

FORSKNINGSRAPPORT

DrengeAkademiet 2016

Årets resultater

Indholdsfortegnelse

Indledning	2
Metode.....	3
Hvem deltog?	3
Hvad og hvordan målte vi?	4
Hvordan kom vi frem til resultaterne?	4
Hvad betyder det? Effektstørrelse og læringsår.....	5
Hvad bør man ellers være opmærksom på? – andre metodiske forbehold.....	7
Resultater fra DrengesAkademiet 2016	7
Hovedresultater summeret	7
Fremgang i de faglige test	8
Udvikling i faglige niveauer.....	9
Effektstørrelse og læringsår	12
De syv karaktertræk	13
Fysisk aktivitet.....	15
Sammenhæng mellem udvikling på de syv karaktertræk og den faglige udvikling	16
Opsamling og perspektivering.....	17
Referencer.....	19

Indledning

I sommeren 2016 blev endnu en runde af DrengAkademiet vel gennemført. Den generelle holdning blandt både deltagere og undervisere var, at deltagerne havde gennemgået en faglig og social meget positiv udvikling i perioden på akademiet. Det er dog centralt for LøkkeFonden at kunne dokumentere resultaterne med henblik på at sikre sig, at disse var reelle og ikke alene oplevede samt forsøge at kvantificere den formodede progression blandt deltagerne.

I den forbindelse blev der, som i tidligere år, tilknyttet en konsulent med speciale indenfor dokumentation og evaluering med det formål at følge indsatsen i forhold til at dokumentere resultaterne fra akademiet samt at skabe gennemsigtighed om, hvordan resultaterne blev frembragt.

Dette papir har til formål at præsentere og analysere resultaterne fra DrengAkademiet 2016 samt at redegøre for den anvendte metode, og de overvejelser, der er blevet gjort i den sammenhæng.

Data og resultatet er analyseret af Alexander Forbes McClellan, der er uafhængig af LøkkeFonden.

Metode

Det følgende afsnit har til formål at skabe klarhed og gennemsigtighed om, hvordan data og resultater fra DrengAkademiet 2016 er blevet indsamlet og analyseret.

Hvem deltog?

Rekrutteringen af deltagere til dette års akademi foregik, som det var tilfældet i runderne fra 2013 til 2015, 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 ligeledes bidraget til rekruttering af nye deltagere. For at komme i betragtning til DrengAkademiet udfyldte drenge og forældre et ansøgningskema om baggrund, fagligt niveau og motivation for deltagelse. På den baggrund blev ansøgerne sorteret efter følgende tre udvælgelseskriterier:

- 1) Ansøgeren skulle være dreng skulle være i aldersgruppen 14 til 16 år, og ideelt set gå i 8. klasse og være på vej i 9. klasse. 8. klasses drenge blev foretrukket, da erfaring viser, at de er mere fokuserede under indsatsen. Dog blev drenge, der gik i 7. klasse, også optaget, samt en enkelt dreng fra 6. klassetrin.
- 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 2015 ikke tilbudt optagelse.

Dertil blev det så vidt muligt tilstræbt at sammensætte deltagergruppen således, at den var så demografisk repræsentativ som mulig, hvilket selvfølgelig afhang af ansøgerfeltets sammensætning.

Blandt de ansøgere, der opfyldte kriterierne, blev der udvalgt 107, som takkede ja til en plads. 102 mødte op til introcampen, hvor 2 deltagere faldt fra. Dertil faldt 5 deltagere fra under selve hovedcampen. Således gennemførte 95 drenge dette års akademi, heraf 5,3 % af anden etnisk herkomst end dansk. Aldersgennemsnittet blandt drengene var 14,6 år.

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

Tabel 1: Geografisk fordeling og klassetrin

	Sjælland og øer	Fyn og øer	Jylland og øer	6. klasse	7. klasse	8. klasse
Andel	40,0 %	7,4 %	52,6 %	1,1 %	46,3 %	52,6 %

N=95

Det bemærkes, at der blandt deltagerne var en relativ mindre andel 8. klasses elever i forhold til DrengAkademiet 2015, hvor der var 69,4 % 8. klasses elever (LøkkeFonden, 2015).

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 (18. juni-19. juni) samt i slutningen af akademiets første og anden uge (27. juni – 10. 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. Resultaterne fra testene afrapporteres her i procent rigtige svar og forskellen mellem før og efter målingerne i procentpoint. 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 personlighedsmæssige udbytte, i form af indikatorer for de syv non-kognitive egenskaber eller karaktertræk, der arbejdes med på DrengAkademiet. Inspirationen til det systematiske arbejde med og evaluering af de 7 karaktertræk kommer fra KIPPskolerne i USA og fra positiv psykologi³. Yderligere suppleredes målingerne med indikatorer for deltagerens oplevelse af egen fysiske form og samspillet med fysisk aktivitet og trivsel/sundhed/læring.

Hvert af de syv karaktertræk samt indikatorerne for fysisk aktivitet blev respektivt målt med tre spørgsmål på en Likert-skala med vurdering fra 1 til 5. Den samlede scoring på karaktertrækket foregik derefter ved at tage gennemsnittet af besvarelserne på de tre spørgsmål, der blev brugt til at måle det pågældende karaktertræk. Karaktertrækkene blev målt fire gange henover hovedcampen, ved start (27. juni), ved afslutning (9. juli) og to gange derimellem (30. juni og 6. juli). Kun elever, der har besvaret spørgeskemaet ved den første og sidste måling indgår i analyserne.

Det var i den forbindelse en væsentlig prioritet at reducere antallet af spørgsmål mest muligt, idet deltagerne på DrengAkademiet allerede bliver testet meget på kort tid, og erfaringer fra tidligere afholdte DrengAkademier indikerede, at mange blev testtrætte og opgav at gennemføre mere omfattende sociale test. Det endte med, at man satte grænsen ved tre spørgsmål for hvert træk, fordi det blev vurderet, at det samtidig gav mulighed for at indfange en vis mængde variation, uden at spørgeskemaerne blev for omfattende.

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 eftermålingerne 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-scorerne var normalfordelte med tre forskellige test-metoder⁴, og der blev foretaget en visuel vurdering⁵.

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 Der henvises til LøkkeFonden for en beskrivelse af det teoretiske grundlag for, og operationaliseringen af, de syv karaktertræk.

⁴ Skewness-Kurtosis test, Shapiro-Wilk's test samt Shapiro-Francia's test

⁵ I den sammenhæng blev der anvendt QQ-plots, hvor der ses efter visuelle afvigelser fra den estimerede normalfordeling.

Som en ekstra sikkerhed blev alle ændringer mellem før og efter ligeledes testet med en Wilcoxon Signed Rank Test, som er mere velegnet, hvis test forudsætningen om normalfordeling ikke er overholdt. 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 interventioner⁶.

Endelig blev der testet for sammenhæng mellem udviklingen i karaktertræk og fysisk aktivitet set i forhold til den faglige udvikling som målt i de benyttede test. Dette blev gjort ved en visuel vurdering af værdierne plottet mod hinanden samt via en korrelationsanalyse. Der blev beregnet en Pearson korrelation⁷ og denne blev testet i forhold til at være signifikant forskellig fra værdien 0, som indikerer ingen sammenhæng mellem data, altså mellem karaktertræk, og faglig udvikling. Populært sagt fortæller korrelationen noget om, hvor meget data ændrer sig "i takt". Altså hvor meget udviklingerne følges ad. Værdien falder mellem -1 og 1, hvor -1 indikerer en perfekt omvendt sammenhæng (den ene stiger og den anden falder), 0 indikerer ingen sammenhæng og 1 indikerer perfekt sammenhæng (begge stiger eller falder). Konkret blev der analyseret på differenserne mellem slutscore og startscore i forhold til faglig udvikling og karaktertræk.

Det måtte på baggrund af korrelationsanalysen konkluderes, at der ikke kunne findes umiddelbar evidens for en stor grad af sammenhæng. Hverken når man så på de enkelte karaktertræk eller på den samlede udvikling (summen af scorene på de enkelte karaktertræk). Det er dog et åbent spørgsmål, om dette betyder, at den forventede sammenhæng ikke er til stede, eller at måleredskaberne, der blev benyttet ikke i tilstrækkelig grad har kunnet indfange de personlige forandringer, som deltagerne måtte have oplevet ved deltagelse i DrengAkademiet.

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

⁷ Data indikerede visuelt ikke non-lineære sammenhænge og derfor blev en standard pearson korrelation valgt som sammenhængsmål.

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, ifølge Hattie, til cirka det udbytte eleverne gennemsnitligt får af et læringsår, forstået som effekten af ét års almindelig undervisning i skolen, jf. nationale tests i en række lande (Hattie, 2009; 2012). I forhold til Hatties undersøgelse er det værd at bemærke, at den i de senere år har imødegået en del kritik i nogle dele af forskningsverdenen, blandt andet i forhold til betydningen af den metode Hattie bruger til at syntetisere viden på tværs af studier i forhold til "pejlemærket" på de 0,4 (Se fx Cheung & Slavin, 2015).

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 særdeles 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, som man med rette kan kritisere brugbarheden af bl.a. i en intervention med en meget specifik målgruppe, som det er tilfældet med DrengAkademiet.

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

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 2015.

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. Det kan betyde, at noget af den observerede fremgang skyldes, at eleverne bliver bedre til at tage tests og ikke en dybere eller mere almen faglig progression.

For det andet er det vigtigt at understrege, at DrengAkademiets 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-grupes 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 DrengAkademiet. For at undersøge dette kunne man have gennemført et lodtrækningsforsøg, hvor deltagerne blev tilfældigt fordelt mellem DrengAkademiet og en anden sammenlignelig indsats samt eventuelt ingen indsats.

Det er dog værd at bemærke, at det nu er femte gang DrengAkademiet er blevet gennemført med entydigt positive resultater. Det øger alt andet lige sandsynligheden for, at indsatsen ville have en positiv effekt – også i en anden kontekst.

Resultater fra DrengAkademiet 2016

I dette afsnit vil resultaterne fra dette års DrengAkademi 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 i forhold til de nævnte karaktertræk og fysisk aktivitet. Endvidere vil eventuelle ændringer på de målte karaktertræk blive sammenholdt med udviklingen i de faglige test.

Hovedresultater summeret

- Der er markant fremgang på alle tre faglige mål, størst inden for læsning og matematik (matematiske færdigheder – tal og algebra⁹).

⁹ Fremadrettet bruges for nemheds skyld blot betegnelsen "matematik", men der refereres i alle tilfælde til begrebet matematiske færdigheder – tal og algebra.

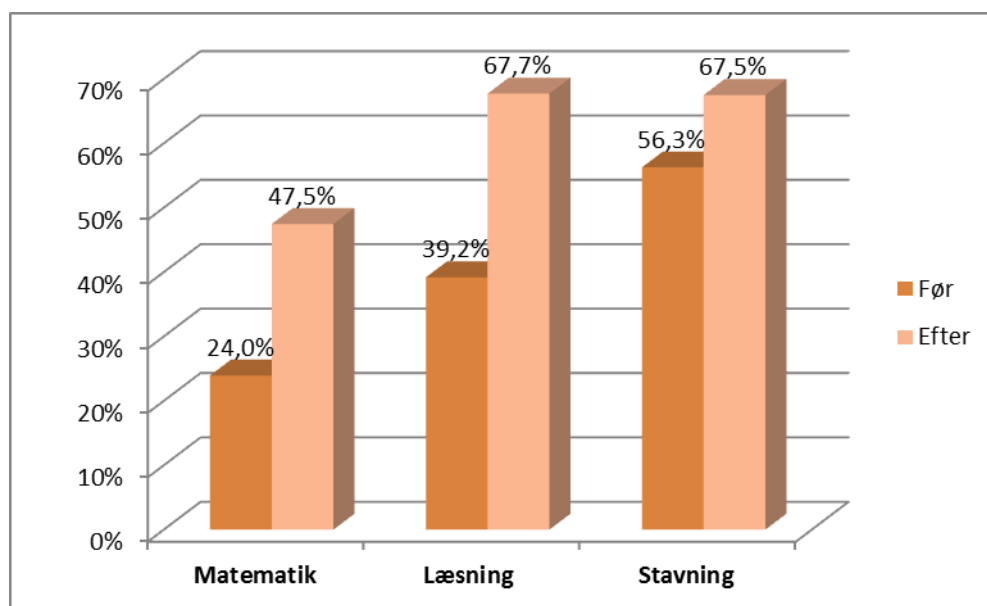
- Fremgangen i deltagernes samlede faglige niveau i læsning og stavning er størst for bund- til middelniveauet. I matematik ses den største fremgang for middel- og topniveauet.
- Store effektstørrelser, målt ved cohens d , inden for læsning og matematik og moderat inden for stavning.
- Alle faglige resultater 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 de syv målte karaktertræk. Fremgangen ligger på mellem 6 % og 33 % i forhold til start-niveauet for deltagerne.
- Der lader ikke til at være en statistisk sammenhæng mellem udviklingen i forhold til karaktertræk og den faglige udvikling.
- Der er signifikante fremgange i forhold til to ud af de benyttede tre mål for fysisk aktivitet: 1) selv-vurdering af egen fysiske formåen samt 2) tro og indstilling til, hvad man selv kan klare. Der blev ikke fundet en signifikant fremgang i forhold til drengenes kobling mellem fysisk aktivitet og følelse af trivsel/sundhed/læring.

Fremgang i de faglige test

Det centrale mål for DrengAkademiets succes er, at eleverne opnår faglig fremgang gennem forløbet. Konklusionen er, at der umiddelbart er tale om en markant fremgang på alle tre fagområder. Figuren nedenfor viser gennemsnitlig andel rigtige svar i de tre faglige test før og efter campen.

På dette års camp ses fremgang på cirka 23,5 procentpoint i matematik, 11,3 procentpoint i stavning, og hele 28,5 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 = 95

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

Tabel 2: Den faglige udvikling i procentpoint de fire seneste år

	Matematik før-efter	Stavning før-efter	Læsning før-efter
2016	+ 24	+ 11	+ 29
2015	+ 23	+ 13	+ 31
2014	+19	+10	+27
2013	+24	+15	+28

N(2013)=94, N(2014)=85, N(2015)=98, N(2016)=95

Overordnet set er fremgangen markant og sammenholdt med de fire sidste års resultater tyder meget på, at man (som i 2014 og 2015) har opnået en stabil fremgang på kun to ugers camp i 2016 versus de tre uger 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 samt dette års camp. Dette mønster bliver ikke mindre interessant, når man tager i betragtning, at DrengAkademiet 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 $p < 0,01$ niveau¹⁰, 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 percentil-analyse, 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. Percentil-analysen 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 deltagernes 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).

¹⁰ Denne testsandsynlighed er konsistent på tværs af t-test for matchede par og Wilcoxon's Signed Rank Test

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 læsetest, under 6,0 % rigtige svar, hvorimod de 5 % der klarede sig bedst fik over 84,0 %, som det fremgår af tabellen nedenfor:

Tabel 3: Udviklingen i deltagernes faglige niveauer fordelt på percentiler

Percentiler	StavFør	StavMidt	StavEfter	LæsFør	LæsMidt	LæsEfter	MatFør	MatMidt	MatEfter
5	15,7 %	20,0 %	30,0 %	6,0 %	16,0 %	36,0 %	2,1 %	8,3 %	14,6 %
10	24,3 %	30,0 %	35,7 %	10,0 %	20,0 %	44,0 %	4,2 %	12,5 %	18,8 %
25	41,4 %	47,1 %	52,9 %	18,0 %	44,0 %	58,0 %	10,4 %	16,7 %	33,3 %
50	60,0 %	62,9 %	70,0 %	38,0 %	62,0 %	70,0 %	20,8 %	22,9 %	45,8 %
75	72,9 %	74,3 %	82,9 %	58,0 %	78,0 %	82,0 %	33,3 %	39,6 %	60,4 %
90	81,4 %	82,9 %	90,0 %	72,0 %	86,0 %	88,0 %	45,8 %	54,2 %	75,0 %
95	84,3 %	88,6 %	92,9 %	84,0 %	88,0 %	92,0 %	60,4 %	68,8 %	89,6 %

N=95

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 60,4 % rigtige svar i den første test mod 89,6 % i den afsluttende.

I stavning derimod er der kun mindre forskel på topniveauet før og efter; de 5 %, der klarede sig bedst fik 84,3 % rigtige i før-testen – et tal, der steg til 92,9 % 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 procent-point 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) fra før til eftertest fordelt på percentiler

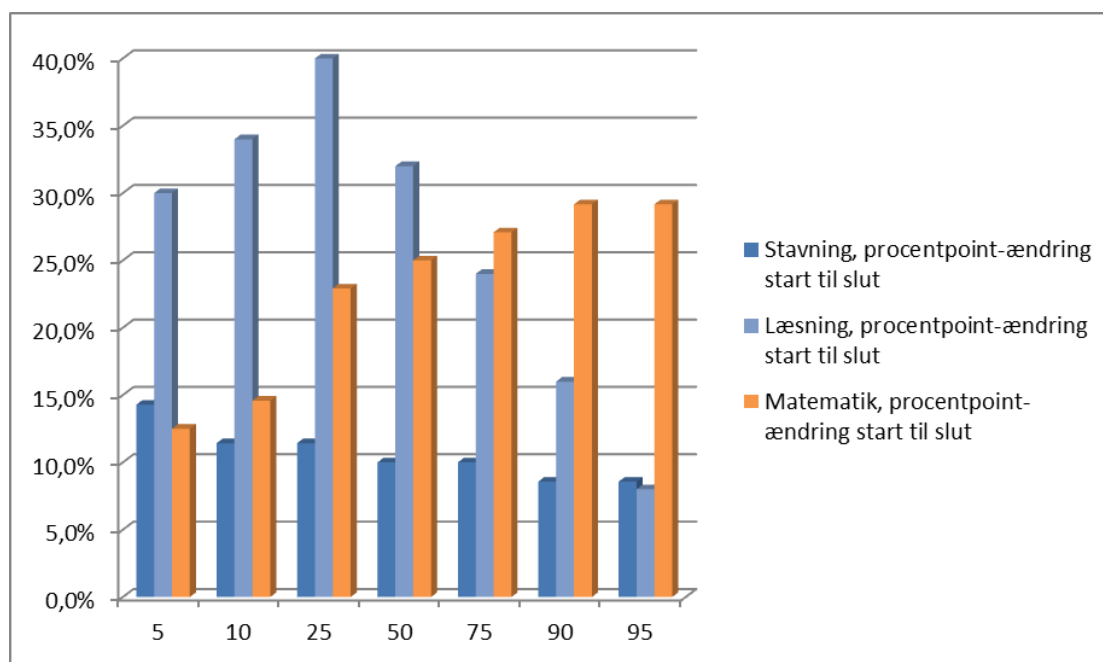
Percentil	Stavning ændring i PP	Læsning ændring i PP	Matematik ændring i PP
5	14,3	30,0	12,5
10	11,4	34,0	14,6
25	11,4	40,0	22,9
50	10,0	32,0	25,0
75	10,0	24,0	27,1
90	8,6	16,0	29,2
95	8,6	8,0	29,2

N=95

Tabellen ovenfor viser igen, at der er sket faglige forbedringer inden for samtlige niveauer. Fx ligger barren for de bedste 5 % 29,2 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 middelniveauet er blevet markant hævet for gruppen, når vi ser på læsning og stavning. Udviklingen i matematik er en smule anderledes, idet den generelle fremgang især sker ved et løft af midt- og topniveauerne. Man kan dog sige, at de bedste i matematik også havde et større rum for udvikling end de bedste i læsning og stavning; de 5 % bedste i læsning og stavning scorede mindst omkring 84 % rigtige, mens de i matematik scorede omkring 60 % rigtige i før-testen, jf. tabel 3 ovenfor.

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



N=95

Effektstørrelse og læringsår

Det er nu blevet statistisk sandsynliggjort og dokumenteret, at deltagerne på DrengAkademiet 2016 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 deltagernes faglige præstationer¹¹.

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.

Der er derfor beregnet effektstørrelser i form af Cohens d ¹². Beregningerne er foretaget i statistikprogrammet STATA og fremgår af tabellen nedenfor:

Tabel 5: Beregning af effektstørrelser

	Stavning	Læsning	Matematik
Effektstørrelse (Cohens d)	0,556	1,352	1,213

N=95

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 DrengAkademiet 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, 2012):

¹¹ Med de forbehold, der er nævnt i metodeafsnittet.

¹² Cohens d uddybes i metodeafsnittet.

Tabel 6: Effekten omregnet til læringsår*

	Stavning	Læsning	Matematik
Læringsår	1,39	3,38	3,03

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

N=95

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 DrengAkademiet har haft en stor effekt på deltagernes faglige resultater, der i den ovenstående fortolkningsramme, forsigtigt, kan antages at svare til mellem 1,39 og 3,38 læringsår. Eller med andre ord: 1,39 til 3,38 års gennemsnitlig skolegang, fortolket med de nævnte forbehold¹³.

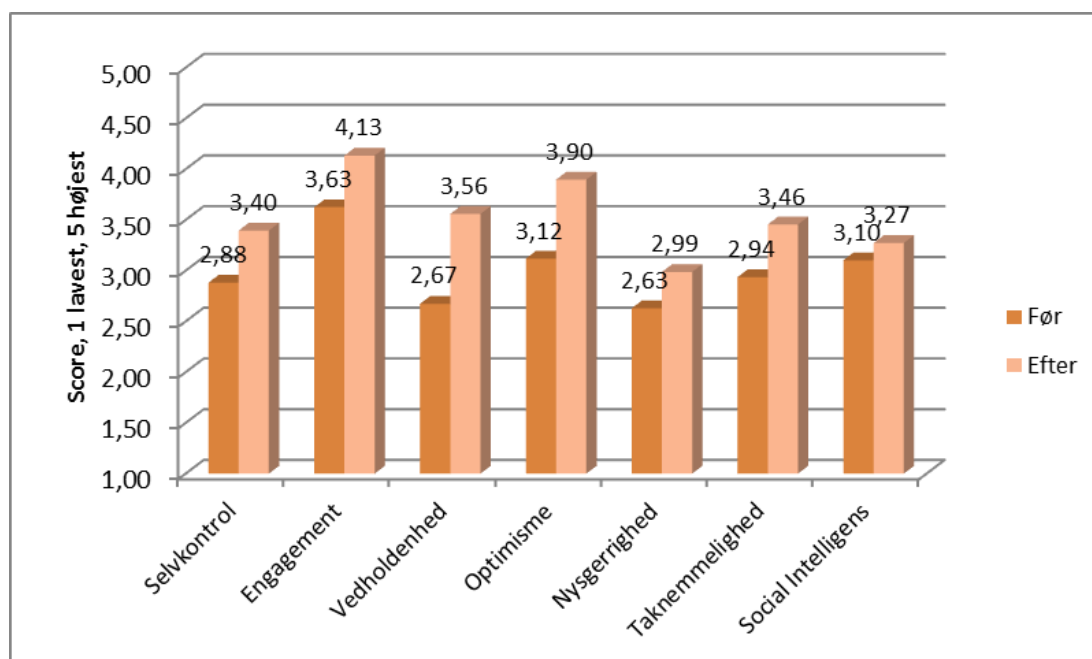
De syv karaktertræk

I DrengAkademiet arbejdes der målrettet med at forbedre deltagerne i forhold til syv generelle karaktertræk. Disse er: selvkontrol, engagement, vedholdenhed, optimisme, nysgerrighed, taknemmelighed og social intelligens. Alle disse karaktertræk måles på en samlet skala, der går fra 1 (laveste niveau) til 5 (højeste niveau). Det er vigtigt at påpege, at den enkelte deltagers score i sig selv ingen substantiel fortolkning har. Skalaen er udelukkende relationel, således at vi kan sige, at værdien 4 ift. selvkontrol angiver mere selvkontrol end værdien 3. Men *ikke* at 4 svarer til 'god selvkontrol' og 3 svarer til 'ringe selvkontrol'.

Figuren nedenfor viser udviklingen fra start til slut i de syv karaktertræk:

¹³ Her er det værd at bemærke, at effektstørrelserne i form af Cohens d er sammenlignelige med DrengAkademiet 2015 om end en anelse lavere. Det skyldes angiveligt 1) en meget lille fald 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 standardafvigelseerne forskellige. Det er vigtigt, at det bemærkes, at effektstørrelserne afhænger af både forskellen mellem start og slutmåling og spredningen i data. Det vil kræve en særlig analyse, der er uden for denne rapport's formål, at se på hvorledes spredningerne i data har ændret sig henover årene.

Figur 3. Udviklingen i de syv karaktertræk, der arbejdes med på DrengAkademiet



N=87/88. For nysgerrighed, taknemmelighed og social intelligens var der 88 gyldige besvarelser, for de resterende 87.

Der ses positive fremgange på alle syv karaktertræk om end i forskellig grad. Forskellene er testet med t-test for matchede par og Wilcoxons Signed Rank Test. I begge tilfælde er konklusionen, at fremgangene er statistisk signifikante. Social intelligens, hvor der ses den mindste fremgang, er signifikant på $p < 0,05$ -niveau, mens de seks resterende træk er signifikante på $p < 0,01$ -niveau.

Tabellen nedenfor illustrerer ændringerne i både absolutte størrelser (skalapoint) og relative størrelser (procent). Der ses størst fremgang for vedholdenhed og optimisme med relative fremgange på henholdsvis 33 % og 25 %. Mindst udvikling ses for Social intelligens med en relativ fremgang på 6 %.

Tabel 7. Absolut og relativ ændring på de syv karaktertræk

Karaktertræk	Absolut ændring (skalapoint)	Relativ ændring*, i pct.
Selvkontrol	0,51	18 %
Engagement	0,51	14 %
Vedholdenhed	0,89	33 %
Optimisme	0,78	25 %
Nysgerrighed	0,36	14 %
Taknemmelighed	0,52	18 %
Social intelligens	0,17	6 %

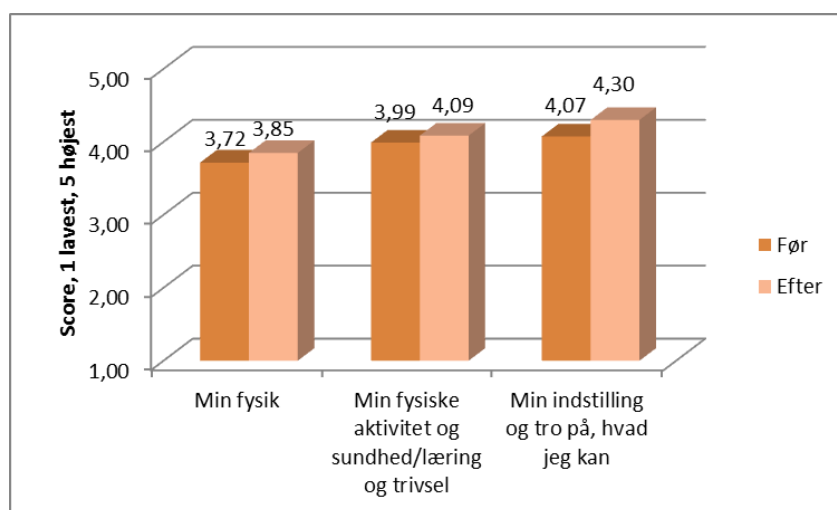
N=87/88. *Den relative ændring er et udtryk for den absolutte ændring set i forhold til udgangspunktet (start-niveauet).

Fysisk aktivitet

På DrengAkademiet er fysisk aktivitet en central del af indsatsen. Deltagerne er løbende blevet målt på deres oplevelse af koblingen mellem fysisk aktivitet og sundhed/trivsel/læring, vurdering af deres egen fysik, samt deres tro på, hvad de kan klare rent fysisk. Som nævnt ovenfor, blev hver af disse områder af-dækket med tre spørgsmål, der – som det var tilfældet med karaktertrækkene – blev omregnet til en samlet score.

Figuren nedenfor præsenterer udviklingen fra start til slut.

Figur 4. Udviklingen i forhold til fysisk aktivitet



N = 81/82. I forhold til 'min fysik' har 81 deltagere svaret på begge målinger, mens det var 82 for de resterende

Af tabel 8 nedenfor ses det, at der ikke er tale om de store fremgange. De ligger i intervallet 2 % - 6 %.

Tabel 8. Absolut og relativ ændring i forhold til fysisk aktivitet

	Absolut ændring skalapoint	Relativ ændring* i pct.
Egen fysiske formåen	0,13	4 %
Indstilling og tro på egne fysiske evner	0,10	2 %
Fysisk aktivitet og sundhed/trivsel/læring	0,23	6 %

N=81/82. *Den relative ændring er et udtryk for den absolutte ændring set i forhold til udgangspunktet (start-niveauet).

På samme vis som tidligere i rapporten har vi testet om fremgangene, der fremgår af figur 4, er statistisk signifikante¹⁴. Konklusionen er, at to ud af tre af dimensionerne af fysisk aktivitet er forbedret i signifikant grad. Det gælder 1) deltagernes egen indstilling og tro på egne fysiske evner ($p < 0,01$) samt 2) deltagernes vurderinger af egen fysiske form – dog er der tale om en lav-signifikant testsandsynligheden på $p < 0,1$, hvor der er mindre end 10 % sandsynlighed for, at der blot er tale om tilfældig variation i data. Dette er dog

¹⁴Igen har vi testet med t-test for matchede par samt Wilcoxon Signed Rank Test

mere end de gængse 5 %, man ofte opererer med, og derfor konkluderes det, at forbedringen er til stede, men er usikker i statistisk forstand. I forhold til den tredje dimension: sammenhængen mellem fysisk aktivitet og læring/trivsel/sundhed er der ikke sket en statistisk signifikant forbedring for deltagerne ($p < 0,194$).

Sammenhæng mellem udvikling på de syv karaktertræk og den faglige udvikling

Det er nu interessant at undersøge, om der kan sandsynliggøres en sammenhæng mellem udviklingen i de syv forskellige karaktertræk 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 syv karaktertræk og et samlet indeks for udviklingen på de syv karaktertræk⁴⁵. 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 udvikling af karaktertræk og deres faglige udvikling, bør korrelationskoefficienten være signifikant større end 0. I tabellen angives kun korrelationer, der er statistisk signifikant forskellig fra 0. Tomme celler angiver altså ingen statistisk signifikant sammenhæng.

Ud fra denne analyse, er det svært at se empirisk evidens for en positiv sammenhæng mellem udviklingen i karaktertræk og den faglige udvikling. Samlet set vurderes der ikke at være grundlag for at mene, at der er en sammenhæng mellem udviklingerne i karaktertræk og faglige test. Men dermed ikke sagt, at karaktertrækkene ikke har betydning for evnen til faglig indlæring generelt set; blot vurderes det, at udviklingen henover campens to uger ikke umiddelbart kan relateres til den faglige udvikling.

Bemærk endvidere, at karaktertræk er svære størrelser at indfange empirisk. Især med det relativt begrænsede måleredskab, der blev benyttet i dette tilfælde. Således kan den grundlæggende sammenhæng sagtens være til stede, men de benyttede målinstrumenter er ikke gode nok til at indfange denne sammenhæng i dette tilfælde.

Tabel 9: Resultat af korrelationsanalyse mellem udviklingen i trivselsindeks og de faglige test. Kun signifikante korrelationer afrapporteres.

	Læsning	Stavning	Matematik
Selvkontrol			
Engagement			
Vedholdenhed			0,203*
Optimisme			0,271**
Nysgerrighed			
Taknemmelighed	- 0,287**		
Social intelligens			
Samlet indeks for udvikling på de syv karaktertræk			

*Signifikant $p < 0,1$ niveau **Signifikant på $p < 0,05$ niveau..

N=87/88. Analysen kræver for hver deltager gyldige før og eftermålinger på karaktertræk såvel som faglige test.

⁴⁵ Dette indeks konstrueres som et additivt indeks, hvor værdierne på hver af de syv karaktertræk summeres til én samlet score

Det skal dog samtidig understreges, at drengene har haft en væsentlig positiv fremgang på de anvendte mål for karaktertræk sideløbende med den store markante faglige fremgang. Der kan blot ikke påvises en direkte statistik valid sammenhæng mellem graden og retningen af de to fremgange.

Opsamling og perspektivering

Samlet set er det konklusionen på den foregående analyse, at DrengesAkademiet 2016, på linje med de tidligere år, har leveret væsentlige positive fremgange i forhold til det faglige niveau i både læsning og stavning såvel som i matematiske færdigheder (tal og algebra). Samtidig har drengene oplevet positive fremgange i forhold til de syv overordnede karaktertræk, som DrengesAkademiet arbejder med. Da dette er det femte akademi i rækken med konsistent gode resultater, er det en rimelig konklusion, at der i DrengesAkademiet synes at være tale om en særdeles effektiv indsats. Det vil dog samtidig være fordelagtigt med yderligere studier af hvor effekten stammer fra, og i hvilket omfang den kan forventes at holde ved fremadrettet og eventuelt i andre, måske meget forskellige, sammenhænge.

DrengesAkademiet er en indsats, der lægger vægt på rammer, struktur, selvevaluering, feedback, engagement, motivation og synlig læring. Om end evalueringsdesignet ikke involverer en egentlig kontrolgruppe, og det dermed er vanskeligt at undersøge, hvorvidt DrengesAkademiet har en større effekt end andre tilsvarende indsatser, peger resultaterne på, at det tilsyneladende er lykkedes at flytte deltagernes faglige niveau markant på meget kort tid. Det ville stadig være fordelagtigt at gennemføre et egentligt kontrolgruppeforsøg eller et match- studie, med henblik på at skabe sikkerhed om, hvor stor en del af den faglige progression, der kan tilskrives henholdsvis DrengesAkademiet som specifik indsats og/eller den træningseffekt, der sker i og med, at drengene bliver øvede i de faglige test. Dertil er det naturligvis forventeligt, at ekstraordinære indsatser, hvor der tilføres væsentlige læringsressourcer til børns hverdag, ofte vil have en eller anden effekt alene på grund af de ekstra ressourcer.

I forhold til, hvorvidt den faglige fremgang holder ved over tid – også efter deltagerne har forladt DrengesAkademiet, kunne man med fordel undersøge dette fremadrettet. Der bør være gode muligheder for at benytte de omfattende danske registre til en sådan analyse - ikke mindst da drengene fra de første runder af DrengesAkademiet efterhånden er nået et godt stykke videre i deres livs- og evt. uddannelsesforløb. Det er dog værd at bemærke, at man bør forvente den såkaldte "boost-effekt", der blandt andet dækker over, at korte interventioner tit involverer, at deltagerne tages ud af deres normale dagligdagsmiljø og at korte interventioner ikke nødvendigvis påvirker den dybere læringsevne, og at effekten heraf derfor kan ses at aftage over tid (Paris et al., 2006). Det er derfor vigtigt at vedligeholde indsatsens effekter gennem systematisk opfølgning, hvis man vil øge sandsynligheden for at en størst mulig del af effekten fastholdes i deltagernes videre liv og 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.

Yderligere kunne man foretage en metaanalyse af publicerede studier, der undersøger effekten af faglige interventioner på elever, der er sammenlignelige med målgruppen. Dette med henblik på at identificere den gennemsnitlige effektstørrelse og sammenligne denne med DrengesAkademiet. Det er i den forbindelse værd at bemærke, at der i dansk kontekst er udgivet et mindre litteraturstudie om intensive læringsforløb i og udenfor Danmark (Dyssegaard et al. 2015), som peger på, at der stadig findes meget lidt empiri fra Danmark og Norden at trække på. Dette er i særdeleshed tilfældet i forhold til de korte og intensive læringsforløb, som DrengesAkademiet tilhører. På den baggrund kunne man med fordel anbefale, at

der blev gennemført større og flere danske studier inden for feltet. Således vil man kunne opnå viden generet i den kontekst, som DrengAkademiet bliver implementeret i.

Afsluttende bør det bemærkes, at resultaterne ikke nødvendigvis bør tolkes som, at DrengAkademiet alene øger deltagernes faglige niveau. Der er ligeledes tale om, at DrengAkademiet aktiverer allerede eksisterende kompetencer, der anses for vigtige i forbindelse den videre læring og fremmer deltagernes tiltro til egen læringsevne og til, at de kan bruge det, de har lært. Disse faktorer bør anses for vigtige for deltagernes evne til at få mest muligt ud af deres fremtidige uddannelsesforløb 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.
- Cheung, A., & Slavin, R.E. (2015). *How Methodological Features Affect Effect Sizes in Education*. Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.
- Dyssegaard, C.B., Bondebjerg, A., Sommersel, H.B. & Vestergaard S. (2015). *Litteraturstudie om intensive læringsforløb*, Dansk Clearinghouse for Uddannelsesforskning, DPU, Aarhus Universitet.
- 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.
- LøkkeFonden (2015). Forskningsrapport. DrengAkademiet 2015. Årets resultater. Lokaliseret d. 07.09.2016 på <http://www.loekkefonden.dk/WP/wp-content/uploads/2016/01/DA-rapport-2015-leveret-1.pdf>
- 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_tolaererforsoeget_og_forsoegsprogrammet_om_modersmaalsbaseret_undervisning.pdf