

# Undervisningsevalueringens dilemmaer: kontrol og udvikling – to verdener?

*Hanne Leth Andersen, prorektor og professor, Roskilde Universitet.*

*Jakob Ravn, evalueringschef, Copenhagen Business School.*

*Thomas Colerick, stud.mag., Roskilde Universitet.*

## Reviewet artikel

*Evalueringsskaber skifter løbende, og der eksperimenteres med nye tilgange, metoder og procedurer. Artiklens fokus er på dilemmaet imellem kontrol og udvikling. Der rapporteres fra en undersøgelse af holdninger til forskellige tilgange til evaluering og udbyttet af disse. Undersøgelsen er en kvalitativ undersøgelse baseret på interviews med centrale aktører på forskellige niveauer (ledelse, lærere, studerende, studieledere og institutledere) og dokumentariske studier af materiale, på tre danske universiteter. Vi sammenligner, hvordan forskellige aktører opfatter mål med evaluering, praksis og udbytte. Alle tre tilfælde viser, at ydre krav og forventninger fører til større centralisering og mere standardisering. Selv om ledelsen forsøger at kombinere udvikling og kontrol, sker der ofte en opdeling i to parallelle processer. Undersøgelsen dokumenterer, at kombinationen af kontrol og udvikling i den samme undervisningsevaluering er u hensigtsmæssig: Såvel studerende som undervisere mister engagementet og savner muligheder for feedback og opfølgning. Det fremgår, at standardisering, centralisering og formalisering af undervisningsevalueringer fører til mindre udvikling og regnes for mindre nyttigt af såvel undervisere som studerende.*

## Indledning

Undervisningsevaluering er omdrejningspunkt for såvel kvalitetssikring som kvalitetsudvikling. Det er en aktivitet fyldt med modstridende hensyn og ønsker: De studerendes organisationer ønsker evaluering for at sikre indflydelse på undervisningen, men de studerende er evalueringstrætte; undviserne ønsker feedback på undervisningen, men de frabeder sig overvågning og kontrol; ledelserne ønsker dokumentation for kvaliteten af undervisningen til brug for eksempelvis akkrediteringer, men de vægrer sig ved udelukkende at evaluere med henblik på eksterne krav.

Der gennemføres derfor forskellige typer af evaluering: Centralt versus lokalt, skriftligt versus mundtligt, kvantitativt versus kvalitativt. Spørgsmålet er, hvordan de forskellige metoder, processer og institutionelle praksisser opleves. Vi søger i denne artikel at belyse, hvordan det grundlæggende dilemma i undervisningsevaluering mellem kontrol og udvikling kommer til udtryk og opleves af de involverede parter. Artiklens primære formål er at dokumentere og beskrive oplevelserne af de forskellige evalueringspraksisser. Det ledende spørgsmål i undersøgelsen har været at afdække og dokumentere, hvordan forskellige aktører oplever praksis i forhold til dimensionen kontrol/udvikling. Vi vil således belyse, om og i hvilken grad forskellige aktører oplever ejerskab til evalueringerne (Dahler-Larsen, 1998) samt om og i hvilken grad de oplever, at de får feedback og information om evalueringresultaternes anvendelse (Watson, 2003) og effekt.

### Metode

Det empiriske materiale, der danner baggrund for analysen i denne artikel, er en undersøgelse foretaget i 2009 af det daværende CBS Learning Lab på CBS. Den omhandler undervisnings- og uddannelsesevalueringer på tre fakulteter på tre danske universiteter med henblik på at sammenligne praksis og ekstrahere fælles udfordringer, sådan som de opleves af ledelser, undervisere og studerende. De forskellige aktørperspektiver skaber en spændvidde i materialet, som giver mulighed for at komme bag om formulerede strategier og procedurer og afdække forskelligheder og flertydighed i evalueringspraksis på universiteterne.

Undersøgelsen er en kvalitativ undersøgelse baseret på en række interviews samt dokumentariske studier af materiale fra de tre deltagende fakulteter: Det Humanistiske Fakultet på Aarhus Universitet (HUM AU), Det Naturvidenskabelige Fakultet på Københavns Universitet (NAT KU) og Copenhagen Business School (CBS), som er et monofakultært universitet med en samfundsvidenskabelig forankring. Inden for hvert fakultet fokuserer undersøgelsen på et udvalgt institut. Her er der talt med respondenter på en række forskellige niveauer: Institutledere, studieledere, undervisere og studerende. Derudover deltager fakulteternes prodekaner (øverste uddannelsesansvarlige myndighed) som respondenter. Der er foretaget individuelle interviews på fakultets-, institut- og studieledelsesniveau, mens der på underviser- og studenterniveau hovedsageligt er lavet fokusgruppeinterviews for at medtage så mange perspektiver som muligt.

Undersøgelsen baserer sig på i alt 15 interviews, hvoraf 5 er fokusgruppeinterviews, og 10 er individuelle interviews. Interviewene er fordelt således, at der knytter sig fem til hvert fakultet. De er ikke transskriberet fuldstændigt, men selektivt ud fra undersøgelsens omdrejningspunkter:

- Formål og intentioner med evaluering af undervisningen.
- Proces og metode (hvordan og hvornår konstrueres informationen?).
- Indhold (hvad opfattes som relevant og gyldig information?).
- Anvendelse (intenderet og praktiseret).

Disse omdrejningspunkter dannede baggrund for hvert enkelt interview og fungerede som temaer til at guide samtalen. Denne tilgang er inspireret af Steinar Kvales semistrukturerede interviewguide (Kvale, 1997). Udvalget af respondenter er foregået ved at inddrage interne nøgleaktører på de enkelte fakulteter, som i kraft af deres position havde adgang til relevante informanter. Eksempelvis bestod undersøgelsens baggrundsgruppe af en repræsentant fra hvert af de tre fakulteter. Samtlige anvendte citater i undersøgelsen er godkendt af de pågældende respondenter.

Det anvendte dokumentariske materiale er indsamlet ved henvendelse hos ledelserne på de deltagende fakulteter og institutter, idet vi udbad os dokumenter, der beskrev fakultetets og instituttets officielle politik for evaluering eller kvalitetsudvikling.

Det komparative aspekt i analysen kommer til udtryk ved, at undersøgelsen arbejder på at anskueliggøre og sammenligne dels de rammer, som de enkelte fakulteter, institutter og uddannelser har etableret for studenterevaluering af undervisning, dels kendskabet til og fortolkningen af rammerne i de konkrete evalueringspraksisser.

Følgende analyseramme blev udviklet i baggrundsgruppen efter en række indledende interviews på de tre fakulteter. Analyserammen er således udtryk for intersubjektive fortolkninger og diskussioner af de centrale temaer i hvert interview:

<b>Kvantitativ/kvalitativ tilgang</b>	Hvilken metode er dominerende, når fakultetet studenterevaluerer, og hvorfor?
<b>Historik</b>	Hvorfor er nye evalueringsmodeller i færd med at blive indført på de tre fakulteter – hvorfor er feltet i forandring?
<b>Ansvarsfordeling</b>	Hvordan ser den fakultetsinterne ansvarsfordeling ud? Hvordan udfoldes <i>Universitetsloven</i> ?
<b>Styringsmekanisme</b>	Hvad styrer man på – rammer og/eller indhold? Hvilke argumenter ligger bag?

Figur 1

For at muliggøre en samlet analyse var det nødvendigt at have indsigt i praksis på hvert af de tre fakulteter. Denne indsigt er faciliteret af tre 'vertikale analyser'. Her arbejdede vi ud fra arbejdsspørgsmål som: Hvilke forbindelser og relationer kan

identificeres mellem de forskellige organisatoriske niveauer? Hvad lægger relationerne op til? Er det homogent? Hvilke konsekvenser har en centraliseret evalueringsspraksis?

### **Kvalitetssikring og kvalitetsudvikling: Et klassisk dilemma**

Evalueringens dobbelte formål har altid voldt problemer. Evaluering finder sted i et spændingsfelt mellem den kritiske refleksion og en udefrakommende bedømmelse. Den kritiske refleksion finder sted ved at deltagerne i tilbageblik, men ikke nødvendigvis ved undervisningens afslutning, vurderer værdien af processer og/eller resultater. Den udefrakommende bedømmelse finder sted på foranledning af andre aktører end deltagerne, fx studieledelse, fakultetsledelse eller institutionseksterne aktører. Denne bedømmelse kan være mere eller mindre fremadrettet, men den vil næsten altid indeholde et kontrolelement, fordi evalueringen foretages på foranledning af andre end de involverede selv, og fordi disse 'andre' besidder magt i forholdet til de(t) evaluerede.

I litteraturen er evalueringernes dobbelte formål behandlet i en mangfoldighed af forskellige kontekster og med forskellig delvist overlappende terminologi som fx følgende begrebspar: Formativ versus summativ (Derry, 1979; Henkel, 1998), udvikling versus kontrol (Perlberg, 1979; O'Hanlon & Mortensen, 1980), intern versus ekstern (Borgnakke, 2008), top-down versus bottom-up (Albæk, 1998) og institutionel versus individuel (Braskamp & Ory, 1994). Fælles er iagttagelsen af, at evaluering som fænomen og praksis er udspændt og modsætningsfuld, og at spændingerne ikke lader sig reducere til et spørgsmål om entydige valg af givne evalueringsmetoder til opfyldelse af givne formål. Undervisningsevalueringer er 'både-og'.

I forhold til universiteters og andre videregående uddannelsesinstitutioners brug af undervisningsevalueringer er det primært evaluering af den enkelte undervisers præstation, der har påkaldt sig opmærksomhed. Spørgsmålet om udviklingsorienteret versus kontrolorienteret evaluering er særligt påtrængende, når kontrollen kobles til personalemæssige beslutninger som for eksempel løn, forfremmelser, fyringer mv.

Spørgsmålet er, om der er en pointe i at samle eller adskille kontrol og udvikling. Mills & Hyle argumenterer for, at separationen af formål i to adskilte sæt evalueringssprocedurer vil opleves som påtrængende og uproduktiv (1999, s. 369). To adskilte sæt evalueringssprocedurer med hver sit formål kan også af de studerende opleves som tidskrævende og uproduktivt, og dermed kan man risikere, at begge evalueringsskemaer vil lide under manglende deltagelse og engagement fra de involverede undervisere og studerende. Som Müllen konstaterer: "De studerende bliver evalueringstrætte af to evalueringer pr. semester i hvert fag, som de ikke kan se effekten af. Underviserne føler sig udsatte, eller at de spilder deres tid" (Müllen, 2005, s. 4).

På den ene side er det forbundet med betydelige vanskeligheder at favne både kontrol- og udviklingsformålet i én og samme evalueringsform, og på den anden side kan en opdeling i to adskilte processer medføre manglende deltagelse og engagement.

### **Evalueringspraksis: formål og metoder**

Det dobbelte formål afspejles i de tre fakulteters procedurer og praksisser. Der anvendes alle steder et miks af kvalitative og kvantitative metoder, skriftlige og mundtlige metoder samt højt standardiserede og mindre standardiserede metoder og processer. Der evalueres efter undervisningens afslutning med brug af evalueringsskemaer. Skemaerne varierer en smule i standardiseringsgraden, men der anvendes som udgangspunkt et fælles skema. Prototypen på evaluering af undervisning på universiteterne ser således ud til at være helt eller delvist standardiserede evalueringsskemaer udsendt efter undervisningens ophør. På alle fakulteter suppleres denne evalueringsform med andre metoder og processer.

Nogle steder vælges der en klar opdeling mellem overvejende standardiserede, kvantitative evalueringsskemaer udsendt efter undervisningens ophør og udvikling af redskaber til at sikre udviklingsperspektivet. Man skal imidlertid være opmærksom på en særlig historik: NAT KU var på undersøgelsestidspunktet ved at udvikle to nye metoder til den kvalitative del, nemlig interne peer review og fokusgrupper. Prodekanen lægger vægt på at adskille de to metoder og tilgange, fordi der er berøringsangst forbundet med evaluering. Dilemmaet med det dobbelte formål fører her til etablering af to parallelle evalueringssystemer.

På CBS kan man ligesom på KU NAT identificere to overordnede typer af studenterevaluering: En formel skriftlig evaluering i form af elektroniske spørgeskemaer, og en uformel mundtlig evaluering, som den enkelte underviser opfordres til at lave, uden styring fra fakultetsniveau.

Ved HUM AU er metoderne mindre standardiserede end ved de to andre fakulteter. Her anvendes både midtvejs- og slutevaluering ud fra forskellige traditioner på fagene. Der anvendes hyppigt et fælles grundskema, der kan tilpasses det enkelte undervisningsforløb. Spørgeskemaets resultater danner grundlag for en efterfølgende diskussion. Alle evalueringsprocesser munder ud i en skriftlig afrapportering som dokumentation til studienævnet.

### **Oplevelser af evaluering**

På alle tre fakulteter kan der konstateres en oplevelse af manglende anvendelse af evalueringerne – først og fremmest blandt studerende, men interessant nok også blandt underviserne. Der er god grund til at antage, at etableringen af et nationalt akkrediteringssystem samt krav om offentliggørelse af evalueringerne rent faktisk påvirker *måden*, der evalueres på.

*Fakultetsledelsernes vinkel: Det dobbelte formål*

Der er enighed om, at det vigtigste formål er at udvikle undervisningen og øge kvaliteten:

*"Udviklingsperspektivet er vigtigt. Evaluering er ikke en kontrolforanstaltning, men et målrettet redskab i kvalitetsudviklingen."* (Prodekan, NAT KU).

Evaluering skal helst kunne fungere som et redskab for dialog mellem ledelse, underviser og studerende og som en indikator for, hvorvidt visse aspekter af undervisningen ikke fungerer hensigtsmæssigt. Samtidig har fakulteterne behov for at dokumentere at al undervisning kvalitetskontrolleres. Derfor påpeges det flere gange i undersøgelsen, at selve afrapporteringen fra det enkelte kursus både har til formål at dokumentere at evalueringen er foregået og at pege på muligheder for udvikling:

*"Generelt ser vi evaluering som et redskab til både kontrol og udvikling."* (Prodekan, CBS).

På dette niveau er der tale om nogle ret klare ambitioner om evalueringer som 'både-og' i forhold til kontrol og udvikling, hvilket også afspejler sig i de forskellige metodiske tilgange og i styringsmekanismen. Der styres på de overordnede rammer, men på en måde der muliggør, at undervisere kan påvirke indholdet af en evaluering-proces.

*Institutedelsernes vinkel: Sikring og ansvar*

Oplevelsen af det dobbelte formål genfindes hos institutlederne. Det fremhæves, at evaluering handler om dialog omkring undervisningskvalitet samtidig med, at den skal kunne afdække større problemer i forbindelse med undervisningen:

*"Vi prøver at skabe en tosidet platform. Den ene handler om, at den modulansvarlige for hvert kursus bruger evalueringerne til opfølgingsmøder. Den anden om, at der er en dialog mellem undervisningsudvalget som repræsenterer instituttet og den modulansvarlige."* (Institutleder, KU NAT).

Personaleansvarets placering hos institutledelsen påvirker perspektivet på evalueringerne i retning af, at de kvantitative evalueringer er mest brugbare. Meningsfulde evalueringer er dem, der kan anvendes til kontrol af og dialog om medarbejdernes undervisning i forbindelse med eksempelvis MUS-samtaler:

*"Det jeg først og fremmest kigger på, er den del, der handler om underviserens performance."* (Institutleder, CBS).

Institutlederne nuancerer dog opfattelsen af kvantitative afrapporteringer, idet de er forsigtige med at tillægge evalueringerne for stor gyldighed, og der tages væsentlige forbehold i forhold til at anvende evalueringerne som egentlige udtryk for medarbejdernes præstationer:

*"Jeg har undervisere, som jeg af forskellige grunde ved, ikke er fremragende undervisere, men det er ikke sådan, at jeg kan gå ind og se på evalueringerne, at de er dårlige undervisere." (Instituttleder, HUM AU).*

Når instituttlederne ved, at der er problemer med undervisningen, er det ofte ikke fordi, de har fundet ud af det gennem evaluering, men fordi det er en viden, som findes i organisationen, og fordi det er et emne, der tales om blandt kolleger og studerende.

*Studieledernes vinkel: Dialog og kvalitet*

Også for studielederne har evalueringerne to forskelligt rettede formål, men til forskel fra fakultetsledelserne og instituttlederne oplever de evalueringerne som udtryk for to separate størrelser med forskelligartede målsætninger. De kvalitative, dialogorienterede evalueringer opleves som de mest værdifulde, og muligheden for at præge deres indhold værdsættes:

*"Den del af evalueringen, der skal kunne standardiseres og offentliggøres, over for den konkrete mundtlige dialog mellem underviser og studerende, det bliver to forskellige verdener." (Studieleder, HUM AU).*

Samme studieleder laver på baggrund af dialog med studenterrepræsentant og fagleder et resumé af evalueringsrunderne, som adskiller sig en del fra skemaresultater:

*"De standardiserede evalueringer vil blive en helt anden ting i forhold til vores værdifulde resuméer. Arbejdsomt er det som om, vi skal leve i to forskellige spor. De ikkestandardiserede vil jeg ikke opgive, for det er der, dynamikken ligger." (Studieleder, HUM AU).*

Der er ikke nogen strukturel eller indholdsmæssig sammenhæng mellem de to typer evalueringer. Det er, som en studieleder beskriver det, "to forskellige verdener". Den ene verden er fortrinsvis obligatorisk, standardiseret, kvantificeret og skriftlig, og den bidrager ikke til den anden verden, som er frivillig, mindre standardiseret, kvalitativ og fortrinsvis mundtlig.

Der nævnes et tredje formål, som handler om evaluering af undervisningen som redskab til udvikling af de studerendes forståelse af læreprocessen. Det er igen den mundtlige dialog der fremhæves:

*"Et andet meget vigtigt formål er at få en dialog med de studerende om, hvad der foregår. Nogle gange er det sådan set meningskabelse, der er behov for, frem for at kunne krydse af i et kvalitetsskema." (Studieleder, CBS).*

Studieledernes vinkel på evaluering synes i ret høj grad at være pædagogisk orienteret. De fremhæver, at evaluering som kontrol og kvalitetssikring er en helt anden tilgang og proces. Ansvarsfordelingen manifesterer sig tydeligt mellem personalelederens fokus på den enkelte undervisers kompetencer og uddannelseslederens fokus, hvor indholdet af evalueringen er uddannelsens kvalitet.

#### *Undervisernes oplevelse af formål*

Også blandt underviserne kan der identificeres en forståelse af, at evalueringsskemaer og løbende dialogisk evaluering opfylder forskellige formål. Dekoblingen mellem de to opleves dog i endnu højere grad på dette niveau. Underviserne adskiller tydeligvis de centrale eller officielle evalueringer, der foreligger som elektroniske eller papirbårne skemaer og afrapporteringer, fra den daglige praksis og fra etableringen af en mundtlig midtvejsdialog:

*"De officielle evalueringer fungerer lidt som en vagthund." (Underviser, NAT KU).*

*"Der går lang tid, før vi ser de officielle evalueringer, så jeg tager den ofte mundtligt med de studerende halvvejs i forløbet. For hvis først de studerende bliver utilfredse med noget, så eskalerer det, og det kommer til udtryk til slut igennem nogle enormt negative evalueringer." (Underviser, CBS).*

Underviserne lægger vægt på at kunne komme i dialog med de studerende. De er mest interesserede i en umiddelbar feedback på og dialog om undervisningen. De anerkender behovet for kontrol og dokumentation, men nogle ville hellere være foruden:

*"Den optimale evalueringsskemaer ville ikke have skemaer og afrapporteringer. Det ville være en gensidig refleksion over, hvordan det går." (Underviser, AU HUM).*

Underviserne fokuserer på den direkte anvendelighed, og der er delte meninger om hvordan og i hvilket omfang, de centralt drevne evalueringer bliver anvendt af underviserne:

*"Anvendelsen sker for usystematisk." (Underviser, NAT KU).*

*"Der kommer ting frem, som ikke ellers ville komme frem." (Underviser, HUM AU).*

Underviserne oplever at få størst udbytte af deres egen mundtlige dialog med de studerende undervejs i forløbet. De fleste er overvejende skeptiske med hensyn til anvendeligheden af de centralt drevne evalueringsskemaer, om end disse også anvendes som retrospektivt pejlemærke for, om undervisningen var vellykket.



### *Studerendes vinkel: Dialog og direkte kontakt*

Ligesom studielederne har de studerende oplevelsen af, at evaluering kan tilhøre to meget forskellige verdener med to forskellige modtagergrupper, ofte manifesteret som mundtlig vs. skriftlig evaluering:

*"Man kan tale om to helt forskellige ting. De mundtlige er til underviserens egen brug. De andre er mere overordnede."* (Studerende, NAT KU).

De oplever ingen sammenhæng mellem de to forskellige typer evalueringer, og der er ingen tvivl om, hvad de oplever som mest meningsfuldt:

*"Den mundtlige evaluering med læreren er klart at foretrække, fordi du er i direkte kontakt med læreren."* (Studerende, CBS).

*"Mange studerende er i tvivl om, hvorvidt de her skemaer overhovedet bliver læst. [...]"* (Studerende, AU HUM).

De studerende foretrækker klart den direkte dialog. De oplever den centrale og kontrollerende del med spørgeskemaer som fjern og overordnet, muligvis uden betydning eller virkning.

### **Evalueringemetoder**

#### *Skriftlig, formel og mundtlig, uformel evaluering*

Det er fortrinsvis undervisere og studerende, der forholder sig til forskellene mellem den direkte mundtlige dialog mellem undervisere og studerende og den skriftlige dialog mellem studerende og 'institution'. Blandt underviserne lægges der vægt på, at den mundtlige dialog med de studerende kan blive konkret og brugbar i sammenhængen:

*"Hvis der er noget jeg synes skal diskuteres, lægger jeg op til det mundtligt. [...] Evaluering med skemaerne kan desværre blive lidt mekanisk."* (Underviser, AU HUM).

Den mundtlige evaluering kan ikke alene anvendes umiddelbart til forbedringer af den enkeltes undervisning men også i højere grad bringes videre til et generelt niveau:

*"Den mundtlige, løbende evaluering, vi alle sammen laver, er den bedste erfaring, vi har, med evaluering. Det burde vi tage op på næste niveau og bruge på hele uddannelsen."* (Underviser, NAT KU).

Selv om skemabaseret, skriftlig evaluering skulle kunne bruges til at gøre evalueringresultaterne mere legitime og statistisk korrekte, viser det sig svært at anvende

den type resultater konkret. Det, som kan bruges, er de kvalitative resultater, hvad enten de produceres i en mundtlig dialog eller i kommentarfeltene til en skriftlig evaluering:

*"Meget af det, de skriver i de kvalitative kommentarer, er det samme som, hvad de siger, når man spørger dem mundtligt. Det er der, man får de gode inputs. Jeg synes, at der er meget lidt råderum for at vurdere, hvad resultaterne fra skemaerne betyder." (Underviser, CBS).*

De studerende oplever den mundtlige evaluering som et godt redskab til justering af undervisningen. De tror på, at underviseren har tænkt sig at bruge evalueringen til noget, mens de skriftlige evalueringer i større grad opleves som henvendt til 'andre' end underviseren. Elektronisk distribuerede evalueringsskemaer bliver upersonlige og 'institutionelle', mens dialog med underviseren virker mere 'seriøs', også selv om de studerende er kritiske:

*"Gennemgående tror jeg, at ret mange var kritiske over for undervisningen og kursernes forløb, men alligevel var det de færreste, der valgte at besvare spørgeskemaerne." (Studerende, NAT KU).*

De studerende fremhæver den mundtlige dialog med det argument, at en kommentar i et tekstfelt kan fortolkes vidt forskelligt. Derudover medvirker den mundtlige dialog til, at de føler sig lyttet til:

*"På et andet kursus blev vi engang bedt om at udfylde et skema halvvejs i forløbet. Til den sidste undervisningsgang brugte vi så en halv time på at samle op på evalueringen og få uddybet nogle af kommentarerne. [...] Man følte, at man trængte igennem, og at hun tog det til sig." (Studerende, NAT KU).*

De studerende har behov for at opleve underviseren i modtagelsen af deres feedback. Dialogsituationen giver dem mulighed for at uddybe, hvad de mener, og føle, at deres budskab når frem. Situationen giver gode forudsætninger for at begge parter kan bidrage til, at der følges op og foretages ændringer der, hvor det blev påpeget, at der var brug for det. Den mundtlige dialog virker forpligtende og udvikler de studerendes evalueringskompetence.

#### *Standardiserede skemaer*

Det er oplevelsen blandt både ledere, undervisere og studerende, at standardiserede spørgeskemaer let bliver for overfladiske eller ligefrem indholdsløse:

*"De generelle evalueringsskemaer vi har, forholder sig bredt til det hele, og det er jo ikke altid det, der siger mest." (Instituteder, HUM AU).*

De studerende er i undersøgelsen enige om, at et spørgeskema til en vis grad skal være tilpasset det konkrete undervisningsforløb, og at indholdet i de standardiserede skemaer bliver for diffuse og fjerne fra den konkrete sammenhæng. De antyder, at et standardiseret skema kan opfattes nærmest useriøst, og de finder ikke, at spørgsmålene rammer ind i den konkrete kontekst eller inspirerer til konstruktiv deltagelse:

*"De virker meget standardiseret de spørgeskemaer. Det får mig til at tænke på, om der overhovedet er nogen der lytter til, hvad vi siger. Bliver det rent faktisk taget op?"* (Studerende, NAT KU).

Blandt underviserne deles denne holdning til en vis grad, men der er dog positive sider:

*"De kvalitative kommentarer indeholder ofte gode forslag. De nye skemaer er kortfattede og præcise. Det betyder meget for de studerende, at skemaet er hurtigt og nemt at svare på."* (Underviser, CBS).

De studerende finder skemaerne uinspirerende, fordi de ikke får udtrykt det, der kunne være relevant i forhold til det enkelte kursus. Underviserne er lidt mere positive, men det gælder især muligheden for kvalitative kommentarer. Også studielederne oplever, at netop de kvalitative kommentarer er anvendelige, når det gælder om at skabe sig et overblik over, om undervisningen fungerer som den skal.

#### *Midtvejs- eller slutevaluering*

Helt overordnet er studerende, undervisere og studieledere enige om, at en mundtlig midtvejsevaluering er den mest nyttige og motiverende form. Typiske argumenter er, at evalueringen kan anvendes til forventningsafstemning og målrettet dialog med de studerende med det formål at følge op og lave justeringer:

*"På andre fag har de typisk en mundtlig midtvejsevaluering, som faktisk viser sig at være næsten den vigtigste af dem alle sammen – fordi underviser og studerende i fællesskab forventningsafstemmer ved denne."* (Studieleder, AU HUM).

Midtvejsevalueringer kan antage forskellige former. Der kan anvendes spørgeskemaer som udgangspunkt for en åben dialog med holdet, eller man kan vælge at gennemføre et fokusgruppeinterview eller en uformel dialog med en udvalgt gruppe af de studerende. Det sidste giver ifølge underviserne mulighed for at komme vidt omkring:

*"Jeg kalder 6-7 studerende fra holdet ind til en round-table discussion. Det er fordi, det giver dig muligheden for at lave justeringer undervejs."* (Underviser, CBS).

*"Da jeg tilbød mine studerende at deltage i en fokusgruppe lige efter en undervisning, var der 10, der meget gerne ville deltage. Her kom der nogle kommentarer frem, som slet ikke fremgik af de skemaer, der blev lavet."* (Underviser, NAT KU).

Nogle studerende mener ikke, de har overblik nok midtvejs, og virker som om, de føler, det er et tidspres. Et andet argument mod midtvejsevaluering er frygt for, at en evt. negativ evaluering skaber dårlig stemning eller bringer den kritiske studerende i et konfliktforhold til underviseren og, at vedkommende dermed løber en risiko for at få dårligere bedømmelse:

*"Hæmskoen ved det her er, at man ikke gider bruge tid på at rakke en ned, som man skal til eksamen hos. Vi ved heller ikke, om den direkte kritik af underviseren kommer videre."* (Studerende, CBS).

Kritikken synes dog også at være forbundet med en evalueringsform, der sætter fokus på underviseren snarere end på undervisningen. Sådanne former er hyppige i slutevalueringer, som reelt ofte er kontrol af underviseren og ikke et oplæg til konstruktiv dialog mellem underviser og studerende.

En af de primære årsager til, at slutevalueringer ikke er særligt populære hos de studerende og dermed ofte har meget lav besvarelsesprocent, især når den udsendes elektronisk efter undervisningsforløbet, er, at den bliver for irrelevant og fjern i de studerendes bevidsthed. På det tidspunkt er eksamen mere aktuel end undervisningen samtidigt med, at opfølgning ikke kan nå at få indflydelse:

*"Når evalueringen først kommer efter, har man lagt kurset bag sig. Dette kombineret med, at man udelukkende gør det for andres skyld – det ville være bedre, hvis man selv fik noget ud af det undervejs."* (Studerende, NAT KU).

Helt grundlæggende er ønsket om og behovet for feedback på evaluering. Det gælder for begge typer af evaluering, men er mest oplagt, når det evaluerede forløb endnu ikke er afsluttet.

*"Hvis man så resultater af sin evaluering og fik feedback på den, ville det klart være en motivationsfaktor for at deltage fremover."* (Studerende, NAT KU).

Endelig angives problemet med at give feedback og evaluere sit undervisningsforløb også som uklarhed omkring hvem, der er modtager af evalueringens resultat.

### **Konklusion: Undervisningsevaluering er dialog**

Der findes nærmest uendeligt mange mulige metoder inden for evaluering, hver for sig med forskelligt fokus og mål. Det handler om tidspunktet, mediet, valg af indhold samt om afsendere og modtagere. Alle muligheder kombineres med underviserens og de studerendes holdning til evaluering, tillid til processen og tilgang til dialog og forandring. Det viser sig, at de evalueringstyper, der oftest arbejdes med i de undersøgte cases, er mundtlig, uformel midtvejsevaluering og skriftlig, standardiseret slutevaluering.

Vi kan konstatere, at det fortsat er en stor udfordring at skabe rammer for en udbytterig evaluering. De studerende på de tre undersøgte fakulteter/universiteter oplever ikke, at de centralt drevne slutevalueringer af undervisningen har nogen reel betydning eller effekt. De ved ikke, hvad der sker med resultaterne af disse evalueringer, og de modtager ikke nogen feedback eller information om opfølgning eller anden anvendelse. Dermed bidrager undersøgelsen med yderligere empirisk belæg for velkendte pointer vedrørende oplevelsen af ejerskab til evalueringen (Dahler-Larsen, 1998) og vedrørende væsentligheden af "closing the feedback loop" (Watson, 2003).

Yderligere viser det sig, at underviserne mangler både feedback og fornemmelse af formål og anvendelse af de centralt valgte evalueringsmetoder, enten fordi de ikke er tilpasset den konkrete undervisning eller fordi, de får tilbagemeldingen for sent til at anvende den. Denne uvished blandt underviserne er identificeret i en anden dansk undersøgelse (Moldt, 2007), hvor underviserne endvidere peger på, at uvisheden medfører frygt for ledelsesmæssigt 'misbrug'.

Intentionerne fra ledelsernes side er på den ene side at tilfredsstille udefra kommende krav, som handler om offentliggørelse, tilsyn med undervisningen og opfølgning på eksternt initierede foranstaltninger, og på den anden side at sikre, at evalueringen anvendes som et udviklingsredskab. Undersøgelsen peger på, at det grundlæggende paradoks skinner tydeligt igennem i aktørernes oplevelser af evalueringerne. På trods af forskellige evalueringsmetoder er undervisere og studerende overraskende enslydende i deres faktiske oplevelser af evalueringerne. Ledelsernes velmente forsøg på at kombinere de to formål bidrager til den usikkerhed, der er fremherskende blandt undervisere og studerende. Dette resultat understøtter til dels idéen om, at evaluering af undervisningen på universiteter er løst koblet til praksis (Dahler-Larsen, 1998). Der er dog ikke noget i undersøgelsen, der tyder på, at der er tale om evaluering som en *rituel refleksion* uden reel interesse for indhold, metode eller resultat (ibid.).

Hvis man skal tage undersøgelsens resultater alvorligt, må man acceptere, at der er to forskellige spor i evaluering, der ikke kan forenes i det samme evalueringsforløb, heller ikke når det deles op i en intern del og en eksternt rettet afrapporteringsdel. Men hvis de to formål skal adskilles i to tempi, skal der skæres i begges omfang og hyppighed. I modsat fald vil evalueringstrætheden blive endnu mere dominerende, end den allerede er.

*Hanne Leth Andersen er prorektor og professor i Universitetspædagogik ved Roskilde Universitet, ph.d. i Fransk Sprog. Hun har erfaringer som studieleder og prodekan for uddannelse fra det Humanistiske Fakultet ved AU og som direktør for CBS Learning Lab. Hun arbejder med forskning og udvikling i forhold til kvalitet i uddannelse, herunder evaluering og dialog med studerende.*

Jakob Ravn er chefkonsulent på CBS. Han har i 15 år arbejdet med udvikling af universitetets undervisning, uddannelser og undervisere. De seneste 5 år med særligt fokus på evaluering som redskab til kvalitetssikring og -udvikling.

Thomas Colerick er specialestuderende ved Institut for Samfund og Globalisering, Roskilde Universitet. Han har en B.Sc. i HA(psyk.) fra CBS.

## Litteratur

- Albæk, E. (1998). Knowledge, Interests, and the Many Meanings of Evaluation. A Developmental perspective. *Scandinavian Journal of Social Welfare*, 7(2), 94-98.
- Borgnakke, K. (2005). *Læringsdiskurser og praktikker*. København: Akademisk Forlag.
- Borgnakke, K. (2008). Evalueringsstrategier – i den pædagogiske kontekst. I: Borgnakke, K. (red.): *Evalueringens spændingsfelter*. Århus: Forlaget Klim.
- Dahler-Larsen, P. (1998). *Den rituelle refleksion*. Odense: Odense Universitetsforlag.
- Derry, J.O. (1979). Can Students Ratings of Instruction Serve Rival Purposes? *The Journal of Higher Education*, 50(1) (Jan.-Feb. 1979), 79-88.
- Henkel, M. (1998). Evaluation in Higher Education: Conceptual and Epistemological Foundations. *European Journal of Higher Education*, 33(3) ["The Evaluative State Revisited: 20<sup>th</sup> Anniversary Issue of Review of trends in Higher Education"] (Sep. 1998)], 85-297.
- Lincoln, Y. S. & Guba, E. G (1980). The Distinction between Merit and Worth in Evaluation. *Educational Evaluation and Policy Analysis*, 2(4) (July-August 1980), 61-71.
- Mills, M. & Hyle, E. A. (1999). Faculty evaluation: A prickly pair. *Higher Education*, 38(3) (Oct. 1999), 351-371.
- Moldt, C. (2007). Fra meningsløse til meningsfulde performancemålinger – en case om kvalitetsmåling i undervisningssektoren. *Økonomistyring og Informatik*, 22(5) (2006/2007).
- Müllen, R. von (2005). *Undervisningsevaluering som redskab til pædagogisk udvikling på KU*. Speciale i retorik, Københavns Universitet, Institut for Medier, Erkendelse og Formidling.
- O'Hanlon, J. & Mortensen, L. (1980). Making Teacher Evaluation Work. *The Journal of Higher Education*, 51(6), 664-672.
- Perlberg, A. (1979). Evaluation of Instruction in Higher Education: Some Critical Issues. *Higher Education*, 8(2) (March 1979), 141-157.
- Shadish, W. R., Cook, T. D. & Levitan, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, Calif.: Sage Publications.
- Watson, S. (2003). Closing the feedback loop: Ensuring effective action from student feedback. *Tertiary Education and Management* 9, 145-157.