

Om forholdet
mellem evaluering med
formative og
summative formål



Wynne Harlen

Visiting Professor at the Graduate School of Education of the University of Bristol

Artiklen er en oversættelse af kapitlet "On the Relationship between Assessment for Formative and Summative Purposes" fra bogen *Assessment and Learning: An introduction*, Edited by John Gardner 2006, SAGE Publications



Denne artikel diskuterer forholdet mellem evaluering med to af dets hovedformål – at støtte læring og at opsummere det lærte. Begge formål er centrale i en effektiv uddannelsespraksis, men der er aspekter, der drejer sig om, hvorvidt det er brugbart at anse dem for at være begrebsmæssigt og pragmatisk adskilte. Som udgangspunkt for denne diskussion vil første del af kapitlet fremstille denne skelnen ved hjælp af modeller for evaluering til læring og evaluering af læring. Anden og tredje del behandler spørgsmål om, hvorvidt dokumentation, der er indsamlet til det ene formål, kan bruges til det andet. Den fjerde del overvejer, om der er en skala af formål frem for et modsætningspar, og om denne distinktion mellem formativ og summativ evaluering er brugbar eller redundant. I en tid hvor praksis først lige er ved at indhente teori i relation til formativ evaluering, er det uundgåeligt, at dette kapitel vil stille flere spørgsmål, end det besvarer. Slutteligt vil det konkludere, at distinktionen mellem formative og summative formål med evaluering for nuværende bør opretholdes, mens evalueringssystemer bør inkludere visse forholdsregler, der gør det muligt at bruge dokumentation, der er indsamlet af lærere, til begge formål. Det skal understreges, at begreberne 'formativ evaluering' og 'summativ evaluering' og 'evaluering til læring' og 'evaluering af læring' bruges synonymt i dette kapitel.

Hvordan evaluering tjener formative og summative formål

Der er generel enighed om, at evaluering i uddannelsesmæssig kontekst indebærer at beslutte, indsamle og bedømme dokumentation, der relaterer til de læringsmål, der evalueres. Dette refererer ikke til brugen af dokumentationen, hvem der bruger det og hvordan. Dette er kernen i forskellen mellem formativ og summativ evaluering. Det, vi nu kalder summativ evaluering (evaluering af læring), har været en del af uddannelse gennem århundreder. Nogen vil indvende, at formativ evaluering (evaluering til læring) går lige så langt tilbage fra Rousseaus opfordring i 1762 til at "Begynde med at undersøge de lærende nøjere, da det er klart, at I ikke ved noget om dem" (Rousseau 1961, Forord: 1). Det at beskrive formativ evaluering som et særskilt formål med evaluering, der kræver en præcis definition, er dog af nyere dato.

Selv om det ikke er nyt, blev det at definere evaluering i forhold til formål udbredt, da Task Group on As-

essment and Testing (TGAT) udgav en rapport i 1988, der skelnede mellem fire formål med evaluering: formativt, diagnosticerende, summativt og bedømmende (DES/WO, 1988a). Ordet 'formativ' blev brugt til at definere evaluering, der støtter læring ved at benytte dokumentation om, hvad eleverne har opnået i forhold til deres læringsmål, for at kunne planlægge de næste skridt i deres læring, og hvordan de skal tages. I realiteten indebærer formativ evaluering den diagnosticerende evaluering, som ofte forbindes med problemerne med læring, eftersom formativ evaluering indeholder både de problematiske og de positive præstationer. Summativ evaluering tilvejebringer, som begrebet antyder, en opsummering af præstationer på et bestemt tidspunkt. Dette er en nødvendig del ethvert evalueringssystem, da det tilvejebringer informationer til dem, der måtte have en interesse i elevers præstationer: hovedsagligt forældre, andre lærere, arbejdsgivere, senere og højere uddannelsesinstitutioner og eleverne selv. Evaluering får et bedømmende formål, når en gruppe elevers præstationer bruges til at rapportere om klassens, lærerens, skolens arbejde eller arbejdet i en anden del af uddannelsessystemet. De informationer om elever, der bruges til dette formål, er nødvendigvis baseret på summativ evaluering. Der er dog bred enighed om, at evaluering af en uddannelsesbestemmelse må basere sig på en hel del informationer, der ligger ud over dem, der kommer af evaluering af elever.

Hvad angår de enkelte elever, ligger den væsentlige distinktion mellem evaluering med formative og summative formål. Brugen af begreberne 'formativ evaluering' og 'summativ evaluering' kan dog give det indtryk, at

disse er forskellige former for evaluering, eller at de er forbundet med forskellige metoder til at indsamle dokumentation. Dette er ikke tilfældet; det væsentlige er, hvordan informationerne bliver brugt. Det er derfor, at udtrykkene 'evaluering til læring' og 'evaluering af læring' til tider foretrækkes. Den centrale forskel er, at evaluering til læring bruges til at træffe beslutninger, der påvirker undervisning og læring fremadrettet på kort sigt, hvorimod evaluering af læring bruges til at registrere og dokumentere, hvad der er lært i fortiden. Forskellen afspejles i figurerne 6.1 og 6.2.

Figur 6.1 fremstiller formativ evaluering som en cyklus af begivenheder. I løbet af en aktivitet A indsamles dokumentation, som fortolkes i forhold til fremskridt hen imod lektionens mål. Det er nødvendigt for fortolkningen med et eller andet begreb om progression i forhold til målet, således at elevernes nuværende niveau kan være med til at udpege de næste, passende skridt. Det at hjælpe eleverne med at tage de næste skridt, at nå til aktivitet B, er den måde, hvorpå dokumentationen af den nuværende læring ledes tilbage til undervisning og læring. Denne feedback bidrager til at regulere undervisningen, således at man kan justere hastigheden, hvormed man bevæger sig mod et læringsmål, for at sikre elevernes aktive deltagelse. Som det er tilfældet med alle regulerede processer, er feedback ind i systemet den vigtige mekanisme, der sikrer effektiv drift. Ligesom feedback fra en termostat i en radiator bevirker, at temperaturen i et rum opretholdes, så hjælper feedback af information om læring med at sikre, at nye erfaringer ikke bliver for svære eller for lette for eleverne.

Når det drejer sig om undervisning, skal både læreren og eleverne have feedback. Lærerne skal have feedback, for at de kan vurdere, hvilke skridt der skal tages fremover, og hvad der kan hjælpe eleverne med at tage disse skridt. Feedback til eleverne fremmer læringen mest effektivt, hvis de tages med på råd, når beslutningen om det næste skridt tages, sådan at de ikke bare bliver passive modtagere af lærerens bedømmelse af deres arbejde. Derfor er eleverne i centrum for processen, og pilene med to spidser i Figur 6.1 indikerer, at de spiller en rolle i indsamlingen, fortolkningen og brugen af dokumentationen af deres arbejde.

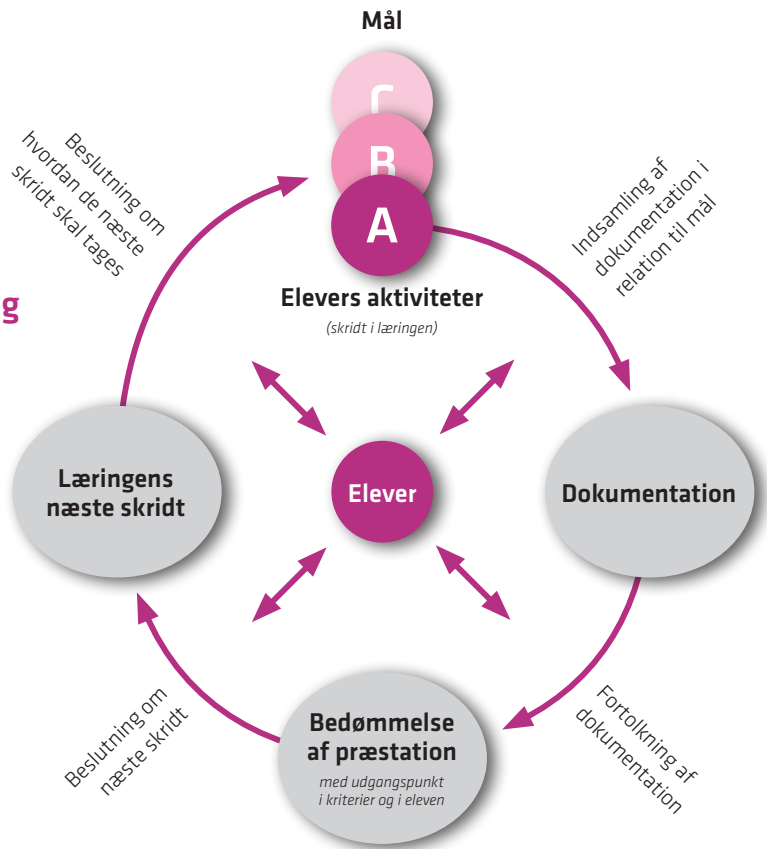


Figur 6.1

Evaluering til læring som en cyklus af begivenheder

(Tilpasset fra Harlen, 2000)

Ideen bag figuren er en cirkel, der viser en forbundet proces.



Handlingerne, der er beskrevet i boksene i Figur 6.1, er ikke 'faser' i en lektion eller bevidste beslutninger truffet af læreren. De repræsenterer en struktur, der kan bruges til at overveje, hvad det indebærer, når man fokuserer på, hvad og hvordan elever lærer, og som kan bruges til at støtte den videre læring. I nogle tilfælde kan lærere og elever sammen tage en beslutning om øjeblikkelig handling. Hvis en lærer fx ser, at en gruppe elevers hypoteser om en videnskabelig begivenhed, de undersøger, ikke stemmer overens med den videnskabelige forklaring, kan vedkommende hjælpe eleverne med at afprøve deres idéer, og de kan på den måde selv opdage, at der er behov for alternative forklaringer. I andre tilfælde kan læreren notere sig, hvad der er behov for, og først senere bidrage med det. Når man implementerer formativ evaluering, betyder det, at ikke alt i en lektion kan planlægges på forhånd. Hvis elevernes nuværende læring skal tages med i overvejelserne, vil nogle beslutninger per definition afhænge af, hvad denne læring er. Nogle idéer kan forudses på baggrund af lærerens erfaring, men ikke alle. Læreren har altså ikke brug for et foruddefi-

neret indhold i lektionen, men en række strategier, der kan tages i brug, når det er passende i den konkrete situation.

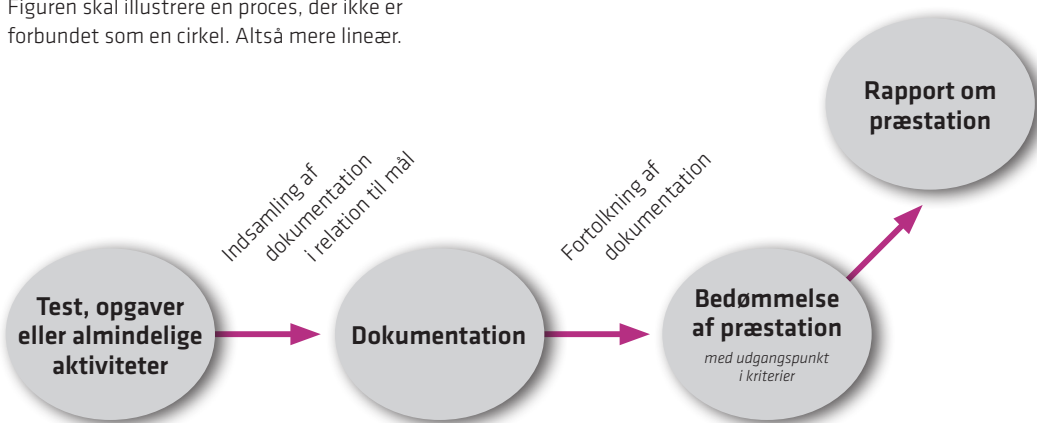
Når man sammenligner med Figur 6.1, viser Figur 6.2 den grundlæggende forskel mellem evaluering til læring og evaluering, der benyttes til at dokumentere præstationer. I Figur 6.2 kan dokumentation i relation til læringens mål indsamles gennem almindelige aktiviteter eller gennem særlige evalueringsopgaver eller test. Fortolkningen foregår ud fra opnåelsen af bestemte færdigheder, forståelser og holdninger gennem et antal aktiviteter. Denne vil referere til bestemte kriterier, der er ens for alle elever, eftersom formålet er at dokumentere præstationer, sådan at der kan sammenlignes eleverne imellem. Her er der ingen feedback til undervisningen - i hvert fald ikke omgående, som det er tilfældet i cyklussen, der viser reevaluering til læring. I denne model spiller eleverne ingen rolle i evalueringen, selv om flere argumenterer for, at de kan og burde have en rolle (se fx Frederiksen og White, 2004).

Evaluering af læring udført af lærere

(Tilpasset fra Harlen, 2000)

Figur 6.2

Figuren skal illustrere en proces, der ikke er forbundet som en cirkel. Altså mere lineær.



Disse modeller repræsenterer to begrebsmæssigt forskellige måder at bruge dokumentation på, hvilket giver mulighed for en klar skelnen mellem evaluering til læring og evaluering af læring. Men er denne skelnen helt så skarp i praksis? Kan dokumentation, der er indsamlet med det formål at opsummere læring, bruges til at støtte læring? Kan dokumentation indsamlet med formative formål for øje også bruges til summative formål? Hvis dette er tilfældet, hvordan kan det så gøres, og hvad siger det om en reel skelnen mellem formativ evaluering og summativ evaluering? Dette er de spørgsmål, der behandles i de næste to dele af dette kapitel, som trækker på materiale fra Harlen (2005).

At bruge dokumentation fra summativ evaluering til at støtte læring

Spørgsmålet er her, om dokumentation, der er indsamlet for effektivt og pålideligt at tjene summative formål, også kan bidrage til at støtte læring. Denne dokumentation kan fx være indsamlet ved hjælp af test eller særlige opgaver eller en opsamling af arbejder i et fag for at se, i hvilken udstrækning eleverne har opnået bestemte færdigheder eller forståelser. Kan denne dokumentation bruges til at støtte den videre læring? Der findes eksempler på praksis fremlagt af Maxwell (2004) og af Black m.fl. (2003), der understøtter et positivt svar på dette spørgsmål.

Maxwell beskriver den evaluering, der bruges på uddannelsestrinnet Senior Certificate i Queensland, hvor dokumentation samles over tid i en elevportfolio, som 'progressiv evaluering'. Han skriver at:

"Al progressiv evaluering indebærer nødvendigvis feedback til eleven angående kvaliteten af deres præstation. Dette kan udtrykkes i form af elevens fremskridt hen imod de ønskede læringsresultater og i form af forslag til udvikling og forbedring fremover..."

Hvis denne tilgang skal fungere, er det nødvendigt at udtrykke forventningerne til læringen i form af fælles aspekter ved læringen (kriterier). Derudfra kan man så diskutere, om eleven er på rette vej i forhold til forventningerne til læring, og hvad der skal gøres for at forbedre præstationen i fremtidige evalueringer, hvor de samme aspekter gør sig gældende.

I takt med at eleven opbygger en portfolio med dokumentation af hans eller hendes præstationer, kan tidligere evaluering erstattes af senere evaluering, der vedrører de samme underliggende aspekter ved læring. Målet er at dokumentere, 'hvad eleven har opnået' på sin læringsrejse, ikke hvor de begyndte, eller hvor de lå i forhold til gennemsnittet gennem hele faget." (Maxwell, 2004: 2-3)

Centralt for denne tilgang er definitionen af mål og evalueringskriterier som "fælles aspekter ved læring". Beskrivelsen af disse aspekter ved læring må være detaljeret for at kunne vejlede, men ikke så foreskrivende, at lærerens medejerskab i forhold til læseplanen krænkes. Som den forskning, som Harlen (2004) har sammenfattet, viser, så fremmes pålideligheden af evaluering, når lærerne har en gennemgribende forståelse af målene og af den karakter, som progressionen hen imod dem har. I Queensland gøres dette lettere, ved at skolerne har mulighed for at træffe beslutninger omkring deres egne arbejdsplaner, og ved at lærerne regelmæssigt deltager i denne moderationsproces. Det er også vigtigt at prioritere tiden til denne deltagelse og respekten for lærernes professionalisme (Cumming og Maxwell, 2004). Forhold, der fremmer pålideligheden, er tydeligvis essentielle, når lærernes evaluering har stor betydning for de enkelte elever. Et markant træk ved systemet i Queensland er dog, at evaluering af elever på Senior Certificate er uafhængig af skolens og lærernes ansvarsprocedurer.

Black m.fl. (2003) inkluderede den formative brug af summativ evaluering som en ud af fire praksisser, som lærere fandt effektive i forhold til at implementere formative evaluering (de andre var at udspørge, feedback gennem karaktergivning og elev-, kollega- og selvevaluering). Disse praksisser blev alle udtænkt og udviklet af lærere, mens de arbejdede sammen med forskere om at indføre ændringer i deres klasser, sådan at evaluering bedre kunne støtte læring. I forhold til at bruge summative test med formativt formål udtænkte lærerne tre primære måder at bruge test i klasseværelset til at udvikle elevernes forståelse, ud over bare at evaluere resultater. Den første gik ud på at hjælpe elever med at forberede sig til test ved at vurdere deres arbejde og se på tidligere testspørgsmål for at finde de områder, hvor der var usikkerhed i forståelsen. Ved at reflektere over de svage områder blev det muligt for eleverne at målrette deres repetition. Den anden nyskabelse var at bede eleverne om at formulere testspørgsmål og udtænke karaktergivningssystemer. Dette hjalp dem med "både at forstå evalueringsprocessen og at målrette den videre indsats for forbedring" (Black m.fl. 2003: 54). Den tredje forandring var, at lærerne skulle bruge te-

stresultaterne diagnosticerende og involvere eleverne i at rette hinandens test, i nogle tilfælde efter selv at have udarbejdet karaktersystemet. Dette har visse lighedspunkter med tilgangen beskrevet hos Carter (1997), som hun kaldte 'testanalyse'. Dette indebærer, at læreren gav testene tilbage til eleverne med markeringer af, hvor der var fejl, og så var det op til eleverne at finde og rette fejlene. Elevernes endelige karakterer afspejlede derved både deres respons på testanalysen og deres oprindelige svar. Carter beskrev dette som at lægge ansvaret for læringen over på eleverne, som blev opfordret til at arbejde sammen for at finde og rette deres fejl.

Mens der er klare fordele ved at bruge test på disse måder, er der også adskillige begrænsninger i disse tilgange. I hvor stor udstrækning denne information kan hjælpe lærerne med at vejlede eleverne i forhold til specifikke mål, afhænger for det første af, hvor ofte test og særlige opgaver blev stillet. Hvis disse opgaver for det andet er konstrueret til at opsummere læring i relation til generelle kriterier, som fx udsagn om resultater i nationale læseplaner, vil de ikke være detaljerede nok til at være diagnosticerende i en sådan grad, som er nødvendig for at støtte den specifikke læring. Selv om det er muligt at bruge visse eksterne test og eksaminationer på denne måde, fx ved at få fat i allerede bedømte opgavebesvarelser og diskutere dem med eleverne, så er tilgangen i praksis en, som lærere kan bruge principielt i sammenhæng med test i klasseværelset, som de har fuld kontrol over. Black (1993) pointerede, at når eksterne test er involverede, kan processen bevæge sig fra at udvikle forståelse til at 'undervise efter testen' [teaching to the test]. Mere generelt set er presset fra de eksterne test- og evalueringskrav ikke i fuld overensstemmelse med god formativ praksis (Black m.fl. 2003: 56). Der vil være en stærk tendens til hyppigt at indsamle, hvad der i bund og grund er summativ dokumentation, som erstatning for dokumentation, der kan bruges formativt. Mens de lærere, som Black m.fl. (2003) beskriver, benyttede deres kreativitet til at skabe formativ værdi ud fra summative procedurer, er det nødvendigt med en mere grundlæggende ændring, hvis evaluering skal konstrueres til at tjene begge formål fra begyndelsen.

De ti principper for evaluering til læring (ARG, 2002a) tilbyder en måde at undersøge, i hvilken udstrækning dokumentation fra summativ evaluering virkelig kan være formativ. Før man formoder, at den slags dokumentation kan støtte læringen, kan en lærer evt. bedømme det i forhold til disse principper ved fx at spørge:

- Fokuserer det på, hvordan elever lærer?
- Er det nuanceret og konstruktivt?
- Fremmer det motivation?
- Fremhæver det forståelsen af mål og kriterier?
- Hjælper det de lærende med at vide, hvordan de forbedrer sig?
- Udvikler det evnen til selvevaluering?
- Anerkender det alle uddannelsesmæssige præstationer?

Portfoliosystemet i Queensland kan siges at ramme de fleste af disse krav ganske godt; det er bestemt konstrueret med både et formativt og summativt formål for øje. Men det samme kan ikke siges om brugen af test og eksaminationer. Disse kan bidrage til at definere den videre læring, men kan aldrig være fyldestgørende i forhold til at møde kravene til evaluering til læring, fordi:

- Indsamlingen af summativ dokumentation finder ikke sted tilstrækkeligt ofte;
- Informationen er ikke tilstrækkeligt detaljeret til at være diagnosticerende;
- Det finder i virkeligheden kun sted i relation til opgaver, der er valgt, styret og bedømt af lærere;
- Der er en fare for fejlagtigt at antage, at summativ dokumentation, der indsamles hyppigt, kan bruges formativt, og derved forsømme den ægte evaluering til læring;
- Der er en risiko for at undervise med henblik på testen, når man bruger eksterne test på denne måde;
- Det lever sjældent op til principperne for evaluering til læring.

At bruge information fra formativ evaluering til summativ evaluering

Tilgangene, der er beskrevet ovenfor, er forbundet med summativ evaluering, hvilket er en begivenhed, der indtræffer lejlighedsvist. Ind i mellem test i klassen og karaktergivning af arbejde i faget foregår der utallige andre begivenheder i klasseværelset, hvor lærere og elever kan indsamle dokumentation om sidstnævntes igangværende bedrifter. Dette gør de ved at observere, spørge, lytte til diskussioner blandt eleverne, bedømme skriftligt arbejde og bruge elevernes selvevaluering (Harlen og James 1997). Som nævnt tidligere kan denne dokumentation, når den bliver brugt til at tilpasse undervisning og læring, give eleverne øjeblikkelig hjælp, eller den kan indgå i senere overvejelser og blive brugt til at planlægge efterfølgende læringsmuligheder.

Dokumentation, der er indsamlet på denne måde, er ofte ikke endegyldig og kan være selvmodsigende, da elevernes kunnen højst sandsynligt påvirkes af den specifikke kontekst. Denne variation, som ville være et problem i summativ evaluering, er brugbar information for de formative formål, da den antyder, i hvilke kontekster eleverne har brug for hjælp til at udvikle deres idéer og evner. Eftersom den indsamles i løbet af undervisningen, relaterer dokumentation med denne detaljeringsgrad per definition til alle aspekter af elevernes læring. Det bliver værdifuldt, da det relaterer til målene for den specifikke lektion eller aktivitet, og den kan derved bruges til bestemme de næste skridt for den enkelte lærende eller gruppe. Et vigtigt spørgsmål er: Kan denne fyldige, men til tider inkonsistente, dokumentation tjene formålene med summativ evaluering såvel som formålene med formativ evaluering, som den passer så godt til? Hvis ikke, bliver det nødvendigt med en adskilt summativ evaluering.

Et bekræftende svar på ovenstående spørgsmål gav Harlen og James (1997), som fremsatte, at både formative og summativeformål kan tjenes, hvis man skelner mellem dokumentationen og fortolkningen af dokumentationen. I forhold til formativ evaluering fortolkes dokumentationen i relation til elevens fremskridt hen imod målene for et bestemt stykke

arbejde, og de næste skridt planlægges i henhold til, hvor langt eleven er nået. Informationen støtter på den måde beslutningen om, hvad der skal gøres for at støtte den videre læring, og handler ikke om hvilket niveau eller hvilken karakter eleven har opnået. Der er to forbundne, men forskellige emner, der skal overvejes, når man bevæger sig fra at bruge dokumentation på denne måde til at bruge den til at opsummere det lærte. Disse er de involverede mål og grundlaget for bedømmelsen af dokumentationen.

Mål på forskellige niveauer

For det første vil lektionens mål, der deles med eleverne, være defineret i forhold til det specifikke emne for den lektion. Arbejdet med disse specifikke mål vil bidrage til udviklingen af en mere generel forståelse eller forbedret evne, altså i forhold til mål på et mere generelt niveau end lektionens specifikke mål. Målene for en specifik lektion kan fx inkludere en forståelse af, hvordan sneglels struktur og form passer til de steder, hvor snegle kan findes. Dette kan bidrage til en forståelse af, hvordan dyr generelt passer ind i deres habitater, men for at opnå dette må man se på et bredt udvalg af dyr, hvilket vil foregå i en række andre lektioner med deres egne specifikke mål. Evner som at planlægge en videnskabelig undersøgelse er ligeledes udviklet, ikke på én lektion, men i flere forskellige kontekster i forskellige lektioner.

Så mens lærere og elever kan bruge dokumentationen til at fremme læringen i forhold til specifikke mål, kan dokumentation fra flere lektioner bidrage til et mere generelt læringsresultat. For at opnå dette er det vigtigt, at lærere fokuserer på progressionen i forhold til de forståelser og færdigheder, som de ønsker, eleverne skal opnå i de enkelte lektioner. Progressionen kan udtrykkes gennem indikatorer, som både retter opmærksomheden mod relevante aspekter ved elevernes adfærd og sætter lærerne i stand til at se, hvor eleverne befinder sig i udviklingen. I Boks 6.1 er et eksempel på indikatorer for udviklingen i at kunne planlægge videnskabelige undersøgelser i natur/teknik i folkeskolen.

Et eksempel på indikatorer for udvikling

Boks 6.1

Hvad eleverne gør som indikatorer på planlægning af videnskabelige undersøgelser:

- Foreslår en brugbar tilgang til at besvare et spørgsmål eller teste en formodning gennem undersøgelse, selv hvis der mangler detaljer eller yderligere overvejelser;
- Kommer med forslag til, hvad der sker, når man laver bestemte ændringer;
- Identificerer den variabel, der skal ændres, og de ting, der skal forblive uændret, for at en test bliver fair;
- Identificerer, hvad man skal se efter eller måle for at skaffe sig et resultat i en undersøgelse;
- Vælger og bruger udstyr og måleapparater, der er egnede i den foreliggende opgave;
- Succesfuldt planlægger en fair test gennem støtte fra spørgsmål eller et planlægningskema;
- Spontant konstruerer en plan til at identificere variabler og tiltag, der kan gøre resultaterne så nøjagtige som muligt.

(Harlen, 2005)

Disse indikatorer er blevet udviklet med udgangspunkt i, hvad man ved om progression fra forskning og praksis, men de er på ingen måde definitive. Det er ikke sandsynligt, at der findes en nøjagtig og uforanderlig sekvens, der kan bruges i forhold til enhver elev, men det er nyttigt at have en grov skitse. Eksempler på lignende lister er blevet udgivet i Australien (Masters og Foster, 1996) og udviklet i Californien (som fx systemet kaldet Berkeley Evaluation and Assessment Research (BEAR), Wilson, 1990; Wilson m.fl., 2004). På disse lister indikerer de første udsagn forståelser, færdigheder og holdninger, der sandsynligvis udvikles før de udsagn, der følger senere listen. Der er ingen grund til at foreslå 'niveauer', klassetrin eller faser;

udelukkende en sekvens, der forventes af elever inden for en bestemt aldersgruppe (som aldersgruppen, der har natur/teknik i eksemplet ovenfor). I formativ evaluering er det ikke nødvendigt at forbinde disse indikatorer med forventninger til et klassetrin; det eneste, der er nødvendigt, er at bemærke, hvor eleverne befinder sig, og hvad det næste skridt i deres videre fremskridt skal være. Dette er i overensstemmelse med den slags feedback til eleverne i formativ evaluering, som bør være ikke-bedømmende og fokuseret på læringens næste skridt.

Det er vigtigt, at lærerne grundigt forstår, tilegner sig og føler medejerskab over for disse indikatorer, hvis de skal blive nyttige i det daglige arbejde med eleverne. Dette medfører dog et potentielt problem i forhold til at sikre, at de er omhyggeligt og strengt defineret for at sikre validitet, og at lærerne samtidigt forstår og føler medejerskab i forhold til dem. Det er ikke realistisk, at lærerne selv skal udfærdige sådanne indikatorer til sig selv, ej heller ville det skabe valide resultater; det kræver bidrag fra forskere og udviklere, der har ekspertise inden for elevers udvikling såvel som erfaring fra klasseværelset. Men det er samtidig nødvendigt at undgå, at en samling indikatorer pålægges ovenfra, hvilket kunne resultere i, at brugen af dem bliver en mekanisk proces. Forskning antyder dog, at dette kan undgås, og at medejerskab kan fremmes gennem øvelse, som inkluderer erfaringer med at konstruere nogle indikatorer, uden at de skal dække hele spektret og derved kræver den grundighed, der er nødvendig, hvis de skal kunne bruges af alle lærere i forskellige omstændigheder.

Hvad angår afrapporteringen af opnåelser, skal der tages et andet skridt, eftersom de progressive kriterier er for detaljerede for de summative formål. Den opsummerende afrapportering kræver en overordnet bedømmelse af, hvad der er blevet opnået i form af fx 'viden om og forståelse af livsprocesser og levende ting' eller 'færdigheder inden for videnskabelig undersøgelse'. Det er nødvendigt med en yderligere sammenlægning af dokumentation i forhold til det omfang af idéer og processer, som disse globale begreber indebærer. For nu at tage et eksempel på færdigheder inden for videnskabelig undersøgelse så vil

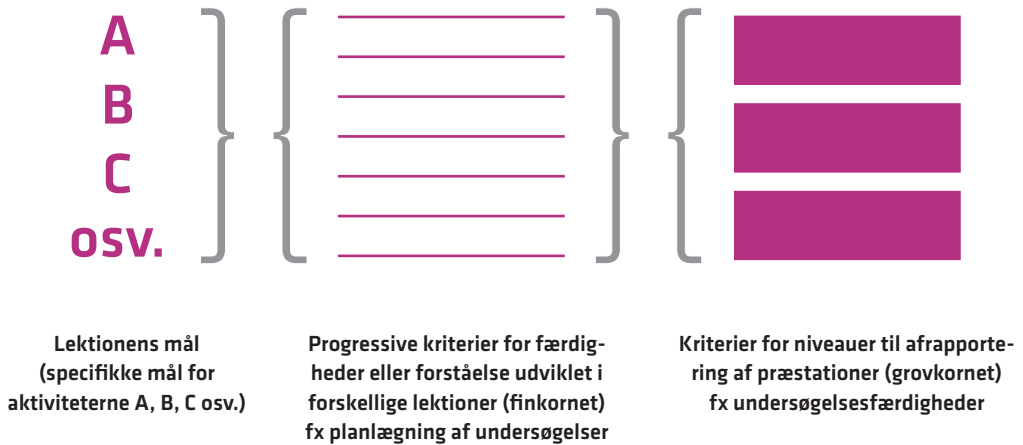
dokumentation fra adskillige lektioner, hvor målet har været at støtte eleverne i at planlægge videnskabelige undersøgelser, kunne påvise, hvor langt eleverne er nået i forhold til at kunne planlægge undersøgelser. Denne færdighed er igen en del af en større intention om at udvikle evnen til at undersøge, hvilket er et mål i den nationale læreplan, og som lærerne behandler i en sammenfatning i slutningen af året eller i rapporter på centrale stadier. Figur 6.3 afbilder således en tretrins-proces i brugen af information fra de enkelte lektioner til at lave en opsummering i forhold til bestemte kategorier. Men denne sammenlægning er ikke lige til, hvilket vi ser i relation til det andet spørgsmål, der må overvejes i opsummeringen af dokumentation: grundlaget for bedømmelse.

Ændring af grundlaget for bedømmelse af dokumentation

Når dokumentation indsamles i en lektion (A, B, C, osv. i Figur 6.3), kan det som nævnt benyttes med det samme eller senere for at støtte elever eller grupper i at nå målene for lektionen. For at gøre dette må læreren fortolke dokumentationen i forhold til de involverede elevers fremskridt, således at bedømmelsen tager udgangspunkt i eleven og samtidig er relateret til kriterierne. Hvis dokumentationen også skal bruges til at dokumentere bredere aspekter ved færdigheder eller viden, illustreret i bevægelsen fra venstre mod højre i Figur 6.3, så må det bedømmes ud fra kriterierne alene. Således kan dokumentationen bruges med to formål, hvis altså det omfortolkes i henhold til kriterier, der er ens for alle elever. Dette betyder, at hvis information, der allerede er indsamlet og brugt formativt, skal bruges i summativ evaluering, må det vurderes i forhold til de bredere kriterier, der definerer niveauer eller klassetrin. Dette gøres ved at finde det bedste match blandt den indsamlede dokumentation for hver elev og et af niveauerne. I denne proces kan man tage højde for udviklingen over tid, sådan at man foretrækker den dokumentation, der viser fremskridt i løbet af perioden, der er i fokus for densummative evaluering, som ved portfolioevalueringen i Queensland.

Mål på forskellige detaljeniveauer

Figur 6.3



Disse overvejelser viser, at det er muligt at benytte dokumentation, der er indsamlet som en del af undervisningen, både til at støtte læring og til at afrapportere. Men ligesom det er tilfældet ved brugen af summativ data til formative formål, har processen visse begrænsninger. I denne kontekst udføres bedømmelsen per definition af læreren, og i forhold til de summative formål må der være en sikring af pålideligheden. Derfor må der være nogle procedurer for kvalitetssikring. Jo mere betydning den summative bedømmelse tillægges, des mere stringent må kvalitetssikringen være. For at moderere bedømmelserne af dokumentationen kan den evt. inkludere samarbejde eksternt mellem skoler eller internt på skolen. Dette er dog svært i forbindelse med resultater, hvor dokumentationen er flygtig. En konsekvens af dette kan være, at den endelige brug af dokumentationen påvirker, hvilken dokumentation der indsamles og hvordan. Dette kunne resultere i, at dokumentationen indsamles med udgangspunkt i en færdig tjekliste eller i en række særlige opgaver, der gi-

ver konkret dokumentation, hvilket gør formativ evaluering til en efterfølger til summativ evaluering.

Der ligger en yderligere begrænsning i, at hvis summativ evaluering udelukkende er baseret på dokumentation indsamlet i en kontekst af almindelige aktiviteter i klasseværelset, så vil denne dokumentation være begrænset af omfanget og fyldigheden af de uddannelsesmæssige bestemmelser og lærernes effektivitet i indsamlingen af dokumentation. Under nogle omstændigheder kan det være nødvendigt at supplere den dokumentation, der skal til for at opsummere læring, med særlige opgaver, hvis en lærer fx af en eller anden grund har været ude af stand til at indsamle, hvad der er nødvendigt for at bedømme alle eleverne. Til de formative formål er det ofte passende at vurdere en gruppes fremskridt frem for de enkelte elevers. Da kan man så få brug for yderligere dokumentation, når der skal skrives en rapport over de enkelte elevers præstationer.

Opsummerende kan man sige, at begrænsningerne ved at bruge dokumentation, der er indsamlet til formativ evaluering, men som skal leve op til kravene for summativ evaluering, er:

- Det er essentielt at omfortolke dokumentationen i forhold til de samme kriterier for hver enkelt elev;
- Det er vigtigt, at lærerne er helt på det rene med, hvornår kriterier med udgangspunkt i niveau er passende, og at de ikke bruger dem til at give eleverne karakterer, når formålet med evalueringen er formativt;
- Eftersom den formative brug af dokumentation afhænger af læreres bedømmelse, vil det være nødvendigt med yderligere kvalitetssikringsprocedurer, når dokumentationen benyttes til et andet formål;
- Lærere kan få brug for at supplere dokumentation fra aktiviteter i klasseværelset med særlige opgaver for at sikre, at den nødvendige dokumentation for alle eleverne bliver indsamlet;
- Det besværlige ved at have med flygtig dokumentation at gøre kan føre til, at man bruger en færdig tjekliste eller en række summative opgaver;
- Dette kan meget vel ændre den formative evaluerings væsen og altså gøre den mere formel.

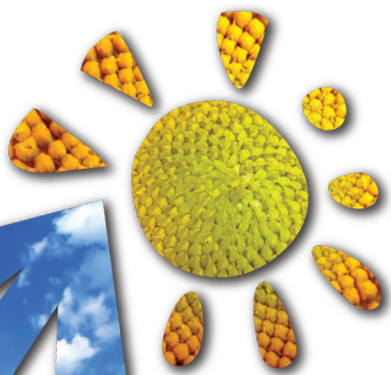
Genovervejelse af forholdet

Et modsætningspar eller en skala?

Diskussionen i de foregående to dele af dette kapitel indikerer, at der ikke er et skarpt skel mellem evaluering til læring og evaluering, der rapporterer om læring. Mere specifikt er det muligt at bedømme dokumentation både formativt som støtte til beslutninger om de næste skridt og summativt for at indikere, hvor langt eleverne er nået i forhold til de progressive kriterier i midterspalten i Figur 6.3. Dette antyder, at forholdet mellem formativ og summativ evaluering kunne beskrives som 'en skala' frem for 'et modsætningspar'. I Figur 6.4 indikeres nogle punkter på denne skala (taget fra Harlen, 1998).

Yderpunkterne markerer den praksis, der kendetegner evaluering til læring og evaluering af læring mest. I den rent formative ende finder vi evaluering, der er en del af interaktionen mellem elev og lærer, og som også er en del af elevens rolle. Læreren og eleven vurderer arbejdet i forhold til de mål, der er sat for den specifikke lærende, således at bedømmelserne i bund og grund tager udgangspunkt i eleven. Det centrale formål er at sætte lærer og elever i stand til at træffe beslutning om næste skridt i forhold til læringen og at vide, hvordan disse skridt skal tages. I den rent summative ende af skalaen er formålet at dokumentere, hvad der er blevet opnået på bestemte tidspunkter. Evalueringen til dette formål burde resultere i en pålidelig rapport over hver enkelt elevs præstationer. Selv om selvevaluering kan være en del af denne proces, ligger det endelige ansvar for en retfærdig vurdering af, hvordan den enkelte elevs læring skrider frem i forhold til kriterierne eller standarderne, hos læreren.

Mellem disse yderpunkter er der en stor spændvidde af procedurer, der spiller forskellige roller i undervisning og læring. Mange lærere ville fx indlede et nyt emne med at finde ud af, hvad eleverne allerede ved, for bedre at kunne tilrettelægge undervisningsplanen i stedet for med det formål at finde de punkter, hvor hver enkelt elev kan udvikle sig. På samme måde laver lærere ofte en uformel test mod slutningen af et forløb for at vurdere, hvorvidt nye idéer har fundet grobund eller har brug for yderligere forankring.



En mulig skala for formål og praksisser i evaluering

Figur 6.4

	Uformel formativ	Formel formativ	Uformel summativ	Formel summativ
Vigtigste fokus	Hvad er næste skridt i læringen?		Hvad er blevet opnået indtil nu?	
Formål	At støtte beslutning om næste skridt i læring	At støtte beslutning om næste skridt i undervisning	At overvåge fremskridtet i forhold til planerne	At dokumentere individernes præstationer
Hvordan indsamles dokumentation?	Som en almindelig del af arbejdet i klassen	Indføres i det almindelige arbejde i klassen	Indføres i det almindelige arbejde i klassen	Adskilt opgave eller test
Bedømmelsens grundlag	Udgangspunkt i eleven	Udgangspunkt i eleven og kriterier	Udgangspunkt i kriterier	Udgangspunkt i kriterier
Bedømmes af	Elev og lærer	Lærer	Lærer	Lærer og ekstern censor
Medfølgende handling	Feedback til elever og lærer	Feedback til undervisningsplaner	Feedback til undervisningsplaner	Rapport til elev, forældre, andre lærere osv.
Tilnavn	Evaluering til læring	Tilpasning	Temperaturmåling	Evaluering af læring

Formativ ←————→ **Summativ**

Der er her nogle referencer til mellemliggende formål med evaluering, som andre har defineret. Cowie og Bell (1999) observerede ti læreres evalueringsskikker i New Zealand og udledte heraf to former for formativ evaluering: planlagt og interaktiv. Planlagt formativ evaluering vedrører hele klassen, og lærerens formål er at finde ud af, hvor langt læringen er skredet frem i forhold til, hvad der forventes i læreplanen. Indsamlingen af information, som fx kan være en kort test i klassen eller en særlig opgave, er planlagt og forberedt på forhånd, og resultaterne ledes tilbage til undervisningen. Dette ligner 'uformel summativ' ovenfor. Interaktiv formativ evaluering er ikke planlagt på forhånd på denne måde; den udgår fra læringsaktiviteten. Dens funktion er at støtte individernes læring, og den rækker ud over de kognitive aspekter ved læring til social og personlig læring. Feedback sker øjeblikkeligt og er både rettet mod lærer og de lærende. Den har kendetegnene fra 'uformel formativ'.

Cowie og Bells interaktive formative evaluering ligner den evaluering i klassen, som Glover og Thomas (1999) beskriver som 'dynamisk'. Ligesom Black og Wiliam (1998a) fremhæver de elevernes involvering i læringen og taler endda om at 'overdrage magt til de lærende' og antyder, at uden dette er dynamisk evaluering ikke muligt. I modsætning til Cowie og Bell påstår de dog, at al evaluering må planlægges.

Der er også forskellige grader af formalitet hen mod den summative ende. Det, der beskrives som 'uformel summativ', kan indebære en praksis, der ligner 'formel formativ', hvilket er illustreret i Figur 6.4. Den væsentlige forskel er dog måden, man bruger dokumentationen på. Hvis ringen er sluttet, som i Figur 6.1, og dokumentationen bruges til at tilpasse undervisningen, da er det 'formel formativ'. Hvis der ikke er feedback til undervisningen, som i Figur 6.2, falder det ind under kategorien 'uformel summativ', selv om dokumentationen måske er selv samme test.

Frem for et forsøg på at lave endnu flere distinktioner mellem evalueringsprocedurer burde denne analyse måske ikke tages for andet end en indikation på, at der findes forskellige måder at praktisere og bruge formativ evaluering og summativ evaluering. Hvis dette er tilfældet, har vi så overhovedet brug for denne distinktion?

Findes der formativ og summativ evaluering eller bare god evaluering?

Nogle af dem, der arbejder med udvikling af evaluering, har argumenteret for at selve distinktionen mellem formativ og summativ ikke er nyttig, og at vi alene skal stræbe efter 'god evaluering'. God formativ evaluering vil støtte læreres gode bedømmelser af elevers fremskridt og opnåelser, og god summativ evaluering vil bidrage med feedback, der kan bruges til at fremme læring. Maxwell (2004) beskriver progressiv evaluering som en sløring af grænsen mellem formativ og summativ evaluering.

Diskussionen af Figur 6.4 tyder bestemt på en sløret grænse. I tillæg hertil synes anerkendelsen af, at dokumentation kan bruges til begge formål, ved første øjekast at støtte en ophævelse af distinktionen mellem formativ og summativ evaluering. I begge tilfælde er der begrænsninger i den todelte brug af dokumentationen, men ved nærmere eftersyn viser disse sig at være af forskellig art. Begrænsningen ved at bruge dokumentation, der oprindeligt var indsamlet med et summativt formål for øje, til at fremme læringen ligger i dokumentationens validitet; det er bare ikke tilstrækkelig fyldigt og let tilgængeligt til at være fyldestgørende i formativ brug. Begrænsningen ved at bruge dokumentation, der oprindeligt var indsamlet til at fremme læring, til at rapportere om læring ligger i dokumentationens pålidelighed. I dette tilfælde er der dog skridt, der kan tages for at komme denne begrænsning i møde og øge pålideligheden: øvelse kan sikre, at lærere indsamler dokumentationen systematisk og med integritet, mens moderation kan optimere mulighederne for sammenligning.

Når sådanne procedurer til at sikre kvaliteten er på plads, kan dokumentation, der er indsamlet af lærere, med en detaljeringsgrad, der passer til at kunne

fremme læringen, også bruges til pålidelig evaluering af læring. Dette er, hvad der sker på Senior Certificate i Queensland, hvor alle de nødvendige informationer til at give eleverne karakterer indsamles og bedømmes af lærere. Den omvendte situation kan dog ikke findes; der er ingen eksempler på, at den nødvendige viden til evaluering til læring kommer fra dokumentation, der er indsamlet med summative formål for øje. Det er selvfølgelig ikke logisk, at det kunne være tilfældet.

Asymmetrien i den todelte brug synes at være et stærkt argument for at bibeholde distinktionen mellem formål. Vi har brug for at vide, til hvilket formål dokumentationen var indsamlet, og til hvilket formål det er brugt. Først derefter kan vi vurdere, om det er 'godt' eller ej. Man kan udføre den samme evaluering og bruge den til forskellige formål, ligesom man kan rejse mellem to steder med forskellige formål. Ligesom formålet er grundlaget for at vurdere rejsens succes, sætter evalueringens formål os i stand til at vurdere, om formålet er blevet opfyldt. Hvis vi forbinder eller forveksler formative og summative formål, tyder erfaringen kraftigt på, at 'god evaluering' vil betyde god evaluering af læring, og ikke til læring.

Konklusion

Begrebet progression er et af to nøglebegreber i diskussionen i dette kapitel. For at udvikle elevers forståelse og færdigheder må lærere være bevidste om nogle kriterier for udvikling for at kunne se, hvordan målene for den specifikke lektion er forbundet med progressionen af mere generelle begreber og færdigheder. Hvad angår formative formål, behøver disse kriterier ikke være forbundet med niveauer; de skal bare vejlede i forhold til læringens næste skridt. Hvad angår formålene med summativ evaluering, hvor en opsummering er påkrævet, er brugen af niveauer, standarder eller karakterer en måde at kommunikere, hvad en elev har opnået i forhold til de kriterier, der er repræsenteret gennem disse niveauer, standarder eller karakterer. Denne proces fortætter dokumentationen og medfører nødvendigvis et tab af detaljer. Når man bruger summativ information, kræver det også, at niveauerne og deslige er de samme for alle elever. Hvis elev A opnår et niveau eller karakteren X,



betyder det, at denne elev har opnået stort set det samme som elev B, der også opnåede dette niveau eller karakteren X – hvis man for et kort øjeblik ser bort fra muligheden for at fejlvurdere, hvilket diskuteres i kapitel 7. Det er kun for nemheds skyld, at vi benytter bestemte niveauer; i teorien kunne vi dokumentere præstation i form af en profil lavet ud fra en række kriterier for progression, men dette ville sandsynligvis generere for mange detaljer til de fleste formål, hvor man efterspørger en kortfattet opsummering.

Det er de forskellige formål med informationen, hvilket er det andet nøgleaspekt i dette kapitel, der netop skaber distinktionen mellem formativ og summativ evaluering. Vi har argumenteret for, at dokumentation, der er indsamlet som en del af undervisning og læring, kan bruges til både formative og summative formål. Det bruges til at støtte læring, når det fortolkes i forhold til de enkelte elevers fremskridt hen imod lektionens mål. Den samme dokumentation kan også fortolkes i forhold til generelle kriterier, der bruges, når man skal rapportere om præstationer ved hjælp af niveauer eller karakterer. Vi har dog bemærket, at dokumentation, der allerede er en opsummering, når det indsamles, som i en test eller eksamination, generelt mangler de nødvendige detaljer til at identificere og støtte beslutninger om læringens næste skridt. Dette betyder, at vi ikke kan bruge enhver dokumentation til ethvert formål. Dette taler for at bibeholde en klar distinktion mellem formativ og summativ i forhold til brugen af dokumentationen.

Selv om der er forskellige grader af formalitet i måderne at bedrive formativ og summativ evaluering på, som udtrykt i Figur 6.4, er der dog stadig forskel på formål. Vi kan ikke antage, at måden, dokumentatio-

nen er indsamlet på, bestemmer brugen i forhold til læring; en test i klassen kan bruges til at give input til undervisningen uden nogen reference til niveauer, og den kan bruges til at give en karakter eller bestemme et niveau i slutningen af et forløb. Det asymmetriske forhold – at dokumentation indsamlet til formativ evaluering kan bruges til summativ evaluering, men ikke omvendt – betyder, at hvis man fjerner disse mærkater og kun refererer til 'evaluering', vil det uundgåeligt favorisere summative formål.

Det er både en svaghed og en styrke, at summativ evaluering, der har udgangspunkt i omfortolkningen af formativ dokumentation, medfører, at begge dele ligger i lærerens hænder. Svagheden ligger i de kendte fordomme og fejl, der forekommer i læreres bedømmelser. Al evaluering indebærer bedømmelse og vil derfor være underlagt visse fejl og fordomme. Dette aspekt har fået en del opmærksomhed i forhold til læreres evaluering med summative formål, men det eksisterer uden tvivl også i læreres evalueringer, der har formative formål. Det er dog ikke nødvendigt at være meget bekymret for pålideligheden af evaluering med dette formål (fordi det sker ofte, og læreren har mulighed for at bruge feedback til at rette fejlagtige bedømmelser). Men jo mere omhyggelig en evaluering udføres, des mere værdi har den i forhold til at fremme læringen. Styrken ligger derfor i, at de procedurer, der sikrer en mere pålidelig summativ evaluering, som netop må være der i et system, der bruger læreres bedømmelser, også vil gavne den formative brug, læreres forståelse af læringsmålene og progressionens karakter i opnåelsen af dem. Erfaring viser, at moderation af læreres bedømmelser, hvilket er nødvendigt, når man bruger eksternsummativ evaluering, kan udføres sådan, at dette ikke alene fungerer som

kvalitetskontrol, men også bliver en kvalitetssikring, der har indflydelse på læreres evalueringsproces (ASF, 2004). Dette vil forbedre indsamlingen og brugen af dokumentation til et formativt såvel som et summativt formål.

Dette kapitel har udforsket forholdet mellem formativ evaluering og summativ evaluering med et perspektiv på brugen af samme dokumentation til begge formål. Vi har set, at der er potentielle farer for formativ evaluering, hvis man antager, at dokumentation indsamlet til summativ evaluering også kan tjene formative for-

mål. Ligeledes må der tages yderligere forholdsregler, hvis summativ evaluering baseret på dokumentation, der er indsamlet og brugt til formative formål, skal være tilstrækkelig pålidelig. Disse aspekter er centrale for beskyttelsen af evalueringens integritet, og i særdeleshed integriteten i formativ evaluering, således at evaluering har en positiv indflydelse på læringen, hvilket er omdrejningspunktet i denne bog.

Referencer

- ARG (2002) *Assessment for Learning: 10 principles*. University of Cambridge: Assessment Reform Group.
- ASF (2004) *Working Paper 2, Draft 3 Summative Assessment by Teachers: Evidence from research and its implications for policy and practise*. Assessment Systems for The Future Project.
- Black, P. ('Formative and summative assessment by teachers'. *Studies in Science Education*, 21: 49-97.
- Black, P., Harrison, C., Lee, C., Marshall, B. og William, D. (2003) *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.
- Black, P og William, D. (1998) *Inside the Black Box: Raising standards through classroom assessment*. London: King's College (see also *Phi Delta Kappan* 80: 139-48).
- Carter, C. R. (1997) 'Assessment: shifting the responsibility'. *The Journal of Secondary Gifted Education*, 9(2): 68-75.
- Cowie, B. og Bell, B. (1999) 'A model of formative assessment in science education'. *Assessment in Education*, 6(1):101-16.
- Cumming, J. og Maxwell, G.S. (2004) 'Assessment in Australian schools: current practice and trends'. *Assessment in Education*, 11(1): 89-108.
- DES/WO. (1988a) *Task Group on Assessment and Testing: A Report*. London: Department of Education and Science and the Welsh Office.
- Frederiksen, J.R og White, B.Y. (2004) 'Designing assessment for instruction and accountability: an application of validity theory to assessing scientific enquiry', in M. Wilson (red.) *Towards Coherence between Classroom Assessment and Accountability: 103th Yearbook of the National Society for the Study of Education Part II*. Chicago: National Society for the Study of Education, pp. 74-104.
- Glover, P. og Thomas, R. (1999) 'Coming to grips with continuous assessment', *Assessment in Education*, 4(3): 365-80.
- Harlen, W. (1998) 'Classroom assessment: a dimension of purposes and procedures'. Paper presented at the annual conference of the New Zealand Association for Research in Education, Dunedin, December.
- Harlen, W. (2000) *Teaching, Learning and Assessing Science 5-12*. 3rd edition. London: Paul Chapman Publishing.
- Harlen, W. (2004) 'A systematic review of the reliability and validity of assessment by teachers used for summative purposes', in *Research Evidence in Education Library Issue1*. London: EPPI-Centre, Social Sciences Research Unit, Institute of Education.
- Harlen, W. (2005) 'Teachers' summative practices and assessment for learning: tensions and synergies'. *Curriculum Journal* 16(2): 207-23.
- Harlen, W. og James, M. (1997) 'Assessment and learning: differences and relationships between formative and summative assessment'. *Assessment in Education*, 4(3): 365-80.
- Masters, G og Foster, M. (1996) *Progress Maps*. Victoria, Australia: Australian Council for Educational Research.
- Maxwell, G.S. (2004) 'Progressive assessment for learning and certification: some lessons from school-based assessment in Queensland'. Paper presented at the third Conference of the Association of Commonwealth Examination and Assessment Boards, March, Nadi, Fiji.
- Rousseau, J.-J. ([1762] 1961) *Emile*. London: Dent.
- Wilson, M. (1990) 'Measurement of developmental levels', in T. Husen og T.N. Postlethwaite (red.) *International Encyclopedia of Education: Research and Studies (Supplementary volume)*. Oxford: Pergamon Press.
- Wilson, M., Kennedy, C. og Draney, K. (2004) *GradeMap (version 4.0) [computer program]* Berkeley, CA: University of California, BEAR Center.

