

Karaktergivning – intuitiv ekspertise eller ‘viden i praksis’?

Nina Bonderup Dohn, adjunkt, PhD, Institut for Fagsprog, Kommunikation og Informationsvidenskab, SDU, har bl.a. beskæftiget sig med Karakterkommissionens arbejde.



Nina Bonderup Dohn er uddannet cand.mag. i filosofi (hovedfag) og fysik (bifag). Hun er PhD i læringsteori med en afhandling, der kobler erkendelsesteori og læringsteori med udgangspunkt i et Heidegger-Wittgenstein-inspireret syn på mennesket som aktør i meningsbærende handlingssammenhænge. Hun

er siden 2003 ansat som adjunkt i humanistisk informationsvidenskab. Forskningsområdet er viden og læring i uddannelsessammenhænge og i arbejdslivet, med særligt fokus på begreber som tavs viden, kropslighed, forståelse og refleksion og disses eventuelle ændrede rolle i IT-baserede læringsammenhænge. Hun har deltaget i den offentlige debat om PISA-undersøgelsen, senest som medlem af ekspertpanelet ved Folketingshøringen den 12. september 2005. Seneste publikationer indbefatter:

»Learning for Tomorrow's World – Assessing What?« (submitted) i *Læring i praksis – fremstruktureringen af et handlingsorienteret perspektiv*, Aalborg Universitet, 2005 (ph.d.-afhandling); »Bare det, at de kan få snakket ud om det – tavs viden som resonansbund for forståelse« i Frostholt (red.): *De Kompetente Frivillige*, 2005; »Læsekompetence – i praksis og i testsituationer« i *Dansk Pædagogisk Tidsskrift*, 2003, nr. 3; »Viden og læring i praktikerens praksis og på universitetet« i Jæger, K. (red.): *Projektarbejde og aktionsforskning*, 2002.

Reviewet artikel

Artiklen anskuer bedømmelse og karaktergivning som en praksis, der udføres af konkrete mennesker i konkrete situationer, og diskuterer på den baggrund, hvori evalueringskompetence består. Tesen er, at det er etisk og kompetencemæssigt problematisk at være en ‘dreyfusiansk’ intuitiv ekspert i karaktergivning, og at evalueringskompetence i stedet bør være en ‘viden i praksis’, forstået som en enhed af sprogligt artikulerbare mål, praktisk kunnen og personlige erfaringer med studenterpræstationer. En udskiftning af 13-skalaen hævdes at kunne være et incitament til igangsætning af refleksion over kriterier for bedømmelse.

Siden Karakterkommissionen i november 2004 fremsatte sit forslag til en ny karakterskala til erstatning af 13-skalaen¹, er spørgsmålet om karaktergivning kommet i fokus i den offentlige debat. Karakterkommissionens udspil er blevet skarpt kritiseret, især fordi det indeholder den negative karakter -3, og fordi det postulerer, at det er muligt med den foreslåede skala at udøve bedømmelse, således at denne på en og samme tid er både absolut og relativ². Disse kritikpunkter er væsentlige og udgør efter undertegnede mening rimelige argumenter for at afvise Karakterkommissionens konkrete oplæg til ny skala. En sådan afvisning indebærer imidlertid ikke nødvendigvis, at 13-skalaen skal fortsætte som bedømmelsesskala – spørgsmålet om, hvorvidt der er gode grunde til at udskifte den nuværende skala, er relativt uafhængigt af spørgsmålet om fordele og ulemper ved specifikke alternative forslag.

Formålet med denne artikel er at diskutere, hvori evalueringskompetence består, med udgangspunkt i anskuelsen af bedømmelse og karaktergivning som en praksis, som udføres af konkrete mennesker i konkrete situationer. Fokus er på spørgsmålet om, hvilken relation der er, og hvilken der bør være, mellem ‘intuitiv ekspertise’ og refleksion hos evalueringspraktikere. Tesen er, at det er etisk og kompetencemæssigt problematisk at være ‘intuitiv ekspert’ i kvantificeringen over andres præstationer, og at stadig refleksion over de kriterier, som kvantificeringen foregår på baggrund af, er nødvendig.

Set i denne sammenhæng får spørgsmålet om det hensigtsmæssige i at udskifte karakterskalaen en ekstra dimension, der samtidig sætter Karakterkommissionens begrundelser herfor i et nyt lys. Argumenterne, nationalt set, for erstatning af den eksisterende skala har været, at skalaen har været brugt uensartet i forskellige fag, at der hersker tvivl, mht. om den skal bruges absolut eller relativt, og at der kan påvises karakterglidning over tid³. Disse argumenter synes svage i sig selv, da problemer i karaktergivningspraksissen ikke *per se* opløses ved at udskifte en gammel skala med en ny. Imidlertid kan man spørge, om ikke en udskiftning kunne

fungere som et konkret incitament til igangsættelsen af lokalt forankret refleksion over og imellem bedømmelsespraksisser: Det synes naturligt i forbindelse med indførelsen af en ny bedømmelsesskala at overveje de nævnte problematiske forhold tillige med mere generelle spørgsmål, såsom hvilken sammenhæng der er mellem en kvantitativ vurdering ud fra en karakterskala og en kvalitativ, sprogligt formuleret bedømmelse; og hvilken relation der består mellem sprogligt udtrykte evalueringskriterier og faktiske evalueringer af konkrete studenterpræstationer.

Afgørende for, om udskiftningen af karakterskalaen kan have en sådan refleksionsinitierende effekt er imidlertid, hvorvidt der sideløbende med dens indførelse afsættes de fornødne ressourcer *til* refleksion – og vel at mærke til refleksion både praksisinternt og mellem fag. Afsættes der ikke sådanne midler, er der en overvejende risiko for, at bedømmere i den første lange tid rent faktisk blot vil foretage deres vurdering efter 13-skalaen, men 'oversætte' til den nye skala inden afsigelsen af karakteren. Sker dette, vil praksis fortsætte relativt ubesværet af udskiftningen – men dermed også relativt ubesværet af den katalyserende funktion, som en sådan principielt set vil kunne have. De problemer, der måtte eksistere i evalueringspraksis, vil derfor formentlig blive videreført, uden at blive taget op til overvejelse; måske endog uden at blive bemærket. Og dermed vil de positive muligheder, som karakterskalaudskiftningen principielt byder på for evalueringspraksis, reelt ikke have en chance for at blive virkeliggjort.

Evalueringssituationen – et indkredsede eksempel

I Keldorffs bog *Tæt på en gruppe* gengives en votering, der finder sted mellem eksaminator og censor efter en 3-timers gruppeeksamination i et projektarbejde⁴. Forud for eksaminationen havde censor på et kort møde med eksaminator udtrykt sig noget negativt om projektet, som han vurderede til et 8-tal. Den samlede karakter for projekt og mundtlig eksamen bliver imidlertid 10 – efter udspil fra censor.

Umiddelbart ville man måske forvente, at en sådan hævelse af karakteren kun kunne ske på baggrund af en lang, detaljeret diskussion af vægtningen af forskellige aspekter ved den skriftlige og mundtlige præstation sammenholdt med studieordningsdefinerede målbeskrivelser og evalueringskriterier. Dette er imidlertid ikke tilfældet. Selve eksaminationen er endt med, at eksaminator og censor hver for sig over for de studerende opridsede nogle positive og negative sider ved deres præstation, og disse formuleringer danner naturligvis til en vis grad baggrund for voteringen, også efter de studerende er gået ud. De danner i hvert fald baggrund i den forstand, at de en gang *er* udtrykte. Så længe et punkt ikke eksplicit modsiges af den anden part under voteringen, har den, der sagde det, derfor

ikke grund til at antage, at den anden er uenig. Men omvendt er der ikke tale om en fælles forhandlet enighed.

Spørgsmålet er, præcis hvilken rolle de opridsede punkter faktisk har i det videre forløb: Hverken lærer eller censor vender tilbage til dem under voteringen, og de sættes på intet tidspunkt i relation til fag- eller målbeskrivelser for det pågældende semesters projektarbejde. Der forekommer således ikke nogen *eksplicit* vægtning af de forskellige styrker og svagheder ved præstationen i forhold til formulerede kriterier, hverken specifikt studierelaterede eller generelt akademiske. Ej heller sættes præstationen eller aspekter ved den i relation til karakterskalaens definition. En mulig fortolkning er, at de forhold, som eksaminator og censor hver for sig betoner over for de studerende inden voteringen, slet ikke udgør enkeltpunkter, hvis værdi bedømmerne kan vurdere forholdsvis uafhængigt af andre enkeltpunkter i relation til fastsatte kriterier, for siden hen at summere dem til en generel, sammenfattende vurdering af præstationen. I stedet er bedømmernes vurdering en helhedsopfattelse af præstationen, og de forhold, de nævner for de studerende, er dem, der træder frem for dem som betydningsfulde, på baggrund af denne helhedsopfattelse.

Denne fortolkning understøttes af de udsagn, der faktisk fremkommer under voteringen. Diskussionen her går primært på, hvorvidt hele projektgruppen skal have samme karakter, eller om der skal differentieres mellem de studerende. Voteringen starter med, at censor siger »Jeg må sige ... jeg synes ... jeg er lige ved at være tilbøjelig til at synes, at de har hævet den op på et tital ... har de ikke det?« Der falder stort set ingen konkrete bemærkninger om eksaminationen; i stedet bruges udtryk som »der er alligevel en uensartethed«; »Leif er meget seriøs. Han er meget taktsom ...«; »Mads og Mette er sådan lidt svagere papirer, ikke?«; »på en eller anden måde synes jeg jo ikke at Mads er kommet ret langt ud over sin hverdagsbevidsthed.«; »og du hører Mads og Mette være lidt svagere, ikke ... men ikke eklatant«; »Ikke et trin svagere end de andre. Det er det, jeg har hørt og da helhedsindtrykket er positivt, fordi der helhedsmæssigt er lagt op ...«; »For det skal springe relativt meget i øjnene for at begynde at differentiere. Skal man differentiere hver enkelt. Nu sad vi og lavede to niveauer, ikke ... så synes jeg vi skulle være gået endnu tættere på...« »Det ville være min fornemmelse«.

Der er flere forhold, der er værd at bemærke i relation til disse kommentarer. For det første understøtter de tesen om, at bedømmerne har en helhedsopfattelse, både af den samlede gruppepræstation og af hver studerendes individuelle præstation, og at det er denne helhedsopfattelse, evalueringen angår, ikke atomistiske enkeltpunkter. For det andet har bedømmerne en helt klar 'føling' med karakterskalaen: De vurderer ikke først, hvilket kvalitativt niveau præstationen ligger på,

hvorefter de tilskriver en kvantitativ karakter; i stedet tænker de niveauerne gennem skalaen. Censor starter således med at foreslå en karakter og taler siden om at være 'et niveau svagere'. Med Merleau-Pontys ord er karakterskalaen et værktøj, som bedømmerne har inkorporeret, og som de derfor opfatter og handler i verden på baggrund af⁵. For det tredje kan de ord, bedømmerne anvender, for en umiddelbar betragtning synes ret vage og upræcise, og man kunne fristes til at spørge, hvorledes en angiveligt fagligt kompetent, rationelt begrundet evaluering kan basere sig på udtryk som 'svagere papirer' og 'hverdagsbevidsthed'. Dette spørgsmål vender jeg tilbage til siden hen. For nuværende er det imidlertid tilstrækkeligt at påpege, at voteringens forløb synes at være i modstrid med en opfattelse af karaktergivning som en proces, der bygger på en eksplicit udtrykt sammenligning af en given præstation og sprogligt formulerede evalueringskriterier.

En oplagt indvending vil her naturligvis være, at det gengivne eksempel af selv samme grund ikke er repræsentativt for evaluering generelt og især ikke for adækvat evalueringspraksis. At enighed i det konkrete tilfælde ikke opnås på baggrund af en diskussion af relationen mellem aspekter ved den givne præstation og studieordningsdefinerede målbeskrivelser er jo ikke et argument for, at evaluering ikke *burde* bestå i en sådan diskussion. Besvarelsen af denne indvending vil for en stor del ske i det følgende, hvor der mere teoretisk redegøres for, hvori evalueringskompetence som eksemplificeret i Keldorffs case består, og hvor der argumenteres for, at kompetente bedømmelser, der forholder sig mere eksplicit til kriterier, også trækker på en ikke sprogligt formulerbar fagforståelse. Det kan dog være rimeligt ganske kort at knytte et par fænomenologisk beskrivende bemærkninger til evalueringssituationen generelt til sandsynliggørelse af, at Keldorffs eksempel ikke er enestående.

Det blev foreslået ovenfor, at bedømmerne i Keldorffs case havde en helhedsforståelse af de studerendes præstation, og at det var på baggrund af denne helhedsforståelse, at de forhold, de valgte at pointere over for de studerende, trådte frem som væsentlige positive og negative sider. Generelt oplever den erfarne bedømmer ved gennemlæsning af en opgave eller ved tilstedeværelsen ved en mundtlig eksamen meget hurtigt en fornemmelse af eller 'føling med' niveauet for præstationen. Bedømmeren vil sædvanligvis tage notater undervejs, men som i Keldorffs eksempel er disse notater ikke enkeltpunkter, som til slut summeres til en bestemt kvantitativ eller kvalitativ vurdering. I stedet er det de punkter, der på baggrund af helhedsforståelsen giver sig som af særlig betydning, vel at mærke ikke uafhængigt af andre punkter, men tværtimod i relation til disse – punkterne giver sig som værende af større eller mindre betydning i forhold til hinanden. Helhedsforståelsen kan her rykkes af diskussionen af enkeltpunkter – men kun begrænset. Og det er der-

for, bedømmere efter en votering, der ikke forløb i lutter enighed, og hvor der herskede en del diskussion netop om enkeltpunkter, kan sidde tilbage med fornemmelsen af, at selvom man blev enige om disse enkeltpunkter, var bedømmelsen alligevel uretfærdig: På baggrund af hans/hendes helhedsforståelse fremstår situationen for bedømmeren som én, hvor der blev tillagt de udslagsgivende enkeltpunkter for stor positiv eller negativ vægt.

Evaluerer er en praksis

I fortsættelse af indvendingen ovenfor kunne man spørge, med hvilken ret det overhovedet betvivles, at adækvat evaluering består i en eksplicit, rationel argumentation om forholdet mellem en given præstation og formulerede evalueringskriterier. Konkrete eksempler på det modsatte, kunne man sige, tjener kun til at problematisere praksis, men rækker ikke ved idealet.

Denne indvending rammer imidlertid en anelse ved siden af. Pointen er ikke, at 'eksplicit, rationel argumentation' ikke kan være et led i adækvat evaluering. Pointen er i stedet den todelte, at en sådan argumentation for det første har en række basale forudsætninger, der må være opfyldt, for at argumentationen overhovedet er adækvat. Og for det andet, at *når* de basale forudsætninger er opfyldt, fungerer disse forudsætninger i højere grad som evalueringsgrundlag end den 'rationelle argumentation', hvorfor en sådan ikke er en *nødvendighed* for adækvat bedømmelse. 'Argumentatorisk sammenligning med evalueringskriterier' fordrer, at den, der foretager sammenligningen, *forstår* evalueringskriterierne, kan *se*, når de er opfyldt, kan *vurdere* graden af manglende opfyldelse, når de ikke er det, og kan *eksplicitere* sin vurdering i en 'rationel argumentation'. Evalueringskriterier viser, som Wittgenstein ville have sagt⁶, ikke selv deres anvendelse, og det kræver fagforståelse at vide, hvordan de skal appliceres på hvert enkelt konkret tilfælde. Og denne 'fagforståelse' på sin side er ikke givet ved yderligere kriterier eller regler – det ville give en regres af kriterier for kriteriers anvendelse⁷ – men er i stedet bundet op med erfaringen af konkrete eksempler på, hvad der i praksissen gælder som forskellige grader af opfyldelse af de beskrevne kriterier. At følge en regel er, som filosofen Wittgenstein siger, en praksis, der fordrer, at man forstår, hvad det *i* denne praksis *er* at 'gøre det samme i den samme situation'⁸. At man altså forstår, hvad 'samme' vil sige her. Og en sådan forståelse består væsentligt i kendskabet til konkrete eksempler og i selve dét at se og handle i forhold til disse *som* eksemplariske for praksissen. Forståelsen er, kunne man sige, i højere grad én, man udviser, end én, man udtrykker. Rationaliteten har med andre ord grund i den praksis, som bedømmerne er udøvere af.

Denne beskrivelse af evaluering som en praksis, der udøves af konkrete mennesker i forhold til konkrete

præstationer på baggrund af fagforståelse og erfaring med tidligere studenterpræstationer har ikke som hovedpointe, at evaluering er en uhjælpelig subjektiv affære. Erfarings- og personbundethed er ikke et usagligt træk – tværtimod er den en forudsætning for enhver evaluering, saglig eller ikke. Ej heller gør den *per se* bedømmelse arbitrær og idiosynkratisk: Bedømmerens evalueringskompetence er opøvet i faglig dialog med andre bedømmere om givne konkrete opgaver, og erfaringerne med, hvad man kan og bør forvente på det pågældende niveau, har derfor en grund, der i større eller mindre grad er socialt medieret. Imidlertid gør opfattelsen en række spørgsmål vigtige at stille, såsom hvilken betydning erfaring mere konkret har i evalueringspraksis, hvordan forståelse af et givet fag tilegnes, hvori karaktergivnings- og bedømmelsesekspertise består, hvordan den opøves, og hvilken rolle 'faglig socialisering' spiller i den forbindelse.

Ekspertise generelt og på evalueringsområdet

Et første oplagt sted at søge evalueringsekspertise belyst, givet et fokus på betydningen af erfaring, er Dreyfus' og Dreyfus' model for tilegnelse af færdigheder, såvel intellektuelle som kropslige, den såkaldte 'skill model'⁹. Ifølge denne forløber færdighedstilegnelse gennem en række trin¹⁰, fra novicens rigide følge af regler fastsat ud fra situationsuafhængige træk til ekspertens forholden sig til den specifikke situation med dens unikke kontekstafhængige aspekter – en forholden sig, der består i en erfaringsbaseret intuitiv begribelse af situationen og umiddelbar adækvat handlen i relation hertil. Væsentligt er her, at denne adækvate handlen ikke uden tab af netop det, der gør den adækvat – dens situationsspecificitet – kan beskrives i regler: Regler udsiger, hvordan man skal handle på samme måde i en række tilfælde, der alle er af samme 'type'. Ekspertens ekspertise, derimod, består netop i anerkendelsen af hver enkelt situation i dens specifikke særegenhed og i evnen til umiddelbart at handle herpå. Den involverer således slet ikke regelfølge, heller ikke på det ubevidste plan, men er i stedet en begribelse af og ageren i forhold til situationens holistiske gestalt, genkendt på baggrund af dennes lighed med gestalten i andre paradigmatiske situationer, som eksperten tidligere har været i¹¹.

Dreyfus & Dreyfus underbygger deres opfattelse med analysen af en række ret forskelligartede eksempler, hvoraf de primære angår at spille skak, at køre bil, at stille en medicinsk diagnose og at flyve et krigsfly. P. Benner har anvendt Dreyfusbrødrenes model i beskrivelsen af niveauer i sygeplejerskens praksisviden¹², og ekspertens 'ordløshed' og mangel på regelfølge er yderligere dokumenteret af forskellige praksisstudier¹³. Karakteristisk for disse forskellige studier er en understregning af, at eksperten ofte kommer i situationer, hvor han/hun ikke udtømmende kan redegøre sprogligt for grundlaget for sine vurderinger, men simpelthen

'kan se', hvordan det forholder sig, og hvilke handlinger der fordres – og hvor det videre hændelsesforløb vel at mærke underbygger den intuitive vurderings adækvathed.

Vender man tilbage til Keldorffs eksempel, synes det ikke urimeligt at beskrive voringen som en udøvelse af ekspertise i dreyfusiansk forstand: Med Dreyfusbrødrene kunne man sige, at bedømmerne intuitivt genkender 'gestalten' i studenterpræstationen, dvs. uden at følge regler umiddelbart handler, som 'det er passende', og vurderer præstationen til at ligge på et 10-tal. Censors udspil af en karakter to niveauer højere end det, der forinden havde været tale om, uden fremførelse af konkrete punkter til grund herfor, er således ifølge denne opfattelse en umiddelbar reaktion på præstationens 'gestalt'. Og den korte diskussion af muligheden af en karakterdifferentiering mellem de studerende kan ses som det, Dreyfusbrødrene kalder 'deliberative rationality': En afprøvning af intuitionen som helhed, ikke mod kriterier eller regler, men mod muligheden af andre synsvinkler, der kunne give en anden helhedsmæssig intuition¹⁴. Helt i overensstemmelse med denne fortolkning siger censor netop, i en let omskrivning af det tidligere citerede, at 'differentieringssynsvinklen for ham indebærer den fornemmelse, at de, hvis der skulle differentieres, skulle »gå endnu tættere på«'.

Tilsvarende er det fænomenologisk velkendt, at eksplicitte overvejelser over, hvorledes en studenterpræstation bør vægtes i forhold til studieordningsdefinerede krav er betydeligt hyppigere, hvis man som bedømmer ikke helt 'er inde i' eller 'har føling med' feltet, enten fordi man er novice som bedømmer, eller fordi man som censor er kaldt til at vurdere præstationer på en studieretning eller et studieområde, der er perifert i forhold til ens eget hovedfelt. For den erfarne bedømmer er henvisninger til studieordningen sjældne, forudsat at dét, der skal bedømmes, ligger centralt inden for vedkommendes undervisnings- og evalueringsområde. Fænomenologisk synes regler således at have en langt mere eksplicit handlingsstyrende rolle for novicen end for eksperten, hvilket er i klar overensstemmelse med den dreyfusianske opfattelse af færdighedstilegnelse som en udvikling fra strikt regelfølge til intuitiv, ikke-regelbaseret ekspertise.

Imidlertid er det spørgsmålet om en sådan ekspertise, i fald den eksisterer, virkelig er en adækvat form for kompetence på evalueringsområdet. Det synes problematisk på den ene side at have studieordningsdefinerede mål og kriterier, hvis grad af opfyldelse en given karakter hævdes at afspejle, såfremt disse mål og kriterier på den anden side i praksis ikke spiller nogen styrende rolle overhovedet for eksperten. Væsentligt er det også, at novicen på grund af sin manglende 'føling med' området reelt kun til dels kan hente handlingsvejledning fra de studieordningsdefinerede krav: Som tidligere diskuteret er novicens problem netop, at han/hun mangler den erfaringsbaserede forståelse

af, *hvad* reglerne udsiger, og *hvordan* man følger dem – en forståelse, som ekspertten til gengæld har. Den adækvate form for ekspertise synes derfor i stedet at måtte involvere en erfaringsrigdom knyttet til, ikke uafhængigt af, målbeskrivelser og evalueringskriterier. Der synes fordret en 'følning' med disses betydning, frem for en blot udvendig forståelse af de ord, der indgår i deres beskrivelse.

En anden mulighed er at forstå ekspertise som en 'viden i praksis' bestående af en enhed af de gensidigt bestemmende aspekter: sprogligt artikulerbar viden (målene og deres grader af realisation), praktisk kunnen (evnen til at skelne styrker og svagheder ved en studenterpræstation) og personlige erfaringer. I en anerkendelse af Dreyfus' pointe om, at ekspertten handler intuitivt, i en begribelse af situationens mening, må denne 'viden i praksis' forstås som et handlingsorienteret perspektiv, der lader forhold i situationen træde frem med den betydning og handlingsfordring, de har. Men i modsætning til Dreyfus skal intuitionen ikke ses som helt uden relation til sprogligt medieret viden. I stedet er sprogligt artikulerbar viden en del af perspektivets baggrund. Den indgår således som et integreret led i rammesættelsen af, *hvad* der kan træde frem i situationen og *på hvilken måde*, det kan gøre det. Samtidig er de personlige erfaringer og den praktiske kunnen med til at give mening til den sprogligt artikulerbare viden – for evaluerings vedkommende konkret ved at give en eksemplarisk forståelse af mål og målopfyldelse. De ikke-sproglige aspekter af ekspertisen er en tavs resonansbund og meningsramme for det sproglige aspekt, der på sin side virker som fokuserende faktor.

Når bedømmere diskuterer i konkrete voterings-situationer, resonerer deres personlige erfaringer og praktiske kunnen således med i deres forståelse af de ord, der bliver sagt. Derved får disse en anden og større meningsfylde, end de har for den, der ikke er evalueringsmæssigt kompetent. For denne kan de synes vage eller intetsigende, men dette skyldes i sådanne tilfælde netop, at de sætter meningsaspekter i spil, som han/hun ikke har tilsvarende at trække på. I en 'viden i praksis'-tolkning af Keldorffs case gør noget sådant sig gældende for de anvendte termer som 'svagere papirer' og 'hverdagsbevidsthed': Disse synes umiddelbart upræcise, men for bedømmerne 'fyldes' de meget muligt med en betydningsrigdom knyttet til deres fag- og præstationsforståelses tavse dimensioner.

'Viden i praksis' er en tilgrundliggende forståelse, der går dybere end hvert af sine aspekter, sammenbinder dem og lader dem gensidigt give hinanden meningsmæssigt indhold. På den ene side bevirker dette, at konkrete forhold kan ses som eksemplificeringer af regler – her, at en given præstation umiddelbart begribes som eksemplificerende en given grad af realisering af et mål. Og på den anden side giver den en 'fingerspidsfølelse' for regulering af reglerne med baggrund i eksemplerne. Den giver med andre ord en

'følning med', *hvilke* eksempler der fordrer ændringer af eller undtagelser fra reglerne, enten fordi det enkelte tilfælde er unikt anderledes, eller fordi det illustrerer nye muligheder for praksis – muligheder, der ikke er indeholdt i reglerne. Konkret betyder dette i forhold til bedømmelse af præstationer, at 'viden i praksis' giver følning med, hvornår og på hvilken måde studieordningsmål må tilpasses for at være adækvate i forhold til studenterpopulationen, til de muligheder, der er til stede i deres præstationer, og til det omgivende samfunds krav¹⁵.

Kan man være intuitiv ekspert i at give karakterer?

Et vigtigt spørgsmål implicit i det ovenstående er, hvilken relation der består mellem evaluering og karaktergivning. Umiddelbart synes det, som om evaluering måtte være meget mere end blot fastsættelsen af en karakter. En sådan siger ganske vist på et meget overordnet plan, hvorvidt en given præstation i høj grad, i rimelig grad, i mindre grad eller slet ikke lever op til fastsatte krav, men den siger intet om den mere præcise måde, hvorpå dette sker. Ej heller udtrykker den, om præstationen har det samme, karakterfastsatte, niveau på alle områder, eller tværtimod er meget bedre i nogle henseender end i andre, således at det pågældende niveau er en form for gennemsnitsvurdering af præstationen. Karakteren siger med andre ord meget lidt nuanceret og kvalitativt om præstationen. Lidt groft formuleret, så er karaktergivning en sammentrængning og ensretning af en mangefacetteret virkelighed¹⁶ med henblik på at sortere denne i et begrænset antal kvantificerende kasser. Evaluering generelt, vil man sige, munder måske nok ud i en sådan kvantificering, men består jo grundlæggende i en omfattende stillingtagen til samtlige kvaliteter ved en given præstation og er derfor ikke en ensretning af virkelighedens mangfoldighed. Hvilket man i mange tilfælde også har mulighed for at give udtryk for gennem tilknytning af en skriftlig eller mundtlig kommentar.

Dette til trods er det en kendsgerning, at hverken eksaminandens familie, venner, studiekammerater eller øvrige undervisere sædvanligvis stiller sig tilfreds med en oprensning af eksaminators kvalitative udsagn som svar på spørgsmålet »Hvordan gik det så?« Forsøger eksaminanden at svare på denne måde, kommer der normalt et opfølgende spørgsmål i stil med »Aha... OK... og hvad fik du så?« Et hyppigt ønske fra studerende i forbindelse med eksamener, der bedømmes bestået/ikkebestået, er tilsvarende at få oplyst, hvilken karakter de ville have fået, hvis bedømmelsen var sket efter 13-skalaen. Den begrundelse, der fremsættes for dette ønske, er sædvanligvis, at vedkommende gerne vil vide, 'hvor god opgaven faktisk var'. Og ser man på Keldorffs case, angår voringen her udelukkende karakterniveauet, ikke kvalitative forhold ved præstationen.

Om end de fleste formentlig principielt vil tilslutte sig, at nogle få tal ikke kan repræsentere en mangesidet virkelighed på fyldestgørende vis, ageres der altså jævnligt i praksis, som om den kvantificerende ensretning faktisk siger *mere* om den grad, i hvilken en præstation lever op til fastsatte mål, end den sproglige formulering af kvaliteter og mangler i forhold til disse. Som om ord er upræcise og utroværdige, mens tal er pålidelige og rammer plet. Og reelt giver eksaminanden og andre den mest præcise forståelse af, hvor god eller dårlig præstationen har været. De sproglige formuleringer synes i praksis blot at blive opfattet som et middel til evaluering, frem for som evalueringen selv¹⁷.

Givet denne tendens kunne man, lidt sat på spidsen, spørge, om eksperten på evalueringsområdet egentlig overhovedet har brug for sprogligt artikulerede begrundelser for sin karakterfastsættelse. Om vedkommende ikke præcis kunne være en dreyfusiansk ekspert, der intuitivt opfatter præstationen i sin helhed som eksemplificerende dette eller hint karakterniveau. Dette, kunne man sige, synes netop at være, hvad bedømmerne i Keldorffs case gør: Karakterskalaen er, som tidligere bemærket, et inkorporeret værktøj, som de opfatter studenterpræstationen gennem. Den sproglige artikulation er, kunne man hævde, måske nok nødvendig over for andre, der *ikke* er evalueringsmæssige eksperter på feltet, men udgør for eksperten selv en begrænsning af hans/hendes muligheder for fleksibel evaluering. Præcis, med H. Dreyfus' billede, som støtthjul reelt virker hæmmende på den erfarne cyklists handlinger.

Alternativt kunne man spørge, hvorvidt det overhovedet er rimeligt, etisk og fagligt set, at være intuitiv ekspert i at ensrette og sammentrænge en mangfoldig virkelighed i kvantificerende kasser? Om det ikke er et alvorligt problem for evalueringens lødighed, hvis den tendens, der er hos eksaminander og undervisere til at agere, som var karaktergivning evalueringens essens frem for dens grovmaskede sammenfatning, forleder til den opfattelse, at en karakter faktisk i tilstrækkelig grad indfanger en præstation kvalitativt. Bør en erfaringsbaseret evne til umiddelbart at se studenterpræstationer som eksemplificerende bestemte tal ikke betragtes med stor skepsis og som et tegn på, at der er noget galt i den praksis, erfaringerne er gjort i?

Den skitserede intuitive 'føling med', hvad en given karakter betyder, og hvordan præstationer derfor skal evalueres, er problematisk af to grunde: For det første er intuitionen alt for snæver og overfladisk i sin relation til en bestemt kvantificering af feltet. Den angår således slet ikke dette i dets kvalitative kompleksitet og mangfoldige realisationsmuligheder. For det andet, i forlængelse heraf, men et niveau dybere, tillader den pga. dels sin overfladiskhed, dels sin entydige retning fra karakterforståelse til præstation, ikke en stadig gensidig bestemmelse og udvikling af mål og konkrete vurderinger i forhold til hinanden. Den muliggør med andre

ord ikke en fortløbende nuancering af bedømmerens tilgrundliggende forståelse af feltet.

Problemet opstår, når karakterskalaen går fra at være ét redskab, som bedømmeren benytter sig af i en standardiseret rapportering af sine gestaltevalueringer, til at være dét inkorporerede redskab, som han/hun opfatter studenterpræstationer på baggrund af: Redskabet er ganske enkelt ikke nuanceret nok til at have denne rolle som perceptuelt udgangspunkt. Et sådant 'perceptuelt udgangspunkt' kan rettelig kun udgøres af 'viden i praksis' med dets interrelaterede fagspecifikke helhed af mål, vurderinger og eksemplificeringer. Karakterskalaen kan være et redskab *for* bedømmeren *i* udøvelsen af 'viden i praksis', men sidstnævnte er en grundlæggende kvalitativ fag- og præstationsforståelse, som ikke kan erstattes af 'intuitiv føling' med en kvantitativ bedømmelsesform. Det ville ikke være adækvat, hverken fagligt eller evalueringsmæssigt, og er derfor heller ikke etisk forsvarligt.

Fagligt socialiserede evalueringspraksisser

Med den skitserede opfattelse af evalueringskompetence som en erfaringsbaseret forståelse af faglige mål og paradigmatisk eksemplificeringer melder sig imidlertid et yderligere spørgsmål, nemlig hvilken rolle 'socialisering' spiller for etableringen *af* denne forståelse. Her er det som det første punkt klart, at bedømmerens fagforståelse i sin grundlæggende form er dannet igennem hans/hendes egen studietid, i arbejdet med fagets tekster og opgaver, praktiske som teoretiske, i diskussioner med medstuderende og undervisere og i de kvalitative evalueringskommentarer, vedkommende selv har fået i den forbindelse. Væsentligt er ikke kun det, der eksplicit positivt blev sagt om faget, dets tematikker, grundspørgsmål og grundarbejdsformer. Mindst lige så vigtigt er det negativt sagte, det usagte, det ignorerede, de afviste spørgsmål, indfaldsvinkler, emner og teoriretninger; kort sagt, den måde faget *i praksis* blev afgrænset for vedkommende på¹⁸.

Fagforståelse etableres i bund og grund i en socialt medieret forholden sig til det faglige felt, hvor både afgrænsningen af feltet, kendskabet til dets metoder og problematikker og forståelsen for, hvad der udgør mulige former for svar, opøves gennem arbejdet med faget i interaktion med andre praksisdeltagere, såvel mere kompetente som jævnbyrdige. Dette gør imidlertid ikke fagforståelsen til en fuldstændigt socialt konstitueret, relativistisk størrelse: For det første er den socialt medierede forholden sig en forholden sig til *noget*, dvs. en tilnærmen til et fagligt felt, der måske nok fortolkes på mange måder, men som ikke kan antage en hvilken som helst vilkårlig form. Og for det andet indebærer social mediering ikke individuel determination: Den studerende overtager ikke bare en socialt konstitueret forståelse, men fremstrukturerer sin egen mening på baggrund af arbejdet med stoffet i interaktion med andre.

Denne fagforståelse etableret igennem studietiden er den baggrund, som evalueringskompetence tager udgangspunkt i. Imidlertid får fagforståelsen et andet fokus i forbindelse med evaluering, idet den her skal kobles til vurderingen af realistiske faglige forventninger til forskellige uddannelsesniveauer. Bedømmeren skal således transformere den mere eller mindre absolutte kvalitative faglige vurdering, han/hun har af forskning inden for feltet, til en forståelse af mål, der nok er absolutte, men kun relativt til et givet uddannelsesstrin. Ligesom den oprindelige tilegnelse af fagforståelsen er denne transformationsproces socialt medieret, men ikke nødvendigvis socialt determineret. Den finder sted igennem dialog med andre bedømmere om konkrete studenterpræstationer og forholder sig således til et specifikt materiale, der klart lader sig anskue fra forskellige vinkler, men hvis kvaliteter eller mangel på samme omvendt ikke kan konstrueres arbitrært.

Væsentligt er her dog igen udviklingen af evalueringskompetencen som en 'viden i praksis', dvs. som en enhed af den sproglige artikulation af mål og den tavse meningsmæssige erfaringsresonansbund for disse mål; en enhed, der er realiseret som bedømmerens eget, personbundne perspektiv: Uden en sådan personlig engageret sætten mål og erfaringer i spil med hinanden i bedømmelsesprocessen vil bedømmelsesparternes faglige dialog ikke kun udgøre et fokuspunkt for bedømmerens udvikling af sin egen erfaringsbaserede forståelse af studenterpræstationer. Dialogen vil i stedet være fikspunkt herfor. I sådanne tilfælde vil den sociale mediering kunne degenerere til social determination.

Også når dette ikke sker, indebærer en udelukkende praksisintern socialt medieret forholden sig til studenterrepræsentationer dog en risiko for udviklingen af en fagidiosynkretisk evalueringspraksis, i hvert fald på visse punkter: Om end vurderingen af, *hvad* der udgør kvaliteter ved en studenterpræstation i mindst lige så høj grad er fagligt, indholdsmæssigt bestemt som socialt konstitueret, kan det *niveau*, som disse kvaliteter bedømmes at ligge på, have en vis fagidiosynkretisk forskydning, afhængigt af om den faglige praksis har en tradition for at sætte sine mål meget højt eller meget lavt, relativt til et givet uddannelsesniveau. Ligeledes kan der være praksisspecifikke tendenser til kun at bruge udsnit af karakterskalaen¹⁹. Der er derfor god grund til at fremme tiltag, der på adækvat fagcentret vis sammenligner evalueringspraksisser på tværs af fag.

Afrundende bemærkninger

Med Keldorffs case er der påpeget en risiko for, at karakterskalaen for bedømmerne kan blive inkorporeret og danne det perceptuelle udgangspunkt for bedømmelse. Dette forhold taler for, at en karakterskala fra tid til anden bør udskiftes, for gennem fremmedgørelsen af

det kvantitative bedømmelsesredskab at modvirke en sådan inkorporeringstendens. Ligeledes viser risikoen mere generelt hen til nødvendigheden af en stadig refleksion over og diskussion af de kriterier, der lægges til grund for bedømmelse. Også faren for udvikling af fagidiosynkretiske evalueringspraksisser peger i denne retning.

Imidlertid kan en sådan refleksion ikke primært være eksternt funderet i f.eks. konsulentbistand, fordi en nødvendig forudsætning for og et essentielt led i refleksionen er den tilgrundliggende forståelse af det pågældende faglige felt. I stedet må den tage sit udgangspunkt i praksis. Mere konkret vil refleksionen naturligt forholde sig til og forsøge at italesætte relationen mellem de forskellige aspekter af evalueringens 'viden i praksis', nærmere betegnet relationen mellem på den ene side de sprogligt artikulerbare mål og på den anden side den tavse resonansbund og meningsramme, som praktisk kunnen og personlige erfaringer udgør²⁰. Herigennem muliggøres nemlig en større fagligt begrundet, sprogligt artikuleret fokusering af den meningsfylde, som erfaringernes tavse resonansbund bidrager med. Dette på sin side nuancerer forståelsen af de kriterier, der er lagt til grund ved konkrete evalueringsopgaver og bidrager dermed til at kvalificere vurderingen af disses adækvathed.

Denne forankring af refleksionen i 'viden i praksis' og i videre forstand i evalueringspraksissen udelukker imidlertid ikke, at den kan være initieret af en ekstern påvirkning, ligesom det ikke forhindrer, at den praksisforankrede refleksion kan hente inspiration fra og være i dialog med andre evalueringspraksisser på tværs af fag. Tværtimod synes dette sidste nødvendigt i afværgelsen af fagidiosynkrasi på evalueringsområdet og konkret f.eks. i sikringen af, at forskelle i karakterfordelinger på tværs af fag er fagligt og/eller målgruppemæssigt velbegrundet. En nødvendig forudsætning for, at en sådan tværgående dialog bliver vellykket er dog, at den formår hos hver enkelt deltager at sætte dennes tavse resonansrum af erfaringer i spil, således at dialogen bevæger sig fra den relativt overfladiske accept eller forkastelse af de samme ord til en reel fælles erfaringsrelateret forståelse af disse. Meget muligt kan dette kun gøres til fulde mellem beslægtede fag, da det erfarings- og eksempelgrundlag, der er at trække på i diskussionen af, f.eks. hvori 'selvstændighed' består, kan tænkes at være for forskelligartet for fag, der har få berøringspunkter. Der vil dog utvivlsomt være tale om et kontinuum her, således at evalueringsdebatten med større eller mindre udbytte kan føres mellem alle fag.

Set i lyset af disse overvejelser kan indførelsen af en ny karakterskala til erstatning af 13-skalaen dels hæmme en eventuel inkorporeringstendens, dels, som indledningsvis foreslået, udgøre en konkret anledning til igangsættelse af den nødvendige refleksion over evalueringskompetence også i praksisser, hvor en sådan

refleksion hidtil har manglet. Det aktuelt bedste argument for at skifte skala er således meget muligt ikke de *specifikke* problemer, som af Karakterkommissionen er blevet påpeget i brugen af den eksisterende, men den åbning, en udskiftning giver, i forhold til at diskutere, hvad der i evalueringspraksis har *ledt til* disse problemer. Og mere specifikt, om de skyldes, at evaluering i mange praksisser sker gennem udøvelse af en grad af intuitiv karaktergivningsekspertise med en fagidiosynkratisk brug af karakterskalaen, begge dele opstået som følge af mangelfuld refleksion over evalueringskriterier²¹.

Litteratur

- Benner (1984): *From Novice to Expert: Excellence and Power in Clinical Nursing Practice*, Addison-Wesley, Massachusetts.
- Caroll, Lewis (1895): »What the Tortoise Said to Achilles«, *Mind*, nr. 4, s. 278-280.
- Dohn, Nina Bonderup (2000): »Viden i praksis«, *Tidsskrift for Socialpædagogik*, nr. 6.
- Dohn, Nina Bonderup (1999): *Hen imod viden som viden i praksis*, Network for Non-Scholastic Learning, Working Papers, nr. 14, Aarhus Universitet, Århus.
- Dohn, Nina Bonderup (2005): *Læring i praksis – fremstruktureringen af et handlingsorienteret perspektiv*, Aalborg Universitet.
- Dreyfus, Hubert & Dreyfus, Stuart (1986a): »From Socrates to Expert Systems: The Limits of Calculative Rationality« I Carl Mitcham & Alois Hünig, (red.): *Philosophy and Technology II*, D. Reidel Publishing Company, Dordrecht, s. 111-130.
- Dreyfus, Hubert & Dreyfus, Stuart (1986b): *Mind over Machine*, The Free Press, New York.
- Dreyfus, Hubert (2001): »Phenomonological Description versus Rational Reconstruction« *Revue internationale de philosophie*, vol. 55, nr. 216, s. 181-196.
- Dreyfus, Hubert (2001): *On the Internet*, Routledge, London.
- Dreyfus, Hubert (1992): *What Computers Still Can't Do*, Harper & Row, New York.
- Göranzon, Bo & Josefson, Ingela (1988) (red.): *Knowledge, Skill and Artificial Intelligence*, Springer-Verlag, England.
- Johannessen, Kjell & Rolf, Bertil (1989): *Om tyst kunnskap – två artiklar*, Uppsala Universitet, Uppsala.
- Josefson, Ingela (1998): *Läkarens yrkeskunnande*, Studentlitteratur, Lund.
- Kanigel, Robert (1986): *Apprentice to Genius. The Making of a Scientific Dynasty*, John Hopkins Univ. Press, Maryland.
- Karakterkommissionen (2004): *Betænkning om indførelse af en ny karakterskala til erstatning for 13-skalaen*, Undervisningsministeriets forlag, København.
- Keldorff, Søren (1996): *Tæt på en gruppe*, Edupax, Aalborg.
- Kuhn, Thomas (1970): *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.
- Kvale, Steinar (1999): »Forskere i lære« Klaus Nielsen & Steinar Kvale, (red.): *Mesterlære – Læring som Social Praksis*, Hans Reitzels Forlag, København.
- Latour, Bruno (1987): *Science in Action*, Harvard Univ. Press, Cambridge, Massachusetts.
- Lave, Jean & Wenger, Etienne (1991): *Situated Learning – Legitimate Peripheral Participation*, Cambridge University Press, New York.
- Merleau-Ponty, Maurice (1962): *The Phenomenology of Perception*, Routledge and Kegan, London.
- Molander, Bengt (1996): *Kunnskap i handling*, Daidalos, Göteborg.
- Polanyi, Michael (1966): *Personal Knowledge – Towards a Post-Critical Philosophy*, Harper & Row, New York.
- Ryle, Gilbert (1990): *The Concept of Mind*, Penguin Books, London.
- Schön, Donald (1983): *The Reflective Practitioner*, Basic Books, USA.

- Wenger, Etienne (1998): *Communities of Practice*, Cambridge University Press, New York.
- Wittgenstein, Ludwig (1984a): *Philosophische Untersuchungen*, Suhrkamp, Frankfurt a.M.
- Wittgenstein, Ludwig (1984b): *Über Gewißheit*, Werkausgabe 8, Suhrkamp, Frankfurt a.M.
- Zuckerman, Harriet (1996): *Scientific Elite: Nobel Laureates in the United States*, Transaction Publishers, New Brunswick, New Jersey.

Noter

- 1 *Betænkning om indførelse af en ny karakterskala til erstatning for 13-skalaen.*
- 2 Sådanne kritiske bemærkninger er fremsat f.eks. i foræet »Ny karakterskala« på www.folkeskolen.dk; på »Seminar om ny karakterskala«, Københavns Universitet, 15. marts 2005, og i artiklerne »Dumpekarakter til ny skala« *Ny Viden*, nr. 9, oktober 2005; »Hård kritik af ny karakterskala«, *Jyllands-Posten*, 26. oktober 2005; »Ny karakterskala er »skudt i hovedet« i *Politiken*, 27. oktober, 2005; »Dumpekarakter til ny karakterskala«, *DR Nyheder*, 26. oktober 2005 samt tilhørende debatforum på <http://www.dr.dk/Nyheder/Politik/2005/10/26/094641.htm?rss=true>.
- 3 Jf. *Betænkning om indførelse af en ny karakterskala til erstatning for 13-skalaen*, s. 31ff.
- 4 Keldorff: *Tæt på en gruppe*, 1996, s. 24ff.
- 5 Merleau-Ponty: *The Phenomenology of Perception*, 1962. Jf. specielt s. 142ff. og s. 174ff.
- 6 Jf. Wittgenstein: *Über Gewissheit*, 1984b, § 139.
- 7 Jf. Ryle: *The Concept of Mind*, 1949, s. 31. Jf. også Carolls fabel i »What the Tortoise Said to Achilles«, hvis pointe er, at logik ikke i sig selv er tvingende, men at det er *forståelsen* af logik, der er det.
- 8 Wittgenstein: *Philosophische Untersuchungen*, 1984a; Wittgenstein: *Über Gewissheit*, 1984b. Jf. også Dohn: *Læring i praksis – fremstruktureringen af et handlingsorienteret perspektiv*, 2005; Dohn: *Hen imod viden som viden i praksis*, 1999; Johannessen & Rolf: *Om tyst kunnskap – två artiklar*, 1989; Molander: *Kunnskap i handling*, 1996.
- 9 Dreyfus & Dreyfus: »From Socrates to Expert Systems: The Limits and Dangers of Calculative Rationality«, 1986a; Dreyfus & Dreyfus: *Mind over Machine*, 1986b; Dreyfus: *On the Internet*, 2001.
- 10 Nærmere bestemt 5 trin. H. Dreyfus har i de senere år udvidet modellen med yderligere 2 trin, mester og verdensoplukker (jf. Dreyfus, 2001). Disse er dog af mindre betydning i denne sammenhæng.
- 11 Man moder af og til den misforståelse af Dreyfusbrødrenes synspunkt, at selvom eksperten ikke eksplicit benytter sig af regler i sin handlen, følger han/hun stadig regler, bare på en 'ubevidst' eller 'intuitiv' måde. Denne misforståelse er ifølge H. Dreyfus en 'rationel rekonstruktion' af ekspertens handlen – en rekonstruktion, som der ikke er belæg for i den fænomenologisk beskrivelse af denne. Han afviser den med følgende udsagn: »[This] is like supposing that, since I once needed training wheels to ride a bicycle, now that I seem to be able to ride without them, I must still be using them even though they can no longer be perceived« (Dreyfus: »Phenomonological Description versus Rational Reconstruction«, 2001, s. 182). I stedet siger Dreyfusbrødrene, »Given our account of the five stages of skill acquisition, we can understand why the knowledge engineers from Socrates, to Samuel, to Feigenbaum have had such trouble getting the expert to articulate the rules he is using. The expert is simply not following any rules! He is doing just what Socrates and Feigenbaum feared he might be doing – recognizing thousands of special cases.« (Dreyfus & Dreyfus, 1986a, s. 121). Jf. også Dreyfus: *What Computers Still Can't Do*, 1992, specielt kapitel 8.

- 12 Benner: *From Novice to Expert: Excellence and Power in Clinical Nursing Practice*, 1984.
- 13 Feks. Göranson (red.): *Knowledge, Skill and Artificial Intelligence*, 1988; Josefson: *Läkarens yrkeskunnande*, 1998; Schön: *The Reflective Practitioner*, 1983.
- 14 Dreyfus & Dreyfus, 1986, s. 36ff.
- 15 Begrebet 'viden i praksis' har jeg behandlet betydeligt mere udførligt i Dohn, 2005; Dohn: »Viden i praksis«, 2000; Dohn, 1999.
- 16 Mangefacetteringen er naturligt nok betydeligt mere udtalt ved større skriftlige opgaver end ved løsning af relativt specifikke opgaver ved eksempelvis en 4-timers skriftlig eksamen.
- 17 Tendensen blandt studerende og undervisere til at fokusere på kvantitativ bedømmelse understøttes naturligvis af forventningen om, at en persons fremtidige karrieremuligheder afhænger af vedkommendes karakterer. Denne eksterne motivationsfaktor må dog forventes at være af relativt mindre betydning for følelsen af, at kun den kvantitative bedømmelse reelt indeholder information om, 'hvor god' en præstation har været.
- 18 Jf. f.eks. Kuhn: *The Structure of Scientific Revolutions*, 1970; Kani-gel: *Apprentice to Genius. The Making of a Scientific Dynasty* 1986; Latour, Bruno: *Science in Action*, 1987; Kvale: »Forskere i lære«, 1999; Zuckerman: *Scientific Elite: Nobel Laureates in the United States*, 1996.
- 19 Undersøgelser foretaget af Karakterkommissionen tyder på, at dette er tilfældet Jf. *Betænkning om indførelse af en ny karakter-skala til erstatning for 13-skalaen*, s. 31ff samt bilag 3. Censor-institutionens krav om, at 1/3 af censorerne skal komme fra aftagerpraksisser, burde måske principielt kunne modvirke de nævnte fagidiosynkratiske tendenser. I hvilken udstrækning dette reelt vil være tilfældet afhænger dog af, om aftagercensorerne via deres praksis har andre instanser (eksempelvis andre fag) at vurdere evalueringniveauet imod. Er dette ikke tilfældet, er deres mulighed for at opdage og modvirke sådanne tendenser i realiteten relativt begrænsede, og sandsynligheden for, at de udvikler en fagidiosynkratisk evalueringskompetence svarende til universitetsbedømmernes, forholdsvis høj.
- 20 Refleksionen vil selvfølgelig også skulle forholde sig til relationen mellem disse mål, deres opfyldelse og de krav, det omgivende samfund stiller til studiets færdige kandidater.
- 21 En stor tak til to anonyme reviewers for mange værdifulde kommentarer til en tidligere version af artiklen.