

# Udvikling og evaluering af undersøgelsesbaseret undervisning<sup>1</sup>



Wynne Harlen

*Originaltitel: EVALUATING INQUIRY-BASED SCIENCE DEVELOPMENTS – A PAPER COMMISSIONED BY THE NATIONAL RESEARCH COUNCIL IN PREPARATION FOR A MEETING ON THE STATUS OF EVALUATION OF INQUIRY-BASED SCIENCE EDUCATION. Artiklen er oversat af redaktionen. Den bringes her med forfatterens tilladelse. Originalen kan ses på [http://www7.nationalacademies.org/bose/wharlen\\_inquiry\\_mtg\\_paper.pdf](http://www7.nationalacademies.org/bose/wharlen_inquiry_mtg_paper.pdf)*

## Introduktion

af Jens Dolin, Institut for Naturfagenes Didaktik, KU

Redaktionen har valgt at oversætte Wynne Harlens artikel om undersøgelsesbaseret naturfagsundervisning – eller IBSE (Inquiry Based Science Education) som det efterhånden hedder på nudansk – som oplæg til MONA-konferencen om netop undersøgelsesbaseret naturfags- og matematikundervisning den 4. oktober 2011.

Artiklen kan ses som et led i den serie af "klassiker"-artikler der i MONA 2008(1) blev introduceret med John Deweys artikel fra 1909 om naturvidenskab som stofområde og som metode. Harlens artikel er langtfra så gammel og kan vel næppe kaldes en klassiker. Men den har en række generelle betragtninger, især vedrørende implementering af IBSE og problemerne ved at evaluere effekten af at indføre IBSE – betragtninger som desuden har værdi ud over IBSE. Og Wynne Harlen selv har været aktiv forsker inden for såvel evaluering som IBSE i en menneskealder, og hun er bl.a. medlem af det videnskabelige rådgivningspanel for Fibonacci-projektet (som er omtalt i Claus Michelsens artikel i dette nummer) – og en af hovedtalerne ved MONA-konferencen.

<sup>1</sup> Oversætterens anmærkning (o.a.): I artiklen oversættes "inquiry-based" konsekvent med "undersøgelsesbaseret". Der er et par steder angivet den originale engelske betegnelse ved den valgte danske betegnelse når det kan lette forståelsen.

En undersøgelsesbaseret tilgang til naturfagsundervisning og -læring er bestemt ikke ny – den kan næsten siges at række tilbage til Dewey og hans insisteren på at færdigheder og viden udvikles ved at den lærende aktivt undersøger sine omgivelser. Men den har aldrig været – og er stadig ikke – ukontroversiel. Undervisere og forskere strider om værdien af induktive kontra deduktive tilgange til undervisning, om hvorvidt man kan have projektarbejde uden et solidt fagligt grundlag, eller om projektarbejde netop er en (motiverende) vej til faglig viden. I den engelske litteratur formuleres det som en modsætning mellem "direct instruction" og "constructivist instruction", og debatten er ophedet.

Virker en undersøgelsesbaseret tilgang, lærer eleverne noget? er det oplagte spørgsmål. Hvad er der evidens for? Harlens artikel påpeger hvor svært det er at undersøge dette spørgsmål, og på konferencen vil bl.a. Robin Miller argumentere for at det ikke er en farbar vej frem. Han støttes af kognitionsforskere som Paul A. Kirschner, John Sweller og Richard E. Clark der på baggrund af den menneskelige kognitions opbygning afviser en undersøgelsesbaseret tilgang. Det gør de i artiklen "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching" (i *Educational Psychologist* 41(2), 2006) der gav anledning til en ophedet debat. Andre undersøgelser finder positive effekter af forskellige former for undersøgelsesbaseret undervisning.

En række metastudier giver ikke afklaring på spørgsmålet. Fx undersøger Minner et al. (2010) 138 studier af undersøgelsesbaseret undervisning, og selvom de viser at undersøgelsesprocesser har positiv indvirkning på faglig læring og fastholdelse af det lærte (undervisning virker!), så konkluderer undersøgelsen bl.a. at "evidens for effekter af en undersøgelsesbaseret undervisning ... er ikke overvældende positive", og "et generelt højt undersøgelsesniveau var ikke associeret med et mere positivt læringsudbytte for eleverne" (s. 493, min oversættelse).

Så måske er pointen at man selv skal prøve og selv skal vurdere om det virker. Og her kan Harlens artikel være en stor støtte.

### Referencer

- Kirchner, P.A., Sweller, J. & Clark, R. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), s.75-86.
- Minner, D.D., Levy, A.J. & Century, J. (2010). Inquiry-Based Science Instruction – What Is It and Does It Matter? Results from a Research Synthesis Years 1984-2002. *Journal of Research in Science Teaching*, 47(4), s.474-496.

## Indledning

Denne artikel har til hensigt at præsentere spørgsmål der har at gøre med evaluering af undersøgelsesbaseret undervisning og læring inden for naturfag, med særlig opmærksomhed rettet mod at vurdere alle effekter på elevernes udbytte i naturfag. Da mange af disse spørgsmål har med karakteren af undersøgelsesbaseret læring at gøre, begynder artiklen med en diskussion af hvad undersøgelse er, af hvad undersøgelsesbaseret undervisning er, og af begrundelserne for at benytte denne. Anden del omhandler spørgsmål om implementering og de vanskeligheder der er ved at vurdere effekten af ændringer på ét område af elevernes erfaringer når der er så mange andre ting der influerer på deres læring. Tredje del ser på nogle spørgsmål om evalueringsdesign. Den fjerde ser på nogle cases med evalueringsundersøgelser der kaster lys over nogle af disse spørgsmål, og det fører frem til sidste afsnit som ser på veje videre frem. Artiklen er skrevet ud fra et perspektiv der er klart positivt over for undersøgelsesbaseret undervisning og læring.

## 1. Om undersøgelse og undersøgelsesbaseret læring og undervisning

### 1.1 Undersøgelse

Det er ikke svært at finde definitioner af begrebet undersøgelse. Den ofte citerede beskrivelse i National Science Education Standards (boks 1) indfanger essensen.

#### *Boks 1*

Undersøgelse er en mangeartet aktivitet der involverer at foretage observationer, stille spørgsmål, undersøge bøger og andre kilder til oplysninger for at se hvad der allerede vides, at planlægge undersøgelser, gennemgå hvad der allerede vides, i lyset af eksperimentel evidens, anvende værktøjer til at indsamle, analysere og fortolke data, foreslå svar, forklaringer og forudsigelser og formidle resultaterne. Undersøgelse fordrer identifikation af antagelser, brug af kritisk og logisk tænkning og inddragelse af alternative forklaringer. (National Research Council, 1996, s. 23)

Forskere bruger undersøgelse til at udvikle forståelse af naturen og den menneskeskabte verden fordi det fører til teorier og idéer som forklarer observerede begivenheder og fænomener. Det indebærer også en erkendelse af at nuværende teorier giver de bedste forklaringer vi kan give på dette tidspunkt, men når der dukker facts op som

modsiges nuværende teorier, skal disse ændres, og alternativer afdækkes. Stephen Hawking forklarer dette i boks 2.

Når elever udvikler deres forståelse af naturen og den menneskeskabte verden omkring dem, kan de, ligesom forskerne, bruge undersøgelse til at nå frem til idéer og teorier der hjælper dem med at forklare hvad de observerer. Elever bliver også nødt til at ændre deres forestillinger når de støder på ny og modstridende viden. Og også ganske som forskerne begynder de ikke med en ren tavle, men med det de allerede ved, og de forestillinger de allerede har.

Men i modsætning til forskere har elever – især de mindreårige vi beskæftiger os med i grundskolen – ikke på forhånd en veludviklet evne til at observere, indsamle data, lave forudsigelser, teste mulige forklaringer og fortolke resultater. Derfor er det et centralt mål for undersøgelsesbaseret undervisning i naturfag på det elementære niveau at hjælpe eleverne med at udvikle disse evner (kaldet spørge-, undersøgelses- eller proceskompetencer). Men det er en fejl at opfatte kompetenceudvikling som det eneste fokus for undersøgelsesbaseret undervisning og læring, for formålet er også at udvikle deres indsigt. For igen at citere National Science Education Standards: “Når elever deltager i undersøgelse ... identificerer de antagelser, de bruger kritisk og logisk tænkning, og de overvejer alternative forklaringer. På denne måde udvikler elever aktivt deres forståelse af naturvidenskab ved at kombinere naturvidenskabelig viden med fornuft og tænkning.” (National Research Council, 1996, s. 2).

## Boks 2

Enhver fysisk teori er altid foreløbig, i den forstand at den kun er en hypotese: Du kan aldrig bevise den. Uanset hvor mange gange forsøgsresultaterne stemmer overens med en teori, kan du aldrig være sikker på at resultatet ikke næste gang vil være i modstrid med teorien. På den anden side kan du modbevise teorien ved at finde blot en enkelt observation der strider mod teoriens forudsigelse ... Hver gang nye eksperimenter konstateres at stemme overens med teoriens forudsigelse, overlever teorien, og vores tillid til den vokser; men hvis der nogensinde bliver foretaget en ny observation der strider mod teorien, er vi nødt til at opgive eller ændre den. I det mindste er det det der bør ske, men man kan jo altid sætte spørgsmålstejn ved kompetencen hos den person der har foretaget observationen. (Hawking, 1988, s. 10)

## 1.2 Undersøgelserbaseret læring i praksis

Der er mange måder at lave undersøgelser på i skolerne: Det er ikke et undervisningsprogram, ikke en arbejdsplan eller en læreplan. Den praktiske udførelse af læring gennem undersøgelse kan variere fra en stribe aktiviteter hen over en måned der så udvikler sig til et helt år hvor eleverne er optaget af at studere en forårspyt, som beskrevet af Rankin (1999), til korte undersøgelser der passer til en eller to lektioner. At have forskelligartede gennemførelser "er ikke blot uundgåeligt, men også ønskeligt fordi de tegner et rigt billede af meningsfuld læring i forskellige situationer. En mangfoldighed af typer af undersøgelsesundervisning og -læring opmuntrer lærere til at engagere sig i at deltage i undersøgelser på måder der passer til deres egne forestillinger og undervisningstilgang" (Keys & Bryan, 2001, s. 632).

Det der knytter disse forskellige implementeringsmåder sammen og gør dem alle genkendelige som "undersøgelserbaseret undervisning og læring", er de aktiviteter som læreren og eleverne er engageret i. Boks 3 opsummerer nogle centrale træk ved disse aktiviteter.

### Boks 3

#### Lærerhandlinger

- Sørge for oplevelser, materialer, informationskilder som elever direkte kan bruge.
- Demonstrere brugen af værktøjer og materialer som eleverne får brug for i undersøgelsen.
- Stille åbne og personcentrerede spørgsmål for at afdække elevers nuværende forståelse og deres måde at forklare det de finder ud af, på.
- Engagere elever ved at foreslå hvordan deres idéer kan afprøves, eller deres spørgsmål besvares gennem undersøgelse, eller hvordan de kan finde viden fra sekundære kilder.
- Hjælpe eleverne når det er nødvendigt med planlægning, så idéer kan testes ordentligt.

#### Elevhandlinger

- Engagere sig i udforskning af materialer, begivenheder, ting.
- Arbejde i grupper der sammen udveksler idéer og konstruerer forståelse.
- Stille spørgsmål og overveje hvordan de kan besvares via undersøgelse.
- Foreslå mulige forklaringer på observationer.
- Foreslå hvordan idéerne bag mulige forklaringer kan testes, eller spørgsmål besvares gennem undersøgelse/aktiv udforskning.
- Planlægge og udføre undersøgelser med relevante observationer og målinger eller på andre måder skaffe data der kan afprøve idéer.

- Lytte til elevernes idéer og tage dem alvorligt.
- Stille spørgsmål der opmuntrer eleverne til at tænke over hvordan man kan forklare det de finder.
- Sørge for muligheder for kollaborativ læring og dialog.
- Stilladsere alternative idéer der kan forklare undersøgelsesdata.
- Indsamle oplysninger om elevernes udvikling i kompetence og opfattelse gennem observation, spørgsmål og interaktion.
- Tage noter og optegne resultater hensigtsmæssigt.
- Knytte resultater til de forestillinger og spørgsmål der er under afprøvelse eller besvarelse.
- Kommunikere hvad de har foretaget sig; lytte til og udveksle idéer med andre.
- Reflektere over undersøgelsesprocessen og overveje eventuelle andre måder at gøre det på.

Aktiviteterne i boks 3 beskriver åben eller selvstændig undersøgelse hvor læreren giver plads til eleverne til at udvikle deres egne spørgsmål og designe egne undersøgelser. I praksis er så stor frihed sjældent mulig, og det er mere sandsynligt at eleverne undersøger "strukturerede" eller "guidede" spørgsmål identificeret af læreren (Windschitl, 2001).

### *1.3 Hvorfor vil vi gerne have børn til at lære på denne måde?*

Svaret her vil nødvendigvis være kort, og det er begrænset til tre vigtige og indbyrdes forbundne pointer. Den første er at læring gennem undersøgelse er i overensstemmelse med moderne læringspsykologi, ifølge hvilken den lærende spiller en aktiv rolle i sin læring. Det "behavioristiske" syn på læring (Skinner, 1974) – baseret på idéen om at den adfærd som regelmæssigt belønnes, vil blive styrket, mens den adfærd der gentagne gange straffes, vil forsvinde – placerer styringen af læring uden for den lærende. Eleven er modtageren af færdiglavet forståelse af hvad der skal accepteres og huskes. Dette synspunkt er helt i modstrid med opfattelsen af læring som konstrueret gennem den lærendes mentale aktivitet der knytter nye og tidligere erfaringer sammen. Evidensen for at den lærende deltager aktivt i læringen, kommer ikke kun fra læringspsykologer, men fra studier af hjernens funktion (Bransford, Brown & Cocking, 1999). Piaget (fx 1929, 1955, 1956) førte an ved at påvise at selv små børn bestræber sig på at finde mening i det de ser omkring sig. Nyligere studier af babyer viser også hvordan de kan skelne mellem forskellige detaljer ved ting længe før de kan tale (Gopnik, Meltzoff & Kuhl, 1999). Så når børn kommer i skolen, er deres hoveder på ingen måde "tomme kar".

Den anden pointe drejer sig om karakteren af naturvidenskabelig viden. Meget af naturfagsundervisningen i skolen formidler en gammeldags opfattelse af videnskab som værende værdifri, en række fakta der skal læres, og som afspejler "sandheder" som på en eller anden måde eksisterer og skal opdages af dem der er kloge nok til at gøre dét. Men naturvidenskab ses i dag (fx Hawking, 1988) som et resultat af menneskelig tænkning der resulterer i teorier som accepteres så længe de passer til den til rådighed stående viden. Teorierne kommer fra fælles tænkning blandt videnskabsfolk og er, som historien viser, udtryk for hvad der er socialt og kulturelt acceptabelt. (Det tog Kopernikus' idéer [om at jorden og planeterne bevæger sig om solen] hundrede år at blive accepteret selvom denne model stemte langt bedre overens med observationer end Ptolemæus' verdensmodel, primært fordi den geocentriske model var blevet accepteret af den kristne kirke. Og der er nutidige eksempler på teorier baseret på kulturmønstre der i nogle samfund foretrækkes frem for videnskabelige teorier). Hvis vi således betragter det at elever lærer naturvidenskab som en parallel til at forskere "laver naturvidenskab", så er det vigtigt at eleverne diskuterer og sammenligner deres teorier med andre og kombinerer deres tænkning om hvordan evidens gives mening. En elev er en del af en gruppe, bidrager til gruppens tænkning og vinder meget fra det. Nu om stunder er "gruppen" ikke nødvendigvis fysisk samlet på ét sted, men selv når læring foregår på afstand, lærer en person af og med andre. Idéen om "social konstruktion af idéer" stammer fra Vygotskys arbejde (1978) hvor han argumenterede for at forståelse eksisterer i den sociale interaktion når lærende som gruppe diskuterer resultater ("evidence", o.a.) og forsøger at forklare og beder andre om at retfærdiggøre deres forklaringer; hver enkelt lærende opbygger sin forståelse gennem denne proces.

Begge disse pointer knytter an til den tredje begrundelse for undersøgelsesbaseret undervisning, nemlig at den sætter os i stand til at tage højde for de forestillinger som elever allerede har. Den betydelige mængde af forskning der afdækker elevers forforestillinger om naturvidenskab, bevidner at børn stræber efter at skabe mening i verden omkring dem uanset om de bliver undervist i naturfag eller ej. De forestillinger de bruger til at forklare tingene, giver mening for dem ud fra deres (begrænsede) erfaring og kompetencer. Selvom disse forestillinger ofte virker mærkelige og ulogiske for voksne, kræver det kun en smule eftertanke at se hvordan eleverne kan have dem. Disse idéer afspejler hvad eleverne har oplevet, og viser tydeligt deres bestræbelser på at give mening til deres erfaringer. Som et eksempel beskriver en otteårig i boks 4 hvordan rust på et søm opstår, ud fra sine erfaringer og observationer af form og forekomst af rust.

### Boks 4

Der er en væske i sømmet som siver ud af det.

Derved dannes store klumper efterhånden som det siver ud.

Denne væske kommer kun ud når det er vådt.

Der må være en slags signal der fortæller det at det skal lække.

### Boks 5

En eftertænksom 12-årig pige sagde at hun syntes det var svært at tro på noget af det hendes lærer sagde. Fx forklarede læreren forekomsten af dug på græs på en kold morgen ved vanddamp i luften. Pigen havde allerede sin egen forklaring på det, nemlig at kulde i græsset skabte vandet. Denne idé passede også med hendes oplevelse af vanddråber på en flaske som lige er taget ud af køleskabet. Selvom lærerens forklaring ikke gav mening for hende, kunne hun ikke bare afvise den så den forblev en forklaring hun kendte til og kunne huske hvis hun blev spurgt, men den var ikke en del af hendes forståelse.

Børnenes egne forestillinger giver mening for dem og bliver foretrukket frem for det "naturvidenskabelige synspunkt" når der ikke er nogen overbevisende grunde til at deres egne idéer er fejlagtige, eller at den alternative forklaring skulle give bedre mening. I boks 5 var pigen ikke i stand til at teste hverken sin egen eller lærerens forklaring på en måde der kunne overbevise hende. Der er en hel del evidens for at børn beholder deres egne ikkevidenskabelige synspunkter sideløbende med de naturvidenskabelige når sidstnævnte blot er superponeret på førstnævnte i stedet for at erstatte dem som et resultat af viden ("evidence", o.a.) opnået gennem undersøgelse.

Naturvidenskabelige forklaringer passer ikke blot i en bestemt situation, men forklarer også tilsvarende fænomener. Et vigtigt mål for naturfagsundervisning er at give eleverne mulighed for at udvikle mere bredt anvendelige idéer ("store" tanker) da disse har større forklaringskraft end "små" idéer som kun forklarer konkrete tilfælde. Ved at starte med de forestillinger børn har fra tidligere erfaringer, kan undersøgelsesbaseret forbinde disse med nye erfaringer og derved danne "større" idéer. Ved således at hjælpe eleverne med at bruge og udvikle deres forestillinger sætter læreren dem i stand til at forstå en bred vifte af fænomener. Men læring kan ikke finde sted i den omvendte rækkefølge, ved at kortslutte det undersøgende og undervise direkte



i "store" idéer, fordi disse er for abstrakte og for fjerne fra elevernes erfaringer til at give mening. Fx udvikler elever forestillinger om hvordan dyr og planter i deres egen have eller park eller å sameksisterer med hinanden, og ud fra dem udvikler de mere generelle idéer om levende organismers indbyrdes afhængighed. Men hvis de store idéer er udgangspunktet, kan det være at forståelsen ikke stikker dybere end slogans, og det vigtige begreb bliver måske aldrig virkelig forstået.

Men selv med en forudsætning om at det at sætte børn i stand til at lære gennem undersøgelse er et værdifuldt mål, skal der i mange klasserum betydelige ændringer til for at tilvejebringe den slags lærer- og elevhandling der er angivet i boks 3. Så nu vender vi os imod måder det kan gøres på, og hvordan vi kan konstatere hvor effektive de er.

## 2. Gennemførelse og evaluering af undersøgelsesbaseret undervisning og læring

### 2.1 Om at skabe forandring i klasseværelset

At skabe forandring i klasselokalet udefra sker ved en kombination af en eller flere af følgende:

- a) Direkte ændringer i elevernes oplevelser ved brug af særlige oplevelser eller "lærersikrede" materialer ("teacher-proof", forklaring følger, o.a.), online eller papirbaserede
- b) Ændringer i undervisningen gennem tilvejebringelse af undervisningsmaterialer eller programmer med lærervejledning og andre materialer
- c) Ændringer hos lærere gennem faglig udvikling, online eller ansigt til ansigt.

a) Direkte ændringer i elevernes oplevelser kan tilvejebringes ved særlig undervisning eller oplevelser som ikke er en del af skolens sædvanlige aktiviteter, såsom naturfags-sommerlejr. Noget tyder på at undersøgelsesbaserede oplevelser af denne art kan påvirke læring (selvom det meste af forskningen er lavet med gymnasieelever) og, måske endnu vigtigere, kan præge mere varige holdninger til naturvidenskab (Gibson & Chase, 2002). Men den slags aktiviteter berører kun et lille antal elever, og for at opnå store forandringer er det vigtigt at påvirke den almindelige undervisning. Det er formentlig en myte at der kan laves materialer til elever som i det store og hele kan køre uafhængigt af en lærer og direkte kan influere på eleven ("lærersikrede"). Det kan godt være at man kan stille regneopgaver på den måde, men man kan ikke stimulere den slags elevhandling der nævnes i boks 3. Det ligger i undersøgelsesbaserede aktiviteter natur at deres forløb er vanskelige at forudsige, og selvom it-programmer kan være mere interaktive end papirbaserede materialer, så er en lærers facilitering

nødvendig for at give den samkonstruktion af forståelse som ligger i at lære naturfag gennem undersøgelse. Derfor vil vi ikke yderligere omtale denne direkte vej til elever.

b) Udvikling af programmer der for lærere udstikker procedurer for undersøgelsesbaserede aktiviteter (og ofte også indeholder elevmaterialer) er en almindelig måde at skabe forandring på; tiltalende fordi den åbner mulighed for at nå alle de skoler der er villige til at bruge programmet. Men det er svært udelukkende skriftligt at formidle begrundelserne for de klasserumsinteraktioner der skal til hvis lærerne skal udvikle elevernes forståelse gennem undersøgelse, og der er meget der tyder på at mange lærere bruger disse tilgange uden engagement og opbakning til de underliggende værdier. Af den grund begrænser nogle programudviklere brugen til dem der også har deltaget i efteruddannelsesaktiviteter, selvom disse ofte er minimale.

c) At bruge faglig udvikling (efteruddannelse) som det vigtigste middel sker i erkendelse af at forandring i undervisningen sandsynligvis fordrer mere end bare tilvejebringelse af egnede undervisningsmaterialer, og at der kan være behov for ændringer i lærernes opfattelse af naturvidenskab, af undersøgelse og af hvordan elever lærer. Mange efteruddannelsesprogrammer (fx Institute for Inquiry på Exploratorium) forsøger gennem deres tilgang at modellere de undervisningsaktiviteter som lærerne skal bruge. De giver også lærerne lejlighed til at opleve undersøgelse på deres eget niveau og skaber dermed et solidt grundlag for forståelse af den intendede elevlæring. De afspejler Supovitz & Turners (2000) seks kritiske komponenter af høj kvalitetsefteruddannelse opregnet i boks 6. Disse erfaringer kræver tid, ganske som refleksion over dem og diskussionerne om konsekvenserne for lærerens rolle gør. Den nedenfor refererede forskning viser at 80 timers efteruddannelse er nødvendig for at dette bliver virkningsfuldt. Dette understøttes af erfaringer hos efteruddannelsesudbydere, der finder at praksisforandringen hos deltagere der tager mere end ét kursus, stiger.

### *Boks 6*

For at være af høj kvalitet skal faglig udvikling:

- lade deltagerne fordybe sig i undersøgelse
- være både intensiv og vedholdende
- involvere lærere i konkrete undervisningsopgaver og være baseret på læreres erfaringer med elever
- fokusere på faglig viden og uddybe lærernes faglige kompetencer
- være baseret på et fælles sæt af standarder for faglig udvikling
- være forbundet med andre aspekter af skolereform.

(Fra Supovitz & Turner, 2000, s. 964-965)

## 2.2 Hindringer for evaluering af virkningen af programmer og efteruddannelse

Som nævnt ovenfor skal storskalaforbedringer af elevers læring foregå i en tottrinsproces hvor input i form af programmer eller efteruddannelse medieres af læreren (boks 7). Dette bringer en række ubekendte ind i billedet som slører ("confound", o.a.) evalueringen af effekterne, og det rejser spørgsmål om design og de anvendte instrumenter.

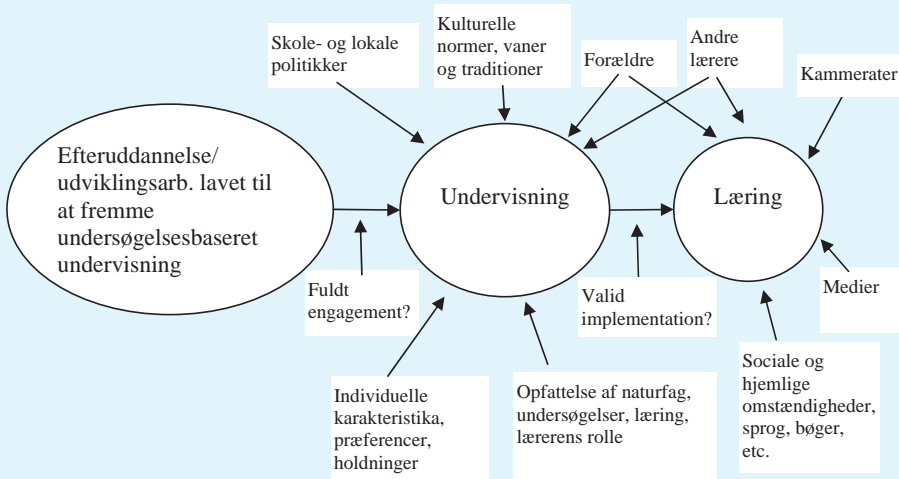
### Slørende variable

Spørgsmålet om hvor effektive undersøgelsesbaserede udviklingsprogrammer og efteruddannelse er til at gennemføre ændringer i undervisning og læring, er notorisk vanskeligt at besvare. Det kræver kun en smule eftertanke at indse at der er rigtig mange faktorer der påvirker processerne på hvert enkelt trin fra oprettelsen af programmet eller efteruddannelsesdesignet frem til den lærende. Dermed kan det synes som om en hvilken som helst effekt kan forventes at blive for udvandet til at kunne måles, selv før man begynder at overveje hvordan denne effekt overhovedet skulle kunne konstateres. I boks 7 er der en liste over nogle af de mulige begrænsende faktorer for implementering af et udviklingsprogram.

Forbindelsen mellem et program eller de professionelle udviklere og klasserumlæreren er den første trussel mod implementeringen; læreren er måske ikke helt og fuldt engageret eller kan have misforstået nogle af de intenderede signaler. Læreres muligheder for at føre noget ud i livet kan være begrænset af skolepolitikker, fx skemalægning, eller skolevæsenets politikker, fx krav om at forberede elever til visse test. Andre af klassens læreres undervisningspraksis kan være i strid med det der kræves for undersøgelsesbaseret læring, og kan være forskellig, fx hvad angår omfanget af forventet elevsamarbejde og indbyrdes diskussion af deres lektier. På den måde får elever blandede signaler om hvilke aktiviteter læring indebærer. Forældre kan også have visse forventninger til hvad lærerne laver, som kan adskille sig fra hvad et program kræver. Hertil kommer effekterne af lærernes egen baggrund, erfaringer og overbevisninger som kan forvandle programmet, måske ubevidst. Derfor kan det være et spørgsmål i hvilken grad programmet kan implementeres efter hensigten.

Selvom alle disse andre indvirkninger på undervisningen ikke er så store at de hæmmer de elever der får undersøgelsesbaserede erfaringer, så bliver disse kun en del af deres samlede erfaringer. Andre lærere, kammerater, forældre og medier har direkte eller indirekte indflydelse på deres læring.

## Boks 7



### 3. Spørgsmål vedrørende evalueringsdesign

#### 3.1 Variabelkontrol

Hvis virkningen af undersøgelsesbaserede oplevelser skal kunne ses i sammenhæng med alle de andre faktorer der har indflydelse på elevernes læring, er det nødvendigt at kigge på en situation hvor undersøgelseserfaring er den eneste forskel mellem to grupper hvis læring så kan sammenlignes. Det betyder at grupperne skal være ens fx mht.:

- indholdet af naturfagsundervisningen
- elevernes karakteristika og baggrund (køn, etnicitet, modersmål osv.)
- lærernes baggrundsforståelse af og holdning til naturvidenskab og undersøgelsesundervisning

for blot at nævne nogle få variable der vides at være forbundet med forskelle i læring. Hvis man skal kunne kontrollere disse variable for at få sammenlignelige "eksperimentel"- og "kontrol"-grupper, skal man kunne måle dem, og det er i sig selv meget vanskeligt. Desuden er de ikke indbyrdes uafhængige, så en gruppeudvælgelse på grundlag af én kan forårsage forskellighed i en anden.

En måde at undgå dette på er at bruge en gruppe som sin egen kontrol, det vil sige at måle ændringer et stykke tid før interventionen (den undersøgelsesbaserede undervisning) begynder, og derefter måle ændringerne over en tilsvarende periode efter interventionen. Det er klart nok at der er problemer her på grund af de forandringer der

under alle omstændigheder finder sted med eleverne over tid. Alternativt kan elever og lærere, i teorien, placeres tilfældigt i forsøgs- og kontrolgrupperne, men det ødelægger de normale elev-elev- og elev-lærer-relationer og skaber en kunstig situation hvis resultater ikke kan overføres til normalklasser. Det kan også være at lærerne er tilbageholdende med at acceptere den forventede rolle eller ikke har tiltro til den gennemførte intervention, og det kan føre til udviskning af afvigelserne mellem de to grupper.

Disse komplikationer og det tilsyneladende umulige i at afdække eventuelle læringsændringer der kan henføres til en intervention, har betydet at program- og efteruddannelsesudbydere kun i ringe omfang har forsøgt at vurdere ændringer i elevernes resultater. De har snarere været tilbøjelige til at bedømme validiteten af deres programmer ud fra det indhold og de frembringelser der indgår, og fra reaktionerne fra brugere og kursusedtagere. På den måde bliver spørgsmål om hvordan man vurderer den opnåede læring, undgået.

Men spørgsmålet er ikke blot om evaluering bør fokusere på undervisningsforandring *eller* læringsforandring – begge dele er nødvendige. At vide *at* elevernes læring er forbedret, fortæller os ikke noget om hvordan denne ændring er sket. I ekstreme tilfælde kan de observerede forandringer have haft meget lidt direkte at gøre med de anvendte programmer eller den gennemførte efteruddannelse. Endvidere kan det hævdes at medmindre der sker ændringer i undervisningen, er ændringer af den intenderede slags læring usandsynlige. Man må også vide hvad slags undersøgelsesbaseret der bliver benyttet (åben, guidet eller struktureret) for at kunne vurdere effekten af tiltaget.

### 3.2 Valg af troværdige mål for elevers udbytte

På trods af hindringerne for at kunne se de ændringer i læring der kan tilskrives virkningen af undersøgelsesbaserede programmer, er det fortsat sådan at den endelige bedømmelse af værdien af en intervention når denne er gennemført, skal være om elevernes læring er blevet forbedret. Så vi *skal* forholde os til spørgsmålet: Hvilke målinger ville give evidens for “succes”? De vigtigste alternativer er:

- om man skal vurdere de resultater der er specifikke mål for undersøgelsesbaseret naturfagsundervisning, såsom brug af proceskompetencer og evnen til at kunne anvende begreber
- om man skal bruge almindelige summative evalueringsmetoder (normalt test) som forventes at lægge større vægt på faktuel viden og negligere kompetencer.

Hvis kun den første af disse bruges, kan resultaterne måske ikke overbevise skeptikerne. Det vil med rette kunne hævdes at hvis man sammenligner med en kontrolgruppe der ikke har haft undersøgelsesbaserede oplevelser, vil alle forvente de bedste

resultater hos forsøgsgruppen. Desuden er der på nuværende tidspunkt ikke mange målemetoder til rådighed til vurdering af resultaterne af undersøgelsesbaseret læring som har den fornødne validitet og pålidelighed.

Hvis den anden mulighed vælges – at bruge sædvanlig evaluering eller test – er det muligt at der i det mindste på kort sigt ikke kan konstateres nogen læringsforbedring. Men i betragtning af påstanden om at undersøgelsesbaseret undervisning forbedrer elevlæring i naturfag generelt, er det et stærkt argument at det at påvise bedre resultater med de gængse læringsmålemetoder giver mere overbevisende evidens.

I begge tilfælde er der et spørgsmål om hvor store ændringer der skal til i forsøgsgruppens resultater fra præ- til postundersøgelsesbaseret undervisning for at de bliver opfattet som tilstrækkelige til at berettige at læreplaner og praksis ændres. Er en smule mere læring hos forsøgsgruppen nok – eller skal det være væsentlig mere? Hvem skal fælde dommen? At basere sig på statistik i denne beslutning er problematisk da det statistisk signifikante ikke nødvendigvis er uddannelsesmæssigt signifikant.

### 3.3 Tidshorisont

Et andet vigtigt designspørgsmål vedrører tidsrammen for en evaluering: Hvilken tidshorisont giver overbevisende dokumentation for reelle læringsfordele? Her er det vigtigt at skelne mellem kun at holde øje med umiddelbare læringsforbedringer og forbedringer som stadig kan ses flere år senere. Der kan argumenteres for værdien af umiddelbare ændringer uanset om de holder sig eller ej, eftersom eleverne kan agere på et højere niveau end ellers, og deres læring vil nødvendigvis have fordel af dette.

Der er flere ting der taler til fordel for en længere tidshorisont før man vurderer læringsgevinsten. For det første er der den tid det tager for læreren at blive fuldt fortrolig med reformen/ændringen og at blive forpligtet på den (dette punkt vender vi tilbage til senere). For det andet det at forskellene mellem naturfagsundervisning gennem undersøgelse og andre tilgange til undervisning er temmelig subtile, så "behandlingen" skal køre over ret lang tid for at give de langsigtede virkninger en chance på præstationsniveauet. Det tredje er at i starten af en implementering kan lærere føle sig usikre på deres rolle. De nødvendige omstillinger kan reducere deres selvtillid (som tidligere har været baseret på en traditionel opfattelse af lærerrollen) og forårsage et midlertidigt kompetencetab indtil han/hun begynder at føle sig dygtig nok i en ny rolle. Dette understøttes af erfaringer (se CASE-eksemplet senere) om at det varer nogen tid før undersøgelsesbaserede oplevelser dukker op i resultatet, og der kan være den fare som Ruiz-Primo et al. (2002) peger på, at konklusioner drages for tidligt i og med at "effekter ikke bliver konstateret selvom de er der, men er små, og den reform der har frembragt dem, skulle have været plejet."

Veje videre frem i forhold til nogle af disse spørgsmål kan udstikkes af undersøgelser af klasserumsaktiviteter og af tegn på forandring i lærerforståelsen. Næste afsnit

giver derfor eksempler fra nylig forskning der er særdeles informativ hvad angår de faktorer der påvirker forandring, også selvom de indskrænker sig til at se på ændringer i undervisningen. Bagefter betragter vi nogle af det begrænsede antal undersøgelser der har set på elevernes udbytte.

## 4. Eksempler på evalueringsundersøgelser

### 4.1 *Undersøgelser af virkningen på undervisning*

Supovitz & Turners (2000) undersøgelse af The Local Systemic Change Initiative indsamlede i 1997 oplysninger om 24 projekter fra hele USA. Spørgeskemaer gav selvrapporterede data om lærerholdninger, lærerforestillinger og undervisningsformer. Skoleledere besvarede også spørgsmål om deres støtte til initiativet. Ud fra det store antal svar (787 skoler og 4.903 lærere) kunne forskerne udvikle skalaer for indikatorer for undersøgelsesbaseret undervisningspraksis. Via en flerniveaumodel knyttede de omfanget af gennemført efteruddannelse (blandt andre lærervariable) til indikatorer for lærernes undersøgelsesbaserede undervisningspraksis og deres niveau for undersøgelsesmæssig klasserumskultur. De fandt at “det var først efter ca. 80 timers efteruddannelse at lærernes svar angav at de betydelig hyppigere end gennemsnitslæreren benyttede undersøgelsesbaseret undervisning” (ibid., s. 973). Deres konklusion var at der var “stærk og signifikant sammenhæng mellem faglig udvikling og lærernes praksis og klasserumskultur” (ibid., s. 975), men varigheden af de fleste efteruddannelsesforløb var for kort til at være effektiv. Deres undersøgelse henledte også opmærksomheden på andre faktorer der påvirker ændring. Fx bruger lærere på skoler hvor eleverne har lav socioøkonomisk status (SES), mere traditionelle undervisningsmetoder end dem med elever med højere socioøkonomisk baggrund. En anden vigtig faktor var skoleledelsens støtte. Et studie af TERC/Lesley Universitys<sup>2</sup> online masterprogram for grund- og mellemskolenaturfagslærere viste at efteruddannelse i undersøgelsesbaseret naturfag kan udbydes online med samme virkning som face-to-face hvis det sker på et omhyggeligt udformet onlinekursus designet til at give læringserfaringerne gennem undersøgelse. Harlen & Altobello (2003) sammenlignede forandringer hos deltagerne på introduktionskurset i programmet (Try Science) med forandringer hos deltagerne på et kursus med samme mål og indhold som blev afholdt som tilstedeværelseskursus. Omkring halvdelen af kursustiden blev brugt på at lære naturfag vha. undersøgelse, og halvdelen på at studere aspekter af undersøgelsesbaseret undervisning. I alle dele af kurserne arbejdede deltagerne i grupper. Onlinedeltagerne blev placeret i grupper på seks eller syv, og deltagerne

---

2 TERC, Technical Education Research Centers, er en amerikansk organisation til fremme af naturvidenskab og matematik. Den har en hjemmeside: [www.terc.edu](http://www.terc.edu) (o.a.).

udvekslede asynkrone mails med hinanden hver uge. Onlinedeltagere gennemførte undersøgelser eller analyse af dokumentation for undervisningen ved hjælp af videoer eller casestudier offline, fulgt af onlinerapportering af resultater eller forslag til deres gruppe. Forskerne indsamlede data om erfaringerne med kurset, ændringer i forståelsen af naturvidenskabens indhold, ændringer i forståelsen af undersøgelse og af undersøgelsesbaseret undervisning, ændring i selvtillid ift. naturfagsundervisning og anvendelse af strategier for undersøgelsesbaseret undervisning i deres egen undervisning. Selvom det primære formål med undersøgelsen var at sammenligne ændringer der havde med online- og face-to-face-undervisning at gøre, er det relevante her det omfang hvori der skete ændringer hos deltagerne på begge kurser.

De vigtigste resultater er opsummeret i boks 8. Hvad angår forståelse af undersøgelse og af undersøgelsesbaseret undervisning, rapporterede begge sæt lærere en større forandring end der kunne observeres ud fra mere objektive kriterier, såsom de definitioner de blev bedt om at angive i præ- og postkursusspørgeskemaer. Der var således en opfattelse af en udvikling af forståelse som ikke var til stede i deres praksis. Det var også slående at der var større succes med at udvikle lærernes naturfagsforståelse end med at skabe forandring af praksis i klasseværelset (bedømt for alle deltagere ud fra deres lektionsplaner og oversigter over disses omsætning til praksis og for fem lærere ud fra observation af lektioner).

## Boks 8

### **Virkingen af de to forskellige undersøgelsesbaserede kurser:**

- Try Science-kurset, både i online- og on-campus-versionen, havde succes med at involvere deltagerne i naturvidenskabelig udforskning hvor de regelmæssigt gjorde brug af undersøgelseskompetencer.
- Begge sæt kursister ændrede deres forståelse af kursets naturfaglige indhold; ændringen var signifikant større for onlinedeltagerne.
- Der var kun få synlige ændringer i deltageres forståelse af definitioner af det undersøgelsesbaserede i deres præ-/postsvar, men deltagerne i begge kurser mente at deres forståelse af undersøgelse i naturfag var blevet forbedret.
- De fleste deltagere havde velbegrundede meninger om undersøgelsesundervisning, og dette ændrede sig kun lidt i kursets løb.
- De lektionsplaner som deltagerne i begge kurser havde udarbejdet, gav eleverne mulighed for hands-on-aktiviteter og for at lave forudsigelser, men levede ikke op til forventningerne til elever der skal anvende begreberne i undersøgelser af egne spørgsmål.



*Forskelle mellem online- og on-campus-efteruddannelse:*

- Den vigtigste forskel mellem online og on-campus i deltageres erfaring var at førstnævnte hyppigere end sidstnævnte reflekterede over deres læring og over undersøgelsesprocessen.
- Onlinedeltagere oplevede, værdsatte og kommenterede deres kollaborative læring; de følte ikke at de arbejdede alene.
- Den tillid til egen undervisningskompetence i naturfag gennem undersøgelse som lærerne gav udtryk for, steg i kursets løb, og betydelig mere for online- end for on-campus-deltagerne.
- Det gennemsnitlige tidsforbrug på kurset for onlinedeltagere var ca. 90 timer og for face-to-face-deltagerne omkring 66 timer (heraf 36 timer som undervisning).

Baseret på Harlen & Altobello (2003).

Undersøgelser af ændringer hos lærere har påpeget forskellige stadier i implementering af nye strategier, startende med anerkendelse af og tilslutning til forandring uden megen forståelse og sluttende med fuld forståelse og tilslutning (Rudduck & Kelly, 1976; Doubler, 1991). En bredt valideret række af stadier for gennemførelse udvikledes af Hall & Loucks (1977) baseret på de overvejelser lærere gør sig når reformer skal føres ud i livet. Ud fra denne teori har Hord et al. (1987) udviklet de stadier der er anført i boks 9.

**Boks 9**

0	Opmærksomhed	Minimale overvejelser om eller deltagelse i en reform
1	Information	Almindelig opmærksomhed over for reformen
2	Personlig	Usikkerhed om reformens krav og personlig formåen til at opfylde disse krav
3	Styring	Opmærksomhed fokuseret på effektivitet, organisation, forvaltningskrav
4	Konsekvens	Fokusering på konsekvenserne af reformen for eleverne
5	Samarbejde	Koordinering og samarbejde med andre i udnyttelsen af reformen
6	Nyfokusering	Udforskning af mulighederne for større ændringer og alternativer

Selvom der måske ikke er almindelig enighed om hvordan de specifikke stadier skal beskrives, er der accept af at lærere gennemløber en gennemførelsesproces der således ikke er én enkeltstående begivenhed, men en gradvis ændring af forståelse og af praksis. Det kan også være forskelligt for de forskellige aspekter hvad der skal til. Fx kan nogle lærere lettere arbejde sig frem til fase fire hvad angår ændringer i egen spørgepraksis, hvorimod de kan befinde sig på et tidligere stade hvad angår fx at sætte elever i stand til at undersøge egne spørgsmål.

For deltagerne i Try Science-kurser fremgik det at de kun befandt sig på begyndelsesstadier af forandring. Deltagernes forståelse for og engagement i at foranstalte den slags oplevelser for eleverne som kræves for undersøgelsesbaseret læring, var ikke tilstrækkelig god til at ændre deres eksisterende praksis. I et introducerende kursus kan man naturligvis ikke forvente dét. Anekdotisk evidens tyder på at de senere i onlinemasterprogrammet blev bedre til at foretage de radikale ændringer i deres rolle der skal til for at gennemføre undersøgelsesbaseret læring i klasseværelset.

#### *4.2 Studier af virkninger på elevlæringen*

Spørgsmålet om hvor tæt vurderingen af læring bør svare til læringsaktiviteterne, blev behandlet i Ruiz-Primo et al.s (2002) forskning om ændringerne hos elever i 5. klasse der fulgte emner i FOSS<sup>3</sup>-pensum. Deres multiniveau-tilgang var baseret på testscorer som indikator for reformens virkning. De fremførte at hvis der er ændringer, vil de først vise sig i præstationer i klasserumsaktiviteter, derefter i overførsel til nært beslægtede aktiviteter med et andet indhold, derefter i mere fjernt forbundne opgaver og så videre. På dette grundlag identificerede de metoder til vurdering af læring på fem niveauer af afstand fra klasserumslæring: umiddelbar, nær, proksimal, distanceret og fjern ("immediate, close, proximal, distal and remote", o.a.). De udviklede mål for forandringer på disse niveauer i to 5.-klasser-naturfagsmoduler af FOSS-programmet (boks 10).

Forskerne brugte elevernes noter, indsamlet ved afslutningen af de otte måneders undersøgelse, som datakilde for forløbet af modulet såvel som et mål for "umiddelbar" læring. Vurderinger af "nær" og "proksimal" blev lavet ved præ- og posttest. Efter posttest blev "distanceret"-vurdering udført.

Resultaterne viste at der i gennemsnit for de 20 involverede klasser var signifikante forskelle mellem præ- og posttestresultater for både den "nære" og den "proksimale" vurdering. Som ventet var forskellene større for "nær"-målene end for de "proksimale", og der var et betydeligt spænd på tværs af klasser og mellem de to enheder der blev undervist. Bedømt ud fra elevernes noter bemærkede forskerne at "kravene i de lærerstillede opgaver ... var generelt lave. Læreren var tilbøjelig til at bede eleverne

3 FOSS Full Option Science System, [www.fossweb.com](http://www.fossweb.com) (o.a.).

registrere forsøgsresultaterne (ikke at fortolke dem) eller til at gentage definitioner. Disse opgaver kan næppe i sig selv hjælpe eleverne med at forbedre deres forståelse.” (ibid., s. 383). Da der ikke var nogen prætestresultater for den distale måling, blev korrelationer mellem de distancerede og de nære og proksimale posttestscorer brugt til at estimere effekten på “distanceret”-vurderingerne. Disse korrelationer var signifikante og, imod forventning, af nogenlunde samme størrelsesorden alle sammen.

### Boks 10

#### Variabelenheder for niveauerne umiddelbar, nær, proksimal, distanceret og fjern for FOSS (Ruiz-Primo et al., 2002)

- Umiddelbar: vurdering af artefakter i løbet af læseplansaktiviteter; specifikt hvor godt variable blev identificeret og manipuleret i pendulsvingninger.
- Tæt: vurdering ifm. en modificeret pendulaktivitet, med variation i de anvendte materialer og den måde den afhængige variabel blev målt på.
- Proksimal: vurdering baseret på at bede eleverne om at forklare hvad der får flasker til at flyde eller at synke, hvor eleverne er nødt til at bruge viden om udregninger med variable og om hvordan resultater kan fortolkes i en noget anderledes situation.
- Distanceret: en test udviklet til storskalavurdering baseret på National Science Education Standards og administreret af California Systemic Initiative Assessment Collaborative, med “affald” som emne.
- Fjern: ingen specifikke oplysninger, bortset fra at det ville være en generel vurdering af uddannelsesniveaut.

Selvom konklusionerne i denne undersøgelse var foreløbige, viser de at en vurderings “afstand” fra klasserummet er af betydning.

The Cognitive Acceleration through Science Education (CASE)-projekterne i Storbritannien er egentlig ikke undersøgelsesbaserede forløb, men deres formål med at udvikle tankeprocesser gør dem særdeles relevante i denne sammenhæng. Der er tre CASE-projekter der alle har som mål at udvikle niveauet af børns tænkning. I det første projekt var målgruppen de 11- og 12-årige, og målet for klasserumstiltagene var at fremme tænkning fra Piagets konkret-operationelle niveau til det formelt-operationelle niveau. Adey & Shayer (1990) udformede en række aktiviteter der skulle udvikle elevernes identificering af variable og deres brug af variabelrelationer til at

foretage forudsigelser for at stimulere kognitiv konflikt og tilskynde til metakognition og eksplicit sammenkædning af strategier til sammenhænge uden for dem hvori de blev udviklet.

Lærere der gjorde brug af disse materialer under evalueringens testfase, fik særlig træning. I testfasen fik de kontrolbesøg om den måde selve forløbet kørte på. Forsøgs- og kontrolklasser blev testet før og efter de to års forsøg med en forskerudviklet Piaget-ræsonnementstest. Posttesten omfattede også en naturfaglig test udviklet i samarbejde med lærerne som kunne tjene som skolens årsprøve. Der var også en umiddelbar posttest, en posttest ét år efter forløbet og én efter endnu et år.

Resultaterne af de nationale test (GCSE) i naturfag der tages af alle 16-årige, indgik også. Disse målinger kan altså beskrives som "distancerede" med de glosor der anvendes i Ruiz-Primo et al. (2002).

Resultaterne umiddelbart efter de to års forløb udviste kun forskelle i Piaget-ræsonnementstesten, hvor signifikante forskelle fra kontrolgruppen kun blev fundet for 12-årige drenge. Et år senere var disse forsvundet, men forskelle begyndt at vise sig i naturfagsposttesten. Naturfagspræstationerne var tilsyneladende forbedret, og effekten var endnu større i GCSE-resultaterne senere. Forsøgsgruppen var signifikant foran i naturfagstesten, men der var også fremgang i matematik og i engelsk, hvilket muligvis kan forklares ved at det er eksaminer der kræver analyse af figurer og forståelse af tekster, og den kompetence kan godt have været påvirket af forløbets aktiviteter. Resultaterne synes at vise at generel tænkning kan læres. Men de viser også at effekten meget vel kan komme senere.

Succesen for CASE på grænsen mellem konkret og formel tænkning opmuntrede til en indsats på de lavere alderstrin, hvor børn ifølge Piaget rykker fra præoperationel til konkret-operationel tænkning. CASE blev derfor udviklet til fem- til syvårige, i et projekt der hedder Let's Think. Her er der udviklet aktiviteter ud fra Piagets skema for konkrete operationer, dvs. seriation, klassificering, synspunkter, kausalitet, regler og tidssekvensering (Adey et al., 2002). Resultaterne af det ét år lange forsøg viste at de elever der brugte denne slags aktiviteter, havde væsentlig større fremgang på test for rumlig forståelse og på en bevarelestest ("conservation test" jf. Piaget, o.a.) end tilsvarende kontrolgrupper. Da bevarelsesbegrebet ikke indgår i programmet, viser dette overførsel og antyder at disse aktiviteter kan hæve elevernes generelle tænkeniveau (Robertson, 2004).

Som forventet var der variation på tværs af skoleklasser, og undersøgelser af årsagerne til dette øger forståelsen af hvad der har indflydelse på implementering. En vigtig og forventet faktor var støtte fra skoleledelsen og den centrale ledelse. Presset på engelske skoler for at opfylde de nationale pensumkrav kan begrænse lærernes frihed til at give aktiviteter af denne art den tid de har brug for for at blive korrekt implementeret, og til at gøre det på det tidspunkt på dagen hvor børn er fokuserede.

Forskerne fandt også at aktiviteterne fungerede bedst når eleverne arbejdede i grupper af blandet køn, etnicitet, evner og personlighed. Eleverne havde også kognitiv udbytte i de klasser hvor læreren var opmærksom på hvad der betinger godt gruppearbejde, samtidig med at han/hun gjorde det klart at succeskriteriet for dét er at "hjælpe hinanden med at tænke bedre".

Et tredje CASE-projekt der bruger materialer udviklet for otte- og niårige, er i gang for tiden. Det bruger de nationale test som gives til 11-årige og omfatter naturfagstest, til at evaluere indvirkningen på børns læring.

Et andet engelsk projekt (Black et al., 2003) hvor efteruddannelse var det eneste input, brugte National Tests i naturfag som et mål for effektivitet. Dette projekt fokuserede på at udvikle læreres brug af formativ evaluering. Selvom det ikke direkte fokuserede på undersøgelse, kræver formativ evaluering ganske tilsvarende tilgang af læreren – specielt skal han/hun finde og bruge oplysninger om elevernes forestillinger og færdigheder, dele læringsmål med eleverne og opmuntre dem til at tage ansvar for egen læring og at reflektere over hvordan de lærer, så de kan vurdere deres eget arbejde. Projektet involverede efteruddannelsesdata for en toårig periode, hvor den vigtigste drivkraft for ændring var samarbejde i lærergrupper snarere end input fra forskerne hvis rolle var at reagere på lærernes behov. De 19 lærere der deltog, besluttede selv hvilke aspekter af deres praksis de ville ændre for at kunne foretage formativ evaluering, hvoraf de vigtigste var at spørge, at give feedback via rettelser, peer- og selvevaluering og formativ brug af summative test. Mellem efteruddannelsesmøderne besøgte en forsker klasserne for at observere undervisningen og give feedback til lærerne på disse observationer. Efter to års kollaborativt arbejde blev resultaterne for hver klasse sammenlignet med en tilsvarende klasse på samme skole. De fundne forskelle kan udtrykkes ved Nationalt Curriculum-niveauer (ét niveau svarer stort set til to år): knap et halvt niveau for elever i 11-åriges test og lidt over et halvt niveau for eleverne i 13-åriges test. Det er stor fremgang.

Også dette projekt fandt variationer på tværs af klasser og identificerede faktorer der hæmmede implementeringen af ønskelige tilgange. Disse omfattede en lektiepolitik der kræver karakterer på alt arbejde, hvor projektet dokumenterede at det at give kommentarer øger indsatsen (Butler, 1988). Et andet eksempel var målfastsættelse som kræver hyppig repetition og hæmmer ændringer i læringstilgang. Derfor er det vigtigt at skolens ledelse er indstillet på at slække på den slags krav så læreren kan have mulighed for at afprøve fornyelser.

## 5. Veje frem

Disse cases er kun eksempler, men de kaster lys over de spørgsmål der blev stillet i afsnit 2. Dette afsnit indledes med et sammendrag af disse punkter, fulgt af nogle konsekvenser for fremtidige evalueringer.

### 5.1 Forskning i lærerændringer

Hovedpunkterne er:

- Forandring af undervisning kræver ændring både hos lærere og i skolepolitik. Derfor er støtte fra skoleledere og øverste ledelse afgørende for nyskabelsers ordentlige gennemførelse og evaluering.
- Forandring sker mere sandsynligt når lærerne har et vist ejerskab af implementeringsprocessen i klasserummet.
- Nyskabelser bør være til stede i klassen i mindst et år og helst længere før elevernes læringsudbytte vurderes.
- De fleste efteruddannelser er for korte til at have varig effekt.
- Evalueringsmetoder der kun bruger læreres selvrapportering, kan antyde lærerpåvirkninger som ikke nødvendigvis omsættes til praksisændringer.

Disse punkter viser at man overordnet skal være opmærksom på at identificere hvor langt lærerne er kommet i deres forståelse og gennemførelse af undersøgelsesbaseret undervisning. Som uddannede fagfolk besidder lærere faste synspunkter og praksisser om naturfagsundervisning som udspringer af deres forståelse af hvordan elever lærer, og af hvad naturvidenskab er (Harlen & Osborne, 1985). Holdninger ændrer sig i en gradvis proces som det er beskrevet, hvor der er variation fra person til person og hos den enkelte ud fra i hvor høj grad visse aspekter af reformen opfattes og internaliseres. Når virkningerne af reformen på eleverne undersøges, er det derfor afgørende at fastslå hvilket punkt på implementeringsbanen læreren er nået til, da dette sandsynligvis vil påvirke elevernes udbytte.

### 5.2 Forskning i elevændringer

De vigtigste punkter er:

- Data om processen fra klasserummet bør fokusere på forekomsten af kollaborativ læring og af elever der bruger deres undersøgelseskompetence til at forfølge egne spørgsmål under anvendelse af evidens.
- Data om udbyttet skal muliggøre sammenligning af observerede ændringer med hvad der ville være sket i samme tidsrum uden undersøgelsesbaseret naturfagsundervisning.

- Valget af vurderingsinstrumenter for læringsresultaterne kan afgøre om forandringer bliver fundet eller ej; afstanden for overførsel ("transfer", o.a.) er en faktor der skal overvejes.
- Nogle erfaringer er lang tid om at have effekt på elevers tænkning og deres udnyttelse af denne i videre læring.
- Resultatmålinger som er overbevisende for aftagerne af evalueringen, bør anvendes; vurderinger af mål som kun kan opnås af dem der oplever selve nyskabelsen, er nyttige forskningsmæssigt, men ikke politisk.
- Omfanget af læringseffekten skal kunne kommunikeres i termer som aftagerne er bekendt med.

Alt i alt understreger ovenstående punkter hvilke beslutninger om forskningsdesign der har indflydelse på evalueringsresultaterne. Når disse beslutninger træffes, skal man være opmærksom på formålet med evalueringen, så oplysningerne er relevante for brugerne. Listen her påpeger også behovet for at give lærerændringerne tid så lærerne først kan forårsage ændringer i elevernes klasserumsoplevelser og dernæst i deres kompetencer og forestillinger. De mest værdifulde og varige læringsudbytter er sandsynligvis dem der tager længst tid om at vise sig i vurderinger af læringsresultater.

### 5.3 Konsekvenser

Hvis vi forudsætter at småskalaforsøg gennemført i udviklingen af efteruddannelsesprogrammer eller undervisningsmaterialer har godtgjort at disse initiativer kan give læring gennem undersøgelse, så retter opmærksomheden sig mod evalueringen af deres implementering i almindelige skoler og klasser. Som skitserne af evalueringer ovenfor har vist, er der mange faktorer der spiller ind når det skal afgøres i hvilket omfang et elevudbytte kan påvises. Vi kan inddele dem i to hovedgrupper: dem der vedrører den "uafhængige variable", dvs. omfanget af elevernes erfaring med undersøgelsesbaseret læring, og så dem der vedrører måling af den afhængige variabel, dvs. effekten på eleverne.

Hvad angår den uafhængige variabel, tyder eksisterende praksis på at seriøs opmærksomhed skal rettes mod at sikre at vi ved *hvad* resultaterne skyldes. Det betyder at der skal indsamles data om lærernes forståelse af reformen, om deres undervisningstilgange og om elevernes resultater. Alt for mange evalueringsundersøgelser (som fx Ruiz-Primo et al., 2002) har rapporteret at aktiviteterne i klasserummet kun trak på nogle få af de praksisser der er forbundet med undersøgelsesbaseret læring. I den slags tilfælde giver resultaterne ikke noget grundlag for en gyldig vurdering af værdien af programmet. De stærke beviser for skolelederens indflydelse indikerer at ledelsen skal være involveret i beslutninger om initiativer for at sikre en mere effektiv implementering.

Når det står klart at den uafhængige variabel er på plads, skal den være i drift i mindst et år før der foretages resultatmålinger. Denne tidsramme tilgodeser alle de andre påvirkninger af elevresultaterne og at det tager tid at etablere de ændringer i måder at lære på som undersøgelsesbaseret sigter efter.

Med hensyn til den afhængige variabel er der vigtige beslutninger at træffe om arten af de anvendte målemetoder uanset om de skal vurdere holdninger, præstationer, valg af videre undersøgelser osv. Som diskuteret tidligere skal disse målinger overbevise dem der skal informeres om evalueringen og/eller træffe beslutninger på baggrund af den. De skal også sættes ind over et stykke tid så spørgsmål om kort- hhv. langsigtede ændringer i læring kan besvares.

Evalueringdesignet skal gøre det muligt at sammenligne de observerede ændringer med hvad der ville være sket uden indgreb. Men størrelsen af eventuelle forskelle skal udtrykkes meningsfuldt; det er usandsynligt at procentuelle forskelle er til nogen hjælp – og statistisk signifikans er ikke ensbetydende med uddannelsesmæssig signifikans, især ikke når de involverede tal er store. Black & Wiliam (1998) fandt fx at det var med til at gøre indtryk på de politiske beslutningstagere at påpege at hvis formativ evaluering blev indført i hele landet, ville den potentielle gevinst have løftet England fra midten af de 41 lande i TIMSS-undersøgelsen til en placering i top-5.

Og endelig skal evalueringdesign tage hensyn til de dybtgående ændringer der skal ske i lærernes og elevernes tænkning. Kun hvis de ændringer får nok tid så de kan blive indlejret i praksis, kan evalueringerne forsyne de politiske beslutningstagere og praktikere med valide oplysninger om eleveffekten.

## Referencer

- Adey, P., Robertson, A. & Venville, G. (2002). Effects of a Cognitive Acceleration Programme on Year 1 Pupils. *British Journal of Educational Psychology*, 72, s. 1-25.
- Adey, P. & Shayer, M. (1990). Accelerating the Development of Formal Thinking in Middle and High School Students. *Journal of Research in Science Teaching*, 27(3), s. 267-285.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for Learning. Putting it into Practice*. Maidenhead: Open University Press.
- Black, P. & Wiliam, D. (1998). *Inside the Black Box*. School of Education, King's College London.
- Bransford, J., Brown, A.L. & Cocking, R.R (red.). *How People Learn*. Washington DC: National Academy Press.
- Butler, R. (1988). Enhancing and Undermining Intrinsic Motivation: The Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance. *British Journal of Educational Psychology*, 58, s. 1-14.
- Doubler, S. (1991). *Change in Elementary School Teachers' Practice in the United States*. Upubliceret PhD thesis. University of Liverpool.



- Gibson, H.L. & Chase, C. (2002). Longitudinal Impact of an Inquiry-Based Science Program on Middle School Students' Attitudes towards Science. *Science Education*, 86, s. 693-705.
- Gopnik, A., Meltzoff, A.N. & Kuhl, P.K. (1999). *The Scientist in the Crib*. New York: William Morrow.
- Hall, G.E. & Loucks, S.F. (1977). A Developmental Model for Determining Whether the Treatment Is Actually Implemented. *American Educational Research Journal*, 14(3), s. 263-276.
- Harlen, W. & Altobello, C. (2003). *An Investigation of Try Science Studied On-line and Face-to-Face*. Cambridge, MA: TERC.
- Harlen, W. & Osborne, R. (1985). A Model for Learning and Teaching Applied to Primary Science. *Journal of Curriculum Studies*, 17(2), s. 133-146.
- Hawking, S.W. (1988). *A Brief History of Time*. London: Bantam Press.
- Hord, S.M., Huling-Austin, L., Hall, G.E. & Rutherford, W. (1987). *Taking Charge of Change*. Alexandria, VA: ASCD.
- Keys, C.W. & Bryan, L.A. (2001). Co-constructing Inquiry-Based Science with Teachers: Essential Research for Lasting Reform. *Journal of Research in Science Teaching*, 38 (6), s. 631-645.
- National Research Council. (1996). *The National Science Education Standards*. Washington DC: National Academy Press.
- Piaget, J. (1929). *The Child's Construction of the World*. London: Routledge and Kegan Paul.
- Piaget, J. (1955). *The Child's Construction of the Reality*. London: Routledge and Kegan Paul.
- Piaget, J. (1956). *The Child's Construction of Space*. London: Routledge and Kegan Paul.
- Rankin, L. (1999) Lessons in a Pond: A Year-Long Inquiry Investigation. *Foundations 2*, Arlington, VA: NSF, s. 71-78.
- Robertson, A. (2004). Let's Think: Two Years On. *Primary Science Review*, 82, s. 4-7.
- Rudduck, J. & Kelly, P. (1976). *The Dissemination of Curriculum Development: Current Trends*. Slough: The National Foundation for Educational Research.
- Ruiz-Primo, M.A., Shavelson, R.J., Hamilton, L. & Klein, S. (2002). On the Evaluation of Systemic Science Education Reform: Searching for Instructional Sensitivity. *Journal of Research in Science Teaching*, 39(5), s. 369-393.
- Skinner, B.F. (1974). *About Behaviourism*. New York: Alfred A. Knopf.
- Supovitz, J.A. & Turner, H. M. (2000). The Effects of Professional Development on Science Teaching Practices and Classroom Culture. *Journal of Research in Science Teaching*, 37(9), s. 963-980.
- Vygotsky, L.S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- Windschitl, M. (2003). Inquiry Projects in Science Teacher Education: What Can Investigative Experiences Reveal About Teacher Thinking and Eventual Classroom Practice? *Science Teacher Education*, 87, s. 112-143.

Haymount Coach House, Bridgend, Duns, Berwickshire TD11 3DJ, Scotland, UK  
wynne@torphin.freeserve.co.uk