



FOTO: MIKAL SCHLOSSER

Evaluerings af elevers faglige kompetencer

Af Tomas Højgaard og Jan Sølberg



FOTO: DANIEL HIORTH/POLITIKEN/RITZAU SCANPIX

Tema: Evaluering af elevers faglige kompetencer

Pædagogisk indblik
Nr. 29
Juni 2026

Pædagogisk indblik:
Evaluering af elevers faglige kompetencer
© forfatterne, DPU og Aarhus Universitetsforlag

Litteratursøgninger:
Det Kgl. Bibliotek, Vibeke Jarloft

Design:
Toke Bjerneboe

Layout:
Ib Jensen

ISBN
978 87 7645 0229

ISSN
2596-9528

dpu.au.dk/paedagogiskindblik

03
Pædagogisk indblik

04
Evaluering af elevers faglige kompetencer

06
Overordnede kommentarer til litteraturen på feltet

07
Forskningen i evaluering af elevernes faglige kompetencer kort fortalt

09
Hvad er en god karakteristik af kompetencemål?

16
Hvordan indsamler man kvalificerede evalueringsdata?

22
Hvordan vurderer og bruger man evalueringsdata konstruktivt?

30
Om forfatterne

31
Referencer





Pædagogisk indblik

03 - 34

Forskningsoversigten, du sidder med, er en del af en serie forskningsoversigter, der hedder Pædagogisk indblik. Med Pædagogisk indblik vil vi gerne give praktikere inden for henholdsvis dagtilbud, grundskole, ungdomsuddannelse og videregående uddannelse et lettilgængeligt indblik i og overblik over den eksisterende forskning på forskellige områder – i dette tilfælde evaluering af elevers faglige kompetencer (alle fag inkluderet). Hver forskningsoversigt udarbejdes af forskere fra DPU, Aarhus Universitet, der selv forsker på det pågældende område, i dette tilfælde i samarbejde med en forskerkollega fra et andet forskningsmiljø. Forskningsoversigterne fagfællebedømmes. Det betyder, at to andre forskere, der har indgående kendskab til forskningsområdet, har læst forskningsoversigten kritisk, hvorefter den er gennemskrevet igen med henblik på at imødekomme kritikken. Formålet hermed er at sikre den videnskabelige kvalitet af forskningsoversigterne. Forskningsoversigterne bliver også læst og kommenteret af to praktikere. I dette tilfælde har Klaus Bovbjerg Sørensen og Line Ryberg-Jensen bidraget med værdifulde kommentarer. Tak for hjælpen. Også tak til Christian Højbjerg Hansen for hjælp med databasesøgningerne og til Thomas Teglgård Olsen og Kristine Ørntoft for studentervedhjælp undervejs.

Vores ambition med forskningsoversigterne er, at de kan hjælpe praktikere til at opnå den indsigt i et givet forskningsfelt, der skal til for selv at kunne tage stilling til forskningsresultaterne. Praktikere vil i denne sammenhæng sige lærere, lærerstuderende, læreruddannere og konsulenter i kommuner og ministerier, der arbejder med grundskolen og/eller ungdomsuddannelser. Derfor har vi lagt vægt på, at forskningsoversigterne ikke bare skal formidle et overblik over eksisterende forskningsresultater, men også give indblik i de grundantagelser, der ligger bag de forskellige studier. Afhængigt af temaet for forskningsoversigten kan det være, hvordan en given forståelse af læring bag en specifik undersøgelse

øver indflydelse på forskningsspørgsmål, forskningsdesign og resultater. I denne forskningsoversigt handler det fx om, hvordan en henholdsvis bred og smal forståelse af kompetencebegrebet sætter rammen for, hvad man opfatter som gode tilgange til evaluering af elevers faglige kompetencer.

Forskningsoversigterne i Pædagogisk indblik indeholder også en eller flere arbejds- og refleksionsspørgsmål og et produkt (fx en podcast eller en grafisk illustration), som praktikere kan arbejde med, fx på pædagogiske temadage. Vi håber, at Pædagogisk indblik dermed kan danne afsæt for at diskutere og videreudvikle praksis inden for de temaer, som forskningsoversigterne beskæftiger sig med. Hvis du får lyst til at læse mere, kan du lade dig inspirere af referencerne i referencelisten.

Forskningsoversigten om evaluering af kompetencer er baseret på et litteraturstudie, hvor i alt 117 publikationer blev inkluderet. Hvis du vil læse mere om fremgangsmåden og se, hvilke publikationer der er inkluderet i forskningsoversigten, kan du gå ind på dpu.au.dk/paedagogiskindblik. Her finder du bl.a. en såkaldt logbog, det vil sige et dokument, som indeholder oplysninger om undersøgelsesspørgsmål, definitioner og afgrænsninger, søgeord og søgestreng, litteraturstudiets design, kriterier for in- og eksklusion af studier og informationer om, hvordan studierne er læst.

I teksten refererer vi naturligvis til de studier, som indgik i litteraturstudiet, men vi tillader os også at inddrage yderligere tekster og pointer fra vores erfaringer uden for dette studie. Vi gør dette for bedre at kunne formidle pointerne og for at skabe et overblik, der kan benyttes af målgruppen. I teksten refererer vi til litteraturen, når vi trækker på litteraturen fra studiet, og bruger mere generelle vendinger, når vi ønsker at fremhæve pointer, som rækker ud over litteraturen.

God læselyst.



Evaluering af elevers faglige kompetencer

04 - 34

God evaluering er afgørende for læring, fordi den giver et præcist indblik i, hvor eleverne befinder sig fagligt, og hvad der skal til, for at de kan komme videre. Forskningen peger på, at systematisk evaluering og især formativ feedback har markant betydning for elevernes læring, fordi den gør deres næste skridt i læringsprocessen tydelige og mulige at hand-

le på. Denne forståelse ligger tæt op ad nyere tendenser i skoleudvikling, hvor undervisningen tilrettelægges med udgangspunkt i dokumenteret viden om elevernes læring. I professionelle læringsfællesskaber arbejder lærere netop med fælles data, elevprodukter og observationer for løbende at udvikle deres praksis, og datainformeret udvikling mere generelt forudsætter pålidelige evaluering-

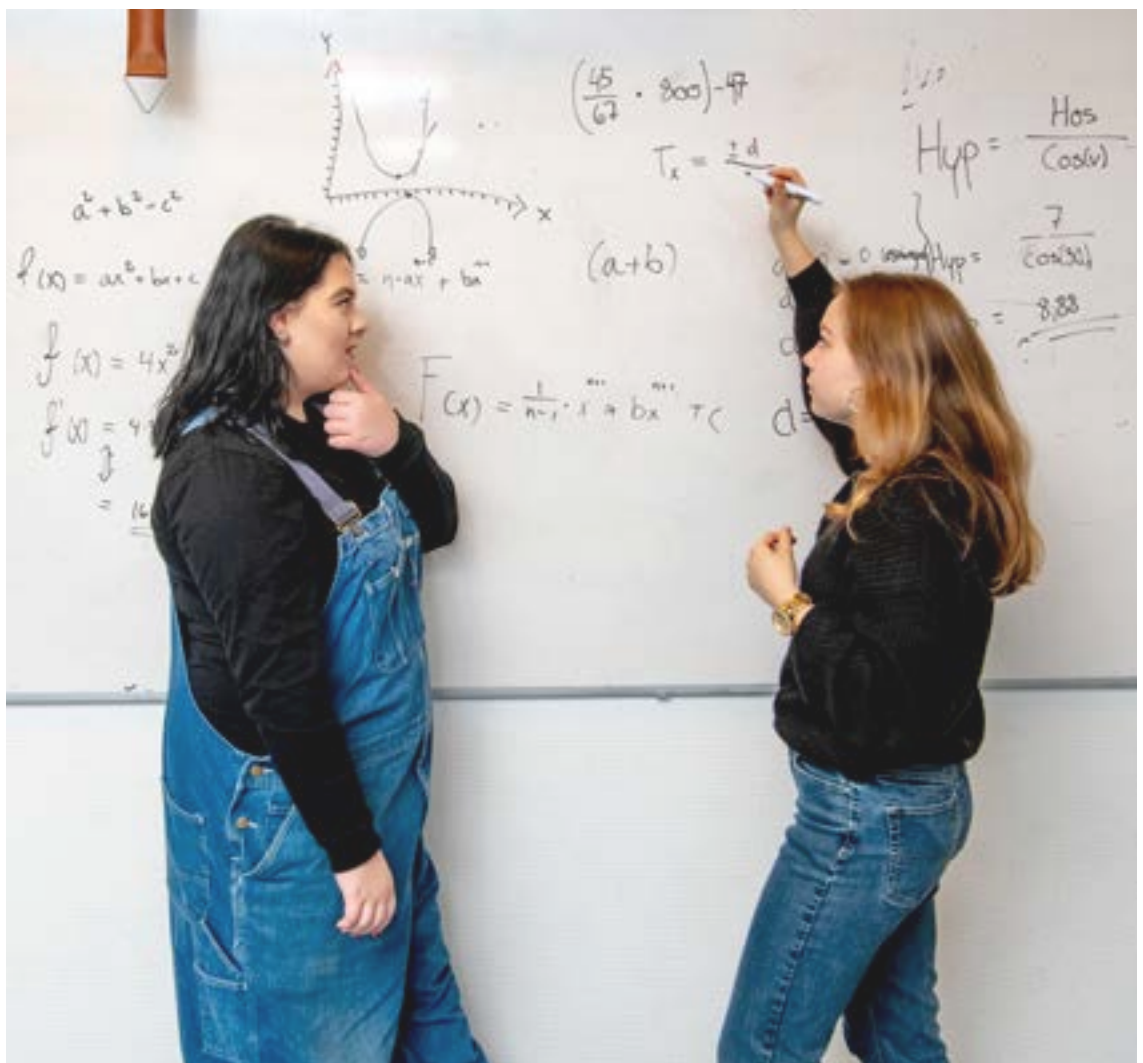


FOTO: SØREN BIDSTRUP/RTZAU SCANPIX



former. Formativ feedback bliver dermed ikke blot et redskab til elevernes umiddelbare læring, men også en central drivkraft i den kollektive og systematiske forbedring af undervisningens kvalitet.

Evaluering er et omfattende felt med mange nuancer og mange interesser. Derfor har vi valgt at afgrænse os til studier, som beskæftiger sig specifikt med evaluering af elevers faglige kompetencer i grundskolen og på ungdomsuddannelserne. Til gengæld har vi valgt at inkludere alle fag i vores litteraturstudie, da vi gerne ville afsøge mange forskellige tilgange, selvom vores fokus har været naturfag og matematik. Af samme grund har vi også inkluderet tekster, som ikke er peer-reviewed i klassisk forstand, såsom grundlagsdokumenter bag store internationale undersøgelser som PISA, da disse har haft stor indflydelse på feltet og samtidigt er baseret på grundig forskning.

Vi har valgt at fokusere på faglige kompetencer, fordi det har vist sig vanskeligt for mange lærere at arbejde med den slags læringsmål i modsætning til fx vidensmål, som lærerne er mere vant til at vurdere. En faglig kompetence kan beskrives på mange forskellige måder. Vi bruger begrebet som betegnelse for indsigtfuld parathed til at handle hensigtsmæssigt i situationer, som rummer en bestemt slags faglige udfordringer (jf. Niss & Jensen 2002). Denne definition rummer nogle centrale pointer, som adskiller kompetencer fra færdigheder og viden. Fx at kompetencer vedrører handlinger i bestemte situationer. Det er centralt for mange definitioner af kompetencer (se fx Ropohl et al. 2018), og det peger på, at eleverne

ikke blot skal vide noget, men også kunne gøre noget, som rækker ud over simple "opskriftbaserede" handlinger (færdigheder).

En konsekvens af denne begrebsforståelse er, at besiddelse af kompetence ikke kan evalueres uden at tage hensyn til de situationer, som man ønsker, at eleven skal kunne håndtere. I en uddannelseskontekst betyder det ofte, at man må skabe undervisningssituationer, som efterligner den kompleksitet, eleverne skal kunne håndtere. På den måde bliver selve undervisningen en oplagt anledning til at vurdere elevernes kompetencer (Buckley et al. 2010). Eksaminer og tests uden for undervisningen kan også bruges til at evaluere kompetencer (fx til en køreprøve). De er dog ofte mindre komplekse af praktiske og økonomiske grunde og kan derfor give et mere begrænset billede af elevernes faktiske kompetencer.

Generelt peger litteraturen på, at der i en uddannelsessituation altid bør være en tydelig sammenhæng mellem læringsmål (som, afhængigt af typen af ambition, fx kan opdeles i kompetencemål, vidensmål og færdighedsmål), undervisningsaktiviteter og evalueringsmetoder (Dolin 2020). Vi ved, at evalueringer har en kraftig indvirkning på kvaliteten af undervisning og læring – den såkaldte backwash-effekt (William 1994). Hvad enten det drejer sig om standardiserede tests og eksaminer eller evaluering gennemført af læreren som del af undervisningen, så påvirker det læreres undervisning og elevers læring. Derfor ønsker vi her at pege på nogle af de vigtigste opmærksomhedspunkter og eksempler fra vores litteraturstudie, som kan være med til at øge kvaliteten af evaluering af elevers kompetencer.



Overordnede kommentarer til litteraturen på feltet

06 - 34

Evaluering af elevers faglige kompetencer udgør et bredt og komplekst forskningsfelt præget af både teoretisk mangfoldighed og praksisnære tilgange. En grundlæggende udfordring er, at begrebet "kompetence" ikke har en entydig definition, og forståelsen varierer betydeligt på tværs af både fagområder og nationale kontekster. Forståelserne spænder fra forståelse af kompetence som relativt simple færdigheder og viden til komplekse læringsdispositioner, hvor sociale og andre former for kontekstuelle faktorer spiller ind (Carr 2001; Gerrevall 2003).

En væsentlig iagttagelse i forskningsoversigten er, at en overvejende del af litteraturen på området stammer fra lande uden for Skandinavien. Kun 34 ud af de 117 inkluderede publikationer har skandinavisk oprindelse, hvilket betyder, at mange evalueringsmodeller er udviklet i internationale evalueringskulturer, hvor standardiserede tests spiller en større rolle end i den danske. Det er derfor nødvendigt at tage kritiske forbehold, når denne viden anvendes i danske skoler, hvor evaluering i højere grad varetages af læreren.



FOTO: THOMAS VILHELM/RITZAU SCANPIX



Forskningen i evaluering af elevernes faglige kompetencer kort fortalt

07 - 34

Forskningsoversigten er opbygget i tre dele, som afspejler tre sammenhængende grundelementer i evaluering, der tilsammen omtales som 'evalueringstrekan-ten' (Shavelson 2010; Wilson & Carstensen 2007). Evalueringstrekan-ten handler om, at der skal være sammenhæng mellem læringsmål, undervisning og evaluering, hvis man vil optimere mulighederne for, at eleverne når de læringsmål, man sigter efter (Dolin et al. 2017). De tre grundelementer i evalueringstrekan-ten kan omsættes til tre relaterede spørgsmål, som er centrale for at kunne gennemføre en kvalificeret evaluering. Disse spørgsmål udgør strukturen i denne oversigt.

- Hvad er en god karakteristik af kompetencemål? (læringsmål)
- Hvordan indsamler man kvalificerede evalueringsdata? (undervisningsaktivitet)
- Hvordan vurderer og bruger man evalueringsdata konstruktivt? (evaluering)

Vi udfolder, hvad forskningen kan give os af svar på disse spørgsmål, og det udfoldes i detaljer i hver af de tre første dele af forskningsoversigten. Men kort fortalt, så er svarene på de tre spørgsmål følgende:

Hvad er en god karakteristik af kompetencemål?

En god karakteristik af kompetencemål kan tage udgangspunkt i enten en opdelt eller en holistisk tilgang. Den opdelt tilgang nedbryder kompetencen i delmål og kriterier, hvilket giver præcis feedback og mulighed for at vægte forskellige aspekter forskelligt. Holistiske tilgange fastholder kompleksiteten

og overlader en stor del af fortolkningen til læreren, hvilket kan være mere fleksibelt, men samtidig mere krævende for både lærere og elever.

Hvordan indsamler man kvalificerede evalueringsdata? Vores analyse af litteraturen peger på fem overordnede metoder til at indsamle kvalificerede evalueringsdata om elevernes faglige kompetencer:

- portfolio
- dialogbaseret
- opgavebaseret
- autentisk
- digital.

De forskellige metoder har hver deres fordele og ulemper, og valget af metode afhænger af de aktiviteter, evalueringen finder sted i forbindelse med, og de læringsmål, der sigtes efter.

Hvordan vurderer og bruger man evalueringsdata konstruktivt? Evalueringsdata kan bruges til forskellige formål. Hvad de bruges til, har stor betydning for elevernes læring og kvaliteten af undervisningen. I nogle tilfælde betyder det, at eleverne ikke lærer noget af at blive evalueret. Derfor er tydelighed omkring formålet med evaluering vigtig. Litteraturen i studiet har særligt fokus på potentialet i evalueringer til at understøtte elevernes læring. Afhængigt af formålet kan enten kvantitative eller kvalitative tilgange til vurderingerne af de indsamlede data være nyttige. Hvis formålet med evalueringerne er at fremme elevernes kompetenceudvikling, er kvalitati-



08 - 34

ve vurderinger mest effektive, men det øger kompleksiteten af evalueringerne.

Vi udfolder som sagt disse svar mere detaljeret nedenfor. Det er vigtigt at sige, at læseren kan springe direkte til de dele, som er mest oplagte for ham eller hende. Selvom

der naturligvis er en tanke bag rækkefølgen af afsnittene, så er forskningsoversigten i vid udstrækning tænkt til at kunne være en inspiration til at komme godt i gang med at udvikle sin evalueringspraksis.



FOTO: CHRISTIAN LINDGREN/RITZAU SCANPIX



Hvad er en god karakteristik af kompetencemål?

9 - 34

Vi fandt store forskelle i de inkluderede studiers forståelse af faglige kompetencer og af, hvilken grad af kompleksitet de indeholder. Det er vigtigt at være bevidst om, hvilken forståelse man benytter, da det hænger sammen med, hvordan man kan evaluere kompetencerne på en konsistent og kvalificeret måde.

Derfor indleder vi med at beskrive nogle af de forskellige forståelser af faglig kompetence, vi fandt i studierne. Carr (2001 s. 5) beskriver eksplicit forskellige grader af kompleksitet i de forskellige forståelser af det, hun kalder læringsdispositioner, hvilket i vid udstrækning svarer til vores valgte definition af kompetence. Inspireret heraf, kan man se faglige kompetencer som et kontinuum af læringsudbytte med stigende kompleksitet:

1. Færdigheder og viden
2. Færdigheder og viden + hensigt = læringsstrategier
3. Læringsstrategier + sociale partnere og praksisser + redskaber = situerede læringsstrategier
4. Situerede læringsstrategier + motivation = læringsdispositioner.

I den såkaldte KOM-blomst, som Niss & Jensen (2002) bruger til at visualisere deres bud på en kompetencebeskrivelse af matematisk faglighed, bruges beherskelse af et fags sprog og redskaber som samlebetegnelse for den ene halvdel af de foreslåede kompetencer. Hvis man sigter på evaluering af kompetence, der vedrører situationer med lav kompleksitet, kan man sætte eleverne i en situation, hvor de får mulighed for at demonstrere en sådan beherskelse. Det kan fx bestå i evaluering af matematisk hjælpemiddelkompetence, som beskrevet i

Højgaard & Niss (2023), eller i den begrebsorienterede, kognitivt baserede evalueringstilgang, som Graf et al. (2009) beskriver. Disse situationer er relativt nemme at gå til som lærer.

I evalueringsaktiviteter, der sigter på evaluering af kompetencer i situationer med høj kompleksitet, fokuserer man på fagkarakteristiske måder at handle situeret på, svarende til den anden halvdel af KOM-blomsten. Fx beskriver Stille, Jang & Wagner (2015) en model til evaluering af engelskfaglig kompetence, der både indeholder personlige elementer (som evnen til at kommunikere forskellige positioner i en tekst) og sociale og kontekstuelle elementer (som at kunne udpege relevant information og samarbejde med andre). Evaluering af sådanne handlemådekompetencer vil ofte foregå i forbindelse med mere komplekse undervisningsforløb, fx længerevarende problemorienterede projektforsløb. Den definition af kompetence, vi har valgt at lægge til grund i forskningsoversigten, læner sig op ad den mere komplekse kompetenceforståelse.

Tilgange til beskrivelse af kompetencemål

Der findes som sagt mange måder at arbejde med at beskrive kompetencer på, og i dette afsnit vil vi beskrive forskellige grundlæggende tilgange til dette. Inspireret af Pettersson et al. (2010) kan man bruge følgende inddelinger til at beskrive forskellige måder at forstå og at arbejde med kompetencer på. Det handler om, hvorvidt man tilgår kompetencer ud fra en mere holistisk forståelse eller ønsker at dekomponere kompetencer i mindre og tydeligere dele. Samtidig kan man skelne mellem kompetencer forstået som noget, der går igen på tværs af kontekster og evt. også fag eller noget, som er mere (fag-)specifikt.



	Specifik	Generel
Opdelt	Fx redskaber knyttet til konkret undervisning som Evidence-Centred Designs Hutchins et al. (2020)	Fx rubric over modelleringskompetence til brug i flere sammenhænge Christiansen et al. (2020)
Holistisk	Fx løbende helhedsvurdering af undersøgelseskompetence Østergaard (2013)	Fx lokalt udviklede skemaer til vurdering af kompetence Højgaard & Niss (2023)

Tabel 1. Oversigt over forskellige tilgange til beskrivelse af kompetencemål.

Opdelt tilgang

Langt størstedelen af litteraturen læner sig op ad en opdelt tilgang. Den har fokus på at nedbryde den enkelte kompetence i mindre enheder baseret på en grundig analyse af den.

Et eksempel på dette er de såkaldte Evidence-Centred Designs (ECD, Graf et al. 2009). ECD lægger vægt på grundig analyse af forskningslitteratur og curriculære dokumenter for at udfolde, hvad der forstås ved en given kompetence, hvorefter man udvikler evalueringsaktiviteter, der kan indfange den. Et eksempel på et ECD-design findes i Hutchins et al. (2020), der beskriver hele processen fra operationalisering af, hvordan elevers kompetence inden for STEM og computational thinking skal forstås, til design af konkrete undervisningsaktiviteter, der hver især er velegnet til, at lærere kan evaluere elevernes kompetencer inden for området. På den måde er ECD et eksempel på en meget opdelt tilgang rettet mod meget specifikke undervisningsforløb.

Et andet eksempel på en opdelt tilgang findes i Christiansen et al. (2020). Forfatterne starter som Hutchins et al. (2020) med at gennemføre en grundig afsøgning af relevant litteratur inden for modellering i naturfagene. Derefter udvikler de et detaljeret bud på delmål, kriterier og forskellige grader af målopfyldelse, som kan bruges til at evaluere udskolings elever i forskellige undervisningssituationer. Her er således tale om en opdelt, men generel tilgang, hvor evalueringen ikke er bundet op på konkrete aktiviteter, men kan bruges på tværs af flere forskellige undervisningssituationer knyttet til modellering i naturfagene.

Den oplagte fordel herved er, at man får detaljeret viden om, hvad der menes med en given kompetence. Det kan bruges både til at præcisere, hvordan en elev klarer sig, og som hjælp til at kommunikere, hvad elever forventes at kunne. Derudover giver en opdelt tilgang mulighed for at vægte forskellige aspekter af kompetencen, afhængigt af hvad der er vigtigst i en given situation. Ulempen



11 - 34

er, at det kan være både tids- og kapacitetsmæssigt krævende at opdele kompetencer i kvalificerede delementer. Desuden risikerer man, at elever og lærere fokuserer så meget på enkeltdele, at de mister forståelsen af, hvad den samlede kompetenceambition var.

Holistisk tilgang

I modsætning hertil er udgangspunktet for den holistiske tilgang at fastholde kompleksiteten af kompetencemålene. Den operationaliserer dem uden at opdele dem i mindre dele. I stedet overlades en stor del af fortolkningen af kompetencen til lærerne, og det

er op til dem at finde meningsfulde måder, hvorpå de kan hjælpe eleverne med at udvikle (og evaluere) kompetencerne uden at være bundet til forskere eller andre eksperter beskrivelse af en given kompetence. Således lægger holistiske tilgange i høj grad op til, at lærerne skal tage ansvar for evalueringen, modsat de opdelte tilgange, der typisk er defineret af andre end lærere.

Et eksempel på en specifik holistisk tilgang findes hos Østergaard (2013). Forfatteren argumenterer her for, at evaluering af undersøgelseskompetence i en konkret situation ikke kan forstås uden for den sociokulturelle kontekst. Som en konsekvens heraf kan



FOTO: VOLUROLICOLIBBOX.COM



undersøgelseskompetencen ikke defineres entydigt, men skal fortolkes ud fra den konkrete situation. Det kan ske ved at indsamle data fra elevernes mundtlige, skriftlige og praktiske arbejde, så læreren (gerne sammen med andre) kan foretage en helhedsvurdering af, hvor godt den enkelte elev arbejder i den konkrete undervisningssituation.

Et eksempel på en generel holistisk tilgang ses hos Højgaard & Niss (2023), som beskriver udarbejdelsen af modeller af matematiske kompetencer. I deres eksempel udfoldede en matematiklærer sammen med en forsker hjælpemiddelkompetence i et nemt, overskueligt skema (jf. Højgaard 2017). Skemaets fire felter indikerer, om der er tale om mere eller mindre avanceret brug af hjælpemidler, og om der er et eller flere hjælpemidler i spil på samme tid. I hvert felt står et verbum, som beskriver, hvad eleven skal kunne i forhold til hjælpemidler: anvende, vurdere, skifte og vælge. Denne model viste sig at være tilstrækkelig generel til, at læreren kunne bruge den på tværs af årgange og undervisningskontekster, hvilket er en af styrkerne ved den holistiske tilgang.

Fordelene ved denne tilgang er, dels at den eksplicit fastholder fokus på den læringsmæssige ambition i sin helhed, dels at den muliggør en forholdsvis hurtig bedømmelse af elevers arbejde, fordi man ikke behøver redegøre for alle de forskellige elementer, som indgår heri. Ulempen er, at hverken lærere eller elever opnår entydig viden om, hvad der mere specifikt er styrker og svagheder ved elevernes præstationer, da holistisk evaluering i sig selv ikke kan være så detaljeret, som tydeligt opdelt læringsmål tillader. Det kan gøre det vanskeligt at følge

op på evalueringen. Denne tilgang fordrer således god dømmekraft og evalueringsfaglighed hos den enkelte lærer, hvilket kan være en udfordring for lærerne og derigennem mindske kvaliteten af evalueringerne (Andersen 2014).

Operationalisering af kompetencemål

Uanset om man vælger en opdelt eller holistisk tilgang, og uanset om man ønsker at evaluere en specifik situation eller ej, så er det fundamentalt, at man operationaliserer det, man vil evaluere, til noget, som kan indfanges relativt entydigt. En sådan operationalisering følger typisk en række trin (Dolin 2016):

1. Beskriv de vigtigste delmål, der er relevante for kompetencen, herunder også kriterierne for hvert delmål, som angiver, hvilke former for elevhandling der tages som udtryk for, at delmålet er nået.
2. Beskriv, hvordan forskellige grader af målopfyldelse ser ud for hvert delmål, så man kan skelne mellem forskellige præstationer og vise, hvad der skal til for at leve op til kriterierne for delmålet (en progressionsbeskrivelse).

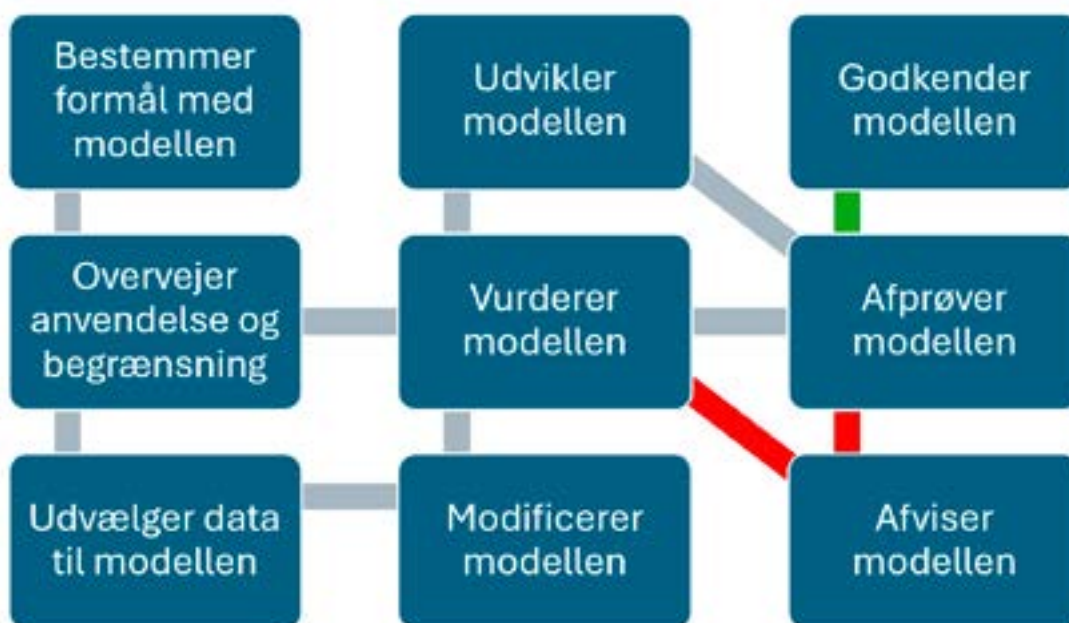
Det er ofte op til lærerne selv at operationalisere kompetencer, så de kan arbejde konkret med dem i undervisningen og evaluere dem i forhold til kriterierne. Det kan være en stor opgave (Andersen 2014), og i det følgende viser vi nogle eksempler på, hvordan sådan en operationalisering kan se ud, når det gøres grundigt.



Delmål og kriterier

Christiansen et al. (2020) har lavet et bud på, hvordan man kan operationalisere modelleringskompetence i naturfag i udkolingen (se boks 1). Først analyserede de curriculære dokumenter og relevant forskningslitteratur for at sammenfatte, hvad der var skrevet om modelleringskompetence. Resultatet af deres analyse førte til en beskrivelse af ni delprocesser, der tilsammen giver et bud på, hvad det vil sige at modellere inden for naturfagene. Hver delproces udgør på den måde et delmål for

modelleringskompetencen. De formulerede delmålene som observerbare handlinger, eleverne skulle kunne udføre som tegn på, at de opfyldte målene (kriterier). De ni delprocesser blev også illustreret i en model, så man kunne se, hvordan de hænger sammen og tilsammen udgør kompetencen. Modellen viste sig at kunne bruges til at planlægge undervisningen med fokus på enkelte delmål et ad gangen, samtidig med at eleverne forstod, hvordan de enkelte dele indgik i den større helhed. På den måde blev kompetencemålet håndgribeligt for både lærer og elev.



Boks 1. Forenklet oversigt over modelleringskompetence i naturfagene (Frit efter Christiansen et al., 2020: 50). Boksene indikerer ni forskellige delmål. Hver boks i oversigten indikerer en vigtig del af processen, som læreren kan forsøge at indfange i sin evaluering. Stregerne mellem boksene viser forskellige veje i processen fra idé til færdig model. De røde streger betyder, at processen skal starte forfra, og den grønne streg indikerer, at eleverne er nået i mål med at udvikle deres model.



Grader af målopfyldelse (progressionsbeskrivelser)

Det er vigtigt at kunne skelne mellem forskellige grader af målopfyldelse, så både lærere og elever kan få en klar idé om, hvad der forventes af dem, og hvad de skal gøre for at udvikle sig i retning af målene. Det kan også være nødvendigt at kunne skelne i situationer, hvor der skal foretages en vurdering af elevernes evne på et givet tidspunkt (fx ved en eksamen). Progressionsbeskrivelser kan hjælpe med dette. De er systematiske beskrivelser af, hvordan forskellige grader af målopfyldelse ser ud. Lærere kan bl.a. bruge dem til at give feedback og vise eleven, hvad næste skridt er i forhold til at opfylde målet.

Det kan dog være svært at beskrive forskellige grader af målopfyldelse inden for alle tænkelige kompetencer. Det kan også være meget tids- og ressourcekrævende at udarbejde progressionsbeskrivelser. Fx brugte Liu et al. (2013) data fra lærere, fagdidaktiske eksperter, klasserumseksperimenter og psykometriske analyser i deres arbejde med at udarbejde progressionsbeskrivelser for tre mindre områder af naturfaglig kompetence. Grundige beskrivelser af læringsprogression er således ofte meget specifikke i den forstand, at de kun omhandler en lille del af en given faglig kompetence, hvilket kan være en udfordring, hvis man vil dække alle aspekter af den.

Som alternativ til at beskrive meget specifikke faglige mål i en progressionsbeskrivelse kan man finde mere generiske tilgange, som ikke er bundet til specifikke emner, processer eller fag (Dolin 2016). Disse er typisk baseret

på taksonomier, såsom Blooms taksonomi eller SOLO-taksonomien, der har mere fokus på elevernes evne til at arbejde på et bestemt kognitivt niveau end på mestring af et konkret fagligt emne. Et eksempel på denne tilgang ses hos Rönnebeck et al. (2017), som forklarer, at deres "metadefinition" af undersøgelsesbaseret undervisning er så generel, at den kan anvendes i såvel naturfagene og matematik som inden for teknologi og innovation.

Rubrics

Et meget udbredt værktøj til at visualisere både delmål, kriterier og grad af målopfyldelse er såkaldte rubrics. De kan se meget forskellige ud, men består grundlæggende af en liste over udvalgte delmål, der omhandler centrale elementer af det, man ønsker at evaluere. Ud fra hvert delmål opstilles så en række beskrivelser af, hvordan forskellige grader af målopfyldelse ser ud. På den måde opnås en todimensionel oversigt over en given kompetence, som kan bruges til at kommunikere, hvordan man definerer kompetencen, og hvordan man forestiller sig, at den kan komme til udtryk hos de elever, man ønsker at evaluere.

I forlængelse af eksemplet i boks 1 udfolder Christiansen et al. (2020) i deres artikel et eksempel på udarbejdelse af en rubric. Den består af et skema, som detaljeret beskriver mulige elevpræstationer på fem forskellige niveauer inden for hver kompetencemæssig delproces. Fx blev delmålet udvalgt data til modellen udfoldet i følgende fem niveauer:



	Niveau 1	Niveau 4	Niveau 3	Niveau 4	Niveau 5
Delmål: udvælger data til modellen	Udvælger irrelevante data	Udvælger mangelfulde eller ikke-relaterbare data	Udvælger og kombinerer forskellige data	Udvælger, kombinerer og relaterer forskellige data	Udvælger ud fra den fremsatte hypotese data og relaterer til eksisterende teorier

Tabel 2. Mulige elevpræstationer på fem niveauer inden for hver kompetencemæssig delproces.

Således kan rubrics give et overblik over delmål og grader af målopfyldelse, hvilket kan hjælpe lærerne med at danne sig et overblik over, hvordan en given kompetence kan forstås. Et sådant overblik kan være en hjælp til at planlægge undervisning såvel som til at kommunikere kompetencemålene til eleverne. Samtidig er rubrics et redskab til at vurdere konkrete elevpræstationer, hvilket vi udfolder senere.

Refleksionsspørgsmål

- ✓ Kan du redegøre for de vigtigste delmål inden for de forskellige kompetencer, du arbejder med? Kan du beskrive den enkelte kompetence på en måde, så du, dine kolleger og dine elever kan genkende kompetencen i elevernes arbejde? Hvordan adskiller du kompetence fra andre former for læringsmål som fx viden eller færdigheder?
- ✓ Ved dine elever, hvad du betragter som de vigtigste kompetencer, de skal udvikle i din undervisning? Ved de også, hvad det vil sige at være kompetent, og hvad de skal gøre for at blive det? Hvornår og hvordan bliver eleverne klar over, hvad de forventes at kunne på forskellige trin?



Hvordan indsamler man kvalificerede evalueringsdata?

16 - 34

Nu vender vi os mod, hvordan man kan indsamle data til en evalueringsproces med kompetencemål som læringsmæssigt omdrejningspunkt. Tabel 3 er en oversigt over forskellige tilgange til dataindsamling, som vi har valgt at kategorisere litteraturens beskrivelser efter, og som vi efterfølgende beskriver lidt nærmere. Fælles for disse tilgange er, at de foregår enten i forbindelse med eller lige efter undervisningen (og ikke helt afkoblet fra undervisningen, som fx PISA-tests), hvilket er et fælles træk ved de fleste evalueringstilgange i vores forskningsoversigt.

Portfolio

En portfolio er en samling af materiale, som en eller flere elever har udvalgt til at repræsentere deres læringsproces (Andersen 2014; Rönnebeck et al. 2017). Det kan være opgavebesvarelser, refleksioner, videoklip, rapporter etc. (Dolin 2020). Det kan også gennemføres, ved at eleverne i slutningen af en lektion skriver deres reaktion på undervisningen ned eller i slutningen af et forløb skriver et lille essay om deres udbytte af undervisningen (Dolin 2020). Portfolioler kan dermed give eleverne mulighed for

Tilgang	Kort karakteristik
Portfolio	En samling af det, som en eller flere elever har udvalgt til at repræsentere deres læringsproces: opgavebesvarelser, refleksioner, videoklip, rapporter etc.
Dialogbaseret	Gennem dialog mellem lærer og elever og/eller mellem eleverne gøres elevernes tanker eksplicite, så de foreligger som datagrundlag for evaluering.
Opgavebaseret	Indsamlingen af evalueringsdata tager udgangspunkt i en skriftligt formuleret opgave, som eleverne skal arbejde med.
Autentisk	Indsamlingen af evalueringsdata foregår, mens eleverne gennemfører aktiviteter, som formodes at udfordre den eller de kompetencer, som ønskes evalueret.
Digital	Indsamlingen af evalueringsdata om – og nogle gange også vurdering af – elevers kompetencer foregår ved hjælp af digitale værktøjer.

Tabel 3. Oversigt over forskellige tilgange til dataindsamling i forbindelse med evaluering af elevers faglige kompetencebesiddelse.



at samle deres refleksioner over egen læring i et format, der gør det muligt at bygge videre på deres erfaringer og vende tilbage til dem. De kan også bruges af lærere som grundlag for samtaler med forældre.

Calfee et al. (2014) betoner, at der kan og bør indsamles datainput til evalueringen både i nuet og i tilbageblik på længerevarende undervisningsforløb, men at disse data selvsagt er ret forskellige og bidrager på forskellig vis til portfolioen. Det første kan fx foregå, når man observerer elevernes arbejde og går i dialog med dem, mens det andet fx kan foregå, når man læser en rapport over et længere projektarbejde.

Fordele: Andersen (2014) beskriver, at portfolio understøtter elevens udvikling af evnen til selvevaluering og refleksion over egen læring i forhold til målkriterierne. Som et konkret eksempel herpå beskriver Gorlewski (2010) en proces, hvor nogle elever selvevaluerer ved at gennemgå deres egen portfolio ud fra spørgsmål, der er defineret af læreren og relateret til et bestemt kompetencemål. Portfolio kan også understøtte samarbejde elever imellem, hvis de fx kan kommentere hinandens bidrag (Rönnebeck et al. 2017).

Udfordringer: Portfolioer risikerer at blive et produkt i sig selv, der indeholder ritualiserede, forskønnede dokumentationer frem for refleksioner over de handlinger, der faktisk fandt sted (Rönnebeck et al. 2017). Desuden kræver portfolioer disciplin; det tager tid at samle materiale og at gennemse det individuelt. Digitale platforme kan hjælpe med organiseringen, men læreren skal stadig have tid til at give feedback på indholdet. I praksis kan dette afhjælpes ved at fokusere feedback

på udvalgte dele eller lade eleverne præsentere uddrag af deres portfolio i grupper.

Dialogbaseret evaluering

Dialogbaseret evaluering er karakteriseret ved, at man gennem dialog mellem lærer og elever og/eller mellem eleverne gør elevernes tænkning eksplicit. Dialogen bliver dermed til evalueringsdata, som enten kan bruges til at give direkte feedback eller lagres til senere brug (fx skriftligt, på video eller med lydoptagelse).

Dialogbaseret evaluering kan være meget ritualiseret som i de såkaldte strukturerede evalueringsdialoger (Dolin et al. 2018), hvor klassen samles om en udvalgt elev, som har en kortere samtale (ca. syv minutter) med læreren om et fagligt emne. De øvrige elever giver peer feedback og foretager en selvevaluering på baggrund af dialogen, så alle får noget ud af den.

En anden variant er elevdebatter (McIntosh & Milam 2016), hvor eleverne debatterer et centralt fagligt spørgsmål og derefter revurderer det svar, de indledningsvis er kommet med. Læreren kan så se, hvordan debatten påvirkede elevernes forståelse af de faglige nøglebegreber. Evalueringsdata består her af de forskellige svar på det centrale faglige spørgsmål før og efter debatten samt selve argumenterne under debatten. Ifølge German (2017) kan man med denne tilgang både evaluere den enkelte elevs forståelse af faglige nøglebegreber, faglige praksisser og tværfaglige begreber. Elevbesvarelserne illustrerer deres tænkning, og hvordan den har forandret sig i løbet af diskussionen.



Fordele: Overordnet kan vi på tværs af eksemplerne i litteraturen se, at dialogbaseret evaluering i sin simpleste form er nem at iscenesætte, og den forekommer naturligt i megen undervisning. Det kræver dog en vis struktur for at sikre, at dialogen fører til kvalificeret evaluering (Furtak et al. 2016). Gennem dialog hører man elevens forklaringer, overvejelser og misforståelser, mens de kommer til udtryk. Dette skaber mulighed for umiddelbar feedback, og læreren kan justere sin undervisning eller stille opfølgende spørgsmål for at fremme elevens læring (Black 2016).

Udfordringer: Det kan være en udfordring for læreren at orchestrere dialogen sådan, at

den faktisk er udtryk for elevens forståelse af det faglige (Dolin et al. 2018). Det kræver en balance mellem autoritativ og dialogisk diskurs (Furtak et al. 2016), hvor læreren ikke stiller spørgsmål, der er så åbne, at vigtige faglige pointer bliver uklare. Omvendt påpeger forfatterne, at spørgsmål, der er for lukkede (autoritative), ikke giver indsigt i elevens tænkning. Furtak et al. (2016) påpeger desuden, at det kræver en tryghed i klassen, hvis eleverne skal turde "tænke højt" og turde begå fejl foran hinanden. Det kan tage tid at etablere en sådan kultur. Kort sagt er dialogbaseret evaluering et stærkt redskab til formativ evaluering, men det kræver øvelse, struktur og en støttende atmosfære i klassen.





Opgavebaseret evaluering

Opgavebaseret evaluering er karakteriseret ved, at indsamlingen af evalueringsdata tager udgangspunkt i skriftligt arbejde. Det kan være skriftlige opgaver, cases, opgavesæt m.m. Man kan fx designe opgavesæt med stigende sværhedsgrad, så man kan vurdere elevernes kompetencer ud fra, hvor langt de kan komme i opgavesættet (Burkhardt 2018).

En anden variant er at designe opgaver, der undervejs rummer en række relevante deludfordringer i forhold til den kompetence, som udgør hovedudfordringen. På den vis kan man arbejde med både dele og helheder undervejs i evalueringen og både på kollektivt og individuelt niveau (Deane & Sparks 2019). Ofte kombineres dialog- og opgavebaseret evaluering, så elevernes arbejde fastholdes i det skriftlige, samtidig med at man får adgang til elevernes dialogbaserede refleksioner undervejs (Martin et al. 2017).

Fordele: Formuleringen af skriftlige opgaver i disse tilgange giver evalueringen en konkret genstand, som både elev og lærer kan forholde sig til, og som gør det muligt at holde sammen på og fokusere evalueringsprocessen. En veludformet opgave kan fokusere evalueringen på de kriterier, man ønsker at vurdere, og gøre det lettere at sammenligne elevpræstationer (fx Burkhardt 2018). Opgaver anses generelt for at være tidsbesparende i den forstand, at man kan evaluere en hel klasse relativt effektivt, særligt hvis man efterfølgende vurderer dem eller bruger objektive scoringsmetoder. Dette gælder dog i mindre grad åbne opgaver, som kan kræve en del arbejde at vurdere.

Udfordringer: Deane & Sparks (2019) peger på to udfordringer ved design af evalueringsopgaver, som forstærkes, når der er tale om evaluering af kompetencer: Den ene er, at lærere som regel hverken har tilstrækkelige kvalifikationer eller tid til at designe opgaver, som kan indfange kompetencer. Den anden er, at hvis opgaven består af ét slutprodukt, fx en tekst, så er der en risiko for, at evalueringen fokuserer alene på produktet og overser betydningen af processen bag. Slutproduktet kan sjældent bruges til at afdække delmålene i kompetencen.

Autentisk evaluering

Autentisk evaluering er karakteriseret ved, at indsamlingen af evalueringsdata foregår, mens eleverne gennemfører aktiviteter, som udfordrer en given kompetence – og derved giver mulighed for at demonstrere den (Spektor-Levy et al. 2009; Anker-Hansen & Andrée 2015). Det skal ses i modsætning til kun at basere evalueringen på de skriftlige produkter, mundtlige dialoger mv., som genereres i forlængelse af (og eventuelt før) kompetenceudøvelsen. Autentisk evaluering er gerne knyttet til praktisk arbejde såsom laboratoriarbejde, feltarbejde, fysiske øvelser, håndtering af instrumenter mv., hvor eleverne skal løse specifikke opgaver i interaktion med den fysiske verden. Ofte er der tale om aktiviteter, som involverer flere trin, som hver især kan gøres til genstand for evaluering (Kruit et al. 2020). Frejd (2013) fremhævede tilsvarende i sit review af evaluering af matematisk modellering, at projektarbejde var særligt velegnet til at indfange elevernes arbejde med autentiske problemer, lige fra identificering af problemet til løsning af det.



Fordele: Autentisk evaluering giver lærere mulighed for at få et nuanceret indblik i forskellige aspekter af elevens kompetencer, som ikke indfanges ved mere traditionelle evalueringer (Harrison 2014). Det giver bl.a. mulighed for at observere elevers diskussioner undervejs i arbejdsprocessen (Anker-Hansen og Andréé 2015) og at opdage, hvilke trin i processen eleverne har svært ved, så man kan udvikle undervisningen og fremme elevernes læring (Kruit et al. 2020). Eleverne oplever ofte autentisk evaluering som meningsfuld og engagerende, fordi der som regel findes flere mulige løsninger på det pågældende problem, og det inviterer eleverne til at tage ejerskab over deres egne løsninger (Spektor-Levy et al. 2009).

Udfordringer: En af de største udfordringer er, at læreren kan have svært ved at indfange alle de relevante handlinger hos eleverne, når

de arbejder med forskellige dele af arbejdsprocessen eller i grupper med forskellige fokusområder (Anker-Hansen & Andréé 2015). Det kan føre til, at evalueringen af elevernes kompetencer bliver vilkårlig og fx baserer sig på det mest iøjnefaldende. En anden central udfordring er, at det kan være tidskrævende at udvikle og implementere autentisk evaluering (Kuo et al. 2015). Det kan også kræve særligt udstyr eller omgivelser, såsom laboratorier, at give eleverne mulighed for at demonstrere deres kompetencer i en autentisk faglig kontekst (Abrahams et al. 2013; Erduran et al. 2020).

Digital evaluering

Digital evaluering refererer til brugen af digitale værktøjer til at indsamle data om og vurdere elevers kompetencer. Dette omfatter bl.a.



FOTO: MIKAL SCHLOSSER



brug af visualiseringer, simuleringer, lydfiler eller andet, som eleverne skal forholde sig til (Ketelhut et al. 2013). Digital evaluering giver mulighed for nemt at indsamle evalueringsdata på individ- og/eller klasseniveau eller i endnu større målestok (fx OECD 2003). Fx kan eleverne ved hjælp af computerprogrammer og simuleringer i et virtuelt rum afprøve hypoteser ved at udforske, hvilke konsekvenser deres handlinger har. Der indsamles data, mens eleverne arbejder, som kan bruges til at give eleverne feedback, men som også kan give lærerne en oversigt over elevernes udvikling (Kuo et al. 2015). Andre programmer er baseret på scenarier, hvor eleverne skal håndtere komplekse problemstillinger (Liu et al. 2013).

Fordele: En af de mest markante fordele er muligheden for at indsamle data nemt og hurtigt og give eleverne (automatiseret) feedback, hvilket kan styrke både motivation og læringsprocesser (Liu et al. 2013; Ketelhut et al. 2013). En anden fordel er, at det er muligt at sætte eleverne i situationer, der ellers ville være vanskelige, farlige eller dyre at opleve, såsom et vulkanudbrud. Digital evaluering kan således spare både tid og ressourcer og kan potentielt skabe helt unikke læringsmuligheder (Buckley et al. 2010).

Udfordringer: Det kan være svært at sikre en troværdig måling af elevers kompetencer gennem digitale evalueringer (Dolin & Krogh 2010). En central overvejelse er, i hvilken udstrækning en virtuel simulering eller problemstilling reelt udgør en autentisk situation, hvor eleverne kan demonstrere deres kompetencer til at håndtere virkelige udfordringer, hvor bl.a. sociale forhold og den fysiske virkelighed altid spiller ind. Digitale platforme kan også være statiske eller kon-

tekstløse og dermed hurtigt miste deres kompetencemæssige relevans. Dertil kan der være tekniske udfordringer forbundet med brugen af digitale evalueringer (Deane & Sparks 2019).

Som en sidste bemærkning vil vi kort nævne, at vi ikke i litteraturen fandt væsentlig omtale af AI (kunstig intelligens) og dens konsekvenser for evaluering af og beskrivelse af kompetencer. Dette er et område, som er eksploderet og forsat udvikler sig hastigt, og forskningen i evaluering og AI er ikke moden på dette tidspunkt. Derfor omtales AI ikke yderligere her, men i praksis er det en væsentlig faktor allerede nu.

Refleksionsspørgsmål

- ✓ Hvordan indsamler du data om dine elevers kompetencer? Indsamler du data om de samme kompetencer over tid? Hvilken rolle spiller dialogen med eleverne i vurderingen af deres kompetencer?
- ✓ I hvilke undervisningssituationer oplever du, at det er lettest eller mest naturligt at vurdere elevernes kompetencer? Hvordan får eleverne mulighed for at vise deres kompetencer i din undervisning? Har du mulighed for at indsamle systematiske data om elevernes kompetenceudvikling over tid? Har du et klart billede af, hvordan progressionen af den enkelte kompetence ser ud over tid?



Hvordan vurderer og bruger man evalueringsdata konstruktivt?

22 - 34

Efter diskussionen af, hvordan man kan karakterisere kompetencemål, og hvordan man kan indsamle evalueringsdata, er vi nu nået til, hvordan man kan bruge de indsamlede data til vurdering af elevlæring. Her er det en vigtig pointe, at evalueringsdata kan bruges til forskellige formål, og afhængigt af formålet er der forskellige måder at komme frem til en endelig vurdering på, hvilket vi vil udfolde i det følgende.

Fire typer formål med evaluering

Man skelner ofte mellem summative og formative evalueringer (Dolin et al. 2017). Formative evalueringer handler om at hjælpe eleven med at lære, mens summative evalueringer handler om at vurdere, hvad eleven har lært. I litteraturstudiet fandt vi, at særligt nyere studier i stedet for disse to grundtyper af evaluering snarere forholder sig til fire forskellige typer af formål med evalueringen: tre formål rettet mod eleverne og et fjerde formål rettet mod udvikling af undervisningen, jf. tabel 4 og udfoldningen herunder.

Med **evaluering for læring** evalueres den enkelte elevs kompetence med det formål, at eleven får klare, konstruktive tilbagemeldinger, som vedkommende kan handle på (Jönsson et al. 2013). Dette kræver, at læreren præcist og effektivt kan bestemme elevernes kompetenceniveau på baggrund af deres præstationer og ud fra det give dem kvalificeret feedback (Kruit et al. 2020). Hvordan man giver kvalificeret feedback, er centralt for evaluering for læring, men det vil føre for vidt at forsøge at give et overblik over dette emne her, og der er mange gode ressourcer om dette andetsteds (Se fx Lauvås & Bruun 2021; Wiliam 2011; Keeley 2015).

En måde at arbejde med evaluering for læring på er såkaldte performance assessments (Kruit et al. 2020). Eleverne får en autentisk opgave, som involverer tydelige handlinger, der kan evalueres ud fra forudbestemte kriterier. Hvis opgaven er kompleks nok, vil evalueringen afsløre typiske misforståelser og fejl samt give indsigt i elevens tankeproces.

Evaluering som læring handler om, at elever aktivt indgår i evalueringen ved at indsamle og anvende evalueringsdata til at forstå

Formål med evaluering	Hvem foretager evalueringen	Perspektiv ift. brug af evalueringsdata
Evaluering for læring	Læreren	Formativt/fremadskuende
Evaluering som læring	Eleverne	Formativt/fremadskuende
Evaluering af læring	Læreren, eksterne bedømmere	Summativt/bagudskuende
Evaluering til udvikling af undervisning	Læreren	Formativt/fremadskuende

Tabel 4. Fire forskellige formål med evaluering.



og forbedre deres egen læreproces. Denne tilgang er typisk karakteriseret ved evalueringsformer som selvevaluering eller peer feedback. Studier viser, at selvevaluering kan styrke elevernes præstation, ansvarsfølelse og tro på egne evner (Dochy et al. 1999; Topping 2003), og Earle & Davies (2014) anbefaler selvevaluering som et middel til, at elever løbende kan følge med i deres egen udvikling. En gennemgående forudsætning for, at selvevaluering kan fungere, er, at eleverne har adgang til klare og forståelige kriterier at vurdere sig selv ud fra (Bolden et al. 2020). Derfor anbefales det, at eleverne får mulighed for at blive trænet i selvevaluering, og at de får god tid til at revidere deres arbejde på baggrund af selvevalueringen.

Et eksempel på undervisning, som kombinerer selvevaluering og peer feedback, findes hos Valle et al. (2016). Forfatterne brugte såkaldte scoring guides og rubrics i musikundervisningen i 1.-4. klasse til at udvikle en evalueringspraksis, hvor eleverne kunne dygtiggøre sig på egen hånd. Ved hjælp af simple tjeklister og mere avancerede skemaer (rubrics) blev eleverne løbende bedre til at forstå og selv forbedre deres kompetencer.

Evaluering af læring har til formål at bruge data til at bedømme elevers præstationer. Denne form for evaluering betegnes ofte som summativ. Det er en proces, hvor data anvendes til at dokumentere elevers kompetencebesiddelse på et givet tidspunkt (Dolin et al. 2017). Et typisk eksempel på denne type evaluering er eksaminer. Evaluering af læring er nyttigt til at beskrive en elevs niveau på et givet tidspunkt (typisk efter et bestemt forløb) og afspejler resultatet af en læringsproces (Nicolaou & Constantinou 2014). En

fare ved denne evaluering af læring er, at lærere og elever bliver så optaget af, at eleverne opnår gode resultater, at det egentlige mål, at fremme læring, træder i baggrunden. Det gælder særligt, hvis evaluering af læring ikke er designet til at indfange kompetencer, men reelt peger på mindre komplekse læringsmål (Dolin 2020).

Et eksempel på evaluering af læring er Nielsen (2015), der undersøgte, hvordan autentiske problemstillinger kunne anvendes i eksamenssituationer i gymnasiet på en måde, så evalueringen på samme tid kunne indfange elevernes innovationskompetence og begrebsmæssige faglige niveau. Det er typisk vanskeligt at indfange kompetencer ved en eksamen, da der sjældent er tid til at nå rundt om alle aspekterne. I dette tilfælde var det dog muligt at opnå en nuanceret vurdering af elevernes kompetencer, fordi eksaminatorerne til den mundtlige eksamen havde adgang til at følge elevernes arbejdsprocesser fra undervisningen. Det viser, at selv ved en eksamen kan man opnå en mere nuanceret bedømmelse af elevens kompetence.

Evaluering af undervisning fokuserer på at bruge evalueringsdata til at udvikle og forbedre undervisning. Formålet er ikke i første omgang at give formativ feedback til eleverne, men derimod at forbedre et undervisningsforløb (Jönsson et al. 2013). Selvom der ikke umiddelbart er fokus på den enkelte elev, men på at forbedre undervisningen, vil man dog typisk sikre, at evalueringen også fremmer elevernes udvikling (Højgaard & Niss 2023). Læreren bruger evalueringsdata til at se, hvor undervisningen virker, og hvor den kan justeres. Perspektivet er stadig formativt, men modtageren af feedback er lære-



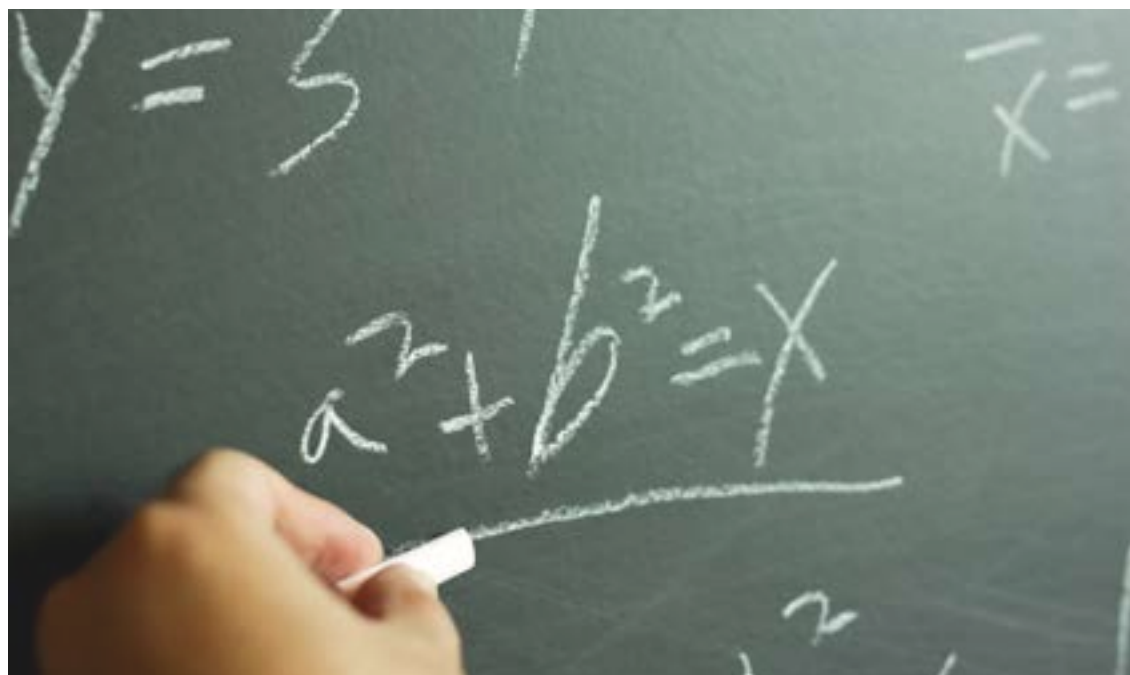
ren selv. Dette kan ske løbende eller efter et forløb med henblik på næste gang, man skal forberede undervisning med samme sigte.

Flere studier fremhæver potentialet i at gøre evalueringen af undervisningen til et kollektivt anliggende, hvorved lærerne sammen gennem systematisk indsamling og vurdering af elevdata får værdifuld feedback om deres egen undervisning – samtidig med at de bliver bedre til at evaluere deres elever (Foster et al. 2007).

Validitet og reliabilitet som kvalitetskriterier

Vi afrunder vores analyse af, hvordan evalueringsdata kan vurderes og bruges, ved at sætte fokus på to klassiske evalueringsmæssige kvalitetskriterier, som diskuteres flittigt i litteraturen: validitet og reliabilitet.

Validitet (også kaldet gyldighed) er et kvalitetskriterie, der handler om, hvorvidt en evaluering fanger det (og kun det), som den hævder at evaluere (Højgaard & Niss 2023). I litteraturstudiet er vi stødt på især to forhold, der eksemplificerer, at det kan være en udfordring at opnå høj validitet. Den ene vedrører opgavebaserede opgaver. Her kan elevernes læsevne blive afgørende, og det kan være en udfordring for validiteten, når det ikke er læsevnen, der skal evalueres (Ketelhut et al. 2013). Det kan være en særlig evalueringsmæssig udfordring i forhold til tosprogede elever eller elever med læsevanskeligheder. I samme boldgade fremhæver OECD (2003), at kulturelle forskelle også kan udfordre validiteten. Nogle opgaver kan være fremmedgørende eller vanskelige for forskellige grupper af elever af årsager, der ikke har noget at gøre med målet for evalueringen.





Det andet forhold er gruppearbejde. Her er der risiko for, at ikke alle elever når det, som evalueringen sigter mod. Enkelte elevers præstationer kan være stærkt påvirket af, hvad andre gør, og er derfor ikke nødvendigvis udtryk for den enkelte elevs kompetencer (Harrison 2014). Fx viste Nielsen (2015), at i forbindelse med en øget brug af gruppearbejde i gymnasieundervisningen (som følge af indførelsen af innovationskompetence som læringsmål) var det vanskeligt for bedømmerne at foretage en valid evaluering af elevernes samarbejdskompetence, da det var tilfældigt, hvad eleverne foretog sig, når bedømmerne var i nærheden.

Reliabilitet (også kaldet pålidelighed) handler om, hvorvidt samme præstation vurderes til samme niveau, uafhængigt af hvem der er bedømmer, hvornår evalueringen sker eller andre faktorer, som kan skabe støj i bedømmelsen (Højgaard & Niss 2023).

Pettersson et al. (2010) peger på, at et mere struktureret bedømmelsesgrundlag i form af kriterier mv. øger evalueringens reliabilitet. Grundlæggende kan man øge reliabiliteten ved at gøre vurderingerne mere opdelte, fx ved at øge antallet af delmål, man ønsker at måle kompetencen ud fra, eller ved at bruge detaljerede bedømmelseskemaer såsom rubrics.

En anden måde at øge reliabiliteten på er ved at samarbejde om at udvikle en fælles vurderingstilgang. Thompsons (2017) beskriver fx, hvordan et fagteam på en skole gennemførte fælles vurderinger af skriftlige elevbesvarelser af kompetenceorienterede opgaver. Her diskuterede lærerne forskellige elevbesvarelser og sammenlignede dem med beskrivelserne af fem forskellige niveauer

af målopfyldelse i stil med dem beskrevet i afsnit xx [Grader af målopfyldelse][S2.1]. Gennem diskussion af konkrete elevprodukter blev lærerne mere konsistente i, hvilket niveau de enkelte præstationer lå på, og dermed øgedes reliabiliteten af deres vurderinger.

En udfordring i forhold til at opnå høj reliabilitet kan være, hvis bedømmerne ikke forstår kriterierne eller forskellene på niveauer af præstation på samme måde (Turner et al. 2015). Dette er særligt relevant for evaluering af kompetence, da kompetencemål naturligt lægger op til, at opgaverne er relativt komplekse, hvilket øger risikoen for, at vurderingen bliver mindre ensartet.

Nødvendigheden af at prioritere validitet

Ved enhver evaluering er det oplagt at tilstræbe både høj validitet og høj reliabilitet, men det kan være en udfordring at opnå begge dele, og man må ofte gå på kompromis med den ene for at opnå den anden (Jensen 2007; Højgaard & Niss 2023).

Hvis evalueringen fx er en eksamen, vil man ofte prioritere reliabilitet over validitet for at sikre, at forskellige elever får samme karakter for samme præstation. Men som mange har påpeget, kan eksaminer eller standardiserede opgaver have svært ved at indfange et validt billede af elevernes kompetence, da man ofte må give køb på kompleksitet for at gøre opgaverne ens for alle (Carlsson et al. 2007).

Omvendt vil undervisningen ofte have gavn af, at man prioriterer evalueringens validitet højest. Går man på kompromis med



validiteten, risikerer man, at undervisningen og elevernes læring rettes mod det forkerte, da evaluering har en kraftig tilbagevirkende effekt på begge dele (Black 2016).

Her vil vi gerne dvæle lidt ved betydningen af validitet og tilføje nogle refleksioner fra vores øvrige arbejde med evaluering af kompetencer, som ikke dukkede op eksplicit i den fundne litteratur, men som er vigtigt at vide. Vi har omtalt mange udfordringer for lærere i forhold til at evaluere faglige kompetencer, og det kan være fristende at forsøge at gøre vurderingen mere enkel ved at nedbryde kompetencemålene til mindre delelementer, som er nemme at vurdere hver for sig. Men det indebærer en alvorlig risiko for at miste hele pointen med kompetencemål.

Opdeling af en kompetence i mindre dele og fokus på detaljer, når man skal vurdere den, kan potentielt give mindre pres på lærerne, fordi en sådan "atomisering" (Højgaard & Niss 2023) er forholdsvis nem at forholde sig til og gennemføre i praksis. Det var fx en del af rationalet, da man i forbindelse med folkeskolereformen i 2014 udviklede de såkaldte Forenklede Fælles Mål (Højgaard & Sølberg 2023).

En sådan opdelt tilgang kan potentielt også bidrage til at højne reliabiliteten, da mindre komplekse evalueringsmål er nemmere at evaluere entydigt. Udfordringen er dog, at evalueringens validitet bliver kritisk lav. Hvis forenklingen af kompetencemålene fører til, at delmålene ikke længere indfanger kompleksiteten og kernen i helheden, så har



FOTO: MIKAL SCHLOSSER



man mistet pointen med kompetencemålene. Så er man gået i det, vi vil betegne atomiseringsfælden.

Vi er stødt på denne problematik mange gange i vores litteraturstudie (fx Berggren 2013; Marshall 2004; Sadler 2013; Shavelson 2010). Graf et al. (2009) beskriver det fx i det såkaldte CBAL-projekt. Her var ambitionen at designe et redskab til at evaluere matematikkompetence ved at udpege delkompetencer med dertilhørende opgaver og svartyper til at beskrive specifikke præstationsniveauer for hver delkompetence. Forfatterne endte dog med at opgive at være tro mod de nationale kompetencemål, fordi pilotafprøvninger pegede på, at opgaverne i så fald ville blive for svære. På den måde lykkedes det at nedbryde kompetencen i mindre dele, men essensen af kompetencemålene gik tabt i processen.

For at navigere uden om atomiseringsfælden, når man designer og implementerer en given evalueringspraksis, foreslår vi, at man går eksplicit holistisk til værks. Det vil sige, at man bevidst fastholder fokus på helheden og kompleksiteten i sin evaluering frem for at stirre sig blind på isolerede delmål. Man har stadig brug for en god karakteristik af kompetencemålet, så man kan sikre sammenhængen mellem kompetencemål, undervisningen og evalueringen, men især vurderingen bør tænkes holistisk for at holde fast i ambitionen bag kompetencemålene.

Et eksempel på, hvordan man praktisk kan opnå en holistisk vurdering, er gennem projektarbejde, som er en form for autentisk evaluering. Frejd (2013) peger med afsæt i et litteraturstudie på, at projektarbejde som evalueringsform har høj validitet, men lav reliabilitet. Da projektarbejde integrerer både

faglige, sociale og kommunikative elementer, er det potentielt muligt at opnå en holistisk vurdering, der afspejler den form for kompleksitet, som kompetencemål typisk sigter mod.

Et andet eksempel er ved at arbejde med såkaldte læringshistorier, som Carr (2001) har udviklet og anvendt i en børnehavekontekst. Det er en form for porteføljeevaluering, baseret på observationer i hverdagssituationer og udformet med henblik på at skabe en kumulativ række af kvalitative "øjebliksbilleder" eller skriftlige vignetter fra enkelte børn. Vurderet samlet kan disse nedslagspunkter give et billede af, hvilke aspekter af en given faglig kompetence et barn har demonstreret at være i besiddelse af. Ifølge Carr kan denne narrative tilgang indfange kompleksiteten af kompetencemålene i form af sociale, kognitive og affektive aspekter, ligesom den giver vægt til betydningen af deltagelse og kultur og inkorporerer børnenes stemme i evalueringsprocessen.

Mangel på tid og lærerkompetencer

Mange lærere nævner manglende tid som en af de væsentligste barrierer for evalueringsarbejdet. Fx peger Willis & Adie (2014) på, at rubrics, som ellers kan være nyttige redskaber til systematisk evaluering, tager for lang tid at udvikle og anvende ordentligt.

Selve indsamlingen og analysen af evalueringsdata kan også være så tidskrævende, at det hindrer systematisk evaluering (Østergaard 2013). Indsamlingen alene kan være svær, hvis læreren skal nå at registrere evalueringsdata i forbindelse med, at eleverne arbejder, og samtidig sørge for, at undervisningen fungerer (Albanese & Case 2016).



Dertil kommer udfordringen med at finde tid til at bruge evaluering konstruktivt i forbindelse med feedback til eleverne og til at udvikle undervisningen (Baartman et al. 2011). Når eleverne så har fået feedback, har de brug for tid til at bearbejde den, så de ved, hvad de kan gøre fremadrettet, hvilket også kan lægge yderligere tidspres på undervisningen (Hill & Edwards 2019).

Der kan således være mange tidsmæssige udfordringer, men som vores litteraturstudie viser, findes der efterhånden også mange ressourcer (fx skemaer, digitale platforme, undervisningsforløb med designet evaluering) til at understøtte lærere. Mange af ressourcerne er dog ikke tilpasset det danske uddannelsessystem og skal derfor først oversættes og afstemmes til danske forhold. Dertil kommer, at selv veldesignede digitale platforme, som kan facilitere dataindsamling og automatisere en del af behandlingen af evalueringsdata, kræver træning, før lærere kan gennemføre evalueringen pålideligt (Deane & Sparks 2019).

Mange af ressourcerne er også specifikke og afdækker kun enkelte dele af de kompetencer, lærerne skal forholde sig til. Samtidig begrænser de læreres mulighed for at tilpasse evalueringen til eleverne, hvilket kan skade "fagets potentiale til at motivere og engagere eleverne" (Dolin 2020: 28).

Mere grundlæggende har det vist sig at være en udfordring for mange lærere at forholde sig til kompetencemål frem for andre typer læringsmål (Stille, Jang & Wagner 2015), da evaluering af kompetencemål rummer en større kompleksitet og kræver en anden tilgang end evaluering af fx videns- eller færdighedsmål (Sølborg 2020). Således peger litteraturstudiet på en stor udfordring i form af nødvendig

(efter)uddannelse til lærere, så de kan arbejde mere systematisk med evaluering i deres undervisning. Dette er ikke et udtryk for manglende evne eller vilje fra lærernes side. Det er et udtryk for en manglende opmærksomhed på evalueringens betydning, som betyder, at lærerne generelt set ikke er klædt på til eller får støtte til at arbejde systematisk med evaluering.

Der findes heldigvis eksempler på projekter, hvor det lykkes at udvikle læreres evalueringspraksis ved at danne fællesskaber, hvor lærere sammen kunne udvikle og understøtte hinandens praksis. Fx FADC-projektet (Furtak et al. 2016), hvor ni biologilærere i gymnasiet arbejdede sammen i tre år om at udvikle formative evalueringsmetoder. Gennem samarbejde på tværs af skoler udviklede de kompetence til at designe og anvende formative evalueringsopgaver, som kunne indfange elevernes ideer og tænkning omkring bestemte emner.

Wrenn (2004) beskriver tilsvarende et historiefagteam, der i fællesskab rammesatte et problemorienteret projektarbejde. Ved sammen at udpege og operationalisere centrale historiekompetencer fik lærerne en fælles forståelse af evalueringsgrundlaget. Og MacDonald (2007) beskriver, hvordan indsamlingen af pædagogisk dokumentation som middel til at evaluere elevernes læring krævede et samarbejde mellem lærere og pædagoger, som der ellers ikke var tradition for i den beskrevne kontekst.

Inddragelse af elever i evaluering

Afslutningsvis vil vi fremhæve, at eleverne ikke kun behøver at være genstand for evaluering, men også kan være aktive deltagere.



Fordelene ved at inddrage eleverne aktivt i evalueringer kan være mange, fx:

- Eleverne lærer at give, modtage og anvende feedback og at udvikle selvstændige læringsstrategier (Berggren 2013).
- Eleverne lærer at sætte mål (Dolin et al. 2017).
- Det styrker elevernes selvtillid og motivation (Hill & Edwards 2019).
- Det fremmer refleksive praksisser hos eleverne (Andrade & Valtcheva 2009).
- Det kan fremme selvindsigt hos eleverne i forhold til læringsbehov og dermed styrke deres evne til selvregulering (Brindley & Marshall 2015).

Der er imidlertid også en række udfordringer ved at inddrage eleverne i evalueringer. Fokusering på elevernes evaluering af sig selv og andre kan føre til, at de oplever både stress, præstationsangst og eksklusion i undervisningen (Kaas, Berg & Grumløse 2020). Desuden fremhæves det gang på gang, at det kræver klare rammer og læringsmål, hvis eleverne skal kunne evaluere sig selv og hinanden på en kvalificeret måde (Andrade & Valtcheva 2009). Det tager tid at opnå den nødvendige rutine, så man undgår, at elevernes evalueringer bliver for subjektive (Berggren 2013). Selv da kan de være tilbøjelige til at fejlvurdere egne præstationer og være overvenlige i deres feedback til andre – særligt, hvis de ikke har det nødvendige faglige overblik (Dolin et al. 2017). Dertil kommer, at eleverne skal lære at bruge og indgå i evalueringer, og at det kun er gennem løbende dialog, at de bliver i stand til at få det fulde ud-

bytte af evalueringerne (Willis & Adie 2014).

På trods af disse udfordringer peger vores samlede gennemgang på, at elevinddragelse i evaluering rummer et betydeligt potentiale for at styrke både læring og motivation, når inddragelsen sker systematisk. Det kræver tid, tydelige rammer og en vedvarende indsats fra læreren at opbygge en kultur, hvor eleverne kan indgå i evalueringsprocesser på en konstruktiv og kvalificeret måde. Med tiden kan elevinddragelse udvikle sig til et værdifuldt redskab, der ikke blot dokumenterer læring, men også aktivt bidrager til at forme den. Hermed understøtter den både den enkelte elevs faglige udvikling og de overordnede mål for undervisningen. På den måde udgør elevinddragelse et vigtigt element i en helhedsorienteret evalueringspraksis.

Refleksionsspørgsmål

- ✓ Hvad bruger du primært dine evalueringsdata til – at fremme læring, registrere udvikling eller forbedre undervisningen? Arbejder du med tydelige kriterier, når du vurderer elevernes kompetenceniveau? Hvis ja, deler du dem med eleverne?
- ✓ Hvordan sikrer du dig, at eleverne faktisk lærer af den feedback, de får? Hvordan fastholder du det, som eleverne siger og gør, så du kan bruge det i din evaluering i situationen såvel som senere? Hvad vejer tungest, når du evaluerer dine elevers kompetencer: validitet eller reliabilitet?



Om forfatterne

30 - 34



Tomas Højgaard er lektor ved DPU, Aarhus Universitet, og forsker i matematikdidaktik. Forskningen handler for en stor dels vedkommende om at

undersøge muligheder og vanskeligheder ved at bruge faglige kompetencebeskrivelser som didaktisk udviklingsværktøj, både generelt i forhold til faglige kompetencer og med særligt fokus på matematisk modelleringskompetence. Herudover forsker han i vilkår for og vanskeligheder ved at etablere et gensidigt frugtbart samspil mellem forskningsfeltet matematikkens didaktik og matematikundervisningens praksis.



Jan Sølberg er lektor ved Institut for Naturfagenes Didaktik, Københavns Universitet, og forsker i naturfagsdidaktik.

Han har arbejdet med naturfaglige kulturer og andre faktorer, der er afgørende for udvikling af naturfagsundervisning. Gennem sin involvering i mange forskellige forsknings- og udviklingsprojekter har han vist, at det oftest tager mange år at opnå reelle ændringer i praksis. I de senere år har hans arbejde kredset om, hvordan man definerer, arbejder med og evaluerer naturfaglige kompetencer.



Referencer

- Abrahams, I., Reiss, M.J. & Sharpe, R. M. (2013): The assessment of practical work in school science. *Studies in Science Education* 49(2), 209-251.
- Albanese, M. & Case, S.M. (2016): Progress testing: critical analysis and suggested practices. *Advances in Health Sciences Education* 21, 221-234.
- Andersen, T.B.L. (2014): *Evaluering for læring i naturfagene*. Astra.
- Andrade, H. & Valtcheva, A. (2009): Promoting learning and achievement through self-assessment. *Theory Into Practice* 48(1), 12-19.
- Anker-Hansen, J. & Andrée, M. (2015): Affordances and constraints of using the socio-political debate for authentic summative assessment. *International Journal of Science Education* 37(15), 2577-2596.
- Baartman, L.K.J., Prins, F.J., Kirschner, P.A. & van der Vleuten, C.P.M. (2011): Self-evaluation of assessment programs: A cross-case analysis. *Evaluation and Program Planning* 34(3), 206-216.
- Berggren, J. (2013): *Learning from giving feedback: Insights from EFL writing classrooms in a Swedish lower secondary school* [unpublished master's thesis]. Stockholm University.
- Black, P. (2016): The role of assessment in pedagogy – and why validity matters. I: D. Wyse, L. Hayward & J. Pandya (red.): *The SAGE handbook of curriculum, pedagogy and assessment*, 725-739. Sage.
- Bolden, B., DeLuca, C., Kukkonen, T., Roy, S. & Wearing, J. (2020): Assessment of creativity in K–12 education: A scoping review. *Review of Education* 8(2), 343-376.
- Brindley, S. & Marshall, B. (2015): Resisting the rage for certainty: Dialogic assessment. *English Teaching: Practice & Critique* 14(2), 121-139.
- Buckley, B.C., Gobert, J.D., Horwitz, P. & O'Dwyer, L.M. (2010): Looking inside the black box: Assessing model-based learning and inquiry in BioLogica. *International Journal of Learning Technology* 5(2), 166-188.
- Burkhardt, H. (2018): Ways to teach modelling: A 50-year study. *ZDM—Mathematics Education* 50, 61-75.
- Carr, M. (2001): *Assessment in early childhood settings: Learning stories*. Sage.
- Calfee, R., Wilson, K.M., Flannery, B. & Kapinus, B. (2014): Formative assessment for the Common Core literacy standards. *Teachers College Record* 116(11), 1-32.
- Carlsson, C.-G., Gerrevall, P. & Pettersson, A. (2007): *Bedömning av yrkesrelaterat kunnande*. HLS förlag.
- Christiansen, J.L., Lilius, K.M., Thynebjerg, K.A., Jensen, M.A.S., Andersson, J. & Kinnerup, L.B. (2020): Evaluering af modelleringsprocessen i naturfagsundervisningen. *MONA – Matematik- og Naturfagsdidaktik* 4, 46-64.
- Deane, P. & Sparks, J.R. (2019): Scenario-based formative assessment of key practices in the English language arts. *Handbook of formative assessment in the disciplines*, 68-96. Routledge.
- Dochy, F.J.R.C., Segers, M. & Sluijsmans, D. (1999): The use of self-, peer- and co-assessment in higher education: A review. *Studies in Higher Education* 24(3), 331-350.
- Dolin, J. (2016): Idealer og realiteter i målorienteret undervisning. *CURSIV: Sammenlignende fagdidaktik* 19, 67-87.
- Dolin, J. (2020): *Evaluering på godt og ondt*. Aarhus Universitetsforlag.
- Dolin, J., Bruun, J., Nielsen, S.S., Jensen, S.B. & Nieminen, P. (2018): The structured assessment dialogue. *Contributions from Science Education Research*, 4, 109-140. Springer.
- Dolin, J. & Krogh, L.B. (2010): The relevance and consequences of PISA science in a Danish context. *International Journal of Science and Mathematics Education* 8(3), 565-592.
- Dolin, J., Nielsen, J.A. & Tidemand, S. (2017): Evaluering af naturfaglige kompetencer. *Acta Didactica Norge* 11(3), artikel 2.
- Earle, S. & Davies, D. (2014): Assessment without levels. *Education in Science* 258, 30-31.
- Erduran, S., El Masri, Y., Cullinane, A. & Ng, Y.P.D. (2020): Assessment of practical science in high stakes examinations: a qualitative analysis of high performing English-speaking countries. *International Journal of Science Education* 42(9), 1544-1567.
- Frejd, P. (2013): Modes of modelling assessment—A literature review. *Educational Studies in Mathematics* 84(3), 413-438.
- Foster, D., Noyce, P. & Spiegel, S. (2007): When assessment guides instruction Silicon Valley's mathematics assessment collaborative. In *Assessing mathematical proficiency*, 137-154. Cambridge University Press.
- Furtak, E.M., Kiemer, K., Circi, R.K., Swanson, R., de León, V., Morrison, D. & Heredia, S.C. (2016): Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science* 44(3), 267-291.
- German, S. (2017): Teacher to teacher: Vote, discuss, revote: A formative assessment classroom technique. *Science Scope* 41(2), 26-29.
- Gerrevall, P. (2003): Bedömning av demokratisk kompetens: En pedagogisk utmaning. *Utbildning & Demokrati* 12(3), 41-66.
- Gorlewski, D.A. (2010): Research for the classroom: Overflowing but underused: Portfolios as a means of program evaluation and student self-assessment. *The English Journal* 99(4), 97-101.
- Graf, E.A., Harris, K.R., Marquez, E., Fife, J.H. & Redman, M. (2009): *Cognitively based assessment of, for, and as learning (CBAL) in mathematics: A design and first steps toward implementation* (ETS Research Report No. RR-09-xx). Educational Testing Service.
- Harrison, C. (2014): Assessment of inquiry skills in the SAILS project. *Science Education International* 25, 112-122.
- Hill, L. & Edwards, F. (2019): Student perceptions of their involvement in formative assessment feedback practices: "I can do it myself" *Assessment Matters* 13, 6-43.
- Hutchins, N.M., Biswas, G., Maróti, M., Lédeczi, K., Grover, S., Wolf, R., Blair, K.P., Chin, D., Conlin, L., Basu, S. & McElhaney, K. (2020): C2STEM: A system for synergistic learning of physics and computational thinking. *Journal of Science Education and Technology* 29(1), 83-100.



- Højgaard, T. (2017): Hjælpe middelkompetence. YouTube. <https://www.youtube.com/watch?v=0J-xZm6jRdc&t=53s>
- Højgaard, T. (2025): Teaching for competence. *The Curriculum Journal*. Manuskript accepteret til udgivelse.
- Højgaard, T. & Niss, M. (2023): Om formativ evaluering af matematiske kompetencer. *Sammenlignende Fagdidaktik* 7, 37-56.
- Højgaard, T., & Sølberg, J. (2023): Fostering competence: a narrative case study of developing a two-dimensional curriculum in Denmark. *Journal of Curriculum Studies*, 1-28.
- Jensen, T.H. (2007): Assessing mathematical modelling competency. I: *Mathematical modeling (ICT-MA 12): Education, engineering and economics*, 141-148. Horwood Publishing.
- Jönsson, A., Ekborg, M., Lindahl, B. & Löfgren, L. (2013): *Bedömning i NO: Grundskolans tidiga år*. Gleerups Utbildning.
- Kaas, L.A., Berg, M.S. & Grumløse, S.P. (2020): Når skolens børn evaluerer sig selv. *Studier i Lærerdannelse og Profession* 5(2), 56-75.
- Ketelhut, D., Nelson, B., Schifter, C. & Kim, Y. (2013): Improving science assessments by situating them in a virtual environment. *Education Sciences* 3(2), 172-192.
- Keeley, P. (2015): *Science formative assessment, volume 1: 75 practical strategies for linking assessment, instruction, and learning*. Corwin Press.
- Kruit, P., Oostdam, R., van den Berg, E. & Schuitema, J. (2020): Performance assessment as a diagnostic tool for science teachers. *Research in Science Education* 50(3), 1093-1117.
- Kuo, C.-Y., Wu, H.-K., Jen, T.-H. & Hsu, Y.S. (2015): Development and validation of a multimedia-based assessment of scientific inquiry abilities. *International Journal of Science Education* 37(14), 2326-2357.
- Lauvås, P. & Bruun, J. (2021): *Ren formativ evaluering i skolen*. Klim.
- Legé, J. (2007): To model, or to let them model? That is the question! I: *Modelling and applications in mathematics education* 10, 425-432. Springer.
- Liu, L., Rogat, A. & Bertling, M. (2013): A CBAL science model of cognition: Developing a competency model and learning progressions to support assessment development. *ETS Research Report Series* 2013(2), 1-54.
- MacDonald, M. (2007): Toward formative assessment: The use of pedagogical documentation in early elementary classrooms. *Early Childhood Research Quarterly* 22(2), 232-242.
- Marshall, B. (2004): Goals or horizons—The conundrum of progression in English: Or a possible way of understanding formative assessment in English. *The Curriculum Journal* 15(2), 101-113.
- Martin, C.S., Polly, D. & Kissel, B. (2017): Exploring the impact of written reflections on learning in the elementary mathematics classroom. *The Journal of Educational Research* 110(5), 53-553.
- McIntosh, J. & Milam, M. (2016): Competitive debate as competency-based learning: Civic engagement and next-generation assessment in the era of the Common Core learning standards. *Communication Education* 65(4), 420-433.
- Nicolaou, C.T. & Constantinou, C.P. (2014): Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review* 13, 52-73.
- Nielsen, J.A. (2015): *Rapport fra arbejdsgruppe for prøveformer der tester innovationskompetencer i gymnasiet*. Institut for Naturfagenes Didaktik, Københavns Universitet.
- Niss, M. & Jensen, T.H. (2002): *Kompetencer og matematiklæring: Ideer og inspiration til udvikling af matematikundervisning i Danmark*. Undervisningsministeriet.
- OECD (2003): *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. OECD Publishing.
- Pettersson, A., Olofsson, G., Kjellström, K., Ingemansson, I., Hallén, S., Björklund Boistrup, L. & Alm, L. (2010): *Bedömning av kunskap: För lärande och undervisning i matematik*. Institutionen för matematikämnet och naturvetenskapsämnenas didaktik, Stockholms Universitet.
- Ropohl, M., Nielsen, J.A., Olley, C., Rönnebeck, S. & Stables, K. (2018): The concept of competence and its relevance for science, technology and mathematics education. I: *Transforming assessment: Through an interplay between practice, research and policy*, 3-25). Springer.
- Rönnebeck, S., Nielsen, J.A., Olley, C., Ropohl, M. & Stables, K. (2017): The teaching and assessment of inquiry competences. I: *Contributions from Science Education Research* 4, 27-52. Springer.
- Sadler, D.R. (2013): Making competent judgments of competence. I: *Modeling and measuring competencies in higher education*, 11-27. Brill.
- Shavelson, R.J. (2010): On the measurement of competency. *Empirical Research in Vocational Education and Training* 2(1), 41-63.
- Spektor-Levy, O., Eylon, B.S. & Scherz, Z. (2009): Teaching scientific communication skills in science studies: Does it make a difference? *International Journal of Science and Mathematics Education* 7(5), 875-903.
- Stille, S.V.V., Jang, E. & Wagner, M. (2015): Building teachers' assessment capacity for supporting English language learners through the implementation of the STEP language assessment in Ontario K-12 schools. *Canadian Journal of Education* 32, artikel 1.
- Sølberg, J. (2020): To metaforer for læring og kompetenceorienteret undervisning. *MONA – Matematik- og Naturfagsdidaktik* (2), 7-22.
- Sølberg, J., Bundsgaard, J., & Højgaard, T. (2015): Kompetencemål i praksis – hvad har vi lært af KOMPIS? *MONA - Matematik og Naturfagsdidaktik* (2), 46-59.
- Thompson, M. (2017): Skab fokus på mestring. *MONA – Matematik- og Naturfagsdidaktik* (1), 7-21.
- Topping, K. (2003): Self and peer assessment in school and university: Reliability, validity and utility. I: *Optimising new modes of assessment: In search of qualities and standards*, 55-87. Springer.
- Turner, R., Blum, W. & Niss, M. (2015): Using competencies to explain mathematical item demand: A work in progress. I: *Assessing mathematical literacy*, 85-115. Springer.



Valle, C., Andrade, H., Palma, M. & Hefferen, J. (2016): Applications of peer assessment and self-assessment in music. *Music Educators Journal* 102(4), 41-49.

William, D. (1994): Assessing authentic tasks: alternatives to mark-schemes. *Nordic Studies in Mathematics Education* 2(1), 48-68.

William, D. (2011): *Embedded formative assessment*. Solution tree press.

Willis, J. & Adie, L. (2014): Teachers using annotations to engage students in assessment conversations: Recontextualising knowledge. *The Curriculum Journal* 25(4), 495-515.

Wilson, M. & Carstensen, C. (2007): Assessment to improve learning in mathematics: The BEAR assessment system. I: *Assessing mathematical proficiency*, 311-332. Cambridge University Press.

Wrenn, A. (2004): Making learning drive assessment: Joan of Arc-Saint, witch or warrior? *Teaching History* 115, 44-51.

Østergaard, L. (2013): Multimodal evaluering af deltagerstyret problem- og undersøgelsesbaseret aktiviteter i børnehave og indskoling. *NorDiNa* 9, 125-144.



Kommende udgivelser

Børn og unges skrivelyst

Skærmbrug i skolen

Matematikangst

Fritidspædagogikken i skolefritidsordningen og på fritidshjem

Læringsmålsorientering i matematikundervisningen



dpu.au.dk/paedagogiskindblik

