

Evaluering af evalueringer

Gerd Christensen, lektor, Institut for Medier, Erkendelse og Formidling, Københavns Universitet



vedrørende humanistisk og samfundsvidenskabelig forskning.

Gerd Christensen er cand.mag. i psykologi og filosofi samt exam. pæd. Er ansat som lektor i pædagogik på Institut for Medier, Erkendelse og Formidling, Københavns Universitet. Forsker i problemorienteret projektarbejde i grupper samt i videnskabsteoretiske og metodologiske forhold

Reviewet artikel

Studerendes evaluering af undervisningen på universiteterne er et fænomen, der har været stigende interesse for igennem de senere år. Dels stilles der ministerielle krav om, at alle landets uddannelser skal evaluere undervisningen, dels kræves der i stigende grad åbenhed om evalueringernes resultater. I forbindelse med den igangværende akkreditering af universitetsuddannelserne er de anvendte evalueringsmetoder derfor en af de parametre, uddannelserne vurderes på.

Hensigten med nærværende artikel er at diskutere evalueringens konsekvenser, samt hvorvidt de måder, hvorpå konkrete evalueringsprocedurer udformes, kan have betydning for deres resultater i form af, hvordan de studerende vægter deres evaluering, og hvad de i realiteten evaluerer, når de evaluerer undervisning. Det vil foregå i form af analyser af tre evalueringsprocedurer fra det Humanistiske Fakultet, Københavns Universitet¹. Afslutningsvis vil jeg give konkrete forslag til forskydning af evalueringernes fokus imod de studerendes evaluering af egen læring.

Indledning

Præcisering: Studerendes evaluering af undervisning vs. eksamen som evaluering

I en del af den forskningslitteratur, der igennem de senere år er produceret om evaluering, omtales den stigende interesse for evaluering som et tidstypisk fænomen (f.eks. Dahler-Larsen, 1998; Borgnakke (red.), 2008; Moldt, 2007). Evaluering er således ikke et nyt fænomen, men et fænomen der særligt siden 1970'erne har været i stigning i Danmark (se f.eks. Rasborg m.fl. (red.), 1977), og som igennem de senere år har antaget nye former. Dette udmønter sig i, at der evalueres overalt: I og af organisationer og medarbejdere, af den offentlige sektor og dens ydelser og i

undervisningssammenhæng, ligesom selvevalueringen, som man f.eks. finder den i folkeskolens elevplaner, logbøger og porteføljer, er i stigning (Borgnakke, 2008; Dahler-Larsen, 1998; Dahler-Larsen 2002; Dahler-Larsen, 2003).

Følger man litteraturen om evaluering (f.eks. Borgnakke, 2008; Miller, 2005; Dahler-Larsen, 1998) opdager man således hurtigt, at begrebet 'evaluering' anvendes om en række forskellige aktiviteter, der alle etableres med henblik på vurdering eller bedømmelse. Aktiviteter, der er så forskellige som evaluering af undervisning og eksamen, betegnes begge som 'evaluering' (Borgnakke, 2008). Dette sker, idet man skelner imellem *formativ* og *summativ* evaluering. Mens den formative evaluering er de typer evaluering, der afholdes løbende med henblik på at ændre på undervisning eller indsats, er den summative en afsluttende evaluering, der, i og med det tidspunkt hvorpå den afholdes, får karakter af en bedømmelse.

I en historisk redegørelse over evaluering som fænomen beskriver Tanja Miller, hvordan vurderinger af, hvad eleverne har tilegnet sig i undervisningen, betegnes som *bedømmelse* frem til 1970'erne for derefter i et betydeligt omfang at blive betegnet som evaluering (Miller 2005, s. 66)². Hun fremfører endvidere at: »Skiftet i ordbrug er en konsekvens af et skift i synet på undervisning og læring. Et meget grundlæggende skift i antagelser om, at undervisning og læring ikke kan sættes på den samme ligning« (ibid.).

Dertil kommer, at den progressive pædagogik, der på dansk universitetsniveau satte sig igennem i etableringen af universitetscentre i Aalborg og Roskilde, betragtede eksamen som 'direkte skadelig for den pædagogiske og videnskabelige virksomhed', som det hedder i *Betænkning om den Samfundsvidenskabelige Basisuddannelse* på Roskilde Universitetscenter (1972; her efter Hansen, 1997, s. 68).

Her skelnede man imellem *intern* og *ekstern* evaluering. Mens den interne evaluering blev foretaget de involverede parter imellem og ofte løbende, fik eksamen betegnelsen *ekstern evaluering*, hvilket blandt andet hænger sammen med, at den indbefatter ekstern censur. Samtidig blev ekstern evaluering udråbt som 'det onde', mens intern evaluering blev om ikke 'det gode' så dog en rimelig foranstaltning (ibid.; Miller 2005: 70). Retorikken kan stadig genfindes på universitetscentre, hvor projekteksamen også i dag betegnes 'evaluering', ligesom man frem til nu ihærdigt

vendelse af evaluering til udøvelse af ledelsesopgaver omfatter det forhold, at nogle undervisere er bedre til at undervise end andre (ibid.s. 449). Endvidere anfører Moldt det forhold, at evalueringer og det 'kunde-forhold', som de studerende herigennem tildeles, kan betyde, at undervisere oplever at blive ansvarliggjort for de studerendes ansvar for egen læring (ibid.s. 465).

Analysegenstand og metode

Mens Moldt fortrinsvis betragter disse problemer som forankrede i ledelsens måder at introducere og anvende undervisningsevalueringer på, mener jeg, at en stor del af problemerne er forankrede i udformningen af de konkrete evalueringsmetoder, der anvendes. Dette mener jeg illustreres igennem analyser af tre forskellige evalueringsskemaer, der alle anvendes på Det Humanistiske Fakultet, Københavns Universitet. De tre evalueringsskemaer er udvalgt ud fra, at de på forskellig vis er typiske i den forstand at mange andre evalueringsskemaer på andre institutter har stor lighed med disse. De tre procedurer er anonymiserede.

I analysen forholder jeg mig til tre parametre: 1) Den anvendte metode, hvilket vil sige, *hvordan* der evalueres. 2) Evalueringens genstand, hvilket vil sige, *hvad* der evalueres. 3) Evalueringens formål, hvilket vil sige, *hvorfor* der evalueres. Analysen inddrager hverken den undervisning, der evalueres på eller intentionerne bag de enkelte evalueringsskemaer, men fokuserer alene på det, der kan fremanalyses igennem læsninger af de nedfældede procedurer. I relation til spørgsmålet om formålet med og anvendelsen af evalueringen, knytter jeg an til Dahler-Larsens skel imellem kontrol-evaluering og læringsevaluering (Dahler-Larsen, 1998 s. 99) som to forskellige formål for og anvendelser af evaluering.

Dahler-Larsen anfører således, at det kan gøre en forskel, hvad man sætter som formålet med evalueringen: Læring eller kontrol. Læringsevaluering har til formål at undersøge, hvad eleverne har lært, hvad de selv har forståelsen af at have lært eller med andre ord, hvordan det kursus, der evalueres, kan indplacere i deres forståelse af personlig progression på studiet. Formålet med læringsevaluering er endvidere at bidrage til udvikling af den aktivitet, der evalueres. Kontrol-evaluering har derimod til formål at kontrollere forløbet ved f.eks. at undersøge, om underviseren har passet sit arbejde, og om de studerende har deltaget i undervisningen. Der kan vel næppe herske tvivl om, at det primære fokus for en uddannelsesinstitution bør være læring og udvikling af såvel deltagerne som af aktiviteterne (Andersen og Søndergaard, 2006).

Mens Dahler-Larsen diskuterer skellet imellem læring og kontrol som spørgsmål, der bør medtænkes i udformningen af evalueringsskemaet, hvilket vil sige præskriptivt, vil jeg her anvende skellet analytisk til at undersøge, hvilke typer evalueringsskemaer der

i realiteten er iværksat i forbindelse med undervisning på Humaniora.

Evaluering ved institut X: Individuelle evalueringsskemaer med tilfredshedsskalering

Den første evalueringsskema, jeg vil diskutere, er en type, der er meget udbredt generelt og ikke kun på Københavns Universitet. Den består af et skema, der uddeles til alle de studerende, sådan at de kan besvare det individuelt. Her bliver den studerende bedt om at vurdere forskellige elementer i undervisningen og indplacere bedømmelsen i Likert-skalaer på fem trin fra 'meget godt' til 'meget dårligt' (se fig. 1). Derudover er der mulighed for at tilføje uddybende kommentarer til de enkelte elementer. Skemaet skal udfyldes midtvejs i semestret under selve undervisningen, hvorefter et par af de studerende påtager sig (eller bliver bedt om) at sammenfatte skemaerne og fremlægge resultatet til uddybende diskussion mellem holdet og underviseren. Den endelige sammenfatning indleveres til Studienævnet.

Metoden med individuelle besvarelser af et spørgeskema giver de studerende optimale muligheder for anonymitet. Der opfordres således ikke til at oplyse andet end kursus/holdnummer og underviser. Endvidere giver kombinationen af Likert-skala og opfordring til uddybende kommentarer mulighed for både kvantitativ og kvalitativ feedback. Den kvalitative feedback kan ligeledes tilføjes under fremlæggelsen af sammenfatningen på holdet og den diskussion, der her kan rejses. Mens det kvantitative design letter sammenfatning og opsummering, giver det kvalitative islet mulighed for at indfange de studerendes subjektive forståelser og tilføje nuancer. Disse er som oftest de mest oplysende for underviseren.

Imidlertid kræver det en relativt høj grad af indsigt og refleksion at kunne besvare spørgsmålene i skemaet fra Institut X. Der spørges således til dels, hvorvidt kursets formål er klart for den studerende, dels til hvordan man vurderer det valgte tekstmateriale ud fra graden af relevans. Mens det første spørgsmål kræver, at den studerende har sat sig ind i studieordningens formulering af faget og enten løbende holdes opdateret med eller selv kan vurdere det, man er blevet præsenteret for i relation hertil, kræver det andet spørgsmål et overblik, som man vanskeligt kan forvente hos særligt nye studerende. Det er således sjældent umiddelbart gennemskueligt, hvordan hver enkelt tekst relaterer sig til et overordnet formål.

Endvidere skal denne evalueringsskema ses i lyset af, at evalueringen afholdes midtvejs i forløbet. Det betyder, at de studerende skal evaluere forløbet, før de har haft mulighed for at kunne vurdere det i sin helhed. For det første må dette anses for at være en metodologisk svaghed, der bidrager med et betydeligt

Kursus _____ Hold nr. _____
 Underviser _____
 Søst Ærlyde: _____

1. Beskriv hvad du synes er godt ved kurset:

2. Beskriv hvad du mener kunne forbedres ved kurset:

3. Er kursets **formål** klart for dig?
 meget klart klart hverken/eller uklart meget uklart
 Uddybende kommentarer (herunder om kursets relevans for studiet som helhed):

4. Svarer det **faglige niveau** i undervisningen til dine behov?
 alt for højt for højt passende for lavt alt for lavt
 Uddybende kommentarer (herunder om niveauet er for højt eller lavt):

5. Hvordan vurderer du det valgte **tekstmateriale**?
 meget relevant relevant hverken/eller irrelevant meget irrelevant
 Uddybende kommentarer:

Vend

6. Finder du **undervisningsformen** passende i forhold til kursets faglige indhold?
 meget passende passende mindre passende
 Uddybende kommentarer (f.eks. holdundervisning vs. forelæsning vs. studenteroplæg)

7. Hvordan vurderer du **underviserens indsats**?
 meget god god middel mindre god dårlig
 Uddybende kommentarer:

8. Hvordan vurderer du **din egen indsats** (herunder fremmøde og forberedelse)?
 meget god god middel mindre god dårlig
 Uddybende kommentarer:

9. Hvordan vurderer du **holdets indsats** (herunder fremmøde og konstruktiv deltagelse i undervisningen)?
 meget god god middel mindre god dårlig
 Uddybende kommentarer:

10. Hvordan vurderer du **alt i alt kurset**?
 meget højt højt Middel lavt meget lavt
 Uddybende kommentarer:

Evt. yderlige kommentarer:

Fig. 1

usikkerheds-element i evalueringen. I realiteten kan det være meget vanskeligt at vurdere et kursus, før det er afsluttet og – ifølge mange studerende – reelt for eksamen. For det andet kan det give de studerende den (oftest fejlagtige) opfattelse, at deres evaluering vil kunne resultere i, at der umiddelbart foretages ændringer – f.eks. at tekstmateriale og emner udskiftes i forhold til de oprindelige planer.

Evalueringens genstand er i dette tilfælde i fremtrædende grad undervisningens kvalitet. Spørgsmål som ‘Svarer det faglige niveau i undervisningen til dine behov?’ og ‘Hvordan vurderer du det valgte tekstmateriale?’ lægger således op til de studerendes tilkendegivelser af, hvorvidt de er tilfredse med undervisningen. Derudover peger flere af spørgsmålene på, at evalueringens genstand her tillige er undervisernes kompetencer og arbejdsindsats. Dette ses af spørgsmål som ‘Finder du undervisningsformen passende i forhold til kursets faglige indhold?’ og ‘Hvordan vurderer du underviserens indsats?’. Mens den pågældende evalueringsskema således fokuserer på undervisningen og underviseren ud fra de studerendes tilfredshed, er der ingen spørgsmål, der opfordrer de studerende til at reflektere over egen læring. Det nærmeste, man kommer, er spørgsmål vedrørende den studerendes egen arbejdsindsats og holdets samlede indsats (fig. 1).

Anskuet ud fra spørgsmålet om, hvorvidt evalueringen kan bidrage til ændring og udvikling af det pågældende modul, foreligger der imidlertid et vist

læringspotentiale. Evalueringen kan således give oplysninger om, hvorvidt der er forhold som sværhedsgrad i tekster og undervisning, som man måske skal overveje med henblik på planlægning af næste semester. Ikke desto mindre er der her primært tale om en kontrol-evaluering, der, ud fra de studerendes tilkendegivelser af tilfredshed eller mangel på samme, kan anvendes til at kontrollere, hvorvidt underviseren og undervisningen har levet op til nogle forventninger, som de studerendes kan have mere eller mindre belæg i f.eks. studieordningen for at have.

Evaluering ved institut Y: Mundtlig evaluering varetaget af studerende.

Den anden evalueringsskema, der her skal analyseres, adskiller sig væsentligt fra den foregående. Her er der tale om en mundtlig evaluering, der afholdes midtvejs og afslutningsvis på kurset. Til dette formål har det pågældende institut formuleret en række spørgsmål, der skal diskuteres af de studerende (fig. 2). Diskussionen sammenfattes på tavlen af studerende fra holdet, sædvanligvis mens underviseren er til stede, og den endelige sammenfatning indleveres til Studienævnet.

Metodisk kan der være problemer med denne evalueringsskema i og med, at mundtligheden umuliggør anonymitet. Det kan forhindre de studerende i at tilkendegive deres oprigtige mening, såfremt denne er negativ eller ikke er i overensstemmelse med den

SPØRGSMÅL TIL BRUG VED MIDTVEJS- OG SLUTEVALUERING	
1)	Lever kurset op til kursusbeskrivelsen og den faglige målsætning?
2)	Hvordan har forholdet været mellem forelæsning og øvelsestimer (sammenhæng, mangel på sammenhæng?)
3)	Hvordan er sammenhængen mellem læsepensum og undervisning?
4)	Det faglige niveau – for højt, for lavt, tilpas? Begrund hvorfor.
5)	Hvordan har studenteroplæggene fungeret i forhold til kurssets overordnede, faglige målsætning? Hvad kan evt. gøres bedre? Underviserens indsats i forhold til at informere de studerende om, hvordan man holder oplæg, underviserens efterfølgende respons, de studerendes indsats i forhold til at formidle stoffet?
6)	Hvor lang tid har de studerende brugt på at forberede sig? Hvordan forbereder de studerende sig? Læsegrupper, opgavegrupper e. lign.?
7)	Hvordan har dialogen været mellem de studerende og underviseren? Har underviseren givet plads for de studerendes aktive deltagelse? Har de studerende deltaget aktivt i og taget ansvar for undervisningen dér, hvor det ifølge studieordningen er påkrævet (dvs. i alle kursusformer bortset fra forelæsningsrækker)?
8)	Hvad synes de studerende/underviseren, at der skal laves om?

Fig. 2

forståelse, som der i øvrigt er skabt konsensus om på holdet. Mundtlige evalueringer er således under stærkere socialpsykologisk indflydelse end skriftlige. Derudover kan der være de førnævnte problemer med den evaluering, der ligger midtvejs i forløbet: Det kan være vanskeligt for de studerende at overskue forløbet i dets helhed, og de kan få en fejlagtig opfattelse af at kunne få indflydelse på kurssets fortsatte forløb. Endvidere kan det stærkt kvalitative design give problemer med sammenfatning og opsummering af evalueringens resultater. Udsagnene kan således være vanskelige at gruppere på en måde, der kan illustrere hovedtendenserne i evalueringen.

Evalueringens genstand er i dette tilfælde i mindre grad end ved evalueringsproceduren fra Institut X undervisningens kvalitet. I stedet er der, med spørgsmål om forhold og sammenhæng mellem forelæsninger og øvelsestimer samt imellem læsepensum og undervisning, mere fokus på sammenhæng imellem forskellige elementer i undervisningen. Underviserens kompetencer og arbejdsindsats er dog genstand for evalueringen på en anden måde end i den foregående procedure. Mens der i førstnævnte procedure er fokus på underviserkompetencer og forberedelse, opfordres de studerende i dette tilfælde primært til at evaluere underviserens indsats i forbindelse med information og dialog (fig.2). Men som det er tilfældet i evalueringens proceduren fra Institut X spørges der heller ikke på Institut Y til de studerendes forståelse og indsats i relation til egen læring. De spørgsmål, der stilles til de studerendes egen deltagelse, har stor lighed med dem, der stilles på Institut X og opfordrer alene de studerende til at reflektere over egen og holdets samlede arbejdsindsats.

Genstandsmæssigt adskiller evalueringsproceduren på Institut Y sig fra evalueringsproceduren på Institut X

ved at vægte de studerendes feedback på, hvad der skal ændres frem for tilkendegivelser af grad af tilfredshed med tingenes aktuelle tilstand. Ændringsperspektivet i spørgsmålene kan tolkes som, at evalueringen har til formål at bidrage til udvikling af modulet. I princippet kunne det pege på et læringsperspektiv i form af udvikling af modulet. I realiteten betyder det dog ikke, at der i mindre grad er tale om en tilfredshedsevaluering end på Institut X. Her vurderes underviserens kompetencer ud fra, hvorvidt det har ladet sig gøre at skabe og formidle sammenhæng og overholde formalia. Dermed forskydes underviserens primære opgave til orientering om formalia frem for indholdsmæssig undervisning. De studerende opfordres ikke til at reflektere over egen læring og det, der ydes – eller skal ydes – i denne sammenhæng. Derimod formidles en forståelse af, at læring understøttes af procedurer og information, og at det er underviserens opgave at varetage dette. De studerende er passive – og vurderende – brugere.

Evaluering ved Institut Z: 'Delfi-metoden'

Den tredje og sidste evalueringsskema, der skal evaluere i denne sammenhæng, præsenteres af det pågældende institut som evaluering efter 'delfi-metoden' (fig.3).

Dette indbefatter en skriftlig evaluering, der foretages på ét samlet ark for hele holdet. Her opfordres de studerende til at tilkendegive deres mening i form af udsagn, som deres medstuderende kan tilslutte sig ved at sætte streger på arket. Når papiret har været to gange rundt på holdet, har alle haft mulighed for at læse og evt. tilslutte sig alle udsagn. Metodisk har denne evalueringsskema den fordel, at udsagnene som udgangspunkt kan holdes anonyme. Dog kender de studerende formentlig hinanden tilstrækkeligt godt til at kunne identificere, hvem der er afsender af de enkelte udsagn. Samtidig kombineres det kvalitative og det kvantitative i designet, hvor udsagn kun i et begrænset omfang er styret af instituttets formalia.

Problemet er imidlertid, at udsagnenes relevans ikke er sikret med denne metode. Det bliver således noget grumset, hvad der egentlig evalueres. Holdets dynamik og de studerendes indbyrdes fordeling af popularitet kan være afgørende for ikke alene hvilke udsagn, der fremkommer, men også for de øvrige studerendes grad af tilslutning. Det er heller ikke utænkeligt, at de første udsagn, i og med den svage styring fra instituttets side, kan være retningsgivende for, hvorvidt evalueringen bliver udpræget positiv eller negativ.

Evalueringens genstand er her kun i begrænset omfang afgjort af formalia, hvilket vil sige, at det er de studerende (holdet), der definerer denne. Dog har studienævnet indført enkelte kategorier, som f.eks. semesterplan, kursusplanlægning, brug af AV-udstyr m.m. samt information om eksamen, som man ønsker belyst i evalueringen ud fra, hvad der er godt, og hvad der kan forbedres. Når dette er anført som punkter for evalu-

ering, må det antages at være ud fra overvejelser over, at oplysninger herom vil kunne bidrage til fremtidig regulering og udvikling af modulet.

Mens undervisningens kvalitet og underviserens kompetencer som underviser som omtalt på forskellig vis var genstand i evalueringsprocedurerne fra Institut X og Y, er dette således ikke nødvendigvis tilfældet her. Her er det i et betydeligt omfang op til de studerende at afgøre, hvad der skal fokuseres på. Dette gør det dog også ret usandsynligt, at de studerende vil fokusere på deres egen indsats i relation til egen læring. Det løse design til trods fremstår proceduren derimod, qua de indledende spørgsmål vedrørende opfyldelse af formalia på kurset, som kontrolevaluering, der muligvis, men ikke nødvendigvis, kan anvendes til udvikling af det pågældende modul.

Evaluering som rituel refleksion

Som det fremgår af de tre analyser af evalueringsprocedurer fra tre institutter, er det meget begrænset, hvor meget der opfordres til, at de studerende evaluerer deres egen indsats og læring. På denne vis friholdes de så at sige af evalueringen. Den bliver 'gratis' for dem; de skal alene bedømme deres undervisere. Derefter er det op til den enkelte underviser eller det enkelte studienævn at udlede, om evalueringen kan anvendes til at ændre på og udvikle det pågældende modul fremover. Først her afgøres det, om evalueringen kan give anledning til læring forstået som udvikling af undervisningens form og indhold, og hvordan evalueringen anvendes som ledelsesredskab.

Ligesom opfølgningen på evalueringerne ikke på forhånd er givet, er der ikke nødvendigvis en sammenhæng imellem det, man må formodes at ønske at evaluere og det, der reelt evalueres med de pågældende foranstaltninger. Evalueringshandlingerne fremstår således mere end noget andet som det, Dahler-Larsen betegner 'rituelle refleksioner'. I bogen *Den rituelle refleksion* redegør Dahler-Larsen for, hvordan megen evaluering i organisationer, herunder på universitetet, i realiteten er rituelle handlinger, der kun effektueres, fordi man skal, men uden reel interesse for hverken hvad, man ønsker at undersøge, hvordan evalueringen skal udformes eller hvad den viser: »Pointen er, at evalueringsproceduren skyldes et normativt krav om en evalueringsprocedure, ikke et rationelt krav om at forbedre programmets funktion og målopfyldelse.« (Dahler-Larsen, 1998 s. 92).

Oftentimes anvendes evalueringerne således ikke til noget, men arkiveres, ligesom de drøftelser, de eventuelt måtte afføde, ofte ikke kommer til de involveredes kendskab. Dahler-Larsen omtaler dette som, at evalueringen er løst koblet til praksis i organisationen. Her bruger han bl.a. universitetet som eksempel, idet han kritiserer, at man kun evaluerer på undervisning, men ikke på de akademiske medarbejders øvrige aktiviteter (ibid.s. 83). Samtidig fremhæver han, at mange evaluerings-

procedurer er begrænsede af institutionelle regler og konventioner for, hvad man kan, og hvad man ikke kan evaluere. På Københavns Universitet synes de studerendes læring at falde uden for kursusevalueringen og således alene være et forhold, der bedømmes til eksamen. Her evalueres den kun summativt, mens den formative evaluering af læreprocessen, der kunne have været inddraget i form af spørgsmål til de studerendes egen indsats i relation hertil, ikke berøres.

Problemet er samtidig, at evalueringerne ikke er neutrale i den forstand, at de ikke har en effekt for praksis. Som Dahler-Larsen påpeger, er de anvisende for praksis i den forstand, at positive eller negative evalueringer eller forventningerne til samme kan fremme eller blokere for nye tiltag i undervisningen. Dette fænomen omtales tillige af Rikke von Müllen: »Evalueringer man år efter år tilfredshed og lader disse evalueringer få konsekvenser for genansættelser mv., kan man tværtimod opdrage de studerende til, at deres tilfredshed eller velbehag er kriteriet for god undervisning ...« (von Müllen, 2006 s. 17). Von Müllen påpeger hermed problemet i at designe evalueringer, der har en form, som de tre her omtalte.

Et af de presserende problemer i denne forbindelse omtales af Hanne Leth Andersen og Louise Søndergaard, der i forbindelse med en undersøgelse af evalueringspraksis på Århus Universitet omtaler den måde, hvorpå evaluering af undervisningen i realiteten oftest forskydes til at være en evaluering af *underviseren* (Andersen og Søndergaard, 2006):

»Det er undersøgelsens mest markante resultat, at de spørgsmål, der blev stillet, i 40 % af tilfældene lagde op til, at de studerende skulle vurdere og bedømme underviserens faglige og pædagogiske formåen. Den oplysning rejste en række principielle spørgsmål, der tematiserede rollefordelingen mellem underviser og studerende og deres forventninger til hinanden i forbindelse med undervisningen.« (ibid.s. 23)

Dette svarer i stor udstrækning til det, jeg har fundet i nærværende analyser, hvilket er interessant, eftersom der er tale om to forskellige universiteter. Noget tyder på, at der kunne være tale om generelle problemer i udformningen af evalueringsprocedurer og deres relevans i forbindelse med det, de er foranstaltet til at evaluere. Dette problem svarer til det, Dahler-Larsen omtaler som en løs kobling imellem praksis og evaluering af praksis. Med henblik på at undgå den svage relation, anbefaler Dahler-Larsen, at man undersøger evalueringskriterierne validitet ved at overveje, om de ville fungere, hvis de var retningsanvisende for et 'program', f.eks. et kursus/en uddannelse (Dahler-Larsen, 1998 s. 96).

I de her omtalte tilfælde, hvor genstanden for evalueringen i betydeligt omfang er identificeret som kontrol og identifikation af de studerendes tilfredshed med undervisningen og underviserens indsats, må eva-

lueringskriterierne vurderes ud fra, hvorvidt det vil være hensigtsmæssigt at lade dette være styrende for planlægning af undervisning. Det er muligt, at disse parametre vil kunne anvendes af den enkelte underviser med henblik på tilretning af elementer i et modul. Det kunne f.eks. være udskiftning af enkelte tekster eller ændringer i undervisningsformen. I forlængelse heraf vil jeg imidlertid anbefale, at evalueringen i større omfang retter sig imod, at de studerende skal overveje deres egen indsats og læring. Først herigennem gøres de ansvarlige for samme, ligesom den formative evaluering heraf ville kunne anvendes af underviseren med henblik på udvikling af undervisningen. Hermed øges også dette læringspotentiale i evalueringen.

Konklusion: Udformning af læringsevaluering.

Mit bud på, hvordan man kan inddrage evaluering af de studerendes læring, indebærer først og fremmest en opfordring til at overveje, hvad det er, der spørges til i evalueringerne. Dertil kommer spørgsmålenes konkrete udformning. Som det understreges af Andersen og Søndergaard: »Hvis læring og bevidsthed om læreprocesser er i fokus, er det nødvendigt at bede de studerende om at reflektere over undervisningen med udgangspunkt i dette overordnede mål med undervisningen: deres egen læring og faglige udvikling« (Andersen og Søndergaard, 2006 s. 23).

Dette kan blandt andet gøres ved at spørge til nogle af de parametre, der sædvanligvis konnoteres med læring: Refleksion, tilegnelse, motivation og progression. Et forslag til udformning af en evaluering, der retter sig imod de studerendes egen læring, kunne således starte ved at formulere spørgsmål som nedenstående:

På hvilken måde har undervisningen/forelæsningerne bidraget til, at du føler, at du har lært noget på dette modul?

- Har du oplevet, at du bliver bedre til det faglige – f.eks. har lettere ved at læse teksterne eller markere dig i diskussioner på holdet?
- Hvad har du selv gjort for at lære mest muligt? – kvantitativt: læst pensum, løst opgaver, deltaget i diskussioner, mødt op til undervisningen?
- På hvilken måde kan du sammentænke det, du har lært på dette kursus med tidligere eller samtidige moduler?
- På hvilken måde har modulet givet dig lyst til at fortsætte dine studier?
- Osv.

Som det fremgår, må de spørgsmål, der anvendes i evalueringen, lægge op til de studerendes selvrefleksion frem for tilkendegivelser af tilfredshed, og hvorvidt underviseren har overholdt formalia, passet sit arbejde og været passende underholdende og ikke for krævende samt udvalgt tekster af et passende sværhedsniveau.

Selv om det kan forekomme i overensstemmelse med idealer om studenterdemokrati, er det således heller ikke hensigtsmæssigt at overlade det til de studerende frit at vurdere, hvad der skal evalueres på, sådan som tilfældet er i 'delfi-metoden'. Samtidig må der gives mulighed for både kvantitative og kvalitative svar, samt at de studerende kan indgå i en dialog med underviseren eller, hvis de ønsker det, kan forblive anonyme i selve evalueringen. På denne vis vil det være muligt for både studerende og undervisere at øge værdien af undervisningsevaluering i form af udvikling af egne kompetencer og undervisningens kvalitet.

Referencer

- Andersen, Hanne Leth og Louise Søndergaard (2006). Undervisningsevaluering som redskab til kvalitetsevaluering af undervisning. I: *Dansk Universitetspædagogisk Tidsskrift* nr.2/2006. pp.22-35.
- Borgnakke, Karen (red.) (2008). *Evalueringens spændingsfelter*. Århus: Klim.
- Borgnakke, Karen (2008). Evalueringsstrategier i den pædagogiske kontekst. I: Borgnakke, Karen (red.). *Evalueringens spændingsfelter*. Århus: Klim. pp.9-66.
- Christensen, Gerd (2008). Eksamen som en kamp om positioner. I: *Unge Pædagoger* nr.5/2008. pp.17-23.
- Christensen, Gerd (2006). Projektpædagogikkens didaktik – en kritisk diskussion. I: *Nordisk Pedagogik*, 1/2006 pp.30-47.
- Dahler-Larsen, Peter (1998). *Den rituelle refleksion*. Odense: Odense Universitetsforlag.
- Dahler-Larsen, Peter (2002). *Evaluering kortlagt*. Århus: System Academic.
- Dahler-Larsen, Peter (2003). *Selvevalueringens hvide sejl*. Odense: Syddansk Universitetsforlag.
- Hansen, Else (1997). *En koral i tidens strøm*. Frederiksberg: Roskilde Universitetsforlag.
- Lejre, Thomas & Susanne Murning (2004). *Undersøgelse af karaktergivningens betydning for projektarbejdet på Humbas*. RUC.
- Miller, Tanja (2005). Hvordan synes du selv, det går. I: *Unge Pædagoger* 2/3, 2005. pp.64-79.
- Moldt, Christian (2007). Fra meningsløse til meningsfulde performancemålinger – en case om kvalitetsmåling i undervisningssektoren. I: *Økonomistyring og Informatik*, 22. Årgang, 2006/2007 nr.5.
- Rasborg, Finn, Jesper Jensen, Gustav Leunbach (red.) (1977) *Evalueringproblemer*. København: Munksgaard (i samarbejde med Danmarks pædagogiske Institut).
- Von Müllen, Rikke (2006). Undervisningsevaluering og udvikling af underviserkompetencer. I: *Dansk Universitetspædagogisk Tidsskrift* nr.2/2006. pp.16-21.
http://www.globalisering.dk/multimedia/Kap_7.pdf. d. 2/6-2008.

Noter

- 1 Artiklen er blevet til i forlængelse af et oplæg, jeg holdt på Det Humanistiske Fakultets kursus i adjunktpædagogikum. Den er således blevet til efter forespørgsel fra flere af adjunkterne, der mente, at der skulle større fokus på de problemstillinger, der her omtales.
- 2 Miller nævner imidlertid ikke, at man frem til midten af det 20. århundrede betegnede bedømmelsen af elevernes læring som 'skudsmål'; et begreb, som i større omfang refererer til elevens personlighed og moral, end de senere anvendte betegnelser.

har kæmpet imod 'eksterne evalueringer' (Christensen, 2008; Hansen, 1997; Lejre & Murning, 2004).

I denne værdisætning genfinder man til dels skellet imellem formativ og summativ evaluering, men uden at det er muligt at sætte absolut lighedstegn imellem intern og formativ og ekstern og summativ evaluering. Derudover tilføjer man med projektarbejdet fokus på skellet imellem *proces-* og *produktevaluering* (Borgnakke, 2008; Christensen, 2006). Mens traditionelle eksaminer vurderer produktet, er projektpædagogikken traditionelt forpligtet til at vægte en evaluering af processen (ibid.).

I og med at der er foregået et skred i forståelsen af læring, som noget elever og studerende selv tager aktivt del i, er der, som Miller udtrykker det, foregået et skift til et konstruktivistisk lærings syn og dermed skabt en øget fokusering på det individuelle subjekt (Miller, 2005 s. 73). Dermed bliver elevernes og de studerendes oplevelser af undervisningen og de muligheder, denne stiller til deres rådighed, anset for mere væsentlige end tidligere (ibid.). Dette kan forklare, hvordan der er sket et skred fra, at det alene var underviserne, der havde retten, magten og pligten til at evaluere de studerende i form af 'vidnesbyrd' (årskaraktør) og eksamen, til at en stor del af den nuværende evalueringsaktivitet på universitetsniveau foregår igennem de studerendes evaluering af undervisningen – hvilket vil sige den form for evaluering, der er genstand for nærværende artikel.

Evalueringens betydning

Igennem de senere år er de officielle krav om, at studerende skal evaluere undervisningen og at evalueringerne skal offentliggøres på universiteternes hjemmesider, steget. Således hedder det i regeringens globaliseringsstrategi: »De studerendes evalueringer af undervisningen skal offentliggøres: Systematiske og sammenlignelige kursus- og undervisningsevalueringer skal offentliggøres på universiteternes hjemmesider.« (http://www.globalisering.dk/multimedia/Kap._7.pdf s. 68). I takt hermed har institutterne på Det Humanistiske Fakultet på Københavns Universitet øget formaliseringen af praksisser for evaluering og dokumentation heraf i form af forskellige evalueringsprocedurer og -skemaer samt efterfølgende afrapportering til det pågældende Studienævn.

Set fra studenterside kan der være fordele ved at få mulighed for at udtrykke sin tilfredshed eller mangel på samme ved en netop afholdt undervisning. Set fra underviserside kan der ligeledes være fordele ved at få indblik i de studerendes opfattelser af, hvorvidt undervisningen har været på et passende niveau, og om de studerende mener at have lært noget. I bedste fald kan dette anvendes til at ændre undervisning eller indhold sådan, at niveauet er bedre tilpasset de studerendes ønsker og kvalifikationer.

Problemet er imidlertid, at mens niveauet for eva-

lueringspraksis er intensiveret, forekommer refleksionerne over praksis at ligge på et relativt lavt niveau. Egentlige diskussioner af institutternes evalueringsprocedurer forekommer således relativt sjældent og mig bekendt aldrig i en sammenhæng, hvor der diskuteres på tværs af studienævn og institutter. Det betyder blandt andet, at erfaringsudvekslingen er begrænset og at frustrationer i forbindelse med evaluering individualiseres og bliver den enkelte undervisers problem. På denne vis kan uhensigtsmæssige evalueringsprocedurer vedblive med at figurere i systemet uden, at *de* bliver evalueret. Dette er et problem, fordi en uhensigtsmæssig evalueringsprocedure kan betyde, at undervisning og undervisere evalueres på måder, som gør det vanskeligt at anvende evalueringen til forbedring af undervisningsindsatsen. Mens formålet med evalueringen fortaber sig i det uvisse, bliver evalueringen nyttesløs (Borgnakke 2008, s. 55). Jeg vil ligeledes hævde, at skellet imellem positive ('gode') og negative ('dårlige') evalueringer ikke kun er et spørgsmål om undervisningens kvaliteter men også om evalueringsprocedurens udformning.

Samtidig med at en uhensigtsmæssig evalueringsprocedure altså kan gøre selve evalueringens resultat uanvendeligt, er der etiske forhold, som man må tage med i betragtning, når der vælges evalueringsmetode. Det er således ikke uden omkostninger for underviserne at lade sig evaluere. Ikke mindst unge og mindre erfarne undervisere lader sig påvirke af, hvad de studerende mener om deres undervisning – for ikke at tale om, hvad de mener om dem som undervisere (Moldt, 2007 s. 448).

Mens en positiv evaluering kan bekræfte underviseren i, at vedkommende har valgt den rigtige metier, kan en negativ evaluering for det første give anledning til, at underviseren begynder at tvivle på sig selv og egne kompetencer som underviser. Det rejser således spørgsmål som: Kan jeg overhovedet undervise? Egner jeg mig til mit job? For det andet betyder det forhold, at der skal aflægges rapport til uddannelsens studienævn, at underviseren kan blive usikker på, hvad evalueringen skal bruges til. Det rejser spørgsmål som: Bliver jeg nu fyret eller...? Denne usikkerhed stiger naturligvis i takt med det øgede krav om synlighed: Vil den negative evaluering være synlig for alle på nettet? I denne sammenhæng må det tages i betragtning, at det at evaluere i princippet er ganske 'gratis' for de studerende. Evalueringen er oftest anonym, hvilket betyder, at den studerende kan skrive hvad som helst uden at skulle stå til ansvar for det.

Den angst, som de uklare formål med evalueringer kan medføre, har Christian Moldt identificeret i et forskningsprojekt om evalueringer på mellemlange uddannelser (Moldt, 2007). Et af de forhold, som underviserne pegede på, var således 'frygt for mikro-politisk misbrug af resultaterne' (ibid.: 439). Moldt anfører her en række tabuer omkring evaluering, der ud over an-