

Tobias Bøgeskov Eriksen, Andrej Christian Lindholst
og Søren Valgreen Knudsen

Elevernes oplevelse af nationale test: En spørgeskemaundersøgelse af elever fra strategisk udvalgte folke- skoler i Aalborg Kommune

I denne artikel komplementerer vi tidligere forskning og undersøger, hvordan folkeskoleelever i 8. og 9. klasse opfatter de nationale test, samt hvordan opfattelsen hænger sammen med faglighed, (mis)trivsel og køn. Undersøgelsen bygger på en spørgeskemaundersøgelse og fokusgruppeinterviews gennemført i 2019 med elever i Aalborg Kommune. Vi finder, at store elevgrupper har en overvejende negativ opfattelse af de nationale test, og at elevernes humør påvirkes negativt før og efter en test. Samtidig finder vi signifikante sammenhænge mellem en mere negativ opfattelse af de nationale test og henholdsvis en svagere faglighed, større mistrivsel og køn (piger i forhold til drenge). Sammenhængene indikerer, at det tilsyneladende er hos elevgrupper, hvor behovet for faglig understøttelse er størst, at holdningen til de nationale test er mest negativ. Samlet set påpeger vi relevansen af at inddrage bredere perspektiver på effekterne af de nationale test, sætter spørgsmålstejn ved værdien af de nationale test som læringsredskab og underbygger behovet for yderligere undersøgelser af, hvordan de nationale test påvirker eleverne – personligt, socialt såvel som fagligt.

Søgeord: De nationale test, evaluering, folkeskolen, kønsforskelle, mistrivsel, faglighed.

Indledning

Jeg synes, at de er håbløse, og jeg har hørt det der med, at hvis man svarer de første tre forkert, så har man fucket hele prøven op. Det er lidt tamt (folkeskoleelev i 8. klasse om de nationale test).

Fra sidste halvdel af 2000'erne og frem til i dag har »de nationale test« (DNT) været afprøvet og implementeret i den danske folkeskole som et nyt obligatorisk test- og evalueringsredskab af elevernes faglige kunnen. DNT repræsenterede en nyskabelse i forhold til de i forvejen eksisterende redskaber, dels teknisk, fordi DNT er digitale og adaptive test, dels politisk, fordi det er obligatorisk for alle skoler at teste eleverne med DNT på fastlagte tidspunkter i deres skoleforløb. De politiske begrundelser for implementeringen henviser til ønsker om at løfte fagligheden, især for de fagligt svageste elever, styrke evalueringskulturen i folkeskolen samt styrke mulighederne for en central styring af folkeskoleområdet (Børne- og Undervisningsministeriet 2020a). Der er i forbindelse med implementeringen udarbejdet et stadigt større vidensgrundlag om erfaringerne, virkningerne og anvendeligheden af DNT, som blandt andet har medvirket til, at DNT har været stærkt omdiskuteret i både faglige og politiske kredse (VIVE 2020; Andersen & Nielsen 2019; Bundsgaard og Kreiner 2019; Wandall m.fl. 2018; Beuchert & Nandrup 2017; Beuchert & Nandrup 2014; Rambøll 2013).

En samfundsmæssigt vigtig, men samtidig overset dimension i det eksisterende vidensgrundlag er dog en belysning af DNT, der tager udgangspunkt i folkeskoleelevernes egne perspektiver og oplevelser. Folkeskolen er én af de sociale sammenhænge, hvor en stor del af børns udvikling finder sted. Her danner eleverne forståelser af sig selv og omverdenen og får erfaringer, som tages med videre igennem hele livet (Andreasen 2019). Under den nuværende indretning af test- og evalueringsredskaber i folkeskolen vil samtlige elever stifte bekendtskab med DNT. Forskningen påpeger, at disse test kan spille en meget direkte rolle i forhold til det enkelte barns etablering af syn på såvel egne som andres evner og potentialer (Andreasen 2019). De få eksisterende undersøgelser (VIVE 2020; Wandall m.fl. 2018; Andreasen 2015; Rambøll 2013; Kousholt 2012), som i en eller anden grad belyser DNT ud fra folkeskoleelevernes egne perspektiver, er dog delvist modstridende eller bygger på mindre udsnit af folkeskoleelever. Endvidere finder vi, at der mangler danske undersøgelser, som belyser sammenhænge mellem folkeskoleelevernes oplevelse af DNT og vigtige baggrundsforhold såsom køn, (mis)trivsel og faglighed.

Med et eksplorativt udgangspunkt har denne artikel til formål at øge vores eksisterende forståelse af DNT set ud fra folkeskoleelevernes egne perspektiver ved at belyse de to følgende forskningsspørgsmål:

- 1) Hvad er elevernes generelle holdning og oplevelse af DNT?

	Tobias Bøgeskov Eriksen Konsulent i Ældre- og Handicapforvaltningen i Aalborg Kommune E-mail: tobias@grover.dk		Andrej Christian Lindholst Lektor, Institut for Politik og Samfund, Det Samfundsvidenskabelige Fakultet Aalborg Universitet E-mail: acl@dps.aau.dk	
				Søren Valgreen Knudsen Læge, ph.d.-studerende, Klinisk Institut, Det Sundhedsvidenskabelige Fakultet, Aalborg Universitet E-mail: soerenk@dcu.aau.dk

- 2) Er der væsentlige forskelle i elevernes oplevelse af DNT i forhold til køn, faglighed og (mis)trivsel?

Grundlaget for artiklen er data fra en spørgeskemaundersøgelse og fokusgruppeinterviews gennemført i 2019 blandt folkeskoleelever i 8. og 9. klasse i strategisk udvalgte folkeskoler i Aalborg Kommune. Ambitionen med artiklen er at bidrage med et komplementært perspektiv i det nuværende vidensgrundlag om anvendelsen og virkninger af DNT i den danske folkeskole – et perspektiv, som også bør indgå i pædagogiske og politiske overvejelser om folkeskolens generelle udvikling.

Baggrund

Traditionelt har den danske folkeskole været præget af et grundtvigiansk værdisæt, som har betonet historisk bevidsthed, værdifællesskab, fortælling og folkelighed (Dahler-Larsen 2008). Fra begyndelsen af 1990'erne har den grundtvigianske folkeskoletradition dog været under opbrud. Opbruddet har særligt været inspireret af en angelsaksisk skoletradition, hvis grundpiller er baseret på test og en mere summativ evalueringsform (Schou m.fl. 2015). Forandringen ligger i tråd med en mere generel udvikling i den offentlige sektor, hvor organisering, ledelse og styring siden 1980'erne i stigende grad har været præget af New Public Management (NPM)-inspirerede reformer med vægt på valgfrihed, konkurrencemekanismer, resultatmålinger og decentralisering af ansvar (Torfing & Triantafillou 2017).

Internationale læseundersøgelser dokumenterede i 1990'erne, at danske elever læste langsommere end elever fra andre lande (Dahler-Larsen 2008). Det blev anledningen til en transformation i den danske folkeskole, hvor definitionen af kvalitet i højere grad har beroet sig på summative evalueringsformer og kvantitative målinger, der kan sammenlignes nationalt som internationalt (Ibid). I 2003 blev det vedtaget ved lov, at folkeskolernes afgangskarakter skulle offentliggøres for at fremme det frie skolevalg og sikre større gennemsigthed, så forældrene kunne vælge, hvilken skole deres barn skulle gå på ud fra karaktergennemsnittet (Undervisningsministeriet 2005). I 2004 anbefalede OECD, at Danmark skulle etablere en bedre evalueringskultur, fordi læringsmålene ikke blev indfriet i tilstrækkelig grad. Desuden viste PISA-undersøgelser fra 2000, 2003 og 2006, at danske folkeskoleelever ikke var i nærheden af toppen (Schou 2009). I 2010 blev anvendelsen af DNT gjort obligatorisk i alle folkeskoler. En del af den politiske hensigt var at styrke folkeskolens evalueringskultur ved at tilvejebringe bedre pædagogiske værktøjer og derigennem styrke elevernes faglige niveau (Børne- og Undervisningsministeriet 2020a).

Der har været rejst flere kritikpunkter mod DNT af blandt andre Bundsgaard & Kreiner (2019), og Bundsgaard & Puck (2016). Det første kritikpunkt omhandler testens måletekniske egenskaber og prædiktive validitet. Kritikken går på, at testen sværhedsgrad ikke har været tilstrækkeligt tilpasset den enkelte elev; at testresultaterne ikke giver tilstrækkeligt præcise målinger af en elevs faglige niveau; at der har manglet opgaver med tilstrækkelig sværhedsgrad for de dygtigste elever, samt at den statistiske usikkerhed på testresultater har været for stor (Bundsgaard & Kreiner 2019). I forlængelse af denne kritik viser en analyse af de statistiske sammenhænge mellem resultater i DNT i 8. klasse og folkeskolens afgangsprøver i 9. klasse i VIVE's evalueringsrapport fra 2020 (VIVE 2020c), at trods relativt stærke generelle statistiske sammenhænge forklarer resultaterne i DNT for nogle fag (fysisk/kemi) mindre end 25% af variationen i afgangsprøvens resultater. I den forbindelse fremhæver evalueringsrapporten, at en relativ lav forklaringsgrad er almindelig for lignende test i andre lande (eksempelvis Sverige); at der findes en række muligheder for målefejl; at der måles på forskellige delaspekter inden for et fag, samt muligheden for, at elever »lærer at tage de nationale test«. Det andet kritikpunkt vedrører det pædagogiske og formidlingsmæssige aspekt af DNT. Kritikken går på, at fejlbehæftede testresultater formidles videre til elever, samt at testresultater er »pædagogisk ubrugelige«, fordi de er »uforståelige« og ikke giver oplysninger, der er relevante for læreren (Bundsgaard & Kreiner 2019). Det tredje kritikpunkt forholder sig til, at lærerne observerer lavere trivsel hos eleverne i forbindelse med DNT (Bundsgaard & Puck 2016). Det fjerde kritikpunkt omhandler en tvivl om, hvorvidt DNT kan løfte de fagligt svage elevers faglige niveau. I VIVE's evalueringsrapport fra 2020 beskrives det mere overordnet, at udfordrede elevgrupper overordnet set ikke har fået et fagligt løft i årene 2014-2018, samt at der samlet set ikke er nogen klar

positiv udvikling i elevernes læring og trivsel i løbet af reformårene (VIVE 2020b). Andersen & Nielsen (2019) konkluderer derimod mere specifikt og på baggrund af et kvasiekperiment, der sammenligner folkeskoleelever fra 2. til 6. klasse, som i skoleåret 2009/2010 henholdsvis gennemførte og ikke gennemførte DNT, at det faglige niveau er højere for gruppen, der har gennemført DNT – særligt i gruppen af fagligt svagere elever. Studiet er i faglige kredse blandt andet kritiseret for, at nogle grupper af folkeskoler/lærere af forskellige årsager delvist *selv* valgte ikke *at* gennemføre DNT (selektionsbias); at der ikke er sikkerhed for, at de to grupper har fået de samme spørgsmål stillet i DNT, og at de statistisk signifikante forskelle i realiteten er små og usikre (Ravn 2019). På baggrund af de beskrevne kritikpunkter konkluderer Bundsgaard & Kreiner (2019), at den obligatoriske brug af de nationale test trækker folkeskolen og fagene i en »forkert retning«.

På trods af – eller måske netop på grund af – den overordnede kritik af DNT og den begrænsede dokumentation for en sammenhæng mellem DNT og det faglige niveau, er det relevant at spørge mere direkte til, hvordan elever oplever DNT. Den eksisterende viden på området er, som det fremgår nedenfor, dog begrænset og delvist modstridende.

Eksisterende forskning

I det følgende afsnit redegør vi for den eksisterende danske og internationale forskning, der belyser anvendelsen af standardiserede test generelt samt DNT specifikt ud fra folkeskoleelevernes egne perspektiver og oplevelser.

Elevernes oplevelse af standardiserede test

International og dansk forskning viser generelt forskellige variationer om elevernes oplevelse af standardiserede test. Nordenbo m.fl. (2009) fremhæver i deres gennemgang af eksisterende forskning, at højt præsterende elever på baggrund af testresultater i standardiserede test får en øget motivation for læring, mens de fagligt svage elever mister motivation. Desuden fremhæver Nordenbo m.fl. (2009), at der er en stærk evidens for, at testresultater har en indvirkning på elevens fremtidige motivation og selvværd.

I den internationale forskning er sammenhængen mellem standardiserede test og elevernes trivsel ikke undersøgt specifikt. En litteraturgennemgang af Harlen & Deakin (2002) fremhæver dog en sammenhæng mellem standardiserede test og elevers selvværd. Særligt viser spørgeskemabaserede studier af Davids & Brember (1998;1999), at der ingen sammenhæng er mellem elevers selvværd og faglige præstation før introduktionen af nationale test i England (National Curriculum Test), men efter indførelsen havde lavt præsterende elever lavere selvværd end højtpræsterende elever. Ligeledes fremhæver Nordenbo m.fl. (2009), at tilsvarende konklusioner ses i studier af statsstyrede test i USA. Ferman (2004) finder, at tre ud af fire elever oplever følelsen af angst og ængstelse for testresultat – særlig hos de fagligt svage elever. Desuden viser

et studie af McDonald (2001), at elevers præstationsangst overordnet set har en negativ indflydelse på deres præstationer til test, og at præstationsangsten for 20 % af eleverne er udslagsgivende for, hvorvidt de bestod eller dumpede testen. Et casestudie af 11-årige elever i en folkeskole i London viser, at når elever udsættes for test gentagne gange, bliver de meget bevidste om, hvad de mener, at de selv kan opnå, hvilket medfører, at de får lavere forventninger til deres egne evner (Reay & Wiliam 1999). Samlet set viser international forskning, at elevers oplevelse af standardiserede test varierer efter faglighed, hvor lavt præsenterende elever i overvejende grad påvirkes i en mere negativ grad end de fagligt dygtige (Nordenbo m.fl. 2009; Scott 2007; Harlen & Deakin 2002; David & Brember 1998; 1999).

Elevernes selvvaluerede oplevelse af de nationale test

Elevernes generelle oplevelse af DNT er undersøgt ved hjælp af spørgeskemaer, interview og observationer i en række rapporter, som når frem til meget varierende resultater (VIVE 2020a; Wandall m.fl. 2018; Bundsgaard & Puck 2016; Rambøll 2013; Kousholt 2012). En webbaseret spørgeskemaundersøgelse med 1.110 elever fra 4. til 8. klasse fra 2018 konkluderer, at eleverne i overvejende grad har en positiv oplevelse af DNT (Wandall m.fl. 2018). Endvidere viser undersøgelsen, at knap en femtedel af alle eleverne beskriver, at de føler »ubehag« ved DNT, og at denne andel stiger til mere end 40% blandt elever på 8. og 9. klassetrin. På trods af rapportens overordnede konklusion er det derfor centralt, at den positive oplevelse således er stærkt formindsket blandt ældre elever. Omvendt viser undersøgelsen også, at jo yngre og mere ressourcetsvage eleverne er, desto bedre oplevelse har de af DNT. At eleverne overordnet har en positiv oplevelse med testene, konkluderes også i Rambølls rapport fra 2013, der er baseret på casestudier af elever fra 2. til 8. klasse (Rambøll 2013). VIVE's evaluering fra 2020 (2020a), som blandt andet trækker på data fra observationer og interviews, viser derimod, at eleverne oplever test-situationen meget forskelligt, men også at hovedparten af elevernes oplevelse af DNT imidlertid hverken er særligt negativ eller særligt positiv. Det gælder eksempelvis en spørgeskemaundersøgelse blandt folkeskolelærere af Bundsgaard & Puck (2016), der viser, at lærerne vurderer, at elevernes oplevelse er overvejende negativ. Et studie fra 2012 af Kousholt (2012) baseret på elevinterviews og observationer af elever fra 6. klasse og 8. klasse på to folkeskoler konkluderer ligeledes, at elevernes oplevelse af DNT overvejende er negativ.

Samlet set finder vi, at der er betydelige variationer i resultaterne fra danske studier, samtidig med at den eksisterende viden er relativt sparsom. Hertil kommer, at der mangler mere forskningsmæssig viden, der afdækker elevernes egne oplevelser af DNT. Vi finder kun en enkelt større kvantitativ undersøgelse (Wandall m.fl. 2018), der afdækker sådanne oplevelser, hvorimod de øvrige danske studier er baseret på kvalitative elev- og lærerudsagn samt observationer (VIVE 2020a; Bundsgaard & Puck 2016; Rambøll 2013; Kousholt 2012).

Elevernes oplevelse af de nationale test og faglighed

Der er ligeledes kun en begrænset og delvis modstridende viden om sammenhænge mellem faglighed og oplevelsen af DNT. Wandall m.fl. (2018) finder, at jo højere elevernes karaktergennemsnit er til afgangsprøverne i folkeskolen desto mere positiv er oplevelsen af DNT. Omvendt finder Bundsgaard & Puck (2016), at lærerne oplever, at både fagligt dygtige og fagligt svage elever har en dårlig oplevelse af DNT. Dog finder de samtidigt, at især de fagligt svage elever »taber modet«, når de skal gennemføre DNT.

Elevernes oplevelse af de nationale test og (mis)trivsel

Ministeriet for Børn, Undervisning og Ligestilling gennemførte i 2015 en undersøgelse af sammenhængen mellem skolernes trivselsmålinger og testresultater i DNT. Undersøgelsen fandt en positiv sammenhæng mellem elevernes testresultater i 4. til 9. klasse og deres trivsel og konkluderede, at en øget trivsel styrker elevernes læring (Børn-, Undervisning- og Ligestillingsministeriet 2016). Studiet belyser dog ikke en direkte kausal sammenhæng mellem trivsel og oplevelse af DNT, men indikerer en (positiv) korrelation/samvariation mellem trivsel og testresultater.

Vi har ikke fundet yderligere studier, som direkte undersøger trivsel og DNT, men der findes undersøgelser af DNT og elevernes selvforståelse, som kan give en indikation på trivsel. En række danske studier baseret på elevinterviews og observationsdata finder, at DNT påvirker elevernes selvopfattelse, hvilket kan føre til, at nogle elever får lavere forventninger til deres egne potentialer og muligheder for fremtiden (Andreasen 2015; Andreasen & Rasmussen 2014; Andersen 2012). Andreasen (2019) beskriver, at »nervøsitet« kan ses som et udtryk for, at der er noget »på spil«, og at det derfor tillægges en stor betydning, hvor godt man præsterer. Dét, at man konfronteres med sine præstationer, kan være en meget stressende og angstfyldt proces for elever, hvis de gentagne gange oplever at få lave karakterer. Det kan medvirke til en selvopfattelse præget af lave forventninger til egne evner og potentialer, hvilket kan give anledning til yderligere nervøsitet (Andreasen 2019).

Vores litteraturgennemgang viser overordnet set, at der mangler empirisk viden om sammenhænge mellem DNT og elevers (mis)trivsel i en dansk sammenhæng. Der er ikke foretaget studier internationalt om mistrivsel, men forskning indikerer, at fagligt svage elevers selvværd i overvejende grad påvirkes i en mere negativ grad af de standardiserede test end hos de fagligt dygtige (Nordenbo m.fl. 2009; Scott 2007; David & Brember, 1998;1999).

De nationale test og kønsforskelle

Der findes ikke danske undersøgelser af kønsforskelle i elevernes oplevelser af DNT. Eksempelvis undersøger evalueringsrapporter af DNT fra VIVE ikke kønsforskelle eller refererer til andre studier, der gør dette (VIVE 2020a; VIVE 2020b). Kønsforskelle er dog vigtigt at inddrage i et børne- og ungdomsper-

spektiv. Eksempelvis ved vi, at piger har et markant dårligere mentalt helbred end drenge (Sundhedsstyrelsen 2018; SIF 2017). Kønsforskelle er ligeledes markante og vigtige, når det gælder tiltro til egne evner, læring og motivation. I en rapport fra VIVE (2019) omhandlende PISA-undersøgelserne dokumenteres det, at danske drenge overordnet set har større tiltro til egne evner end danske piger og er mærkbart mindre bange for at begå fejl. Endvidere findes der betydelige kønsforskelle i motivationen for læring. Piger har et større indre ønske om at forbedre egne færdigheder og lære, mens drenge motiveres af autonomi og status (Egelund, Nordahl, & Andersen 2018). Den internationale litteratur fremhæver ligeledes, at drenge og piger har forskellige udgangspunkter for læring afhængig af, hvilken tilgang der anvendes i klasselokalet (Harlen & Deakin 2002). Et studie af Johnson & McClune (2000) finder, at drenge er mere tilbøjelige end piger til at foretrække problemløsende og praktiske opgaver, hvorimod piger er mere tilbøjelige til at foretrække en mere systematisk og lærerstyret undervisning. Ifølge studiet har piger et højere selvværd, når den dominerende undervisningsstrategi er baseret på test, og en mere sekventiel undervisningstilgang (Johnson & McClune 2000). Andre studier viser samtidigt, at piger i højere grad end drenge giver udtryk for at have præstationsangst i forbindelse med test (Benmansour 1999; Evans & Engelberg 1988; Reay & Wiliam 1999; McDonald 2001), og at piger oftere vender deres succeser eller fejl indad, hvilket påvirker deres selvværd (Evans & Engelberg 1988). Det vil sige, at pigerne tilskriver test en større betydning – en tilskrivning, der kan forventes at påvirke deres selvværd, alt efter hvor godt de klarer sig. Den internationale forskning indikerer således, at piger i overvejende grad påvirkes i en mere negativ retning end drenge af standardiserede test og derfor kan forventes at have en mere negativ oplevelse af DNT.

Metode og empiri

Artiklen bygger på et »mixed method«-forskningsdesign (Doyle, Brady & Byrne 2019), hvor der i analysen trækkes på kvalitative data indsamlet via fokusgruppintervjuer og kvantitative data indsamlet via en spørgeskemaundersøgelse. I artiklen bidrager den kvalitative analyse til at illustrere og uddybe resultaterne i den kvantitative analyse. Den kvantitative del af analysen bygger på beskrivende statistik, signifikanstest af sammenhænge/forskelle samt vurdering af effektstørrelser. Den kvalitative del af analysen bygger på meningskondensering af indholdet i de gennemførte fokusgruppintervjuer og udvælgelse af udsagn fra informanterne, der illustrerer og uddyber resultaterne i den kvantitative del.

Fokusgruppintervjuer

Fokusgruppintervjuene blev gennemført med henblik på at kvalificere temaer og formuleringen af spørgsmål i spørgeskemaundersøgelsen samt understøtte den efterfølgende belysning af resultater. Interviewene blev gen-

nemført med strategisk udvalgte elever fra 8. og 9. klasse. Udvælgelsen af elever sket i samarbejde med skolerne og havde fokus på, at der indgik en bredere sammensætning af ressourcestærke og -svage elever. Formålet var at sikre, at der ikke var en overrepræsentation af en bestemt type (typisk ressourcestærke) elever. De deltagende elever blev lovet anonymitet. Interviewene havde en løs struktur for at skabe et rum for åbne og brede diskussioner med mindst mulig indblanding fra intervieweren. Temaerne og diskussionerne i interviewene var organiseret omkring: 1) elevernes forståelse af trivsel og mental sundhed, 2) elevernes oplevelse af at blive testet mere generelt, 3) elevernes oplevelse af, om DNT styrker deres faglighed, samt 4) elevernes oplevelse af DNT. Databehandlingen bestod af en åben kodningsproces, hvor formålet var at forholde sig åbent over for de temaer i empirien, som ikke i forvejen var belyst i den eksisterende litteratur.

Spørgeskemaundersøgelsen

Spørgeskemaundersøgelsen bygger på svar indhentet i foråret 2019 blandt folkeskoleelever fra 19 strategisk udvalgte folkeskoler ud af i alt 48 i Aalborg Kommune. Spørgeskemaet indeholdt i alt 36 spørgsmål opdelt i seks undertemaer, der afdækkede forhold vedrørende 1) baggrundsoplysninger, 2) faglighed, 3) trivsel, 4) selvværd, 5) selvvurderet helbred og 6) de nationale test. Spørgsmålene, der indgår i spørgeskemaet, blev udformet eller kvalificeret på baggrund af fokusgruppeinterviewet. Ligeledes har to folkeskolelærere kvalificeret spørgeskemaet for at sikre, at spørgsmålene var letforståelige for eleverne.

De 19 folkeskoler var udvalgt således, at de repræsenterede variationer i forhold til geografi, urbanitet og socioøkonomiske forhold indenfor Aalborg Kommune. Den strategiske udvælgelse blev valgt, da det i forbindelse med projektet ikke umiddelbart var muligt at indhente et udsnit baseret på en tilfældig stikprøve. Den strategiske udvælgelse gav desuden mulighed for at samarbejde med de enkelte folkeskoler i forhold til dataindsamlingen og herigennem understøtte en højere svarprocent.

Spørgeskemaundersøgelsen er gennemført som et webbaseret spørgeskema i anonymiseret form og distribueret til alle elever i 8. og 9. klasse i samarbejde mellem den centrale skoleforvaltning og skoleledere og faglærere på de strategisk udvalgte skoler. Spørgeskemaet blev distribueret til i alt 1.472 elever, hvoraf 888 gennemførte spørgeskemaet. Spørgeskemaundersøgelsens svarprocent er således på 60 %. En liste over de enkelte folkeskoler, deres karakteristika og svarprocenter er gengivet i appendiks 1.

Analysestrategi

I vores analyse fokuseres der på centrale variable og deres indbyrdes statistiske sammenhæng. De udvalgte variable bygger på spørgsmål om elevernes generelle oplevelse og følelsesmæssige oplevelse af de nationale tests, deres køn og deres selvvurderede faglige niveau. Herudover indgår en række

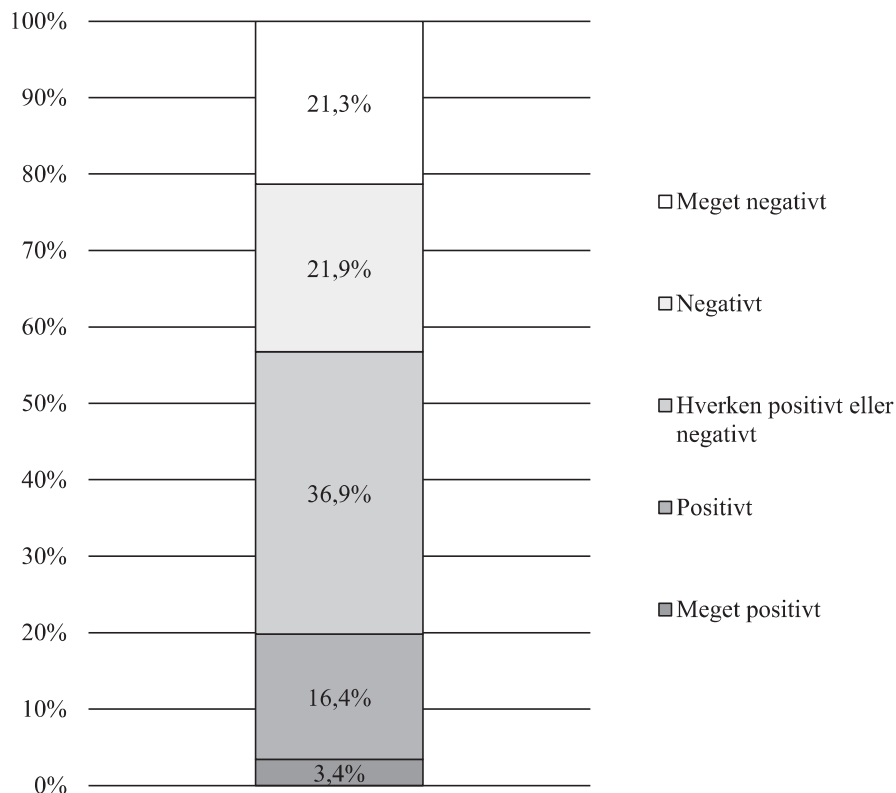
spørgsmål, der anvendes som indikatorer på graden af elevernes (mis)trivsel. Appendiks 2 indeholder en oversigt over de i analysen anvendte spørgsmålsformuleringer og svarkategorier samt fordeling af svar i de enkelte kategorier. Operationaliseringen af vores tre variable vedrørende opfattelse og oplevelse af DNT tager udgangspunkt i de dominerende beskrivelser, der fremkom blandt de deltagende elever i vores fokusgruppeinterviews. Den efterfølgende afprøvning af spørgeskemaet viste heller ikke forståelsesproblemer med spørgsmålene hos elever (eller lærere). De tre variable fremstår derfor umiddelbart med en høj grad af »face validity«. Vi har endvidere valgt at præsentere de tre variable individuelt, da gengivelse af svarfordelingerne i procent for hver variabel giver en umiddelbar og let fortolkning af centrale resultater. Anvendelse af en alternativ reflektiv indeksvariabel konstrueret ud fra flere variable (indikatorer) under en betragtning om, at de er tilnærmelsesvist intervalskalerede, ville teoretisk set bedre kunne indfange en underliggende latent variabel. Værdier for variabelen ville dog være vanskeligere at fortolke umiddelbart i forhold til vores første forskningsspørgsmål. χ^2 -tests (ikke vist) viser endvidere, at der er statistisk signifikante forskelle i svarfordelingerne mellem de tre variable. De anvendte variable for henholdsvis opfattelse og oplevelse af de nationale test, deres køn og faglige niveau indgår i analysen som kategoriske variable. Mistrivsel er konstrueret som en indeksvariabel, der indgår som en intervalskaleret variabel i analysen. I konstruktionen af variabelen for mistrivsel indgår svar på fem spørgsmål, som tidligere er anvendt i danske spørgeskemaundersøgelser, herunder »Skolebørnsundersøgelsen« fra 2018 (Rasmussen, Kierkegaard & Rosenwein m.fl. 2019) og »Stress blandt unge« fra 2007 (Nielsen, Vinther-Larsen, Nielsen & Grønbæk 2007). Fordelelen ved at anvende et indeks er, at de fem spørgsmål tilsammen – teoretisk set – bedre indfanger den underliggende (latente) variable for (mis)trivsel, end spørgsmålene ville gøre enkeltvis. Svarkategorierne for de fem spørgsmål er delvist forskellige, og på baggrund af en rangordning og tilskrivning af værdier for de enkelte svar (eksempelvis fra 1 for »aldrig/næsten aldrig« til 5 for »dagligt«) er der foretaget en transformation af svar til standardiserede værdier (z-scores). Det antages således, at de oprindelige ordinaltskalerede data tilnærmelsesvist kan betragtes som intervalskalerede. En tillægsanalyse af sammenhænge mellem den generelle holdning til de nationale test og de enkelte spørgsmål i indekset finder tilsvarende sammenhænge som i hovedanalysen. Eksempelvis viser en χ^2 -test, at der er en statistisk signifikant sammenhæng mellem en mere negativ oplevelse af de nationale test og en højere frekvens i oplevelsen af at være irriteret ($\chi^2 = 43,98$ (df = 16); $p < 0,01$ og Cramers V = 0,12). Beregning af Cronbachs alpha (=0,77) indikerer, at indekset for mistrivsel er kendetegnet ved en tilfredsstillende intern reliabilitet. Det endelige indeks er transformeret til standardiserede værdier for at muliggøre en umiddelbar tolkning af effektstørrelser, såsom forskelle i gennemsnit mellem grupper.

En mulig skævhed i de anvendte spørgsmålsformuleringer i vores analyse er, at de overvejende er negativt ladede og derfor bidrager til svar, som bliver mere negative, end hvis mere positivt ladede formuleringer var anvendt. En alternativ analyse (se appendiks A3) af sammenhænge mellem holdning til og oplevelse af de nationale test og andre variable baseret på mere positivt ladede spørgsmålsformuleringer om trivsel mere generelt og i forhold til skolen (eksempelvis: »Er du glad for at gå i skole?«) underbygger dog de samme konklusioner, som vi præsenterer i vores hovedanalyse.

Resultater

I dette afsnit gengiver vi hovedresultaterne fra vores kvantitative og kvalitative analyser af folkeskoleelevers generelle oplevelse og følelsesmæssige oplevelse af DNT og sammenhængene med faglighed, (mis)trivsel og køn.

Figur 1 a. Elevers generelle oplevelse af DNT



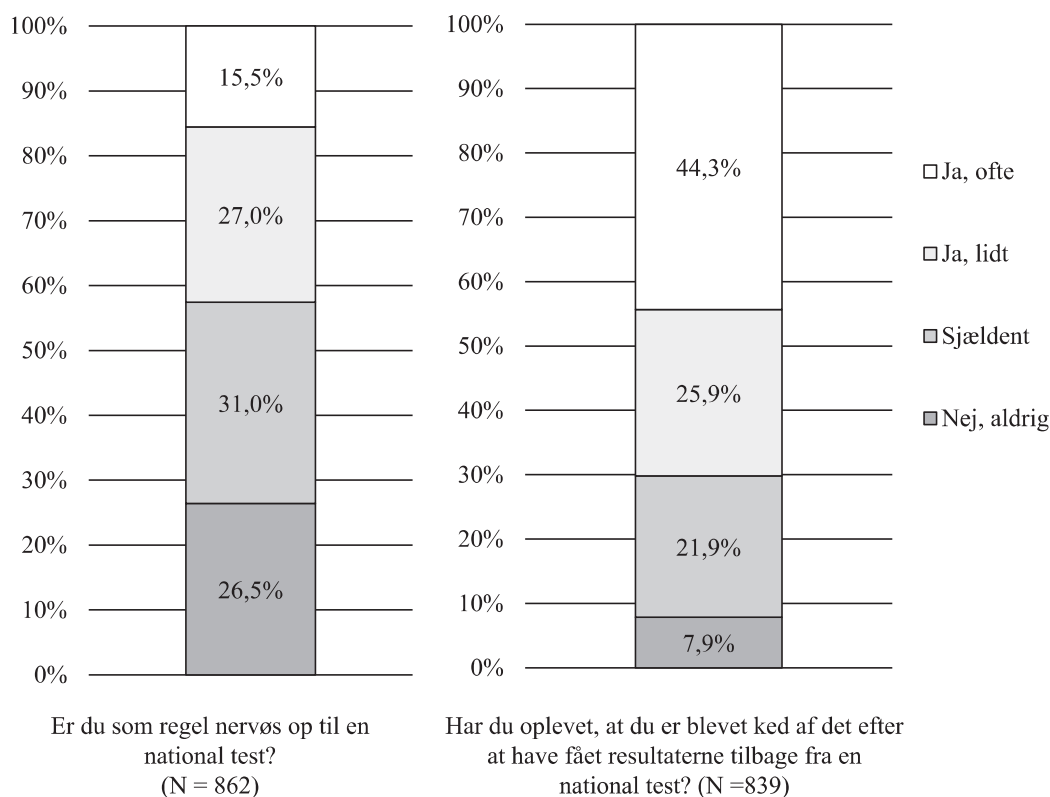
Hvordan forholder du dig generelt til de nationale test?

N = 858

I figur 1a gengiver vi fordelingen af svar på spørgsmålet om, hvordan folkeskoleeleverne forholder sig generelt til DNT. Et markant resultat er, at omkring dobbelt så mange elever forholder sig til »negativt« eller »meget negativt« til DNT (43,2 %) i forhold til folkeskoleelever, der forholder sig »positivt«

eller »meget positivt« (19,7 %). Samtidig er der en stor gruppe (37 %) der forholder sig »hverken positivt eller negativt« til DNT.

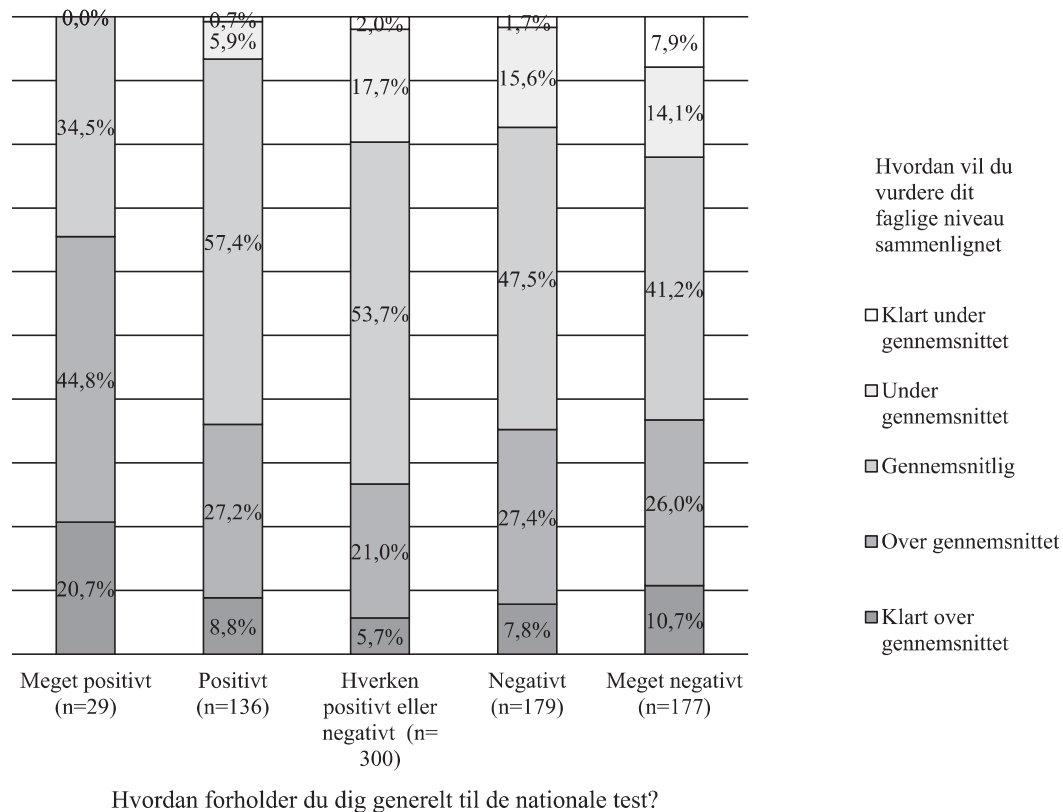
Figur 1 b. Følelsesmæssig oplevelse i forbindelse med DNT



I figur 1b gengiver vi fordelingen af svar på spørgsmål om, hvorvidt folkeskoleeleverne oplever at være henholdsvis »som regel nervøs« op til en national test og »ked af det« efter at have fået testresultater tilbage. Et markant resultat er, at meget store grupper har svaret »ja, ofte« (44,3 %) eller »ja, lidt« (25,9 %) til spørgsmålet om, hvorvidt de er »kede af det« efter at have fået testresultater tilbage. Grupperne, der svarer »ja, ofte« og »ja, lidt« til spørgsmålet om at være nervøse op til en test, er samlet set relativt mindre, men udgør stadig markante andele. Samlet indikerer resultaterne, at den følelsesmæssige oplevelse af DNT overvejende er negativ umiddelbart efter en test, samt at denne oplevelse er relativt almindelig blandt folkeskoleelever.

I figur 2 gengiver vi fordelingerne for selv vurderet faglige niveau opdelt i elever, der forholder sig mere eller mindre positivt (negativt) til DNT. En statistisk test (χ^2) viser, at der overordnet set er en størrelsesmæssig mindre (Cramers V = 0,13) men statistisk signifikant ($p < 0,01$) forskel i fordelingerne. Et markant resultat er, at der findes væsentlige forskelle mellem fordelingerne af den selv vurderede faglighed på tværs af grupperne, der forholder sig

Figur 2. Generel oplevelse af DNT og selv vurderet fagligt niveau



N = 821. χ^2 (Pearsons) = 56,73 (df = 16); $p < 0,01$. Cramers V = 0,13.

mere eller mindre positivt (negativt) til DNT. Tendensen på tværs af forskellene indikerer samtidigt, at der er en sammenhæng mellem en mere negativ oplevelse af DNT og et lavere (selvvurderet) fagligt niveau. En visuel inspektion af fordelingerne for selv vurderet faglighed i de enkelte grupper (oplevelse af DNT) indikerer samtidigt en række interessante forskelle. Umiddelbart synes elever i gruppen, der forholder sig meget positivt til DNT, at vurdere deres faglige niveau højere end i de øvrige grupper. Denne gruppe udgør dog samtidigt en meget lille del af alle elever (3,4 %, jf. figur 1). Det er derfor mere interessant, at gruppen, der forholder sig positivt til DNT (16,4 % af alle elever), har en markant mindre andel af elever, der eksempelvis vurderer deres faglige niveau »gennemsnitligt« (57,4 %) samt »under« (5,9 %) eller »klart under« (0,7 %) gennemsnittet i forhold til de grupper, der forholder sig relativt mindre positivt til DNT. Denne forskel er mest udtalt i forhold til gruppen, der forholder sig »meget negativt« til DNT (21,3 % af alle elever, jf. figur 1), hvor henholdsvis 41,2 %, 14,1 % og 7,9 % vurderer deres faglige niveau som »gennemsnitligt« samt »under« eller »klart under« gennemsnittet.

Tabel 1. Generel oplevelse af DNT og selvvrurderet mistrivsel

Opfattelse af NAT	Selvvrurderet mistrivsel ^a	95 % konfidensinterval		N	Signifikantest af gruppeforskelle (p-værdier) ^b				
		Nedre	Øvre		Meget positiv	Positivt	Hverken /eller	Negativt	Meget negativt
Meget positivt	-0,41 (0,73)	-0,72	-0,10	24	-				
Positivt	-0,36 (0,89)	-0,53	-0,19	110	1,00	-			
Hverken positivt eller negativt	-0,03 (0,98)	-0,16	0,11	214	0,16	0,02	-		
Negativt	0,23 (0,96)	0,07	0,38	147	0,00	0,00	0,11	-	
Meget negativt	0,13 (1,11)	-0,05	0,31	149	0,03	0,00	0,65	0,93	-

N = 644. Welchs t-test = 8,72 (4, 147), $p < 0,001$.

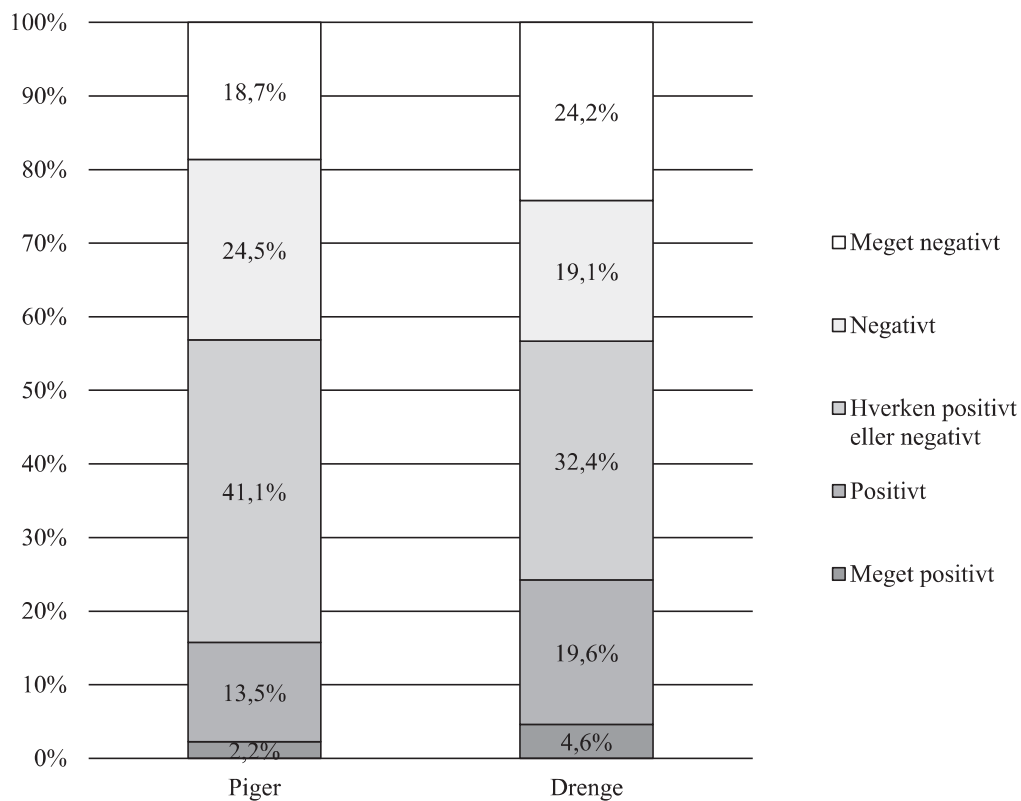
a Z-scores: gennemsnit og standardafvigelse i parentes

b Signifikantest baseret på Games-Howell (Welchs) post-hoc tests.

I tabel 1 gengiver vi hovedresultaterne af vores analyse af sammenhængen mellem den generelle oplevelse af DNT og selvvrurderet mistrivsel. Et markant resultat er, at der er væsentlige forskelle i graden af mistrivsel på tværs af grupperne med elever, der forholder sig mere eller mindre positivt (eller negativt) til DNT. Eksempelvis viser vores analyse, at den relativt større selvvrurderede mistrivsel i gruppen, der forholder sig »negativt« til DNT på 0,26 standardafvigelser i forhold til gennemsnittet for alle, er statistisk signifikant (vrurderet i forhold til et 95 % konfidensinterval). Forskellene i gennemsnit for mistrivsel blandt de grupper, der forholder sig »positivt« eller meget »positivt« til DNT, er samtidig statistisk signifikante ($p < .0,05$). Samlet set finder vi en markant sammenhæng mellem en større grad af mistrivsel og en mere negativ oplevelse af DNT.

I figur 3 gengiver vi, hvordan henholdsvis piger og drenge oplever DNT. Et væsentligt resultat er, at gruppen af piger generelt set synes at forholde sig lidt mindre positivt til DNT end gruppen af drenge, og at drengene fremstår lidt mere ekstreme i, hvordan de forholder sig til DNT. Eksempelvis er grupperne af piger, der forholder sig henholdsvis »positivt« (13,5 %) og »meget positivt« (2,2 %) mindre end de tilsvarende grupper for drenge (19,6 % og 4,6 %). Yderligere delanalyser (ikke vist) viser, at de overordnede sammenhænge mellem generel oplevelse af DNT og henholdsvis fagligt niveau og mistrivsel, som fremkommer i de ovenstående analyser, også genfindes, når analysen opdeles på køn (Eriksen, Lindholst & Knudsen, 2020). For både drenge og piger hænger en større mistrivsel og et lavere selvvrurderet fagligt niveau således sammen med en mere negativ oplevelse af DNT. Delanalyserne viser desuden, at mistrivslen blandt piger generelt set er markant større end blandt drenge. Eksempelvis er forskellen i mistrivsel størrelsesmæssigt på omkring en halv standardafvigelse (forskæl = 0,56, $p < 0,01$). Dette kunne indikere, at den psykologiske betydning af sammenhængene er relativt mere alvorlig for piger end drenge.

Figur 3. Generel oplevelse af DNT og køn



N = 858 (hvoraf piger udgør 445, og drenge udgør 413). $\chi^2 = 18,70$ (df = 4); $p < 0,001$. Cramers V = 0,15.

Diskussioner

Primære resultater

I denne undersøgelse har vi vist, at der er en stor andel af elever i 8. og 9. klasse, som oplever DNT negativt eller meget negativt (tilsammen 42 %). Denne tendens er forstærket i elevgrupper, der er fagligt svage (selvvurderede) og som mistrives. Desuden viser vores undersøgelse, at piger i lidt højere grad end drenge har en negativ oplevelse af DNT. Dette fund ligger i forlængelse af to tidligere studier (Bundsgaard & Puck 2016; Kousholt 2012), hvorimod det strider imod tidligere undersøgelser, der konkluderer, at eleverne overordnet set har en positiv oplevelse (Wandall m.fl. 2018; Rambøll 2013) eller indifferente (VIVE 2020). I forhold til den eksisterende forskning bygger den indeværende undersøgelse på et bredt datagrundlag i form af 888 respondenter fra 8. og 9. klassetrin og underbygger således en konklusion, der tilsiger, at en væsentlig andel af de danske elever i de sene klassetrin har en negativ oplevelse af DNT.

Den kvalitative del af vores undersøgelse bidrager til at nuancere forståelsen af elevernes oplevelser af DNT. I et af vores fokusgruppeinterviews fremhæver en elev fra 8. klasse, at vedkommende oplever læringsværdien af DNT som meget lav, og kobler dette på en gengivelse af lærerens negative oplevelse til DNT som læringsredskab:

Jeg føler ikke, at jeg lærer noget af dem. Og så føler jeg ikke, at man kan bruge dem til noget, fordi, jo, man får at vide, hvor man ligger henne, men man lærer ikke noget af dem. Og jeg kan ikke se, hvad de skal gøre godt for, fordi vores lærer kan ikke rigtig bruge dem til noget.

En anden nuance i elevernes negative oplevelse af DNT kan hænge sammen med tekniske problemer og selve organiseringen af, hvordan test gennemføres i det praktiske. Dette afspejles i følgende citat:

Det er jo også det der med, at det så tit går galt med DNT. Der er altid en, som ikke kan komme ind, så er der dem, som skal have ekstra tid. Og så skal resten af klassen, som ikke skal have ekstra tid, de skal bare sidde og lave ingenting.

Et tredje aspekt af, at eleverne har en negativ oplevelse med DNT, kan skyldes et manglende læringspotentiale, fordi eleverne ikke får at vide, hvilke fejl de har lavet. Eleverne udtrykker ligeledes i fokusgruppeinterviewene, at de overordnet er tilfredse med at blive testet, men at der er noget særligt problematisk ved DNT:

Jeg tænker også sådan, at DNT, der kan man ikke se, hvilke fejl man har lavet, hvilket vil sige, at man ikke kan gå igennem de fejl, man så har lavet, og derfor kan man ikke forbedre sig på dem. Jeg ved, at de færdighedsprøver, vi har i matematik, der går vores lærer så igennem de fejl, vi så har haft, og fortæller os så, okay, det er sådan, den skal laves, og så viser os det, og så lærer man noget. Det gør man ikke af DNT.

Sammenhæng mellem elevernes oplevelse af DNT og selvvurderet faglighed

Vores undersøgelse indikerer, at elever med et svagere fagligt niveau også har en mere negativ oplevelse af DNT. Resultatet afviger fra resultater i undersøgelsen af Wandall m.fl. (2018), der viste en svag sammenhæng mellem lavere karaktergennemsnit hos eleverne i skolens afgangsprøver og en mere positiv holdning til DNT. Omvendt er vores resultater i overensstemmelse med en række nationale og internationale studier, som også finder, at fagligt svage

elever generelt har en dårligere oplevelse af DNT (Nordenbo 2009; Bundsgaard & Puck 2016; Scott 2007).

At de fagligt svage elever særlig har en negativ oplevelse af DNT, er et vigtigt fund, da ét af de politiske formål med testene netop er at løfte de fagligt svage elever.

Hvis det forholder sig sådan, at de fagligt svage elever er den gruppe, som udtrykker den største modvilje imod dem, er det relevant at spørge, hvad det har af betydning for deres hverdag og trivsel. Dette kom til udtryk i fokusgruppeinterviewene, hvor flere udsagn illustrerede og supplerede den generelle sammenhæng, som vi fandt i vores kvantitative analyse mellem elevens faglige niveau og oplevelse af DNT. Herudover illustrerede en del udsagn, at sammenhængen mellem et svagere fagligt niveau og en dårligere oplevelse af DNT hænger sammen med en påvirkning af elevens trivsel. En elev i 9. klasse, der vurderede sig selv som fagligt svag, udtrykte det eksempelvis således:

Jeg føler, at det er en selvtillidsknuser, hvis man ikke kommer over middel. Man får bare at vide, du er under målet, eller langt under middel. Og så tænker man lidt, hvad kan jeg bruge det til? Udover at det gør en lidt sur, ja, man kommer bare i dårligt humør.

En andet elevudsagn udtrykker det således:

Ligesom med karakter føler man, at man bliver smidt i en bås. Hvis man kommer op i noget, som bare ikke er ens stærkeste side, så er det også sådan, altså jeg er god til det her, men det, jeg kommer op i, kan jeg ikke finde ud af. Så bliver man virkelig sat i en bås – det er du bare dårlig til, og det er ikke så fedt.

Udsagnene supplerer den kvantitative analyse og illustrerer således en mulig forklaringsmodel på, hvordan og hvorfor DNT påvirker særligt fagligt svage elevers trivsel.

Sammenhæng mellem de nationale test og elevernes mistrivsel i folkeskolen

Vores undersøgelse viser en markant sammenhæng mellem større mistrivsel og en mere negativ oplevelse af DNT blandt de udspurgte 8. og 9. classes elever. Der er ikke udarbejdet samme indeksering af mistrivsel i andre studier, men en negativ sammenhæng mellem DNT og elevernes selvforståelse og selvværd er fremhævet i tidligere danske kvalitative undersøgelser (Andreasen 2019; Andreasen 2014; Andreasen 2015; Andersen 2012; Andreasen & Rasmussen 2014). Samme tendens ses også i den internationale litteratur (Reavy & Wiliam 1999; Harlen & Deakin 2002). Desuden påpeges en sammenhæng mellem en lavere faglig præsentation i standardiserede test og et dårlig-

gere selvværd (Nordenbo m.fl. 2009; Harlen & Deakin 2002; David & Brember 1998; 1999). Følgende udsagn af en elev i 8. klasse giver et indblik i en mulig årsag til denne sammenhæng:

Jeg ved, at jeg altid går ind og yder mit 100% bedste, og hvis det så ikke er godt nok for den her computer, som sidder og laver den her test, så er det ligesom at få en hammer oveni hovedet. Det bedste, du har ydet, det var ikke godt nok. Og der synes jeg godt, at det kan være en fejl, at man får at vide, ja, du er under middel, men hvis man så har gjort sit 100% bedste, så føler man bare, nå, så er det jo bare forgæves.

Ovenstående elevudsagn giver anledning til refleksion, hvor vi anskuer det at være »under middel« som et udtryk for en sårbarhed over ikke »at være god nok« og »leve op til kravene«. Det var ligeledes den kontekst, eleven svarede ud fra i fokusgruppeinterviewet. I et andet fokusgruppeinterview fortalte en pige i 9. klasse mere uddybende, hvordan »nervøsitet« spiller ind i forbindelse med at tage testen, hvilket illustreres i dette udsagn:

Jeg tror, jeg tænker meget over det, når jeg tager testen, eksempelvis om det her spørgsmål virker nemmere end det forrige, fordi så tænker man, har jeg lige svaret forkert på en hel masse spørgsmål. Og på den måde bliver jeg usikker på det endegyldige svar/resultat. Det er tiden mellem resultatet og testen, at jeg er usikker og nervøs.

Hvilke kønsforskelle ses der ved brug af standardiserede test i folkeskolen?

I vores litteraturgennemgang fandt vi, at kønsforskelle er potentielt vigtige i forhold til at forstå elevers oplevelse af DNT. Vigtigheden af at inddrage kønsforskelle ses også i vores undersøgelse, hvor mistrivlsen er markant større blandt piger end drenge. Mere centralt for vores forskningsspørgsmål finder vi også, at piger har en lidt mere negativ oplevelse af DNT. Det er i overensstemmelse med den internationale litteratur, der påpeger, at piger i forbindelse med standardiserede test har en større tilbøjelighed til at vende fejl og nederlag indad på en måde, som kan påvirke pigernes selvværd (Harlen & Deakin 2002; McDonald 2001; Johnson & McClune 2000; Evans & Engelberg 1988). Det er således nærliggende at overveje, om DNT bidrager til og forstærker pigernes mistrivsel. Resultaterne i vores undersøgelse underbygger vigtigheden af et kønsperspektiv på DNT.

Generaliserbarhed

Denne undersøgelse er foretaget på et antal skoleelever i 8. og 9. klasse i Aalborg Kommune. Generaliserbarheden skal således ses i den kontekst, og re-

sultaterne kan således ikke direkte overføres til hele den danske folkeskole og bør betragtes som overvejende eksplorativ i forhold til vores forståelse af, hvordan folkeskoleelever i Danmark oplever de nationale test. Den generelle relevans af undersøgelsen for særligt elever på 8. og 9. klassetrin underbygges af flere argumenter. For det første bygger vores undersøgelse på et stort kvantitativt datamateriale. I forbindelse hermed er udvælgelsen af skoler sket gennem en stringent strategisk udvælgelse baseret på geografi og socioøkonomisk social profil for at sikre et repræsentativt elevudsnit af populationen. For det andet viser eksisterende undersøgelser af den selvvaluerede trivsel, afgangsprøver og resultater i DNT, at elever i Aalborg Kommune ligger tæt på landsgennemsnittet (Aalborg Kommune 2020).) Elevernes selvvaluerede trivsel i Aalborg Kommune ligger dog lidt over det nationale gennemsnit. Da vi samtidigt finder, at mistrivsel er signifikant associeret med en mere negativ oplevelse af DNT, rejser denne forskel en spekulation om, hvorvidt oplevelsen af DNT er relativt mere negativ på landsplan end i Aalborg Kommune. Samlet set forventes det dog, at resultaterne tilnærmelsesvist repræsenterer det generelle billede blandt elever på 8. og 9. klassetrin i folkeskolen.

Konklusion

Vi har i de seneste årtier set, at folkeskolen har udviklet sig fra et grundtvigiansk værdisæt til en folkeskole med et større fokus på individet, faglige præstationer og test. Indførelsen af DNT har været en del af denne bevægelse. Hensigtsmæssigheden af denne type af test er dog omdiskuteret, og den nuværende viden om elevernes egen oplevelse af DNT har været sparsom og har peget i forskellige retninger. Denne undersøgelse viser på et bredt datamæssigt grundlag indhentet blandt folkeskoleelever i 8. og 9. klasse i Aalborg Kommune, at store elevgrupper har en negativ oplevelse af DNT, og at denne negative oplevelse hænger sammen med mistrivsel, svagere faglighed og er lidt mere udbredt blandt piger end drenge. Disse resultater bliver ekstra relevante, i perspektivet af at tidligere studier ikke på overbevisende måde har kunnet finde en association mellem DNT og læringsudbyttet generelt og i særlig grad hos fagligt svage elever. Tilbage står spørgsmålet om, hvorvidt behovet for central styring af folkeskolen kan retfærdiggøre en anvendelse af de nationale test – et behov, der også må afvejes i forhold til validitetsproblemer og den relativt begrænsede sammenhæng, der forefindes mellem resultater i de nationale test og øvrige prøver i folkeskolen.

Som denne undersøgelse illustrerer, kan inddragelsen af elevernes perspektiv på DNT udgøre et væsentligt supplerende bidrag til den eksisterende forståelse af effekten af DNT, og elevernes perspektiver og behov bør indgå i dialogen om fordele og ulemper ved et nationalt bedømmelses- og evalueringssystem. Vores resultater påpeger dog et behov for yderligere omfattende undersøgelser af, hvordan danske folkeskoleelever i flere klassetrin og i

endnu større skala oplever de nationale test i forhold til deres egen vurdering af trivsel og læring.

Taksigelse

Forfatterne vil gerne takke følgende: Skoleforvaltningen og folkeskolelærere i Aalborg Kommune for at bistå med indsamling af data, elever i 8. og 9. klasse i Aalborg, som tog sig tid til at deltage i fokusgruppeinterviews og svare på vores spørgeskema, samt kollegaer ved Center for Organisation, Management og Administration, Aalborg Universitet for kommentering på artikeludkast.

Litteratur

- Aalborg Kommune (2020): Kvalitetsrapport 2020. Fundet den 19. februar 2021 på hjemmesiden:
www.aalborg.dk/usercontrols/AalborgKommune/Referater/Pdf.aspx?pdfnavn=21175893-22352311-23.pdf&type=bilag&pdfid=114756
- Andersen, D. (2012): *Skoleerfaringer – fra skolestart til klassetrin som 15-årig*. I Ottesen, M.H (red): *15-åriges hverdagsliv og udfordringer*. SFI – Det nationale forskningscenter for velfærd. København.
- Andersen, S.C. & Nielsen, H.S. (2019): Learning from Performance Information. *Journal of Public Administration Research and Theory*, 2019, 1-17.
- Andreasen, K.E., & Rasmussen, A. (2014). »I was Just an Average Girl«: Student Identities in the Context of Testing. I: Rasmussen, A., Gustafsson, J., & Jeffrey, B. (red.), *Performativity in Education: An international collection of ethnographic research on learners' experiences* (s. 131-152). United Kingdom, Essex: Ethnography and Education.
- Andreasen, K.E. (2015): *De Nationale Test – nogle overordnede perspektiver på deres betydninger for pædagogikken og eleverne*. CEPRA-Striben, (18), 14-21.
- Andreasen, K.E. (2019): *Nationale Test set fra børneperspektiv*. Viden om Literacy nummer 25, april 2019. Nationalt Videncenter for Læsning.
- Benmansour, N. (1999): *Motivational orientations, self-efficacy, anxiety and strategy use in learning high school mathematics in Morocco*. Mediterranean Journal of Educational Studies, 4, 1-15.
- Beuchert, L.V. & Nandrup, A.B. (2017): *The Danish National Tests at a Glance*. Aarhus Universitet: Institut for Økonomi, TrygFondens Børneforskningscenter.
- Beuchert, L.V. & Nandrup, A.B. (2014): *The Danish National Tests: A practical Guide*. Aarhus Universitet: Institut for Økonomi, TrygFondens Børneforskningscenter.
- Bundsgaard, J. & Puck, M.R. (2016): *Nationale test: Danske lærere og skolelederes brug, holdninger og viden*. Aarhus Universitet. <https://doi.org/10.7146/aul.145.124>
- Bundsgaard, J. & Kreiner, S. (2019): *Undersøgelses af nationale tests måleegenskaber*. Aarhus Universitet, København. <https://doi.org/10.7146/aul.319.217>
- Børne- og Undervisningsministeriet (2020a): *Om Nationale Test*. Fundet den 01. marts 2020 på hjemmesiden: www.uvm.dk/folkeskolen/elevplaner-nationale-test--trivselsmaaling-og-sprogproever/nationale-test <https://doi.org/10.7146/aul.145.124>
- Børne- og Undervisningsministeriet (2020b): *Ny aftale skaber tryghed om kvaliteten af elevernes bedømmelsesresultater*. Fundet den 01. marts 2020 på hjemmesiden: www.uvm.dk/

- aktuelt/nyheder/uvvm/2020/feb/200221-ny-aftale-skaber-tryghed-om-kvaliteten-af-ele-
vernes-bedoemmelsesresultater <https://doi.org/10.6027/9789289334815-11-da>
- Børne- og Undervisningsudvalget (2020): *Rådgivningsgruppen for evaluering af de nationale test – anbefalinger*, bilag 82, januar 2020.
- Børn-, Undervisning- og Ligestillingsministeriet (2016): *Sammenhæng mellem elevernes trivsel og elevernes nationale testresultater*. Ministeriet for Børn, og Ligestilling. Styrelsen for It og læring, Notat. <https://doi.org/10.6027/anp2014-716>
- Dahler-Larsen, P. (2008): *Kvalitet i skolen*. I Bjørnholt, B. Andersen, J. G & Lolle, H. (red): *Bag kulissen i konstruktion af kvalitet*. Syddansk Universitetsforlag.
- Davies, J. & Brember, I. (1998): *National curriculum testing and self-esteem in year 2 the first five years: a cross-sectional study*. Educational Psychology, 18:365-375. <https://doi.org/10.1080/0144341980180401>
- Davies, J. & Brember, I. (1999): *Reading and mathematics attainments and self-esteem in years 2 and 6: an eight-year cross-sectional study*. Educational Studies, 25:145-157. <https://doi.org/10.1080/03055699997873>
- Doyle, L., Brady, A. & Byrne, G. (2019): *An overview of mixed methods research – revisited*. In *SAGE mixed methods research*. Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781526498137.
- Egelund, N. Nordahl, T. & Andersen, P.G. (2018): *Piger og drenge i skolen*. Aalborg Universitetsforlag. FULM: Forskningsinformeret udvikling af læringsmiljøer.
- Eriksen, T.B, Lindholst, A.C. & Knudsen, S.V. (2019): *Resultater fra en spørgeskemaundersøgelse om folkeskoleelevers opfattelse/oplevelse af de nationale test*. Aalborg Universitet. DOI: 10.5278/nbmh-w514.
- Evans, E. & Engelberg, R. (1988): *Pupils' perceptions of school grading*. Journal of Research and Development in Education, 21:44-54.
- Ferman, I., Watanabe, Y., & Curtis, A. (2004): *The Washback of an EFL National Oral Matriculation Test to Teaching and Learning*. In *Washback in language testing: Research contexts and methods* (pp. 191-210). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. <https://doi.org/10.1093/applin/ami015>
- Ravn, K. (2019). *Nationale test: Da »ikke statistisk signifikant« blev til »markant« fremgang (inklusive diskussion)*. Folkeskolen.dk: fagblad for undervisere. Den 4. november 2019: www.folkeskolen.dk/937080/nationale-test-da-ikke-statistisk-signifikant-blev-til-markant-fremgang. <https://doi.org/10.4045/tidsskr.18.0847>
- Harlen, W. & Deakin, C. (2002): *A systematic review of the impact of summative assessment and test on student's motivation for learning*. London: EPPI. Institute of Education, University of London.
- Johnston, J. & McClune, W. (2000): *Selection project sel 5.1: pupil motivation and attitudes – self-esteem, locus of control, learning disposition and the impact of selection on teaching and learning*, in: *The Effects of the Selective System of Secondary Education in Northern Ireland: Research Papers Volume II* (Bangor, Co. Down, Department of Education).
- Kousholt, K. (2012): *De danske nationale test og deres betydninger set fra børnenes perspektiver*. Pædagogisk psykologisk tidsskrift, Årg. 49, 4:273–290.
- McDonald, A. (2001): *The prevalence and effects of test anxiety in school children*, Educational Psychology, 21, pp. 89-101. <https://doi.org/10.1080/01443410020019867>
- Nielsen, L. Vinther-Larsen, M. Nielsen, N.R & Grønbæk, M. (2007): *Stress blandt unge*. Statens Institut for Folkesundhed.

- Nordenbo, S.E. m.fl. (2009): *Pædagogisk brug af test – et systematisk review*. København: Danmarks Pædagogiske Universitetsforlag & Dansk Clearinghouse for Uddannelsesforskning.
- Rambøll (2013): *Evaluering af De Nationale Test i folkeskolen*. Undervisningsministeriet (Kvalitets- og Tilsynsstyrelsen).
- Rasmussen, M. Kierkegaard, L. & Rosenwein, S.V. (2019): *Skolebørnsundersøgelsen 2018 – Helbred, trivsel og sundhedsadfærd 11-,13- og 15-årige skoleelever i Danmark*. Statens Institut for Folkesundhed, Syddansk Universitet.
- Reay, D. & Wiliam, D. (1999) »I'll be a nothing«: structure, agency and the construction of identity through assessment, *British Educational Research Journal*, 25:343–354. <https://doi.org/10.1080/0141192990250305>
- Robinson, K. (2011): *Out of our Minds – Learning to be Creative*. Capstone Publishing Ltd.
- Scott, C. (2007): *Stakeholder perceptions of test impact*. Policy & Practice. 14, 1:27-49
- Schou, L.R. (2009): *Evalueringskultur I Danmark*. Sprog & Integration, nr. 2.
- Schou, L.R. m.fl. (2015): *Når nationale test tager livtag med folkeskolen*. CEPRA-Stripen, (18):38-45.
- SIF (2006): *Sundhed og sygelighed i Danmark – & udviklingen siden 1987*. Statens Institut for Folkesundhed, Syddansk Universitet.
- SIF (2010): *Sundhed og sygelighed i Danmark – & udviklingen siden 1987*. Statens Institut for Folkesundhed, Syddansk Universitet.
- SIF (2017): *Mental sundhed blandt unge – resultater fra Ungdomsprofilen fra 2014*. Statens Institut for Folkesundhed, Syddansk Universitet.
- Sundhedsstyrelsen (2018): *Danskernes sundhed – den nationale sundhed*. Sundhedsstyrelsen.
- Torring, J. og Triantafillou, P. (red.) (2017): *New Public Governance på dansk*. Akademisk Forlag, København.
- Undervisningsministeriet (2005): *LBK nr. 880 af 19/09/2005: Bekendtgørelse af lov om gennemsigtighed og åbenhed i uddannelserne m.v.*
- Undervisningsministeriet (2018): *Notat: Køn og forældreuddannelse har stor indflydelse på elevernes nationale testresultat*. Undervisningsministeriet – Af Center for Data og Analyse.
- VIVE (2020a): *Anvendelsen af de nationale test – Delrapport 5: Evalueringen af de nationale test*. Viden til Velfærd. Det Nationale Forsknings- og Analysecenter for Velfærd. <https://doi.org/10.7146/aul.145.124>
- VIVE (2020b): *Elevernes læring, trivsel og oplevelser af undervisningen i folkeskolen – en evaluering af udviklingen i reformårene 2014-2018*. Viden til Velfærd. Det Nationale Forsknings- og Analysecenter for Velfærd. <https://doi.org/10.7146/lom.v5i9.6146>
- VIVE (2020c): *De nationale tests samvariation med karakterer – Delrapport 2: Evalueringen af de nationale test*. Viden til Velfærd. Det Nationale Forsknings- og Analysecenter for Velfærd.
- VIVE (2019): *Pisa 2018 – Danske unge I en international sammenhæng*. Viden til Velfærd. Det Nationale Forsknings- og Analysecenter for Velfærd.
- Wandall, J. m.fl. (2018): *Elevernes syn på de nationale test*. Nordic Metrics.

The pupils' experience of National Tests: a questionnaire survey from strategically selected schools in Aalborg Municipality

In this article, we complement earlier research and study how pupils in Danish public schools perceive national tests and how this perception correlates with skill, (lack of) well-being and gender. The study is based on a questionnaire and focus group interviews carried out in 2019 with pupils in public schools in Aalborg Municipality. The study shows that a majority of pupils has predominantly negative views of the national test and that a more negative perception is associated with negative well-being before and after the test. Also, we find significant associations between more negative views of the national test and lower skills, less well-being, and girls (compared to boys). The associations indicate that the perception of the national test is most negative for those groups of pupils where the need for learning support is greatest. Overall, our study points to needs for bringing in broader perspectives on the effects of national tests, question the value of national tests as a learning tool, and identify a need to study further, how national tests affect pupils – personally and socially and their learning.

Keywords: National tests, evaluation, public schools, gender, well-being, skills.