

## **Traditioner for test og prøver – et komparativt perspektiv**

af Jakob Wandall

*Moderne test og prøver er blevet til gennem en lang udvikling og med komponenter taget fra forskellige tider, kulturer og systemer – formålet med denne artikel er at belyse de to mest markante hovedtræk, og samtidigt drage paralleller til baggrunden for og udviklingen af de danske nationale test.*

*Der er traditionsmæssigt to grundlæggende forskellige måder at udforme og at anvende test og prøver på. Disse forekommer ikke i ren form, men kommer i forskellig grad til udtryk i de enkelte landes systemer. Danmark og England i den sammenhæng eksponenter for nogle yderpunkter, og disse to systemers historie er derfor gode til at illustrere forskellene.*

*De nationale test i Danmark er designet med udgangspunkt i at plukke det bedste fra hver af disse verdener*

### **Skolens prøver stammer fra kirken i Danmark**

I Danmark udspringer skolen og traditionen med prøver af en kirkelig tradition, hvor hovedformålet med skolegang var at forberede eleverne til konfirmationen. Frem til starten af forrige århundrede omhandlede "examen" i underskolen kun kristendomsundervisning. I 1814 blev der beskrevet et lidt bredere fagformål, som dog stadig var koblet til konfirmationsforberedelsen. I anordning om skolevæsenet (forløber for folkeskoleloven) lød formuleringen således:

*Børnene skulde 2de Gange om Aaret, nemlig i April og October Maaneder, i den for Stedet anordnede Skolecommissions Nærværelse, offentligen overhøres og gjøre Rede for, hvad de have lært.*

*Ved saadan offentlig Prøve, som bør afholdes paa den af Skole-Directionen opgivne Maade og i alle Skoler være fuldført, forinden den 1ste Maii og 1ste November, skal det paaskønnes, hvilke der i hver Classe have lært meest, og hvis Fliid, Sæder og Opførsel have været de bedste saavel som hvorvidt de, der ville udskrives af Skoleprotocollen, have erhvervet de fornødne Kundskaber, eller om de behøve endnu at forblive i Skolen.*

Princippet var dengang at havde man lært tilstrækkeligt kunne man forlade skolen. Selv om dette princip og meget andet er ændret siden da, er der en del der er fælles for systemet dengang og nu, eksempelvis, at vurderingssystemet baseres på en – indenfor de givne rammer - subjektiv vurdering af den enkelte elevs præstation og standpunkt. Dette gælder stadig i vidt omfang ved en række selvstændige præstation, fx ved mundtlige/praktiske og andre former for ikke-standardiserede prøver (fx dansk stil). At inddrage den selvstændige præstation er i dansk pædagogisk tænkning vigtigere, end at målingen af kundskaber og færdigheder bliver objektiv. Og det er også bemærkelsesværdigt, at der ofte lægges større vægt på processer end på resultater.

Selvom konfirmationen (og senere folkeskolens afgangsprøve) er vigtige begivenheder - så er konsekvenserne af resultatet ikke afgørende for eleverne videre muligheder i uddannelsessystemet eller for skolernes økonomi og/eller eksistens.

Med den stigende globalisering er det uundgåeligt, at der ske kultursammenstød – selv i så nationalt definerede institutioner som grundskolen er. Et af de største kulturchoke i nyere tid var da danske skoleelever for første gang deltog i internationale målinger tilrettelagt på præmisser, der var ganske fremmede for den danske skole.

Det var den IEA-test, der blev gennemført i 1991 og blev offentliggjort i 1994. Der lå danske elever væsentligt under det niveau, som det var forventet.

Senere har PISA-undersøgelserne givet den samme oplevelse af, at verdens bedste folkeskole ikke har elever der præsterer i overensstemmelse med ressourceindsatsen.

### **Udvikling af moderne test - England lærte af Kina**

I England er testene baseret på en helt anden tradition, hvor udgangspunktet er, at man i videst muligt omfang objektivt skal måle, hvad eleven kan og ved. Objektivitet og ligebehandling er overliggende begreber. Dette baserer sig på en tradition som kan spores tilbage til Kina.

Det system som Englænderne var inspireret af var det kinesiske kejserlige eksamenssystem (The Imperial examination system – på kinesisk kaldet Keju). Kineserne havde udviklet et og raffineret et testsystem, hvor formålet bl.a. var give grundlag for rekruttering til den offentlige sektor i Kina. Frem til år ca. 600 havde man i Kina alvorlige problemer med nepotisme ved ansættelse i den kinesiske statsadministration. Det var baggrunden for, at kejseren lod udvikle et eksamenssystem baseret på tankerne af filosofen Kong Fuzi, et system baseret på ligebehandling, objektivitet og fairness – fx måtte eksaminator ikke kende identiteten af den, der blev eksamineret. Alle havde ret til at prøve lykken til disse skriftlige eksaminer, og der var hverken knyttet undervisning eller anden form for skolegang hertil – det var et rent rekrutteringssystem. At bestå disse eksamener var helt afgørende for eksaminandens fremtidsmuligheder i den kinesiske centraladministration.

Dette system kom til at virke indtil 1906 - i næsten 1.300 år. Der blev afholdt eksamener på tre niveauer efter meget omhyggeligt udformede og detaljerede regler/vurderingskriterier. Eksamenerne drejede sig om at kende og mestre konfucianismens[1] tankemåder og teknikker. Beståelse på første niveau gav ret til skattefritagelse og offentlig ansættelse. Kun en af en million bestod på tredje niveau, som var det højeste. At bestå her var en forudsætning for at kunne blive minister. Da der var så meget på spil, var indsatsen for at begrænse snyd helt central. Og sanktionerne ved snyd var voldsomme – i visse perioder henrettelse.

Det var gennem kristne missionærer at kendskabet til de kinesiske principper kom til englændernes kundskab. Omkring 1830 kunne hjemvendte missionærer rapportere, at den kinesiske centraladministration havde nogle helt særlige kvaliteter (ingen nepotisme og korrupsion, forvaltningsfagligt højt niveau), som blev tilskrevet deres rekrutteringssystem.

Dette fandt englænderne meget interessant, især da Storbritannien som kolonimagt stod overfor den udfordring at skulle opbygge centraladministrationer i kolonierne (især Indien). Men det fik kort tid efter også effekt i England. I 1855 blev der, med udgangspunkt i en rapport (den såkaldte Northcote-Trevelyan Report) der var udarbejdet med inspiration fra Kina, nedsat en "Civil Service Commission" der fik til opgave at sikre, at den offentlige administration i Storbritannien er effektiv og upartisk, og at udnævnelser foretages på grundlag af kvalifikationer.

Efter englændernes systemimport spredte principperne sig til andre områder. Det oprindelige testsystem havde ingen statistisk dimension – der fik det omkring 1860 gennem arbejde af englænderen Francis Galton (grundlægger af psykometrien, pionér inden for statistisk korrelation og regression) og belgieren Quetelet (forsker i måling/statistik om menneskelige egenskaber, herunder intelligens). I 1890 bragte den amerikanske studerende James Cattle Galtons teorier om intelligens til USA. I mellemtiden i Frankrig, var Alfred Binet (en psykolog der var meget optaget af test og måling af menneskelige formåen) travlt optaget med at udforme tests til bedømmelse af barnet intelligens. I 1904 blev han bestilt af den franske regering til at finde en metode til at skelne mellem børn, som var intellektuelt normale, og dem, der var underlegne. I Paris fik skolebørn Binet's test og der blev skabte en standard baseret på hans data. I USA udviklede militærmanden og psykologen Lewis Terman

---

1 [1] Konfucianismen en stærkt hierarkisk statsideologi, en kinesisk filosofi og religion hvor bl.a. fromhed, næstekærlighed og indordning i familie- og samfundsstrukturen er fremtrædende elementer. Den indeholder en mængde detaljerede leveregler for forholdet mellem over- og underordnede (mand/kone, fyrste/undersåt mv.). Den er baseret på læresætninger overleveret fra den kinesiske filosof Kong Fuzi (551-479 BC).

omkring 1919 (ifm. Første verdenskrig) en generel egnethedstest til hæren – baseret på en intelligensstest. Terman var ophavsmand til intelligenskvotienten (IQ) som måleenhed.

I perioden 1920-1950 (primært i USA og England), skete der en betydelig udvikling og raffinering af metoder til test i uddannelsessystemet og anvendelse af metoderne på faglige områder – især sprogtest – den lingvistiske testtradition. Sprogtest fik allerede der en central rolle som sorteringskriterium ifm immigration.

I 1950-1975 kom uddannelse på toppen af især den amerikanske politiske dagsorden og test integreres i diskussionen, som kontrol, som et styringsværktøj, som et vidundermiddel, som et problem, som løsningen på problemer, som et mål, som indikator.

I 1975 - 1995 giver computeren i stadigt stigende omfang mulighed for behandling af testdata, og teorier om testmodeller, der tidligere ikke havde været mulighed for at anvende i praksis (fx Rasch-modeller og andre beregningskrævende test-data-behandlinger) bliver mulige at anvende. Her bliver også grundlaget for de store sammenlignede undersøgelser af skoleelevers testresultater iværksat (NAPE, PISA, PIRLS, TIMMS m.fl.).

Fra 1995 begynder internettet at vinde frem og computerbaserede test bliver udviklet og Computere bliver dels kommunikationsmiddel og dels hvermandseje.

## **Den dansk og engelsk/amerikanske testtradition**

Der er store forskelle på forskellige landes motiver til at teste og traditioner for at anvende testresultater. Forskellene beror hovedsageligt på en kombination af forskellige historiske baggrunde og forskellige kulturer. I Danmark er der meget begrænset tradition for at teste, og da Undervisningsministeriet i 2004 fik i opdrag at udvikle grundlaget for nationale test, kunne man have forventet at Danmark ville gå til en af de store udbydere af testsystemer i den engelsktalende eller asiatiske verden. Men de systemers design og filosofi var det vurderingen ville være uforenelige med den danske tradition for skole- og evalueringskultur.

Undervisningsministeriet måtte derfor udvikle en helt egen testmetode, som forsøger at bruge det bedste fra den engelske tradition, uden at give køb på de særlige kvaliteter og kendetegn, der præger den danske skole.

### **Hvori består forskellene?**

Det kan siges meget kort: I Danmark tester vi for at eleverne skal blive klogere. Hvorfor skulle man dog ellers teste, vil de fleste danskere sige? Det betyder også, at det forekommer naturstridigt for de fleste danskere ikke at give eleverne en tilbagemelding på, hvad de har gjort rigtigt og forkert. Især der hvor de har lavet fejl, skal de have at vide hvad de har gjort galt – ellers lærer de jo ikke noget af det. Der arbejdes med en formativ tilgang.

Omvendt i England. Her tester man for at kontrollere, om eleverne har lært det de skal. Hvis ikke, skal læreren gøre noget i forhold til eleverne, ledelsen i forhold til læreren og de tilsynsførende i forhold til skolen. Senere tester man igen for at finde ud af om eleverne har lært noget i mellemtiden. – og gerne med den samme test/de samme opgaver, så man kan få en sammenlignelig måling. Hvis ikke de er blevet dygtigere, laver eleverne de samme fejl næste gang. At fortælle eleverne hvilke opgaver de har svaret forkert på og hvad det rigtige svar er, ville være at ødelægge testen som evalueringsværktøj.

### **Test - kontrol og pædagogik**

Formålet med test for englænderne og amerikanere er at overvåge skolerne, og testresultaterne anvendes ved beregning af løn og tilskud. I Danmark – ligesom i fx Norge og Sverige - anvender vi test som pædagogisk værktøj. Eleverne testes, for at de skal blive dygtigere - snarere end for at tilvejebringe information om elevernes faglige niveau til administrative formål.

Når englænderne anvender testresultater som grundlag for tilskud og løn må de nødvendigvis lægge stor vægt på, at resultaterne er uigendrivelige og objektive. Det medfører, at testene

snarere skal give en fair bedømmelse - uden at lægge fingrene imellem - end at tage hensyn til at eleverne skal have en positiv oplevelse og en fremadrettet og konstruktiv tilbagemelding. Fairness er en helt overordnet værdi i de engelsksprogede skolesystemer – filosofien er at både de svage og stærke elever skal have en krystalklar tilbagemelding. Man belønner den gode præstation og straffer den dårlige.

Omvendt i norden. Når testresultaterne ikke får indflydelse på hverken elevens eller lærerens situation, når processen er lige så vigtig som resultatet, når det er vigtigere at eleverne får noget ud af det end at lærerne får præcise tilbagemeldinger, så er det også klart, at præcisionen og objektiviteten i testværktøjet ikke bliver så central som den er fx i England. Her er omsorg en overliggende værdi – man må passe på ikke at sige de negative ting for direkte, så eleverne bliver kedede af det. Dette sympatiske træk har i visse tilfælde taget overhånd – eleverne kan ikke altid selv regne ud, at "du er dygtigst til at læse" betyder at eleven har alvorlige problemer med matematik. I nogle tilfælde betyder denne omsorg, at eleverne først for sent i skoleforløbet opdager hvor galt det står til.

Nogle vil bebrejde lærerstanden den manglende tilbagemelding. Men det er faktisk et kendetegn ved den skandinaviske model, der er i overvejende grad kan henføres til systemniveau. Et meget markant tegn på denne mekanisme er at der i folkeskoleloven er forbud mod at give karakterer før 8'ende klasse. Det er udtryk for et ønske om at beskytte de svagere elever på indskolings- og mellemtrin mod oplevelsen af nederlag, som kan virke demotiverende. Men det har i mange tilfælde en utilsigtet effekt, og det kan det være vanskeligt for skoleeksperter fra andre kulturer at forstå dette hensyn. Også mange elever hungrer efter mere klart feedback, så i mange tilfælde omgår lærerne reglerne ("Hvis jeg måtte give dig karakter ville du have fået 7") for at imødekomme elevernes behov.

Forskellen på de to tilgange til test kan illustreres således:

<b>Traditioner</b>	<b>Nordisk/kontinental</b>	<b>Engelsk-amerikansk</b>
Anvendelse	Pædagogik	Kontrol
	Didaktik / Undervisning	Styring
	Formativ / Low-stake	Summativ / High-stake
Formål	Læring --> Eleven	Information/Analyse --> "Systemet"/Læreren
	Lighed og solidaritet	Ambitioner og elite
	Omsorg	Fairness

Der er altså en stærk sammenhæng mellem formål med testen og hvor fagligt sikre og stærke test og evalueringsværktøjer, der anvendes, hvilket illustreres ved flagene mellem de to punkterede parallelle linjer i figuren nedenfor. Det som OECD fortalte Danmark var, at det ikke behøver at hænge sådan sammen. Beviset herpå er Finland.

### **Finske erfaringer – inspiration for danske nationale test**

I Finland har man særligt uddannede specialister (6-årig universitetsuddannelse), der gennemfører test i Finsk og Matematik hvert år på alle klassetrin i stort set alle finske skoler. Den finske stat giver vejledende retningslinjer, men det er kommunerne, der står for driften af skolen.

Men det finske system kunne ikke overføres til danske forhold. Selv om der er mange lighedspunkter mellem den danske og finske skoletradition, er der også store forskelle. Det ville ikke passe til den danske pædagogiske tradition, hvis arbejdet med pædagogiske test blev flyttet helt væk fra faglæreren.

Det finske system er grundigt beskrevet af Frans Ørsted, DPU, der blandt andet har skrevet phd-afhandling, som netop handler om en sammenligning af det danske og finske skolevæsen.

Læs også om hvad chefen for den finske evalueringsenhed selv vurderer, er forklaringen på den finske succes.

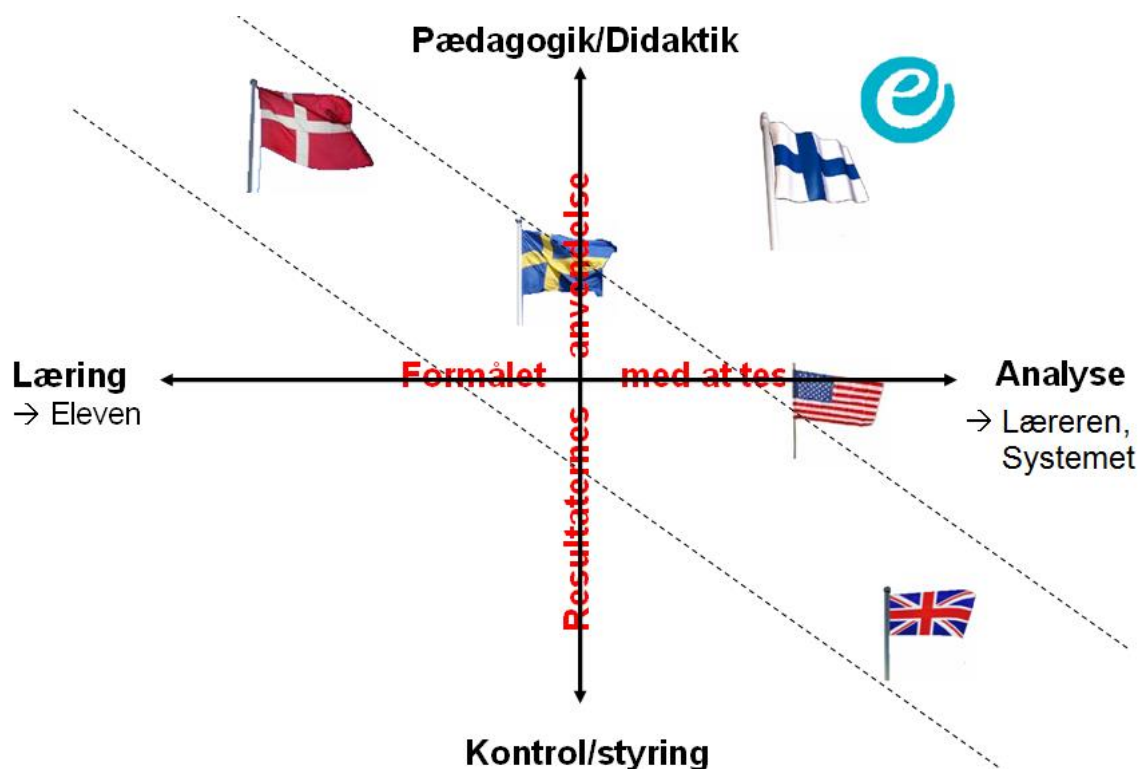
Det var altså nødvendigt at udvikle en helt egen model til de danske nationale test, hvor betingelserne har været:

- Testen skal kunne anvendes (gennemføres og fortolkes) af faglærerne,

- de skal være objektive og valide måleredskaber til elevernes faglige niveau og
- de skal kunne anvendes indenfor rammerne af den almindelige undervisning – dvs. kunne afvikles når det passer ind i lærerens skema og indenfor rammerne af en lektion.
- At testen **både** bliver et redskab til lærerens evaluering af undervisningen **og** et redskab for **pædagogisk** ledelse – i forholdet mellem lærer og leder og mellem kommunen og skoleledelsen. Men ikke bliver brugt fx i resultatkontrakter/-aflønning og offentlig rangordning af skoler

Det har altså været målet at udvikle test, der forener kvaliteterne i testtraditionerne og som er så let tilgængelige, at det ikke kræver særlig uddannelse at benytte dem. Ambitionen – som har vakt international opmærksomhed - er markeret på figuren, hvor de nationale test er symboliseret med "det smilende e".

Figur 1 De enkelte lande er placeret efter formål med testene og hvordan resultaterne bliver anvendt. De danske nationale test er markeret med "det smilende e" øverst til højre, lige ved siden af det finske flag.



## Om de nationale test

De nationale test adskiller sig på en række områder fra andre test, der anvendes i folkeskolen, blandt andet ved at de er:

- It-baserede, hvilket vil sige, at de skal besvares på en computer med netforbindelse
- Selvscorende, hvilket vil sige, at man som lærer ikke skal rette dem og udregne resultater, men får resultaterne i forskellige former for resultatvisninger, der hentes på nettet
- Adaptive, hvilket vil sige, at de tilpasser sig den enkelte elevs faglige niveau.

## Det adaptive princip

At en test er konstrueret efter det adaptive princip betyder, at opgaverne bliver udvalgt specielt til den enkelte elev. Princippet er enkelt: Eleven starter med en middelsvær opgave. Svarer eleven rigtigt på opgaven, vil næste spørgsmål blive sværere. Svarer eleven forkert, vil næste spørgsmål blive lettere. Efter nogle spørgsmål, der er udvalgt på denne måde, beregnes elevens dygtighed ud fra de besvarede opgaver. Da opgaverne vælges specielt til den enkelte elev afhængig af elevens besvarelse, vil eleven få opgaver, der har en passende sværhedsgrad – uanset hvor stærk eleven er i faget. Erfaringsmæssigt giver det et mere udfordrende og effektivt testforløb for eleven, når eleven ikke møder opgaver, der enten er alt for lette eller alt for svære.

Det adaptive princip betyder, at man kan afdække den enkelte elevs kundskaber og færdigheder

med under 1/3 af de opgaver, der skal buges ved en normal test. Det betyder også, at der sandsynligvis ikke vil være to elever i en klasse, som vil få de samme spørgsmål i den samme rækkefølge. Men det kan forekomme, at den ene elev bliver testet særligt grundigt i nogle opgaveemner, som kun lige akkurat forekommer i en anden elevs test.

Der er altså et dobbelt perspektiv med den adaptive test:

- At sørge for at eleverne får spørgsmål på et passende fagligt niveau, hvor eleverne hverken bliver overudfordrede eller keder sig
- At der med relativt få spørgsmål gives et pålideligt testresultat for den enkelte elev.

Hver elev får således i praksis sin egen skræddersyede test med en særlig kombination af opgaver. Resultatvisningerne til læreren rummer en beskrivelse af hver enkelt elevs præstation. Det støtter læreren, når denne efterfølgende tilrettelægger undervisningen og ønsker at tilgodese både fagligt svage og fagligt stærke elever efter princippet om undervisningsdifferentiering og holddannelse.

For at testens resultater skal blive så retvisende som muligt, skal selve testafviklingen forberedes og gennemføres omhyggeligt. Gode råd og vejledning hertil er tilgængelige i brochurer og vejledninger.

Et væsentligt formål med den adaptive test er, at den giver et pålideligt resultat for den enkelte elev for den samlede test, men også et separat resultat for de forskellige faglige områder, som testen omhandler. Disse områder kaldes profilområder.

## Profilområder, opgaveemne og trinmål

Hvert af de fag, der skal testes i, er blevet opdelt i indholdsmæssige profilområder, som hver danner ramme om en del af testforløbet.

Der er tre profilområder i hver test, jf. tabellen.

Fag/klasse	Profilområde 1	Profilområde 2	Profilområde 3
Læsning (dansk), 2., 4., 6. og 8. klasse	Sprogforståelse	Afkodning	Tekstforståelse
Matematik, 3. og 6. klasse	Tal og algebra	Geometri	Matematik i anvendelse
Fysik/kemi, 8. klasse	Energi og energiomsætning	Fænomener, stoffer og materialer	Anvendelser og perspektiver
Engelsk, 7. klasse	Læsning	Ordforråd	Sprog og sprogbrug
Geografi, 8. klasse	Naturgrundlaget	Kulturgeografi	At bruge geografien
Biologi, 8. klasse	Den levende organisme	Levende organismers samspil	At bruge biologien
Dansk som andetsprog, 5. og 7. klasse	Ordforråd	Sprog og sprogbrug	Læseforståelse

Man kan læse mere om det faglige indhold i profilområderne i profilområdebeskrivelserne for de nationale test og i opfølgingsartiklerne for hver test på undervisningsministeriets evalueringsportal [www.evaluering.uvm.dk](http://www.evaluering.uvm.dk)

I henhold til Fælles Mål for folkeskolen skal undervisningen lede frem mod, at eleverne har tilegnet sig kundskaber og færdigheder, der sætter dem i stand til at forstå . . . , kende til . . . , udføre . . . mv. på forskellige tidspunkter i skoleforløbet. Målene er formuleret, så de ligner mål for elevernes læring, og det har ført til mange misforståelser.

Trinmålene er mål for undervisningen – en form for overligger, en checklist for læreren om hvad læreren skal nå igennem i sin undervisning. Trinmål kaldes derfor **undervisningsmål**, der skal anvendes som pejlemærker for lærerens undervisning. En evaluering af trinmålene ville derfor være en evaluering af lærerens indsats og ikke af elevernes faglige niveau.

Elever er forskellige, lærer på forskellige måde på forskellige områder med forskellige hastigheder og har forskellige potentialer. Der er mange klassetrins forskel på de fagligt dygtigste og svageste elever i normale klasser, også selvom de har haft samme lærer.

Eksempelvis vil der normalt være et par elever i en normal 2. klasse der klarer en læsetest bedre end den tunge del af eleverne i en normal 8. klasse. Gennem resultater fra de nationale test kan man udlede, at der som regels er mere end 10 klassetrins spring mellem de 2-3 bedst præsterende elever og de 2-3 dårligst præsterende elever i den samme normale 8. klasse i folkeskolen. Denne afstand vil overraske de fleste, men svarer fuldstændigt til hvad der er normalt i de fleste andre lande (fx USA og New Zealand).

Derfor er det også af afgørende betydning, at læreren differentierer sin undervisning på grundlag af den løbende evaluering, og at den løbende evaluering afdækker elevernes faglige niveau. Evalueringen danner på denne måde grundlag for, at læreren i samarbejde med eleven kan fastsætte individuelle faglige mål - **læringsmål**. Det er vigtigt at læreren fokuserer på elevernes individuelle fremskridt- altså på elevernes progression – ikke blot end på det absolutte niveau.

Testenes funktion er først og fremmest at vurdere elevernes faglige niveau inden for hvert af de områder, som testes - profilområderne. Testene kan ikke måle, hvor eleven har lært det, som eleven kan – om det er i undervisningen, i hjemmet eller af kammerater. Testopgaverne i profilområdet er måleredskaberne. Hver opgave skal afprøve en lille del af profilområdet, og det, der mere specifikt afprøves, kaldes for opgavens testmål. Eksempler på testmål inden for tal og algebra er: "Addition af encifrede tal med tierovergang", "Enkel omsætning af brøk til decimaltal", "Division af flercifrede tal med afrunding".

Når der bliver besvaret mange opgaver med et forskelligt indhold, kan testen give et meget sikkert billede af elevens kundskaber og færdigheder inden for et profilområde. Der kan endvidere være megen information at hente ved at gennemgå elevens besvarelse af opgaverne – men dette er overordentligt tidskrævende. For at kunne danne sig et overblik over elevens niveau på et detaljeringniveau mellem profilområde- og opgaveniveau, er der koblet to oplysninger (metadata) til hver opgave, nemlig et opgaveemne og et trinmål.

**Opgaveemnet** beskriver i almindeligt sprog de kundskaber og færdigheder, der afprøves med opgaven, fx som her eksemplificeret ved profilområdet tal og algebra inden for faget matematik: talkendskab, 10-talsystemet, regneregler, addition eller subtraktion, multiplikation eller division, flere regningsarter, ligninger, decimaltal, procent, omskrivning og talmønstre. Denne oplysning skal blandt andet bruges til at fortælle forældrene om, hvilke opgaver deres barn har besvaret.

Opgaveemnet er således defineret med henblik på formidling af resultatet og skal ikke forveksles med opgavens testmål.

**Trinmål:** Alle opgaver kan man lære at løse på mange måder – i og udenfor skolen – og gennem mange forskellige dele af undervisningen både hvad angår fag og trinmål. Det trinmålet, der er koblet på opgaven, er et udtryk for den del af undervisningen, hvor eleven med størst sandsynlighed vil kunne blive bedre til at løse netop denne opgave.

Hvis eleven svarer forkert på en opgave, skal opgavens trinmål altså ses som et forslag til, hvor læreren kan styrke sin undervisning, for at eleven får de kundskaber og færdigheder, der

skal til for at besvare den pågældende type af opgaver. Det til opgaven knyttede trinmål refererer således til den del af undervisningen, som eleven med det gode resultat har brug for ekstra udfordringer i (i forhold til klassen), og som den elev med det dårlige resultat har brug for, at der gøres en ekstra indsats indenfor.

Det er vigtigt at understrege, at opgaverne ikke tester trinmålene. Det betyder også, at ændringer i Fælles Mål ikke medfører behov for ændring i opgaverne, men at der måske vil skulle tilknyttes nye trinmål til de eksisterende opgaver.

Kunsten at anvende testresultaterne i forbindelse med fastsættelse/revidering af individuelle læringsmål for eleven og opfølgning herpå, er en markant udfordring for læreren. Det forudsætter forståelse for hvad testresultater kan fortælle, viden om hvad man skal stille op rent undervisningsmæssigt med den nye viden og evne til at implementere ændringer i den daglige undervisning. Men det er ikke desto mindre det som læreren - med sin fagdidaktiske og pædagogiske viden, kendskab til eleven og andre resultater af den løbende evaluering - skal praktisere ved tilrettelæggelse af den daglige undervisning.

### Lærerens testresultater

For hver elev måles elevens faglige niveau som beskrevet på tre områder – profilområderne. Alle testresultaterne er forskellige visninger af disse tre tal for den enkelte elev. Testresultatet giver læreren oplysninger om, hvordan eleven har besvaret opgaverne på dagen. Profilområderne er udvalgt, så testresultaterne giver information om kundskaber og færdigheder, der er vigtige for faget og undervisningen. Men testresultaterne giver hverken en forklaring på hvorfor noget er gået godt/galt eller på hvordan undervisningen skal tilrettelægges.

Når der skal følges op på testens resultater, kan læreren dels anvende sin viden om testforløbet:

- Fungerede alt som det skulle?
- Kendte eleven teknikken godt nok?
- Var eleven koncentreret og fokuseret?
- Hvilke hjælpemidler var til rådighed?
- Fik eleven hjælp?

og dels sammenligne med sin viden om eleven:

- Hvordan svarer testresultaterne til forventningerne?
- Var der positive og/eller negative overraskelser?
- Giver resultaterne behov for yderligere afdækning?

For at give læreren den letteste adgang til den information der er skabt, er resultaterne præsenteret på forskellige måder i form af flere forskellige rapporter

I disse resultatvisninger ligger der mange oplysninger, der kan støtte arbejdet med at planlægge og tilrettelægge en undervisning, der rummer udfordringer for såvel klassen som for den enkelte elev.

Klikker man på Testresultater, får man adgang til følgende rapporter med test resultater:

- Klassen/tabeloversigt - er en samlet oversigt over de gennemførte test for lærerens klasser (eller elevgrupper). Resultaterne vises som et gennemsnit for hvert af profilområderne for den enkelte klasse samt et gennemsnit af elevens resultat i de tre profilområder (samlet vurdering).
- Klassen/grafisk oversigt - viser grafisk, hvordan den samlede elevgruppe eller klasse i den valgte test fordeler sig på fem niveauer inden for hvert profilområde. Til sammenligning vises det forventede niveau, der er den norm som testen måles op mod (referencen for de nationale test er resultatet af de



test, der blev gennemført de første uger i foråret 2010, hvor testene blev obligatoriske første gang. Ca. 10.000 elever pr. test udgør referencegruppen).

- Alle elever/tabel. I denne oversigt vises hver elevs score inden for hvert profilområde samt samlet vurdering. Elevens score angives i percentiler i forhold til hvordan referencegruppen har klaret sig (den dårligst præsterende procentdel får 1, middeleven får 50 og den bedst præsterende procentdel af eleverne får 100) er både angivet som percentil - et tal fra 1-100 - og i tekst, henholdsvis klart under middel (1-10), under middel (11-35), middel (36-65), over middel (66-90) eller klart over middel (91-100).
- Alle elever – detaljer. Ønsker man en mere detaljeret resultatvisning kan man klikke videre på:
  - Se resultater i tekst. Denne visning er den beskrivelse i ord af elevens resultater, som ved de obligatoriske test skal formidles til hjemmet.
  - Se besvarelse. Her vises en tabel over alle de testopgaver, eleven har mødt i testen. Tabellen viser detaljer for testopgaverne, dvs. hvilke opgaveemner testopgaven relaterer sig til, elevens score (R= rigtig eller F= forkert eller B= blank, ej besvaret) og opgavens sværhedsgrad (vist på en skala fra 1-5 stjerner). Ved et klik på Se opgave ses selve opgaven og elevens besvarelse samt det rigtige svar. Desuden vises opgavens bagvedliggende data: Det trinmål, opgaven relaterer sig til, opgavens sværhedsgrad og opgaveemnet. I læsning vil opgavens normerede tidsforbrug ligeledes fremgå.

### Hvad måler de nationale test?

Testresultaterne kan bidrage til en mere nuanceret vurdering af klassens faglige niveau. Men det er vigtigt at understrege, at test alene giver en tilbagemelding på klassens samlede resultat og den enkelte elevs præstation. Test er for læreren, hvad blodtryksmålere og termometre er for lægen: De kan give en aktuel status på nogle konkrete områder som er vigtige for læreren at vide noget om. Men de kan ikke fortælle, hvad årsagen til problemer er, og endnu mindre, hvad der er den rigtige kur. Dette beror på lærerens fortolkning af resultaterne, viden om eleven og andre evalueringresultater. De nationale test er specielle ved, at de på en række områder kan give et velfunderet, hurtigt og systematisk overblik over, hvad eleverne kan på en række centrale områder af faget.

Læs om adaptive test: <http://www.rasch.org/memo69.pdf>

Læs om teknikken i de Danske nationale adaptive test:

- [Journal of Applied Testing Technology, Vol 12, Article 3](#)
- [National Tests in Denmark – CAT as a Pedagogic Tool](#)