

De nationale test – positiv eller negativ wash- back på undervisningen?

Faglig evaluering i folkeskolen har først og fremmest til hensigt at afdække elevernes kundskaber og færdigheder i forhold til de trinmål, der er opstillet for de forskellige fag. De evalueringsredskaber, der indgår som en integreret del af undervisningen, skal derfor kunne give et billede af, hvad eleverne har lært, og hvad de ikke har lært, og give et indblik i deres faglige progression. De nationale test i folkeskolen er udviklet med henblik på at give lærere et systematisk evalueringsredskab, der skal benyttes som ét blandt mange redskaber i den løbende evaluering af elevernes kundskaber og færdigheder. Det er imidlertid ikke nationale test og andre evalueringsformer, der i sig selv gør eleverne fagligt dygtigere – det gør derimod lærernes pædagogiske opfølgninger på test- og evalueringresultaterne. Netop derfor er det vigtigt, at lærere har et solidt indblik i forholdet mellem læringsmål og testmål i de nationale test, så de ved, hvad testen tester, og hvad den *ikke* tester. Denne viden er en forudsætning for at kunne anvende resultaterne fra de nationale test som et pædagogisk redskab i undervisningen og dermed for positiv washback.

Hensigten med denne artikel er at give et lille indblik i forholdet mellem læringsmål og testmål i den nationale test i engelsk og diskutere nogle af de faktorer, der har indflydelse på, om de nationale test får en positiv eller negativ indvirkning på lærernes undervisning og elevernes læring. Artiklen er et uddrag fra kapitlet »De nationale test i et pædagogisk perspektiv« (Stæhr 2009), som optræder i bogen *Test i folkeskolen*.



LARS STENIUS STÆHR
Projektleder, Novo Nordisk
larsstaehr@gmail.com

Summativ og formativ evaluering

Når man vil se nærmere på nationale test i folkeskolen, er det hensigtsmæssigt at skelne mellem to overordnede formål med at evaluere: *summativ* evaluering og *formativ* evaluering.

Summativ evaluering er, som navnet angiver, en opsummerende bedømmelse, der finder sted ved afslutningen af et kortere eller længere undervisningsforløb. Denne form for evaluering har som formål at afdække, hvor meget eleverne har lært i undervisningsforløbet, og i hvor høj grad de lever op til de mål, der er opstillet for undervisningen. Den er således en slags status over eller kontrol af, hvad eleverne har lært. Et eksempel på summativ evaluering er Folkeskolens Afgangsprøve, der netop har som formål at teste, i hvilken grad elevernes viden og færdigheder inden for et givent fag lever op til slutmålene for faget.

Hvor summativ evaluering således finder sted ved afslutningen af et undervisnings- eller uddannelsesforløb og kan beskrives som *produktorienteret*, foregår formativ evaluering løbende under et undervisningsforløb og er i højere grad *procesorienteret*. På engelsk skildres denne distinktion i udtrykkene »assessment of learning« og »assessment for learning«, altså som forskellen mellem evaluering af læring og evaluering med henblik på læring. Formativ evaluering refererer således til lærerens løbende bedømmelser af elevernes udvikling, og den har bl.a. som formål at identificere deres styrker og svagheder inden for et givent område. Resultaterne fra formative evalueringer giver læreren et grundlag for løbende at tilrettelægge sin undervisning på en måde, der tilgodeser elevernes forskellige forudsætninger og behov. Derudover kan den feedback, som gives i forlængelse af formative evalueringer, bidrage til, at eleverne bliver bevidste om deres styrker og svagheder, således at de kan fokusere deres indsats mod de områder, hvor de har problemer.

Trinmål og de nationale test

De nationale test i folkeskolen er tænkt som et formativt evalueringsredskab. Testene adskiller sig imidlertid fra andre formative redskaber ved at være centralt tilrettelagt og obligatoriske, og alene af den grund er det essentielt, at lærere bliver bevidste om, hvilke læringsmål der testes, og hvordan testresultaterne kan anvendes pædagogisk i undervisningen og bidrage til, at elevernes differentierede behov tilgodeses.

De nationale test er it-baserede, selvscorende og adaptive¹, og de er udarbejdet med udgangspunkt i de trinmål, som gælder for fagene på de pågældende klassetrin. Trinmålene er udformet som en række undervisningsmål, der beskriver, hvad undervisningen skal lede frem mod. De angiver de kundskaber og færdigheder, som eleverne forventes at have tilegnet sig på et specifikt klassetrin, og som undervisningen derfor skal planlægges ud fra. Hvis vi tager udgangspunkt i faget engelsk, er læringsmålene både rettet mod engelsk som et kommunikationsmiddel og engelsk som et indholds-fag, og dette dobbelte formål er afspejlet i beskrivelserne af de fire centrale kundskabs- og færdighedsområder (CKF): *kommunikative færdigheder, sprog og sprogbrug, sprogtilegnelse, kultur- og samfundsforhold*.

Når man kigger nærmere på de enkelte trinmål for de fire CKF-områder efter 7. klasse, hvor testen gives, bliver det tydeligt, at langt fra alle trinmål er direkte målbare i en it-baseret test. Trinmål som fx at »fremlægge et forberedt emne«, »udtrykke sig spontant mundtligt« eller »skrive tekster med begyndelse, midte og slutning« vil ikke kunne testes i de nationale test, idet der kræves, at eleven har mulighed for at producere et længere mundtligt eller skriftligt produkt, og dette produkt er ikke automatisk scorbart, men skal bedømmes kvalitativt af en eksaminator og en censor ud fra veldefinerede relevante kriterier.

Grunden til at en lang række af trinmålene ikke kan testes direkte, er, at de er *undervisningsmål* og ikke *testmål*, og de er dermed ikke formuleret med henblik på at kunne blive operationaliseret i en it-baseret test. Det betyder, at de nationale test langt fra vil kunne evaluere den brede vifte af kundskaber og færdigheder, som indgår i de forskellige fag. Netop derfor er det vigtigt, at læreren er bevidst om præcist, hvilke trinmål der helt eller delvist, direkte eller indirekte indgår i testen, og hvilke der ikke gør, og at læreren ligeledes anvender andre evalueringsmetoder i den løbende, formative evaluering.

I udviklingen af de nationale test har det været hensigten at udvælge de trinmål (eller dele af trinmål), som rent faktisk *kan* testes i en it-test, og på denne baggrund dække faget så bredt som muligt. I forhold til målbeskrivelserne for engelsk i 7. klasse kunne det eksempelvis involvere trinmål som at »forstå skrevne tekster herunder skønlitterære og enkle sagprosattekster«, »stave centrale og hyppigt forekommende ord« og »forstå (...) centrale sproghandlinger«. Derudover kunne man ligeledes forestille sig, at dele af et trinmål som fx »tale og skrive engelsk, således at grammatiske regler om især ordklasser, ental og

flertal og tidsangivelse ved hjælp af verbernes former følges« kunne testes indirekte, dvs. ved at afdække elevernes viden om og evne til at anvende grammatiske regler i korte sætninger.

De trinmål eller dele af trinmål, som direkte eller indirekte indgår i de nationale test for faget engelsk, er fordelt på tre profilområder, som er tilstræbt at afspejle forskellige centrale aspekter af fagets kundskabs- og færdighedsområder. Det drejer sig om profilområderne 1) Læsning, 2) Ordforråd samt 3) Sprog og sprogbrug.

Profilområder og kendeord

Det overordnede formål med de nationale test er således at afdække elevernes styrker og svagheder inden for disse tre profilområder. Hver enkelt testopgave i de nationale test er forsynet med oplysninger om tilknyttet profilområde og trinmål og indeholder ligeledes en angivelse af, hvilket konkret aspekt af elevernes kundskaber og færdigheder opgaven tester. Denne angivelse sker i form af de såkaldte *kendeord*. Eksempler på kendeord inden for profilområdet *læsning* er: Hovedindhold i tekster; formål og funktion i dagligdags tekster; logiske følgeslutninger; genrer; specifikt indhold i tekster; fakta og holdninger i tekster. Kendeordene kan betragtes som en slags testmål, og de er centrale, når læreren skal give sine elever deltagende feedback på deres styrker og svagheder inden for et givent profilområde. Det er således vigtigt, at læreren har overblik over kendeordene inden for de forskellige profilområder, eftersom de angiver, hvordan et profilområde konkret er blevet operationaliseret. Ved at danne sig et overblik over kendeordene inden for profilområderne kan læreren dermed få et indblik i, hvad den pågældende test tester, og hvad den *ikke* tester.

Tilbage melding på testresultater

Når eleverne har taget en test, kan lærerne få adgang til deres resultater i form af en række forskellige rapporter. Disse rapporter rummer en stor mængde information om elevernes præstationer, som kan støtte lærerne i deres tilrettelæggelse af undervisningen. Rapporterne giver eksempelvis læreren adgang til klassens samlede præstation for hele testen og for de enkelte profilområder. Rapporterne giver ligeledes læreren adgang til informationer om den enkelte elevs præstation i forhold til de tre profilområder – og i forhold til specifikke faglige områder, kundskaber eller færdigheder.

Ifølge Undervisningsministeriets evalueringsportal² er hensigten med de forskellige rapporter, at de skal opfylde tre overordnede funktioner. De skal kunne danne udgangspunkt for 1) overvejelser over faglige mål, 2) overvejelser over lærings- og undervisningsformer og 3) refleksion over og handling i forhold til tidligere og fremtidig undervisning.

Positiv eller negativ washback?

Der er empirisk belæg for, at formativ evaluering fører til læring og faglig progression (OECD, 2005). Spørgsmålet er imidlertid, om de obligatoriske nationale test, der er tiltænkt en formativ funktion, vil fremme elevernes faglige udvikling. Det er med andre ord nødvendigt at forholde sig til, hvorvidt indførelsen af nationale test vil have en positiv eller negativ indflydelse (*washback*) på lærernes undervisning og elevernes læring. Når man indfører nationale testsystemer, der ensarter måden, hvorpå man evaluerer elevernes kundskaber og færdigheder, og som gør det muligt at sammenligne elevers, klassers og skolers resultater med hinanden, vil testene uundgåeligt blive bestemmende for de krav, man stiller til eleverne i undervisningen. Også selvom man gør det tydeligt for lærerne, at testene kun kan teste et begrænset udvalg af de trinmål, som elever og lærere skal arbejde hen mod. Der er dermed stor sandsynlighed for, at testene vil have en eller anden form for washback, dvs., at de bliver styrende for, hvad der kommer til at foregå i undervisningen. Man kan tale om positiv washback, hvis testenes indflydelse medfører en række hensigtsmæssige ændringer i undervisningspraksis, som fremmer elevernes læring inden for en række tilsigtede områder. Der er derimod tale om negativ washback, hvis testene har en række u hensigtsmæssige påvirkninger på undervisningen, og hvis de ikke afspejler det lærings syn, de principper og de mål, der er for undervisningen. Det er imidlertid ikke altid indlysende, om en given test har en positiv eller negativ indflydelse på undervisningen og elevernes læring, og det vil i nogen udstrækning afhænge af øjnene, der ser. Diskussionen om, hvorvidt en tests washback er positiv eller negativ, vil naturligvis være præget af forskellige parters holdninger til, om de intendede virkninger er ønskværdige eller ej.

I artiklerne »For og imod nationale test« og »Nationale test og læring – nogle empiriske argumenter« redegør Nordenbo (2007; 2008) for en række af de centrale positioner i den internationale debat om indførelsen af nationale testningssystemer. En oversigt

over internationale undersøgelser af washback og pædagogisk brug af test kan ligeledes findes i et forskningsreview foretaget af Dansk Clearinghouse for Uddannelsesforskning (Nordenbo m.fl. 2009). Her vil jeg dog nøjes med at pege på nogle enkelte forhold, som potentielt kan føre til negativ washback i forbindelse med de nationale test i Danmark.

Der er en risiko for, at der ikke er overensstemmelse mellem de former for viden og færdigheder, som kan testes i de nationale test, og de former for viden og færdigheder, der er mest repræsentative for det pågældende fagområde. Dette kan så medføre, at netop de former for viden og færdigheder, som bliver testet i de nationale test, opprioriteres i undervisningen på bekostning af andre, mere centrale aspekter af fagområdet, dvs., at testen bidrager til at skævvride undervisningens fokus på en u hensigtsmæssig måde (se fx undersøgelser af Alderson & Hamp-Lyons (1996) og Ferman, Watanabe & Curtis (2004)). Netop derfor er det vigtigt, at lærere har et indgående kendskab til, hvilke aspekter af fagets mål de nationale test kan teste, og i høj grad hvilke aspekter de *ikke* kan teste. Som jeg tidligere har nævnt i denne artikel, er det kun et begrænset udsnit af trinmålene for faget engelsk, der bliver dækket i den engelske test. Når lærere skal anvende resultaterne fra de nationale test til at tilrettelægge og differentiere undervisning, opstille læringsmål og indsatsområder og give feedback til elever og forældre, er det centralt, at de er bevidste om, at de nationale test langt fra giver hele billedet. Hvis ikke lærere handler ud fra denne bevidsthed, kan de nationale test få en meget uheldig virkning på undervisning og læring i folkeskolen. Hvis en skævvridning af undervisningens fokus skal undgås, skal lærere således have et præcist indblik i forbindelsen mellem profilområder, trinmål og kendeord i de nationale test, og de skal på den baggrund anvende andre formative evalueringsformer, som kan bidrage til at give et mere komplet billede af elevernes faglige progression, fx inden for mundtlig og skriftlig kompetence.

Hvis man fra Undervisningsministeriets side vil sikre en positiv washback på undervisning og læring, er det således vigtigt, at de rapporter, som lærerne har adgang til, bliver mere detaljerede, og at lærerne får en mere konkret og fagspecifik indføring i, hvordan disse informationer kan tolkes og anvendes pædagogisk. Kun hvis dette sker, kan lærerne anvende de nationale test som et formativt evalueringsredskab med pædagogisk værdi. Dette vil ligeledes kunne bidrage til at forhindre, at der sker *teaching to the test*, dvs., at

lærere ensidigt underviser for at forberede eleverne til testen ved bl.a. at anvende testopgaver som undervisningsmateriale, og at dette sker på bekostning af andre centrale fagområder – som for eksempel arbejdet med elevernes mundtlige og skriftlige kompetencer.

Risikoen for at undervisningen bliver testrettet og dermed potentiel risiko for skævvridning af undervisningens fokus er større, hvis der er tale om en såkaldt *high-stakes*, som eksempelvis en afgangsprøve, hvorimod risikoen er mindre, når der er tale om en test, der skal bruges som formativt evalueringsredskab. Så hvis de nationale test udelukkende var udviklet som et pædagogisk redskab til brug i den løbende interne evaluering af eleverne, ville det mindske risikoen for negativ washback og øge chancen for positiv washback. Der er imidlertid lagt op til, at testresultaterne også skal have en anden anvendelse, idet både skoleleder og kommunalbestyrelse skal have adgang til resultaterne på forskellige måder. Grunden til dette er bl.a., at testresultaterne for den enkelte skole ligeledes skal anvendes i kommunens årlige kvalitetsrapport for kommunens skoler. Ud over at skulle bruges i pædagogisk øjemed er testresultaterne dermed tiltænkt en anvendelse, der har med kontrol og kvalitetsmåling at gøre. Set fra et washback-perspektiv er denne anvendelse problematisk, idet den kan medføre et pres på lærere om, at deres elever klarer sig godt i de nationale test. Konsekvensen kan være, at de nationale test flytter sig fra at være et *low-stakes* evalueringsredskab med pædagogisk funktion til at blive en form for *high-stakes* test med deraf følgende øget risiko for negativ washback på grund af testrettet undervisning og skævvridning af undervisningens fokus.

Hvis de nationale test i Danmark skal have en positiv virkning på elevernes læring, er det i lyset af ovenstående en forudsætning, at lærere har et indgående kendskab til forholdet mellem trinmål, kendeord og profilområder, og at de dermed ved, hvad testen tester, og hvad den ikke tester. Dette er nødvendigt for at lærerne kan tilgode alle målene for faget og ikke kun dem, der bliver afdækket i de nationale test. Derudover er det en forudsætning, at lærerne supplerer med andre formative evalueringsformer, som kan følge op på de elevmålsætninger som de nationale test ikke tester. Og sidst, men ikke mindst, er det en forudsætning, at lærerne har konkret indsigt i, hvordan de kan tolke de forskellige testrapporter og anvende testresultaterne som udgangspunkt for det videre pædagogiske arbejde med elevernes læring.

Noter

1 Det betyder, at sværhedsgraden af testspørgsmålene tilpasser sig den enkelte elevs faglige niveau i testforløbet. Hvis en elev svarer rigtigt på to opgaver i træk, får eleven en sværere

opgave næste gang, og hvis eleven svarer forkert på to opgaver i træk, stilles efterfølgende en lettere opgave.

2 *Opfølgning på testresultaterne*, side 8.

Litteratur

- Alderson, J. C., & Hamp-Lyons, L. (1996): TOEFL Preparation Courses: A Study of Washback. *Language Testing* 13(3), 280-297.
- Ferman, I., Watanabe, Y., & Curtis, A. (2004): The Washback of an EFL National Oral Matriculation Test to Teaching and Learning. I: L. Cheng & Y. Watanabe (red.): *Washback in language testing. Research contexts and methods* (s. 191-210). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Nordenbo, S. E. et al (2009): *Pædagogisk brug af test – Et systematisk review. I: Evidensbasen*. København: Dansk Clearinghouse for Uddannelsesforskning, DPU, Aarhus Universitet.
- Nordenbo, S. E. (2008): Nationale test og læring. *Kognition og pædagogik*, 67, 6-10.
- Nordenbo, S. E. (2007): For og imod nationale test. I: R. Engelhardt, C. Mehlsen & F. Stjernfelt (red.): *Tankestreger – tværvenskabelige nybrud*. København: Danmarks Pædagogiske Universitetsforlag.
- OECD (2005): *Formative assessment: Improving learning in secondary classrooms*. OECD Publishing.
- Stæhr, L. S. (2009): De nationale test i et pædagogisk perspektiv. I: C. Bendixen & S. Kreiner (red.): *Test i folkeskolen*. København: Hans Reitzels Forlag.