

## Ti svage elever og en fyring

I New York bliver elevers testresultater brugt til at evaluere lærere med. Det sker på en måde, som både er misvisende, og har en række utilsigtede konsekvenser.

Et af de mere kontroversielle skolepolitiske tiltag, som Barack Obama har gennemført i løbet af hans første præsidentperiode, er at fremme udviklingen af lærerevalueringssystemer, der primært evaluerer lærere ud fra deres elevers præstationer i standardiserede test. Men er det overhovedet muligt at vurdere kvaliteten af en lærers undervisning ud fra sådanne test?

Tilbage i 2002 implementerede Bushadministrationen The No Child Left Behind Act. Loven påbyder, at hver stat skal holde dets skoler ansvarlige på grundlag af standardiserede test i læsning og matematik. Obama, og hans uddannelsesminister, Arne Duncan, har taget skridtet videre: ikke bare skoler, men også hver enkelt lærer skal evalueres årligt. Oplysningerne skal, lyder argumentet, først og fremmest oplyse skoleledere om, hvilke lærere der har brug for at forbedre sig. Men tanken er ligeledes, at evalueringerne skal bruges til at rangere lærere, så gode lærere kan belønnes og dårlige lærere lettere kan afskediges.

New York er en af de stater, der tidligere i år udviklede et lærerevalueringssystem, i henhold til at kunne modtage et \$700 mio. stort føderalt bloktilskud gennem "Race To The Top"-planen. Lærere i New York vil, når loven træder i kraft næste år, blive bedømt på en skala fra 1-100. Det tungestvejende parameter i karaktergivningingen bliver lærerens elevers resultater i de statslige test. Formår læreren ikke at hæve sine elevers testresultater år for år, vil vedkommende overordnet set bedømmes som "ineffektiv", uanset udfaldet af de andre evalueringsparametre, der bl.a. inkluderer klasseobservationer. Hvis en lærer to år i streg får tildelt betegnelsen "ineffektiv", står vedkommende til at blive fyret.

Den bagvedliggende virkelighedsopfattelse er med andre ord, at der er en direkte kausal sammenhæng mellem kvaliteten af lærerens undervisning og *hvor meget* eleverne lærer. Upåagtet af forskelle i elevernes forudsætninger for læring. Og at det derfor udelukkende er lærerens ansvar, hvis vedkommendes elever ikke forbedrer sig nok mellem hver test. Problemet er, at dette ikke er rigtigt. Der er ingen tvivl om, at der selvfølgelig er en sammenhæng mellem kvaliteten af en lærers undervisning, og hvor meget eleverne lærer. Men det er vigtigt at holde sig for øje, at mange andre faktorer også spiller ind. Fx har økonomien Dan Goldhaber fra University of Washington påvist, at faktorer som socioøkonomisk baggrund og elevsammensætning, ud over selvfølgelig at påvirke elevernes generelle kundskabsniveau, også har stor betydning for deres relative forbedringer i testresultater over tid.

Med dette in mente er det i øvrigt kun så meget desto mere langt ude, at navngivne lærere fra netop New Yorks skolesystem tilbage i februar i år blev hængt ud i den lokale presse, fordi deres elever havde scoret lavt i de statslige test. Nogle lærere var endda blevet bedømt ud fra kun 10 elevers testresultater. Alligevel valgte blandt andet avisen New York Post efterfølgende at publicere oversigten, og samtidig bringe interviews med fortørnede forældre, der krævede de pågældende lærere fyret.

Ovenstående er et grelt eksempel. Men er det egentlig muligt at konstruere en mere hensigtsmæssig og gennemtænkt måde, hvorpå elevers testresultater kan bruges til at evaluere lærere?

Det afgørende er selvfølgelig at få isoleret lærerens indsats. Dette gøres statistisk ved at kontrollere for andre relevante variable der influerer elevernes testresultater. Et første skridt på vejen kan være at korrigere for socioøkonomiske forhold. Det sker fx herhjemme i forbindelse med de nationale test, hvor der for hver skole bliver udarbejdet en såkaldt socioøkonomisk reference, der i udregningen af resultatet tager højde for skolens specifikke elevgrundlag. I beregningen indgår en række let målelige faktorer, som fx køn, etnisk oprindelse, og forældrenes uddannelse og indkomst.

Herudover kan man udforme testene, som det er tilfældet i New York, så der bliver taget højde for hver enkelt elevs tidligere testresultater. I stedet for et øjebliksbillede af elevens kundskabsniveau, kan denne metode hjælpe med at give et indblik i, hvor meget eleven har forbedret sig siden sidste test. Dette giver unægtelig en bedre mulighed for at isolere effekten af en lærers undervisning.

Flertallet af forskere er dog enige i, at eksisterende metoder ikke er præcise nok til at kunne give et retvisende og fair billede af en lærers indsats. For det første er der en række statistiske indvendinger. Det er fx umuligt at kontrollere for *alle* andre relevante faktorer, som den pågældende lærer ikke har indflydelse på, men som også er med til at influere elevernes testresultater. Det kan være faktorer, der ikke er lige så lette at kvantificere, som dem der bliver brugt i udregningen af den førnævnte socioøkonomiske reference. Fx hvor motiverede forældrene er på elevens vegne. Det kan være en udslagsgivende faktor, der ikke nødvendigvis hænger sammen med hvad man tjener, og hvor man kommer fra. Flere studier i USA har fx sammenlignet testresultater mellem elever, der umiddelbart havde samme socioøkonomiske baggrund, men gik på forskellige skoler: offentlige og privat-offentlige skoler. Eleverne i de privat-offentlige skoler scorede højere i de statslige test. En nærliggende konklusion var derfor, nu der var taget højde for socioøkonomiske faktorer, at de privat-offentlige skoler var bedre end de offentlige skoler. Dette var dog ikke tilfældet. For sammenlignede man mere specifikt eleverne fra de privat-offentlige skoler, med elever der havde søgt om optagelse på disse, men havde fået afslag (gennem lodtrækning), og i stedet gik i de offentlige skoler, var der ikke forskel på testresultaterne. Det afgørende var altså, at både eleverne i de private-offentlige skoler, og eleverne der havde fået afslag, kom fra hjem, der aktivt engagerede sig i deres børns uddannelse, på trods af, at de på en række målelige økonomiske og uddannelsesmæssige punkter var at betragte som ressourcetsvage. En pointe er her også, at sådanne elever ikke er ligeligt fordelt på hver skole. Skoler med et godt ry tiltrækker flere elever, hvis forældre er engagerede på deres vegne. Et helt andet eksempel kan være, at der geografiske områder imellem er forskel på, hvilke muligheder børn og unge har for at blive læringsmæssigt stimuleret uden for skoletiden. Er der fx et bibliotek i nærheden? Pointen er i begge tilfælde, at der er en sandsynlighed for, at testresultaterne enten under- eller overestimerer lærerens indsats, hvis der ikke korrigeres for disse forskellige (fx geografiske) forudsætninger for læring. Og hvordan måler man fyldestgørende det?

At det kun er en enkelt lærer der evalueres, og ikke fx en hel skole, skaber også et andet statistisk problem. Da det er forholdsvis få elevers testresultater, som evalueringen derved er funderet på,

vil den statistiske usikkerhed i mange tilfælde simpelthen være for stor til, at man kan konkludere noget signifikant om sammenhængen mellem lærerens præstation og elevernes læring.

Samtidig skal man holde sig for øje, at standardiserede test kun måler en del af det, som lærere underviser i, og som eleverne skal beherske. Ikke så let målbare kompetencer, som fx elevernes samarbejdsevne, kommer derfor ikke til udtryk i testene. En fyldestgørende evaluering af en lærers virke kan testene derfor ikke give alene.

Et andet og mere generelt forbehold er, at en sådan testbaseret evaluering af lærere ofte har en u hensigtsmæssig indvirkning på undervisningen. Erfaringer fra fx USA viser, at der er en tendens til, at undervisningen bliver mere snævert målrettet mod det, som eleverne bliver testet i. Teaching to the test, med andre ord, hvor basale færdigheder og testtagningsteknikker opprioriteres, på bekostning af mere abstrakte og ikke så let målelige kompetencer. Denne dynamik er veldokumenteret. Herudover viser forskning også, at testbaserede lærerevalueringer, især de der potentielt kan påvirke en lærers løn- og ansættelsesforhold, historisk har givet anledning til snyd.

Ovenstående refleksioner omkring brugen af elevens testresultater til at evaluere lærere viser, hvor mange forbehold der skal tages. Når det er sagt, kan resultater i elevtest selvfølgelig godt bruges som en indikator på, at en lærer ikke gør det godt nok. Især hvis en lærers skiftende elever over tid konsekvent underpræsterer. Det bliver, i min optik, også først for alvor problematisk, når resultaterne bliver brugt til at rangere skoler og lærere, eller til at belønne og straffe skoler og lærere.

Måden det bliver gjort på i New York er u hensigtsmæssig. Den er ikke bare misvisende, men også, som jeg ser det, et udtryk for en form for institutionaliseret mistillid til lærerne. Samtidig kan man se det som en måde, hvorpå ansvaret isoleres og individualiseres, i den forstand at det hurtigt kan komme til at fungere som en slags ansvarsfralæggelse for især politikere.

**Af Mikkel Wittenburg Petersen**  
*Masterstuderende i sociologi*  
*The City University of New York*