

Ledelsesperspektiver på skolens brug af data

Jørgen Søndergaard
Forskningsleder, SFI – Det Nationale Forskningscenter for Velfærd
Herluf Trolles Gade 11, 1052 København K, E-mail: js@sfi.dk

Marts 2016

1. Indledning

Formålet med denne artikel er at give et overblik over, hvordan skolens elevdata kan anvendes i arbejdet med at udvikle skolen, så eleverne lærer mere¹. Elevdata kan på en systematisk måde fortælle noget om elevernes læring – elevernes progression – og den information er af afgørende betydning for skolens bestræbelser på at skabe øget læring hos eleverne. Ikke sådan at forstå, at skolen ikke har viden om elevernes læring uden data. Undervisere danner sig naturligvis et billede af elevernes udbytte af undervisningen. Uden data er denne viden om eleverne i vidt omfang privatiseret og vanskelig at kommunikere til andre, der ikke selv har oplevet, hvordan de forskellige elever har udviklet viden og færdigheder under undervisningsforløbet. Elevdata giver mulighed for en større grad af systematik i vurderingerne af elevernes læring og en større præcision i kommunikationen om elevernes læring til andre på skolen.

Der er grund til at understrege, at data i denne sammenhæng er et bredt begreb og især at skolen selv har en betydelig indflydelse på, hvilke data den vil skabe og hvilke metoder, herunder evalueringsformer, den vil anvende. Skolen kan således udover de obligatoriske nationale test og trivselsmålinger, vælge at evaluere eleverne på mangfoldige andre måder, herunder f.eks. også indsamle information, hvor eleverne evaluerer undervisningen eller evaluere elevernes personlige udvikling og almene dannelse undervejs i skoleforløbet. I denne artikel anvendes begrebet læring i bredeste forstand, dvs. det omfatter ikke kun de faglige kompetencer, men også udvikling i personlige og sociale kompetencer, almen dannelse mm.

Artiklen vil særligt fokusere på ledelsesaspekter af brugen af data i skolen. Elevdata kan være af værdi til mange forskellige formål, som skolens ledere har (med)ansvar for, herunder f.eks.:

- Løbende vurderinger af udviklingen på skolen i forhold til politiske mål og i forhold til skolens egne mål
- Kvalitetssikring af undervisning
- Evaluering af undervisere, MUS og kompetenceudvikling
- Læringsmålsstyret undervisning
- Undervisernes selvevaluering, refleksion og samtale om udfordringer i undervisningen
- Undervisernes aktionslæring

¹ Der findes en omfattende udenlandsk litteratur om dette emne, se f.eks. Datnow and Pack (2014), Boudett et. al. (2013), Bambrick-Santoyo (2010) og Mandinach and Honey (2008). Man kan læse om brug af data i danske skoler i en rapport fra Danmarks Evalueringsinstitut (2014).

Hvor meget og hvordan den enkelte leder selv vil være involveret i brugen af data afhænger af lederens ansvarsområde og opgaver på skolen. Populært sagt: jo tættere på den daglige undervisning, lederen er, jo mere detaljerede data vil være relevant for vedkommende og omvendt, jo mere overordnet ledelsesopgaven er, jo større er værdien af data, der giver overblik over hele skolen. Skolelederen på en stor skole med mange elever og medarbejdere vil formentlig primært interessere sig for de store linjer, herunder ikke mindst udviklingen over tid i data, der siger noget overordnet om elevernes læring. Afdelingsledere, der har til opgave at vejlede og coache underviserne og deltage i drøftelser i lærerteams om aktionslæring, teamets selvevaluering af elevernes udbytte af undervisningsforløb osv., vil derimod have behov for at arbejde med detaljerede elevdata, så de kan deltage i analyse og diskussion af, f.eks. hvordan læringen kan øges hos de elever, der har haft lavt udbytte af den hidtidige undervisning.

Ledere har imidlertid også en anden opgave, der handler om at iscenesætte og tilrettelægge, hvordan skolens undervisere bruger evalueringer og evalueringsresultater og andre data i deres løbende arbejde med at tilrettelægge undervisning og reflektere over elevernes læringsudbytte.

2. Hvad fortæller test- og prøveresultater?

Inden vi går nærmere ind på disse forskellige ledelsesopgaver og måder at bruge data på, er der behov for at se på, hvad data egentlig fortæller. Lad os først se på en enkelt elevs resultat ved en enkelt prøve.

Hvis en elev f.eks. har fået resultatet "god" i de nationale test i læsning i 6. klasse, er det så sikkert, at "god" er et retvisende udtryk for elevens læsefærdigheder? Svaret er nej, og det er der en række grunde til². For det første kan prøven tilfældigvis have indeholdt elementer, som eleven er særlig dygtig til eller særlig dårlig til. For det andet spiller elevens motivation den pågældende dag en rolle for prøveresultatet. Var eleven særligt motiveret eller særligt demotiveret/uoplagt, så har det en positiv eller negativ indflydelse på prøveresultatet³. For det tredje findes der ingen test eller prøver, der måler f.eks. læsefærdigheder i deres fulde dybde og bredde. Derfor kan en bestemt prøve give såvel et overvurderet som et undervurderet billede af elevens "sande" læsefærdigheder.

² Der ses her bort fra illegitime grunde som f.eks. at eleven uretmæssigt har fået hjælp eller haft mulighed for at kigge hos sidemanden.

³ Hans Henrik Sivertsen på SFI har sammen med to kolleger undersøgt, om det betyder noget for resultater i de nationale test, hvornår på dagen, de tages. De påviser, at resultaterne er (lidt) dårligere jo senere på skoledagen testen er taget, samt at resultaterne er (lidt) bedre, hvis testen tages lige efter et stort frikvarter. Det underbygger, at veloplagthed/uoplagthed har indflydelse på testresultater, jf. Sivertsen et al. (2016)

Hvor stor er denne usikkerhed? Det er det svært at give et præcist billede af, men man kan få et indtryk ved f.eks. at se på den frivillige brug af nationale test. En del skoler bruger de frivillige test med nogle få ugers mellemrum, således at de samme elevers resultater ved to på hinanden følgende test kan sammenlignes, jf. Styrelsen for It og Læring (2016). Ser vi f.eks. på resultatet af læsetesten i 6. klasse opgjort på den femtrins skala, der hidtil har været anvendt til formidling af resultater fra nationale test, viser det sig at 60 pct. af eleverne er placeret på samme trin, når testen tages to gange med nogle ugers mellemrum, mens 12 pct. ligger et trin lavere i test nr. 2 end de gjorde i test nr. 1, og 26 pct. ligger et trin højere i test nr. 2 end i test nr. 1. Endelig er der 2 pct. der ligger to trin højere i test nr. 2. Det er forventeligt at nogle elever ligger højere ved den anden test, fordi de er blevet dygtigere i mellemtiden. Det viser sig da også, at andelen, der ligger højere i test nr. 2 er større jo længere tid, der går mellem test nr. 1 og test nr. 2. Undersøgelsen viser desuden, at der er en forholdsvis god overensstemmelse mellem resultater i nationale test og de samme elevers resultater ved folkeskolens afgangsprøver. F.eks. opnår 80 procent af de elever, der samlet opnår vurderingen 'rigtig god' i nationale test i læsning i 8. klasse, mindst karakteren 7 i dansk læsning i folkeskolens afgangsprøve i 9. klasse. Men der er altså 20 pct., der klarede sig væsentligt dårligere til afgangsprøven end i de nationale test i 8. klasse.

Der er med andre ord en vis sandsynlighed for, at elevens "sande" læsefærdigheder er enten noget dårligere eller noget bedre end det opnåede resultat ved én prøve. Det gælder ikke blot for resultater i nationale test – det gælder for enhver form for evaluering af elevens færdigheder – også den, den enkelte underviser selv laver. Hvordan kan man opnå et mere sikkert billede af elevens færdigheder? Svaret er flere opgaver og prøver! Hvis samme elev har deltaget i andre prøver af læsefærdigheder og/eller lavet skriftlige opgaver, der siger noget om elevens ordforråd og sprogforståelse, og resultaterne af disse ligger på omtrent samme niveau som det nationale testresultat, - evt. nogle gange lidt bedre og andre gange lidt dårligere - så øger det i høj grad sandsynligheden for, at elevens "sande" niveau er "god". Har eleven derimod klart dårligere resultater i de fleste andre opgaver og prøver, så øges sandsynligheden for, at resultatet "god" i de nationale test overvurderer elevens "sande" læsefærdigheder. Bemærk at den øgede sikkerhed i vurderingen ikke nødvendigvis består i at gentage de samme opgaver eller prøver én gang til. Det kan lige så godt bestå i, at eleverne laver nogle opgaver, som underviserne selv har lavet og bedømmer med udgangspunkt i en rubric eller bedømmelsesvejledning.

Pointen ved at inddrage resultater fra flere test og prøver er, at det tilfældige element i hver enkelt prøve tenderer at udligne hinanden, når flere prøver inddrages. Sandsynligheden for

f.eks. at være "særlig motiveret og veloplagt", hver gang der skal laves en opgave eller prøve, er meget mindre end sandsynligheden for, at dette sker en gang imellem.

Lad os herefter se på en klasses elevresultater. Hvis en 6. klasse i nationale test i læsning har opnået resultater, så f.eks. kun 40 pct. af eleverne har fået mindst resultatet "god" er det så – i lyset af den nationale målsætning om, at 80 pct. af eleverne skal opnå dette - et sikkert bevis for på, at underviseren ikke har gjort sit arbejde godt? Nej naturligvis er det ikke et bevis – der er ikke nogen enkel og sikker sammenhæng mellem en lærers undervisning og elevernes resultater. På den anden side er det en indikation for, at der er noget, der bør undersøges nærmere. Dels er det naturligt at undersøge: Hvordan er klassens resultater i sammenligning med den socioøkonomiske reference? Hvordan er klassens resultater i andre fag og prøver? Er der særligt mange udfordrede elever med særlige behov i netop denne klasse. Har klassen et særligt højt fravær? Dels er det naturligt at se på, hvad der ellers er af viden om underviserens elevresultater. Har samme underviser haft tilsvarende lave resultater med andre klasser? Igen er pointen ved at inddrage flere sæt af data, at det tilfældige element i den enkelte prøve bliver udlignet. Har en underviser eksempelvis svagere elevresultater end andre undervisere på skolen i f.eks. 8 eller 9 ud af 10 nationale test, er det et indicium for, at der kan være nogle mangler ved den pågældendes undervisning. Har den pågældende derimod i gennemsnit resultater på samme niveau som andre lærere på skolen, er dårlige resultater i en enkelt prøve, næppe et udtryk for generelle mangler ved den pågældendes undervisning, men snarere for tilfældige faktorer.

Vi ved, at elevernes sociale baggrund spiller en betydelig rolle for deres karakterer og testresultater i skolen. Derfor kan der være andre grunde end forskelle i undervisningen til at klasseresultater varierer fra år til år – nemlig at elevsammensætningen i klassen er forskellig fra år til år. Det kan der til dels tages højde for ved at anvende den socioøkonomiske reference, som udtrykker hvad elever med samme sociale baggrund på landsplan opnår i den samme prøve eller test. Men til at evaluere læringsudbyttet af skolens undervisning er det bedre at måle elevernes progression frem for at måle deres niveau. Man kan skelne mellem absolut og relativ progression. Absolut progression kan kun måles ved at bruge samme prøve og samme måleskala på to forskellige tidspunkter⁴. Relativ progression kan derimod måles ved at bruge to forskellige prøver på to forskellige tidspunkter. Fordi prøverne er forskellige siger

⁴ Brug af samme prøve eller prøvetype er ofte kun muligt med en forholdsvis kort tidsafstand mellem prøverne. Der er dog udviklet en metode - kaldet Beregneren -, som en del kommuners skoler har taget i anvendelse, som gør det muligt at udlede et absolut progressionsmål fra de nationale testresultater over såvel korte perioder som over næsten hele skoleforløbet. Se evt. <http://www.nordicmetrics.com/beregneren/>

de ikke noget om den absolutte progression, men de kan afdække den relative progression – altså f.eks. om en elev har større eller mindre fremgang end klassens gennemsnit.

Selvom progressionsmål fjerner noget af den sociale baggrunds betydning og noget af den individuelle faktor i evalueringen, er der dog ikke tale om fuldstændig sikre målinger. Så det gælder fortsat, at en enkelt måling af en elevs progression både kan overvurdere og undervurdere, hvor meget eleven rent faktisk har lært i den pågældende periode. Men de er bedre egnede til at vurdere elevernes udbytte af undervisningen end de absolutte testresultater og derfor mere relevante som pædagogisk redskab for underviserne.

Endelig kan vi kigge på skolens samlede resultater. De nationale mål er som nævnt, at mindst 80 pct. af eleverne skal opnå mindst resultatet "god" i de nationale test i læsning og matematik. For den øverste ledelse er det centrale spørgsmål om andelen, der når det nationale mål øges år for år eller ej. Hvis man kigger på de enkelte elevers resultater, kan man imidlertid ikke "se skoven for bare træer". Detaljen fjerner overblikket, og med tanke på, at den enkelte elevs resultat er ganske usikkert, jf. ovenfor, kan man måske komme i tvivl om, om man overhovedet kan udlede noget af testresultaterne. Men det kan man! Hvis man hvert år samler alle testresultater i et samlet tal for hvor stor en procentdel af eleverne i de klasser, der er omfattet af nationale test, der har opnået et resultat på mindst "god", og sammenligner med foregående års resultater opgjort på samme måde, har man en måling af *udviklingen* på skolen med en ret lille usikkerhedsmargin.

Læg mærke til, at det vi har betonet i dette afsnit, er to ting:

- 1) Det tilfældige element, som er ganske betydeligt for den enkelte elevs resultat i en enkelt prøve, bliver meget mindre, når der inddrages data fra flere opgaver eller prøver, fordi det tilfældige element i de enkelte prøver eller evalueringer tenderer imod at udligne hinanden. Så flere prøveresultater for samme elev reducerer det tilfældige element. Tilsvarende er en classes gennemsnit langt mindre påvirket af tilfældigheder end hver enkelt elevs resultat.
- 2) Find et relevant sammenligningsgrundlag. Det kan f.eks. være samme elevs resultater i andre prøver, samme undervisers resultater i andre prøver eller andre undervisers resultater i samme prøve eller i samme klasse, og det kan f.eks. være skolens resultater de foregående år. På den måde får man et bedre grundlag for at vurdere det aktuelle resultat. Dog er der et tilfældigt element, når man sammenligner forskellige klasser på samme klassetrin, som følge af "tilfældige" forskelle i elevsammensætningen i klasserne. Dette sidste tilfældige element er igen mindre, når man sammenligner hele

årgange på en stor skole med hinanden, end når man på en ensporet skole sammenligner to årgange med hinanden.

3. Iscenesættelse af undervisernes brug af evalueringer og elevresultater

Hvis skolens data skal anvendes til f.eks. målstyret undervisning og til aktionslæring, analyser og refleksioner over elevernes læring og over sammenhænge mellem undervisning og læring, kræver det, at ledelsen iscenesætter dette. Hvad vil det sige? Det konkrete svar vil selvfølgelig afhænge af den eksisterende praksis på skolen, men typisk vil der være behov for en ledelsesindsats på i hvert fald disse fire områder:

- Politik og retningslinjer for skolens brug af test, opgaver og evalueringer
- Organisationens kultur
- Undervisernes kompetencer til dataanalyse
- Datamanagement

3.1. Politik og retningslinjer for skolens brug af test og evalueringer

Dette punkt handler om, hvilke data og informationer, man ønsker at bruge i skolen. Skolens politik og retningslinjer for test og evalueringer bør fastlægge, hvilke test og opgaver, der anvendes i hvilke fag og på hvilke klassetrin, hvordan evalueringen af eleverne foregår, og hvornår i skoleåret, opgaverne/prøverne mv. finder sted. Politikken kan med fordel også beskrive hvordan evalueringerne anvendes til feed back til eleverne og til at evaluere undervisningens betydning for elevernes læring.

Behovet for politik og retningslinjer hænger sammen med, at systematik er helt afgørende for værdien af de data, der skabes. Hvis f.eks. den enkelte lærer selv kan beslutte, hvordan, hvor ofte og hvornår i skoleåret elevernes læring evalueres, er der ingen mulighed for at sammenligne med andre evalueringer, fordi resultaterne pr. definition er usammenlignelige.

Det centrale her er altså systematik. OECD-reviewet om evalueringskulturen i Folkeskolen, jf. OECD (2011), fandt ret stor variation i praksis på dette område mellem skoler og kommuner, og kom med en lang række anbefalinger til at styrke brugen af evalueringer i skolen, herunder anbefales en stærkere systematisk struktur for og fastlagte procedurer i brugen af evalueringer, bedømmelser og feed back mm. Pointen er, at elevevalueringer uden systematik mht. indhold, placering i skoleåret, rette/vurderingsvejledning, har meget lille værdi til analyse af undervisning og elevs læring, jf. også afsnit 2. Hvordan systematikken skal se ud, afhænger naturligvis af, hvad skolen vil med sine data. Derfor bør det være en del af skolens politik og retningslinjer, hvordan undervisere, teams m. fl. anvender skolens data til analyse,

refleksion og evt. aktionslæring. Det kan naturligvis også omfatte andre informationer end elevevalueringer og testresultater som f.eks. elevfravær, lærerfravær/vikartimer i klassen og trivselsmålinger. Desuden kan det med fordel omfatte en systematiseret feed back fra eleverne, hvor der indsamles oplysninger om deres vurdering af undervisningen. Og det kan også omfatte en systematiseret opsamling af iagttagelser fra observationer af undervisningen.

I forlængelse heraf, er der ét yderligere emne, der skal tages stilling til, nemlig hvad skal evalueres? I forlængelse af skolereformens fokus på læringsmål, foregår der jo en til tider heftig offentlig debat om kompetencer versus dannelse, hvor især nogle kritikere af skolereformen beskylder denne for at være præget af instrumentel tænkning og kompetencefokusering på bekostning af skolens rolle i forhold til elevernes dannelse.

I forhold til brug af elevdata bør det give anledning til at overveje, om man "får mere af det, man måler" og "mindre af det, man ikke måler"? Det er der formentlig en risiko for. Derfor er det en god ide, at skolen overvejer, hvordan man i sit brug af evalueringer kan opnå en balance mellem de forskellige opgaver skolen har ifølge folkeskolelovens formålsparagraf, som både handler om faglige kompetencer, alsidig udvikling og dannelse. Det kræver naturligvis, at skolen fastlægger, hvad man forstår ved, at elevernes opnår alsidig udvikling og dannelse, samt hvad der vil være "tegn på læring" i forhold hertil. Har man det, er det også muligt at tilrettelægge opgaver og undervisningsforløb, der er egnede til at danne grundlag for at vurdere eleverne i forhold til sådanne brede mål for skolens virke.

Et eksempel på opstilling af trinmål for elevernes personlige og sociale udvikling i danske skoler findes i Gladsaxe Kommune (2004). Når skolen udvikler egne opgaver og evalueringer er det generelt en god ide, at underviserne udarbejder en eller flere "rubrics" – et skema, der både definerer, hvilke elementer/dimensioner, der indgår i evalueringen, og giver en karakteristik af de forskellige niveauer i den anvendte vurderingsskala. En sådan rubric udgør dermed samtidig en form for vurderingsvejledning, som bidrager til, at forskellige undervisere har et fælles redskab til evalueringen og dermed til en vis homogenitet i evalueringer på tværs af undervisere og klasser, hvilket er af væsentlig betydning for værdien af evalueringerne og de data, der fremkommer. Inspiration til "rubrics" for f.eks. samarbejdsevne fra udenlandske skoler er let at finde på nettet: søg på "collaboration rubric".

3.2 Organisationens kultur

Det netop omtalte OECD-review fandt at der var sket en positiv udvikling i skolernes brug af evalueringer fra 2004 til 2011, men pegede også på, at der mange steder fortsat er en ret svag evalueringskultur og fravær af traditioner for at bruge skolens elevdata til analyser og

vurderinger af skolens undervisning. Det peger på, at der er en vigtig ledelsesopgave i at udvikle skolens kultur, så analyse, refleksion, aktionslæring osv. bliver en mere integreret del af skolens praksis, som det f.eks. er på skoler med stærke professionelle læringsfællesskaber.

At udvikle en ændret kultur er en længerevarende ledelsesopgave. Det er afgørende, at alle i ledelsesteamet er enige om og selv efterlever de normer og værdier, det vil fremme. Det indebærer at lederne selv anvender data og er gode rollemodeller for, hvordan det kan gøres på en respektfuld og tillidsskabende måde. Det indebærer også, at ledelsen synliggør, anerkender og belønner den adfærd, der ønskes fremmet. Og overfører ressourcer til de aktiviteter, der ønskes fremmet - her undervisere og teams, der går foran mht. analyse, refleksion, aktionslæring mv., hvor de bruger skolens data.

Endelig kan der med fordel tænkes i processer, der involverer medarbejderne i de mange beslutninger, der skal træffes for at iscenesætte brugen af test- og evalueringsresultater og andre data i den daglige praksis. (Se artikel om involverende beslutningsproces til komplekse udfordringer, Søndergaard (2015)). Undervisere vil have forskellige erfaringer med og holdninger til brug af data som redskab, og det kan spille en vigtig rolle for, hvordan især de skeptiske medarbejdere forholder sig til at analysere egen praksis med brug af data, at de ikke blot oplever, at der lyttes til deres bekymringer, men at det også kan ses i de konkrete beslutninger. En involverende proces indebærer ikke, at ledelsen pantsætter sin ledelsesret, men at man har en fleksibilitet i tilgangen til udformning af den praksis, man ønsker at fremme.

3.3 Undervisernes kompetencer til dataanalyse

Der kan selvsagt være flere grunde til, at undervisere ikke anvender skolens elevdata til at analysere deres undervisning og elevernes læring. Men en oplagt grund er, at det ikke har været en del af den lærer- eller pædagoguddannelse, de pågældende har taget. Derfor vil en god iscenesættelse af skolens brug af data også indeholde en plan for opbygning af medarbejdernes kompetencer på dette område. Det kan med fordel ske ved i en periode at facilitere f.eks. teammøder, der har analyse og brug af skolens data på dagsordenen, så kompetenceopbygningen får et praksisnært indhold. Det kan også omfatte et udviklingsforløb for teamkoordinatorer og skolens vejledere, så de klædes bedre på til at tilrettelægge arbejdet i teamene.

3.4 Datamanagement

Endelig bør det overvejes, hvordan skolens elevdata kan gøres tilgængelige for skolens undervisere på en brugervenlig måde. Det har stor betydning for brugerne, at data gøres

tilgængelige med et brugervenligt software. Brugervenligheden kan desuden understøttes ved at nogle af skolens ansatte får det som en opgave at "processe" elevresultater og andre elevdata til formater, der er brugervenlige og/eller, at skolen får bistand fra kommunens konsulenter eller køber sig konsulentbistand udefra til at tilrettelægge databaser, dataudtræk, grafisk præsentation af resultater mv.. Det er meget banalt, men hvis det er besværligt at finde og/eller at forstå data, så bliver de ikke brugt til ret meget af ret mange.

4. Ledere som brugere af skolens data

4.1 Løbende vurdering af udviklingen på skolen

Når man vil overvåge, om udviklingen på skolen viser tegn på fremgang, kan man med fordel starte "så højt oppe med helikopteren" som muligt. Tænk på en 3-sporet skole. Der vil der hvert år være 12 klasseresultater fra de nationale test i læsning og 9 i matematik. Nogle klasseresultater vil være bedre end året før, nogle vil være på samme niveau og nogle vil være dårligere. Sådan vil det være hvert år – uanset om skolens resultater bliver bedre og bedre, er omtrent uændrede eller går tilbage. Og noget af dette kan skyldes tilfældigheder, der ikke siger noget om kvaliteten af undervisningen. Disse tilfældigheder ophæver imidlertid i store træk hinanden, når man lægger alle tallene sammen og undersøger om andelen med mindst "god" går op eller ned.

Lad os se på et tænkt eksempel. Resultatet – andel elever, der når mindst resultatet "god" i nationale test i læsning og matematik - på to skoler A og B kunne f.eks. se således ud:

År	2014	2015	2016	2017	2018	2019
Skole A	60	64	63	66	68	70
Skole B	60	64	61	62	59	62

Man kan ikke uddrage noget væsentligt ved at se på et enkelt års resultater. Alle skoler kommer en gang imellem ud for et år, hvor resultaterne er bedre end i tidligere år og alle kommer en gang imellem ud for et år, hvor resultaterne er dårligere end i de foregående år. Det interessante her er derfor, om der er en tendens til bedre resultater over tid eller ej. Resultaterne for skole A bliver tydeligvis bedre og bedre, selvom der er et enkelt år med et lille tilbagefald. Resultaterne for skole B bliver tydeligvis ikke bedre over tid, idet resultaterne svinger op og ned fra år til år. Skole B har med andre ord nogle udfordringer, og næste skridt i analysen kunne passende være at undersøge, om det både gælder for læsning og matematik og om det gør sig gældende på alle klassetrin. Altså især når udviklingen ikke er som ønsket, kan

man med fordel sænke helikopterhøjden og inddrage lidt flere detaljer i vurderingen af skolens udvikling.

For den leder, der har ansvaret for en bestemt del af skolen – f.eks. udskoling, hvor nationale test kun omfatter 8. klasse – kan resultaterne naturligvis være et fingerpeg om udviklingen i undervisningens kvalitet, selvom det tilfældige element her er større end for skolen som helhed - især når der er få spor på skolen. Det enkelte års resultat kan svinge noget uden at det behøver være udtryk for tilsvarende variation i undervisningens kvalitet. Derfor har man som leder brug for at følge udviklingen i resultaterne over længere tid for at få en mere sikker fornemmelse af, om udskoling trods årlige udsving er på vej frem eller ej. Hvis gennemsnittet af de sidste tre års resultater er klart bedre end det tilsvarende gennemsnit for 4-6 år siden, så er dette et mere sikkert tegn på øget læring hos eleverne, end blot at resultaterne i år er bedre end sidste år.

Endelig er der grund til at minde om, at det ikke kun handler om testresultater. Udvikling i elevfravær og elevtrivsel er også vigtige indikatorer at overvåge – især udviklingen i den andel, der ikke trives helt godt i skolen og/eller har stort fravær.

4.2 Evaluering af undervisernes resultater, MUS og kompetenceudvikling

At anvende skolens data som et led i vurderingen af underviserne er meget lidt udbredt i Danmark, jf. OECD (2011). Sådanne data kan selvsagt heller ikke stå alene, men de kan på den anden side udgøre en del af grundlaget for udviklingssamtalerne og planlægning af kompetenceudvikling for skolens undervisere. Det indebærer, at skolen samler data for de klasser, hver enkelt underviser har undervist. Både ved at sammenligne over tid og ved at sammenligne på tværs af undervisere kan både leder og underviser på denne måde få et bidrag til et billede af den enkelte undervisers bidrag til elevers læring, som naturligvis skal holdes sammen med den information, der fås fra andre kilder - f.eks. observationer af undervisningen.

Umiddelbart vil mange vælge at se på elevernes gennemsnit, men det kan være mindst lige så relevant at se på spredningen i elevresultater eller endnu bedre på elevernes progression, når lærerens styrker og svagheder skal afdækkes. Man kan sagtens have en underviser, der ikke adskiller sig målbart ud fra elevernes gennemsnit, men som skaber gode resultater for de svageste elever i klassen, ligesom man kan have en underviser, der scorer højt målt på klassens gennemsnit, men det sker på bekostning af de svageste elever.

4.3 Kvalitetssikring af undervisning

Ledernes involvering i kvalitetssikring af undervisning kan ske på flere niveauer:

- Tilrettelægge strukturer, retningslinjer og processer, der skaber kvalitetssikring. Kan bl.a. omfatte retningslinjer for læringsmålsstyret undervisning og skolens brug af evalueringer og evalueringsresultater, jf. afsnit 4 nedenfor.
- Overvåge at retningslinjer overholdes, processer gennemføres og at alle elever gør gode fremskridt. Indebærer bl.a. overvågning af test-og evalueringsresultater.
- Selv deltage aktivt i udvalgte processer som f.eks. klassekonferencer og observation af undervisning.

Den centrale pointe her er, at skolen gennem valg af strukturer, retningslinjer og processer kan skabe kvalitetssikring, uden at ledere nødvendigvis behøver selv være aktive udførere af kvalitetssikringen. Vælger skolen en høj grad af gennemsigtighed, hvor mange har let adgang til data, kan det i sig selv have en positiv kvalitetssikringseffekt. Herudover kan kvalitetssikring være noget alle skolens undervisere medvirker til, og hvor ledelsesopgaven primært består i at iscenesætte og overvåge på et mere overordnet niveau. Men i nogle situationer vil ledere også selv være aktive deltagere i kvalitetssikringen f.eks. gennem deltagelse i klassekonferencer, hvor de sammen med underviserne og de relevante vejledere fortolker elevresultater og aftaler, hvordan underviserne skal forsøge at skabe øget progression på områder eller for udvalgte elever, hvor resultaterne ikke er tilfredsstillende. Ledelsesopgaven ved klassekonferencer består i høj grad i at holde fokus på behovet for at "gøre noget i undervisningen", hvis elevernes resultater tyder på, at der er faglige områder med lav progression, eller at udvalgte elever har haft et svagt læringsudbytte. Det er omsætningen fra analyse af data til handling, der er det springende punkt. For at mestre den ledelsesopgave kræves en vis træning i at fortolke testresultater og evalueringer af elevernes opgaver og en evne til at skabe en åben dialog om de udfordringer, der tegner sig for undervisningen i den pågældende klasse. Vivianne Robinsons "Open to learning conversations" kan være en god inspiration hertil, jf. Robinson (2014).

4.4 Undervisning med udgangspunkt i læringsmål

Undervisning med udgangspunkt i læringsmål kan lyde ret mekanisk, men det behøver det ikke at være. Tankegangen bag er jo, at undervisningen skal tilrettelægges ud fra en bevidst stillingtagen til, hvad eleverne skal tilstræbe at lære i et undervisningsforløb, og hvordan undervisningen kan give alle elever de bedste betingelser for at opnå dette. Det er på en vis måde et skift i "mindset" fra, at undervisning handler om at gennemgå/arbejde med et bestemt stof/pensum, til at indholdet i undervisningen tænkes i forhold til, hvad der skal til, for at eleverne får de bedste betingelser for at nå læringsmålene i tilfredsstillende grad. Hvad er så tilfredsstillende grad? I nogle situationer kan målet sikkert være det samme for alle i

klassen, men i andre tilfælde er det ikke realistisk, at alle elever kan nå lige langt, og i disse situationer er der behov for differentierede mål (sikkert ikke 24 forskellige mål i en klasse med 24, for eleverne vil sandsynligvis falde i et mindre antal grupper mht. forudsætninger)⁵. Bemærk differentieringen handler ikke om individuelle læringsmål i betydningen, at eleverne skal lære noget helt forskelligt, men om en graduering af hvor langt, den enkelte kan forventes at nå givet elevens startforudsætninger. Denne form for differentiering er væsentligt for at kunne tilrettelægge en undervisning, der skaber læringsmuligheder for elever med forskellige forudsætninger.

Ministeriet for børn, undervisning og ligestilling har udgivet et inspirationshæfte om anvendelse af læringsmål i skolen, som bygger på, at valg af læringsmål for undervisningsforløbet, valg af undervisningsaktiviteter, tegn på læring og evaluering hænger tæt sammen i alle faser af undervisningen. I det følgende vil vi alene beskrive, hvordan evalueringer og data kan bringes i anvendelse inden for denne ramme.

Ved starten af et nyt undervisningsforløb er der behov for at kende elevernes faglige niveau eller forudsætninger i forhold til klassens læringsmål for det nye forløb (nogle gange også kaldet en før-test). Det kan i nogle tilfælde bestå i resultaterne fra en prøve, en opgave eller nationale test taget i slutningen af foregående forløb. I andre tilfælde vil det være mest hensigtsmæssigt at bruge en prøve eller opgave, der specifikt siger noget om elevernes forudsætninger for det, klassen nu skal i gang med.

Denne indledende evaluering kan give et bedre grundlag for at vurdere behovet for differentiering i undervisningen. Er der væsentlige forskelle mellem eleverne mht. forudsætninger, er der som regel behov for at tænke i differentierede læringsmål, som på den ene side indebærer, at alle ikke bliver lige dygtige, men på den anden side sikrer, at alle elever lærer noget i forløbet. Denne form for differentiering kan f.eks. tage udgangspunkt i Blooms taksonomi, SOLO taxonomien eller lignende gradueringer af læringsudbytte, så læringsmålene differentieres svarende til de forskellige trin i taksonomien, hvilket samtidig bidrager til, at der vil være udfordringer for alle elever uanset deres initiale forudsætninger. En veltilrettelagt undervisning vil naturligvis altid bygge på en forventning om, at alle elever gør fremskridt. Uanset hvordan læringsmålene differentieres, udtrykker de en forventet progression for den enkelte elev. Måske er der tillige behov for, at de forskellige elevs vej til det endelige læringsmål overvejes. Nogen skal måske have vejen delt ind i mange etaper med delmål, så

⁵ Folkeskoleloven indeholder alene et krav om individuelle læringsmål i elevplanerne, men ikke for hvert enkelt undervisningsforløb. Det er således skolens opgave at vurdere behovet for differentiering i de enkelte undervisningsforløb.

udfordringen virker overkommelig, mens andre kan håndtere et læringsmål, der ligger et godt stykke fra elevens aktuelle niveau.

Det er desuden en fordel sammen med fastlæggelse af "tegn på læring" at overveje, hvilke opgaver/prøver, der er egnede til at evaluere elevernes progression undervejs (hvis der er tale om et længere undervisningsforløb), samt ved afslutningen, hvor evalueringen skal bidrage til at vurdere, om de enkelte elever har opnået deres forventede udbytte af undervisningen, jf. afsnittet 3.1 om politik og retningslinjer for skolens brug af test og evalueringer. Formålet med at evaluere såvel løbende som ved afslutningen af længerevarende forløb er at opsamle viden om elevernes progression, så undervisningen evt. kan justeres undervejs, hvis der er tegn på at (nogle) elever har et utilfredsstillende udbytte.

Det er ikke hensigten her at komme nærmere ind på, hvordan undervisningen kan tilrettelægges i undervisningsforløb med differentierede læringsmål. Et inspirerende forslag til, hvordan dette kan gøres, findes i Griffith og Burns (2014) under overskriften "Teaching Backwards".

Konceptet "Assesment for learning", som først blev lanceret af Black og Wiliam (1996), er på mange måder tænkt som en evalueringstilgang til læringsmålsbaseret undervisning. Denne evalueringstilgang til udvikling af undervisning og undervisningsdifferentiering findes efterhånden i mange udgaver. Et eksempel er The Teaching-Learning Critical Pathway (TLCP), som bygger på Fullan et. al. (2006). TLCP er udtænkt som tilgang i et professionelt læringsfællesskab. (Læs om professionelle læringsfællesskaber i Søndergaard (2016)).

Et af værktøjerne i TLCP er klassens datavæg, som "hænger" på væggen i det lokale, hvor klassens lærerteam holder sine møder. Nedenfor er vist et eksempel på en sådan datavæg for læsefærdigheder i en tredje klasse i en skole i Ontario. Cirklen viser hver elevs resultat i en prøve ved begyndelsen af forløbet. I eksemplet opnåede 38 per cent af eleverne niveau 3 eller 4. Firkanten viser underviserens forventninger til hver elevs resultat ved den afsluttende prøve i undervisningsforløbet. I eksemplet er forventningen at 67 pct. af eleverne vil nå niveau 3 eller 4. Tankegangen er, at dette billede af elevernes niveau ved starten og underviserens forventninger til hver enkelt elevs resultat ved slutningen danner grundlag for underviserens tilrettelæggelse af undervisningen. Trekanten viser så hver enkelt elevs faktiske resultat ved afslutningen af forløbet, hvor 81 pct. i det viste eksempel nåede niveau 3 eller 4, dvs. væsentligt flere end forventet ved starten af forløbet.

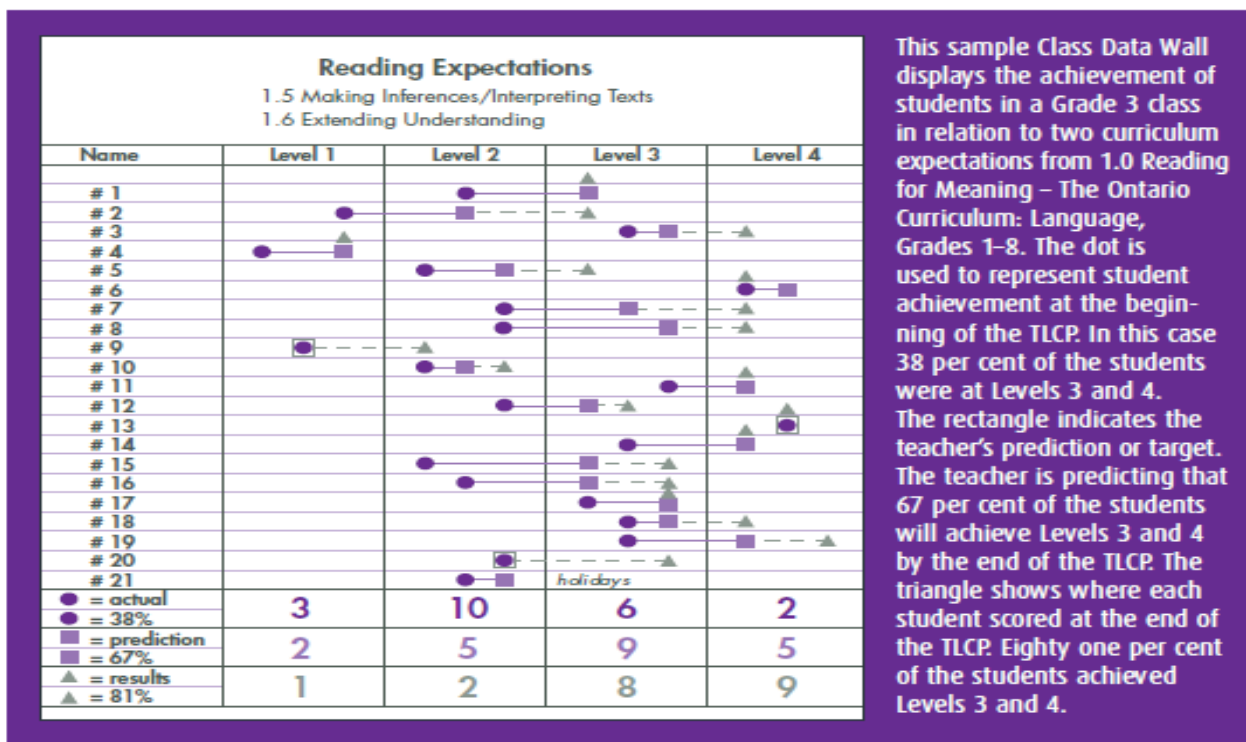


Figure 1 Class Data Wall

Kilde: The Literacy and Numeracy Secretariat Capacity Building Series: The Teaching-Learning Critical Pathways.

Ideen med at have en sådan datavæg er, at den understøtter en løbende samtale mellem klassens undervisere om, hvordan undervisningen kan tilrettelægges, så den understøtter alle elevers mulighed for læring. En datavæg som det viste eksempel fra Ontario vil let kunne skabes i danske skoler. Det stiller, som vi har været inde på, nogle krav til, hvordan skolen evaluerer eleverne, men der er stor fleksibilitet mht. valg af indhold og metode. Datavægen kan således f.eks. handle om personlige og sociale kompetencer, hvis det er det, der især er i fokus i en klasse.

Ledelsesopgaven i forhold til undervisning med udgangspunkt i læringsmål omfatter dels en række forhold på et overordnet niveau så som skolens brug af test og evalueringer, software til understøttelse af læringsmål i undervisning mv., dels samtaler med undervisere og teams om f.eks. ambitionsniveauet og den konkrete udformning af læringsmål, behovet for differentierede læringsmål, og hvilken betydning differentierede læringsmål skal have for undervisningens form og indhold. I forhold til brugen af elevresultater, er den vigtigste lederkompetence i denne sammenhæng fortolkning af elevdata i forhold til elevernes forudsætninger og ambitiøse, men realistiske læringsmål.

4.5 Undervisernes aktionslæring, selvevaluering og refleksion

Hvilken glæde kan den enkelte underviser eller det enkelte team have af elevdata? Først og fremmest er evalueringresultater ikke kun værdifulde for eleverne (når de ledsages af konstruktiv feedback), de er også værdifulde for undervisere, der er optaget af at udforske, hvilken betydning undervisningen – valg af læringsmål, undervisningsmaterialer, undervisningsformer, feedback, opgaveformer osv. – har for forskellige elevers læring. Det kan være i systematisk form som f.eks. aktionslæring eller mere uformelt i form af refleksioner og samtaler med kollegaer. I alle tilfælde drejer det sig om at finde ud af, om der sker noget med de forskellige elevers læring, når der introduceres en eller anden form for forandring i undervisningen.

De data, der er behov for til sådanne undersøgelser af egen praksis er ikke meget forskellige fra de data underviserne har behov for til undervisning med udgangspunkt i læringsmål. Med én vigtig undtagelse: man har behov for at sammenligne elever, der har deltaget i forskellige undervisningsformer. Det kan f.eks. være med data fra samme klassetrin i tidligere år. På den måde kan man undersøge, om de undervisningsformer, der er valgt i år, ser ud til at føre til mere eller mindre læring end de undervisningsformer, der blev anvendt i tidligere år. Det gælder især, hvis man anvender målinger af elevernes progression i løbet af en bestemt periode, fordi elevernes progression er mindre følsom overfor forskelle i elevforudsætninger end elevernes absolutte niveau.

5. Afslutning

Data kan være en vigtig ressource til mange forskellige formål. Styrken ved en databaseret tilgang i skolens praksis er den systematik og gennemsigtighed, der skabes. Men data er aldrig hele sandheden. De skal fortolkes af indsigtfulde professionelle, der også har en masse ikke-formaliseret viden om eleverne og om den gennemførte undervisning.

Litteratur

Bambrick-Santoyo, P. (2010), *Driven by Data: A Practical Guide to Improve Instruction*, John Wiley&Sons

Black, P. & D. Wiliam (1998), *Assessment and Classroom Learning*. *Assessment in Education* 5(1) pp. 7–71.

Boudett, K.P., E.A. City and R.J. Murnane (eds.) (2013), *Data Wise, Revised and Expanded Edition. A Step-by-Step Guide to Using Assessment Results to Improve Teaching and Learning*. Harvard Education Publishing Group

Bambrick-Santoyo, P. (2010), *Driven by Data: A Practical Guide to Improve Instruction*, Jossey-Bass.

Danmarks Evalueringsinstitut (2014), *Et bevidst blik på alle eleveres læring. Systematisk brug af data til udvikling af undervisningen*.

Datnow, A. and V. Park (2014), *Data-Driven Leadership*, Jossey-Bass

Fullan, M., Hill, P., & Crévola, C. (2006), *Breakthrough*. Thousand Oaks, CA Corwin Press.

Gladsaxe Kommune (2004), *Trinmål for børnenes alsidige, sociale og personlige udvikling*
http://www.gladsaxe.dk/Files/Filer/dokumenter/byraad%20og%20kommune/Informationsafdelingen/Trinmaal_for_boerrenes_alsidige_sociale_og_personlige_udvikling.pdf

Griffith, A. and M. Burns (2014), *Outstanding Teaching: Teaching Backwards*, Crown House Publishing.

Mandinach E.B. and M. Honey, (2008), *Data-Driven School Improvement: Linking Data and Learning*, Teachers College Columbia University

Ministeriet for børn, undervisning og ligestilling (2016) *Læringsmål – Inspiration til arbejdet med læringsmål i undervisningen*,
<http://www.emu.dk/sites/default/files/Inspiration%20til%20arbejdet%20med%20l%C3%A6ringsm%C3%A5l%20i%20undervisningen.pdf>

OECD (2011) *OECD Reviews of Evaluation and Assessment in Education DENMARK*, Claire Shewbridge, Eunice Jang, Peter Matthews and Paulo Santiago
<http://www.oecd.org/denmark/47696663.pdf>.

Robinson V. (2014), Open to learning conversations,
https://www.youtube.com/watch?v=0_I5-HKIR1s

Sievertsen H. H. et al. (2016), *Cognitive fatigue influences students' performance on standardized tests*, PNAS Early Edition,
<http://www.pnas.org/content/early/2016/02/09/1516947113.full.pdf?with-ds=yes>.

Styrelsen for It og Læring (2016), *Undersøgelse af de nationale tests reliabilitet*,
[http://www.uvm.dk/Aktuelt/~UVM-DK/Content/News/Udd/Folke/2016/Mar/160303-Ny-undersogelse-viser-sammenhaeng-mellem-resultater-ved-to-test-i-traek](http://www.uvm.dk/Aktuelt/~/UVM-DK/Content/News/Udd/Folke/2016/Mar/160303-Ny-undersogelse-viser-sammenhaeng-mellem-resultater-ved-to-test-i-traek)

Søndergaard, Jørgen (2015), *Involverende beslutningsproces til komplekse udfordringer*,
<http://www.emu.dk/modul/introduktion-til-beslutningsproces-komplekse-udfordringer>

Søndergaard, J. (2016), *Professionelle læringsfællesskaber i skolen: Hvad, Hvorfor og Hvordan*,
<http://www.emu.dk/modul/introduktion-til-professionelle-l%C3%A6ringsf%C3%A6llesskaber>