

Evaluering

for læring i naturfagene

Af Thomas Bech Lillelund Andersen

Abstract

Evaluering har de seneste årtier indtaget en mere central rolle end tidligere, dette gælder både på globalt samfundsplan og i uddannelsesinstitutionerne. Denne artikel handler om vigtigheden af at overveje, hvilke evalueringsformer der anvendes til hvilke former for målsætninger. Læring er en kompleks størrelse, og evaluering i form af test er kun sjældent svaret. Det gennemgås her, hvorfor det er problematisk at fokusere for meget på test, og hvilke former for evaluering der kan gøres brug af i stedet. Der argumenteres for, at valget af test hviler på en for spinkel forståelse af, hvad viden er, og mangel på indsigt i, hvilke slutninger der kan drages på baggrund af denne viden.

Menneske eller samfund først?

Evaluering virker som et nyt fænomen, men i virkeligheden har det at vurdere ens omgivelser nok altid været en del af menneskets natur. Den tidlige form for evaluering har været i forbindelse med uddannelse til forskellige håndværk, hvor en øvet praktiker og en eller flere mestre har bedømt den ikke færdiguddannedes færdigheder. Terrakottamestre i Kina, guldsmede i Ægypten og bronzestøbere i Danmark har sikkert evalueret, hvorvidt de nye håndværkere var udlært eller ej. I dag findes denne evalueringsform stadig primært hos de traditionelle håndværksfag.

Det er først under industrialiseringen sidst i 1800-tallet, at evaluering for alvor bruges til sammenlignelighed og konkurrence og ændrer form, først til IQ-test og senere til skriftlige prøver (Broadfoot, 2007, s. 20). Det er vigtigt at gøre sig klart, at dette skete for at sikre sammenlignelighed og ensartethed, så krav og niveau i forhold til uddannelser og arbejdsgivere var gennemsigtige og pålidelige. Man gik fra godkendelse af praktikere til systemer bygget op om at kunne konkurrere på lige vilkår. Det er dette fokus, der stadig dominerer evalueringspolitikken.

Samtidig med behovet for sammenligning og ensartethed har et mere menneskecentreret syn på uddannelse gjort sig gældende. Tankerne om den enkeltes mulighed i verden, dannelsestanken, har blomstret siden 1800-tallet og er i dag det fundament, vi bygger vores folkeskolelov på – ja, i princippet det, alle demokratier hviler på. Disse to strømninger – det menneskecentrerede og det samfundscentrerede – har udviklet sig side om side, men til stadighed i konflikt. I det følgende citat af den tyske dannelsesfilosof Johann Friedrich Herbart ses denne gryende konflikt.

”En degradering kan opstå, hvis den erhvervede viden tjener til at blive stillet til skue for bifalds skyld eller på anden måde bringe ydre fordele; dette er en uheldig side ved mange offentlige eksaminer” (Herbart 1841/1980, s. 53).

Dominerende tendenser – evalueringsparadigmer

Interessen for ensartethed og sammenlignelighed har historisk set været dominerende. Opfattelsen af, at man ved hjælp af ’objektive metoder’ har kunnet evaluere hvad som helst, var i Danmark fremherskende helt frem til midten af 1970’erne (Miller, 2007, s. 187). Inden for evaluering taler Tanja Miller om tre forskellige ’paradigmer’.

- Det tekniske paradigme: står for, at alt kan testes – det er bare et spørgsmål om metode og teknik.
- Det politiske paradigme: evaluering som en del af en politisk værdikamp. Denne kamp tager udgangspunkt i øget opmærksomhed på utilsigtede effekter af evaluering.
- Det kulturelle paradigme: evaluering, som er mere værdineutral og har sin egen kultur, forskellige former for evaluering passer til forskellige læringsituationer og kan forbedre læring (Miller, 2007, s. 194-195).

Det kulturelle paradigme udviklede sig op igennem 80’erne, men slog aldrig ordentligt igennem. Det skyldes blandt andet sen lovgivning på området (1993)(Miller, 2007, s. 203). Denne pragmatiske måde at betragte evaluering på tilgodeser, at man må vælge form efter, hvad der skal evalueres, og at denne form ofte bruges i en læringsproces, men disse former egner sig ikke nødvendigvis til sammenlignelighed. Vi er med det kulturelle paradigme ved at være fremme ved en individcentreret lærings- og dannelsesorienteret anvendelse. Det hele fik dog en brat ende med folkeskoleloven i 2006, hvor det ’tekniske paradigme’ genvandt sin dominans. Et udtryk for det tekniske paradigmes sejr ses i naturfagene i form af de nationale test og afgangsprøver som test (biologi og geografi). Vi er på denne måde tilbage ved udgangspunktet: tiden omkring og før industrialiseringen i 1800-tallet (Miller, 2007, s. 207).

Da der i 70'erne blev gjort op med det tekniske paradigme, dukkede der to væsentlige begreber op, der blev brugt til at skelne mellem forskellige former for anvendelse af evaluering. I tiden op til 70'erne havde man kun brugt evaluering på en måde, vi i dag kalder summativ. *Summativ evaluering* er ikke en speciel form for evaluering, men en måde at tænke og anvende forskellige former for evaluering på. At en evaluering er anvendt summativt angiver blot, at den anvendes som slutevaluering, og at det, man er interesseret i, er det, der ligger bagud i tiden – altså det, eleven har lært op til evalueringen.

Den klassiske summative situation er afgangsprøver og eksamener. Mange af de summative former har været velegnede til sammenligning og ensartethed. Tiden op til 70'erne har netop været præget af den summative anvendelse af evaluering, så det nye er den formative anvendelse af evaluering. *Formativ evaluering* kan også betegnes som en procesevaluering (Broadfoot, 2007, s. 7). Det vigtige ved denne anvendelse er, at evalueringen anvendes som en del af en læringsproces. Ved formativ evaluering anvendes resultaterne af et evalueringssværktøj til at vurdere den lærendes læring med henblik på at tilrettelægge læringsprocessen fremadrettet. Vi vil senere se på disse to anvendelser i forhold til konkrete evalueringsformer i naturfag.

Evaluering er styrende for indholdet i undervisningen

I 70'erne blev man også opmærksom på det, den valgte evalueringsform gør ved underviseren, og de valg, han/hun tager i forhold til målsætninger og tilrettelæggelse af undervisningen (Tyler, 1974, s. 98).

Peter Dahler-Larsen kalder denne evalueringseffekt for 'den konstitutive'. At evaluering har konstitutiv effekt betyder netop, at den ofte vil bestemme metode og indhold i undervisningen, det er altså helt naturligt sådan, at underviseren lader sig påvirke af evalueringsformen og på denne måde planlægger 'baglæns' – altså at målsætningen for faget og valg af metoder ikke kommer først, men sidst. Denne effekt kan derfor skade naturfagsundervisningen og tilsidesætte både folkeskoleloven, Fælles Mål for fagene samt et århundredes didaktisk forskning.

"Virningen kan også være umærkeligt snigende, når lærere mere eller mindre bevidst begynder at planlægge undervisningens indhold og metodevalg ud fra forestillinger om, hvad der giver pote evalueringsmæssigt" (Dahler-Larsen, 1999, s. 159).

Alene denne konstitutive effekt gør, at vi nøje må overveje, hvordan vi evaluerer, så vi kan sikre os en sammenhæng mellem målsætning, metode og evalueringsform. Vi kan tale om kohærens, når der netop er denne sammenhæng mellem det, der ligger forud for evalueringen, og selve evalueringen, ligesom at en evaluering kun kan være valid, altså gyldig, hvis den netop evaluerer på en måde, som tager hensyn til det, der faktisk er foregået

i undervisningen. Kohærens og validitet er derfor vigtige begreber, når man planlægger hele undervisningsforløb.

”Der er en tendens til at fokusere på, om man kan iagttage en eller anden adfærd eller kvalitet, snarere end på den proces, hvorigennem adfærden eller kvaliteten er erhvervet” (James, 2009, s. 47).

Problemet er, at evaluering med summative test let fører til, at underviseren vælger én bestemt metode, hvor virkeligheden er, at vi i dag står med en vifte af læringsteorier, der begrundes forskellige metoder og målsætninger. James anvender en tredeling i læringsteorier:

Behaviorisme: Læring ses som respons på ydre stimuli. Bevidstheden, intelligensen og jeget har ingen plads. Evalueringsformen vil være test – summativ.

Kognitive, konstruktivistiske teorier om læring: Læring forudsætter aktiv indsats af den lærende og afhænger af, hvad der foregår i elevernes hoveder. Allerede tilegnet viden er vigtig for muligheden for at lære nyt. Evaluering vil være formativ anvendelse af for eksempel begrebskort.

Sociokulturelle, situerede og virksomhedsorienterede teorier om læring: Læring er socialt afhængig. ”Tænkning udføres gennem handlinger, som ændrer situationen, og situationen ændrer tænkningen.” (James 2009, s.53). Læring hjælpes ved anvendelse af redskaber som sprog og bøger. Formativ anvendelse af evaluering, for eksempel portfolio.

Den forståelse af læring, der har udviklet sig siden slutningen af 1800-tallet, har taget afstand fra kun at betragte læring som påfyldning af viden, som behavioristerne forstår det. Hvis Mary James har ret i, at evalueringsform og læringsteori hænger sammen, er det et problem, hvis vi kun tester, når vi evaluerer. Hvis vi ser på folkeskolens formål og fagenes formål, kan det være svært at finde begrundelser for udbredt brug af behavioristisk tilgang til læring. Der lægges i høj grad vægt på, at det er eleven, der er den aktive i læringsprocessen. Følgende citat er fra formålsformuleringen for faget natur/teknologi: ”Undervisningen skal i vidt omfang bygge på elevernes egne oplevelser, erfaringer, iagttagelser, undersøgelser og eksperimenter og medvirke til, at de udvikler praktiske færdigheder, kreativitet og evne til samarbejde” (Undervisningsministeriet 2014a). Hverken de overordnede målsætninger eller de specifikke, der beskriver, hvordan man skal arbejde i fagene, egner sig til en behavioristisk dominans. Vi kan ikke umiddelbart teste de helt centrale dele af naturfagene. Hvis der ikke er en sådan sammenhæng mellem den metode, fagene evaluerer, og den målsætning, naturfagslæreren planlægger sin undervisning ud fra, ja, så er der ikke kohærens mellem målsætning, undervisning og

evalueringsform. Hvis der ikke er kohærens, kan resultatet af en sådan evaluering ikke have validitet. Vi kan ikke bruge test til at evaluere på hovedparten af naturfagenes målsætninger.

PISA, test og afgangsprøver

Det, der har givet magten tilbage til det 'tekniske paradigme', er uden tvivl de store internationale evalueringer. I forhold til naturfag er det i nogen grad TIMSS og i høj grad PISA, der har været de toneangivende. TIMSS-testen tager udgangspunkt i de målsætninger, der findes nationalt, mens PISA helt ser bort fra disse (Jensen, 2011, s. 86), men til gengæld opnåer større international sammenlignelighed. Både nationalt og internationalt har PISA på denne måde haft en kolossal indflydelse på uddannelsespolitiske beslutninger omkring evaluering af blandt andet naturfagene i folkeskolen (Bendixen, 2009, s. 65).

Det er Organisation for Economic Co-operation and Development (OECD), der har lavet PISA hvert tredje år siden 2000. PISA havde i 2012 deltagelse af 74 lande. I takt med den tiltagende internationale konkurrence har der været en tiltagende interesse for at få redskaber til at måle, i hvor høj grad de enkelte lande får udbytte af deres investeringer i uddannelse – det, der kaldes "human kapital" (Jensen, 2011, s. 84). PISA har indtaget en position, så politikerne i de enkelte lande kan måle, om økonomisk input i uddannelse giver output i form af human kapital som viden, færdigheder og kompetencer. Især fra uddannelsesforskning har der undervejs været flere former for kritik, blandt andet gående på PISA's manglende sammenhæng med undervisningen i folkeskolen (Jensen, 2011, s. 88). PISA har i takt med kritikken udviklet sig til et avanceret redskab, der nu forsøger at rette sig mod elevernes fremtid, deres kompetencer og interesser.

Disse elementer kan være meget interessante at interessere sig for, men det er et grundlæggende problem, at PISA ikke tester efter nationale mål for undervisningen, men mere på nogle antagelser om, hvad der vil være vigtigt at kunne i naturfag i en fremtidig moderne verden (Jensen, 2011, s.86). Der er mange interessante resultater i PISA, men der er grundlæggende problemer med sammenhæng og derfor også validitet – derfor skal man være varsom med konklusioner og tiltag for 'at gøre det bedre'.

Nationale test og afgangsprøver

Når vi tester i naturfagene, er det under inspiration fra PISA. Den måde, der testes på i naturfagene, er dog mere simpel og interesserer sig overvejende for det, vi kan kalde kundskabsviden. Nationale test og afgangsprøver opfylder tydeligvis ikke krav om kohærens, hverken i forhold til indholdet eller i forhold til formen. Den viden, eleverne testes for, står oftest ikke i målsætningen for faget, Forenklede Fælles mål. Igen er problemet, at en del af

de centrale evalueringsformer i naturfag ikke frembringer valide resultater, da der ikke er kohærens mellem krav til undervisning og de nationale evalueringer.

”Hvis testene ikke afspejler undervisningen, er validiteten slet ikke eksisterende” (Miller, 2007, s. 203).

Problemet med manglende kohærens og validitet kunne være et mindre problem, hvis det ikke var for evalueringers ’konstitutive effekt’ – den baglæns fungerende effekt, hvor testens form og indhold bestemmer undervisningens indhold og tilrettelæggelse. Den konstitutive effekt er her ekstra problematisk, da PISA og de nationale test ikke er af samme type. Det vil sige, at hvis læreren forsøger at planlægge undervisningen efter, at eleverne skal testes i et ukendt emne, er der alligevel ingen sammenhæng i forhold til det, der evalueres, når PISA tester.

Lande som England og USA har lange traditioner med test af elever. Erfaringer med de konstitutive effekter af test fra disse lande viser en overhængende fare for at skade naturfagsundervisningen (Nichols & Berliner, 2007) og dertil elevernes generelle trivsel (Miller, 2007, s. 203). Det bliver ’teach to the test’ og ikke elevernes læring, der får betydning. Følgende kan sammenstilles med Herbarts 170 år gamle advarsel som er omtalt tidligere.

”Når man som lærer eller skoleleder bliver målt – og måske endda belønnet eller straffet – på udfaldet af en test, inviterer det til umoralsk adfærd og defensiv undervisning. Man begynder at anskue eleverne ikke som elever, men som testscoreere ...” (Berliner 2008, s. 13).

Det kulturelle paradigme – evalueringskultur

Broadfoot mener også, at disse testformer ikke er gavnlige og ofte gør skade. Hun fremhæver, at alle omkring undervisningsinstitutionerne bør tage ansvar og tage disse negative virkninger op til overvejelse (Broadfoot, 2007, s. 163). Broadfoot taler om ”New Assessment Regime” (Broadfoot, 2007, s. 161), hvor første punkt på dagsordenen er, at læringsperspektivet står først, evaluering vælges efter, hvad der skal læres og kan styrke denne læring. Dette genkendes netop fra det, Miller kaldte ’det kulturelle paradigme’, der indebærer, at vi igen skal styrke skolernes evalueringskulturer.

Det, der er i fokus, når vi taler om evalueringskultur, er netop evaluering som et redskab i elevernes læring, dvs. at der er vægt på formativ anvendelse af evaluering. Det er en kultur, hvor eleverne ved, hvad der forventes, hvor man i undervisningen arbejder efter disse mål og evaluerer på en måde, der svarer til målsætningen og undervisningens tilrettelæggelse. Det er ikke nødvendigvis Forenklede Fælles Mål, der evalueres på, men de målsætninger, læreren – eventuelt i samarbejde med eleverne – har formuleret ud fra Fælles Mål eller folkeskolens formål. I det følgende vil vi se på forskellige evalueringsredskaber, hvor der er fokus på den formative anvendelse, men hvor alle redskaber også kan bruges summativt.

Begreber og begrebskort

En måde at målsætte på ud fra Forenklede Fælles Mål kan være at sætte mål for de vigtige faglige begreber, der skal arbejdes med i naturfagsundervisningen. Dette skal gerne give eleverne et bedre overblik og en bedre forståelse af sammenhæng mellem begreberne. Nedenstående er et eksempel fra en læseplan udviklet under det nationale udviklingsarbejde 'Naturfagene i bevægelse'; læseplanen er netop bygget op omkring begrebsmålsætning. Her er temaet 'Atomere, mad og kroppen' i en 8. klasse. Følgende målsætninger er fra fysik/kemi, tilsvarende findes under samme tema for de øvrige naturfag, biologi og geografi.

Begreber	Faglige generalisationer
<ul style="list-style-type: none">• Kilojoule• Kulhydrater• Glukose• Proteiner• Aminosyrer• Fedt• Mættede og umættede fedtsyrer• Forbrænding• Kovalent binding	<p>Man måler madenergi i joule.</p> <p>Kulhydrater og fedt består af C-, H- og O-atomer. Kulhydrater findes i mange størrelser, de korteste hedder glukose (og fruktose), de længste er stivelse og cellulose.</p> <p>Kulhydrater findes især som energi i planter.</p> <p>Proteiner er kæmpemolekyler bygget af forskellige aminosyrer. Proteiner findes overalt i kroppen og har vigtige funktioner som enzymer og i DNA.</p> <p>Fedt består oftest af tre fedtsyrer. Fedtsyrer kan indeholde enkelt- eller dobbeltbindinger og kaldes mættede eller umættede fedtsyrer.</p> <p>Molekylers bindinger kaldes kovalente og består af et 'elektronfællesskab' mellem atomer.</p> <p>Kulhydrater, fedt og protein kan forbrændes i kroppen. Ved forbrænding i kroppen forbruges ilt, og der dannes kuldioxid og vand.</p>

Figur 1: Begreber fysik/kemi 8. klasse

Til venstre er det beskrevet, hvilke faglige begreber der skal arbejdes med, til højre hvilken forståelse af begreberne eleverne gerne skal opnå. Valget af begreber er begrundet i Fælles Mål. Begreberne udleveres til eleverne ved opstart på nyt tema. De kan eventuelt tidligt i forløbet formulere deres forforståelse af begreberne til højre i skemaet og så igennem forløbet skrive ændringer til.

Evaluering af begrebsmålsætning kan foregå på en række måder, for eksempel ved hjælp af såkaldte 'begrebskort'. Begrebskort er udviklet i 1970'erne med inspiration i den læringsteoretiske retning, der tidligere i teksten er blevet kaldt konstruktivisme, her repræsenteret ved David Ausubel (Dolin, 2001, s. 192). Begrebskort blev brugt til at vise

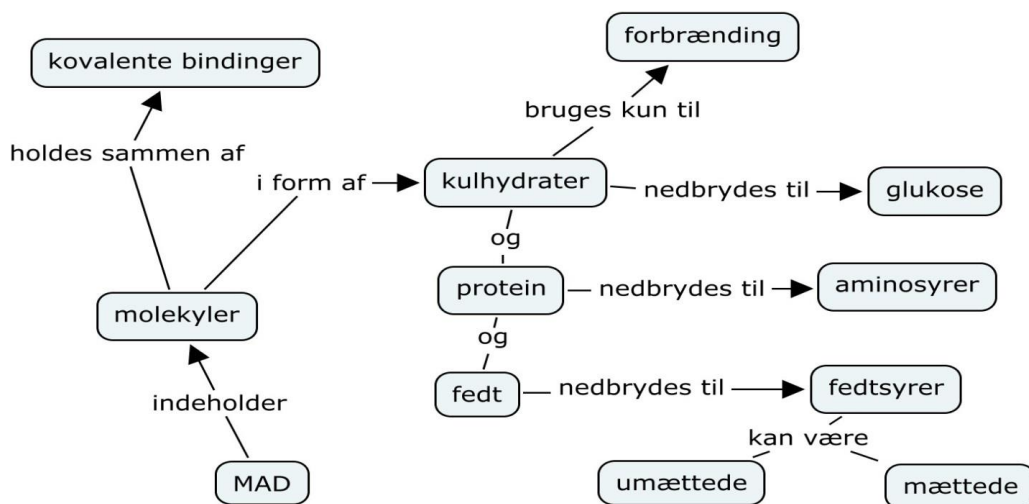
underviseren, hvad eleven allerede vidste, altså en slags billede af elevens tankemæssige struktur som udgangspunkt for læring – med andre ord en formativ anvendelse (Cañas & Novak, 2008).

”The most important single factor influencing learning is what the learner already knows” (Ausubel, 1968).

Begrebskort har altså været tænkt brugt i forhold til de enkelte elever for at kortlægge deres forforståelse, men kan også lægges sidst i et forløb og bruges summativt.

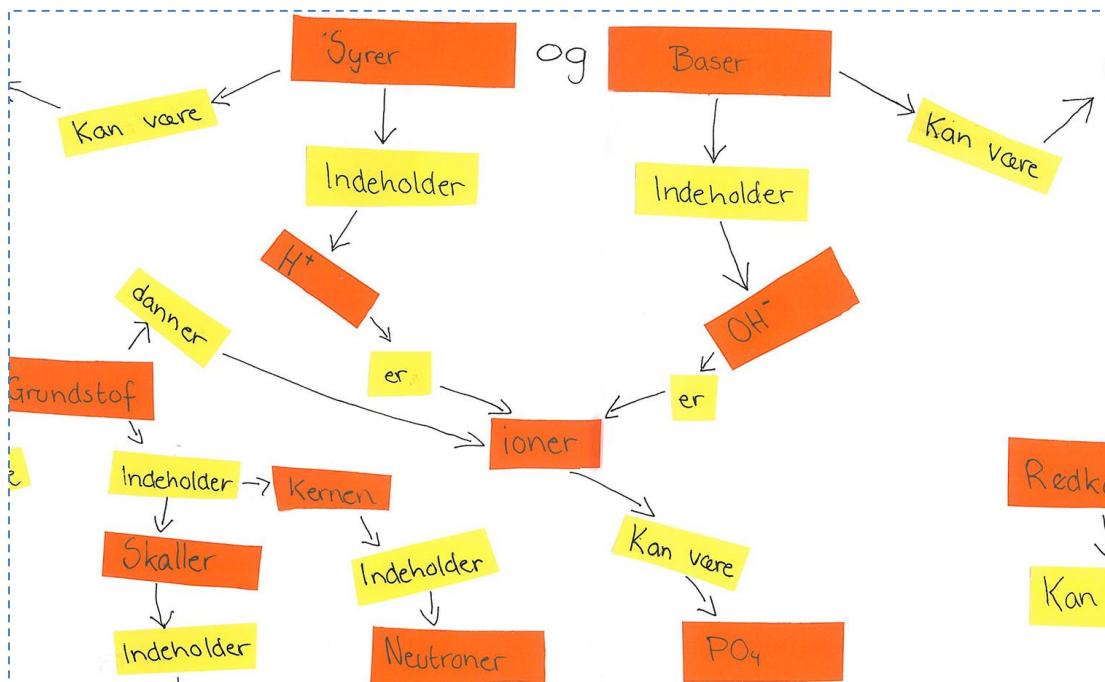
Begrebskort kan desuden bruges i grupper med fokus på mundtligheden i gruppen, her er diskussionen af begrebernes sammenhæng det mest interessante. Ved en sådan anvendelse vil der fremkomme flere forskellige erfaringer og viden i forhold til de enkelte begreber, og den enkelte elev får mulighed for mundtligt at prøve egen forståelse i forhold til de øvrige gruppemedlemmers – denne anvendelse har vi tidligere kaldt socialkonstruktivistisk.

Begrebskort forveksles ofte med brainstorm/mindmaps/tankekort, men de er væsentligt forskellige. Det, der på væsentlig måde adskiller begrebskort fra brainstorm/mindmaps/tankekort, er det, der kaldes cross-links. Cross-links skal forstås som de ord eller sætninger, der binder begreberne sammen, og det er derfor vigtigt, at de begreber, der vælges i målsætningerne, netop kan forklares ved hjælp af hinanden (Andersen, 2007, s. 35-36). Nedenfor vises et eksempel, der kunne passe til ’Atomer, mad og kroppen’:



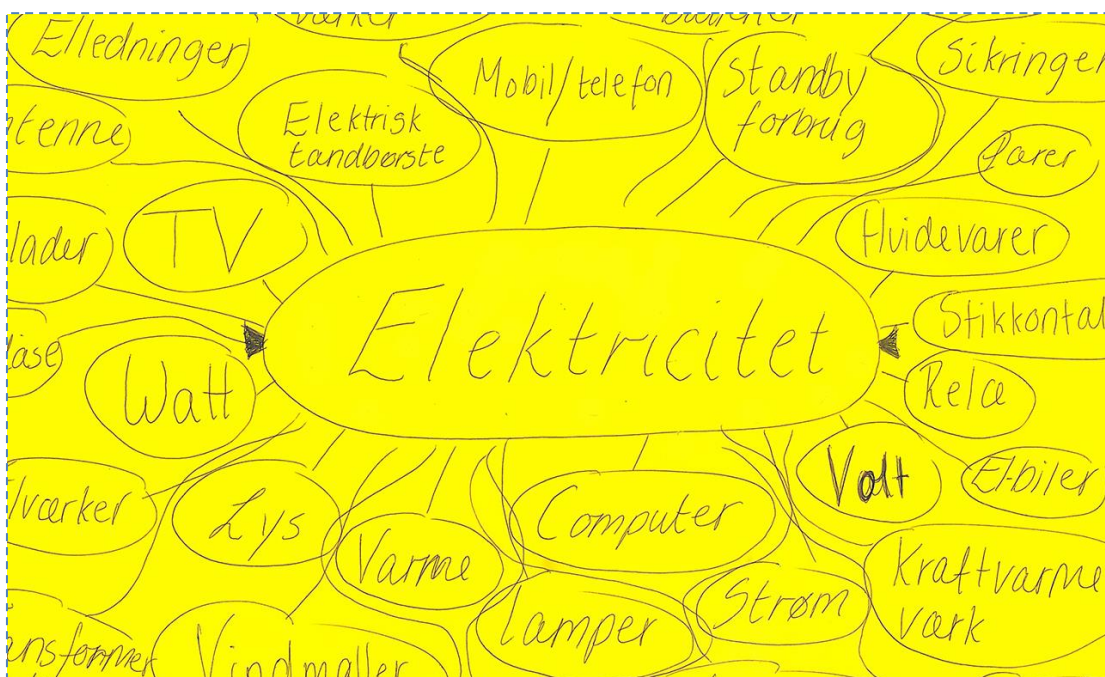
Figur 2: Begrebskort i 'cmap tools'

Begrebskortet ovenfor er lavet i det gratis program cmap tools, men det kan anbefales, at grupper arbejder med konkrete materialer som pap og papir og eventuelt efterfølgende fotograferer eller scanner resultatet. Nedenfor er der indscannet eksempler fra elevarbejder. Det første er et begrebskort fra 8. klasse i et forløb om syrer/baser:



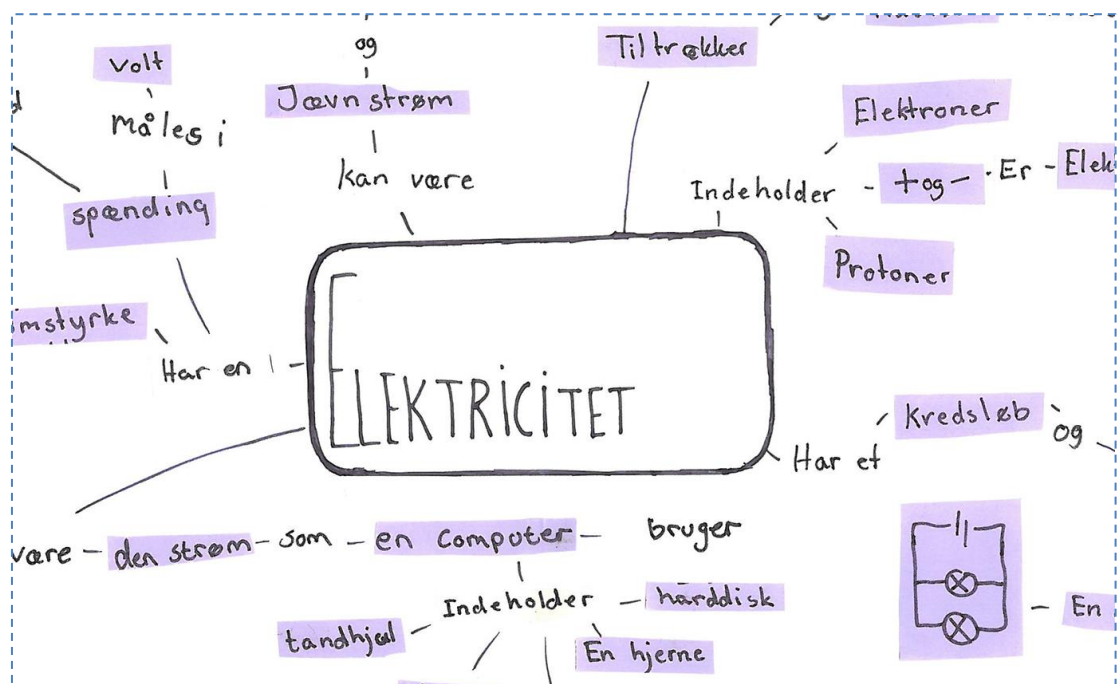
Figur 3: Udsnit af begrebskort fra 8. klasse

Der tages udgangspunkt i et centralt begreb, for eksempel 'elektricitet'. Brainstormen kan her bruges for at se, hvilke begreber eleverne i klassen forbinder med dette centrale begreb, altså elevernes forforståelse. Nedenfor ses en brainstorm lavet af en elevgruppe i opstartsfasen i 7. klasse:



Figur 4: Udsnit af 'brainstorm' fra 7. klasse

Når arbejdet med begrebskort går i gang, er det en god ide at lave brainstorm med klassen, så alle begreber fra forløbet kommer frem – læreren kan eventuelt tilføje enkelte vigtige, der ikke bliver nævnt. Herefter bygges begrebskortet op med cross-links og begreber. Nogle grupper vil inddrage flere begreber end andre, ligesom det ikke kun vil være begreberne fra skemaet, der må anvendes. Nedenfor vises et begrebskort, der er udarbejdet senere i forløbet i 7. klasse:



Figur 5: Udsnit af begrebskort fra 7. klasse

Når man anvender begrebskort som evalueringsmetode, sikrer man sig et begrebsmæssigt højt 'taksonomisk niveau' (se næste afsnit). Det er en mere udviklet forståelse end at kende begreberne hver for sig. Begrebskort anvendes især fra 5.-6. klasse og op, men med både yngre og ældre elever kan man benytte sig af samtaler om elevernes tegninger, fotos fra ture eller undersøgelser, små film, stop motion-film osv. I disse situationer kan læreren få eleverne til at finde fotos, der passer til et begreb, for eksempel 'insekt'. Når de har fundet billeder af forskellige insekter og andre smådyr, kan gruppen diskutere, hvilke billeder der hører ind under begrebet (Andersen, 2007, s. 39).

For ældre elever kan der være tale om en tegning af fordøjelsessystemet eller nitrogens kredsløb. Stop motion-film kan laves fra 3.-4. klasse eller måske tidligere. Eleverne kan lave enkeltbilledfilm ved hjælp af modellervoks, kulørt karton, fotografier m.m. og fremvise betydningen af begreber som 'fødekæde' eller vise udvikling fra frø til plante. Der findes simple programmer både hos Microsoft og Apple. I alle tilfælde vil produktet ikke være det



primære, det vigtigste i en evalueringssammenhæng er det, læreren kan iagttage og lytte sig frem til i elevernes proces.

Biggs' SOLO-taksonomi

Begrebsforståelse kan som nævnt foregå på flere niveauer. Inden for niveaudeling af læring er det i høj grad Benjamin Blooms taksonomi, der har været anvendt. Her fremhæves dog en anden, nemlig John Biggs' SOLO-taksonomi (Biggs, u.å.). SOLO er en engelsk forkortelse af: **Structure of the Observed Learning Outcome**. Blandt fordelene ved Biggs' taksonomi er, at der lægges vægt på en sammenhængende forståelse af for eksempel faglige begreber, symboliseret illustrativt i nedenstående skema (Figur 6). Det er desuden let at læse Blooms taksonomi som en trappe, hvor eleven over tid udvikler sig op ad trappen til højere niveauer. Biggs lægger derimod vægt på, at der ikke er tale om en udvikling over tid, men at en elev eller en elevgruppe kan arbejde på ét taksonomisk niveau med et emne og et andet niveau med et andet emne.

“Assessment is about how well they achieve the intended outcomes, not about how well they report back to us what we have told them” (Biggs, u.å.).

Skemaet her er en fordansket udgave, der svarer til principperne og formuleringerne i den engelske. Der er i skemaet angivet fem niveauer af forståelse (1-5), kompleksiteten af forståelsen stiger fra venstre mod højre (Biggs & Tang, 2011, s. 91).

0. niveau	1. niveau	2. niveau	3. niveau	4. niveau
Lykkedes ikke Misser pointen	Identificerer navne Følger simple procedurer	Kombinerer og beskriver Sammenligner egenskaber Sorterer	Analysere Anvender Argumenterer Sammenligner Kritiserer Forklarer årsager Relaterer Retfærdiggør	Skaber Udformer Genererer Danner hypotese Reflekterer Teoretiserer
				
Ingen relevante	Et relevant	Flere uafhængige	Integreret i en	Generaliseret til

aspekter	aspekt	relevante aspekter	struktur	et nyt domæne
----------	--------	--------------------	----------	---------------

Figur 6: Biggs SOLO taxonomi

- 0: angiver, at der ikke kan genkaldes viden om emnet
- 1: angiver, at der kun kunne genkaldes et enkelt element
- 2: angiver, at der kan genkaldes flere elementer
- 3: angiver, at disse elementer kan forstås i en sammenhæng
- 4: angiver, at denne sammenhæng kan overføres til et andet system

At kunne 'identificere navne og følge simple procedurer' er lavtaxonomisk viden, altså af en lavere kvalitet og kompleksitet end for eksempel at kunne 'analysere og relatere'. Det er eksempelvis ikke muligt at kunne 'analysere, sammenligne og relatere', hvis man ikke er i stand til at forstå strukturen i det emne, man arbejder med. Der er et afgørende kvalitetspring fra niveau 2 til 3, blandt andet det at kunne relatere begreberne til hinanden og ikke bare kende en stor mængde. I arbejdet med begrebskort sigtes der netop mod dette 3. niveau. Det sidste og højeste niveau – at kunne overføre en sammenhæng fra ét system til et andet – er en meget høj målsætning. Et eksempel på niveau 4 kan være at overføre en forståelse af en sø som økosystem til en skov som et andet økosystem – forskellige, men med samme typer funktioner som rovdyr, planteæder, planter, nedbrydere m.m. (om PISA og taksonomi, se Dolin, 2010).

En naturfagscase

En naturfagscase kan give læreren mulighed for at kunne vurdere, hvilket taksonomisk niveau eleverne befinder sig på (Biggs & Tang, 2011, s. 163). Hvis en klasse for eksempel har arbejdet med emnet 'jord' på tværs af naturfagene, kan de gruppevis få udleveret en case som:

"I skal på grund af global befolkningstilvækst undersøge mulighederne for at dyrke fødevarer i den lokale granplantage."

Eller hvis man ønsker at evaluere bredt i økologi:

"I skal anlægge en koloni på Mars, hvad tager I med? Der må kun medbringes enkeltelementer." (f.eks. bestemte fisk, men ikke et hav).

Sidstnævnte kan præciseres alt efter fokus i undervisningen, for eksempel "hvordan vil I få vand?" eller "hvad gør I med affald?". Ved at cirkulere rundt mellem grupperne og spørge ind til elevernes overvejelser og udfordre deres forståelser vil læreren kunne vurdere, hvor eleverne befinder sig i det taksonomiske hierarki. Umiddelbart passer fremgangsmåden fint til begrebsviden, det, vi også kan kalde kundskabsviden, men SOLO-taksonomien kan også bruges på andre områder.

I det første eksempel kan vi forestille os, at eleverne bliver bedt om at undersøge jorden i plantagen. Så vil det være elevernes færdigheder i jordundersøgelser, der kan vurderes ud fra. Hvis eleverne har færdighed i at måle pH-værdi, humus- og bakterieindhold, kan disse undersøgelser forstås integreret og for eksempel danne grundlag for, at gruppens medlemmer kan analysere, sammenligne og relatere ud fra de resultater, undersøgelserne fremviser, men igen er processen i dialog med læreren.

Casearbejde som formativ evaluering egner sig godt til at være netbaseret. På hjemmesiden www.viten.no vises dette bedst i casen 'Radioaktivitet'.

Feedup, feedback, feedforward

Netop lærerens dialogiske forhold til eleven er i fokus i det, der kaldes feedback. Flere evalueringsteoretikere beskæftiger sig med begrebet, men den væsentligste bidragsyder til anvendelse af feedback i undervisning er John Hattie. For Hattie er Biggs' niveauer særdeles vigtige – han taler om at gå fra overfladisk læring til dybdelæring. Det er vigtigt, at disse mål er klare og opnåelige – eleverne hjælpes med strategier til at opnå disse mål, og læreren giver feedback i processen.

"Students also need feedback that is encouraging and build their confidence, that has a positive emotional as well as intellectual impact" (Broadfoot, 2007, s. 123).

Egentlig er feedback defineret som værende information om afstanden mellem det, man kan, og det, man forventes at kunne, men kun i det tilfælde, at information kan bruges til at formindske denne afstand (Ramaprasad, 1983, s. 4-5). Hattie taler om tre feedbackspørgsmål, der netop tilsammen skal hjælpe eleven med at reducere afstanden mellem det, eleven kan aktuelt, og det mål, der ønskes opnået (Hattie & Timperley, 2013, s. 20-21; Hattie, 2013, s. 185):

- Feedup: Hvor er jeg på vej hen? Hvad er mine mål?
- Feedback: Hvordan klarer jeg mig? Hvilke fremskridt gør jeg hen imod mine mål?
- Feedforward: Hvor skal jeg hen herfra? Hvilke aktiviteter må nu iværksættes for at gøre større fremskridt?

Det er altså alle tre elementer, der tilsammen skal forstås som feedback. Når der gives feedback, er der jo tale om en form for evaluering af, hvordan det går eleven. I feedbacktanken er der også en stor del selvevaluering, hvor eleven bliver bedre til at vurdere egen læringsproces. Det er som hos Biggs især kundskabsviden, der er i centrum, men hele selvvurderingstanken gør det lettere at evaluere på andet end kundskabsviden. Et målsætnings- og selvevalueringsskema med Hattie-inspiration kunne se ud som nedenstående:

Hvad ved jeg noget om?

	Skal øve	Kan næsten	Kan
Kødædere			
Planteædere			
Ånding			
Fødekæde			
Gælle			
Fysisk gælle			
Insekter			
Fuldstændig forvandling			
Ufuldstændig forvandling			
Ilt			

Hvad kan jeg ved hjælp af apparater og redskaber?

	Skal øve	Kan næsten	Kan
Lave undersøgelser med vandhulsketcher			
Forstørre med stereolup			
Bestemme dyr ved hjælp af bestemmelsesdug			
Måle temperatur med termometer			

Hvad kan jeg uden hjælp af apparater og redskaber?

	Skal øve	Kan næsten	Kan
Stille spørgsmål om vandhullet			
Gætte på noget om vandhullet og dyrene deri			
Undersøge noget om vandhullet			
Beskrive, hvordan jeg undersøger			
Vælge det rigtige udstyr til at undersøge og eksperimentere med			
Sammenligne mine resultater med andres			
Vise og fortælle andre, hvad jeg har fundet ud af			

Læreren kan med fordel fastsætte målsætninger i samarbejde med eleven eller elevgruppen. At arbejdet netop foregår i samtale mellem læreren, eleven eller gruppen af elever skal gerne medføre, at læringssituationen opleves meningsfuld af eleverne (Dysthe, 2009). Når målsætningerne er klare, gives tilbagemeldingerne til eleven eller elevgruppen. "Hellere kort og hurtigt end sent og grundigt" (Dysthe, 2009, s. 55). "Hurtig formativ vurdering er en meget virkningsfuld form for feedback" (Hattie, 2013, s. 199).

Bemærk, at det første skema repræsenterer kundskabsviden som begreber, andet skema repræsenterer naturfaglige færdigheder som 'det, man kan ved hjælp af apparater og redskaber', mens det sidste skema indeholder elementer af kompetencer, som de findes i Forenklede Fælles Mål. Disse to sidste elementer evalueres sjældent, selv om de tilsammen udgør en stor del af naturfagernes kernefaglighed og identitet. Der er netop også tænkt i 'confidence' – elevens selvtillid i forhold til disse kategorier. Det er ikke nødvendigvis de samme elever, der er gode til kundskabsviden som til færdigheder.

Portfolio

Både Hattie og Biggs kan bruges i en portfoliotilgang til evaluering. Portfolio er en metode til netop at tænke evaluering som en integreret del af elevens læringsproces. Igen kan elementerne bruges både til enkelte elever og til læring i grupper. En portfolio er kort sagt en samling af alt det, eleven eller en gruppe af elever har udvalgt til præsentation af vigtige dele af deres læringsproces.

"En af hovedtankerne i det portfoliopædagogiske arbejde er, at eleverne skal udvikle deres evne til at reflektere over egen læring og at vurdere egne præstation i relation til bestemte kriterier og mål" (Madsen, 2004, s. 19).

I "Læseplan for faget fysik/kemi" står der: "Eleverne skal i samarbejde med andre designe, opstille og gennemføre undersøgelser" (Undervisningsministeriet 2014b, s. 4). I portfolioen vil der være udvalgt det, der ud fra elevperspektiv bedst repræsenterer, at dette mål er nået. Her kan være tale om fotos, tekst, video, tegninger, stop motion-film osv. Det vigtige er netop, at eleven har taget stilling, har reflekteret over egen proces og er central i egen vurdering (Madsen, 2004, s. 19). Hvis målsætningen er mere kundskabsorienteret, kan der eventuelt udvælges et begrebskort som en del af portfolioen. Ved at se på elevportfolioer kan læreren vurdere, i hvilket omfang målene er nået. I princippet kan en afgangsprøve bestå af en fremlæggelse af elevens naturfaglige portfolio.

"Such portfolios have the double benefit of providing evidence not only of the outcome of the student's learning, but also of the quality of the learning process itself" (Broadfoot, 2007, s. 138).

Projekt opgaver som evaluering

En, der virkelig har fået betydning for vægtningen af metoden i læringsprocessen, er John Dewey. Han var blandt dem, der omkring år 1900 vendte sig mod behavioristerne og udenadslære af kundskabsviden. John Deweys tanker om 'refleksiv tænkning' har ført til udviklingen af projektarbejdsformen som læringsmetode og evalueringsform. I dag er formen, kaldet PBL (Problem Based Learning), blandt andet gennemgribende på Aalborg og

Roskilde Universiteter (AAU og RUC). I folkeskolen kendes den især fra 'projekt opgaven' i 9. klasse, der inkluderer elementer af selvstændig evaluering. Under arbejdet er der blandt andet mulighed for, at eleven får "en vurdering af sine projektfaglige kompetencer", som der står i bekendtgørelsen (Undervisningsministeriet, 2014c, s. 1).

En projekt opgave tager udgangspunkt i en problemstilling valgt af eleverne. Eleverne undersøger problemstillingen og laver en fremlæggelse og/eller fremstiller produkter fra arbejdet med problemstillingen. Læreren vurderer af eleven foretages også på baggrund af læringsprocessen, som er beskrevet i en logbog.

I en projekt opgave er der meget høj grad af kohærens mellem metode og evaluering, idet selve fremlæggelsen af projektet og logbogen afspejler det arbejde, der er udført, og den læring, der er opnået. Selve metoden er inspireret af naturvidenskab, og derfor vil vi finde elementer fra naturfagernes kompetencemålsætninger i naturfagernes Forenklede Fælles Mål, mens det for eksempel er svært at fastsætte målsætninger omkring kundskabsviden.

Læringstilgangen 'inquiry based science education' (IBSE) kan i nogen grad afhjælpe dette, da læreren her kan afgrænse problemfeltet til at handle om for eksempel 'tryk' eller 'vands tilstandsformer'. Når en lærer således beder en elevgruppe om at finde ud af, hvorfra på skolen der fordampes mest vand, har læreren styret problemet ind i elevernes viden om vands tilstandsformer m.m., mens læreren samtidig evaluerer på elevernes evne til blandt andet at kunne danne hypoteser, tilrettelægge undersøgelser og indsamle og præsentere resultater. Hypoteserne – elevernes gæt – kan enten tegnes eller beskrives og eventuelt indsamles af læreren. Det er dog ofte svært at formulere en hypotese, så det kan også gøres i fællesskab på klassen. De forskellige hypoteser skrives på tavlen, og som eleverne tilslutter sig før og siden efter undersøgelserne (læs mere om evaluering af praktisk arbejde i Millar, 2009).

Tegn på læring

I mange af disse vurderingssituationer kan et begreb som 'tegn' være højest anvendeligt. Tegn blev først beskrevet i planlægnings- og evalueringsmodellen SMTTE (Andersen, 2011, s. 7) og beskriver det, vi ser efter, når vi evaluerer noget af det, der er svært at evaluere. Det, vi kan se efter for at kunne vurdere, om eleven er på rette vej.

"Tegn er den konkrete fremtræden af vores tilskrevne kvalitet og viser måden, hvorpå vi kunne tænke os at se kvaliteten foldet ud i praksis" (Andersen, 2011, s. 27).

Et tegn på, at eleverne er i stand til at tilrettelægge undersøgelser, kan være, at 'eleverne diskuterer variable og kun ændrer på én ad gangen'. Hvis eleverne skal undersøge vands fordampning, skal glassene indeholde lige meget vand og have samme diameter, det er kun placeringen, der skal variere.

Evaluering af naturfaglig kompetence

I Forenklede Fælles Mål fra 2014 er der i sket omfattende forandringer i målsætningerne for alle naturfagene. Der er opstillet kompetencemål som både er fælles for alle naturfagene, men også har indbygget progression fra 1. til 9. klasse. Der er opstillet fire naturfaglige kompetenceområder:

- Undersøgelse
- Modellering
- Perspektivering
- Kommunikation

Selvom kompetencebegrebet indgår nu i folkeskolens naturfag, er der ikke enighed om, hvad begrebet dækker, og de færreste har derfor en velbegrundet ide om, hvordan en kompetence kan evalueres (Andersen, 2012). Det vil føre for vidt at diskutere kompetencebegrebet her, men at det er noget, eleven kan i en given fremtid, synes at være sikkert nok. Inden for naturfagene synes der at være enighed om nedenstående definition på naturfaglig kompetence:

”Evne og vilje til handling, alene og sammen med andre, som udnytter naturfaglig undren, viden, færdigheder, strategier og metaviden til at skabe mening og autonomi og udøve medbestemmelse i de livssammenhænge, hvor det er relevant” (Dolin, Krogh & Troelsen, 2003, s. 72).

Bemærk, at denne formulering er meget dannelsespræget med vægt på ’mening, autonomi og medbestemmelse’ og ikke nødvendigvis samfundsmæssig nytteværdi! Den naturvidenskabelige kompetence består her af fire delkompetencer:

- Eksperimentel kompetence
- Repræsentationskompetence
- Modelleringskompetence
- Perspektiveringskompetence (Dolin, Krogh & Troelsen, 2003, s. 102).

Disse fire delkompetencer skal tilsammen beskrive, hvad naturfagene kan bidrage med i et menneskes liv. De fire delkompetencer minder meget om de kompetenceområder, der nu optræder i ’Forenklede Fælles Mål’. Lidt forenklet kan det siges at *modellerings-* og *repræsentations-kompetencen* slået sammen til kompetenceområdet *modellering* samt at *kommunikation* er tilføjet som et nyt kompetenceområde. De fire naturfaglige kompetenceområder: undersøgelse, modellering, perspektivering og kommunikation kaldes samlet set for *de naturfaglige kompetencer*.

Når det gælder evaluering af kompetencer, er der dog en række problemer. For det første, hvad indebærer for eksempel en eksperimentel/undersøgelses kompetence, og dernæst hvordan forholder man sig til elevens fremtidige vilje? , altså hvordan kan vi vide om en elev engang i fremtiden har vilje til at bruge sine kundskaber og færdigheder (jf. definition). Heldigvis har man i Forenklede Fælles Mål taget nogle valg og beskrevet for læreren, hvad der forstås som naturfaglige kompetencer, og når det drejer om elevens fremtidig vilje må vi nøjes med at se efter 'tegn'. Før der kan være tale om at se efter 'tegn' skal der som minimum gælde følgende:

- At eleverne eksempelvis skal udføre modeller, for at vi kan vide noget om denne kompetence.
- At eleverne ikke gentager det, de har arbejdet med, men at de kan bruge det i en ny situation til at løse en problemstilling, de ikke kendte på forhånd.

Problemstillinger som tidligere nævnte cases kan være en måde, hvor læreren ser og lytter efter 'tegn'. Det, vi skal se efter tegn på, er noget, eleven tilsyneladende kan på et generelt plan – for eksempel evnen til at foretage undersøgelser eller konstruere modeller samt evnen til at vurdere, hvordan disse bruges i en konkret situation. Kompetenceevaluering kan lægge vægt på, at eleven møder en ikke kendt problemstilling og her er i stand til at reflektere og handle ved hjælp af viden og færdigheder. Hele problemstillingen med evaluering af naturfaglig kompetence ved hjælp af tegn er grundigt beskrevet af blandt andet Elmoose (Elmoose, 2010) og forsøgt systematiseret i en matrix for evaluering (Elmoose & Sillasen, 2013).

Hvad skal læreren så gøre?

Der findes altså en mængde muligheder, når det gælder evaluering. De nævnte former virker måske besværlige og tidkrævende, men da tid til evaluering også er tid til læring, hjælper det i væsentlig grad. Det er måske mere besværligt end at teste efter hvert forløb, for det er nemt, men de skridt, man som underviser tager mod bedre formativ evaluering, må gerne være små. Begrebskort – eller bare samtaler om begreber – er en start. Bare tag fat på at lytte til elevernes samtaler om deres forestillinger, eller se på, hvad de foretager sig, når de får stillet en opgave. Formativ evaluering kan udvikles over tid, eventuelt i samarbejde med kolleger. Der vælges i fagteamet centrale begreber, færdigheder og kompetencer inspireret af eller direkte fra Fælles Mål (jf. feedbackskemaerne). Disse mål gøres klare for eleverne, så de ved, hvad meningen er, og hvad der forventes. Biggs' taksonomi og 'tegn' kan bruges i forhold til alle nævnte mundtlige og skriftlige situationer og kan alle hjælpe til indsigt i, hvilket niveau eleverne har viden og færdigheder på. Under alle omstændigheder er det helt

centralt, at læreren er bevidst om, hvad evaluering betyder og ikke betyder. Evaluering skal ikke baseres på konkurrencesituationer – man skal ikke optimere små læringsmaskiner. Det er derimod afgørende, at evalueringen har fokus på elevernes trivsel, hjælper dem til at overskue deres læringssituation og dermed hjælper dem til at lære så meget, de kan, så de bliver klar til livet efter folkeskolen. Hold øje med fagenes og folkeskolens formål. ”At lære skal tjene til at fremkalde interesse. Kundskaber forgår, men interesser og interesserethed skal vare hele livet” (Herbart, 1841/1980, s. 54).

Revideret syn på viden og evaluering for læring

På trods af at der altså er mange muligheder og på trods af de enorme mængder af forskning, der understøtter tanken om formativ evalueringsskiltur, er det stadig det ’tekniske paradigme’, der dominerer. Dominansen af teknisk evaluering i folkeskolen skyldes måske nok det enkle ved test og den internationalt skærpede konkurrence, men måske også vores syn på viden. De her gennemgåede evalueringsskilturformer bærer præg af, at læreren skal vurdere og se efter tegn på, at eleverne kan det, der var målet med undervisningen. Mange af elementerne i disse evalueringsskilturformer har altså et tydeligt subjektivt præg. Vurderingen er jo et skøn, altså ikke nødvendigvis helt eksakt dækkende for virkeligheden. Denne form for viden har i vores kultur ikke haft en særlig høj status i forhold til det, vi har kaldt ’objektiv’ faktuel viden. Dette forhold strækker sig helt tilbage til Platon, Sokrates og Aristoteles. Platon og Sokrates var interesserede i endegyldige sandheder, mens Aristoteles anså studier af enkeltsituationer for værdifulde (Flyvbjerg, 1991, s. 138). Når vi vurderer i klassen, er der tale om sådanne enkeltsituationer.

Vi har i vores platoniske kultur haft en forestilling om, at der findes eksakte svar (Gustavsson, 2001, s. 40-41), men op gennem især det seneste århundrede har den ’eksakte/objektive’ videns eksistens været under angreb. Man kan ikke måle alt, især ikke når det drejer sig om forhold mellem mennesker. Platons videnssyn kan ikke bruges, når det drejer sig om praktisk viden, som vi ser på ved færdigheder og kompetencer (Gustavsson, 2001, s. 37).

Fronten inden for viden står mellem det, vi traditionelt vil kalde viden i henholdsvis den naturvidenskabelige tradition og den samfundsvidenskabelige tradition. De to traditioners viden betegnes også som henholdsvis kvantitativ og kvalitativ viden. Når vi tester, samler vi kvantitativ viden, mens vi taler om kvalitativ viden, når vi vurderer ud fra bestemte kriterier. Inden for den naturvidenskabelige tradition har Thomas Kuhn netop forholdt sig kritisk til objektivitet selv inden for ’rigtig’ naturvidenskab (Gustavsson, 2001, s. 57), mens Gadamer og Husserl (Zahavi, 2001, s. 17; Gustavsson, 2001, s. 67) har insisteret på, at viden efter den samfundsfaglige tradition har fuldt ud lige så stor gyldighed som naturvidenskabelig viden. Samlet set gør det, at de, der arbejder med forskning, er kritiske i forhold til, hvilke

konklusioner der kan drages ud fra den viden, der er indsamlet. Det er derfor, at selv professorer fra det kvantitative felt kan kritisere PISA og de konklusioner, der drages på baggrund af den viden, der er indhentet (Kreiner & Allerup, 2014). Viden er ikke bare viden som kvantitative evalueringer. Vi kan i nogen grad nøjagtigt måle på hændelser i naturen, men når det gælder elevers læring, er det afgørende, at vi lærer at respektere kvalitativ viden – uden denne opnår vi aldrig evaluering for læring i naturfagene.

Diskussions- og refleksionsspørgsmål til gruppearbejde

1. I en gruppe på to-fire personer finder i 10-15 begreber, der skal beskrive, hvad der findes i et vandhul. Lav et begrebskort ud fra disse.
2. Udfyld feedbackskemaerne – hvilke områder kunne du tænke dig at forbedre?
3. Gennemgå en afgangsprøve i biologi eller geografi ([find dem her](#)). Hvilket af Biggs' niveauer kræver disse?
4. Hvad ville du tage med til Mars?
5. Hvilket af Biggs' taksonomiske niveauer kræver Mars-opgaven?
6. Find på en case til evaluering i naturfag.
7. Skal man ændre i den danske lovgivning, så målsætningen passer med det, PISA tester?
8. Skal man overveje at evaluere kommende folkeskolelærere som en slags 'mesterlæreprøve'?
9. Hvordan skal man evaluere ved afgangsprøverne i naturfagene?

Litteratur

- Andersen, F.B. (2011). *Tegn er noget vi bestemmer ...* Aarhus/København: VIA SYSTIME.
- Andersen, T.L. (2007). Evaluering i praksis. I T. Miller, B. Christensen, H.J. Staugaard & S. Holm-Larsen (red.) *Evaluering i fysik/kemi*. Vejle: Kroghs Forlag.
- Andersen, T. L. (2012). *Handlekompetence i naturfag*. Aalborg: Master i lærerprocesser. AAU. Masterspeciale.
- Ausubel, D.P. (1968). *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart and Winston Inc.
- Bendixen, C. (2009). *Evaluering og læring* (3. udg.). Frederikshavn: Dafolo
- Berliner, D.C. & Mortimer, P. (2008) Lær af vores fejl i stedet for at begå de samme. *Asterisk*, 41, s.12-15
- Biggs, J. (u.å.) *SOLO Taxonomy*. Lokaliseret 27. 1. 2014 på <http://www.johnbiggs.com.au/academic/solo-taxonomy/>

- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University* (4. udg.). New York: The Society for Research into Higher Education & Open University Press.
- Broadfoot, P. (2007). *An Introduction to Assessment*. New York: Continuum.
- Cañas, A. & Novak, J. (2008). *The Theory Underlying Concept Maps and How to Construct and Use Them*. Lokaliseret 25. 1. 2014 på <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>
- Dahler-Larsen, P. (1999). Den syvende evalueringsanvendelse – skolens virkelighed ændres umærkeligt. I P.F. Laursen (red.) *Vi lærer for livet – hele livet*. København: Danmarks Lærerforening.
- Dolin, J. (2001). Konstruktivismen – enhed og mangfoldighed. I *At lære fysik*. København: Undervisningsministeriet.
- Dolin, J. (2010). PISA er et farligt succesmål. *Asterisk*, 53, s.12-14
- Dolin, J., Krogh, L.B. & Troelsen, R. (2003). En kompetencebeskrivelse af naturfagene. I H. Busch, S. Horst & R. Troelsen (red.) *Inspiration til fremtidens naturfaglige uddannelser*. København: Undervisningsministeriet.
- Dysthe, O. (2009). Evaluering i klassen til støtte for læring. *KvaN – tidsskrift for læreruddannelse og skole*, 85, s.43-59
- Elmoose, S. (2010). Hvordan ser en kompetence ud? *MONA, tidsskrift for undervisere, forskere og formidlere*, 1, s.7-31
- Elmoose, S. & Sillasen, M.K. (2013). Naturfaglig kompetence og IBSE. *NorDiNa – Nordic Studies in Science Education*, 2, s.99-112
- Flyvbjerg, B. (1991). *Rationalitet og magt*. København: Akademisk Forlag.
- Gustavsson, B. (2001). *Vidensfilosofi*. Aarhus: Klim.
- Hattie, J. (2013). *Synlig læring*. Frederikshavn: Dafolo.
- Hattie, J. & Timperley, H. (2013). Styrken ved feedback. I Andreassen, R., Bjerresgaard, H., Bråten, I., Hattie, J., Hermansen, M., Hopfenbeck, T. N.,...Wille, T. S. (red.) *Feedback og vurdering for læring*. Frederikshavn: Dafolo.
- Herbart, J.F. (1841/1980). *Pædagogiske forelæsninger i omrids*. København: NNF.
- James, M. (2009). Evaluering, undervisning og læringsteorier. *Cepra-striben – tidsskrift for evaluering i praksis*, 4, s.44-57
- Jensen, T.P. (2011). PISA en makroevaluering af uddannelse. I K. Andreasen, N. Friche & A. Rasmussen (red.) *Målt og vejret*. Aalborg: Aalborg Universitetsforlag.
- Kreiner, S. & Allerup, P. (2014). PISA-test har skadet uddannelsessektoren. *Magisterbladet*, 1, s. 20-22
- Madsen, C. (2004). *Portfoliopædagogik*. Frederikshavn: Dafolo.
- Millar, R. (2009). At udvikle og evaluere praktisk arbejde i naturfag. I: S. Tougaard & L.H. Kofod (red.) *Metoder i naturfag*. Hellerup: Experimentarium.

- Miller, T. (2007). Evaluering – paradigmer og historie. I T. Miller, B. Christensen, H.J. Staugaard & S. Holm-Larsen (red.) *Pædagogisk evaluering – en grundbog*. Vejle: Kroghs Forlag.
- Nichols, S.L. & Berliner, D.C. (2007). *Collateral Damage*. Cambridge: Harvard Education Press.
- Ramaprasad, A. (1983). On the Definition of Feedback. *Behavioral Science*, 28, s.4-12
- Tyler, R.W. (1974). *Undervisningsplanlægning* (2. udg.). København: Chr. Ejlers Forlag.
- Undervisningsministeriet (2014a). *Forenklede Fælles Mål*. Lokaliseret 7. september på: <http://ffm.emu.dk/maal-struktur/naturfag/natur-teknologi>
- Undervisningsministeriet (2014b). *Læseplan for faget fysik/kemi*. Lokaliseret 7. september på: http://ffm.emu.dk/sites/default/files/unv_laeseplan_for_faget_fysik-kemi.pdf
- Undervisningsministeriet (2014c, 20. juni). *Bekendtgørelse om projektopgaven i folkeskolens 9. klasse*. Lokaliseret 7. september på: <https://www.retsinformation.dk/Forms/R0710.aspx?id=163936>
- Zahavi, D. (2001). *Husserls fænomenologi*. København: Gyldendal.