



**Nasjonal evaluering av høyere utdanning:
Bakgrunn, effekter og fremtidig organisering**

Rita Karlsen
Bjørn Stensaker

Forord

Dette notatet tar sikte på å gi en foreløpig oversikt over pilotprosjektet "Nasjonal evaluering av høyere utdanning" som avsluttes i løpet av 1996. Notatet er delt i tre: Første del omhandler prosjektets målsetting, prosess og omfang. Andre del skisserer de effekter prosjektet foreløpig har hatt ved å oppsummere de rapporter institusjonene selv har sendt inn til Kirke-, utdannings- og forskningsdepartementet. Tredje del antyder en mulig fremtidig organisering av denne type aktivitet, basert på erfaringer man har høstet gjennom prosjektperioden.

DEL I	PROSJEKTETS MÅLSETTING, PROSESS OG OMFANG	3
	Pilotprosjektets målsettinger	3
	Prosessens videre forløp	3
	Pilotprosjektets omfang og kostnader	5
DEL II	PROSJEKTETS EFFEKTER	6
	Ressurs- og stillingsendring	7
	Kobling til ordinært plan- og styringsarbeid	7
	Endringer i organisasjonskultur og samarbeidsklima	8
	Andre tiltak og utviklingsprosjekt	10
	Evaluering og Norgesnett	11
	Forbedring i metoder og organisering	11
	Erfaringer med evaluering generelt	12
	Konklusjon	12
DEL III	FREMTIDIG ORGANISERING	13
	Ulik grad av systematikk	13
	Utgangspunkt for modellene	14
	Modell 1: "Kvalitetskontroll-modellen"	16
	Modell 2: "Kvalitetsutviklings-modellen"	17
	Modell 3: "Behovs-modellen" (modell 1 + 2)	18
	Vurdering av de tre modellene	19
LITTERATUR		20

DEL I PROSJEKTETS MÅLSETTING, PROSESS OG OMFANG

Pilotprosjektets målsettinger

Initiativet til de nasjonale evalueringer kan tilbakeføres til St.meld. Nr. 40 (1990-91) som etterlyste at institusjonene på det daværende tidspunkt (1990), gjerne i samarbeid, utarbeidet egne evalueringsopplegg og vurderingsrutiner. Stortingsmeldingen understreket at dette var et ansvar som institusjonene hadde, men det ble påpekt at det var et behov for en nasjonal støtteaktivitet til slike lokale initiativ. De nasjonale evalueringene skulle altså komme *i tillegg* og være en *støtte* for tiltak som blir ivaretatt på institusjonsnivå. Det ble også understreket fra departementets side at departementet ikke vurderte seg selv som sentral bruker av evalueringresultatene.

Målet for de nasjonale evalueringene var å *“å få kvalifiserte vurderinger av hva som særpreger de utdanningstilbudene som gis, sterke og svake sider ved disse i en nasjonal og internasjonal kontekst og ikke gi forenklete rangordninger av institusjonene. Vurderingene bør munne ut i en beskrivelse av tiltak som kan gjennomføres for å bedre utdanningstilbud og læringsmiljø”* (St. Meld. Nr. 40 1990-91:57). Stortingsmeldingen understreket videre at *“en av hensiktene må være å utvikle et metodeapparat som også kan utnyttes av andre fagmiljøer”* (op.cit). Opplegget for evalueringene, dvs. at det skal være en selvevaluering, samt en ekstern evaluering, og at NIFU (tidligere NAVFs utredningsinstitutt skulle ha ansvaret for å administrere og lede virksomheten, samt innhente erfaringer fra prosjektet, ble også skissert i meldingen (Jordell 1992).

Prosessens videre forløp

Ut fra det utgangspunkt som ble lagt Stortingsmeldingen satte NIFU i gang evalueringsarbeidet med planlegging høsten 1991. Evaluering av det første fagområdet, økonomisk administrativ utdanning, startet i 1992. I en rapport utarbeidet i 1992 ble forløpet og prosessen for evalueringen av økonomisk-administrativ utdanning skissert (Jordell 1992). Opplegget herfra dannet utgangspunktet for alle senere evalueringer. Noe forenklet kan opplegget for evalueringene fremstilles som følgende:

1. Oppstartskonferanse for evalueringen. Her møtes representanter for institusjonene som skal evalueres, evaluatorene, samt representanter for NIFU

og departementet. Hensikten med konferansen er å informere om prosessen, gi institusjonene retningslinjer for hvordan selvevalueringene skal utføres, samt diskutere og presisere det kriteriegrunnlag den enkelte disiplinevaluering skal ta utgangspunkt i. En liste over generelle momenter, som blir tilpasset det enkelte fagområde, er grunnlaget for disse diskusjonene.

2. Den påfølgende selvevalueringen er institusjonenes eget ansvar. Hensikten med denne prosessen er både å stimulere til interesse for kvalitetsarbeid lokalt, som å utarbeide en rapport som skal danne utgangspunktet for den eksterne evalueringen. I en egen rapport (Karlsen og Stensaker 1993) har NIFU systematisert og vurdert erfaringene med selvevalueringer, og antydnet nærmere hvordan slike evalueringer bør utføres (erfaringsgrunnlaget for denne rapporten ble hentet fra øk.adm evalueringen).
3. I en egen konferanse for de eksterne evaluatorene får de eksterne evaluatorene "opplæring" i hvordan evalueringsarbeidet bør/skal foregå. Dette skjer ved hjelp av andre evaluører, og på grunnlag av de systematiserte erfaringer NIFU sitter med. En egen rapport om hvordan eksterne evalueringer foregår og hvordan en slik prosess kan/bør utformes er også utarbeidet, med utgangspunkt i erfaringer fra øk.adm (Stensaker og Karlsen 1994).
4. De eksterne evaluatorene utfører sine besøk på institusjonene (etter å ha studert institusjonenes selvevalueringsrapporter) og utarbeider sine rapporter. Rapportene oversendes departementet (som trykker dem) og NIFU (som sender dem ut til alle institusjoner som har deltatt i evalueringen).

Totalt er tidsrammen for en slik prosess rundt 18 måneder. Siden de nasjonale evalueringer er et pilotprosjekt hvor en del av hensikten er å utvikle et metodeapparat og gode evalueringsprosedyrer, justeres evalueringene, i samarbeid med institusjonene og med de eksterne evaluatorene, etter hvert som nye erfaringer gjøres. Eksempler på justeringer som er gjort underveis er innskjerping og (delvis) standardisering av det kvantitative datagrunnlaget som institusjonene skal samle inn og analysere, større vektlegging på det ansvaret institusjonen har for at evalueringsprosessen holder tidsskjemaet (f.eks eget brev til ledelsen der de gjøres direkte ansvarlig for overholdelsen av tidsfrister), endringer i organiseringen av selvevalueringene (f.eks krav om opprettelse av prosjektgruppe), innskjerping av kravene til at de eksterne evaluatorene (f.eks at det skal fremmes konkrete tiltak). Evalueringene har avdekket at mange forutsetninger for en vellykket evaluering

ikke bestandig er tilstede på institusjonene (se nærmere omtale om dette under prosjektets effekter - del II).

Pilotprosjektets omfang og kostnader

Tabell 1 gir en total oversikt over antall institusjoner og evaluatorene (eksterne sakkyndige) som har deltatt i evalueringsarbeidet:

Tabell 1 Antall institusjoner og evaluatorene involvert i nasjonal evaluering

	Øk.adm	Sosiologi	Ing	Matte	Musikk	Totalt
Institusjoner ¹	47	7	24	24	13	115
Evaluatorene	25	4	16	13	8	66

Fagområdet som det gjenstår å avslutte i løpet av 1996 er musikk. I 1994-95 påtok NIFU seg også ansvaret for administrasjonen og gjennomføringen av det såkalte EU-prosjektet, som hadde til hensikt å utvikle et metodeapparat for evaluering av høyere utdanning i en større europeisk kontekst (noe arbeid forløper også i 1996 med oppfølgingsarbeid i den såkalte Advisory Group). For de tre disiplinene som er ferdigstilt fordeler kostnadene seg slik²:

Tabell 2 Nøkkeltall (oppsett i 1000 kr) for øk.adm, sosiologi og ingeniøruvalueringene³

	Øk.adm	Sosiologi	Ingeniør
Driftsutgifter	2.795	832	1.650
Administrasjon/noe dokumentasjon og formidling	1.323	257	578
Kostnader pr. institusjon	88	156	93

Tabell 2 viser at i de tre evalueringene som er ferdigstilt (økonomisk administrativ

¹ Antall miljøer innen hver institusjon som har deltatt i evalueringene er høyere enn 115 fordi en evaluering ofte kan involvere mer enn ett fagmiljø på en institusjon. Dette ble særlig tydelig etter reorganiseringen av høyskolene i 1994.

² Regnskapet over evalueringen i matematikk er ikke ferdigstilt.

³ Kostnader som har påløpt til selvevaluering på den enkelte institusjon er ikke medregnet.

utdanning, sosiologiutdanningen og ingeniørutdanningen) ble de omtrentlige kostnadene pr institusjon henholdsvis 88.000 kr, 156.000 kr og 93.000 kr.

Driftsutgiftene i forbindelse med evalueringene er stort sett knyttet opp til honorarer og reiseutgifter for de eksterne evaluatorene. Kostnadene i forbindelse med administrasjon/noe dokumentasjon og formidling er utgiftene knyttet til NIFUs arbeid på prosjektet. En vesentlig forskjell mellom nasjonale evalueringer i Norge og f.eks Danmark er størrelsen på honoraret til de sakkyndige. I Danmark gis et symbolsk honorar til sakkyndige på ca. 7.000 dkr, mens vi i Norge opererer med et honorar på kr. 40.000 pr sakkyndige (grunnlaget for beregningen er et omfang på ca 1,5 mnd.verk i daværende lønnstrinn 26). Honorarforskjellen skyldes at det i Norge gis liten bistand til de sakkyndige, mens det i Danmark gis betydelig bistand (f.eks er det ansatte ved Evalueringscenteret som skriver rapportene).

I Norge er institusjonenes tidsbruk i forbindelse med evalueringsarbeidet registrert i evalueringen av økonomisk-administrativ utdanning: Hver institusjon som hadde etablert en prosjektgruppe for selvevalueringen brukte i snitt 221 timer. De institusjoner som ikke hadde en prosjektgruppe tilnyttet prosjektet brukte mindre tid, i snitt 119 timer. Imidlertid var det store variasjoner mellom institusjonene (se Karlsen og Stensaker 1993). Kostnadene for prosjektet blir altså høyere hvis institusjonenes arbeid inkluderes i regnskapet.

DEL II PROSJEKTETS EFFEKTER

Denne delen av notatet bygger hovedsakelig på institusjonenes tilbakerapportering til departementet høsten 1995, Karlsen & Stensaker (1993), Stensaker & Karlsen (1994), samt Stensaker (1996).

For å si noe om prosjektets effekter må målsettingene konkretiseres. Som nevnt tidligere, kan målsettingene deles i to: For det første er pilotprosjektet en reell evaluering der formålet har vært å forbedre utdanningstilbud og læringsmiljø, f.eks gjennom selvutvikling på institusjonsnivå. For det andre skal pilotprosjektet systematisere de erfaringer evalueringene bidrar til å opparbeide, samt forbedre metoder og organisering av denne virksomheten (metaevaluering). Våre erfaringer tilsier at forbedringer i utdanningstilbud og læringsmiljø kan illustreres gjennom fire deleffekter (jfr. Stensaker 1996): Ressurs og stillingsendring, kobling til ordinært plan- og styringsarbeid, endringer i organisasjonskultur og samarbeidsklima, samt andre tiltak og utviklingsprosjekt.

Ressurs- og stillingsendring

Mange miljøer hevder at kvalitetsutvikling nødvendigvis må innebære omfordeling av ressurser, tilføring av ressurser, eller endringer i stillingsstruktur e.l. Følgende tiltak er rapportert inn fra forskjellige institusjoner (institusjoner nevnt i parentes er trukket frem som eksempler):

Ressurser

- Som en konsekvens av evalueringen fikk instituttet tildelt 50.000,- kroner som faglig stimulering for å følge opp de foreslåtte tiltak. Det poengteres at disse "honoreringsmidlene" har gitt vellykkede resultater (NTNU).
- Avsatt 100.000 kroner til handlingsplan for studiekvalitet (Høgskolen i Finnmark).
- Nyttet til omfordeling av ressurser internt (eks. Høgskolen i Bergen, Høgskolen i Gjøvik).
- Utskiftning av utstyr (eks. Høgskolen i Bergen, Høgskolen i Sør-Trøndelag).
- Avsetter ressurser til oppfølgingsarbeid som målbevisst, organisert arbeid med kvalitetsutvikling (eks. Høgskolen i Buskerud).

Stillinger

- Instituttet har merket seg en økende interesse for ledige stillinger, noe de tillegger som en konsekvens av evalueringen ved den offentlige omtalen denne har fått i faglige fora (eks. NTNU).
- Tilført nye stillinger til det evaluerte faget (eks. Høgskolen i Nord-Trøndelag, Høgskolen i Sør-Trøndelag).
- Økt fokus på å øke antall doktorander (eks. Høgskolen i Vestfold).

Kobling til ordinært plan- og styringsarbeid

Kvalitetsarbeid er en prosess som det er viktig å integrere i institusjonenes ordinære virksomhet. Å koble oppfølging av evalueringene til ordinært plan- og styringsarbeid kan derfor sees på som et viktig resultat av virksomheten. Følgende type aktiviteter kan skisseres:

Integrert i planarbeid

- I forhold til strategiplanen har instituttet listet opp de temaområder de ønsker å profilere seg innen (eks. Høgskolen i Bergen).
- Arbeidet nedfelt i plan- og målstyringsdokumenter (strategiplan og årsplan) (eks. UiO, Høgskolen i Bergen, Høgskolen i Bodø, Høgskolen i Buskerud, Høgskolen i Finnmark, Høgskolen i Gjøvik, Høgskolen i Lillehammer, Høgskolen i Stavanger, Høgskolen i Telemark, Høgskolen i Vestfold, Høgskolen i Østfold, Rogaland Markedshøyskole, Varehandelens høyskole).
- Retningsgivende i omorganiseringsarbeid (eks. NMH, Høgskolen i Bergen, Høgskolen i Lillehammer).
- Medført til endring i personalpolitikken (eks. Høgskolen i Bergen).
- Endret opptaksprofil og strategiformuleringer (eks. Høgskolen i Buskerud, Høgskolen i Østfold, Varehandelens høyskole).
- Medførte til økt internasjonalisering, f.eks. med engasjement i ERASMUS. (eks. Høgskolen i Molde, Høgskolen i Telemark).

Studieplanrevisjon

- Revisjon av studieplanen for hovedfaget (opplegget for hovedoppgaven og den felles teoriundervisningen) og dr.polit studiet (større forskningstyngde, blant annet ved at ett av kursene i forskningsmetode ble byttet ut). I tillegg ble ett nytt kurs etablert (NTNU).
- Undervisningsopplegg og studieplaner gjennomgått på nytt (eks. UiTø, Høgskolen i Agder, Høgskolen i Bergen, Høgskolen i Buskerud, Høgskolen i Nord-Trøndelag, Høgskolen i Sogn og Fjordane, Høgskolen i Stavanger, Høgskolen i Sør-Trøndelag, Høgskolen i Telemark, Høgskolen i Vestfold, Høgskolen i Østfold, Høgskolen i Ålesund, Varehandelens høyskole).

Endringer i organisasjonskultur og samarbeidsklime

Studiekvalitetsutvalget fremholdt i sin utredning at et viktig element for å høyne kvaliteten innen høyere utdanning var å få i stand en "kulturrendring" - der høyere bevissthet rundt denne type arbeid er en viktig effekt. Mange institusjoner rapporterer om denne type effekter:

Kulturendring

- Et mer generelt resultat er at mange av institusjonene melder om at evalueringen førte til at partene ble tvunget til å skjerpe bevisshetsnivået i forhold til eget studie. Dette bl.a. fordi en måtte beskrive fortrinn og problemer i forhold til en nasjonal evaluering (eks. UiTø, Høgskolen i Bergen, Høgskolen i Buskerud, Høgskolen i Gjøvik, Høgskolen i Nord-Trøndelag, Høgskolen i Vestfold, Høgskolen i Ålesund).
- For et land med liten tradisjon for å bedrive evalueringer er en av de effekter som understrekes mest av institusjonene at man har "lært" å evaluere. Man har økt kunnskapen om hvordan slike prosesser kan organiseres, styres, og ikke minst gjennomføres (Høgskolen i Agder, Høgskolen i Telemark, se også Karlsen & Stensaker 1993).
- I en intervjurunde som er utført i januar 1996 fremhever mange institutt- og avdelingsledere at evalueringsvirksomheten har gitt dem anledning til å etablere rutiner som burde vært innført gjennom virksomhetsplanleggingen, men som ikke har vært gjennomført (Høgskolen i Molde, NLH, se også Stensaker 1996).

Etablert kontaktnett

- Enkeltmedlemmer i staben har fortsatt kontakt med medlemmer av den eksterne komiteen, også i forhold til konkrete samarbeidsprosjekter (eks. NTNU, Høgskolen i Agder, Rogaland Markedshøyskole).
- Kontakt med andre institusjoner som arbeider med like prosjekt (eks. UiTø med Høgskolen i Telemark om prosjekt-basert læring/PBL).
- Metoder for økt studentkontakt (eks. Høgskolen i Oslo).
- Evalueringen brukt for å konsolidere eget fagmiljø, knytte kontakt med fagfeller på andre institusjoner (eks. Høgskolen i Ålesund, Varehandelens høyskole, se også Stensaker 1996).

Avholdt seminarer/konferanser

Ved flere av institusjonene har det vært avholdt seminarer eller konferanser i forbindelse med evalueringsarbeidet som betegnes som vellykkede for den videre prosessen. Disse konferansene varierer både i omfang og tema, noen eksempler er:

- Ved sosiologisk institutt ved NTNU rapporteres om vellykket seminar for staben som har ført til prosesser som bringer faget videre.
- På NTNU ble det også avholdt kurs for nytilsatte seminarledere (hovedsakelig hovedfagsstudenter).
- Ved høgsolen i Sogn og Fjordane ble det avholdt konferanser for økt kontakt med det lokale næringsliv.

Andre tiltak og utviklingsprosjekt

Enkelte tiltak og effekter faller ikke under oppdelingen over. Likevel er dette resultater som er viktige å synliggjøre:

Konkret oppfølging

- Nedsatt arbeidsgruppe som skal utarbeide strategi for videre oppfølging av evalueringene (eks. UiTØ).
- Planlegger å iverksette brukerundersøkelse (eks. NLH).

Bedring i statistikk og kvantitativt datamateriale

- Svært mange institusjoner fremhever at man oppdaget at statistikkgrunnlaget for egen virksomhet er for dårlig. Dette gjelder både mht systematisering av de data som finnes, så vel som standardisering og oppfølging på instituttnivå (eks. UiTØ, Høgsolen i Molde, Høgsolen i Vestfold, se også Karlsen og Stensaker 1993).

Som vist over er effektene av evalueringsarbeidet sprikende. Evalueringene har medført endringer både når det gjelder struktur, ressurser, personellutvikling, faglig innhold, og kultur. Det er derfor ikke mulig å identifisere en "typisk" tendens fra evalueringsarbeidet. At effektene er såpass spredte henger bl.a sammen med at institusjonene og disiplinene har ulike definisjoner av hva som bidrar til kvalitetsutvikling, ulike behov med hensyn hva som trengs å kvalitetsutvikles, samt at det er forskjellig om og i hvilken grad evalueringene kobles til annen aktivitet i utdanningssektoren.

På grunn av at evalueringene var initiert av departementet i en tid som av institusjonene oppfattes som turbulent kan man derfor også identifisere enkelte "uintenderte" effekter av evalueringsarbeidet:

Evaluering og Norgesnett/andre strukturprosesser

- Mange institusjoner/ disipliner har antatt at departementet har "skjulte" motiver med evalueringsarbeidet. Frykten for å eventuelt miste studietilbud, ikke få tildelt ønskede funksjoner i Norgesnett er noen av motivene. Slike institusjoner har også prioritert evalueringene høyt og har brukt dem aktivt for å synliggjøre seg selv eksternt. Mange av disse institusjonene kan fremvise resultater fra evalueringene (Karlsen & Stensaker 1993, Stensaker & Karlsen 1994, Stensaker 1996).
- Institusjonene har også i enkelte tilfeller koblet evalueringsarbeidet til andre prosesser som har foregått enten på institusjonen eller i sektoren. For økonomisk administrativ utdanning var det f.eks kobling til Gjærumutvalget sitt arbeid, mens det i de andre utdanningene har vært Norgesnett som har vært sentralt. Dette kan selvfølgelig ha bidratt til å flytte på "evalueringenes fokus", men siden kvalitetsforbedringsarbeid kan ha mange uttrykk behøver ikke en slik forskyvning nødvendigvis å være negativ. Tvert imot kan kobling til andre prosesser gjøre at evalueringsarbeidet får høyere prioritering hos de den angår (Karlsen & Stensaker 1993, Stensaker & Karlsen 1994, Stensaker 1996).

En annen viktig del av pilotprosjektet består i å systematisere de erfaringer evalueringene bidrar til å opparbeide, samt forbedre metoder og organisering av denne virksomheten (metaevaluering). Selv om denne erfaringsinnhenting ikke er avsluttet kan et foreløpig bilde tegnes slik:

Forbedring av metoder og organisering

Siden evalueringene har et formativt formål har endringer i metode og organisering skjedd kontinuerlig i evalueringsperioden. Det som NIFU opplever som essensielt er at:

- Institusjonsledelsen får tildelt et klarere ansvar for virksomheten, selv om prosessen er organisert som en disiplinevaluering.
- Evalueringsarbeidet på institusjonene må følges av ressurser til frikjøp av

personell/arbeidsinnsats.

- Statistikkinnsamling og bearbeiding systematiseres og standardiseres i sterkere grad enn det gjøres i dag.
- Kostnadene i forbindelse med eksterne evaluatorene kan kuttes betraktelig mht honorarer.
- Det opprettes en styrings- eller prosjektgruppe for den enkelte evaluering hvor representanter for ulike interesser er representert.
- Det etableres rutiner/tiltak for å få opinionen interessert i evalueringsarbeidet, for derigjennom å skape et sterkere ytre press.
- Formålet med virksomheten presiseres ytterligere.

Erfaringer med evaluering generelt

- Svært mange norske utdanningsinstitusjoner har gjennom det nasjonale pilotprosjektet fått sine første erfaringer med evaluering som metode (Karlsen og Stensaker 1993). Den metode og de rutiner som evalueringene er bygd på er fullt mulig å tilpasse alle typer utdanningsinstitusjoner/disipliner, og fungerer godt i en norsk kontekst (European Report 1995). Dette tilsier at evalueringsprosedyrene er universelle, dvs at de kan brukes uansett nasjonal kontekst og akademisk kultur.
- Evalueringsarbeid må ha konsekvenser for de miljøer som deltar. Disse konsekvensene må utformes slik at de bidrar til interesse for evaluerings- og kvalitetssikringsarbeid. Konsekvenser må utformes på forhånd slik at dette ikke forårsaker mistanker om "hidden agendas" underveis og i etterkant.
- De erfaringer som er gjort er av en slik art at de kan brukes mer designmessig i forhold til utforming av fremtidige (styrings- og evaluerings)modeller.

Konklusjon

Man skal alltid være forsiktig med å skissere for bastante konklusjoner mht konkrete resultater av en evaluering: Det er bl.a vanskelig å vite om det er evalueringene eller andre forhold som har ført til visse typer endringer og tiltak, samt at man også bør ha en viss skepsis til den rapportering som skjer fra institusjonene til departementet.

I forhold til de målsettinger som St.meld. Nr. 40 (1990-91) skisserte for prosjektet er det vår klare oppfatning at de nasjonale evalueringer absolutt har bidratt til å sette studiekvalitet på dagsorden. Man har etablert en langt større bevissthet rundt arbeid med evaluering og studiekvalitet, man har satt i gang en lang rekke tiltak som kan bidra til lokal kvalitetsheving/sikring, samt at de miljøer som har deltatt har "lært å evaluere". Prosjektets resultater kan således sies å være i tråd med målsettingene.

Likevel bør det advares mot overdrevet optimisme med hensyn til effekter. For det **første** er det flere representanter fra institusjonene som etterlyser en skikkelig *oppfølging* av prosjektet. Hvis slik oppfølging ikke skjer, er det svært mange som antyder at denne type virksomhet lett vil få et symbolsk preg, eventuelt at engasjementet ikke vil være så stort fordi "evalueringen ikke får konsekvenser". For det **andre** er pilotprosjektet en evaluering av disipliner, ikke av institusjoner. Det *institusjonelle ansvaret* for denne virksomheten må tydeliggjøres, og ledelsen ansvarliggjøres. Mye av den erfaring som nå er gjort lokalt er nemlig på institutt- og avdelingsnivå, og ikke nedfelt på institusjonsnivå. Det er god grunn til å etterlyse et sterkere institusjonelt ansvar for denne type arbeid. For det **tredje** eksisterer det fremdeles miljøer hvor evaluering og kvalitetssikring har lav *prioritering*. Årsakene kan være alt fra faglig uenighet, personstridigheter, strukturproblemer til liten tiltro til evalueringer/evalueringsmetoder. Slike miljøer må følges opp for at kvalitetssikring og -utvikling skal bli prioritert.

DEL III FREMTIDIG ORGANISERING

Denne delen av notatet bygger på tidligere nevnte publikasjoner fra NIFU, European Report fra EU-prosjektet, en rapport som er utarbeidet for Nordisk Ministerråd der ulike skandinaviske lands organisering av og erfaringer med evaluering er beskrevet (Smeby 1996), samt en rapport (under arbeid) som på grunnlag av ulike vinklinger på utdanning (organisering av undervisning, ulike mål på kvalitet og relevans ift arbeidslivet) ser på kvalitetsarbeid i høyere utdanning (Karlsen & Stensaker 1996).

Ulik grad av systematikk

En fremtidig organisering av evalueringsvirksomheten kan prinsipielt variere som følger:

Tabell 3 Valgmuligheter mht fremtidig organisering av evalueringer

	Lav grad av systematikk	Høy grad av systematikk
Lokal forankring og bruk (Org.utvikling/læring)	A	B
Sentral forankring og bruk (Kontroll/styring)	C	D

Evalueringermodeller kan ha ulik grad av **systematikk** innebygd. Med høy grad av systematikk menes her at evalueringer foregår rutinemessig, at datainnsamling og bearbeiding standardiseres, samt at formålene med evalueringene er konstante over tid. Evalueringermodeller kan også variere med hensyn til **lokal eller sentral forankring og bruk**. Med sentral forankring og bruk menes her at det er sentrale utdanningsmyndigheter (eller andre tilsvarende aktører) som er de fremste brukerne av evalueringresultatene, samt at evalueringene utføres etter et sentralt initiativ. En sentral forankring vil ofte ha et anstrøk av kontroll/styring. Plasserer vi Norge inn i denne tabellen kan det nasjonale pilotprosjektet sies å passe inn et sted mellom rute A og B. Variasjonen i grad av standardiseringen fremkommer ved at evalueringen er sentralt initiert og har en felles metode, men samtidig har både institusjonene og evaluatorene relativ stor grad av frihet.

Ofte vil kvalitetssystemer prøve å kombinere/balansere ulike hensyn mot hverandre. Sammenligner vi oss med andre europeiske land, vil f.eks det engelske og svenske auditsystemet kunne plasseres i rute D, mens de nasjonale disiplinerevalueringer som også gjennomføres i disse landene kan plasseres i rute A for Sveriges del og i rute C for England. Finland befinner seg et sted mellom rute A og C, og Danmark mellom D og B. En fremtidig organisering av evalueringsevirsomheten kan imidlertid også variere etter hvilke forutsetninger man legger til grunn.

Utgangspunkt for modellene

Alle de modeller som er beskrevet i det etterfølgende har derfor følgende utgangspunkt:

- Institusjonene er pr dags dato ikke kommet langt nok i arbeidet med å etablere egne evalueringrutiner og kvalitetssikringssystemer (jfr. institusjonsrapporter til departementet). Svært mange representanter for institusjonene som har deltatt i de nasjonale evalueringene etterlyser et fortsatt nasjonalt initiativ på

feltet. Å stimulere til et fortsatt lokalt arbeid med kvalitetsutvikling er derfor viktig.

- Institusjonene har fremdeles hovedansvaret for å drive med evaluering og kvalitetssikring lokalt (Jfr. St.meld Nr. 40 1990-91).
- Departementet trenger mer kvalitativ informasjon om institusjonenes virksomhet for å utfylle informasjon fra f.eks DBH. Skal de årlige styringsdialogene mellom departement og institusjoner fungere trengs det både kvantitativ og kvalitativ informasjon. Nasjonale evalueringer kan bidra til å skaffe slik kvalitativ informasjon, og departementet bør derfor definere seg som bruker av evalueringsresultater.
- Departementet trenger også kvalitativ informasjon som utgangspunkt for å fatte mer generelle beslutninger på sine styringsområder (jfr. St.meld. Nr. 40 1990-91:52) f.eks når det gjelder institusjonsmønster, samt faglig innhold, tildeling av grader og fordeling av fag (funksjoner i Norgesnettet). Nasjonale evalueringer kan bidra til å skaffe slik informasjon.
- Evalueringer koster penger. Det er derfor viktig å øke anvendelsen av evalueringene for å redusere kostnadene forbundet med denne type aktivitet.
- Evalueringer er tids- og ressurskrevende. Virksomheten bør derfor legges til en egen organisasjonsenhet/uavhengig virksomhet. At virksomheten legges til et "uavhengig" organ understrekes både av erfaringene fra det europeiske pilotprosjektet (European Report 1995), men også av andre observatører (jfr. f.eks Van Vught & Westerheijden 1993). Det bør vurderes om ikke disse aktivitetene kan legges til et eventuelt Råd for høyere utdanning.

I forhold til eksisterende praksis når det gjelder de nasjonale evalueringene anser vi altså både institusjonene og departementet som viktige brukere av evalueringsresultater, der departementet er en langt viktigere bruker enn hva som formelt er fremkommet pr i dag. Vårt forslag til modeller for fremtidig evalueringer i høyere utdanning, er knyttet opp til det vi mener er realistisk å få gjennomført i Norge. Erfaringene med de nasjonale evalueringer tilsier at fremtidig modeller imidlertid ikke må bryte med den eksisterende "overordnede" kulturen i høyere utdanning (høy grad av autonomi og selvråderett), dvs at evalueringen må ta utgangspunkt i den kultur og virkelighet som eksisterer på utdannings-

institusjonene. Evalueringer som ikke har nødvendig legitimitet vil ikke bli nyttet på institusjonene, læringseffekten vil bli lik null, og det vil være sløsing med tid og ressurser. Det er også lite sannsynlig at prosedyrer som i for stor grad er knyttet opp til effektivitet og kontroll får gjennomslag på institusjonene (jfr. Christensen 1991).

Med dette som utgangspunkt kan følgende modeller skisseres:

Modell 1 : "Kvalitetskontroll-modellen"

Formål:

Stimulere lokalt kvalitetssikringsarbeid, departementet er sentral bruker av resultatene ved kvalitetskontroll.

Beskrivelse av modellen:

Institusjonene foretar selvevalueringer. Det utføres en ekstern "metaevaluering" av hvorvidt det eksisterer kvalitetssikringssystem og evalueringsrutiner på den enkelte institusjon.

Evalueringsobjekt:

Institusjoner/Fakulteter/Institutter/Avdelinger.

Kjennetegn ved modellen:

Høy grad av systematisering (alle institusjoner må gjennomgå den samme "kontrollen". Det kan vurderes om det også må foreligge kriterier for hva som er utgjør "minstestandard" på kvalitetssikringen). Det er høy grad av sentral forankring og bruk.

Fordeler med modellen:

Modellen har forholdsvis små sentrale driftsutgifter. Dette avhenger derimot av om "metaevalueringen" skal foretas av fagpersoner eller av personer som har "inspeksjon" som oppgave. Modellen bidrar til å etablere et sterkt ytre press på institusjonene, slik at egenvurdering og rapportering vil bli gjennomført. I tillegg vil modellen ansvarliggjøre ledelsen på ulike nivåer.

Ulemper med modellen:

Bortsett fra at departementet ville vite hvilke institusjoner som har etablert kvalitetssikringssystemer eller ikke, er informasjonstilgangen og dybden begrenset. Modellen har få elementer av institusjonslæring/organisasjonsutvikling/kopiering (informasjon og læring fra andre) innebygd.

Modell 2: "Kvalitetsutviklings-modellen"

Formål:

Stimulere lokalt kvalitetssikringsarbeid, departementet er sentral bruker av resultatene.

Beskrivelse av modellen:

Institusjonene foretar selvevalueringer. Det utføres en ekstern evaluering etter hovedlinjene for de nasjonale evalueringene. Denne evalueringen utføres av fagfeller. Disse skriver to typer rapporter: En rapport går til de miljøer som er blitt evaluert. En samlet rapport går til departementet (denne kan f.eks skrives av lederne for evalueringskomiteene i fellesskap).

Evalueringsobjekt:

Disipliner/Fagområder.

Kjennetegn ved modellen:

Noe mindre grad av systematisering (disipliner/fagområder må ikke nødvendigvis evalueres på samme måte/sætrekk ved den enkelte disiplin må aksepteres). Det kan vurderes om det må foreligge kriterier for hva som er utgjør "minstestandard" på en gitt utdanning. Fordi det produseres to typer rapporter er forankringen delt mellom lokalt og sentralt hold.

Fordeler med modellen:

Informasjonstilgangen og dybden er god. Tilbakemeldingen fra de eksterne evaluatorene skal inneholde konkrete forslag. At de eksterne evaluatorene skal skrive en egen rapport til departementet kan øke relevansen. Modellen har også elementer av institusjonslæring/kopiering (informasjon og læring fra andre) innebygd. Modellen vil i tillegg bidra til å etablere et ytre press på institusjonene.

Ulemper med modellen:

Modellen har større driftsutgifter enn "kvalitetskontroll-modellen". Dette avhenger derimot av omfanget på de evalueringer skal foretas (svært omfattende evalueringer koster totalt mer, men gir billigere evalueringer pr institusjon - jfr. forskjellen i kostnader som fremkommer mellom ulike fagområder). Å utarbeide to forskjellige typer av rapporter er noe som bl.a. gjøres ved akkrediteringsprosesser i USA, men dette medfører også at prosessen kan dra mer ut i tid. Modellen ansvarliggjør i utgangspunktet ikke topp- og mellomledelsen fordi det er disiplinen/fagområdet som er i fokus.

Modell 3: "Behovs-modellen" (modell 1 + 2)

Formål:

Stimulere lokalt kvalitetssikringsarbeid, departementet er sentral bruker av resultatene.

Beskrivelse av modellen:

Institusjonene foretar selvevalueringer. Det kan utføres både institusjons- og disiplin/fagevalueringer, etter et på forhånd definert oppsett. F.eks. kan meta- og disiplinvalueringer gjennomføres parallelt, eller man kan tenke seg at det gjennomføres metaevalueringer hvert år (som en del av rapporteringen til departementet) og disiplinvalueringer hvert andre år etc. En kombinasjon av modell 1 og 2 forutsetter en større organisasjonsmessig enhet enn for modellene enkeltvis (det er mulig å splitte den organisasjonsmessige innsatsen). Samtidig kan denne organisasjonsmessige kapasiteten også utnyttes til andre arbeidsoppgaver som f.eks. "akkrediteringer"/NKU-funksjoner etc.

Evalueringsobjekt:

Institusjoner/Fakulteter/Institutter/Avdelinger/Disipliner/Fagområder.

Kjennetegn ved modellen:

Systematisering vil avhenge av formålet (veksling mellom disiplin/metaevalueringer. Det kan vurderes hvorvidt det bør etableres "minstestandarder" for metaevalueringer). Modellen kan tilpasses spesielt både etter institusjonenes og departementets behov.

Fordeler med modellen:

Informasjonstilgangen og dybden er god. Modellen bidrar til både et ytre og indre press på institusjonene. Graden av læring på institusjonsnivå kan variere etter hvilken evalueringsvariant som utføres. Man kombinerer ansvarliggjøring av ledelsen, med en evaluering av faglig innhold.

Ulemper med modellen:

Modellen får høyere driftsutgifter fordi personalbehovet er større.

Vurdering av de tre modellene

De tre modellene som er skissert antyder også variasjoner i ambisjonsnivå. Modell 3 er i realiteten den variant som brukes både i England og i Sverige. I England utføres riktignok meta- (audit) og innholdsevalueringene av to ulike organisasjoner. I Sverige er begge prosesser Högskoleverkets ansvar. Det burde ikke være noe problem i at samme organisasjon utfører begge typer prosesser så lenge forutsetningene og konsekvensene for begge typer prosesser er kjent og akseptert av institusjonene (man bør f.eks være varsomme med å koble evalueringene tett opp til budsjetttildelinger o.l.)(jfr. Smeby 1996). En slik "evalueringsorganisasjon" bør imidlertid ha en forholdsvis uavhengig status for å kunne oppnå en slik aksept (jf. European report 1995). Modell 1 kan på mange måter betraktes som en minstevariant av et nasjonalt system. Ulempen er at graden av institusjons- og organisasjonslæring er fraværende, noe som kan føre til prosesser med et sterkt symbolpolitisk innhold (eksistensen av papirsystemer). Departementet får også lite informasjon ut av et slikt system. Modell 2 er en videreføring av de nasjonale evalueringene, med en overbygning og spesiell rapportering til departementet som nytt element. Dette sikrer departementet relevant informasjon, men modellen vil ha liten grad av forankring hos institusjonenes ledelse (hvis da ikke et spesielt grep gjøres for å initiere dette).

NIFU vil ut fra de tre modellene antyde at modell 3 kanskje er den beste på det nåværende tidspunkt. Kostnadene ved modellen er større, men en utbygd organisasjon på dette feltet kan også tilføre departementet annen relevant styringsinformasjon, som departementet pr dato ikke har tilgang til (NKU-funksjoner, faglig vurdering av knutepunktfunksjoner etc.). En slik modell ville også kunne bidra til å skille "faglige" og "politiske" vurderinger.

Litteratur:

European Report (1995) European Pilot Project for Evaluating Quality in Higher Education, Socrates Programme, Directorate for Education, Training and Youth. Commission of the European Communities. Brussel.

Jordell, Karl-Øyvind (1992) Nasjonal evaluering av økonomisk-administrativ utdanning. Premisser og prosess, NAVFs utredningsinstitutt, Rapport 3/92, Oslo.

Karlsen, Rita & Bjørn Stensaker (1993) Å organisere kvalitet?, NAVFs utredningsinstitutt, Rapport 9/93, Oslo.

Karlsen, Rita & Bjørn Stensaker red. (1996) Kvalitet i høyere utdanning: Teori, empiri og praksis. NIFU-rapport 1/96, Oslo.

Smeby, Jens-Christian ed. (1996) Evaluation of Higher Education in the Nordic Countries, Nordisk Ministerråd, Nord 1996: 6, Copenhagen.

Stensaker, Bjørn & Rita Karlsen; "Å vurdere kvalitet?", Utredningsinstituttet for forskning og høyere utdanning, Rapport 10/94, Oslo.

Stensaker, Bjørn (1996) Organisasjonsutvikling og ledelse: Om bruk og effekter evalueringer på universiteter og høyskoler. NIFU-rapport (under arbeid), Oslo.

St.Meld.Nr. 40 (1990-91) Fra visjon til virke - om høgre utdanning. Oslo.

Van Vught, Frans & D. Westerheijden (1993) Quality Management and Quality Assurance in European Higher Education. Methods and Mechanisms. Commission of the European Communities. Brussels.