



Test i historisk perspektiv, hvorfor tester vi?

JAN MEJDING, LEKTOR, DANMARKS INSTITUT FOR PÆDAGOGIK OG UDDANNELSE, DPU, AARHUS UNIVERSITET

Gennem nogle nedslag i skolehistorien beskrives det, hvordan testning af elevers læsekompetence blev et område, som med tiden fik et øget fokus, fra de første forsøg på at gøre læsningen 'målbar', så det kunne fastholdes, hvad der var normalt, og hvad der var afvigende fra normen, til placeringen af danske elever i en international sammenhæng – alt sammen med målet om at kunne give alle elever den bedst mulige støtte til at opnå en funktionel og flydende læsning.

Test gennem tiden

Så længe man har haft skoler, der ikke kun var for eliten, har man været klar over, at ikke alle kunne det samme. Allerede i 1800-tallet oprettedes særlige institutioner til de "døve, blinde, åndssvage eller vanføre", og fra 1900-tallets begyndelse opstod i de store byer i Danmark de første hjælpeklasser for elever, der havde svært ved at følge med i skolen (Kirkebæk & Skov, 2019). På den tid kunne antallet af elever i klassen på de store købstadsskoler nå op omkring 80, så der kunne være god grund til at skaffe elever med særlige vanskeligheder et andet undervisningstilbud. I første halvdel af 1900-tallet steg interessen for børnepsykologi og psykologisk funderet pædagogik, og i 1934 blev det første skolepsykologiske kontor i Danmark oprettet på Frederiksberg med arbejdsområdet at tage standpunktspøvrer, henvise til specialklasser og holde foredrag for lærerne (Hansen, 2019). I

1936 oprettede Edith Norrie det første danske institut for ordblinde (Norrie, 1960; Hermann, 1955), og i slutningen af 1940'erne tog Carl Aage Larsen, skolepsykolog i Gentofte, initiativ til udviklingen af de første danske læseprøver til brug ved skolepsykologiske undersøgelser.

Hvor skolesystemets fokus i begyndelsen således var rettet mod *elevdifferentiering* – her blev elever med særlige vanskeligheder henvist til undervisning i særlige klasser eller skoler – kom der op gennem 1960'erne og 70'erne en øget opmærksomhed på identifikationen af elevernes vanskeligheder med henblik på at kunne *undervisningsdifferentiere* – tilrettelægge undervisningen efter de behov, grupper af elever kunne have. I Lyngby-Taarbæk Kommune var den ledende skolepsykolog Arne Søgård en af pionererne i udviklingen af en række stillelæsningsprøver, der kunne bruges i den sammenhæng. Allerede i begyndelsen af 1960'erne begyndte udviklingen af en række af de prøver, som kendes endnu i dag: Ordstillelæsningsprøve 400 og S50 A og B-prøven – forløberne for OS64 og OS120 og de senere sætningslæseprøver SL-60 og SL-40. Til denne samling kom i 1991 de individuelle læseundersøgelser, IL-prøverne, som fokuserede på læsning af små historier. Først med folkeskoleloven af 1993 blev undervisningsdifferentiering gjort til et krav i folkeskolen – nu også med henblik på tilrettelæggelsen af undervisningen for de enkelte elever.

Standpunktsprøver

De test, man anvender i skolen/klassen, kan være enten normrelaterede eller kriterierelaterede. De normrelaterede prøver viser, hvad elever i almindelighed kan på de forskellige klassetrin, og de kriterierelaterede prøver fokuserer på, hvilke differentierede færdigheder eleverne har inden for forskellige færdighedsområder. Til de sidste hører også de diagnostiske prøver, hvor man lige så meget er interesseret i at få et vide specifikt, hvad eleverne IKKE kan, eller hvad de er lige ved at kunne, som i at vide, hvad de faktisk kan.

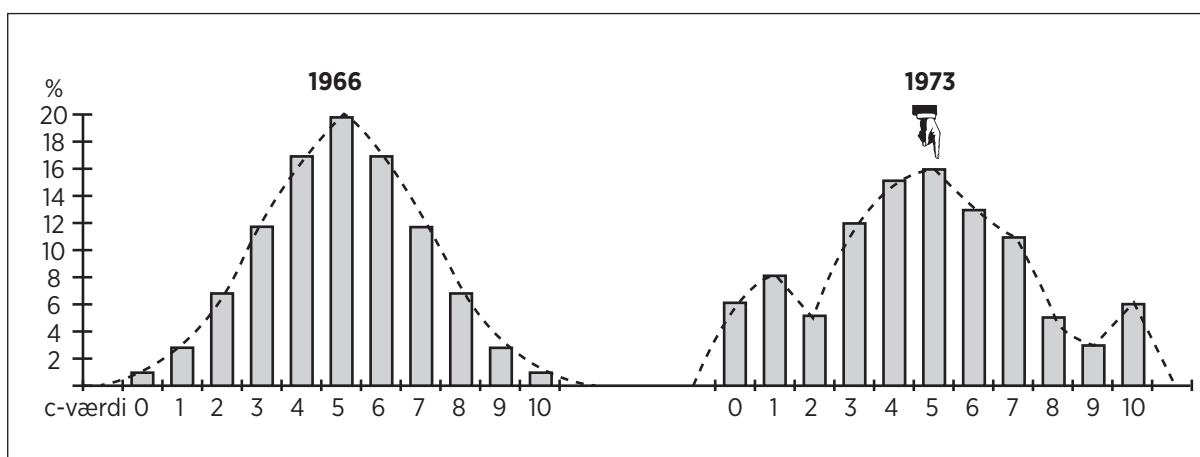
Oftest er der et samspil mellem de to typer af prøver. Det kan være vigtigt for planlægningen af undervisningen af den enkelte elev at vide, hvad eleven endnu ikke er sikker i, men det er også nødvendigt for læreren at vide, om det er noget, elever i almindelighed skal kunne på dette klassetrin, eller om det først kan forventes senere i skoleforløbet. Den viden får man kun ved at kunne sammenligne den enkelte elevs præstation med en større gruppe af elever – hvor gruppens samlede præstation bliver en norm for den enkelte elevs. Normen kan være den enkelte klasse, skole eller kommune, eller det kan være en norm, som bygger på et landsrepræsentativt udvalg af elever på det pågældende klassetrin.

Hvis ikke normerne kontrolleres og revideres med jævne mellemrum, så risikerer [de] at virke konserverende for udviklingen i skolen.

Ideen med standpunktsprøver var, at man skulle kunne give et objektivi mål for, hvad en elev kunne inden for et givet fagområde, og elevens samlede resultat af prøven kunne så sammenlignes med en eksisterende norm for at se, om elevens resultat lå over eller under gruppens gennemsnit. Det har den fordel, at det kan give læreren en fornemmelse for, hvor godt eleverne i klassen klarer sig i forhold til andre elever i skolen, kommunen eller i landet som helhed. Men om dette resultat er retvisende, afhænger helt af, at 'normen' er tidssvarende. Hvis ikke normerne kontrolleres og revideres med jævne mellemrum, så risikerer

sådanne normer at virke konserverende for udviklingen i skolen. Det er imidlertid dyrt og arbejdskrævende til stadighed at ajourføre normer, hvorfor man ofte ser, at normerne fortsætter som referencepunkter, selv om den efterfølgende udvikling har gjort dem forældede.

Med fremkomsten af de standardiserede standpunktsprøver blev det mere og mere almindeligt, at man i kommunerne anvendte disse prøver for at se, hvordan elevernes færdigheder udviklede sig. Ved konstruktionen blev prøverne skaleret, så elevernes præstationer fordelte sig i en normalfordeling, hvor forholdsvis få ikke fik ret mange point, mange klarede sig nogenlunde, og andre forholdsvis få klarede sig rigtig godt (se figur 1). Med fremkomsten af flere og flere kommune-resultater var der dog noget, som tydede på, at elevernes læseresultater ikke i alle tilfælde kunne beskrives ved hjælp af en normalfordelingskurve, men at resultaterne fordelte sig på andre måder. I 1972 blev undervisningspligten udvidet fra syv til ni år, og nogle af de ældste elever, der havde det svært med skolen, havde ikke længere mulighed for at gå ud af skolen til et ufaglært arbejde – og med folkeskoleloven af 1975 blev det fastslået at: *”Til børn, hvis udvikling kræver en særlig hensyntagen eller støtte, gives der specialundervisning og anden specialpædagogisk bistand”*. Begge dele øgede behovet for dels at kunne identificere elever med de særlige behov, også blandt de ældre elever, dels at kunne tilrettelægge den særlige hensyntagende undervisning. I de følgende år oprettedes specialklasser og nogle af de specialskoler, som skulle tage sig af denne del af undervisningen, og det øgede naturligt omkostningerne og lærerforbruget i kommunerne. Senere, for at udgifterne til undervisningen ikke skulle løbe løbsk, skete der en stigende integration i folkeskolens klasser af de elever, der tidligere havde været udskilt til undervisning i særlige specialskoler eller -klasser. Det skete ud fra de bedste motiver om at give alle elever de samme forhold og undervisningsvilkår – men desværre viste det sig oftest, at de fornødne ressourcer til støtte ikke fulgte med ud i klasserne, hvilket stillede både lærere og elever i en uholdbar situation. Læseprøverne kunne dokumentere, at problemet eksisterede, men den kommunale økonomi tillod ikke, at ekstra ressourcer blev anvendt.



Figur 1. Læsefærdighed i 3. klasse, 1966 og 1973

Læseprøveresultater

I Lyngby-Taarbæk (Hansen, Kreiner, & Pagaard, 1987) havde man i længere tid undret sig over, at hvad der i begyndelsen kunne beskrives med normalfordelingskurven, efterhånden bedre kunne beskrives som en 'trykseksten-tendens' eller 'Matthæus-tendensen', hvor grupperne i yderpunkterne af kurven voksede – se figur 1.

Tendensen fra 1966 til 1973 kunne konstateres langt op i 1990'erne, hvor man kunne se, at selv om eleverne i gennemsnit blev lige så dygtige som tidligere, så læste en større gruppe elever dårligere end før, samtidig med at en større gruppe af elever blev dygtigere end før. Denne situation blev en voksende pædagogisk udfordring for dansk lærerne i skolen. Selv om man i Danmark fra 1940'erne og frem havde fået publiceret en lang række børne- og ungdomsbøger, som var læsepædagogisk bearbejdet, og selv om Landsforeningen af Læsepædagoger blev stiftet i 1950 med formålet om at udbrede kendskabet til den gode læseundervisning, så var det tilsyneladende ikke tilstrækkeligt til at sikre, at alle kom godt i gang med deres læseudvikling. Ved revisionen af OS400 og S50-prøverne i slutningen af 1980'erne foranstaltede Dansk Psykologisk Forlag, som dengang stod for prøverne, i 1987 en landsindsamling af prøveresultater, der skulle bruges i forbindelse med etableringen af de nye landsnormer. Også her fandt man, at læseudviklingen ikke længere kunne beskrives med en enkelt norm. Det så snarere ud til, at der var to normer – en øvre og en nedre landsnorm – hvor man ikke

umiddelbart kunne forklare, hvorfor forskellige kommuners resultater adskilte sig fra hinanden (Paludan-Müller & Mejding, 1988; Jansen & Ziegler, 1990).

Tid til dansk

I 1980 ændredes folkeskoleloven i forbindelse med særforborgens udlægning, og henholdsvis kommunerne og amtskommunerne fik med § 19.1 og § 19.2 ansvaret for specialundervisning og anden specialpædagogisk bistand til børn og unge, der havde særlige behov. I forlængelse heraf og i forbindelse med de ændrede resultater i læseprøveresultaterne ønskede man en oversigt over udviklingen i den faktiske undervisningstid i dansk fra 1960'erne og til midten af 1980'erne. Beregnede man den i hele klokketimer, så viste det sig, at der fra 1962/63 til 1984/86 i gennemsnit var skåret omkring 1/3 af dansktimetallet hen over de første syv skoleår (Jansen & Ziegler, 1990). Størst var nedskæringen på mellemtrinnet, hvor dansktimerne i 3., 4. og 5. klasse var skåret ned med henholdsvis 35 %, 37 % og 47 %. Dette var baggrunden for iværksættelsen af "De generelle udviklingsarbejder i dansk – på 3.-4. klassetrin". Evalueringen af projektet blev forestået af Danmarks Pædagogiske Institut og er beskrevet i en rapportrække under fællestitlen: Tid til dansk. En central del af De generelle udviklingsarbejder var at undersøge, om en ugentlig time mere i danskundervisning i 3.-4. klasse betød noget for elevernes læsefærdigheder¹. Som udgangspunkt havde man forestillet sig, at forholdene på skolerne nok kunne variere lidt, men forventningen var, at timetallene nok ville ligge ganske tæt på de vejledende timetal udstukket af Undervisningsministeriet. Først i løbet af

undersøgelsen blev det klart, at såvel timetal som antal elever i klassen varierede langt mere, end man havde troet. I løbet af de to første skoleår var der elever, der havde modtaget mere end dobbelt så megen danskundervisning som de elever, der fik mindst. Allerede ved indgangen til 3. klasse var der en forskel mellem eleverne med hensyn til de forhold, de havde haft i skolen, som kunne betyde noget for deres udbytte af den såkaldte 'gyldne' time – den ekstra time, klasserne fik i forbindelse med udviklingsarbejdet.

Allerede da var det således klart, at den afgørende variabel er kvaliteten af undervisningen i den tid, der er til rådighed, og ikke tiden i sig selv.

Så man nærmer sig på disse forskelle, så viste det sig imidlertid, at de betød langt mindre, end man kunne have forventet. Delte man eleverne ind efter de klasser, som havde haft et lavt gennemsnit af dansktimer i 1.-2. klasse (6-7,5 timer ugentligt), og holdt dem op mod dem, som havde haft et højt gennemsnit af dansktimer (9-10 timer ugentligt), og så på deres resultater i den almindeligt anvendte ordlæseprøve OS120, så var det kun muligt at finde signifikante resultater, hvis man samtidig inddrog klassestørrelsen som en variabel. Hvis klassen var stor (18-29 elever) og havde haft få dansktimer i 1.-2. klasse (6-7,5 timer ugentligt), så var der relativt færre gode og sikre læsere (kategori A1 med 12 %) og relativt flere elever med dårlige resultater (26 % som var dårligere end kategori C1), end der var i de små klasser (7-17 elever), hvor elever med det forudgående samme ugentlige timetal fordelte sig henholdsvis med 22 % (i A1-kategorien) og 17 % (i restkategorierne) (Johansen & Mejding, 1990). Et tilsvarende resultat fandt man med læseprøven SL60 i 3. klasse november – og efter afslutningen af de generelle udviklingsarbejder i 3.-4. klasse var konklusionen den samme: Det er ikke tilstrækkeligt kun at beskæftige sig med rammefaktorer som klassestørrelse og timetal, hvis man ønsker at forbedre præstationerne i skolen. Allerede da var det således klart, at den afgørende variabel er kvaliteten af undervisningen i den tid, der er til rådighed, og ikke tiden i sig selv. En nyttig viden, som med

fordel kunne have været inddraget i forberedelsen af folkeskolereformen i 2013, hvor tiden blev italesat som den afgørende variabel.

De internationale tendenser

I det foregående har vægten i testningen været på de nationale eller regionale normer. Men én ting er at vide, hvordan ens klasse eller kommune klarer sig i forhold til resten af kommunen eller de andre kommuner, noget andet er, om det svarer til, hvad der sker rundt om i verden. Samtidig med den nationale udvikling af skolen som system, med redskaber i form af standpunktsprøver og intelligensprøver, skete der en udvikling i de internationale relationer. 1900-tallets store krige mellem nationerne og industrialiseringens ekspansion gjorde det klart, at man som land ikke længere kunne være sig selv nok. De Forenede Nationer blev oprettet i 1945 på baggrund af 1. og 2. verdenskrig – og i 1957 samlede en række lande i Rom-traktaten, hvilket i 1967 udviklede sig til Det Europæiske Fællesskab (EF). Internationaliseringen var kommet for at blive. Og således også i det pædagogiske perspektiv. I 1994 samlede repræsentanter for en lang række lande i Salamanca i Spanien og vedtog en fælleserklæring om alle børns ret til uddannelse uanset handicap, og denne erklæring og tankerne bag har været grundlag for en lang række skole-reformer i forskellige lande – og afspejles også i den danske folkeskolelov fra 1993. En lang række kommuner fik internationale venskabskommuner, og blandt de skandinaviske lande blev det populært ikke mindst på grund af det sproglige fællesskab. Men selv om man kom på besøg hos hinanden, så var det sjældent, at det gav anledning til større ændringer. Man manglede muligheden for at sætte forskellighederne i relief.

Den internationale læseundersøgelse

Den mulighed kom med "Den internationale læseundersøgelse" iværksat af den internationale forskningssammenslutning IEA – The International Association for the Evaluation of Educational Achievement. IEA var blevet grundlagt i 1958 med ambitionen om at kunne levere internationalt sammenlignelige data på undervisningsområdet, og op gennem 1960'erne og 70'erne kom de første internationale undersøgelser i bl.a. matematik, na-

turfag, medborgerskab og skriftlig fremstilling og i 69-73 også den første undersøgelse i læseforståelse (reading comprehension) – dog kun med 10 lande som deltagere. Danmark blev, med repræsentation fra Danmarks Pædagogiske Institut, først medlem af organisationen i slutningen af 1980'erne. Den internationale læseundersøgelse gennemført i 1991 (Elley, 1992; Mejding, 1994) var den første større internationale undersøgelse, Danmark var med i. Her var 32 lande/skolesystemer repræsenteret og blandt dem alle de nordiske lande.

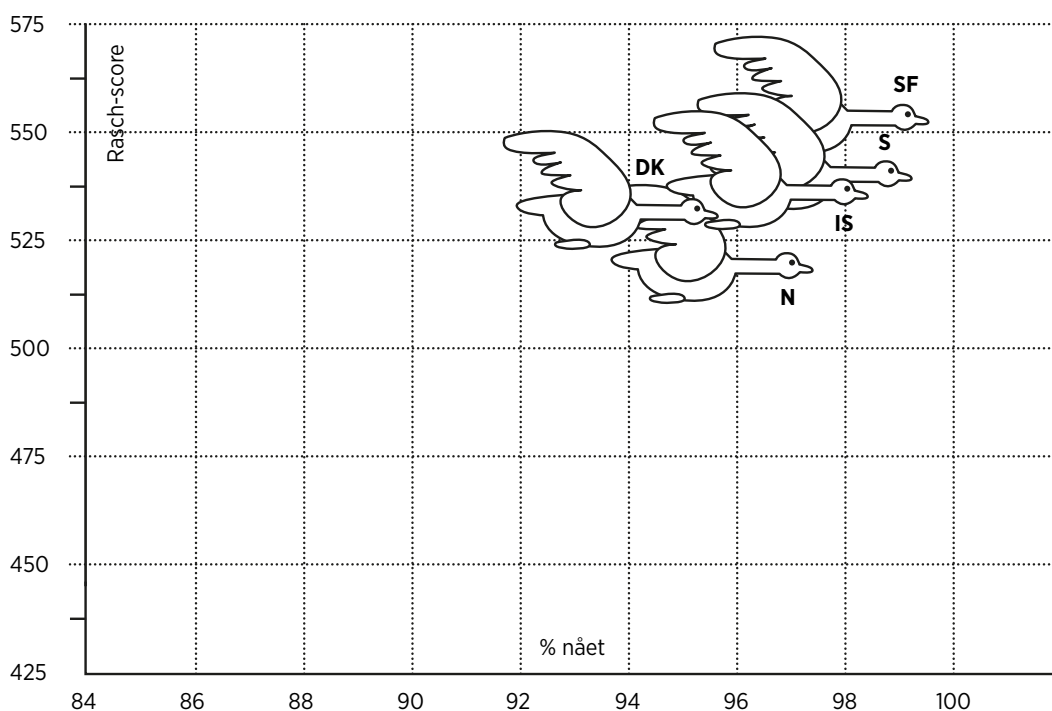
Denne undersøgelse blev noget af en øjenåbner for det danske skolesystem. Hvor vi i Danmark selv betragtede os som foregangsland inden for pædagogikken generelt, så stod vi pludselig i en situation, hvor vi blev sammenlignet – på sammenlignelige data – med elevers præstationer i andre lande. I denne undersøgelse så man på læseforståelsen hos de 9-årige (3. klassetrin) og de 14-årige (8. klassetrin), og man undersøgte hele klasser for også at kunne se på forskellen mellem klassernes resultater (klasserumseffekten) – og dermed gav det også mening at hente information hos modersmålslærerne i form af et spørgeskema. Testen blev udviklet, så man for første gang så på læsning af såvel skønlitterære tekster som faglitterære tekster, men den inkluderede også læsning

af den type tekster, som kunne findes i faglitterære tekster, som fx skemaer, tabeller og grafiske fremstillinger, under et kaldet 'skematiserede tekster'. Endelig omfattede testen for 3. klasse – på dansk foranledning – også en kort, international version af Ordstillelæsningsprøven OS 400 (her OS 40) for også at få en fornemmelse for elevernes afkodningssikkerhed.

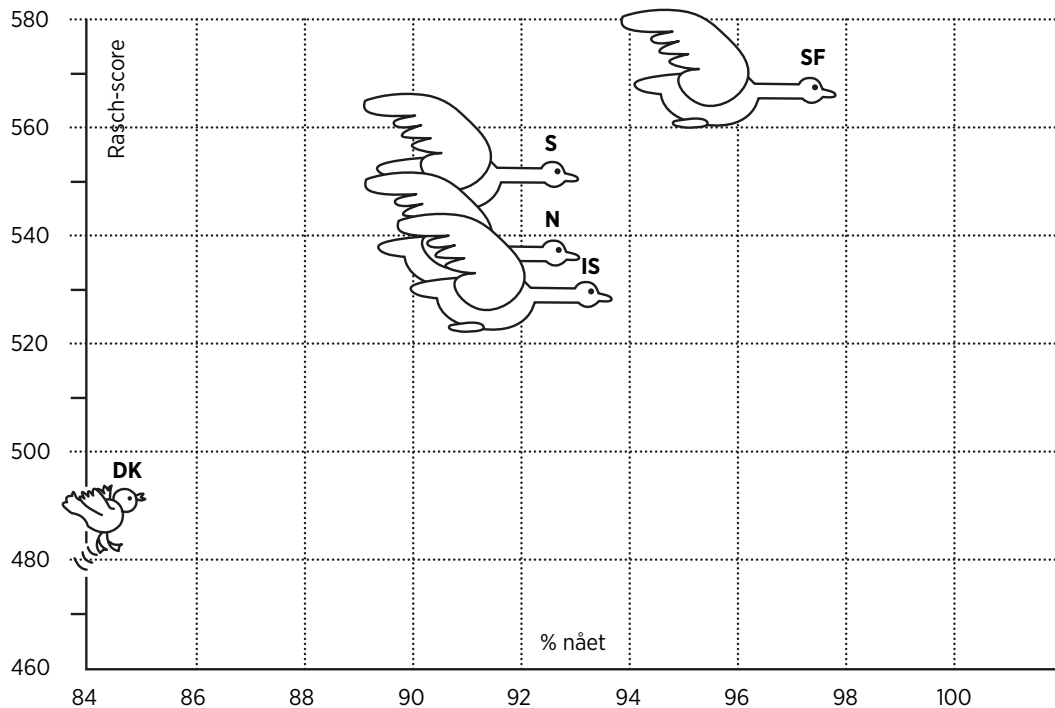
Da de første internationale resultater kom i efteråret 1992 (Elley, 1992), gav de stødet til nedsættelsen af et rådgivende udvalg i Undervisningsministeriets regi, som skulle se på, hvad der kunne gøres. Udvalgets rapport: *Råd til bedre læsning* (Elbro, Hansen, Mejding, Paagaard, & Raahauge, 1993) samlede en lang række anbefalinger på baggrund af de foreliggende resultater til, hvordan man kunne styrke læseundervisningen i skolen, og en række af disse anbefalinger blev siden fulgt.

Resultaterne var så lidt flatterende for Danmark, at de skulle gennemanalyseres, for at vi kunne stole på, hvad de viste. Derfor udkom den danske nationale rapport først efter supplerende analyser i 1994 – og den fik titlen: *Den grimme ælling og svanerne – om danske elevers læsefærdigheder* (Mejding, 1994).

“Sikkerhed” og “hastighed”, population B, resultater for de 14-årige.



“Sikkerhed” og “hastighed”, population A, resultater for de 9-årige.



Figur 2: Grafisk repræsentation af de danske resultater i den internationale læseundersøgelse i 1991 – set i de nordiske svaners perspektiv. Her efter Mejding (1997).

Den korte fortælling er, at de danske 3. klasseelever haltede langt efter de andre nordiske landes elever både med hensyn til læsehastighed og læsesikkerhed – men på 8. klassetrin nåede de danske elever næsten på højde med deres søstre og brødre i Norden, selv om de stadigvæk var lidt langsommere læsere.

Selv om de danske 3. klasseelever klarede sig dårligt i den nordiske sammenligning, så klarede de sig dog bedre, end lærerne forventede det af dem.

En selvstændig dansk undersøgelse af forventningerne til elevernes præstationer på baggrund af IEA-testens tekster viste, at de danske lærere og læseeksperter havde endog endnu ringere forventninger til, hvad et almindeligt niveau på 3. klasse burde være. Selv om de danske 3. klasseelever klarede sig dårligt i den nordiske sammenligning,

så klarede de sig dog bedre, end lærerne forventede det af dem (Mejding, 1991).

Resultatet gav som sagt genlyd både politisk og i skolen – og reaktionen er siden fejlagtigt blevet benævnt ”PISA-chokket”, selv om PISA-undersøgelsen faktisk først blev publiceret syv år senere.

IEA-undersøgelsen gav bl.a. stødet til det store udviklingsprogram i folkeskolen, kaldet Folkeskolen år 2000, hvor en lang række kommuner bød ind med projekter til forbedring af undervisningen på forskellige områder. En del i det program var en gentagelse af IEA's læseundersøgelse fra 1991 for at se, om nogle af de foranstaltninger, man havde truffet i de efterfølgende år, havde haft effekt (Allerup, Mejding, & Zeuner, 2001). Resultatet var meget positivt for 3. klasseeleverne, som nu læste med meget højere sikkerhed – helt på niveau med de svenske og norske resultater fra 1991, men dog stadigvæk lige så langsomt som i 1991. Eleverne i 8. klasse havde til gengæld tabt pusten, så de læste langsommere og mere usikkert end 9 år før. Interessant nok var denne årgang elever den samme

som den, der i år 2000 deltog i den første PISA-undersøgelse, hvor det kunne konstateres, at for stor en del af populationen læste så dårligt, at de måtte have svært ved at leve op til de læsekrav, skolen stillede (se nedenfor).

PISA – Programme for International Student Assessment (15-årige)

Hvor de undersøgelser, der gennemføres i IEA-regi, alle har en forankring i internationale forskningsmiljøer, så meldte der sig nu en anden spiller på banen. OECD, der er et politisk baseret system, kunne godt se perspektivet i de internationale undersøgelser inden for undervisningssektoren. Så de headhuntede en af forskerne bag analyserne af IEA's internationale læseundersøgelse fra Universitetet i Hamburg, Andreas Schleicher, til at stå for udviklingen af OECD's egen undersøgelse. Den skulle ligge på det tidspunkt, hvor eleverne var på vej ud af det obligatoriske skolesystem til enten arbejde eller videreuddannelse, og derfor valgte man at undersøge 15-årige, uanset hvilket klassetrin de gik på. Undersøgelsen skulle dække mindst tre 'fagområder': læsning, matematik og naturfag – og have plads til et fjerde vekslende, supplerende emneområde. Undersøgelsens opbygning gjorde, at man ikke kunne få oplysninger fra elevernes lærere, og der var heller ikke nogen oplysninger om klasserumseffekter, da eleverne blev udvalgt blandt alle 15-årige på skolen uanset klassetilholdsforhold.

Hvor IEA's undersøgelser i testkonstruktionen i forbindelse med matematik og naturfag ofte havde haft et udgangspunkt i landenes curriculumbestemmelser for fagene, skulle PISA-undersøgelsen i stedet tage udgangspunkt i de færdigheder, samfundet forventede, at de unge havde brug for, når de forlod grunduddannelsen. Til det blev der nedsat internationale ekspertkomiteer, som kom med rammebeskrivelser af, hvilke kernekompetencer man anså, at unge ville have brug for i et moderne samfund, og disse rammebeskrivelser blev styrende for konstruktionen af den faktiske test. Da politiske systemer oftere har en kortere tidshorisont end forskningsmiljøer, skulle undersøgelsen gennemføres hvert tredje år, og de tre fagområder skulle skiftes til at være 'hovedområde', mens de to andre blot blev gennemført på en mindre del

af eleverne. På den måde kunne man hurtigere få etableret nogle trendmålinger.

Læsning blev det første hovedområde for PISA-undersøgelsen i 2000, og læsedelen blev udviklet i et samarbejde mellem en nordisk, en australsk og en japansk forskergruppe. Rammebeskrivelsen for læsning kom til at ligge meget tæt på den samme beskrivelse for Den internationale læseundersøgelse – som siden blev til PIRLS (se nedenfor) – hvor man lægger mere vægt på læsekompetencens forskellige delaspekter end på, hvad forskellige skolesystemer har af curriculumbestemmelser.

Den samlede læsekompetence i PISA blev sammensat af resultaterne på de tre skalaer: at finde og gengive information, at fortolke teksten og danne sig sin egen mening på baggrund heraf og at reflektere over og vurdere tekstens oplysninger i relation til hele teksten og til almindelig viden. Da resultaterne blev rapporteret i 2001, fik den danske rapport titlen: *Forventninger og færdigheder*, da ingen på det tidspunkt kendte til PISA-undersøgelsen som begreb. Læseresultatet for de danske 15-årige var tæt på det internationale gennemsnit på 500 point. Bekymrende var det dog, at omkring 18 % af eleverne læste så dårligt, at man måtte forvente, at de ville have svært ved at klare de læsekrav, de ville blive mødt med, når de gik ud af skolen, mens kun 8 % af eleverne læste på det højeste niveau. For Finland, hvis elever var de bedste i undersøgelsen, var procenttallene henholdsvis 2 % (lavt niveau) og 19 % (højt niveau)..

Et af de åbne spørgsmål i forbindelse med PISA-undersøgelserne er, hvorfor de danske 15-årige elevers læsefærdigheder målt med PISA-undersøgelsen stort set ikke har forandret sig siden den første undersøgelse i år 2000, når nu vi ved, at læsefærdigheden for de yngste elever er forbedret meget i den samme periode (Mejding & Rønberg, 2008; Mejding & Rønberg, 2012). Men det er stadigvæk ikke slået igennem for de 15-årige, også selv om den første danske PIRLS-population, som klarede sig pænt internationalt i 2006, fyldte 15 år i 2015, og ellers burde have haft de bedste forudsætninger for også at opnå bedre resultater i PISA-undersøgelsen i 2015.

PISA-undersøgelsen blev mødt med megen kritik, både da den udkom og også sidenhen – og en del af problematikken er, at der er så mange skjulte variable i arbejdet, at det kan være vanskeligt at overskue helheden. Af hensyn til sammenligneligheden med de efterfølgende PISA-undersøgelser må selve testen nødvendigvis være fortrolig, da man ellers kunne risikere, at nogen gav sig til at undervise i netop de tekster og spørgsmål, som anvendes i testen. Derfor er det kun muligt at vise demonstrationstekster og -spørgsmål, og disse er oftest blandt de tekster og spørgsmål, som er sorteret fra, fordi de ikke var helt gode nok til at indgå i undersøgelsen. Men de kan bruges til at vise principperne i typen af tekster og den type af spørgsmål, man primært arbejder med.

PIRLS – Progress in International Reading Literacy Study (10-årige)

Da PISA nu havde overtaget undersøgelsen af de ældste elever, så valgte man i IEA at koncentrere sig om eleverne på mellemtrinnet, da man begyndte videreudviklingen af testen fra Den internationale læseundersøgelse. Hvor denne havde bestået af en række mindre, men uafhængige historier og fortællinger, med spørgsmål, som fortrinsvis blev besvaret med afkrydsning af ét blandt en række mulige svarmuligheder, så flyttede man populationen et år op, så det nu blev de 10-årige på 4. klassesetrin, som blev målgruppen. Samtidig ændrede man testformatet, så eleverne blev præsenteret for længere, sammenhængende tekster, så hver elev kun skulle arbejde med to tekster i alt. Spørgsmålsformaterne blev samtidig ændret, så kun ca. halvdelen af spørgsmålene skulle besvares i et 'flervalgs-format', mens resten var åbne spørgsmål, som eleverne selv skulle formulere et skriftligt svar til. Den første PIRLS-test blev afviklet i 2001 uden dansk deltagelse, da vi i Danmark i 2000 netop havde gentaget 1991-testen i forbindelse med 'Folkeskolen år 2000-projektet'. Men siden har Danmark deltaget i PIRLS i 2006, 2011 og 2016.

Vi har stadigvæk en for stor gruppe svage læsere og for lille en gruppe rigtigt gode læsere.

I PIRLS 2006 (Mejding & Rønberg, 2008) lykkedes det i Danmark at gennemføre en undersøgelse af sammenhængen mellem den oprindelige test fra 1991 og PIRLS 2006-udgaven. Derfor er vi helt unikt i stand til at drage sammenligninger helt tilbage fra 1991 over 2000 til den nye PIRLS-skala. Vi har derfor et rigtig godt billede af, hvordan danske læseresultater har udviklet sig i en international sammenhæng over en 25-årig periode. PIRLS-resultaterne har vist, at den fremgang i læsefærdigheder, vi så oven på det katastrofale resultat i 1991, er fortsat, til vi – i 2011 – var oppe blandt de fem bedste lande. I 2016 gik resultatet lidt tilbage igen, men de danske elever klarer sig stadigvæk rigtig pænt i gennemsnit. Dog har undersøgelserne også vist, at når vi klarer os godt, så er det især, fordi det er lykkedes at forbedre resultaterne for midtergruppen af elever. Vi har stadigvæk en for stor gruppe svage læsere og for lille en gruppe rigtigt gode læsere. Undersøgelserne har også vist, at der stadigvæk er endog meget store forskelle mellem klasseresultaterne – forskelle, som svarer til en spredning over tre skoleår. Igen og igen har resultaterne vist, at der ikke er nogen direkte sammenhæng mellem elevernes læseresultater og den tid, der har været afsat på skemaet til undervisning heri. Ikke mindst i 2012 – forud for skolereformen i 2013 – var dette et kontroversielt resultat, som Undervisningsministeriet forsøgte at skjule i pressemeddelelsen om undersøgelsen, hvor de – på trods af data – postulerede den modsatte konklusion.

I PIRLS 2016 er den danske del udvidet til at undersøge læsningen både på 3. og 4. klassesetrin, ligesom man ser på 4. klasseelevernes læsning af tekster på et simuleret internet. Som en af de eneste internationale undersøgelser har PIRLS i Danmark lagt de frigivne tekster fra alle undersøgelserne ud, så lærerne selv kan anvende testmaterialet i deres klasse, og de kan umiddelbart sammenligne klassens resultat med det danske landsresultat det år, hvor de pågældende tekster blev anvendt. Derved har læreren en mulighed for selv at sætte sig ind i forudsætningerne for de udmeldte testresultater, samtidig med at hun får et overblik over, hvordan hendes klasse kan klare at løse de opgaver, der bliver stillet. Dette materiale kan findes på DPU's hjemmeside www.edu.au.dk/pirls.

Test i dag

Hvordan man anvender test og prøver i skolen, afhænger således helt af, hvad resultaterne skal bruges til. Er fokus på de store linjer og en monitorering af tendenser over tid, så er der brug for de store undersøgelser med repræsentative populationer, som ikke nødvendigvis er de bedste, hvis man er mere interesseret i den enkelte elevs standpunkt. Der er i dag publiceret en række nye læseprøver, som kan give læreren rigtig gode oplysninger om, hvad eleverne magter at læse, og hvor de har deres svagheder. Men uanset hvilken test man ser på, så vil der altid være noget, som kan vække undren, eller som kan kritiseres. Det er derfor vigtigt altid at holde sig for øje, at et testresultat aldrig er hele sandheden. Det er et øjebliksbillede, en struktureret observation, som er påvirket af mange andre forhold end blot selve læsningen.

Det er derfor vigtigt altid at holde sig for øje, at et testresultat aldrig er hele sandheden.

Sætter man sig grundigt ind i de forskellige tests baggrund, formål og fortolkningsmuligheder, så har man det bedste udgangspunkt for at kunne vælge, hvilken test man kan anvende for at få resultater, der kan bruges i undervisningen. Selv rammebeskrivelserne af de internationale undersøgelser, der jo ellers ikke er konstrueret med et didaktisk sigte, kan give en god baggrund for fokuseringen af ens egen læseundervisning i klassen.

Referencer

Allerup, P., Mejding, J., & Zeuner, L. (2001). *Færdigheder i læsning og matematik – Udviklingstræk omkring årtusindskiftet*. F-2000 sekretariatet, København: Undervisningsministeriet.

Allerup, P., Jansen, M., & Weng, P. (2011). *Evaluering i skolen – baggrund, praksis, teori*. Frederikshavn: Dafolo.

Elbro, C., Hansen, E., Mejding, J., Pagaard, J., & Raahauge, J. (1993). *Råd til bedre læsning – Det rådgivende udvalgs rapport om læsefærdighed*. København: Undervisningsministeriet.

Elley, W. B. (1992). *How in the World do Students Read?* Hamburg: IEA.

Glæsel, B. (1992). Arne Søgård – En læseforsker fylder 80. *Psykologisk pædagogisk rådgivning*, 4, 251-253.

Hansen, M., Kreiner, S., & Pagaard, P. E. (1987). *Dansk – et fag i krise? Ålykke*.

Hansen, M. (1995). Dansk på dansk – eller hvorfor er der så mange, der ikke kan læse? *Kvan*, 41, 127-137.

Hansen, M. (u.å.). Skolepsykologi i *Den Store Danske*, Gyldendal. Lokaliseret 8. januar 2019 på: <http://denstoredanske.dk/index.php?side-Id=160273>

Hermann, K. (1955). *Om medfødt ordblindhed*. København: Munksgaard.

Jansen, M., & Ziegler, M. (1990). Projektets skolepolitiske baggrund – og danskundervisningens udvikling. I: Jansen, M., & Ziegler, M. (red.), *Tid til dansk 1 – fire års evaluering af et udviklingsarbejde* (s. 23-49). København: Danmarks Pædagogiske Institut og Folkeskolens Udviklingsråd.

Johansen, J., & Mejding, J. (1990). Læsefærdighed og undervisningsressourcer. I: Jansen, M., & Ziegler, M. (red.), *Tid til dansk 1 – fire års evaluering af et udviklingsarbejde* (s. 123-141). København: Danmarks Pædagogiske Institut og Folkeskolens Udviklingsråd.

Kirkebæk, B., & Skov, P. (u.å.). Specialundervisning i *Den Store Danske*, Gyldendal. Lokaliseret 9. januar 2019 på: http://denstoredanske.dk/Erhverv_karriere_og_ledelse/P%C3%A6dagogik_og_uddannelse/Specialundervisning/specialundervisning

Mejding, J. (1991). Læsefærdighed og danskundervisning. I: Johansen, J. (red.), *Debat om læseudvikling* (s. 33-49). København: Danmarks Pædagogiske Institut, 23.

Mejding, J. (1994). *Den grimme ælling og svanerne – om danske elevers læsefærdigheder*. København: Danmarks Pædagogiske Institut.

Mejding, J. (1997). Læsepræstationer i relation til hjem og skole. I: Frost, J., Sletmo, A., & Tønnessen, F. E. (red.), *Skriften på veggen* (s. 41-67). København: Dansk Psykologisk Forlag.

Mejding J. (2001). Læsekompetence. I: Andersen, A. M., Egelund, N., Pilegaard Jensen, T., Krone, M., Lindenskov, L., & Mejding, J. (red.), *Forventninger og færdigheder – danske unge i en international sammenligning* (PISA) (s. 54-90). København: AKF, DPU og SFI-Survey.

Mejding, J. (2009). Anvendelse af test og prøver i klassen. I: Bendixen, C., & Kreiner, S. (red.), *Test i Folkeskolen* (s. 173-196). København: Hans Reitzels Forlag.

Mejding, J., & Rønberg, L. (2008). *PIRLS 2006 – En international undersøgelse om læsekompetence i 4. klasse*. København: Danmarks Pædagogiske Universitetsforlag.

Mejding, J., & Rønberg, L. (2012). *PIRLS 2011 – En international undersøgelse om læsekompetence i 4. klasse*. Institut for Uddannelse og Pædagogik (DPU), Aarhus Universitet.

Mejding, J. (2013). PISA – ikke så ringe endda. *folkeskolen debat* 20.12.2013. Lokaliseret 12. februar 2019: <https://www.folkeskolen.dk/538349/pisa--ikke-saa-ringe-endda>

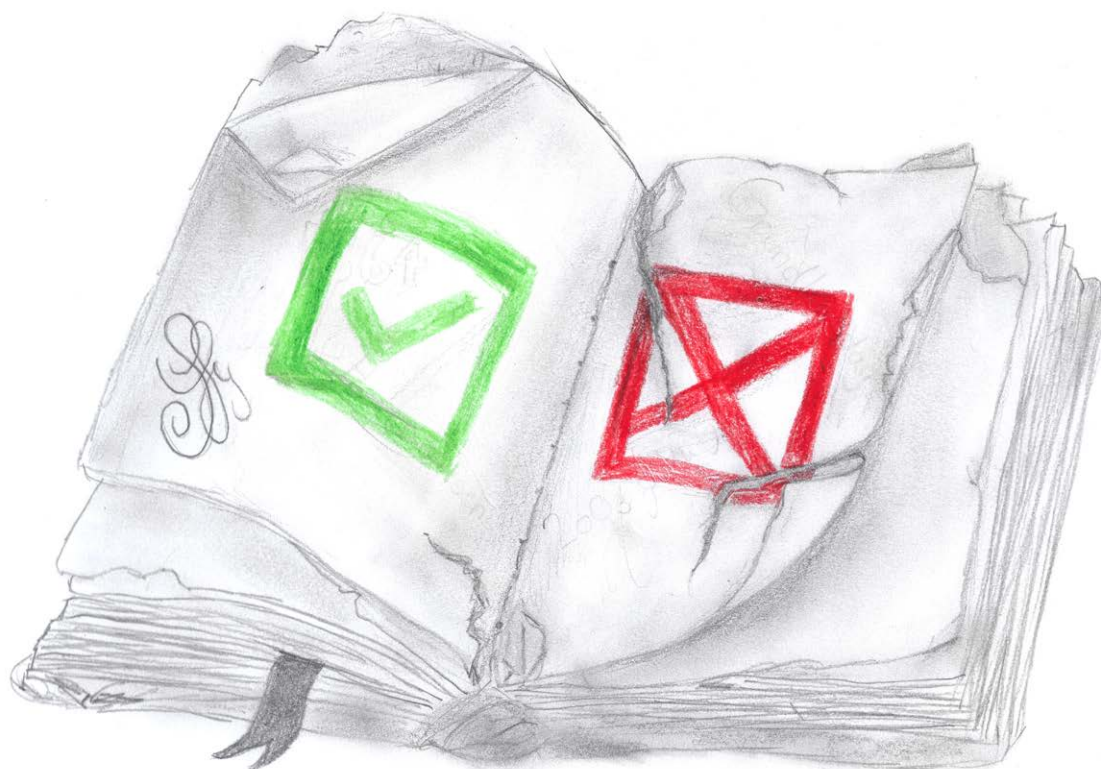
Norrie, E. (1960). *Ordblindeundervisning*. København: Nyt Nordisk Forlag Arnold Busck.

Paludan-Müller, D., & Mejding, J. (1988). Mor syr, og far læser avis. *Skolepsykologi*, 6, 541-546.

Søegård, A., & Bording Petersen, S. P. (1972). *Sætningsstillelæsningsprøven S50 – Vejledning*. København: Dansk Psykologisk Forlag.

Noter

- 1 'Tid til dansk'-rapporterne, Jansen, M., & Ziegler, M. (red.), rummer et væld af oplysninger af både kvantitativ og kvalitativ karakter. Interesserede henvises til den række af publikationer, der kom fra DPI i perioden 1990-1992.



Tegning: Julie Solbjerg-Nielsen Alber