

FORSKNINGSRAPPORT

DrengAkademiet 2015

Årets resultater

Indholdsfortegnelse

Indledning	2
Metode	3
Hvem deltog?	3
Hvad og hvordan målte vi?	4
Hvordan kom vi frem til resultaterne?	5
Hvad betyder det? Effektstørrelse og læringsår	5
Hvad bør man ellers være opmærksom på? - andre metodiske forbehold	6
Resultater fra DrengAkademiet 2015	7
Hovedresultater summeret	7
Fremgang i de faglige test	8
Udvikling i faglige niveauer	9
Effektstørrelse og læringsår	11
Almen trivsel, fagligt orienterede kompetencer og handlekompetencer	12
Sammenhæng mellem udvikling i trivsel og faglig udvikling	13
Opsamling og perspektivering	15
Referencer	16

Indledning

I sommeren 2015 blev endnu en runde af DrengAkademiet vel gennemført. Indtrykket blandt deltagere såvel som personale fra undervisere til holdledere var udpræget positivt, og den generelle holdning var, at deltagerne havde gennemgået en væsentlig faglig og social udvikling. Det er imidlertid vigtigt for LøkkeFonden også at kunne dokumentere resultaterne med henblik på at sikre sig, at disse var reelle og ikke alene oplevede.

I den forbindelse blev der, som i 2014, tilknyttet en metodekonsulent med det formål at følge indsatsen i forhold til at dokumentere resultaterne fra akademiet samt at skabe gennemsigtighed om, hvordan resultaterne blev frembragt.

Nærværende bidrag har således til formål at præsentere og analysere resultaterne fra DrengAkademiet 2015 samt at redegøre for den anvendte metode, og de overvejelser, der er blevet gjort i den sammenhæng.

Denne rapport er udgivet af LøkkeFonden i samarbejde med Jesper de Hemmer Egeberg, metodekonsulent.

Metode

Det følgende afsnit har til formål at skabe klarhed og gennemsigtighed om, hvordan data og resultater fra DrengAkademiet 2015 er blevet indsamlet og analyseret.

Hvem deltog?

Rekrutteringen af deltagere til dette års akademi foregik, som tilfældet var i 2013 og 2014, via flere uafhængige kanaler. Der blev e-mailet informationsmateriale til skoler og ungdomsuddannelser rundt om i landet, ligesom der blev reklameret for årets akademi på LøkkeFondens hjemmeside, facebookside og via fondens nyhedsbrev. Dertil har den generelle medieeksponering af resultaterne fra tidligere års DrengAkademi været en kilde til rekruttering af nye deltagere. For at komme i betragtning til DrengAkademiet udfyldte drenge og forældre et spørgeskema om baggrund, fagligt niveau og motivation for deltagelse. Derefter blev ansøgerne sorteret efter tre udvælgelseskriterier:

- 1) At drengen var i aldersgruppen 14 til 16 år, og ideelt set gik i 8. klasse og var på vej i 9. klasse. 8. klasses drenge blev foretrukket, da de erfaringsmæssigt er mere fokuserede under indsatsen. Dog blev drenge, der gik i 7. klasse, også optaget.
- 2) Ansøgeren skulle have et fagligt efterslæb. De fleste af ansøgerne lå væsentligt under gennemsnittet rent fagligt.
- 3) Det tredje og sidste udvælgelseskriterium var, at ansøgeren skulle have et minimum af sociale og personlige kompetencer, som kvalificerer ham til at være og agere i større og mindre fællesskaber. Kompetencer, der er centrale for at kunne indgå i indsatsen. Endelig blev gengangere fra DrengAkademiet 2014 ikke tilbudt optagelse.

Dertil blev det så vidt muligt tilstræbt at sammensætte deltagergruppen således, at den var så demografisk repræsentativ som mulig.

Blandt de ansøgere, der opfyldte kriterierne, blev der udvalgt 107, som blev tilbudt en plads på akademiet, 102 takkede ja til en plads, 2 endte med at udeblive og 2 faldt fra under forløbet. Således gennemførte 98 drenge dette års akademi, heraf 5,1 % af anden etnisk herkomst end dansk. Aldersgennemsnittet blandt drengene var 14,8 år.

Tabel 1 giver et overblik over deltagerens geografiske fordeling og hvilket klassetrin, de stammer fra.

Tabel 1: Geografisk fordeling og klassetrin

	Sjælland og øer	Fyn og øer	Jylland og øer	7. klasse	8. klasse
Andel	45,9 %	16,3 %	37,8 %	30,6 %	69,4 %

N=98

Hvad og hvordan målte vi?

DrengAkademiets deltagere fik testet deres faglige kundskaber i stavning, læsning og matematik (afgrænset til matematiske færdigheder i tal og algebra¹) i alt tre gange i forløbet. Testene blev udført på introcampen (20.-21. juni) samt i slutningen af akademiets første og anden uge (29. juni-12. juli). Læsning blev testet ved brug af Gullhojs læseprøve for 8. klasse, stavning ved Gullhojs retskrivningsprøve for 7. klasse og matematik med et tilpasset prøvesæt fra Alinea². Stavetesten foregik på Chromebooks, hvorimod matematik og læsning blev testet i papirform. Kun elever, der deltog i samtlige test, indgår i analysen af de faglige resultater.

I forbindelse med dette års DrengAkademi var det et ønske at supplere de faglige test med målinger af deltagerens sociale udbytte, i form af et overordnet mål for faglig og social trivsel. Der blev i den forbindelse søgt metodisk inspiration i forhold til udvalg af testspørgsmål primært i SFIs pilotundersøgelse til udviklingen af de nationale trivselsmålinger i folkeskolen (Keilow et al., 2014) samt den dertil knyttede ekspertgruppes anbefalinger (Undervisningsministeriet, 2014a).

Det var i den forbindelse en væsentlig prioritet 1) at reducere antallet af spørgsmål mest muligt, idet deltagerne på DrengAkademiet allerede bliver testet meget på kort tid, og erfaringer fra DrengAkademiet 2014 indikerede, at mange blev testtrætte og opgav at gennemføre mere omfattende sociale test 2) I forlængelse af 1) at udvælge testspørgsmål fra de dimensioner af trivselsbegrebet, der blev fundet mest relevante i forhold til indsatsen³ og 3) at ordlyden af spørgsmålene blev tilpasset, så de passede til konteksten på DrengAkademiet.

Der blev derfor målt på deltagerens selvvalgte udvikling inden for *almen trivsel*, *generelle handlekompetencer* samt *fagligt orienterede kompetencer* to gange i forløbet. Det foregik på introcampen og i akademiets sidste uge. *Almen trivsel* skal her forstås, som almen velvære eller veltilpasset. *Generelle handlekompetencer* skal forstås som generel tiltro eller tillid til egne evner i forhold til at kunne klare det, man sætter sig for og løse sine problemer. *Fagligt orienterede kompetencer* anvendes i betydningen, at man oplever, man klarer sig godt fagligt, at man er god til at koncentrere sig, når man modtager undervisning, samt at man føler sig i stand til at lære dét, man gerne vil lære. Disse udvalgte trivselsmål blev fundet relevante at måle, da de stemmer godt overens med de dele af DrengAkademiets indsatsfelt, som *ikke* går direkte på det faglige og didaktiske, men som udgør et betydningsfuldt fundament for læring.

Almen trivsel, *generelle handlekompetencer* samt de *fagligt orienterede kompetencer* blev målt med 2-3 spørgsmål for hver af de tre dimensioner, hvor det for hvert spørgsmål var muligt at afgive en vurdering på 1-5⁴. Vurderingerne blev foretaget af drengene selv via spørgeskemaer⁵.

¹ På DrengAkademiet arbejdes der inden for matematik primært med tal og algebra som stofområder, da erfaring fra tidligere har vist, at drengene har mest brug for at arbejde med grundlæggende matematik - herunder de fire regningsarter. Således henviser "matematik" i denne rapport til matematiske færdigheder i tal og algebra.

² Prøvesættet er tilpasset ved at skære de dele fra, som ikke omhandler matematiske færdigheder i tal og algebra.

³ Konkret blev der udvalgt spørgsmål fra dimensionerne "kognitive kompetencer", "non-kognitive kompetencer", "Fysisk og psykisk velbefindende" samt fra "Psykosocialt undervisningsmiljø" fra pilotundersøgelsen. Et enkelt blev tilføjet udefra.

⁴ Afhængig af det enkeltes spørgsmåls formulering var 1 eller 5 bedst. I forhold til indekskonstruktion blev skalaerne vendt samme vej.

⁵ Manglende vurderinger blev kodet til gennemsnittet

Hvordan kom vi frem til resultaterne?

De faglige resultater blev analyseret i flere trin. Først blev der udregnet procentpointvise ændringer mellem før- og eftermålingerne for at få et overblik over, hvor meget deltagerne i gennemsnit havde forbedret sig. Dernæst blev der udført signifikanstest for at undersøge, hvorvidt forskellen på før- og efter målingerne kunne tænkes at skyldes tilfældig variation i data og således ikke en reel fremgang. Til det formål blev der anvendt t-test for matchede par. Videre blev der testet for, om differenserne mellem før- og efter-scoringerne var normalfordelte med Kolmogorov-Smirnovs test, og der blev foretaget en visuel vurdering⁶. Blev der fundet afvigelser fra normalfordelingen blev disse testet med en Wilcoxon Signed Rank Test. Alle faglige resultater blev fundet signifikante på et $p < 0,01$ niveau, og der var ingen væsentlige forudsætningsbrud. Således må det konkluderes, at den observerede faglige fremgang ikke blot skyldes tilfældighed i data.

Derefter blev der beregnet effektstørrelse for de faglige test i form af Cohens d (Cohen, 1988). Effektstørrelse-målet har til formål at undersøge den gennemsnitlige effekt af en indsats i relation til spredningen i data, således at man kan sammenligne på tværs af skalaer. I dette tilfælde mellem matematik-, stavnings-, og læsningstest. Cohens d er blandt de mest anvendte standardiserede effektstørrelsesmål inden for uddannelsesforskning, specielt inden for den gren, der beskæftiger sig med metaanalyser, hvor man samler resultaterne fra et stort antal studier omhandlende samme emne (fx Hattie, 2009; 2012). Således giver Cohens d mulighed for at sammenligne effekter på tværs af interventioner⁷.

Udviklingen i fagligt orienterede kompetencer, handlekompetencer og almen trivsel blev analyseret ved, at vurderingerne for hver dimension blev lagt sammen i såkaldte additive indeks (altså værdien på indekset udregnes som summen af værdierne af de besvarede underspørgsmål) samt som et samlet indeks (summen af scoren på de individuelle indeks). Disse blev signifikant testet efter samme procedure som tidligere beskrevet, og der blev udregnet procentvis ændring fra før- til eftermåling. Signifikantestene viste, at alle ændringer var signifikante på et $p < 0,01$ niveau. Endelig blev ændringerne i indeksene søgt sammenholdt den faglige progression inden for de tre fagområder.

Hvad betyder det? Effektstørrelse og læringsår

Effektstørrelse (Cohens d) tolkes i litteraturen oftest (på Jacob Cohens anbefaling) som, at området omkring 0,2 kan betegnes som en lille effekt, 0,5 som en moderat effekt og 0,8 eller derover som en stor effekt (Cohen, 1988). Flere forskere peger dog på, at en sådan tolkning ikke bør foretages ukritisk, og det foreslås, at man i stedet relaterer effektstørrelsen til det specifikke genstandsfelt for sin intervention/indsats og til den kontekst, hvori indsatsen foregår samt i forhold til indsatsens målgruppe (Hill et al., 2008; Bloom et al., 2008; Hattie, 2009). Fx kan nogle målgrupper eller indsatsområder være langt tungere at påvirke end andre, og i sådanne tilfælde vil effektstørrelser, som fremstår små efter Cohens fortolkning, måske reelt kunne anses for ganske store set i forhold til konteksten (Lipsey et al., 2012).

⁶ I den sammenhæng blev der anvendt QQ-plots, hvor der ses efter visuelle afvigelser fra den estimerede normalfordeling.

⁷ Cohens d beregnes ved at dividere forskellen i gennemsnit mellem før og eftermålingerne med (pooled) standardafvigelsen i data (Cohen, 1988).

Det må dog samtidig slås fast, at det er en stor fordel i forhold til fortolkning og formidling at kunne relatere resultaterne til fremgang både alment i skolesystemet og til andre indsatser, der har til formål at øge faglige resultater.

Et bud på fortolkning kunne være en af nyere tids mest citerede uddannelsesforskere, John Hattie, der har samlet 800 metaanalyser (analyser, der samler resultater fra et stort antal enkeltstudier) med henblik på at opstille et pejlemærke for, hvor stor effekten af en indsats bør være, før det kan betale at tage den i brug og investere i den (Hattie, 2009). Hattie når frem til en effektstørrelse på 0,4, primært da det er den gennemsnitlige effektstørrelse blandt samtlige interventioner, programmer og andre tiltag, der har en effekt på elevers faglighed, som indgår i hans undersøgelse. Ligeledes svarer 0,4 cirka til det udbytte eleverne gennemsnitligt får af et læringsår, forstået som effekten af ét års almindelig undervisning i skolen (Ibid.).

I en dansk kontekst rapporterer Undervisningsministeriet (2014b) i et notat, at den gennemsnitlige effekt af et læringsår, målt i de nationale læsetest, ligeledes er i omegnen af 0,4, og en rapport om effekten af danske forsøg med tolærerordninger fra 2014, udgivet af TrygFondens Børneforskningscenter i samarbejde med Rambøll Management Consulting, anvender samme mål for et læringsår (Andersen et al., 2014).

På den baggrund - og for at lette fortolkningen - vil de faglige effekter af DrengAkademiet blive omregnet til læringsår ved en omregningsfaktor på 0,4.

Man bør dog holde for øje, at målet har en bred karakter, da der er tale om et generelt gennemsnit af alle elever uanset fx baggrund, køn og klassetrin. Dette er specielt relevant, da fx amerikanske studier peger på, at effekten af et læringsår er større end 0,4 på de laveste klassetrin og mindre i de største klasser (Hill et al., 2008; Bloom et al., 2008; Lipsey et al., 2012). Dertil kommer, at piger tilsyneladende får et større gennemsnitligt udbytte af et læringsår end drenge, og at denne forskel vokser gennem skoletiden (ibid.). Man skal naturligvis være varsom med direkte at overføre dette til en dansk kontekst, men det er værd at bemærke, at DrengAkademiets målgruppe netop er drenge fra de øverste klassetrin. Antaget at det ovenstående mønster også gør sig gældende i det danske skolesystem, burde det betyde, at den anvendte enhed for læringsår er et konservativt skøn, når målgruppen tages i betragtning. Det bør dog samtidig understreges, at man ikke med sikkerhed kan fastslå præcis, hvor stor den gennemsnitlige faglige effekt af et læringsår er, målt i effektstørrelse, hverken i en dansk eller i en international kontekst⁸. Det er således at betragte som et kvalificeret bud.

Hvad bør man ellers være opmærksom på? - andre metodiske forbehold

Ud over de nævnte forbehold og opmærksomhedspunkter i forhold til konvertering til læringsår, er der tre andre forhold, man bør være opmærksom på i forhold til fortolkningen af resultater fra DrengAkademiet 2014.

⁸ Fx kan valg af test og kontekstuelle faktorer have en betydning.

For det første tages der forbehold for den såkaldte træningseffekt. Det vil sige effekten af, at deltagerne udfører den samme type test flere gange inden for et relativt kort tidsrum, og således må det forventes, at deltagerne bliver mere øvede for hver testgang.

For det andet er det vigtigt at understrege, at DrengesAkademiets resultater ikke med tilstrækkelig sikkerhed kan overføres direkte til samme målgruppe i hele befolkningen. Det skyldes, at deltagerne, på trods af en pæn geografisk spredning, ikke er tilfældigt udvalgt og dermed ikke nødvendigvis er repræsentative for hele populationen af tilsvarende drenge.

Desuden er der anvendt et standard en-gruppens før- og eftermålingsdesign tilført en midtermåling. Det vil sige, at det ikke er muligt at skabe sikkerhed om, hvor meget af effekten, der specifikt stammer fra DrengesAkademiet. For at undersøge dette kunne man have gennemført et lodtrækningsforsøg, hvor deltagerne blev tilfældigt fordelt mellem DrengesAkademiet og en anden sammenlignelig indsats samt eventuelt ingen indsats.

Det er dog værd at bemærke, at det nu er fjerde gang DrengesAkademiet har kørt med konsistente gode resultater. Det øger alt andet lige sandsynligheden for, at indsatsen ville have en positiv effekt - også i en anden kontekst.

Resultater fra DrengesAkademiet 2015

I dette afsnit vil resultaterne fra dette års DrengesAkademi blive fremlagt og fortolket.

Først præsenteres det mest centrale; de faglige resultater i form af deltagernes fremgang i de faglige test opgjort i procentpoint. Derefter ses der nærmere på, hvordan det faglige top-, middel- og bundniveau blandt deltagerne har ændret sig i løbet af akademiet. Endelig udføres den egentlige statistiske analyse, og der beregnes effektstørrelse, som efterfølgende omregnes til læringsår med henblik på at relatere og perspektivere den faglige fremgang til omverdenen og dermed lette fortolkningen.

Afsluttende vil der blive redegjort for drengenes fremgang på forskellige trivsels- og kompetencedimensioner. Endvidere vil eventuelle ændringer på disse trivsels- og kompetencedimensioner blive sammenholdt med udviklingen i de faglige test.

Hovedresultater sommeret

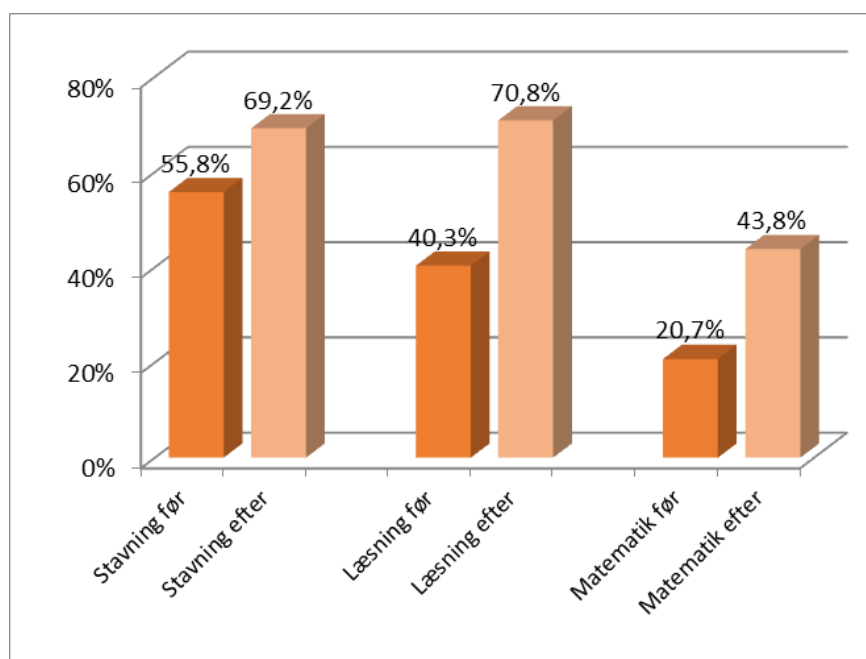
- Der er markant fremgang på alle tre faglige mål, størst inden for læsning og matematik.
- Fremgangen i deltagernes samlede faglige niveau i læsning og stavning er størst for bund- til middelniveauet. I matematik ses den største fremgang for middel- og topniveauet.
- Store effektstørrelser inden for læsning og matematik og moderat inden for stavning.
- Alle faglige resultater er statistisk signifikante med mindre end 1 % sandsynlighed for, at fremgangen blot er udtryk for tilfældighed i data.
- Der er positive statistisk signifikante fremgange på alle tre trivsels- og kompetencedimensioner. Det viser sig, at kun fremgangen på drengenes selvvalgte faglige kompetencer er delvist korreleret med udviklingen i de faglige test. Sammenhængen er størst i forhold til udviklingen i matematik.

Fremgang i de faglige test

Det centrale mål for DrengAkademiets succes er, at eleverne opnår faglig fremgang gennem forløbet, hvilket må siges at være opnået. Figuren nedenfor viser gennemsnitlig andel rigtige svar i de tre faglige test før og efter campen.

På dette års camp ses fremgang på cirka 23 procentpoint i matematik, 13 procentpoint i stavning, og hele 31 procentpoint i læsning fra start til slut.

Figur 1: Den faglige udvikling på de tre områder udtrykt i gennemsnitlig andel rigtige besvarelser, før og efter DrengAkademiet.



N = 98

I tabellen nedenfor vises den gennemsnitlige fremgang mellem første og sidste runde af faglige test målt i procentpoint for de seneste tre afholdte DrengAkademier.

Tablet 2: Den faglige udvikling i procentpoint de tre seneste år

	Matematik før-efter	Stavning før-efter	Læsning før-efter
2015	+ 23	+ 13	+ 31
2014	+19	+10	+27
2013	+24	+15	+28

N(2013)=94, N(2014)=85, N(2015)=98

Overordnet set er fremgangen markant og sammenholdt med de to sidste års resultater tyder meget på, at man (som i 2014) har opnået en stabil fremgang på kun to ugers camp i 2015 versus de tre ugers campen varede i 2013.

Det er også værd at bemærke, at der synes at være et mønster i resultaterne fra DrengAkademierne 2013-2015. Hermed forstås, at deltagerne generelt udviser sammenlignelige forbedringer i de samme fag i 2013, 2014 og 2015. De har således fået det største udbytte inden for læsning,

et lidt mindre inden for matematik og et noget mindre - men dog stadig væsentligt udbytte - inden for stavning i både 2013, 2014 og 2015. Dette mønster bliver ikke mindre interessant, når man tager i betragtning, at DrengeAkademiet 2014 og 2015, som tidligere nævnt, blev afviklet på to uger mod tre uger i 2013. På den baggrund synes det ikke urimeligt at antage, at resultaterne i 2014 og 2015 havde været endnu bedre såfremt indsatsen havde haft samme varighed i 2014 og 2015 som i 2013 - alt andet lige.

De faglige resultater er blevet testet statistisk - før-resultater mod efter-resultater - og blev fundet signifikante på $p < 0,01$ niveau, hvilket vil sige, at der er mindre end 1 % sandsynlighed for, at resultaterne er udtryk for tilfældig variation. Dermed må konklusionen om en reel fremgang blandt deltagerne betegnes som meget valid, i statistisk forstand.

Udvikling i faglige niveauer

Med henblik på opnå en dybere forståelse af resultaterne fra de faglige test, blev der foretaget en percentilanalyse, med det formål at undersøge, hvordan deltagergruppens samlede bund-, midt- og topniveau har ændret sig.

Analysen siger ikke noget om, hvordan den enkelte dreng har forbedret sine resultater, men derimod noget om, hvordan gruppens niveau har udviklet sig som helhed. Percentilanalysen er et supplement til den analyse af gennemsnit og effektstørrelser, der blev præsenteret tidligere, og som der arbejdes videre med efterfølgende. Den giver en mere detaljeret viden om, hvordan deltagerens fordeling henover de faglige test forandres i stedet for blot at beskrive, hvordan et gennemsnit har flyttet sig. Dette kan eksempelvis benyttes til at identificere, om der er grupper med bestemte faglige udgangspunkter, for hvem indsatsen synes at virke bedre end for andre. En percentilanalyse anbefales da således også af en gruppe fremstående uddannelsesforskere som en fordelagtig metode til at fortolke udviklingen i faglige resultater, jf. Lipsey et al. (2012).

Percentiler er et udtryk for, hvor stor en andel af den samlede gruppe, som ligger under et bestemt antal procent rigtige i de faglige test. Ser vi fx på den 95. percentil udtrykker den, hvor mange procent rigtige svar 95 % af gruppen lå under, og 5 % lå over i den respektive test. Eksempelvis fik de 5 %, der klarede sig dårligst i den første stavetest, under 1,4 % rigtige svar, hvorimod de 5 % der klarede sig bedst fik over 88,6 %, som det fremgår af tabellen nedenfor:

Tabel 3: Udviklingen i deltagerens faglige niveauer fordelt på percentiler

Percentiler	StavFør	StavMidt	StavEfter	LæsFør	LæsMidt	LæsEfter	MatFør	MatMidt	MatEfter
5	1,4 %	19,7 %	27,1 %	7,9 %	25,9 %	36,0 %	4,1 %	6,3 %	14,4 %
10	19,7 %	36,0 %	42,6 %	13,8 %	32,0 %	45,6 %	4,2 %	10,2 %	16,7 %
25	44,3 %	52,9 %	61,1 %	22,0 %	48,0 %	58,0 %	9,9 %	14,6 %	27,1 %
50	60,7 %	61,4 %	71,4 %	40,0 %	64,0 %	73,0 %	16,7 %	22,9 %	45,8 %
75	71,8 %	72,9 %	84,3 %	56,5 %	78,5 %	82,5 %	29,2 %	39,6 %	58,3 %
90	84,3 %	81,4 %	88,6 %	74,0 %	86,2 %	90,2 %	43,8 %	52,1 %	71,0 %
95	88,6 %	88,6 %	90,1 %	80,0 %	90,1 %	96,0 %	50,2 %	60,5 %	79,3 %

N=98

Tabellen indikerer, at der er en udpræget tendens i mod at både bund-, midt- og topniveau hæves, dog i forskellig grad i de respektive test. Fx ser topniveauet i matematik ud til at forbedres i højere grad end i de andre fag, da grænsen for de 5 %, der klarede sig bedst, var 50,2 % rigtige svar i den første test mod 79,3 % i den afsluttende.

I stavning derimod er der kun mindre forskel på topniveauet før og efter; de 5 %, der klarede sig bedst fik 88,6 % rigtige i før-testen - et tal, der steg til 90,1 % efter deltagelse på DrengAkademiet.

Tabellen viser også, at de forskellige niveauer generelt udvikler sig forskelligt fra fag til fag.

For at danne sig et overblik over den faglige udvikling, kan man med fordel studere ændringen i procentpoint for hver percentil og, dermed, for hvert niveau. Dette kan tolkes som et udtryk for, hvor meget de respektive faglige niveauer har ændret sig mellem første og sidste testrunde. Fx hvor meget topniveauet er steget fra første test til sidste test. Det bør dog understreges, at der er tale om deltagernes udvikling som samlet gruppe. Med dette menes, at det ikke nødvendigvis er de samme drenge, der klarer sig bedst i den første test og den sidste test, om end sandsynligheden for det må antages at være høj. Man kan derimod, som tidligere nævnt, rent forståelsesmæssigt, tale om, at niveauerne ændrer sig.

Tabel 4: Ændring i procentpoint (PP) mellem før- og eftertest fordelt på percentiler

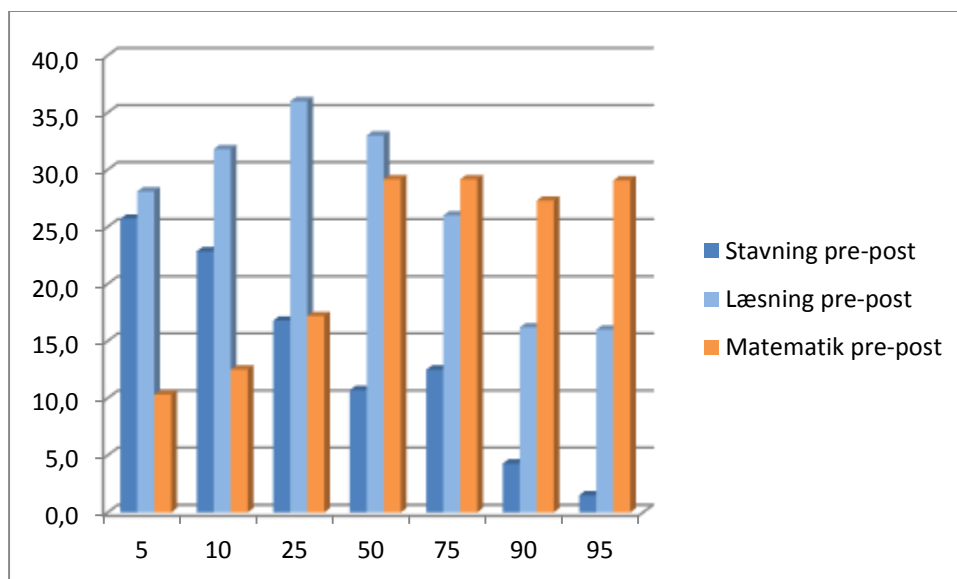
Percentil	Stavning ændring i PP	Læsning ændring i PP	Matematik ændring i PP
5	+25,7	+28,1	+10,3
10	+22,9	+31,8	+12,5
25	+16,8	+36,0	+17,2
50	+10,7	+33,0	+29,2
75	+12,5	+26,0	+29,2
90	+4,3	+16,2	+27,3
95	+1,5	+16,0	+29,1

N=98

Tabellen ovenfor viser igen, at der er sket faglige forbedringer inden for samtlige niveauer. Fx ligger barren for de bedste 5 % 29,1 procentpoint højere inden for matematik i den sidste test sammenlignet med den første test.

Niveauforskellen illustreres tydeligere i figuren nedenfor, der peger på, at særligt bund- og midelniveauet er blevet markant hævet for gruppen som helhed. Dertil synes udviklingen inden for matematik og stavning, og til dels også læsning, at være modsatrettede, idet den generelle fremgang i stavning sker ved et løft af bund- og midelniveauet, hvorimod fremgangen i matematik som nævnt i højere grad er et udtryk for, at topniveauet hæves.

Figur 2: Grafisk fremstilling af ændringen i procentpoint mellem før- og eftertest fordelt på percentiler



N=98

Effektstørrelse og læringsår

Det er nu blevet statistisk sandsynliggjort og dokumenteret, at deltagerne på DrengAkademiet 2015 har forbedret deres faglige resultater fra før-test til efter-test målt i procentpoint, og det er blevet undersøgt, hvordan denne fremgang fordeler sig mellem bund-, middel- og topniveauerne for den samlede gruppe. Gruppen klarer sig altså væsentligt bedre i de faglige test efter at have gennemført LøkkeFondens to ugers sommercamp, og DrengAkademiet har derfor umiddelbart haft en markant positiv effekt på deltagerens faglige præstationer⁹.

Det er imidlertid også væsentligt at undersøge, hvor stor en effekt der er tale om, samt hvordan denne kan relateres til det faglige udbytte målgruppen normalt antages at opnå i løbet af et læringsår (et års undervisning) i folkeskolen generelt.

Sidst, men ikke mindst, er der derfor beregnet effektstørrelser i form af Cohens d ¹⁰. Beregningerne er foretaget i statistikprogrammet STATA og fremgår af tabellen nedenfor:

Tabel 5: Beregning af effektstørrelser

	Stavning	Læsning	Matematik
Effektstørrelse (Cohens d)	0,652	1,587	1,303

N=98

⁹ Med de forbehold, der er nævnt i metodeafsnittet.

¹⁰ Cohens d uddybes i metodeafsnittet.

Hvis vi i første omgang ser bort fra konverteringen til læringsår, så er den gængse metodiske tommefingerregel (fremsat af Jacob Cohen selv i 1988), som tidligere nævnt, at en effektstørrelse på 0,2 kan fortolkes som lille, 0,5 som moderat og 0,8 og derover som stor. Reelt siger effektstørrelsen i form af Cohens d noget om den gennemsnitlige forskel på før- og eftermålingen i forhold til gruppens standardafvigelse (spredning omkring gennemsnittet). Det betyder, at effektstørrelsen også måler, hvor meget deltagerne på DrengesAkademiet gennemsnitligt har flyttet sig ved at se på afstanden til gennemsnittet og ikke kun gennemsnittet alene. Effektstørrelsen siger altså noget om standardiseret effekt og er således uafhængig af den enhed, man målte med i henholdsvis matematik, stavning og læsning.

Jf. det ovenstående kan man dermed konkludere, at der kan ses en stor effekt inden for matematik og læsning og en moderat effekt inden for stavning.

I tabel 6 nedenfor ses effektstørrelsen udregnet på baggrund af forskellen mellem før- og eftermålingerne omsat til antal læringsår, jf. Hattie (2009):

Tabel 6: Effekten omregnet til læringsår*

	Stavning	Læsning	Matematik
Læringsår	1,63	3,97	3,26

*Omregning til læringsår beregnes ved $\frac{\text{Cohens } d}{0,4}$

N=98

Man skal dog, som nævnt, være meget påpasselig med at foretage en sådan konvertering ukritisk. Disse forbehold er beskrevet i metodeafsnittet, men kort fortalt peger amerikanske studier på, at drenge på de øverste klassetrin får et lavere gennemsnitligt udbytte af et års skolegang end drenge på de mindste klassetrin (se fx Hill et al., 2008; Bloom et al., 2008; Lipsey et al., 2012). Omregningen til læringsår kan være rimelig at benytte til at relatere en indsats til omverdenen, så længe man er opmærksom på, at det er et omdiskuteret emne, og at man derfor går til det med forbehold.

Alt i alt kan det konkluderes, at DrengesAkademiet har haft en stor effekt på deltageres faglige resultater, der i den ovenstående fortolkningsramme, forsigtigt, kan antages at svare til mellem 1,63 og 3,97 læringsår. Eller med andre ord: 1,63 til 3,97 års gennemsnitlig skolegang, fortolket med de nævnte forbehold¹¹.

Almen trivsel, fagligt orienterede kompetencer og handlekompetencer

Som supplement til den faglige hovedanalyse blev drengene også testet i forhold til deres oplevede progression inden for almen trivsel, generelle handlekompetencer samt selvvaluerede fagligt orienterede kompetencer. Disse mål er desuden blevet samlet i et indeks, som giver et overordnet

¹¹ Her er det værd at bemærke, at effektstørrelserne i form af Cohens d er væsentlig større end for DrengesAkademiet 2014. Det skyldes angiveligt 1) en mindre stigning i den numeriske størrelse på effekterne (større forskel mellem start og slutmåling) koblet med - og dette er den vigtigste årsag: 2) fordelingsmæssige forskydninger mellem årene, der gør standardafvigelserne forskellige. Dette kunne tale for, at deltagerne i 2015 versus 2014 ikke udgør helt statistisk sammenlignelige grupper, hvorfor der heller ikke præsenteres en sammenligning af Cohens d-størrelserne mellem 2015 og 2014 eller tidligere år i denne rapport.

billede af, hvorledes drengenes trivsel på disse vigtige dimensioner har udviklet sig under deres deltagelse på DrengAkademiet.

De tre indeks samt det overordnede (samlede) trivselsindeks er - for at lette fortolkningen - blevet omregnet til en skala fra 0 til 100, hvor 0 er lavest og 100 er højest.

Udviklingen på de tre indeks samt det overordnede trivselsindeks er angivet i tabellen nedenfor. Der ses fremgange på alle tre indeks og dermed også det samlede trivselsindeks. Der er i alle tilfælde tale om statistisk signifikante fremgange.

Mindst fremgang spores i forhold til almen trivsel med en stigning på blot 4,9 point. Man bør dog bemærke, at drengene fra start allerede lå højt på trivselsindekset set i forhold til de to andre indeks. Indekset for handlekompetencer går 7,1 point frem, mens indekset for de selvvaluerede faglige kompetencer går hele 19,1 point frem - en stigning på næsten 38 %. Det samlede indeks for trivsel er steget med 10,8 point.

Samlet set kan det konkluderes, at det især er drengenes vurdering af egne faglige kompetencer, som er stigende henover campen. Dertil observeres en noget mindre fremgang i drengenes generelle handlekompetencer og almene trivsel.

Tabel 7: Progression i trivsels- og kompetenceindeks og ændring i indekspoint fra start til slut

	Indeks almen trivsel og velvære		Indeks selvvaluerede faglige kompetencer		Indeks handlekompetence		Samlet trivselsindeks	
	Før	Efter	Før	Efter	Før	Efter	Før	Efter
Gennemsnitlig score	82,7	87,6	50,5	69,6	65,0	72,0	66,2	77,0
Ændring, før til efter	4,9		19,1		7,1		10,8	
Relativ ændring	5,9 %		37,8 %		10,9 %		16,3 %	

N=98

Sammenhæng mellem udvikling i trivsel og faglig udvikling

Det synes interessant at undersøge, om der kan sandsynliggøres nogen sammenhæng mellem udviklingen på de forskellige dimensioner af trivsel og den faglige fremgang målt i testene, som blev beskrevet ovenfor.

Til dette formål præsenteres resultaterne af en korrelationsanalyse i tabellen nedenfor, hvor udviklingen i de tre faglige test parvis sammenholdes med de tre trivselsindeks og det samlede indeks. Korrelationen siger noget om, hvor stærk sammenhængen mellem trivslen og den faglige udvikling er i form af en korrelationskoefficient, som går mellem -1 og 1. Hvis der er en positiv sammenhæng mellem drengenes fremgang i trivsel og deres faglige udvikling, bør korrelationskoefficienten være signifikant større end 0.

Tabel 8: Resultat af korrelationsanalyse mellem udviklingen i trivselsindeks og de faglige test

	Læsning	Stavning	Matematik
Almen trivsel	0,04	-0,08	0,03
Generaliserede handlekompetencer	0,10	-0,02	0,03
Selvurderede faglige kompetencer	0,18*	-0,04	0,27**
Samlet trivselsindeks	0,16	-0,07	0,18*

*Signifikant $p < 0,1$ niveau **Signifikant på $p < 0,01$ niveau..

N=98

Resultatet af analysen viser, at der ikke synes at være de store direkte sammenhænge mellem udviklingerne i trivsel og i de faglige test. Den eneste dimension, der er positivt korreleret med fremgangen i de faglige test, inden for rimelig statistisk sikkerhed, er drengenes selvurderede faglige kompetencer. Her ses der en høj-signifikant sammenhæng med udviklingen i matematik og en mindre signifikant sammenhæng med udviklingen i læsning. Det er således også denne sammenhæng, der gør det samlede trivselsindeks lav-signifikant i forhold til matematik. Høj- og lav-signifikant skal i denne sammenhæng forstås som henholdsvis mindre end 1 % sandsynlighed og mindre end 10 % sandsynlighed for, at de observerede sammenhænge blot skyldes tilfældighed i data.

I forhold til de forventninger man kunne have på forhånd er det måske ikke overraskende, at den almene trivsel ikke korrelerer med udviklingen i de faglige test. I det tidligere afsnit blev det godtgjort, at udviklingen i den almene trivsel var lille, idet den allerede var meget god - *i udgangspunktet*. Altså *før* campen. Derudover kan det nævnes som en mulighed, at de udvalgte områder i forhold til trivsel ikke indfanger de udfordringer, som drengene har i forhold til deres faglige udvikling. Altså at de faglige udfordringer har andre årsager, end de der er målt på i denne undersøgelse.

Samlet set vurderes altså, at drengenes egen vurdering af faglige evner i nogen grad hænger sammen med den faglige udvikling, mens der ikke er fundet nogen yderligere sammenhænge i forhold til de andre dimensioner af trivselsbegrebet, der anvendes i denne rapport.

Det skal dog samtidig understreges, at drengene har haft en væsentlig positiv fremgang på de anvendte trivselsmål sideløbende med den store markante faglige fremgang. Der kan blot ikke påvises en direkte statistik valid sammenhæng mellem de to progressioner.

Opsamling og perspektivering

På baggrund af denne rapports gennemgang af metode og resultater fra DrengAkademiet 2015, kan det konkluderes, at der igen i år har været en markant faglig fremgang blandt deltagerne. Samtidig har drengene oplevet øget tiltro til egne fagligt orienterede kompetencer, generelle handlekompetencer samt almen trivsel og velvære. Da dette er det fjerde akademi i rækken med konsistent gode resultater, virker det ikke urimeligt at konkludere, at der synes at være tale om en særdeles effektiv indsats. Det er dog samtidig værd at se på, hvor effekten stammer fra, og i hvilket omfang den kan forventes at holde ved fremadrettet.

DrengAkademiet som indsats lægger vægt på rammer, struktur, selvevaluering, feedback, engagement, motivation og synlig læring, og selvom fraværet af en kontrolgruppe gør det vanskeligt at undersøge, hvorvidt DrengAkademiet har en større effekt end andre sammenlignelige indsatser, peger resultaterne på, at det tilsyneladende har ladet sig gøre igen at flytte deltageres faglige niveau markant på meget kort tid. Det ville dog stadig være fordelagtigt på sigt at gennemføre et kontrolgruppeforsøg eller et match- studie, med henblik på at skabe sikkerhed om, hvor stor en del af den faglige progression, der kan tilskrives henholdsvis DrengAkademiet som specifik indsats og/eller den træningseffekt, der sker i og med, at drengene bliver øvede i de faglige test.

I forhold til, hvorvidt den faglige fremgang er vedvarende, kunne man med fordel undersøge dette fremadrettet, da det endnu ikke har været muligt at dokumentere. Muligheden for at gennemføre en sådan undersøgelse er i øjeblikket ved at blive udforsket, ikke mindst da drengene fra de første runder af DrengAkademiet efterhånden er nået et godt stykke videre i deres livs- og evt. uddannelsesforløb. Det er dog værd at bemærke, at man bør indtænke den såkaldte "boost-effekt", der blandt andet dækker over, at korte interventioner ofte foregår uden for deltageres almindelige kontekst (hvor de ikke altid ville præstere lige så godt), og at korte interventioner ikke nødvendigvis påvirker den dybere læringsevne og derfor kan forventes at fortage sig over tid (Paris et al., 2006). Det er derfor vigtigt at vedligeholde indsatsen, hvis man vil øge sandsynligheden for at så meget som muligt af effekten varer ved i deltageres videre skolegang. For at imødekomme dette, sørger LøkkeFonden for obligatorisk mentorordning for alle drengene i et år, så nye vaner og værktøjer, bliver forankret i drengens hverdag.

Afsluttende bør det bemærkes, at resultaterne ikke nødvendigvis bør tolkes som, at DrengAkademiet alene øger deltageres faglige niveau. Der er antageligt også tale om, at DrengAkademiet aktiverer allerede eksisterende kompetencer og fremmer deltageres tiltro til egen læringsevne og til, at de kan bruge det, de har lært. Begge dele må anses for vigtige i forhold til at give drengene det bedst mulige udgangspunkt for deres videre forløb i uddannelsesverdenen og i livet generelt.

Referencer

Andersen, S. C., Beuchert-Pedersen, L. V., Nielsen, H. S., & Thomsen, M. K. (2014). *2L Rapport: Undersøgelse af effekten af tolærerordninger*. Aarhus universitet

Bloom, H. S., Hill, C. J., Black, A. B., and Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, 1(4): 289-328.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London; New York: Routledge.

Hattie, J. (2012). *Visible learning for teachers: maximizing impact on learning*. London; New York: Routledge.

Hill, C. J., Bloom, H. S., & Black, A. R., Lipsey, Mark W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2(3), 172-177.

Keilow, M., Holm, A., Bagger, S., & Henze-Pedersen, S. (2014). *Udvikling af trivselsmålinger i folkeskolen: en pilotundersøgelse*. SFI.

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). *Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms*. (NCSE 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Paris, S. G., Morrison, F. J., & Miller, K. F. (2006). Academic Pathways From Preschool Through Elementary School, In: Alexander, P. A., & Winne, P. H. (eds.). *Handbook of Educational Psychology*. Erlbaum, 61-87

Undervisningsministeriet (2014a). *Endelige anbefalinger fra ekspertgruppen om elevers trivsel i folkeskolen i relation til nationale trivselsmålinger*. Lokaliseret d. 27.8.2015 på:

http://www.uvm.dk/-/media/UVM/Filer/Aktuelt/PDF14/140620%20Ekspertgruppe_trivselsmaaling.pdf

Undervisningsministeriet (2014b). *Notat om resultaterne i tolærerforsøget og forsøgsprogrammet om modersmålsbaseret undervisning*. Lokaliseret d. 27.8.2015 på:

http://uvm.dk/-/media/UVM/Filer/Udd/Folke/PDF14/Apr/140410_Notat_resultaterne_tolaerer-forsoeget_og_forsoesprogrammet_om_modersmaalsbaseret_undervisning.pdf