter eller sertifisere. Formativ vurdering som favner vidt, beskriver interaksjonen mellom student og lærer der hensikten er å justere arbeidet med læringen (Black et al., 2003). Vurderingsformene kan sees som dimensjoner i vurderingsarbeidet. I henhold til ett perspektiv på eksamen så er ikke hovedformålet å bidra til læring, men å dokumentere kandidaters evne til å svare på en gitt oppgave som ansees å være representativ for det som skal prøves og dokumenteres. Sertifiseringsformålet står altså sterkt. Imidlertid har kravet om andre vurderingsformer preget diskusjonen om og utviklingen av studieopplegg i høyere utdanning i Norge. Mappevurdering, åpenbok/kilde oppgaver, gruppe- og hjemmeoppgaver hevdes å være mer læringsorienterte enn tradisjonell eksamen. Dette er metoder som både brukes underveis og som avsluttende prøving i studieløpene. Brukt underveis har dette som oftest til hensikt å bidra til læring før en endelig eksaminering.

Evalueringen av Kvalitetsreformen viste at mappevurdering oppfattes av de vitenskapelig ansatte som en måte å kople tettere sammen læring og vurdering som er i tråd med intensjonen i reformen. Evalueringen viste også at det er store variasjoner mellom fagene med hensyn til om lærings- eller vurderingsaspektet ved mapper som evalueringsform betraktes som mest viktig (Dysthe et al., 2006: 21 -22).

Økonomi, kunnskap og kompetanse samt tradisjoner er også sterke premissleverandører i diskusjonen om vurderingsformer. Det er ikke sikkert at det alltid er de metodiske vurderingene som kommer først i diskusjoner om hvilke formål det er mest sentralt å ivareta knyttet til læring, vurdering og sertifisering/dokumentasjon.

I hvilken grad en reinnføring av ekstern sensurering vil skape endringer i vektleggingen av læring og vurdering, underveis og avslutningsvis er uvisst. Det er derfor svært interessant å undersøke hvordan informantene forholder seg til spørsmål om forholdet mellom læring og eksaminering, noe vi skal komme tilbake til.

3 Kvaliteter ved sensorordninger

I dette kapitlet diskuterer vi spørsmålet om sensorordninger og kvalitet med basis i resultater fra kartleggingsstudien og intervjuer med studieledere, studiedirektører og tilsynssensorer.

3.1 Kartlegging av sensurpraksiser

3.1.1 Resultater fra en bredt anlagt kartleggingsstudie

For å få en oversikt over omfang og utbredelse av ulike typer sensurpraksis som finnes ved hvert enkelt universitet, vitenskapelige høyskole og statlige høyskole i Norge, og innenfor ulike fagområder, gjennomførte vi en bredt anlagt kartlegging. Her ser vi både på variasjoner i sensurpraksis og på hvilken praksis som er mest utbredt.

Tabell 3-1 Omfang av ulike typer sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde. Prosent.

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Bare intern(e) sensorer	67	48	43	44	45	50
Intern(e) sensorer, ekstern(e) på noen av eksamenene	92	83	96	100	86	83
Bare eksterne sensorer	63	52	43	38	50	46
Tilsynssensor	46	39	43	44	36	42
Antall (N)	24	23	23	16	22	24

Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Tabell 3-1 viser i hvilken grad ulike typer sensurpraksis er utbredt innenfor ulike fagområder på laveregrads studier. Majoriteten av fagområdene innenfor de ulike lærestedene har merket av for flere enn en type sensurpraksis, og ved flere fagområder er alle fire typer sensurpraksis utbredt. Prosentandelene er beregnet ut i fra det totale antallet faggrupper ved de ulike lærestedene som har besvart spørsmålet (se nederste linje i tabellen). Antallet fagområder er naturlig nok lavere enn det samlede antallet læresteder som deltok i kartleggingen, ettersom ikke alle fagområder finnes ved alle læresteder. Samlet deltok 31 læresteder i kartleggingen, mens Tabell 3-1 viser at antallet innenfor de enkelte faggruppene varierer fra 16 til 24.

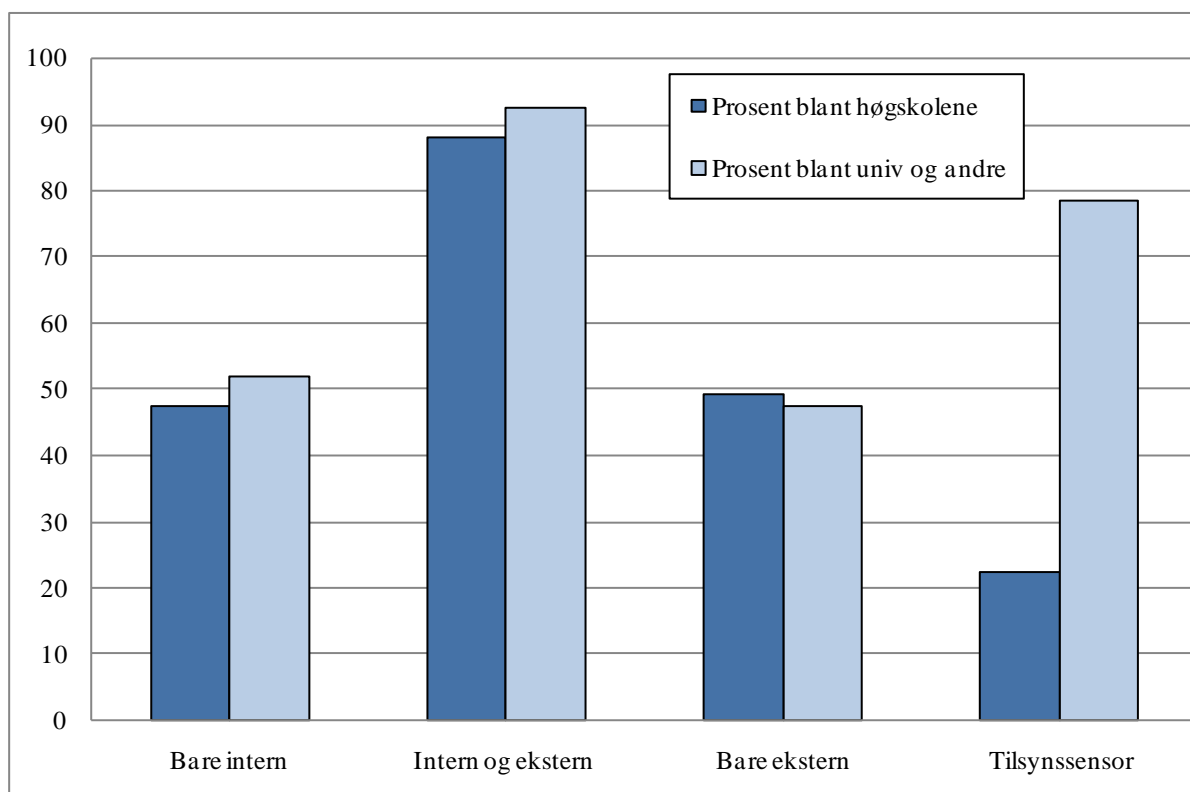
Tabellen viser betydelig variasjon i omfanget av de ulike typene sensurpraksis, i tillegg til variasjon mellom faggruppene. Den mest utbredte formen for sensur på samtlige fagområder er kombinasjonen av intern og ekstern sensor, kategorien med tittelen ”intern sensor, ekstern på noen av eksamenene³”. Kategorien favner en lang rekke ulike typer vurderingsordninger, eksempelvis en rullerende ordning for ekstern sensur på enkeltemner, ekstern sensur på avsluttende oppgaver, ekstern sensur på oppgaver som inngår i mappeinnlevering, ekstern sensur på besvarelser som allerede har fått karakter. Kategorien ”intern sensor, ekstern på noen av eksamenene” forekommer innenfor de humanistiske fagområdene ved samtlige læresteder, og på over 80 prosent innenfor de andre fagområdene. Omfanget av bruk av kategorien ”bare ekstern sensor”, altså der ekstern sensor brukes i alle eksamener på lavere grad, er noe lavere. Andelen varierer fra 38 prosent innenfor humanistiske fag til 63 prosent innenfor helse- og sosialfag. Imidlertid tyder noen av kommentarene fra lærestedene på at enkelte har forstått denne kategorien noe mer bokstavelig enn det som var hensikten. ”Bare ekstern sensor” er ment å beskrive en sensurpraksis der ekstern sensor brukes i alle eksamener innenfor et fagområde, men gjerne sammen med eller i tillegg til intern sensor. Det kan derfor være at enkelte som har svart kategorien ”intern sensor, ekstern på noen av eksamenene”, egentlig bruker en sensorordning med eksterne sensorer på samtlige eksamener. Omfanget av bruk av ekstern sensor kan derfor synes å representere et minimumsanslag, mens omfanget av bruk av ”intern sensor, ekstern på noen av eksamenene” trolig er noe overdrevet. Sammenlagt gir de to kategoriene likevel et inntrykk av et utbredt innslag av ekstern sensor i de fleste fagområder og ved samtlige læresteder. Hvor stort innslaget av ekstern sensor er innenfor de ulike fagområdene, kommer vi tilbake til litt senere i kapitlet.

Omfanget av bruk av kategorien ”bare intern sensor”, altså der det kun brukes intern sensor i alle eksamener på lavere grad, varierer også noe mellom fagområdene. Andelen ligger på 43 prosent innenfor samfunnsfagsområdet og 67 prosent innenfor helse- og sosialfag. Helse- og sosialfag er med andre ord det fagområdet der bruk av kun intern og kun ekstern sensor forekommer i størst grad. Det indikerer en relativt stor variasjon i typer sensurpraksis på lavere grad innenfor dette fagområdet. Når det gjelder bruk av tilsynssensor synes det å ha begrenset utbredelse i de ulike fagområdene. Fra å ligge på 36 prosent innenfor pedagogiske fag og lærerutdanningene, er andelen som oppgir å bruke tilsynssensor 46 prosent innenfor helse- og sosialfag. Det kvalitative materialet støtter inntrykket av at fagmiljøene i begrenset grad har etablert en tilsynssensorordning. Av 20 studier hadde 6 etablert denne ordningen.

Vi undersøker videre i hvilken grad det er systematiske forskjeller mellom ulike typer læresteder. I figur 3.1 sammenlignes sensurpraksis ved de statlige høyskolene med sensurpraksis ved universitetene og de vitenskapelige høyskolene. Til en viss grad innebærer dette en sammenligning av større og mindre læresteder.

³ I kartleggingen ble betegnelsen eksamen og besvarelse benyttet. Enkelte læresteder presiserte at med besvarelse mente man eksamen.

Figur 3.1 Omfang av ulike typer sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde og lærestedstype. Prosent.



Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Figur 3.1 illustrerer mange av de samme tendensene som også kommer frem i Tabell 3-1. Den mest utbredte formen for sensur både innenfor høyskolene og innenfor universitetene og de vitenskapelige høyskolene er kombinasjonen av intern og ekstem sensor, det vi har kalt ”intern sensor, ekstem på noen av eksamenene”. Utbredelsen av ”bare intern” og ”bare ekstem” sensor er også tilnærmet lik på høyskolene som ved universitetene og de vitenskapelige høyskolene. Ser vi på omfanget av bruk av tilsynssensor, er det derimot betydelige forskjeller mellom høyskolene og universitetene samt de vitenskapelige høyskolene. Blant høyskolene opplyser 24 prosent at bruk av tilsynssensor forekommer, mens andelen ved universitetene og de vitenskapelige høyskolene er 78 prosent. Dette samsvarer også med våre funn fra de kvalitative intervjuene. Blant de studiene som var trukket ut i denne delen av undersøkelsen, hadde så langt som bare universitetsfag etablert en tilsynssensorordning. Den kvalitative datainnsamlingen viste imidlertid at tilsynssensorfunksjonen dekkes inn av en del andre typer ordninger. Vi kommer tilbake til dette senere i kapittelet.

3.1.2 Hvilke ordninger brukes mest?

Mens Tabell 3-1 og Figur 3.1 viser utbredelsen av de ulike typene sensurpraksis, ønsket vi også i kartleggingen å få en oversikt over hvilke typer sensurpraksis som er *mest* utbredt

innenfor de enkelte fagområder. Tabell 3-2 viser andeler av de ulike sensorordningene som er mest brukt på lavere grad innenfor ulike fagområder.

Tabell 3-2 Mest utbredt sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde. Prosent.

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Bare intern(e) sensorer	26	41	27	33	38	59
Intern(e) sensorer, ekstern(e) på noen av besvarelsene	57	45	50	53	57	45
Bare eksterne sensorer	26	31	27	27	24	36
Tilsynssensor	17	18	14	7	10	27
Antall (N)	23	22	22	15	21	22

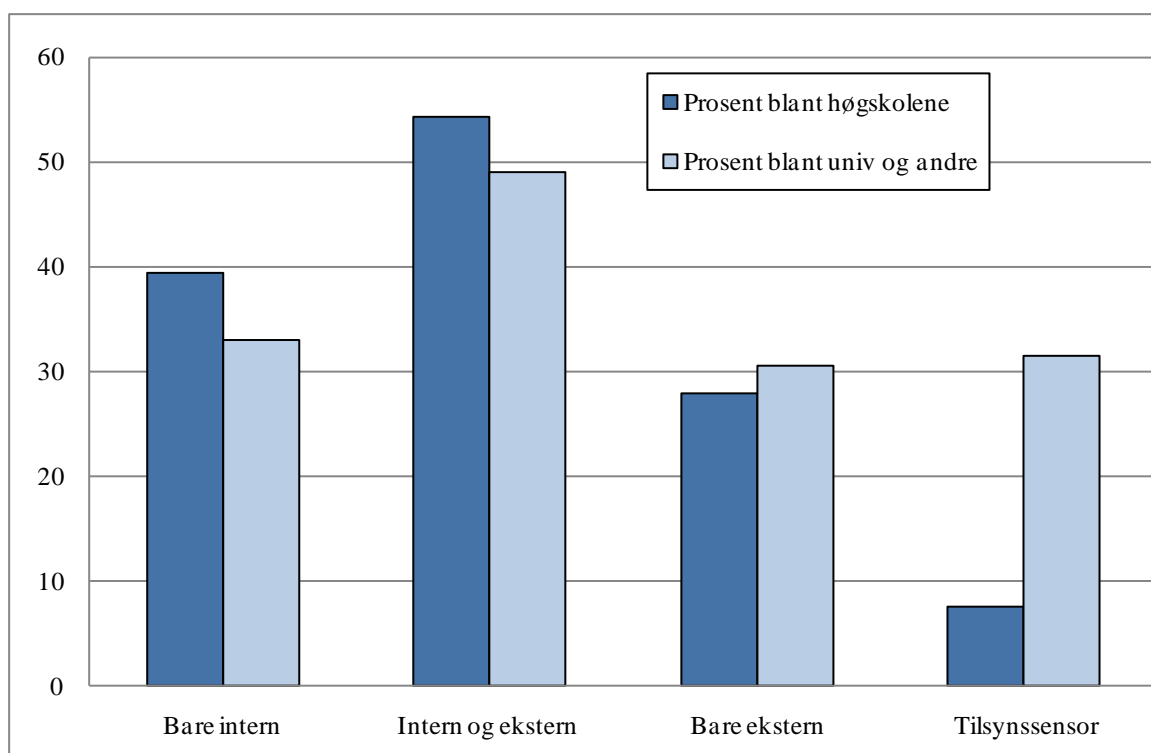
Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Tabell 3-2 viser betydelig variasjon i hvilke typer sensurpraksis som er mest brukt innenfor de ulike fagområdene. Det bør understrekes at noen fagområder har merket av for flere enn en type sensurpraksis også her (andelene kan ikke summeres til 100). I Tabell 3-1 så vi at den mest utbredte formen for sensur på samtlige fagområder er kombinasjonen av intern og ekstern sensor, det vi har kalt ”intern sensor, ekstern på noen av eksamenene”. Dette er også den sensurformen som er mest brukt innenfor samtlige av faggruppene. Andelen varierer fra 45 prosent innenfor de matematisk-naturvitenskapelige fagene, ingeniør- og tekniske fag til 57 prosent innenfor pedagogiske fag og lærerutdanningene. Ettersom kartleggingen har avdekket betydelige variasjoner i sensurpraksis selv innenfor de ulike fagområdene, bør imidlertid ikke tabell 4 tolkes for strengt. Tabellen kan likevel gi en viss oversikt over hvor utbredt de ulike sensorordningene synes å være, innenfor de ulike fagområdene.

I Figur 3.2 sammenlignes andeler mest utbredt sensurpraksis etter lærestedstype. På samme måte som Figur 3.1, viser Figur 3.2 en tydelig forskjell mellom lærestedene i bruk av tilsynssensor. Blant høyskolene opplyser kun 8 prosent at tilsynssensor utgjør den mest brukte formen for sensurpraksis, mens andelen ved universitetene og de vitenskapelige høyskolene er 31 prosent. Mens tilsynssensor forekommer ved et flertall av universitetene og de vitenskapelige høyskolene (78 prosent), er det altså kun et mindretall som oppgir at dette utgjør den mest brukte formen for sensurpraksis. Den kvalitative delen av datainnsamlingen viste at tilsynssensor ofte brukes i kombinasjon med andre

sensorordninger. Andelen som oppgir intern sensor som den mest utbredte sensurpraksis, er noe høyere ved høyskolene enn ved universitetene og de vitenskapelige høyskolene, 39 mot 33 prosent. Det samme gjelder bruk av kombinasjonen av intern og ekstern sensor, her er andelen 54 prosent ved høyskolene og 49 prosent ved universitetene og de vitenskapelige høyskolene. Andelen som oppgir kun ekstern sensor som den mest utbredte sensurpraksis, er tilnærmet lik ved høyskolene som ved universitetene og de vitenskapelige høyskolene, 28 og 31 prosent.

Figur 3.2 Mest utbredt sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde og lærestedstype. Prosent.



Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Som tidligere nevnt, viser kartleggingen at den mest utbredte formen for sensur på samtlige fagområder er kombinasjonen av intern og ekstern sensor, det vi har kalt ”intern sensor, ekstern på noen av eksamenene”. For å få et innblikk i omfanget i bruk av ekstern sensor i disse tilfellene, ble lærestedene bedt om å angi prosentandelen av eksamenene som rettes av ekstern sensor innenfor hver faggruppe der ekstern sensor brukes på noen av eksamener. Tabell 3-3 viser gjennomsnittlig andel av eksamener som rettes av ekstern sensor innenfor hvert fagområde.

Tabell 3-3 Gjennomsnittlig andel av eksamener som rettes av ekstern sensor innenfor hvert fagområde, i tilfeller der ekstern sensor brukes på noen av eksamener ved 31 læresteder. Prosent.

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Intern(e) sensorer, ekstern(e) på noen av besvarelsene	52	36	53	40	38	43
Antall (N)	21	17	18	13	16	15

Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Som Tabell 3-3 viser, er det ingen stor variasjon mellom fagområdene når det gjelder omfanget i bruk av ekstern sensor der ekstern sensor kun brukes på noen av eksamener. Den høyeste andelen finner vi innenfor samfunnsfag, med 53 prosent, mens matematisk-naturvitenskapelig, ingeniør- og tekniske fag har den laveste andelen; 36 prosent. Det bør understrekes at flere av lærestedene hadde vanskeligheter med å beregne en slik gjennomsnittsprosent, ettersom omfanget i bruk av ekstern sensor kunne variere veldig mellom ulike fag innenfor samme fagområde. De kvalitative dataene viser også at sensorordninger varierer innenfor et studieprogram; et institutt; en avdeling; et lærested og mellom fag på tvers av læresteder. Antall fagområder som har oppgitt svar i kartleggingen er også noe lavere her enn i andre deler av kartleggingen. Tallene i Tabell 3-3 bør derfor ikke tillegges for stor vekt, men heller tolkes som omtrentlige anslag over omfang i bruk av ekstern sensor på noen av eksamener innefor de enkelte fagområdene.

3.1.3 Sensorordning en del av kvalitetssikringssystemene?

Et viktig forhold å se nærmere på her er også i hvilken grad dagens sensurering er en del av lærestedenes systemer for kvalitetssikring, eventuelt på hvordan kvalitetssikringsprosesser integrerer sensurerings- og eksaminasjonsarbeid. Siste del av kartleggingen omhandler i hvilken grad de ulike sensorordningene på lavere grad innefor de ulike fagområdene er knyttet opp til lærestedenes kvalitetssikringssystem. Tabell 3.4 viser en oversikt over andel sensorordninger på lavere grad innefor de ulike fagområdene som er knyttet opp til lærestedenes kvalitetssikringssystem etter fagområde.

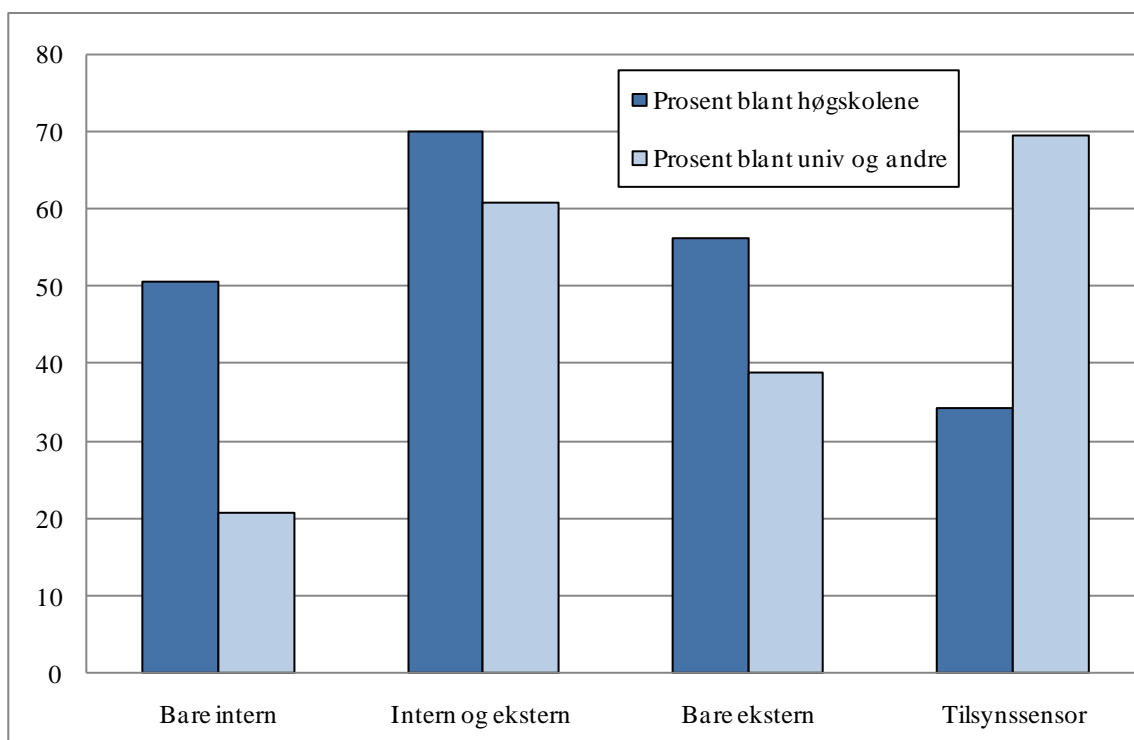
Tabell 3-4 Andel sensorordninger på lavere grads studier knyttet opp til lærestedets kvalitetssikringssystem ved 31 læresteder etter fagområde. Prosent.

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Bare intern(e) sensorer	39	41	41	47	48	35
Intern(e) sensorer, ekstern(e) på noen av besvarelsene	70	73	68	60	71	57
Bare eksterne sensorer	57	55	45	53	52	43
Tilsynssensor	48	45	45	47	43	35
Antall (N)	23	22	22	15	21	23

Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

Tabellen viser til en viss grad tilsvarende mønster som resten av kartleggingen. Sensorordningen ”intern sensor, ekstern på noen av eksamenene” synes å være den som i størst omfang er knyttet opp mot lærestedenes kvalitetssystem. Dette er ikke overraskende, ettersom det også er denne sensorordningen som er mest anvendt innenfor samtlige av faggruppene. Andelen varierer fra 60 prosent innenfor humanistiske fag (57 prosent innenfor gruppen ”Annet”) til 73 prosent innenfor matematisk-naturvitenskapelige, ingeniør- og tekniske fag. Også ordningene med bare ekstern sensur er i betydelig grad knyttet opp mot lærestedenes kvalitetssystem, mens ordningene med kun intern sensor og tilsynssensor i noe lavere grad er knyttet opp mot lærestedenes kvalitetssystem innenfor de fleste fagområder.

Figur 3.3 Andel sensorordninger på lavere grads studier knyttet opp til lærestedets kvalitetssikringssystem ved 31 læresteder etter fagområde og lærestedstype. Prosent.



Kilde: Spørreskjema til studieledere ved hvert lærested (vedlegg 1). Flere svar i hver faggruppe mulig.

I Figur 3.3 sammenlignes andeler sensurpraksis knyttet opp mot lærestedenes kvalitetssystem etter lærestedstype. Sammenlignet med figurene 3.1 og 3.2, synes forskjellene mellom de to gruppene av læresteder å fremtre tydeligere i figur 3.3. En slik forskjell er andelen læresteder der bruk av tilsynssensor er knyttet opp mot lærestedenes kvalitetssystem. Blant høyskolene gjelder dette 34 prosent, mens andelen ved universitetene og de vitenskapelige høyskolene er om lag det dobbelte; 69 prosent. Derimot er andelen læresteder der bruk av kun intern sensor er knyttet opp mot lærestedenes kvalitetssystem betydelig høyere blant høyskolene enn blant universitetene og de vitenskapelige høyskolene; 51 mot 21 prosent. Også de andre sensorordningene, eksternt sensor samt kombinasjonen av eksternt og intern sensor, synes i større grad å være knyttet opp mot lærestedenes kvalitetssystem ved høyskolene enn ved universitetene og de vitenskapelige høyskolene. Vi kan ikke utelukke at noen av disse forskjellene har sammenheng med hvordan denne delen av kartleggingen har blitt forstått og rapportert fra de ulike lærestedene. Nærmere analyser viser for eksempel betydelig spredning mellom lærestedene i antall typer sensorordningene som er knyttet opp mot lærestedenes kvalitetssystem. Enkelte læresteder oppgir at ingen av sensorordningene er omfattet av kvalitetssikringssystemet. Andre oppgir at flere sensorordninger er omfattet av lærestedets/fagområdets kvalitetssystem. Spredningen er større blant høyskolene enn blant universitetene og de vitenskapelige høyskolene.

3.2 Vurderinger av kvaliteten på sensuren

3.2.1 Metodiske hensyn

De kvalitative intervjuene belyser nøkkelaktørers vurderinger av sammenhengen mellom ulike sensorordninger og spørsmål om kvalitet. I det følgende analyserer vi intervjumaterialet i forhold til ulike metodiske aspekter ved sensur (se kapittel 1 og 2).

Sensorordning med ekstern sensur

Et hovedsynspunkt i intervjumaterialet er at sensorordninger som omfatter intern og ekstern sensor bidrar til å sikre kvaliteten på sensuren. Bruk av ekstern sensor legitimeres gjennom å vise til at ordningen bidrar til å sikre nasjonale standarder i karaktersettingen og vurderingene. Dette innebærer at vurderingen gjennomføres på likest mulig måte for alle studentene som skal prøves.

For å sikre den indre validiteten i vurderingene understrekes viktigheten av at sensorene er inneforstått med oppgaveteksten, tolkninger som benyttes i sensurarbeidet og bedømmelseskriteriene. Organiseringen av dette arbeidet varierer fra enkel overlevering av informasjon, til møter mellom eksterne sensorer og faglig ansvarlige ved det enkelte lærested. I forkant av sensurering gjennomgås hovedmål for minstekrav for besvarelsene og hva som kreves i en besvarelse for å oppnå de beste karakterer.

Sensorenes vurderingskompetanse foreslås understøttet gjennom et system for oppfølging og opplæring av sensorene. Det handler om å få de eksterne sensorene til å ta opp i seg den interne sensurkulturen, samtidig som eksterne sensorers objektivitet må ivaretas. Et av lærestedene uttrykte behov for tydelige kvalitetssikringsrutiner og hevdet at tilsynssensor var viktig i dette arbeidet. På dette lærestedet utgjør tilsynssensor en viktig og sentral rolle for ivaretagelse av kvalitetssikring. Det at sensorene må dekke fagområdet ble også oppgitt som et kvalitetsaspekt. Informantene la vekt på reliabiliteten knyttet til sensorordningen ved å understreke viktigheten av samsvar mellom sensorer i vurderingene.

Noen av informantene ved de små høyskolene uttrykte et behov for å velge ekstern sensorordning for å signalisere kvalitet til studenter, samt til andre læresteder og arbeidsgivere. Ekstern sensur som en strategi for å unngå å bli mistenkeliggjort for å gi for ”snille” karakterer og var altså en del av kvalitetstenkningen rundt sensorordninger.

Et sentralt argument for å beholde en ordning med ekstern sensur er formelle hensyn (se kapittel 1 og 2). Ved bortfall av obligatorisk ekstern sensur hevdes det at kvaliteten i sensuren går tapt, truer rettssikkerheten og hindrer transparensen i høyere utdanning.

Enesensur

Tendensen er at enesensur benyttes i enkeltemner, mens ved de store individuelle oppgavene, som Bacheloroppgaven, trekkes ekstern sensor inn. Der det benyttes enesensur

i større skala trekkes enkeltemner ut for ekstern sensur, og den eksterne sensuren benyttes som korrektiv til enesensureringen. I tillegg til ekstern sensur benyttes også tilsynssensor. Tilsynssensor deltar blant annet på vurderingsmøter i forbindelse med sensurarbeidet og planseminarer. Kombinasjonen av ekstern sensor og tilsynssensor benyttes for å ivareta kvaliteten på sensuren.

En av informantene går motsatt vei og hevder at gammel ordning med intern og ekstern sensur ikke borger for kvalitet. Ved å ha to sensorer hevdes det at den ene kan støtte seg til den andre og i liten grad bidra. Det framheves også at enesensur fører til en ekstra skjerpelse gjennom det ansvaret dette gir. Kvalitetssikringen ved enesensur består i dette tilfellet av at ekstern sensor trekkes inn for sensurering av en andel av oppgavene. Denne ordningen mener informanten kan forsvares på grunn av en modularisert utdanning med hyppige eksamener og mindre kurs.

Tilsynssensor

Intervjumaterialet viser at tilsynssensorfunksjonen har forskjellige betegnelser ved de ulike lærestedene. Lærestedene har etablert ordninger som benevnes programsensor, bedømmersensor, eksamenskomiteer og nivåkontrollør. Hensiktene med disse ordningene er å bidra til kvalitetssikringen av sensuren i tilfeller der man har gått bort fra ”den gamle ordningen” med gjennomgående ekstern sensur på lavere grad. En tilsynssensor (programsensor, nivåkontrollør) har til oppgave å vurdere sammenhengen mellom emnebeskrivelsene og undervisningsopplegget, vurdere sammenhengen i emne og ta stikkprøver av karaktergivningen på enkeltbesvarelser. Tilsynssensorer skriver gjerne en rapport i året samt deltar på et årlig møte mellom alle tilsynssensorene på instituttet. Det varierer hvilke av disse oppgavene tilsynssensorene har. En tilsynssensor hadde kun hatt befatning med å jobbe overordnet med faget, en hadde fungert som ekstern sensor, mens en hadde hatt oppgaver som berører alle typer oppgaver som nevnt over.

Andre sensurformer

Flere av aktørene understreket at det ikke var nødvendig med ekstern sensur på all vurdering i alle emner. Ved et av lærestedene ble det av en studieleder fortalt om ulik vektlegging av intern og ekstern sensor, eller at det ikke var samsvar mellom sensorene. Dette til tross for rutiner for å sikre ekstern sensors forståelse for bedømmelseskriteriene. Det ble derfor bestemt å prøve ut en ordning med to interne sensorer der kvaliteten skulle sikres gjennom å samsensurere en besvarelse, for så å gå gjennom den ene besvarelsen, enes om kriterier, nivå med mer for videre sensur. Alle sensorene på emnet samsensurerer for å unngå at det etableres noen mønster og for å gjøre evalueringssituasjonen så lik som mulig for studentene. Dette eksemplet viser noe av bakgrunnen for at det oppstår faglige diskusjoner om kvalitet ved sensur i fagmiljøet som medfører endringer i sensorordninger og at nye former for kvalitetssikringssystemer vokser fram for å sikre en bedre kvalitet enn foregående ordning. Viktigheten av at sensor er fortrolig med kriteriene sensureringen benyttes også som argument for å erstatte ”gammel ordning” med en av de nye variantene

for sensur. Argumentet er at intern sensorer har større forståelse for de aktuelle vurderingskriteriene enn eksterne sensorer.

3.2.2 Praktiske hensyn

Det praktiske hensynet legitimeres gjennom en optimal balanse mellom ressursbruk (både budsjettmessig, arbeidsmengde og tilgjengelige evaluatorene, tid til å utføre evalueringsarbeidet med mer). En av informantene begrunnet bruk av en kombinasjon mellom en intern og en ekstern sensor på samtlige emner på lavere grad med et stort antall studenter som medførte for liten kapasitet blant de faglig ansatte til å håndtere en ren intern sensorordning. Vi skal diskutere økonomiske aspekter ved ulike sensorordninger nærmere i neste kapittel.

3.2.3 Formelle hensyn

Studielederne og studiedirektørene beskriver det eksterne blikket som en kvalitetskontroll av egen virksomhet, og mellom et etablert fagmiljø og ekstern praksis. Bruken av ekstern sensor hevdes å gi legitimitet gjennom å skulle være rimelig og rettferdig overfor studentene. Sensorordning med ekstern sensor som hentes utenfor gjeldende lærested fremmer rettferdig sensurering. Studentenes rettssikkerhet fremmes ved bruk av ekstern sensor. Gjennom bruk av ekstern sensor sikres det også en åpenhet og innsyn i evalueringsprosessen og evalueringsarbeidet, med små muligheter for å ha en praksis som er unndratt innsyn.

3.2.4 Kommunikasjonsmessige hensyn

Det kommunikasjonsmessige hensynet gjelder både i forhold til studentene og mellom vitenskaplige ansatte/sensorer. Det påpekes at studentene skal være inneforstått med på hvilket grunnlag de får karakteren og problemene som oppstår dersom dette ikke oppleves som rimelig og rettferdig. Det kommunikasjonsmessige hensynet består også av dialog mellom sensorene (interne og eksterne) der det oppnås en enighet om nivået, innholdet, det faglige og det vurderingsmessige på sensureringsarbeidet. Men dette handler i like stor grad om dialogen mellom de som utfører sensuren og studentene som får en vurdering ved muntlig eksaminering, eller i form av karakterer på skriftlige arbeider.

3.3 Konkluderende betraktninger om sensorordning og kvalitet

Det er en betydelig bevissthet blant informantene om sammenhengen mellom kvalitet og ulike sensurordninger. Både ved universitets- og høyskolefag settes likhetstegn mellom kvalitet på sensuren og bruk av tradisjonell sensur, eller nyere kombinasjoner av intern og ekstern sensur. Noen uttrykte imidlertid en skepsis mot at ekstern sensur alltid er et kvalitetsstempel. Studiedirektørene, men også noen av studielederne, mente at kvaliteten på sensorordninger kunne ivaretas med intern sensurering, dersom dette omfattet bruk av flere sensorer (til forskjell fra enesensur). Eventuelt med innslag av stikkprøvekontroll eller

andre ordninger. Flere fortalte om pågående prosesser for å finne fram til ordninger som fungerte for sitt fag og som ivaretok kvalitetshensynene ved sensurering. Slike prosesser har sammenheng med praktiske hensyn om optimal balanse mellom ressursbruk og omfanget av emnet.

Det er stor variasjon i utbredelse av ulike former for sensur på lavere grads studier ved de ulike lærestedene, i tillegg til variasjon mellom faggruppene ved det enkelte lærested. På tross av variasjonen i sensurpraksis er kombinasjonen av intern og ekstern sensor den mest utbredte formen for sensur innenfor samtlige fagområder. Dette er også de typene av sensurordning som fagmiljøene i størst grad hevder gir best kvalitet på sensuren.

Det synes å være et utbredt innslag av ekstern sensor i de fleste fagområder og ved samtlige læresteder. Men bruk av ekstern sensor varierer – nesten fra 0 til 100 prosent – innenfor et fagområde. Det er også variasjon i omfanget av bruk av ekstern sensor mellom fag innenfor samme fagområde. Flere av lærestedene hadde derfor vanskeligheter med å beregne en slik gjennomsnittsprosent. Om lag halvparten av lærestedene rapporterer at deler av eksamenene på lavere grad sensureres internt, uten bruk av ekstern sensor. Variasjoner i bruk av sensorordninger kan sees i sammenheng med den modulbaserte oppbyggingen av bachelorgraden, men også fagenes egenart. Flere av studielederne understreket viktigheten av å bruke ulike vurderingsformer. I mindre emner varierer bruken av sensorordninger fra utelukkende bruk av ekstern sensor, til enesensur på utvalgte emner med ulike kombinasjoner av kvalitetssikringsordninger.

Type lærested (høyskole eller universitet) synes også å ha betydning for tenkning rundt kvalitet og sensorordninger. Fag ved mindre høyskoler har uttrykt behov for å velge ekstern sensorordning for å signalisere kvalitet til studenter, samt til andre læresteder og arbeidsgivere. Selv om høyskolene mer enn universitetene vektlegger behovet for ekstern sensur, gjelder dette også fag ved universitetene. Vi finner eksempler på at universitetsmiljø har beholdt ”gammel ordning” i stor utstrekning og endog styrker denne med tilsynssensor.

Tilsynssensor er den minst brukte sensurordningen. Bruk av tilsynssensor forekommer videre i mindre grad ved høyskolene enn ved universitetene. Men det synes som at noe av forskjellen i bruk av tilsynssensor kan ha sammenheng med hva som legges i denne rollen. Flere fag har organisert ekstern sensur på nye måter sammenlignet med ”gammel ordning”. De nye ordningene søker å ivareta et eksternt blikk på vurderingene eller vurderingsordningene, men uten at funksjonen benevnes tilsynssensor.

Når det gjelder hvordan sensurordninger er knyttet opp til lærestedets kvalitetssikringssystem finner vi at ”gammel ordning” og ulike kombinasjoner av intern og ekstern sensur i større grad enn ordningene med kun intern sensor og tilsynssensor er knyttet opp mot lærestedenes kvalitetssystem innenfor de fleste fagområder.

4 Sensurpraksis og kostnader

I dette kapitlet drøfter vi forholdet mellom sensurpraksis og kostnader. Her skal vi se på informantenes vurderinger og erfaringer knyttet til praktiske aspekter ved sensorordninger som omtalt i kapittel 2.

Ressursbruken knyttet til sensurpraksis har vært en tilbakevendende diskusjon. Før Kvalitetsreformen var en utbredt oppfatning at den norske sensorordningen var for ressurskrevende. En undersøkelse fra 2000 kunne verken bekrefte eller avkrefte hvorvidt det ble brukt mye ressurser på eksamen sammenlignet med ressursbruken for andre vurderingsformer (Vabø, 2000: 42). En forklaring på at det norske systemet var særlig ressurskrevende lå i lovverket som påla bruk av ekstern sensor i all avsluttende vurdering. Kostnadene ved ekstern sensur hadde imidlertid vært debattert også i andre land med innbyrdes ulike ordninger for sensur, som Danmark, Storbritannia og Sverige (Vabø, 2000: 69).

4.1 Er ekstern sensur kostnadskrevende?

Det er ikke lett å diskutere kostnader knyttet til ulike sensorordninger. Aktørenes vurderinger er forskjellige og det vanskelig å sammenligne kostnader knyttet til sensur mellom fag og læresteder. Dette skyldes blant annet variasjoner i studietilbudet, fagforskjeller, forskjeller i studentgruppen og studieopplegget. Også tidligere studier av dette spørsmålet har hatt tilsvarende utfordringer (Vabø, 2000). I vår studie finner vi for det første varierende oppfatninger om kostnadene ved ulike sensorordninger. For det andre varierer det hvilke kostnader informantene tar med i betraktningen. For det tredje varierer fagmiljøenes tilpasning til den økonomiske situasjonen.

Det er ulike oppfatninger blant studielederne om sammenhengen mellom sensorordning og økonomi. En hovedoppfatning er at ekstern sensur er en svært kostnadskrevende ordning, men det varierer hvordan fagmiljøene har tilpasset seg til denne oppfatningen. Flere legger vekt på nødvendigheten av å redusere bruken av denne sensorordningen fordi dette er en svært kostnadskrevende ordning. Andre vektlegger at det er nødvendig å opprettholde ekstern sensur av kvalitetsmessige hensyn og argumenterer for at økonomiske hensyn bør underordnes kvalitetshensynet.

Noen argumenterer altså for at ekstern sensur er svært kostnadskrevende. Det er økonomiske besparelser å hente ved å redusere bruken av ekstern sensur. Det presiseres at det ikke er honorarene som representerer den største utgiften i forbindelse med ekstern sensur, men at det er utgifter til reiser med mer, og at dette kan spares inn ved økt bruk av videokonferanser og telefonsensur.

Kostnader er begrunnelsen for at noen fagmiljøer vegrer seg mot ordninger som krever for mye oppfølging. Det hevdes at flere har gått tilbake til tidligere ordninger fordi store ressurser ble benyttet til vurderingsformer som krever mye tilbakemelding til studentene.

Blant dem som har opprettholdt ”gammel ordning” av kvalitetsmessige hensyn på tross av at ordningen oppfattes som svært kostnadskrevende, erfarer noen at ordningen står under press der man oppfordres til å revurdere ordningen for å oppnå økonomiske besparelser. Begrunnelsen for å endre ordning av økonomiske grunner er at man kan opprettholde kvaliteten ved hjelp av andre (mindre kostnadskrevende) sensorordninger.

Andre mener at ekstern sensur er en forholdsvis mindre kostnadskrevende sensorordning. I alle fall sammenlignet med utstrakt bruk av alternative prøveformer underveis, og da gjerne i kombinasjon med et stort innslag av intern sensur. I denne gruppen argumenteres det med at man må forsøke å finne den optimale balansen mellom vurderingsordninger og økonomiske rammer.

Flere uttrykker at om de hadde hatt større ressurser, så hadde de ikke ønsket å bruke disse midlene til sensurering, men heller til å styrke undervisningsopplegget. Dette poenget kommer vi tilbake til i neste kapittel. En stilte spørsmålstegn ved om ekstern sensur var riktig bruk av midler, og var spesielt opptatt av om det var en sammenheng mellom det de brukte av ressurser og det de fikk igjen. I den forbindelse mener informanten at det viktigste er å variere vurderingsformene.

Der man har endret sensorordningene av økonomiske grunner argumenteres det for at de nye ordningene er ”nest-best” løsninger i en presset økonomisk situasjon, der disse dog er av tilfredsstillende kvalitet. Andre hevder at det, på tross av økonomiske hensyn som kan ha påskyndet endringer av sensorordninger, oppnås bedre kvalitet på sensuren ved de nye ordningene.

4.2 Hvilke kostnader skal tas hensyn til?

Utgifter til ekstern sensor synliggjøres i større grad enn kostnader til intern sensor. Her går det et skille mellom dem som hovedsakelig fokuserer på sensorhonorar (til ekstern sensur) og reisekostnader knyttet til ekstern sensors tilstedeværelse under sensuren (reise og opphold), og dem som ser kostnadene knyttet til ekstern sensur i sammenheng med den interne ressursbruken. Intern sensur har også en kostnadsside, som påløper i form av intern sensors tidsbruk på sensur. De som legger hovedvekten i resonnetet på sensorhonorar og reisekostnader, understreker at denne utgiftsposten kan reduseres ved bruk av videokonferanser og telefonsensur. I den andre gruppen, som ser honorar- og reisekostnader i sammenheng med ressursene som går med til intern sensur, fokuserer man på at sensurarbeid er kostnadskrevende uavhengig av hvilken ordning man velger. Her argumenteres det for at kostnadene til sensur sees i sammenheng med kostnader knyttet til å gjennomføre et undervisningsopplegg i et bredere perspektiv.

4.3 Henger erfaringene med kostnader sammen med valg av sensorordning?

Det er ikke en systematisk sammenheng mellom oppfatninger av kostnader og hvilken sensorordning som velges. Ekstern sensur kan ha blitt opprettholdt selv om erfaringene er at det er kostnadskrevende. Det kan ha blitt innført nye ordninger som oppfattes som like kostnadskrevende, men fordi erfaringer tilsier at den valgte løsningen likevel opprettholder kvaliteten. Det kan også være at de nye ordningene medfører andre positive funksjoner som metaevaluering av studieopplegget. Likevel er det gjennomgående en tydelig økonomisk begrunnelse for å endre sensorordning fra ekstern sensur til andre former for sensur.

4.4 Tidsbruk til eksamensarbeid

Våre informanter legger vekt på at det er økonomiske begrunnelser som har drevet frem endringer i sensorordningene etter Kvalitetsreformen. Hovedvekten i resonnetet knyttet til kostnader ved ulike sensorordninger ligger på variable kostnader som sensorhonorarer, reise- og oppholdskostnader. Vi har sett at ressursbruk i form av arbeidstiden som både eksterne og interne sensorer legger i sensurarbeidet, i langt mindre grad er i fokus. Sensur og eksamensarbeid er en del av undervisningsoppgavene ved lærestedene. Vi har innledningsvis argumentert for at diskusjonen om kostnader til sensur i vid forstand handler om prioriteringer mellom ulike sensorordninger og ulike undervisningsopplegg.

Vi har i denne undersøkelsen basert oss på informantenes vurderinger av kostnader knyttet til sensorordninger og vi har ikke andre data på hvordan tidsbruk til sensur fordeler seg før og etter Kvalitetsreformen. Siden sensurarbeid er en del av eksamensarbeidet ved lærestedene kan vi illustrere ressursbruken knyttet til eksamensarbeidet ved hjelp av tidligere undersøkelser. En tidligere undersøkelse ved NIFU STEP (Frølich et al., 2008: 49) diskuterer tidsbruk til eksamensarbeid ved norske læresteder før og etter Kvalitetsreformen. Data fra evalueringen av Kvalitetsreformen viser at både ved universitetene og ved høyskolene oppgir over halvparten av respondentene at de bruker mer tid enn tidligere på eksamensarbeid. Den samme undersøkelsen fant ingen signifikante forskjeller mellom faggruppene ved universitetene når det gjelder tid brukt på eksamensarbeid etter Kvalitetsreformen. Ved høyskolene fant man større variasjon i andelen vitenskapelig ansatte som oppgav at de brukte klart mer tid på eksamensarbeid, mellom 45 og 76 prosent. Helse- og sosialfag var i denne undersøkelsen den eneste faggruppen ved høyskolene som skilte seg signifikant fra de andre faggruppene, og som i *mindre* grad oppgav at de bruker klart mer tid på denne typen oppgaver etter at Kvalitetsreformen ble innført.

I rapporten (Frølich et al., 2008: 50) brukes data fra tidligere undersøkelser av tidsbruk til eksamen ved universitetene til å anslå tid brukt på eksamensarbeid før Kvalitetsreformen. I

1990 brukte det vitenskapelige personalet ved universitetene mellom 10 og 11 dager på denne typen oppgaver (Smeby, 1993). I universitetsundersøkelsen fra 2001 ble det observert en økning i andelen som oppga å bruke tid på eksamensarbeid. Totalt deltok 88 prosent av det faste vitenskapelige personalet i arbeidsoppgaver knyttet til eksamensarbeid i 2000. Den enkelte brukte dessuten gjennomsnittlig 14 dager på eksamensarbeid i 2000, sammenliknet med 10,5 dager i 1990.

Universitetsundersøkelsene viser enkelte forskjeller mellom faggruppene når det gjelder omfang av eksamensarbeid, både når det gjelder andelen som deltar og hvor mye tid den enkelte gjennomsnittlig bruker.

Tabell 4-1 Andel faste vitenskapelige ansatte ved universitetene som oppga å bruke tid på eksamensarbeid i 1990 og 2000 etter faggruppe. Prosent. Gjennomsnittlig antall dager for de som deltok i parentes

	Hum	SV	Nat. vit.	Med	Tekn	Landbruk
1990	82 (13,1)	86 (12,6)	90 (9,4)	74 (6,6)	97 (13,3)	*
2000	86 (14,4)	92 (16)	91 (12,7)	80 (11)	95 (18,3)	87 (12,9)

Kilde Frølich og Waagene (2008: 50)

Ansatte innenfor medisin er den gruppen hvor færrest oppgir å ha brukt tid på eksamensarbeid. Samtidig er de den gruppen som gjennomsnittlig bruker færrest dager på slike oppgaver. Teknologene er den gruppen hvor flest oppgir å ha brukt tid på eksamensarbeid. I tillegg er de den gruppen som gjennomsnittlig bruker flest arbeidsdager på denne typen administrative oppgaver (Frølich et al., 2008: 50).

4.5 Konkluderende betraktninger om sensurpraksis og kostnader

Vår gjennomgang viser at det er ingen enkel sammenheng mellom sensurordninger og kostnader. Det er variasjon i hvilke ordninger som oppfattes som mest kostnadskrevende og hva som vektlegges. En informantenes erfaringer, vurdering og synspunkter tyder på at den gamle ordningen av mange oppfattes som mest kostnadskrevende; mens enesensur oppfattes som minst kostnadskrevende. Det varierer hvordan kostnadene til de andre formene for sensur vurderes, men de rangeres gjerne i en mellomkategori mellom gammel ordning og enesensur.

Tilsynssensor anses av noen å være en nyvinning for å opprettholde kvalitet på sensur, og samtidig redusere kostnadene rundt selve sensureringsarbeidet. Synet på tilsynssensor varierer imidlertid. En av informantene hevder at tilsynssensor ikke medfører noen besparelse i så måte.

Data fra tidligere undersøkelser av vitenskapelig ansattes tidsbruk til eksamensarbeid dokumenterer en viss økning i tid brukt til eksamensarbeid i tiåret før Kvalitetsreformen.

Data fra evalueringen av Kvalitetsreformen viser at 60 prosent av de vitenskapelige ansatte oppgir at de bruker mer tid på eksamen etter reformen sammenlignet med før, 6 prosent oppgir å bruke mindre tid enn tidligere. En tolkning av disse resultatene er at tid brukt på eksamensarbeid ikke er betydelig redusert etter reformen. Dataene gir imidlertid ingen informasjon om hvordan tid brukt på sensur har utviklet seg, for eksempel om fordelingen av tid brukt på intern versus ekstern sensur har endret seg.

5 **Sensur, læring og vurdering**

I dette kapitlet analyserer vi det empiriske materialet med utgangspunkt i studielederne og tilsynssensorenes vurderinger og erfaringer knyttet til sensorordninger og ressurser til læring og vurdering. Vil økt bruk av ekstern sensur flytte ressurser fra læring til vurdering? I andre del av kapitlet drøftes hvorvidt en gjeninnføring av tradisjonell ekstern sensur fører til begrensninger på hvilke typer vurderingsformer institusjonene vil benytte.

5.1 **Sensorordning og ressurser til læring og vurdering**

Det er tre hovedsyn i spørsmålet om ekstern sensur flytter ressurser fra læring til vurdering. Noen hevder at bruk av ekstern sensur flytter ressurser fra læring til vurdering. Andre fremholder at dette ikke nødvendigvis vil skje. Et tredje synspunkt er at læringsorienterte vurderingsformer er mer ressurskrevende enn bruk av ekstern sensur.

De som hevder at bruk av ekstern sensur flytter ressurser fra læring til vurdering argumenterer for at de prioriterer undervisning og oppfølging av studentene fremfor en ordning med gjennomgående ekstern sensur av eksamensoppgaver. Det argumenteres også med at man ønsker å bruke minst mulig ressurser på vurdering og dermed prioritere læring. Det fremholdes at det er et motsetningsforhold mellom læring og vurdering, og at økonomien ikke tillater at begge deler prioriteres like høyt.

Andre hevder at det ikke nødvendigvis er en motsetning mellom ressurser til vurdering og ressurser til læring. Det blir lagt vekt på at man benytter ekstern sensur på eksamensbesvarelser i kombinasjon med vurdering underveis (eksempelvis mappevurdering). Det argumenteres også med at en del vurderingsformer som muntlig eksamen med ekstern sensur bidrar til læring. Det antatte motsetningsforholdet utfordres.

De som argumenterer for at sammenhengen mellom ressurser til læring og vurdering og ekstern sensur fremholder at underveisvurderinger er mer ressurskrevende enn ekstern sensur på tradisjonell eksamen. Bruk av ekstern sensur og mindre bruk av underveisvurdering ville derfor i følge disse informantene frigi ressurser fra vurdering til læring.

Det er ikke utenkelig at det kan være en sammenheng mellom informantenes syn på forholdet mellom ressurser, læring og vurdering og de ordninger for sensurering de kjenner best. Er det så en sammenheng mellom informantenes syn på ressurser til læring og vurdering og den sensorordningen som er mest utbredt?

De som er enig i påstanden om at bruk av ekstern sensur flytter ressurser fra læring til vurdering bruker stort sett en ordning med hovedsakelig interne sensorer gjerne i kombinasjon med en tilsynssensorordning. Blant dem som ikke er enig i at det er et

motsetningsforhold mellom bruk av ekstern sensor og ressursbruk til vurdering og læring finner vi både studier med gjennomgående ekstern sensur og studier med hovedsakelig intern sensur i kombinasjon med en tilsynssensorordning. I den siste gruppen som argumenterer for at sammenhengen er motsatt (undervisvurderinger er mer ressurskrevende enn bruk av ekstern sensur) finner vi både studier som i dag bruker gjennomgående ekstern sensur og studier som hovedsakelig bruker interne sensorer. Det er altså ikke en systematisk sammenheng mellom informantenes syn på spørsmålet om bruk av ekstern sensur vil flytte ressurser fra læring til vurdering og valg av sensorordning.

Blant tilsynssensorene er det to hovedsyn på problemstillingen om bruk av ekstern sensur flytter ressurser fra læring til vurdering. Noen hevder at det ikke nødvendigvis er et motsetningsforhold mellom ressurser til vurdering og ressurser til læring. Andre hevder å ikke ha grundig nok kjennskap til ressurser knyttet til læring og vurdering i faget de er tilsynssensor i til at de kan ha en kvalifisert mening om økt bruk av ekstern sensor vil flytte ressurser fra læring til vurdering.

Heller ikke blant tilsynssensorene er det et en til en forhold mellom synet på ressurser til læring og ressurser til vurdering relatert til bruk av ekstern sensur og den aktuelle sensorordningen ved studiet man er tilsynssensor for.

5.2 Ekstern sensur og typer vurderingsformer

Den siste problemstillingen vi drøfter er hvorvidt en gjeninnføring av tradisjonell ekstern sensur føre til begrensninger på hvilke typer vurderingsformer institusjonene vil benytte? En konklusjon i Solum (2005) var at det var stor variasjon i sensurpraksis mellom læresteder og mellom fag. Det er observasjon som synes å være like gjeldende i 2008. Variasjonen i sensorordninger ser ikke å ha blitt noe mindre i løpet av denne perioden. En tendens til endring kan likevel være verdt å fremheve, bruk av tilsynssensor ser ut til å ha økt noe. Fra å være en ordning som kun fantes ved et lite antall læresteder i 2005, er ordningen med tilsynssensor tilsynelatende mer utbredt i 2008. Kartleggingen dokumenterer imidlertid at ordningen er betydelig mer utbredt ved universitetene enn ved de statlige høyskolene. Vi påpekte i kapittel 3 at lærestedene har etablert ulike kvalitetssikringsfunksjoner av typen tilsynssensor, men også andre betegnelser brukes på denne funksjonen.

5.2.1 Begrunnelser for sensurpraksis og forholdet mellom læring og vurdering

Intervjudataene kan belyse begrunnelser for å endre sensurpraksis. Å opprettholde ”gammel ordning” begrunnes i all hovedsak med at den sikrer en felles nasjonal standard for vurdering av prestasjoner i et fag. Kvaliteten på sensuren brukes som hovedbegrunnelse for ikke å erstatte ”gammel ordning” med en av de nye ordningene. Økonomi er hovedbegrunnelsen for å endre sensorordning bort fra gjennomgående bruk av ekstern sensur. Dette betyr likevel ikke at det oppleves som å fire på kvalitetskravene, enkelte

mener at de nye sensorordningene både gir ”mer for pengene” og ”minst like god kvalitet”. Dette er et rådende syn blant dem som har innført en av variantene av de nye ordningene. Opplevelsen av trang økonomi, i kombinasjon med økt valgfrihet og vurderinger knyttet til pedagogiske hensyn gjør at man mener at en gjeninnføring av ekstern sensur både vil være svært kostnadskrevende, redusere den lokale valgfrihetene og tvinge faget bort fra løsninger som er både pedagogisk begrunnet og kvalitetssikrede.

Vi finner imidlertid ikke at det er et entydig forhold mellom det å velge å beholde ”gammel ordning” og synet på om en gjeninnføring av tradisjonell ekstern sensur vil føre til begrensninger på hvilke typer vurderingsformer det enkelte studium vil benytte. En gruppe av informanter vektlegger at i tillegg til ”gammel ordning”, så benytter man også et bredt spekter av andre vurderingsformer. En gjeninnføring av ekstern sensur vil således ikke med nødvendighet begrense valg av typer vurderingsformer.

En annen gruppe informanter argumenterer ut fra fagspesifikke behov for at man har beholdt vanlig eksamen som vurderes i tråd med gammel ordning. Begrunnelsen er at den kunnskapen og de ferdighetene man har ervervet best vurderes i form av vanlig eksamen. I følge disse informantene vil en gjeninnføring av ekstern sensur heller ikke begrense valg av vurderingsformer, noen fag prøves best ved hjelp av tradisjonell eksamen, andre fag vurderes best gjennom en kombinasjon av vurderingsformer som neppe vil bli påvirket av et direktiv om gjeninnføring av ekstern sensur.

5.3 Konkluderende betraktninger om forholdet mellom sensur, læring og vurdering

Vil en gjeninnføring av tradisjonell ekstern sensur vil føre til begrensninger av hvilke typer vurderingsformer institusjonene vil benytte? En sammenligning av situasjonen og variasjonen i sensorordninger i 2005 og 2008 tyder ikke på mindre variasjonen i sensorordninger i 2008, snarere det motsatte. En tendens kan være verdt å nevne, bruk av tilsynssensor synes å være noe mer utbredt i 2008 enn det var i 2005. Kartleggingen dokumenterer imidlertid at ordningen er betydelig mer utbredt ved universitetene enn ved de statlige høyskolene.

Blant informantene er det tre hovedsyn om hvorvidt økt bruk av ekstern sensur vil flytte ressurser fra læring til vurdering. Noen hevder at økt bruk av ekstern sensur *vil* flytte ressurser fra læring til vurdering. Dette hevdes særlig fra fagmiljø som hovedsakelig bruker interne sensorer, gjerne i kombinasjon med en tilsynssensorordning. Andre er uenige, og hevder at økt bruk av ekstern sensur *ikke nødvendigvis* vil flytte ressurser fra læring til vurdering. Synspunktet finner vi i fagmiljø som benytter tradisjonell ekstern sensur og fagområder som overveiende bruker intern sensur i kombinasjon med en tilsynssensorordning. I følge et tredje synspunkt er det interessante spørsmålet i hvilken grad økt bruk av læringsorienterte vurderingsformer vil flytte ressurser fra vurdering til læring. Disse informantene er uenig i at ekstern sensur vil flytte ressurser fra læring til

vurdering, dessuten hevder de at utviklingen går i *motsatt retning* – ressurser flyttes i større grad fra vurdering til læring.

6 Konklusjon

Hva er viktige kvaliteter ved god sensur? Vår diskusjon av sensorordninger har tatt utgangspunkt i at sensurpraksis er i endring både i Norge og i andre land. Sentrale problemstillinger handler om vurdering for karaktersetting og sertifisering, og om vurdering for "feedforward". Diskusjonen handler også om forholdet mellom fag og administrasjon ved lærestedene. Er sensurpraksis en genuint faglig aktivitet eller kan den også formaliseres og i større grad være et overordnet ansvar for lærestedene?

Sensorordninger bør diskuteres i lys av en rekke hensyn. Etablering av sensorordning innebærer prioriteringer og ivaretagelse av formelle, kommunikasjonsmessige, metodiske og praktiske hensyn ved sensur. Kvaliteter ved god sensur kan diskuteres i lys av disse hensynene.

Før Kvalitetsreformen krevde loven at det skulle være minst en ekstern sensor ved alle prøver og eksamener ved universiteter og høyskoler i Norge. I den nye universitets- og høyskoleloven av 2005 ble dette endret til at det skal være ekstern evaluering av vurderingene eller vurderingsformene (Bernt, 2006: 64). Våre funn tyder på at lærestedene har benyttet den valgfriheten som reformen åpnet for. Kartleggingen viser en fortsatt betydelig bruk ekstern sensur på lavere grad i norsk høyere utdanning. Dette kan sees i sammenheng med at ekstern sensur er en sterk tradisjon i norsk høyere utdanning (Stensaker et al., 2008). Sammenlignet med situasjonen før Kvalitetsreformen har sensurpraksis endret seg. Sensorordningene varierer innen for det enkelte studium; institutt eller avdeling; innenfor faggruppe; mellom faggrupper og mellom læresteder. Den mest brukte sensorordningen er der det kombineres intern sensur og et element av ekstern sensur. Det er mange ulike måter å organisere dette på, og dette kan gjøres i kombinasjon med tilsynssensor. Bruk av tilsynssensor ser ut til å ha økt noe sammenlignet med situasjonen i 2005. Likevel ser vi at det er den tradisjonelle sensurordningen som i større utstrekning er koblet opp til lærestedenes kvalitetssikringssystem fremfor nye ordninger, som kombinasjonen intern sensur med tilsynssensor. Dette kan bety at sensur fortsatt er fagmiljøenes domene i både forvaltningsmessig og utviklingsmessig forstand.

Den sentrale begrunnelsen for endringer av sensurpraksis er hovedsakelig økonomisk, og ikke kvalitetsmessig. Det betyr ikke at fagmiljøene som har valgt andre løsninger enn "gammel ordning" gjennomgående er usikker på kvaliteten i den sensurpraksis de har. Men endringene av sensurpraksis er altså hovedsakelig økonomisk begrunnet. I en slik situasjon har det vært avgjørende for fagmiljøene å sikre kvaliteten på sensuren. Enkelte mener også at kvaliteten på sensuren har blitt styrket gjennom de nye ordningene.

Det må også fremheves at de økonomiske vurderingene av sensurpraksis hovedsakelig legger vekt på de direkte kostnadene knyttet til ekstern sensur. Kostnader knyttet til ulike interne ordninger trekkes i mindre grad med i vurderingene. Vi finner heller ikke en entydig sammenheng mellom sensorordning og ressurser til læring eller ressurser til

vurdering. Det varierer mellom fagmiljøene hvordan man ser denne sammenhengen. Det betyr at det ikke er mulig å konkludere entydig i spørsmålet om ekstern sensur flytter ressurser fra læring til vurdering. Dette kan sees i lys av de to eksamensregimene som Dysthe et al. (2006) diskuterer, der det i følge den ene tilnærmingen er et klart motsetningsforhold mellom læring og vurdering, mens i følge den andre tilnærmingen dreier læring og vurdering seg om samme sak. Vi finner tilhengere av ulike ståsteder i debatten om læring og vurdering ved norske læresteder i dag. Dette vil igjen påvirke hvordan sentralt dirigerte sensorordninger vil bli iverksatt.

De kommunikasjonsmessige hensynene ved sensorordninger er også oppe til debatt i fagmiljøene. I intervjumaterialet kommer det frem at skillet mellom enesensur og det å sensurere i fellesskap (to eller flere personer) er vel så viktig som skillet mellom ekstern og intern sensur. I fagmiljøene legges det vekt på at enesensur medfører et stort ansvar som vitenskapelig ansatte er seg svært bevisst. Enkelte fagmiljø går man nå bort fra denne praksisen.

Det mest utfordrende spørsmålet er hva variasjonen vi finner i sensurpraksis betyr for kvaliteten på sensuren. I fagmiljøene synes det å være en klar oppfatning at ekstern sensur sikrer kvaliteten på sensuren. I de tilfellene der man har endret sensurpraksis i retning av mer intern sensur og bruk av tilsynssensur står hensynet til kvalitetssikring sterkt. Ekstern sensur vurderes å sikre nasjonal standardisering av vurderingene. Samtidig uttrykkes en skepsis mot å oppfatte ordninger med ekstern sensur som det *eneste* rette når det gjelder å sikre kvalitet på sensureringsarbeidet. Denne diskusjonen tyder på at fagmiljøene foretar avveininger av de ulike metodiske aspektene ved sensurering.

Vi returnerer nå til spørsmålene i oppdraget: Vil økt andel tradisjonell ekstern sensur føre til bedre kvalitet på sensuren? Vi har ikke kunnet undersøke sammenhengen mellom karactersetting og sensorordning på en måte som gir grunnlag for robuste konklusjoner. I fagmiljøene er det en utbredt oppfatning at ekstern sensur på best måte sikrer kvaliteten på sensuren. Det betyr imidlertid ikke at de fagmiljøene som velger andre løsninger ønsker seg tilbake til et system med gjennomgående ekstern sensur.

Hva er de økonomiske konsekvensene av økt bruk av ekstern sensur? På den ene siden kan vi anta at det kan ha store økonomiske konsekvenser å øke bruken av ekstern sensur. Dette er basert på at fagmiljøenes tilpasninger har vært motivert av trang økonomi.

Hovedbegrunnelsen for å endre sensurpraksis har vært økonomisk. Det er hovedsakelig utgifter til honorarer og reise- og oppholdskostnader som tas med i betraktningen. På den andre siden fokuseres det i liten grad på ressursbruken internt knyttet til sensur, slik at det er vanskelig å vurdere de samlede kostnadene. Undersøkelser av vitenskapelige ansattes tidsbruk knyttet til eksamensarbeid gir ikke grunnlag for å konkludere med redusert tidsbruk på disse oppgavene før og etter Kvalitetsreformen (Frølich et al., 2008). Men basert på disse dataene er det ikke mulig å skille ut tidsbruk til kun sensurarbeidet, og dermed fastslå hvordan tid til sensurarbeidet har utviklet seg etter reformen. Hovedvekten

av fagmiljøene rapporterer å bruke en kombinasjon av ekstern og intern sensur på lavere grad, og flere synes å ha utviklet tilpassede ordninger for dette. En videreføring av denne problemstillingen kan derfor være å øke kunnskapen om den store variasjonen av sensorordningene som i dag finnes ved de ulike lærestedene. Dette kan gi mulighet for å åpne for mer kompetanseoverføring mellom læresteder og fagområder med hensyn til god praksis på dette feltet. Er det for eksempel læresteder som har utviklet lokale sensorordninger som ivaretar både økonomiske og kvalitetsmessige hensyn på en god måte? Hva er den optimale kombinasjonen av ekstern og intern sensur innenfor ulike fagområder eller enkeltfag?

Vil økt bruk av ekstern sensur flytte ressurser fra læring til vurdering? Vi finner ingen entydig sammenheng mellom ressurser til læring og ressurser til vurdering avhengig av sensorordninger. Det varierer mellom fagmiljøene hvilke vurderingsformer som betraktes som mest hensiktsmessige. Mange har vært opptatt av å styrke undervisningen og bruke mindre ressurser på vurdering og dette har vært en viktig begrunnelse for endring av sensurpraksis. Vi finner også fagmiljø som ikke betrakter det som et motsetningsforhold; man kan ha opprettholdt ”gammel ordning” og også tatt i bruk nye vurderingsformer.

Vil en gjeninnføring av obligatorisk ekstern sensur føre til begrensninger på hvilke vurderingsformer lærestedene vil benytte? Vi påviser at lærestedene bruker et mangfold av vurderingsformer. Bildet synes mer variert enn før Kvalitetsreformen da særlig humanistiske og samfunnsvitenskapelige fag samt jus i større grad benyttet tradisjonell eksamen som vurderingsform. Siden endringene ser ut til å være motivert hovedsakelig av økonomiske grunner, kan det innebære at en gjeninnføring av obligatorisk ekstern sensur vil redusere dette mangfoldet. Men, spørsmålet handler kanskje i vel så stor grad om en gjeninnføring av tradisjonell ekstern sensur vil være den mest hensiktsmessige løsningen for å sikre vurderingsformer med høy kvalitet ved lærestedene. Flere synes å ha innvendinger til en slik løsning. Skillet mellom sensur foretatt av en person alene eller av flere sensorer i samråd fremheves som en vel så viktig utfordring som skillet mellom ekstern og intern sensur.

Tabell- og figuroversikt

Tabeller

Tabell 3-1	Omfang av ulike typer sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde. Prosent.	25
Tabell 3-2	Mest utbredt sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde. Prosent.....	28
Tabell 3-3	Gjennomsnittlig andel av eksamener som rettes av ekstern sensor innenfor hvert fagområde, i tilfeller der ekstern sensor brukes på noen av eksamener ved 31 læresteder. Prosent.....	30
Tabell 3-4	Andel sensorordninger på lavere grads studier knyttet opp til lærestedets kvalitetssikringssystem ved 31 læresteder etter fagområde. Prosent.	31
Tabell 4-1	Andel faste vitenskapelige ansatte ved universitetene som oppga å bruke tid på eksamensarbeid i 1990 og 2000 etter faggruppe. Prosent. Gjennomsnittlig antall dager for de som deltok i parentes.....	40
Tabell V- 1	Oversikt over utvalget av informanter	56

Figurer

Figur 3.1	Omfang av ulike typer sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde og lærestedstype. Prosent.	27
Figur 3.2	Mest utbredt sensurpraksis i 2008 på lavere grads studier ved 31 læresteder etter fagområde og lærestedstype. Prosent.	29
Figur 3.3	Andel sensorordninger på lavere grads studier knyttet opp til lærestedets kvalitetssikringssystem ved 31 læresteder etter fagområde og lærestedstype. Prosent.	32

Litteratur

- Aamodt, P. O., Prøitz, T. S., Hovdhaugen, E., & Stensaker, B. (2007). *Læringsutbytte i høyere utdanning. En drøfting av definisjoner, utviklingstrekk og måleproblemer*. Oslo : NIFU STEP Rapport 40/2007.
- Bernt, J. F. (1996). *Lov om universiteter og høyskoler*. Bergen : Alma Mater.
- Bernt, J. F. (2006). *Universitets- og høyskoleloven av 2005. Med kommentarer*. Bergen : Fagbokforlaget.
- Black, P., & Wiliam, D. (2003). "In praise of educational research": formative assessment. *British Educational Research Journal*, 29(5).
- Bloom, B. S., Hasting, J. T., & Madaus, G. F. (Eds.). (1971). *Handbook in the formative and summative evaluation of student learning*. New York : McGraw-Hill.
- Dysthe, O., Raaheim, A., Lima, I., & Bygstad, A. (2006). *Undervisnings- og vurderingsformer. Pedagogiske konsekvenser av Kvalitetsreformen*. Delrapport 7.
- Frølich, N., & Waagene, E. (2008). *Administrative arbeidsoppgaver for det vitenskapelige personalet ved universiteter og høyskoler i lys av Kvalitetsreformen*. Oslo : NIFU STEP Rapport 10/2008.
- Hannan, A., & Silver, H. (2006). On being an external examiner. *Studies in Higher Education*, 31: 57-69.
- Hersh, R. H. (2004). Assessment and Accountability: Unveiling the student learning factor in higher education. Paper presented at the AAHE National Assessment Conference, Denver.
- Jacobsen, A., & Lauvås, P. (2001). *Examen - eller hvad. Former for summativ evaluering i professionsuddannelser*. Fredriksberg : Samfundslitteratur.
- Lauvås, P. (2000). "Eksamen" innen høgre utdanning i Norge. Vedlegg 13 til NOU 2000: 14 (Mjøs-utvalgets innstilling). Oslo.
- Lauvås, P., Havnes, A., & Raaheim, A. (2000). Why this inertia in the development of better assessment methods? *Quality in Higher Education*, 6(1): 91-99.
- Nilsson, K.-A., & Näslund, H. (1997). *Extern medverkan i examinationen. Nordiska og brittiska traditioner. Svenka försök*. Stockholm : Högskoleverket.
- NOU. 2000: 14. Frihet med ansvar. Om høgre utdanning og forskning i Norge (Mjøs-utvalgets innstilling). *Norwegian Official Report on the Quality Reform*.
- OECD. 1998. Redefining tertiary education. Paris : OECD.
- Silver, H., Stennet, A., & Williams, R. (1995). *The external examiner system. Possible futures*. Report of a project commissioned by the Higher Education Quality Council. London : Quality Support Centre.
- Smeby, J.-C. (1993). *Undervisning ved universitetene*. Oslo : Norges Forskningsråd.
- Solum, N. H. (2005). *Sensorordninger i høyere utdanning - kartlegging av status og utviklingstrekk ved 10 institusjoner*. Oslo : NIFU STEP Arbeidsnotat 17/2005.
- Stensaker, B., Brandt, E., & Solum, N. H. (2008). Changing systems of external examination. *Quality Assurance in Education*, 16: 211-223.
- Vabø, A. (2000). *Eksamens- og evalueringsformer i høyere utdanning: en sammenlignende studie. En rapport for Mjøs-utvalget og for KUF*. Oslo : NIFU Skriftserie 7/2000.

Vedlegg 1 Kartlegging av sensorordninger

NIFU STEP gjennomfører på oppdrag for Kunnskapsdepartementet en kartlegging av lærestedenes praksis knyttet til sensur og sensorordninger. Kvalitetsreformen medførte endringer i bruken av ekstern sensor på lavere grad i Norge. Kravene om obligatorisk ekstern sensur ble gjort om, og lærestedene sto i utgangspunktet fritt til å vurdere omfang og bruk av ekstern sensor på dette nivået.

NIFU STEPs kartlegging gjelder tre forhold:

- Hvilke typer av sensorordninger som nå eksisterer på lavere grads studier fakultets/avdelingsvis.
- I hvilket omfang de ulike typene sensorordninger er fordelt på fakultet/avdeling og typer av læresteder.
- Om sensorordningene er formalisert i forskrifter og eventuelt koplet til institusjonenes kvalitetssikringssystem med sikte på variasjoner mellom fag og typer av læresteder.

Sensorordninger kan variere både mellom fakultet/faggrupper og innenfor den enkelte faggruppe. Hensikten med kartleggingen er både å få en oversikt over **variasjon/bredden** i bruk av sensorer innenfor de ulike faggruppene som finnes på det enkelte lærested, og få kartlegge hvilke sensorordninger som er **mest brukt** innenfor hver enkelt faggruppe.

Oppdraget består både av en kartleggingsdel og en rekke personlige intervju med studiedirektører og tilsvarende, studieledere og tilsynssensorer.

I første omgang gjennomfører vi kartleggingsdelen av undersøkelsen. Henvendelsen retter vi til studiedirektører, studiesjefer, utdanningsdirektører og tilsvarende ved alle statlige høyskoler, universitet og vitenskapelige høyskoler. Under finner du fire spørsmål til å besvare i denne sammenheng. Svar gis selvsagt kun for de faggruppene som finnes ved ditt lærested. Spesifiser gjerne hvilke fakultet/avdelinger som inngår i kategorien ”annet”.

1. Hvilke typer sensurpraksis forekommer på lavere grad ved ditt lærested?

Marker med **X** i tabellen under. (Flere kryss er mulig.)

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Bare intern(e) sensorer						
Intern(e) sensorer, ekstern(e) på noen av eksamenene						
Bare eksterne sensorer						
Tilsynssensor						

2. Innenfor hver faggruppe, hvilken sensorordning er den mest brukte på lavere grad?

Marker med **X** i tabellen under.

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Bare intern(e) sensorer						
Intern(e) sensorer, ekstern(e) på noen av besvarelsene						
Bare eksterne sensorer						
Tilsynssensor						

3. I tilfelle ekstern sensor brukes på enkelte av eksamenene, kan du innenfor hver faggruppe angi prosentandelen av eksamenene som rettes av ekstern sensor?

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Ekstern sensor i prosentandel av besvarelsene						

4. Innenfor hver faggruppe, hvilke av de ulike sensorordningene er knyttet opp til lærestedets kvalitetssikringssystem?

Marker med **X** i tabellen under de ordningene som er knyttet til kvalitetssikringssystemet

	Faggrupper					
	Helse- og sosialfag	Mat. nat. ingeniør- og tekniske fag	Samfunnsfag	Humanistiske fag	Pedagogikk og lærerutdanning	Annet
Kun intern sensur						
Intern sensur, ekstern på noen besvarelser						
Ekstern sensur						
Tilsynssensor						

Har du kommentarer til lærestedets bruk av sensorordning? Skriv gjerne inn kommentarer her:

Vedlegg 2 Intervjuguide

NIFU STEP gjennomfører på oppdrag for Kunnskapsdepartementet en kartlegging av lærestedenes praksis knyttet til sensur og sensorordninger. Kvalitetsreformen medførte endringer i bruken av ekstern sensor på lavere grad i Norge. Kravene om obligatorisk ekstern sensur ble gjort om, og lærestedene sto i utgangspunktet fritt til å vurdere omfang og bruk av ekstern sensor på dette nivået.

Vi har gjennomført en kartlegging av tre forhold:

- Hvilke typer av sensorordninger som nå eksisterer på lavere grads studier fakultets/avdelingsvis.
- I hvilket omfang de ulike typene sensorordninger er fordelt på fakultet/avdeling og typer av læresteder.
- Om sensorordningene er formalisert i forskrifter og eventuelt koplet til institusjonenes kvalitetssikringssystem med sikte på variasjoner mellom fag og typer av læresteder.

I tillegg til kartleggingen gjennomfører vi en rekke personlige intervju med studiedirektører, studieledere og tilsynssensorer. I disse intervjuene ønsker vi å belyse tre hovedspørsmål:

1. Sammenhengen mellom sensorordning og kvalitet
2. Sammenhengen mellom sensurpraksis og kostnader
3. Sammenhengen mellom sensur, læring og vurdering

Sammenhengen mellom sensorordning og kvaliteten på sensuren

- Hva kjennetegner etter din mening høy kvalitet på sensur?
- Vi kan tenke oss stor variasjon i sensurpraksis, i kartleggingen har vi skilt mellom tre hovedformer (bare interne sensorer; interne sensorer, men eksterne sensorer på en del av besvarelsene; bare eksterne sensorer), hva er dine vurderinger av spørsmålet om det er en sammenheng mellom kvaliteten på sensuren og sensorordning?

	bare interne sensorer	interne sensorer, men eksterne sensorer på en del av besvarelsene	bare eksterne sensorer
Kvalitet på sensuren			

- Kan vi dokumentere en slik sammenheng?

Sammenhengen mellom sensurpraksis og kostnader

- Hvilken kunnskap eller oversikt over kostnader knyttet til de forskjellige sensurformene har dere?
- Hva er det mulig å si om sammenhengen mellom ulike sensurpraksiser og kostnader?
- Gjør dere dere vurderinger om kostnader ved valg av sensorordning?

Sammenhengen mellom sensur, læring og vurdering

- Hva tenker dere om sammenhengen mellom sensur, læring og vurdering?
- Flytter en ordning med ekstern sensur ressurser fra læring til vurdering?
- Legger en ordning med ekstern sensur begrensinger på vurderingsformene man benytter?

Vedlegg 3 Oversikt over informanter

Tabell V- 1 Oversikt over utvalget av informanter

Utvalg informanter	Sentralt nivå	Mat.nat- og ingeniørfag	Økonom i	Sykepleie	Språk fag	Jus	Pedagogiske fag	Sum per lærested
Universitetet i Bergen (UiB)	1				2	1	2	5
Universitetet i Stavanger (UiS)	1	2		1				3
Universitetet i Oslo (UiO)	1	2			2			4
Høgskolen i Oslo (HiO)	1	1						1
Høgskolen i Bodø (HBO)	1			1			1	2
Høgskolen i Sogn og Fjordane (HSF)	1		1					1
Høgskolen i Østfold (HiØ)	1			1	1			2
Norges Handelshøyskole (NHH)	1		1					1
Høgskolen i Lillehammer (HiL)	1		1					1
Høgskolen i Nesna (HiNe)	1							
Sum:	10	5	3	3	5	1	3	20
Total sum (alle grupper):	30							