

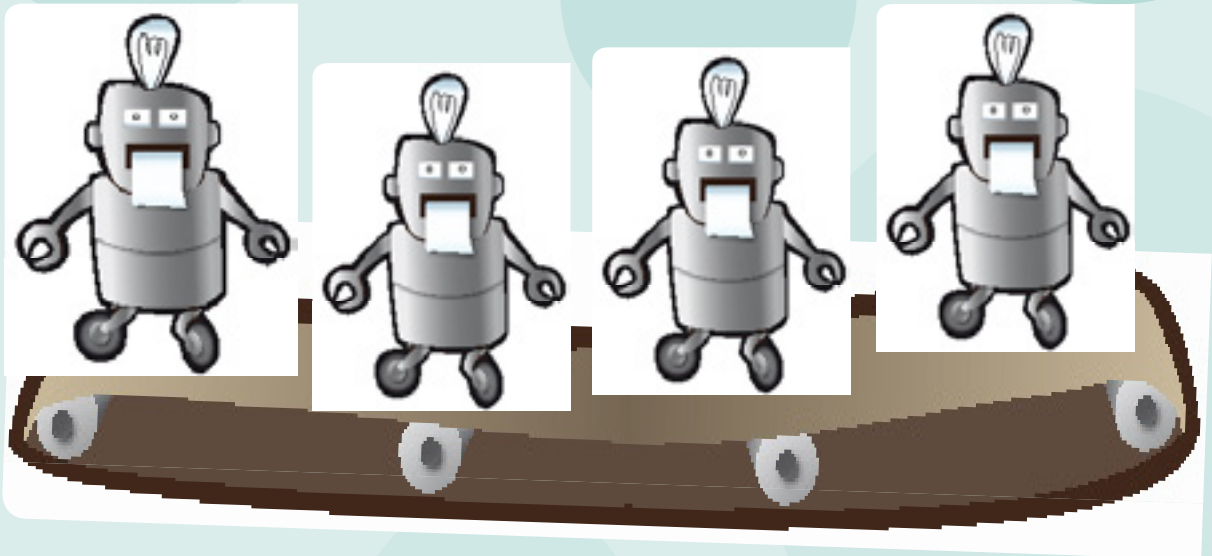


Evaluering og **skole-** **modenhedstest**

på Emdrupborg forsøgsskole

Christian Ydesen

Ph.D. studerende ved DPU



Evaluering og selektion er ikke nogen ny foreteelse i den danske folkeskole. Denne artikel kaster et blik på evaluerings- og selektionspraksissen på Emdrupborg forsøgsskole, der blev etableret i 1948. Artiklen peger på, at evaluering og selektion af børn ved anvendelse af test ikke bør ske på grundlag af blot én test. Endvidere påviser artiklen, at lærerevalueringer let kan komme til at gøre brug af underlødige og problematiske kategoriseringer og forestillinger om årsagssammenhænge. Og endelig fremgår det, at der i menneskelig praksis let kan opstå afsmittende virkninger mellem forskellige evalueringsteknologier, som på forskellig vis vil have betydning for skole- og læringsmiljøet.

Indledning

Det er efterhånden en veletableret myte, at den danske folkeskole historisk set har lidt af evalueringsforskrækkelse. Myten hænger sammen med den megen tale om tidligere tiders rundkredspædagogik og fravær af evalueringskultur i kølvandet på PISA-chokket. Et blik i skolehistorien viser dog et noget andet billede. Fra 1920'erne blev standardiserede standpunktsprøver i stigende grad anvendt til at klarlægge elevers, klassers og skolers faglige niveau, og en kombination af minutøse lærerevalueringer og tests spillede en central rolle i folkeskolens forsøgsarbejder.

Efterkrigstidens højborg for en sådan omfattende evaluerings- og testpraksis var Emdrupborg forsøgsskole, der blev etableret i 1948 under ledelse af skolepsykolog Anne Marie Nørvig.

I denne artikel vil jeg se nærmere på evalueringspraksissen på Emdrupborg forsøgsskole i almindelighed og på den anvendte skolemodenhedsprøve i særdeleshed. Emnet har relevans, fordi Emdrupborg forsøgsskole kom til at spille en afgørende rolle for udviklingen af den danske folkeskole – ikke mindst den banebrydende 1958-skolelov – og fordi et nærmere studie af test- og evalueringspraksis på Emdrupborg formår at kaste et perspektiv på den fremtrædende rolle, som forskellige evalueringsteknologier, såsom logbog og nationale tests, spiller i nutidens folkeskole.

Emdrupborg forsøgsskole

Emdrupborg forsøgsskole var den første forsøgsskole i Danmark, og den kom til at fungere som et foregangseksempel og et pædagogisk fyrtårn ud over Danmarks grænser. Denne rolle skyldtes ikke mindst talrige besøg fra udenlandske og indenlandske undervisere, forskere og pædagoger samt det forhold, at skolen udgav og distribuerede adskillige pædagogiske rapporter, hvori skolens forsøgsresultater nøje blev beskrevet.

Forsøgsskolen arbejdede på et evidensbaseret og naturvidenskabeligt inspireret grundlag. Børnenes evner var genstand for objektive måleresultater. Skolens psykologer forsøgte med andre ord at "(...) finde frem til lovmæssigheder, som kan være af værdi for under-



visningens tilrettelæggelse og arbejdet i klasserne, samt for videre undersøgelser inden for området." (Nørvig, 1955, s. 31). Laboratorieforsøget blev tilstræbt som et erkendt uopnåeligt ideal, og det var en central værdi, at skolens arbejdsmetoder og resultater kunne gøres til genstand for såkaldt objektiv kritik.

Etableringen af en evalueringspraksis

Med det overordnede formål at omsætte pædagogiske og psykologiske teorier til praksis søsatte skolen det hidtil mest omfattende test- og evalueringsprogram i Danmark. Skolens psykologer med Nørvig i spidsen udgjorde skolens ledelse, idet alle sager passede skolepsykologerne eller deres private sekretærer. Samtidig administrerede skolepsykologerne de mange evalueringsskemaer samt børnenes sagsmapper.

Forsøgsskolen gjorde brug af en lang række forskellige faglige og psykologiske tests i samspil med meget omfattende lærer- og psykologbedømmelser. Testene

blev anset for en nødvendig komponent til at supplere evalueringsprogrammets mere subjektive elementer, såsom lærerudtalelser og lærernes udfyldelse af evalueringskemaer.

Ud af disse tests blev Uppsala-skolemodenhedsprøven anvendt med direkte konsekvenser for eleverne. Uppsala-prøven blev brugt til at beslutte, hvorvidt et førskolebarn kunne begynde i skolen, og der kan derfor være god grund til at se nærmere på praksissen omkring anvendelsen af denne test i selektionen af skolebørn.

Uppsala-skolemodenhedstesten

Det at anvende en særlig test til at afgøre et barns grad af skolemodenhed var en relativ ny foreteelse i den danske folkeskole. Begrebet 'skolemoden' var ganske vist blevet brugt allerede i mellemkrigstiden, men det var først efter 2. verdenskrig, at skolemodenhed blev udskilt som et særligt område, hvor test kunne gøres anvendelige. Denne udvikling viser, at test havde undergået en udvikling fra at operere på et generelt niveau med fokus på intelligens, læsefærdigheder mv. til et mere specialiseret niveau med det formål at afdække sammensatte evner hos børn på et bestemt alderstrin.

Uppsala-skolemodenhedsprøven var en prognostisk gruppeprøve, som tog sigte på at afdække evner forbundet med intelligens med henblik på at identificere børn, der med stor sandsynlighed ville klare sig dårligt i skolen. Men prøven blev også anvendt til at undgå overvægt af problembørn i de nye førsteklaser.

Uppsala-prøven var udviklet af den svenske lærer Gösta Levin, og den blev anvendt på forsøgsskolen fra skolens opstart. I de følgende år arbejdede skolens psykologer tæt sammen med Levin med det formål at udarbejde en dansk standardisering af prøven. Dette arbejde skal ses i lyset af, at prøven ikke kun blev anvendt på Emdrupborg, men på mange danske skoler. I 1958 blev således ca. 10 % af danske skolebørn testet med Uppsala-prøven.

Uppsala-prøven tog 45-50 minutter at gennemføre med 15 børn, og den krævede tilstedeværelsen af to lærere –

en instruktør og en observatør. Instruktørens rolle var at instruere muntert og venligt, men efter en nøje fastlagt instruktion. Observatørens rolle derimod var at iagttage børnene og deres arbejdsmetoder og notere sine observationer ned på barnets evalueringskort.

Det var et krav, at begge lærere skulle have stor erfaring med små børn, og de skulle vide, hvornår de skulle hjælpe et barn, og hvornår barnet skulle overlades til sig selv. Dette skabte et klart element af personligt skøn i prøvesituationen; ikke mindst i kraft af observatørens kategoriseringer af barnet, som udelukkende var baseret på iagttagelser af barnets kropssprog, holdning og eksplícitte følelser.

Samtidig afhang prøvens resultater af, at det var lykkedes at skabe en legende, hyggelig og uhøjtidelig atmosfære. Evalueringskortene fra Uppsala-prøven afslører imidlertid, at børn ofte græd som følge af testsituationen. Gråd medførte et negativt resultat på evalueringskortet, da dette blev anset for et tegn på, at barnet ikke var skolemodent. At skabe den rette atmosfære i forhold til hvert enkelt barn skulle derfor ofte vise sig at være en vanskelig foreteelse.

Uppsala-prøven bestod af 15 ikke-verbale opgaver, der ikke krævede nogen læsefærdigheder. Barnet skulle opfatte og skelne mellem forskellige former og figurer, gengive former og figurer, finde ligheder og forskelle på tegninger af diverse genstande, skelne mellem enslydende lyde, arbejde med symboler, huske via synet, huske via hørelsen, opfatte og følge en simpel talrække, udvise initiativ og koncentration, følge en instruktion, færdiggøre et bestemt tankemønster samt arbejde med papir og blyant.

Et centralt element i designet og rækkefølgen af testopgaverne var at sørge for, at børnene ikke blev trætte, og at opgaverne i stedet kom til at virke interessante og sjove for barnet. Levin forsøgte at opnå dette ved at bruge relativt korte og varierede testopgaver. Men Levin bemærkede selv, at træthed generelt afspejlede sig i et fald i korrekte svar i anden halvdel af testopgaverne, hvilket tyder på, at fastholdelse af børnenes koncentration var en betydelig udfordring (Levin, 1951, s. 144).

Den praktiske gennemførelse af Uppsala-prøven tyder derfor på, at der var betydelige fejlkilder til stede i testsituationen, såsom skabelsen af den rette atmosfære, observatørens subjektive bedømmelser, problemer med træthed og manglende koncentration, børnenes kendskab og erfaring med testopgaver og testsituationer samt børnenes forventningsfuldhed og usikkerhed i deres første møde med skolen. Men ud over disse mulige fejlkilder indeholdt prøven også opgaver, hvis løsning i betydelig grad afhang af barnets køn og sociale baggrund. Prøven lagde meget stor vægt på analytiske evner, initiativ og kendskab til figurer og tal. Samtidig opererede prøven med et relativt smalt normalitetsbegreb i forhold til, hvordan et hus ser ud, hvilke genstande der hører til i bestemte situationer, og hvilke genstande der skiller sig ud fra andre genstande.

Evalueringsprogrammet

Som nævnt var Uppsala-prøven en del af et større evalueringsprogram. Når et barn blev indskrevet på forsøgsskolen, blev der registreret meget omfattende informationer. Dette skete ved hjælp af spørgeskemaer til forældrene, barnets sundhedsplejerske og børnehave med det formål udtømmende at afdække barnets første leveår samt afklare risikoen for arvelige skavanker. Al denne information blev indsamlet ud fra ensartede kriterier og indskrevet i barnets sagsmappe. Samtidig blev barnet tildelt en unik sekscifret kode for at udgå forvekslinger og sikre anonymitet i skolens publikationer. Denne minutløse indsamling og registrering af informationer fortsatte gennem hele barnets skoletid.

Til belysning af børnenes udvikling indsamles i løbet af skoletiden følgende oplysninger			Før skoletiden	Klasse										
				1.	2.	3.	4.	5.	6.	7.	8.	9.		
				Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	Efterår Forår	
Standard-prøver	Oplysninger	Fra forældre	•											
		Fra børnehaver	•											
	Grundprøve	Uppsalaprøver	•											
		Individuel-prøver	Binet-Simon		•									
			Startprøve		•									
	Højt læsning			•	•	•	•	•	•	*	•			
	Gruppe-prøver	Stille læsning				•	•	•	•	•	•	•	•	
		Retskrivning				•	•	•	•	•	•	•	•	
		Regning				•	•	•	•	•	•	•	•	
	Vurdering og beskrivelse	Journalist				•	•	•	•	•	•	•	•	
Rapport					•	•	•	•	•	•	•	•		
Prøver med kontrolgrupper	Gruppe-prøver	Raven							*	*	*	*		
		I-prøve							*	*	*	*		
		Blandet regning							•	•	•	•		
		Retskrivning							•	•	•	•		
		Stillelæsning							•	•	•	•		
		Orienteringsfag							•	•	•	•		
		Almen Orientering							•	•	•	•		
		Engelsk								*	*	*		
		Tysk									*	*		
Matematik										*				

* Har været anvendt på forskellige trin - ligger ikke fast

Kilde: Nørvig 1955, s. 32.



Ud over test og prøver bestod forsøgsskolens evalueringsprogram af yderligere to elementer. For det første udarbejdede lærerne to gange om året en evalueringsrapport. Den første rapport gik på arbejdet i klassen, mens den anden rapport gik på den enkelte elevs arbejde og udvikling. For det andet var der den årlige journalliste, der beskrev en lang liste af faktorer med relevans for det enkelte barn, såsom omfang af undervisningstid ydet barnet, barnets interesse for skolen, koncentrationsevne, arbejdstempo, ordenssans, selvtillid, samarbejdsevne, stabilitet, udtrykssevne, manuelle færdigheder, motoriske udvikling, andel i forekommende uro, mangel på præcision, lederevner, fritidsinteresser, initiativ samt bedømmelse af barnets faglige standpunkt i skolefagene.

Det er endvidere bemærkelsesværdigt, at journalisterne også indeholdt en vurdering af hvert enkelt barns hjemmemiljø. Således blev også forældrenes interesse for skolen, forældrenes evne til lektiehjælp, hjemmets økonomiske forhold (herunder om hjemmet bestod af en eneforsørger), hjemmets orientering angående almene forhold, barnets fremtræden, opdragelsen og hjemmets orden (specielt m.h.t. husorden) genstand for lærernes bedømmelse.

Det umiddelbare formål med disse journalister var at sikre en udtømmende evaluering af det enkelte barn, henlede lærerens opmærksomhed på problemområder, identificere børn med problemer, sikre et komplet overblik over børnene samt at beskrive ikke målbare

faktorer. Det ledende princip for udfyldelsen af journalisterne var normalfordelingskurven, hvilket betød, at hvert fokusområde skulle bedømmes separat. Børnene skulle vurderes i relation til deres kammerater og ikke i forhold til eksterne standarder. Men brugen af normalfordelingskurven viser, at et matematisk princip blev ophøjet til at være det ordnende princip. For at sikre en brugbar indplacering på en normalfordelingskurve bestod evalueringsskemaet af fem trin med betegnelserne A til E. I den medfølgende instruktion fra skolens psykologer stod det ligefrem anført, at fordelingen skulle være som følger: A - 2, B - 7, C - 12, D - 7, E - 2.

Et godt eksempel er journalistens kategori syv "barnets fremtoningspræg":

- A** Udseende og adfærd tyder på stor orden og regelmæssighed i hjemmet, så perfekt, at det grænser til det tvangsagtige.
- B** Barnets udseende og adfærd tyder på sirlighed, aldrig eller sjældent svigtende orden og regelmæssighed.
- C** Barnets fremtoningspræg tyder på en pæn, men ikke overdreven orden.
- D** Tyder på en noget sjusket og ligegyldig indstilling.
- E** Udseende og adfærd tyder på elendighedspræget roderi.

Det forhold, at journalisterne indeholdt en så omfattende bedømmelse, indikerer klart, at skolens interesse ikke blot var barnets skolegang. Barnet blev set som en refleksion af en social kontekst med behov for hjælp og kontrol. Dette afspejles videre i lærernes evalueringsrapporter, hvor lærerne kunne beskrive børnene med egne ord. I den følgende tabel har jeg udvalgt nogle eksempler fra evalueringsrapporter i skoleåret 1956-57 om forskellige børn på forskellige klassetrin for at give et indtryk af de værdier og forestillinger, som var fremherskende i evalueringsprogrammet på forsøgsskolen. Det skal således bemærkes, at det ikke er det samme barn, som optræder i hver linje.

Klasse	Begavelse	Krop og fremtoning	Hjemmet	Arbejdsetik
1.	Velbegavet. Jævnt godt begavet. Meget velbegavet. Ret velbegavet. Meget kvik.	En tragisk original. Typisk enebarn. Hans tøj er pjaltet. Let på tå.	Et meget uinteresseret miljø. Forsømt både fysisk og psykisk. Hun er fra et hjem med store modsætninger.	Meget virksom og praktisk. Dygtig og omhyggelig.
2.	Lidt tung. Noget tung. Lidt tung begavelse. Helt godt begavet.	Lidt tøse- og duksedrengeagtig. Synger som en engel. Gødheden lyser ham ud af øjnene. Spilopmager. En rigtig drenge-pige. En lille spirrevip.	Blinde forældre. Enebarn.	Sjusket og hurtig med skriftlige arbejder. Ligeglad med alt. Overfladisk og ugidelig. Passer sit arbejde upåklageligt. Arbejder godt og grundigt.
3.	Velbegavet. Ret gode evner.	En lille spinkel fyr med vat i skuldrene og "he-man" idealer og påtaget grov stemme. Et lidt uheldigt udseende.	Uinteresseret hjemmemiljø.	Yder kun en ringe indsats. Overfladisk indstilling til sit arbejde. Meget passiv og optaget af uvedkommende ting.
4.	Ualmindelig velbegavet. Meget kvik. Velbegavet. Middel-begavet. Sikkert en relativ lille begavelse.	Spejderdrengetype. Don Juan. Flyveører. Blomsterdreng. Har et ubehageligt blik, der virker uærligt. Virker irriterende med sin kæln, gammelkloge snak. En drabelig slagsbroder. En pudsig lille fyr. En lille mimose. Typisk "dreng". Hendes skæbne er tragisk.	Brogede familieforhold. Faderen er alvorlig og tager tingene ret tungt. Ret vanskelige hjemmeforhold Af jødisk herkomst. Påvirkning fra hjemmet har hæmmet ham. Hjemmet er meget harmonisk. Stærkt religiøst hjem.	B bliver en god, solid arbejdsmand. Tjenstvillig og sikkert hvad landboerne forstår ved et rigtig kvikt københavnerbarn. Ukoncentreret. Pligtopfyldende. Flittig. Ikke bange for at bestille noget. Uselvstændig. Arbejder støt og koncentreret.
5.	Begavelse: middel. Begavelse: under middel eller på grænsen af middel. Begavelse: middel eller lidt over. Velbegavet. Ikke ubegavet.	Mærkeligt "feminint" væsen. Tyren Ferdinand. En rigtig dreng. Forhutlet indtryk. Behagelig i enhver henseende. Han er spejder og har derfor et fast greb på tingene.	Får mange formaninger, men eftersom han ligner sin mor, er det svært at ændre.	Flid: stærkt varierende.

Egenskaber

Svinger mellem den yderste barnagtighed og en ofte forbløffende modenhed.

Tiltalende og fornuftig.
Kan ikke tænke fornuftigt.

Uhyre larmende med et kæmpeorgan.

Pylrehoved.

Hendes menneskelige kvaliteter lader noget tilbage at ønske.

Sjov og sprællevende.

Klassens "yndige" pige – og hun ved det.

Meget forvirret og ukoncentreret.

Vigende i sin karakter.

Velmenende og uegenlyst.

Enspændernatur.

Gode lederegenskaber.

Klassens mest urolige.

Meget snakkesalig, nærmest kvænnende.

Sød og flittig, men dødelig og lidt slask.

Lederemne.

Klassens talent med visesang, akrobatik og rulleskøjter på repertoire.

Vælges som tillidsmand.

Godt arbejde i filmklubben.

Valgt til elevrådet næste år.

Forvirret og uden ståsted.

Fuldkommen uden sans for humor.

En lille, lidt farveløst virkende pige.

Fine sjælelige kvalifikationer.

Pattebarnsagtig.



De udvalgte uddrag fra lærerevalueringerne bekræfter billedet af, at lærerne tog sigte på en udtømmende evaluering af hvert enkelt barn, og at de ikke afstod fra at bedømme barnets hjem eller endda lave årsagssammenhænge om barnet og dets forældre baseret på forestillinger om arvelighed.

I kategorien "krop og fremtoning" er det slående, hvordan lærerne abonnerede på alle mulige stereotyper, såsom "en rigtig dreng", "en tøsedreng", "en spejderdrengetype", "Don Juan" og "tyren Ferdinand". Disse stereotyper synes i vid udstrækning at være centreret omkring køn og indikerer dermed, at der var et kønsmæssigt element på spil i lærernes opfattelser og forventninger til det enkelte barn. Lærerevalueringerne afslører dermed, at forestillinger om køn havde betydning for barnets skole- og læringsmiljø. Indholdet af disse evalueringer er slående, for så vidt de afslører et udpræget fokus på fremtoning og udseende blandt mange af forsøgsskolens lærere.

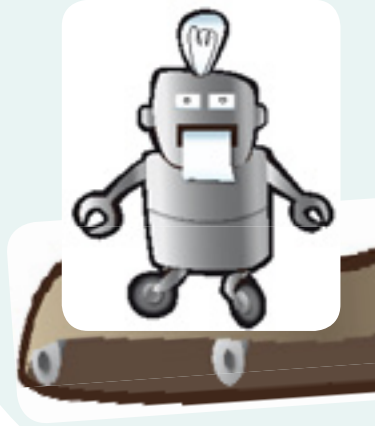
Ud fra de negative bemærkninger om børn, som menes at besidde bestemte egenskaber og en bestemt arbejdsetik, er det klart, at præcision, orden, stabili-

tet og flid fremstår som forsøgsskolens centrale værdier.

Bedømmelsen af begavelse er interessant i lyset af den anvendte Binet-Simon intelligens test. Et gennemsyn af de tilgængelige, halvårslige evalueringer viser, at forestillingen om en intelligenskvotient var et tilbagevendende element i forhold til hvilket, lærerne søgte at positionere sig og argumentere. Især børnenes intelligenskvotienter var noget lærerne orienterede sig imod. Det er især bemærkelsesværdigt, at lærere somme tider forsøgte at fremkomme med en skønnet intelligenskvotient i tilfælde, hvor barnet endnu ikke var blevet intelligentstestet. Dette er bl.a. tydeligt i anvendelsen af udtrykket 'udnyttelsesgraden'. En lærer skrev: "Hans IK er målt til 119, jeg havde vurderet ham til omkring 100, så noget må vel efterhånden vise sig." (Ydesen 2010, s. 186).

Historiske erfaringer med nutidig relevans

Eksemplet Emdrupborg viser, at test og evaluering blev anset for selve forudsætningen for at bedrive skoleforsøg i efterkrigsårene. Evaluering- og testpraksis blev aldrig grundlæggende udfordret



i 1940'erne og 1950'erne, hvilket kan skyldes, at evaluering og test bekræftede den eksisterende skolepraksis. Uppsala-prøven afslørede, at nogle børn var skolemodne, mens andre ikke var, og Binet-Simon testen afslørede betydelige forskelle i børns intellektuelle evner. Begge test var omgærdet af en høj grad af troværdighed. Evaluering og test bekræftede på den måde den konventionelle viden og tilvejebragte videnskabeligt belæg for retfærdigheden i skolepraksis.

Selvom forsøgsskolens psykologer besad betydelig viden om fejlkilder i de anvendte tests, så synes kvantificeringens besnærende kvaliteter, såsom muligheden for umiddelbar sammenlignelighed, at have overtrumfet reservationer i den daglige praksis. Det gennemgående evalueringsprogram gødede jorden for brugen af test, og der kom efterhånden til at eksistere et nærmest symbiotisk forhold mellem evalueringsprogrammets teknologier. Samtidig var evalueringsprogrammet gennemsyret af værdier omkring visse egenskaber og arbejdsetik, såsom præcision, stabilitet og flid. I den forbindelse er det bemærkelsesværdigt, at Uppsala-prøven indeholdt betydelige fejlkilder, såsom lærerens/psykologens personlige skøn, problemer med træthed og koncentration, nervøsitet, antagelser om tidsbegrænsningens neutrale og ensartede virkning på børnene samt børnenes kendskab til og erfaring med testopgaver og testsituationer. På grund af de egenskaber, som krævedes for at klare sig godt i testene, reproducerede Uppsala-prøven sociale og kønsmæssige forskelle. Testpraksis på Emdrupborg var ingenlunde udtryk for en neutral og objektiv evalueringsteknologi.

Fejlkilderne øvede desuden indvirkning på lærerevalueringerne, hvor stereotype forestillinger om det enkelte barn er tydelige. Disse stereotyper afspejler lærerens opfattelser, og dermed fik de betydning for lærerens forventninger til det enkelte barn. Det snævre normalitetsbegreb, som på hver sin måde var til stedet i forsøgsskolens selektionsteknologier, fik betydelig indflydelse på skole- og læringsmiljøet på Emdrupborg.

Hvis man på den baggrund skal forsøge at uddrage nogle nutidige erfaringer fra evalueringspraksissen på Emdrupborg, så tyder eksemplet med Uppsala-prøven på, at evaluering og selektion af børn ved anvendelse af test ikke bør ske på grundlag af blot én test – dertil er der ganske enkelt for mange fejlkilder. Specielt i forhold til skolemodenhed er det særdeles vanskeligt at afgøre, hvem der er modne, og hvem der ikke er, fordi det i forhold til så små børn er umådelig vigtigt med en tryk og afslappet atmosfære, for at barnet kan få gode arbejdsbetingelser.

Erfaringen fra Emdrupborg viser også, at lærerevalueringer – selv inden for meget nøgterne og ordnede rammer – let kan komme til at gøre brug af underlødige og problematiske kategoriseringer og forestillinger om årsagssammenhænge.

Endelig peger eksemplet på, at der i menneskelig praksis let kan opstå afsmittende virkninger mellem forskellige evalueringsteknologier, som på forskellig vis vil have betydning for skole- og læringsmiljøet.

Litteratur

Levin, G. (1951): "Upsala-undersökningen av nybörjare". I Meyer, H.; Gregersen T. (red.): *Skolemodenhedsproblemer* (pp. 141-154). Munksgaard.

Nørvig, A. M. (1955): *Beretning om Emdrupborg skoles første 6 år 1948-1954*. Det Danske Forlag.

Ydesen, C. (2010): *The Rise of High-Stakes Educational Testing in Denmark, 1920-1970*. Ph.d.-afhandling. Danmarks Pædagogiske Universitetsskole, Aarhus Universitet.

