

eRÅDGIVNINGSGRUPPENS INDSTILLING VEDR. EVALUERING AF DE NATIONALE TEST

25. OKTOBER 2018

Rådgivningsgruppen noterer sig, at der i kommissoriet lægges op til brede perspektiver vedrørende DNTs styrker og svagheder. I det lys har rådgivningsgruppen dels formuleret specifikke forslag til evalueringsspørgsmål, dels anslået temaer, som har præget rådgivningsgruppens arbejde indtil nu, og som gruppen kan ønske at vende tilbage til.

Rådgivningsgruppen har afholdt tre heldagsmøder. Rådgivningsgruppen har drøftet forskeroplæg fra Danmark og Norge. Rådgivningsgruppen har været repræsenteret på to dialogmøder med skolefolk og samlet op på synspunkter herfra. Rådgivningsgruppen har med udgangspunkt i kommissoriet tilstræbt den bredest mulige inddragelse af mangfoldige erfaringer og perspektiver afbalanceret med hensynet til et afgrænset og fokuseret evalueringsoplæg.

Rådgivningsgruppen opfordrer ministeriet til i øvrigt at søge inspiration i de forslag til evalueringstemaer og –metoder, som fremgår af deltagernes bidrag til rådgivningsgruppens arbejde, herunder det sammendragende mødepapir dateret 10. august.

I. SPECIFIKKE EVALUERINGSSPØRGSMÅL

Samtidig med, at rådgivningsgruppen i videst muligt omfang ønsker samtlige evalueringsspørgsmål besvaret, giver gruppen i en tentativ prioritering særskilt høj prioritet til spørgsmål 3.1.

Endvidere har gruppen ønsket at give prioritet til 1.1, 1.2, 1.4, 2.1, 3.3, 3.5, 3.6, 4.3, 4.4, 5.2, 5.3 og 6.2. De prioriterede spørgsmål er indrammet herunder.

1.1. Regner DNT rigtigt? De konstaterede problemer vedr. sikkerhed og reliabilitet i DNT kan skyldes, at de sværhedsgrader, som den adaptive algoritme bruger ikke længere er korrekte og i værste fald, at opgaverne ikke længere passer til den såkaldte Rasch model. Det er derfor vigtigt, at evalueringen af DNT omfatter en undersøgelse af, om de sværhedsgrader, som var gældende tidligere stadig er gyldige. En sådan undersøgelse behøver ikke indsamling af nye resultater, men kan gennemføres ved analyse af data fra de seneste års obligatoriske test.

1.2 Kan sikkerheden i målingerne af elevernes færdigheder forbedres ved at kombinere resultater fra forskellige profilområder? En af årsagerne til usikkerheden i DNT er, at der kun afsættes kort tid til hvert enkelt profilområde. Hvis testresultaterne fra de forskellige profilområder kan påvises at måle forskellige aspekter af en og samme bagvedliggende færdighed kan sikkerheden forbedres ved at slå testresultaterne sammen på den rigtige måde. En undersøgelse af denne mulighed kan også på en enkel måde gennemføres ved hjælp af data fra de seneste års obligatoriske test.

1.3 Er den adaptive algoritme så god som den kan blive eller er det muligt at forbedre algoritmen, så usikkerheden i DNT reduceres? En undersøgelse af denne mulighed kræver kun adgang til eksisterende

oplysninger om opgavernes sværhedsgrader og forudsætter ikke analyse af nye eller eksisterende data. Evalueringen af den adaptive algoritme, bør også indeholde en omhyggelig kontrol af om implementeringen af algoritmen er foregået uden programmeringsfejl.

1.4. Findes der test i Danmark eller andre lande, som

- med samme tidsforbrug måler mere præcist end DNT?
- har større målingsvaliditet i fht. hvad der i Danmark er beskrevet i rammedokumenter som skolens formål og målsætninger?

2.1. I hvilket omfang er der sammenhæng mellem på den ene side skolens formål og målsætninger beskrevet i rammedokumenter og på den anden side indholdet i DNT? Her kan spørges om, hvorvidt DNT måler noget vigtigt eller hvilke dele af fagets indhold, som DNT tester og ikke tester. Her kan både tænkes på en gennemgang af fagligt indhold og på en empirisk analyse af samvariation i elevers præstationer i DNT og andre aspekter af fagenes formål.

3.1. Oplever lærerne, at den information som DNT producerer, er nyttig og handlingsvejledende i det pædagogiske arbejde? Herunder: Hvad er informationsværdien af oplysninger knyttet til de nuværende profilområder? Og ligeledes: I hvilket omfang giver de kriteriebaserede resultater oplysninger, som lærerne kan bruge til at vurdere, hvor eleverne ligger i forhold til skolens formål og målsætning?

3.2. Hvad ved lærerne om nationale tests funktionsmåde, fortolkning og anvendelsesmuligheder? Og hvordan har de tilegnet sig denne viden?

3.3. Hvordan forstår lærerne, hvad DNT resultaterne viser? I hvilken grad har de gjort sig bekendt med vejledninger? I hvilken grad svarer vejledningerne til det, som de nationale test viser? I hvilken grad oplever lærerne, at den information som DNT producerer, er nyttig og handlingsvejledende i det pædagogiske arbejde? Og hvad er sammenhængen mellem ovenstående faktorer?

3.4. Oplever lærere særlige, konkrete problemer vedr. testsituationen og/eller testens indhold på særlige klassetrin eller særlige fag? Og: Har det sammenhæng med hvilket klassetrin eller hvilke fag, hvor DNT måler mest upræcist eller i ringest grad svarer til fagenes formål?

3.5. Når DNT fokuserer på delaspekter af specifikke fagområder, hvordan præger valget af disse delaspekter så den måde som det enkelte fag opfattes, realiseres og prioriteres? (jf. punkt 2.1). Hvordan oplever lærerne at DNT (testens eksistens, indhold og form) påvirker lærernes undervisning, dels i klassen, dels i forhold til den enkelte elev? Og på såvel positive som negative måder?

3.6. Hvordan er evalueringspraksis på skolerne (herunder hvilke typer instrumenter bruger lærere til at få indblik i elevernes viden, kunnen, kompetencer og trivsel, hvordan foregår dialoger omkring resultaterne, hvordan indgår resultaterne i den pædagogiske praksis ifht. eleverne)? Hvad er læreres oplevelse af sammenhængen mellem DNT på den ene side og på den anden side det pædagogiske arbejde og skolens evalueringspraksis i øvrigt? Og hvad er brugernes oplevelse af udviklingen af evalueringspraksis over de seneste år?

3.7. I hvilken grad har lærere kendskab til oplysninger om usikkerhed i DNT og andre test, hvordan opfatter de i givet fald de oplysninger og hvilke konsekvenser drages heraf?

3.8. Oplever lærere DNT som støttende og/eller kontrollerende? Og hvordan kan variationer heri forklares?

4.1. Oplever ledere DNT som støttende og/eller kontrollerende? Og hvordan kan variationer heri forklares?

4.2. Hvad ved skolelederne om nationale tests funktionsmåde, fortolkning og anvendelsesmuligheder? Og hvordan har de tilegnet sig denne viden?

4.3. I hvilket omfang og hvordan arbejder skoleledere og lærere sammen, eventuelt under inddragelse af vejledere og konsulenter, om at tolke testresultater og bruge dem i forhold til udvikling af den pædagogiske praksis?

4.4. Hvordan anvender og oplever skoleledelsen DNT i ledelsesarbejdet og i skolens evalueringspraksis i øvrigt?

4.5. I hvilken grad oplever kommunale skolechefer og skoleledere, at DNT, eventuelt sammen med anden evalueringssinformation, leverer relevant og brugbar information, som understøtter samarbejdet om kommunale indsatser og skoleindsatser?

4.6. I hvilken grad understøtter DNT arbejdet på det politiske niveau (i kommunalbestyrelsen og i politiske udvalg)?

5.1. Hvordan oplever elever testsituationen?

5.2. Hvad er elevernes oplevelse af, hvordan lærerne forklarer om DNT, herunder testsituationen, DNTs formål og anvendelse? Og hvordan er sammenhængen med elevernes oplevelse i testsituationen (spg 5.1)?

5.3. Får eleverne feedback efter test, og i så fald, kan de bruge den i forhold til deres egen læring? Indgår elever og lærere i samarbejde om opfølgning på test?

5.4. Hvad er DNTs betydning for elevernes opfattelse af skolen og undervisningen?

6.1. Oplever forældre at få udbytte af information om testresultater?

6.2. Hvordan vurderer forældre (som vidner til elevens reaktioner og lærerens information) fordele og ulemper ved DNT?

7.1. Hvordan vurderes det danske testsystem (herunder det adaptive princip, faglig relevans og kvalitet af opgaver samt muligheder for pædagogisk brug) med internationale øjne i lyset af andre landes erfaringer? Har andre lande, f.eks. nordiske lande, erfaringer med at stille pædagogisk anvendelige test til rådighed for skoler og lærere?

II. DISKUSSIONSTEMAER

Rådgivningsgruppen har i sit arbejde drøftet en række temaer, som knytter sig til DNT. Rådgivningsgruppen udtrykker ikke pt. nogen stillingtagen hertil, men kan ønske at vende tilbage til disse temaer såvel som til andre og nye temaer, når tiden kommer til anbefalinger. Disse temaer er blandt andet:

- a. Behovet for tekniske rapporter og anden dokumentation om DNTs opbygning, opgavetyper og beregningsmåder.
- b. Behovet for at drøfte governance-strukturer angående test, herunder hvordan tilbagemeldinger fra praksis kan bruges i løbende forbedring af DNT.
- c. Behovet for at afdæmpe teaching to the test, herunder forskellige aktørers ansvar i den forbindelse.
- d. Testens tidsmæssige placering i skoleåret.
- e. Muligheden for at benytte afdækningstest, udredningstest og/eller en vifte af test og evalueringsredskaber, der dækker bredere i forhold til de faglige områder og kompetencer/arbejdsformer, og som giver lærerne information de kan bruge i tilrettelæggelsen af undervisningen.
- f. Muligheden for at forholde sig til bestemte former for anvendelse af DNT.
- g. I hvilket omfang DNT fortsat bør baseres helt eller eventuelt delvist på det adaptive princip.
- h. I hvilket omfang DNT er tilpasset eller kan tilpasses til skolens funktioner i et foranderligt samfund.
- i. Behovet for specifikke former for forsknings- og udviklingsarbejde i tilknytning til DNT.

III. INSPIRATION TIL MINISTERIET VEDR TEMAER OG METODER

Det anbefales, at evalueringen inddrager eksisterende danske undersøgelser af DNT.

Ministeriet opfordres til at søge yderligere inspiration vedr. undersøgelsestemaer og metoder i de øvrige bidrag fra rådgivningsgruppens medlemmer, blandt andet som disse fremgår af sammendraget af evalueringstemaer og –metoder dateret d. 10. August.

Der opfordres til, at almindelige metodiske krav overholdes i evalueringsarbejdet.