

## **Præmier og præstationer: skolelederes meningsskabelse<sup>1</sup>**

*Per Nikolaj Bukh, Karina Skovvang Christensen, Morten Lund Poulsen & Anne Kirstine Svanholt \*)*

### **Resumé**

Med henblik på at reducere andelen af lavt præsterende elever ved afgangseksamenerne i folkeskolen vedtog Folketinget i 2016 den såkaldte skolepulje. Formålet var at give udvalgte skoler et økonomisk incitament til at reducere andelen af lavtpresterende elever. Denne artikel fokuserer på, hvordan skoleledere opfattede skolepuljen og de initiativer, der blev taget som følge heraf. Vi fokuserer især på, om der forekom negative tilbageløbseffekter (washback), dvs. uønskede påvirkninger af undervisning eller elever. Artiklen er baseret på semistrukturerede interviews gennemført med skoleledere. I analysen anvendes meningsskabelse (sensemaking) som en teoretisk linse til at undersøge, hvordan ledere skaber nye meninger, når nye initiativer igangsættes som konsekvens af skolepuljen. Resultaterne indikerer, at puljens resultatbaserede finansiering påvirkede undervisningspraksis på samme måde som tilbageløbseffekterne fra standardiserede elevtests. Dog blev skolepuljen og dens virkninger generelt oplevet som positive. Undersøgelsen viser desuden, at skoleledere spiller en vigtig rolle for, hvordan skoler reagerer på resultatbaseret finansiering.

### **Emneord**

Folkeskolen, resultatbaseret styring, resultatbudgettering, sensemaking, skolepuljen, teaching-to-the-test, tilbageløbseffekter, undervisning, økonomistyring.

\*) Per Nikolaj Bukh er cand.oecon. og ph.d. samt professor i økonomistyring ved Aalborg Universitet. Karina Skovvang Christensen er cand.oecon. og ph.d. samt lektor ved Institut for Økonomi, Aarhus Universitet. Morten Lund Poulsen er cand.mag. i filosofi og ph.d.-studerende ved Institut for Økonomi, Aarhus Universitet. Anne Kirstine Svanholt er cand.soc. i udviklingsstudier fra Aalborg Universitet og enhedsleder i socialforvaltningen, Københavns Kommune.

## A. Indledning

Sammenlignet med f.eks. de fleste angelsaksiske skolesystemer har tests og eksamener i Danmark kun i begrænset omfang været anvendt til at holde lærere og skoler ansvarlige for undervisningen, og der har slet ikke været tradition for at anvende økonomiske incitamenter. Niende klasses afgangseksamen og PISA-undersøgelserne<sup>2</sup> var således de eneste systematiske tests af folkeskoleelever indtil 2010, hvor det obligatoriske nationale testprogram blev indført. De danske nationale prøver og afgangseksamener betragtes dog som *low-stakes*<sup>3</sup> tests i relation til skoler, lærere og elever, fordi testene hverken bruges til belønning eller sanktion, selvom eksamensresultaterne kan have betydning for, hvilke skoler der bliver efterspurgt (Andersen & Nielsen 2020). For eleverne kan der være mere på spil, da karaktererne har betydning for muligheden for valg af ungdomsuddannelse. Sammenlignet med ungdomsuddannelseskaraktererne, der har væsentlig betydning for de fremtidige uddannelsesmuligheder, må karakterne i folkeskolen dog anses for også at være *low-stakes* for eleverne.

I 2014 trådte den seneste folkeskolereform i kraft. Den fastlagde tre nationale mål<sup>4</sup>, hvoraf ét af dem blev operationaliseret ved, "at andelen af elever med dårlige resultater i de nationale tests for læsning og matematik skal reduceres år for år." (Undervisningsministeriet 2013, s. 23; 2014, s. 18). Allerede i 2016 konkluderede regeringen, at flere elever stadig stod til at forlade folkeskolen uden tilfredsstillende færdigheder i dansk og matematik. Derfor besluttede Folketinget i 2017 at tilføre 500 mio. kr. over 3 år til 200 udvalgte skoler med en særlig stor andel af såkaldt svage elever. Denne såkaldte skolepulje (Børne- og Undervisningsministeriet 2017) indebærer en resultatbaseret finansieringsmodel (Bukh & Christensen 2020b), hvor ressourcetildelingen var betinget af en reduktion af andelen af svage elever i dansk og matematik. I praksis indebærer det, at skoler fik en bonus<sup>5</sup> svarende til 5%-8% af det årlige budget, hvis de formåede at opnå det fastlagte måltal for reduktion af andelen af svage elever, operationaliseret som det relative antal elever, der ikke opnåede et gennemsnit på mindst 4 i de bundne prøver fra folkeskolens afgangseksamen i dansk og matematik.

### 1. Organisatoriske incitamenter og konsekvenser

Skolepuljen indebærer, at der anvendes en resultatbaseret økonomistyringsmodel, baseret på kollektive økonomiske incitamenter (Andersen *et al.* 2017). Denne form for økonomistyring omfatter både aktivitets- og outputbaserede budgetmodeller, som man kender dem fra finansieringen af de videregående uddannelser herhjemme og i udlandet (Herbst 2008; Dougherty *et al.* 2016), og effektbaserede tildelingsmodeller, hvor tildelingskriteriet er formuleret som mål for effekt (Bukh & Christensen 2020a; 2020b). Mens der i nogle lande er erfaringer med individuelle økonomiske incitamenter i form af resultatløns til lærere (Lavy

2007; Yuan *et al.* 2013), er det for grundskoler mindre udbredt at anvende organisatoriske incitamenter ved at gøre bevillingerne afhængige af elevers præstationer. Vi har således kun kunnet identificeret en enkelt artikel (Al-Samarrai *et al.* 2018), der beskriver resultater, hvor finansieringen er direkte knyttet til elevernes testresultater i grundskolen.

Meget forskning har fokuseret på de såkaldte tilbageløbseffekter (*washback*<sup>o</sup>), der indebærer, at en test påvirker undervisning og læring (Alderson & Wall 1993, s. 120; Cheng 1999; Holm & Kousholt 2019). Sådanne effekter kan opfattes og analyseres som positive, når de resulterer i "god" undervisningspraksis (Holm & Kousholt 2019, s. 921) eller negative, hvis de påvirker undervisnings- og læringsresultaterne negativt. Forekomsten af tilbageløbseffekter afhænger af mange faktorer (Alderson & Wall 1993), men de fleste studier har bekræftet, at *high-stakes* tests påvirker undervisning og læring (Akpinar & Cakildere 2013; Alderson & Wall 1993; Cheng 1999).

Da anvendelsen af en resultatbaseret bevilling øger budgetrisikoen (Bukh & Christensen 2020b; Smith 2007, s. 104ff), hvis beløbet udgør en betragtelig andel af en skoles budget, forventer vi, at skolepuljen relativt set ændrer skolernes syn på eksamenerne fra *low-stakes* til *high-stakes*. Vi forventer, at når ressourcetildelingen knyttes sammen med eksamensresultaterne, kommer der hermed mere på spil for skolerne. Denne artikel undersøger derfor spørgsmålet om, hvorvidt skolepuljen har konsekvenser svarende til dem, der kan observeres, når tests *stakes* generelt øges.

## 2. Skolelederens rolle

Når politiske initiativer introduceres, har lederne på de enkelte uddannelsesinstitutioner en afgørende betydning for, hvordan tiltagene implementeres i praksis (Gunnulfsen & Roe 2018; Paletta *et al.* 2019). Tidligere forskning har vist, at skoleledelser har en vigtig rolle som mediatorer (Koyama 2014; Lambersky 2016; Shirrel 2016), når eksterne krav til øget ansvarlighed skal integreres med den lokale praksis (Lowenhaupt *et al.* 2016). Dette er fordi lokale skoleledere også er nødt til at forholde sig til andre forhold, der ligeledes har betydning (Rutledge *et al.* 2010; Spillane *et al.* 2002). Undersøgelser inden for mange fagområder (f.eks. Bukh *et al.* 2020a, 2020b; Spillane *et al.* 2002) har vist, at sådanne implementeringer indebærer fortolkning, hvor mellemlederne analyserer nye initiativer og deres anvendelighed for at skabe ændringer. Jongbloed & Vossensteyn (2016, s. 591) argumenterer for, at dette ligeledes gør sig gældende for resultatbaseret finansiering. Derfor undersøger vi, hvordan skoleledelsens opfattelse af den resultatbaserede finansiering har betydning for tilbageløbseffekterne.

### 3. Artikkens formål og fokus

Artiklen præsenterer en kvalitativ undersøgelse af 4 danske folkeskoler, der var udvalgt til at deltage i skolepuljen. Vi interviewede skoleledelserne to gange baseret på semistrukturerede interviewguider for at forstå ledelsens meningsgørelse i relation til skolepuljen og inspirationsprogrammet samt eventuelle ændringer i undervisningspraksis på skolerne.

Mere specifikt sigter denne artikel på at besvare følgende to forskningsspørgsmål:

- Hvordan opfatter skolelederne den resultatbaserede ressourcetildeling baseret på eksamensresultater, og hvordan opfatter de ændringerne i *stakes*, når resultatbaseret ressourcetildeling indføres?
- Hvis nogen, hvad er de positive og negative tilbageløbs effekter af resultatbaseret ressourcetildeling baseret på afgangseksamensresultater?

I artiklen anvender vi sensemaking<sup>7</sup> (Weick 1995), eller meningsgørelse på dansk, som en teoretisk linse (Lukka & Vinnari 2014) til at undersøge, hvordan ledere skaber mening, når nye initiativer igangsættes som konsekvens af skolepuljen. Dermed fokuserer vi især på, hvordan aktørerne handler og konstruerer deres omgivelser samt tillægger de nye initiativer mening i handlingerne.

Ved studier af tilbageløbs effekter er der ofte fokus på de potentielt negative konsekvenser i form af blandt andet *teaching-to-the-test* (Madau 1988; Popham 2001), indsnævring af pensum (Amrein-Beardsley 2009; David 2011) og fokusering på de marginale elever (Amrein-Beardsley 2009; Burgess *et al.* 2005). Vi undersøger derfor specifikt, om skolepuljen har tilsvarende konsekvenser.

### 4. Artikkens resultater og bidrag

Tidligere forskning har vist, at der kan forventes tilbageløbs effekter fra “tests and examinations that are used regularly within the curriculum and which are perceived to have educational consequences” (Alderson & Wall 1993, s. 122). Ud over de tilsigtede forbedringer i elevers læring og eksamensresultater har megen forskning dokumenteret, at sådanne tests også kan have utilsigtede negative konsekvenser, især når tests med *high-stakes* anvendes (f.eks. Kelley & Protsik 1999; Lee & Medina 2018). Denne artikel er baseret på en unik kontekst, hvor finansieringen i særligt udvalgte folkeskoler blev gjort delvist afhængig af elevernes resultater ved afgangseksamen, som normalt er *low-stakes* eksamener for den enkelte skole. Artiklen bidrager dermed ved at undersøge, om ressourcetildeling baseret på *low-stakes* eksamener har de samme konsekvenser, som ses ved standardiserede *high-stakes* tests.

Tidligere forskning vedrørende videregående uddannelser (Blankenberg & Phillips 2016; Li & Ortagus 2019; Umbricht *et al.* 2017) har undersøgt utilsigtede konsekvenser af resultatbudgettering. Vi supplerer denne litteratur ved at undersøge, om resultatbudgettering også har utilsigtede konsekvenser for grunduddannelser. Desuden undersøger vi både, hvilke initiativer der igangsættes, og hvordan skolelederne skaber mening ud fra disse initiativer. Artiklen bidrager dermed til litteraturen om skoleledere som mellemlidende (Koyama 2014; Lambersky 2016; Shirrel 2016) ved at vise, hvordan grundskoleledere opfatter, fortolker og reagerer på en resultatbaseret ressourcetildeling.

Resultaterne indikerer, at skolepuljen forøgede skolernes fokus på elevernes afgangseksamen. Selvom skolelederne var skeptiske over for skolepuljen og brugen af økonomiske incitamenter, igangsatte de ændringer i undervisningspraksisser for at nå de mål, der ville give den ekstra ressourcetildeling. Flere af de igangsatte initiativer kan tolkes som *teaching-to-the-test*, indsnævring af pensum og fokusering på de marginale elever. Derfor har resultatbaseret budgettering tilbageløbseffekter og utilsigtede konsekvenser. Disse kan dog opleves enten positive eller negative. Generelt viste det sig, at skolelederne betragtede initiativerne som gavnlige for alle elever. Endvidere viser resultaterne også, at skolelederes italesættelse og handlinger spiller en afgørende rolle for, hvordan skolepuljen opfattes, og dermed om de økonomiske incitamenter opleves som et styringsmæssigt pres. Resultaterne vil være relevante, hvis man overvejer at implementere resultat- eller effektbaserede tildelingsmodeller på andre offentlige udgiftsområder. Det er i særdeleshed vigtigt at være opmærksom på, at sådanne tildelingsmodeller ikke kun påvirker de ønskede resultater men også, om der optræder en uhensigtsmæssig strategisk adfærd, og om der er utilsigtede konsekvenser. Derfor kræver det grundige overvejelser af fordele og ulemper, når tildelingsmodeller skaber *high-stakes* situationer.

## **B. Skolepuljen**

I 2016 besluttede den daværende VLAK-regering at “styrke skolernes indsats for de fagligt svageste elever” (Regeringen 2016). Skolepuljen<sup>8</sup> blev derfor etableret som en incitamentsmodel for skoler med et højt antal elever, der ikke opnåede tilfredsstillende faglige færdigheder (Undervisningsministeriet 2018). Regeringen forventede med skolepuljen, at:

... give skolerne en ekstra tilskyndelse til at løfte de fagligt svageste og f.eks. til at udvikle og afprøve nye undervisningsmetoder. Det vil sætte øget fokus på, at alle børn bliver så dygtige, de kan. (Regeringen 2016)

Skoler fra alle regioner blev udvalgt ud fra regionens andel af elever i folkeskolens 9. klasse. Som udgangspunkt blev skolerne med den største andel af svage elever inden for hver region udvalgt til at deltage; dog betinget af, at skolen i gennemsnit havde mindst 12 elever i målgruppen for de tre skoleår 2013/2014, 2014/2015 og 2015/2016. Det økonomiske incitament var konstrueret således, at den enkelte skoles resultat blev sammenlignet med en baseline beregnet som det gennemsnitlige antal elever, der ikke opnåede karakteren 4 i gennemsnit i dansk og matematik<sup>9</sup> i de tre skoleår forud for igangsættelsen af skolepuljen. Hvis skolen kunne reducere antallet af disse elever med henholdsvis 5, 10 og 15 procentpoint for skoleårene 2017/2018, 2018/2019 og 2019/2020, ville skolen modtage den ekstra finansiering. Selvom incitamentsmodellen var knyttet til resultaterne for fagligt svage elever, var definitionen af "svaghed" prospektivt baseret på årgangens fremtidige eksamensresultater. Det betød, at man på skolerne kunne have forventninger om, hvilke elever der ville opnå under 4 gennemsnit, men de svage elever var i princippet først afgrænset efter at eksamensresultaterne dokumenterede, at de var svage. Da gruppen af svage elever således ikke kunne afgrænses på forhånd, var resultatkriteriet knyttet til skolen generelt fremfor de svage elever.

Deltagelse i skolepuljen var frivilligt. Af 127 skoler, der fik tilbudt deltagelse, tilmeldte 104 sig. Baseline for andelen af fagligt svage elever var 28% - 62% (gennemsnit 40%). Størrelsen af den ekstra ressourcetildeling afhang af antallet af dimitterende elever: Under 50 elever gav ca. 1,3 mio. kr., mellem 50 og 100 elever gav 1,4 mio. kr., og mere end 100 elever gav 1,5 mio. kr. Beløbet ville blive udbetalt i september, få måneder efter eksamenen. For en gennemsnitlig skole, der deltog i skolepuljen, udgjorde beløbet ca. 5%-8% af det ordinære budget og kunne finansiere op til tre lærere i et helt år.

Samtidig med incitamentsmodellen igangsatte ministeriet et inspirationsprogram betegnet "Program for elevløft" (Børne- og Undervisningsministeriet 2017b), herefter betegnet inspirationsprogrammet. Det havde til formål at inspirere lærere fra de deltagende skoler til arbejdet med at løfte de fagligt udfordrede elever. Det fokuserede gennem kurser, workshops og rådgivning på at give lærere værktøjer til samarbejde mellem forældre og lærere, kontinuerlig evaluering og feedback, intensive læringsforløb og læring fra elev til elev. Inspirationsprogrammet var gratis for deltagende skoler men krævede aktiv tilmeldelse. Af de 104 skoler deltog omkring 90.

I oktober 2019 besluttede S-regeringen at afslutte skolepuljeprojektet med virkning fra skoleåret 2019/2020, men inspirationsprogrammet blev videreført. Dermed blev der kun udbetalt bonus for to år. Det første år modtog 61 skoler bonus, mens 54 skoler modtog bonus det næste år. Samlet set blev der udbetalt omkring 160 millioner kroner. Projektet har været

meget omdiskuteret, både da det blev igangsat, og da resultaterne blev offentliggjort. Evalueringer (Andersen *et al.* 2019; 2020) har vist, at de deltagende skoler i gennemsnit har opnået en svag reduktion (0-4 procentpoint) i andelen af elever, der ikke opnår mindst 4 i både dansk og matematik.

## C. Analytisk begrebsramme

### 1. Tilbageløbs effekter

I uddannelseslitteraturen (se f.eks. Alderson & Wall 1993; Biggs 1995; Cheng 1999) er der bred enighed om, at tests og eksamener indebærer tilbageløbs effekter, således at undervisning og læring påvirkes. Det kan lyde som en selvfølgelighed, men tilbageløbs effekt er et komplekst fænomen, hvor både testens opbygning, og en række andre forhold vedrørende f.eks. gennemførelsen af testen kan være afgørende for tilbageløbs effekten. Alderson & Wall (1993, s. 121) fastslog, at “[t]ests that have important consequences will have washback,” mens “[t]ests that do not have important consequences will have no washback.” Derudover har flere undersøgelser (jf. Cheng 1999) vist, at *high-stakes* tests påvirker undervisning og læring. Når ressourcetildeling til skolerne knyttes til test- eller eksamensresultater, forventer vi, at det relativt set skaber en *high-stakes* testsituation, fordi resultaterne får betydelige konsekvenser for skolerne. Derfor forventer vi også tilbageløbs effekter fra skolepuljen.

Som udgangspunkt drejer eksamener sig om at måle elevernes resultater og dermed deres udbytte af skolegangen. Men eksamensresultaterne bliver også resultatmål for skolerne, når de bliver en del af en incitamentsmodel. Forskning viser generelt, at øgede *stakes* relateret til testresultater, især i matematik, kan have en positiv effekt på elevernes resultater (Figlio & Ladd 2015). Tests kan dog også have dysfunktionelle konsekvenser, dvs. effekter, der er i modstrid med intentionerne (Siverbo *et al.* 2019). Meget forskning (f.eks. Amrein-Beardsley 2009; Brill *et al.* 2018; Deming *et al.* 2016; De Wolf & Janssens 2007) har undersøgt, hvordan styringstiltag kan skabe et pres på ledere og medarbejdere, så de på trods af gode intentioner agerer strategisk både ved registrering og rapportering af resultater og i daglige handlinger, således at testresultater, antal elever der fuldfører en uddannelse eller andre indikatorer på elevens succes kan fremstå bedst mulige (Edwards & Mindrila 2019, s. 3). Ofte argumenteres der for, at indsatser som *teaching-to-the-test*, indsnævring af pensum og fokusering på marginale elever har potentielt dysfunktionelle konsekvenser (Amrein-Beardsley 2009; Paletta *et al.* 2019). Indsatserne kan dog ikke i sig selv antages at være dysfunktionelle, da dysfunktionaliteten afhænger af, hvordan indsatserne konkret praktiseres, og om de ønskede resultater opnås.

## 2. Effekter af tests med high-stakes

*Teaching-to-the-test* betyder, at undervisningen tilrettelægges og prioriteres ud fra eksamensformen snarere end de læringsmål, eksamen repræsenterer (Popham 2001). Det omfatter eksempelvis “students spend[ing] hours memorizing facts, learning test-taking strategies, bubbling score sheets accurately, eliminating unlikely distractor responses, [and] making educated guesses” (Amrein-Beardsley 2009, s. 3). *Teaching-to-the-test* kan være problematisk, hvis praksissen fortrænger andre læringsrelevante aktiviteter og dermed giver testresultaterne et for stort fokus, men det kan dog også klarlægge relationen mellem læringsaktiviteterne og eksaminationen og dermed give eleverne en bedre oplevelse og forbedre resultaterne. Hvis eleverne bliver mere komfortable med eksamenssituationen, vil de være bedre i stand til at demonstrere opfyldelsen af læringsmålene, end når resultaterne påvirkes af elevernes nervøsitet (Jennings & Bearak 2014, s. 382). Derfor argumenterer Jennings & Bearak (2014) for, at konsekvenserne af *teaching-to-the-test* bør evalueres som et spektrum med flere muligheder fremfor den dikotomi, der oftest anvendes: Enten anvender lærerne *teaching-to-the-test* eller ej; og enten er det godt eller skidt.

Indsnævring af pensum (*'narrowing the curriculum'*) betyder, at indsatsen fokuseres mod specifikke emner, fag eller områder, hvor centrale dele af pensum udelades, fordi de ikke vægtes højt til eksamen (Amrein-Beardsley 2009). Det kan f.eks. være, fordi der normalt ikke spørges ind til disse emner til eksamen, eller fordi det vurderes, at det er for vanskeligt at lære i forhold til den betydning, det har for karaktergivning. En indsnævring af pensum kan også påvirke den relative tid, der bruges på forskellige fag i forhold til den relative betydning, de har for eksamensresultaterne. Det kan eksempelvis være tilfældet, hvis lærere hjælper med eksamensforberedelser i eksamensfag fremfor at undervise i deres egne, ikke-eksamensfag (Booher-Jennings 2005, p. 242).

Nogle forfattere (f.eks. Au 2007; Brill *et al.* 2018) betragter indsnævring af pensum som en konsekvens af *teaching-to-the-test*, mens andre (f.eks. Amrein-Beardsley 2009) opfatter det som en separat problemstilling, hvor indsnævring af pensum også kan handle om den relative vægt mellem fag og have mere vidtgående konsekvenser, f.eks. hvis dele af læseplanen springes over for nogle elever: Men ligesom *teaching-to-the-test* gælder det også for indsnævring af pensum, at det ikke nødvendigvis har negative konsekvenser, hvis f.eks. læsning opprioriteres i forhold til andre fag, både fordi gode læseevner i sig selv er vigtige, og fordi det også vil styrke elevernes kompetencer i andre fag.

Fokus på marginale elever (*'focusing on marginal students'*) betyder, at indsatsen fokuseres på visse elever, eller rettere sagt deres resultater, fordi de bliver mere værdifulde for skolen, da deres resultater har en relativt stor indflydelse på evalueringerne (Booher-Jennings 2005;



Amrein-Beardsley 2009; Deming *et al.* 2016; Neal & Schanzenbach 2010). Når enkelte elever får stor betydning for, om skolen opnår sine mål eller ej, kan disse børn få større opmærksomhed. Dette kan på den ene side være hensigtsmæssigt og selve formålet med styringen, men det kan også være uhensigtsmæssigt, hvis denne opmærksomhed sker på bekostning af de andre elevers undervisning.

## **D. Meningskabelsesperspektivet**

På grundlag af sin tidligere forskning (Weick 1979) udviklede Weick (1995) begrebet "sensemaking" til ikke blot at være en måde at forstå organisatoriske processer på, men også en mere generel teori om, hvordan vi gennem vores handlinger skaber mening i vores omverden. Ved at anvende meningsskabelse som en teoretisk linse (Lukka & Vinnari 2014) undersøger vi i denne artikel, hvordan ledere skabte mening, når nye initiativer blev igangsat som konsekvens af skolepuljen. Fremfor udelukkende at fokusere på resultaterne af ændringsprocesserne giver meningsskabelse indsigt i, hvordan der skabes mening med begivenheder "that are novel, ambiguous, confusing, or in some way violate expectations" (Maitlis & Christianson, 2014, p. 57). Dette forventer vi var tilfældet, da skolepuljen blev introduceret, og vi studerer derfor, hvordan skoleledere skaber mening samt forsøger at påvirke meningsskabelsen hos deres medarbejdere i retning af en ønsket opfattelse af virkeligheden (Gioia & Chittipeddi 1991), således at der sker ændringer i holdning og adfærd hos både ledere og medarbejdere.

### **1. Mening og handling**

Weick (1995) fokuserede især på, hvordan individer med deres handlinger er med til at skabe det, der samtidig skabes mening om, hvilket Weick beskriver med begrebet *enactment*. Hammer & Høpner (2019) forklarer, at der dermed tale er om "en form for 'at skabe i handling'" (*ibid.* s. 27), når der i en analyse inspireret af meningsskabelse anvendes begrebet 'at skabe'. Det centrale er, at meningsskabelse er en skabelses- eller konstruktionsproces, hvor aktørerne gennem deres handlinger samtidig konstruerer det, de fortolker (Weick & Sutcliffe 2013) - i modsætning til en fortolkningsproces, hvor det fortolkede ikke ændres af fortolkningen. Meningsskabelsen afhænger af en række elementer eller ledetråde (*cues*), som bemærkes (*noticing*) i omgivelserne.

Meningsskabelsen indebærer ikke blot, at der allerede findes en forståelsesramme, som nye ledetråde kan knyttes til, men også at aktører oplever behov for at fortolke inden for denne ramme. Denne proces påvirkes af aktørernes forståelsesrammer i form af mentale modeller, og meningsskabelsesprocesserne er derfor begrænsede af aktørernes evner og færdigheder til at identificere og internalisere ledetråde, dvs. isolere og fokusere på visse specifikke

ledetråde fremfor andre. Derfor vil forskellige personer i samme situation internalisere forskellige ledetråde og dermed handle forskelligt.

I analysen fokuserer vi på, hvordan aktørerne udtrækker ledetråde fra omgivelserne, hvorved specifikke ting udløser (*trigger*) meningsskabelse ved, at de observeres og gives opmærksomhed, så de kan analyseres nærmere (*bracketing*). Derved opbrydes de bestående forståelser, og en ny initial mening skabes baseret på de forhold, opfattelser og ledetråde, der er tilstede i organisationen. Det kan både være metaforer, narrativer og mere konkrete eksempler fra praksis. Ved denne meningsskabelse har handlinger og initiativer en afgørende betydning for meningsskabelse, idet der skabes mening i handlingerne (*enactment*), ved at handlingerne udføres i overensstemmelse med forståelsen af omgivelserne. Se også Hammer & Høpner (2019) for en mere omfattende introduktion til begrebet meningsskabelse.

## 2. Skolelederens rolle

Vi undersøger i artiklen, om skoleledernes meningsskabelse kan bidrage til at forklare, hvordan en ekstern begivenhed, i dette tilfælde skolepuljen, modtages, og hvordan der skabes mening med handlingen, når den implementeres i en organisation. Ledere påvirkes af de omgivelser, de agerer i (Maitlis 2005), og som Maitlis & Lawrence (2007) påpeger, vil ledere blive påvirket og opfange signaler fra de aktører, de interagerer med, hvilket vil påvirke deres fortolkninger og forståelser af forandringen. Sådanne påvirkninger kommer ikke alene fra lærere og elever på skolen, men også f.eks. fra politikere, forældre og lokalsamfund. Skolelederens meningsskabelse er særdeles indlejret i skolernes kultur og traditioner og er integreret med normer, værdier og holdninger på området (Spillane *et al.* 2002). Disse fagprofessionelle forståelser kan i et meningsskabelses-perspektiv opfattes som mentale modeller, der i form af forståelsesrammer etableres gennem arbejdssituationer, uddannelse og livserfaringer.

Selvom mentale modeller guider fagprofessionelle i deres opfattelse af situationer og i den daglige professionspraksis (Weick 1995), kan de mentale modeller blive forstyrrede. Dette sker f.eks. hvis forventninger ikke opfyldes, identiteter trues, eller mål ikke opnås. I sådanne situationer kan nye meningsskabende processer blive initieret. Som anført af Weick & Sutcliffe (2003) viser litteraturen, at organisatoriske og styringsmæssige ændringer ikke nødvendigvis indebærer ændringer i meningsskabelser, hvis gruppenormer og kultur forhindrer det, ligesom medarbejdere aktivt med udgangspunkt i deres egne fortolkninger kan modstå ændringer (Sonenshein 2010). Ledelsens initiering og engagement i meningsgivende processer er derfor afgørende, hvis der skal skabes nye forståelsesrammer og ændringer i skolernes arbejde.

## **E. Metodologi**

Den empiriske del af artiklen er baseret på en kvalitativ metodologi. Kvalitative metoder er velegnede til studier af dynamiske processer relateret til meningskabelse (Gioia & Thomas 1996) samt til at forstå kompleksitet (Flick 2009), når økonomisk styring anvendes i en fagprofessionel sammenhæng. Ved anvendelse af meningskabesperspektivet følger vi de redegørelser og argumentationer, som skoleledere præsenterer ved fortællinger, historier og forklaringer, når de retrospektivt giver mening til de forstyrrede aktiviteter og genskaber mening (Sandberg & Tsoukas 2015; Weick 2001).

### **1. Skolerne i undersøgelsen**

Uafhængigt af skolepuljen foretager Børne- og Undervisningsministeriet hvert år en beregning af den såkaldte undervisningseffekt (også benævnt løfteevne) for de enkelte skoler. Dette gøres ved først at beregne den socioøkonomiske reference<sup>10</sup> for grundskolekaraktererne som et statistisk beregnet udtryk, der viser, hvordan elever med samme socioøkonomiske baggrundsforhold har klaret afgangsprøverne på landsplan. Dernæst beregnes undervisningseffekten som forskellen mellem den socioøkonomiske reference og skolens faktiske karakterer. Hermed etableres et mål for, hvordan skolens elever har klaret afgangsprøverne i forhold til elever på landsplan med samme socioøkonomiske baggrund.

Vi kontaktede fire af de jyske skoler, der deltog i skolepuljen. Skolerne blev udvalgt, så to var fra halvdelen med den relativt højeste procentdel af svage elever, og to var fra halvdelen med den relativt laveste procentdel af svage elever. Inden for de to grupper valgte vi én skole med positiv og en med negativ undervisningseffekt. Valget blev foretaget for at have en vis repræsentativitet, men pga. det lave antal skoler skelner vi ikke i analysen mellem skolerne. Alle fire skoler havde mellem 40 og 50 afgangselever fordelt på to afgangsklasser, hvilket gjorde dem berettigede til en bonus på 1,3 mio. kr., hvis de opnåede målet. Oprindeligt havde alle fire skoler accepteret at deltage i vores undersøgelse, men skolen med en lav andel svage elever og positiv undervisningseffekt faldt senere fra, hvilket efterlod os med tre skoler. Derfor valgte vi efterfølgende en anden skole med tilsvarende karakteristika til at indgå i undersøgelsen.

### **2. Dataindsamlingen**

De oprindelige tre skoler blev interviewet to gange: en gang i det første år af skolepuljen og en gang i løbet af det andet. Den senere tilkomne skole blev interviewet en gang i løbet af det andet år. Venstre side af tabel 1 opsummerer, hvordan de fire skoler A, B, C og D blev valgt. I højre side af tabellen angives, om skolen opnåede bonus for det første år. For at respektere anonymitet angiver vi ikke yderligere detaljer vedrørende skolerne.

Skole	Baseline og undervisningseffekt	Interviewperson		Målsætning opnået
		1 år	2 år	
A	Over 40% og positiv	Skoleleder	Skoleleder	Ja
B	Under 40% og negativ	Viceskoleleder	Skoleleder	Nej
C	Over 40% og negativ	Skoleleder	Skoleleder	Ja
D	Under 40% og positiv		Viceskoleleder	Ja

*Tabel 1: Skolerne i undersøgelsen, karakteristika og interviewpersoner*

Vi interviewede de ledere, der havde ansvaret for deltagelsen i skolepuljen, hvilket på to af skolerne var skolelederne (skole A og C) og på de to andre var viceskolelederne (skole B og D). Viceskolelederen var dog fungerende skoleleder på skole B, da vi gennemførte det andet interview. Skoleledelserne blev valgt som interviewpersoner, fordi de som mellemledere generelt har betydning for, hvordan forandringsprocesser og styringstiltag implementeres og opfattes (Bukh & Svanholt 2020; Martin-Rios 2016). Derudover har tidligere forskning (Ganon-Shilon & Schechter 2018; Spillane & Kenney 2012) specifikt vist, at skoledernes meningsskabelse spiller en vigtig rolle, når nye eksterne politikker og initiativer skal implementeres og koordineres med intern praksis. I den empiriske del vil vi generelt ikke skelne mellem, om interviewpersonen er skoleleder eller viceskoleleder, og når vi analyserer generelle udsagn, vil vi blot bruge betegnelsen skoleledelsen eller lederne. Når der refereres til udsagn, hvor lederne taler om generelle opfattelser, der fortolkes som delte af ledere og medarbejdere på skolen, anvender vi nogle gange formuleringer, hvor skolerne optræder som aktører (f.eks. "skolerne lagde vægt på").

De empiriske data omfatter syv semistrukturerede interviews (Kvale 1996) af to runder. Interviewene var baseret på en interviewguide, der gav mulighed for et indblik i interviewpersoners forskellige opfattelser. Interviewguiden omfattede fire temaer: Det første tema omhandlede beslutningen om deltagelse i skolepuljen, mens andet tema fokuserede på, hvordan incitamentsmodellen blev oplevet. Tredje tema spurgte ind til, hvilke initiativer deltagelsen i skolepuljen havde initieret, og fjerde tema handlede om de forventede langsigtede implikationer af deltagelse. I den første interviewrunde blev der hovedsageligt lagt vægt på tema et og to, mens den anden interviewrunde omfattede alle fire temaer med nogenlunde samme vægt, idet vi i anden runde også bad lederne om at beskrive og reflektere over begivenheder, der havde fundet sted, siden det første interview.

De specifikke spørgsmål blev stillet så åbne som muligt for at give lederne mulighed for at beskrive og forklare de overvejelser, de havde gjort sig, og de bekymringer, de måtte have i forhold til skolens deltagelse i skolepuljen. Interviewene blev således tilrettelagt med fokus på at gøre det muligt for lederne at bidrage med historier, der var vigtige for dem, så vi kunne

analyse de fortolkningsprocesser, der for lederne skabte mening med forandringerne og førte til beslutninger om, hvordan de skulle beskæftige sig med initiativer, der ville påvirke lærerne og være med til at forme skolens udvikling.

Interviewene, der varede cirka 90 minutter, blev optaget og transskriberet. Den første interviewrunde fandt sted i juni-juli 2018, mens den andenrunde fandt sted cirka 6 måneder senere i januar-februar 2019. Interviewet med skole D, som blev inkluderet senere, blev udført i oktober 2018. Skolerne A, B og C blev således alle interviewet i skolepuljens første og andet år, mens skole D kun blev interviewet i løbet af det andet år. Alle interviewene blev gennemført, inden S-regeringens beslutning om at afslutte præmiepuljeprojektet blev offentliggjort i oktober 2019, så dataindsamlingen har ikke været påvirket heraf.

### **3. Analyser og præsentation af data**

Under interviewene var vi opmærksomme på pålideligheden af undersøgelsen ved at undgå ledende spørgsmål (Kvale 1996). Desuden blev pålideligheden styrket ved at transskribere optagelserne ord for ord og lade kodningen blive gennemgået af to af artiklens forfattere. I forhold til validiteten var det centrale spørgsmål, om konklusionerne ville give mening for aktørerne på området samt læserne, dvs. om der blev opbygget et autentisk billede af det, der er undersøgt.

Den indledende analyse af de indsamlede data var baseret på kodning i Nvivo 12. Det blev udført gennem koncept- og i vivo-kodning (Saldaña 2016) med det formål at forankre analysen i interviewpersonernes ordvalg og sprog. Der blev udviklet et case resumé for hver skole, før de mere generelle tendenser blev undersøgt. Alle skoler blev kodet hver for sig, før der blev udviklet et fælles kodningsskema baseret på notater under gennemlæsningen af de transskriberede interview og brugen af display (Miles *et al.* 2014). Udsagn præsenteret i resultaterne er repræsentative for generelle meninger, hvis ikke andet er angivet.

## **F. Analyse**

Lederne havde generelt en pragmatisk tilgang til skolepuljen. De fleste interviewpersoner udtrykte positive men skeptiske holdninger. Enkelte var mere direkte kritiske. F.eks. var skoleleder C indledningsvis positivt indstillet over for puljen men fandt det samtidig provokerende, at skolerne selv skulle finansiere initiativerne, og kun hvis målene blev opnået, ville de "få nogle penge til at betale for indsatserne du egentligt har lavet" (C1). Tilsvarende var andre skoler opmærksomme på, at finansieringsprincippet øgede den økonomiske risiko,

fordi puljen blev udmøntet i slutningen fremfor i starten af et initiativ. Overordnet set var de fleste skoler dog positive over for incitamentsmodellen.

### **1. Mødet med den resultatbaserede finansiering**

Skolepuljen skabte forandringer på skolerne. Skoleleder A bemærkede således, at det tidligere havde været opfattet som problematisk at fokusere på karakterer, fordi det kunne komme i konflikt med elevernes læring, personlige udvikling og trivsel. De bekymringer stemmer overens med kritikken af standardiserede tests i litteraturen (f.eks. Amrein-Beardsley 2009; Booher-Jennings 2005). Puljen havde været med til at ændre denne opfattelse. Skoleleder A fortalte:

Vi har jo i rigtigt mange år sagt, at man skal trives, før at man kan lære noget fagligt. Men altså, faglighed i sig selv, det, at man er dygtig til et eller andet, det giver jo faktisk også trivsel. Så derfor tænker jeg, at det er godt, at det ikke kun er den ene ting men også den anden (A1)

Lederne opfattede skolepuljen som initierende for forståelsen af et forandringsbehov og igangsættelsen af nye initiativer, der skabte ændringer. Samtidig blev de nye muligheder knyttet sammen med den bestående praksis og forudgående behov. F.eks. skabte skoleleder B mening med de initiativer, de satte i værk for at hæve elevens faglige niveau og karakterer som noget de alligevel ”skulle gøre og iværksætte, så der var ikke nogen grund til at sige nej, hvis vi havde muligheden [for at deltage] ” (B1).

Skoleleder A bemærkede yderligere, at puljen var ”meget fokuseret på det faglige, og det er jo det, man gerne vil have op ... når målet er det, og vi gerne vil have de 1,3 mio., så går vi også efter det” (A1). Ved at sætte nye initiativer i gang skabtes mening med skolepuljen igennem handlingerne. Skolelederen fremhævede med bemærkningen om, at ”nu skal det ikke bare lyde skidt det hele” (A1), at han forstod den målrettede indsats positivt. Forståelsen af skolepuljen blev således skabt i både de handlinger, puljen førte til, og ud fra deres sammenhæng med forudgående praksis og elevgruppens behov.

Da vi gennemførte den anden interviewrunde, vidste skolerne, om de havde nået målsætningen og dermed opnået bonus for det første år. Vi var derfor opmærksomme på, om første års resultat påvirkede den nye meningskabelse, f.eks. ved at forudgående fortolkninger eller forventninger blev genovervejet, samt hvilken betydning dette havde for opfattelsen af skolepuljen. Skolerne A, C og D, der opnåede bonussen efter det første år. De opfattede generelt, at interessenter i form af bl.a. politikere, forældre, medarbejdere og lokalsamfundet reagerede positivt. Selvom mediernes omtale var blandet, blev der fokuseret på positive fremfor kritiske historier. I interviewet efter at skole A havde opnået sin bonus,

forklarede skolelederen: "Der har jo været den her stemning med: 'Yes! Det lykkedes den torsdag, hvor vi havde den sidste danskprøve' " (A1). Det blev fremhævet som en sejr for hele skolen. På trods af at bonussen var opnået, måtte lederne dog erkende, at nogle lærere stadig var stadig kritiske i forhold til incitamentsmodellen.

Direkte adspurgt om, hvilke konsekvenser resultatet havde haft på skole B, der ikke nåede målet, blev det fremhævet, at de negative konsekvenser havde været begrænsede. Skole Bs viceskoleleder fortalte:

Altså, skolechefen har bemærket det, naturligvis. Når vi er med i et projekt, så ville de jo gerne, at man opfylder målet. Men der har ikke været noget dunken oven i hovedet eller at skulle stå skoleret, eller hvad man nu skulle. Der er en forståelse for, at det er sådan, det nu er, og vi er videre i processen med det (B2).

Selvom vi ikke spurgte direkte til det, gav interviewene ikke indikation af, at skolens manglende succes med at opnå bonusen førte til yderligere ændringer på skolen eller i sig selv ændrede opfattelsen af skolepuljen.

Traditionelt har folkeskoleområdet været kendetegnet ved en forholdsvis langsigtet budgetsikkerhed for de enkelte skoler, idet de fleste kommuner anvender tildelingsmodeller, hvor budgettet ikke reguleres i løbet af skoleåret, selvom antallet af indskrevne elever på skolen ændrer sig. Ydermere anvender en del kommuner klassebaserede tildelingsmodeller, hvilket indebærer, at tildelingen af ressourcer ligger nogenlunde fast langt ud i fremtiden, medmindre der gennemføres generelle reguleringer af taksterne, f.eks. som følge af besparelsesrunder. I lyset af sådanne praksisser indebærer skolepuljens finansieringsmodel en større budgetrisiko, fordi den resultatafhængige tildeling ikke kendtes på forhånd. Flere af lederne kommenterede direkte på det nye finansieringsprincip og udtrykte bekymringer for, at incitamentsmodellen kunne føre til mere resultatbaseret finansiering i folkeskolen. En skoleleder begrundede det med, at "formålet er jo godt. ... Det er mere måden, man gør det på. ... [man] kunne så gå igennem det samme program og så bare få de her 1,3 mio. kr. på forhånd. ... Det kunne også være, at det havde virket lige så godt." (A2).

Flere ledere påpegede desuden den usikkerhed, der var forbundet med, at faktorer, der ikke var inkluderet i modellen, kunne påvirke resultaterne. Det blev eksempelvis nævnt, at nogle skoler med en stor andel elever, som ikke har dansk som modersmål, muligvis ikke kunne forbedre resultaterne nok til at opnå bonusen. En viceskoleleder fremhævede således, at nogle deltagende skoler kunne have en elevgruppe, der "er så ekstrem, at hvis du skulle differentiere folkeskolebegrebet, så er de måske endda knap folkeskoler; så er vi på vej mod specialklasser." (D2).

## 2. Vurdering af nye initiativer

Vi spurgte eksplicit lederne, om der kunne forekomme snyd, men uden at vi konkretiserede, hvilke former for snyd der kunne være talte om. Skolelederen for Skole A så det som en hypotetisk men ikke sandsynlig mulighed, at skoler ville snyde:

Der er jo nogle, der klarer sig bedre og dårligere til enhver prøve. Der kan man jo ikke gøre noget. Man kan jo sådan begynde at lege med en tanke om at sende 150.000 kr. i flyttepenge til en familie, fordi de har nogle dårlige børn, men det gør vi jo ikke i Danmark, vel! (A1).

Leder A pointerede, at denne form for adfærd ville repræsentere ”en fuldstændig forkvaklet tankegang” (A1), fordi det offentlige skolesystem er for alle elever. Uanset, at alle ledere mere eller mindre klart tog afstand fra snyd, indgik incitamentsmodellens potentielle mulighed for dysfunktionel adfærd som væsentlige ledetråde for ledernes forståelse af skolepuljen, selvom det primært skete ved fortolkning af hypotetiske eksempler. Viceskoleleder B bemærkede eksempelvis:

Hvis vi havde den ene elev, der skulle have et 10-tal til prøven. Altså, forstå mig ret, så fik han det 10-tal ... uden at vi har været i den situation. Men hvis det er rigtigt, min tese, at så bliver det et 10-tal, så er der jo et eller andet, der ikke holder. (B1)

På den måde kunne puljen ”have en tendens til nogle gange at bringe nogle ting i spil i en skolekontekst, hvor de ikke hører hjemme” (B2). Der blev imidlertid ikke i interviewene givet konkrete eksempler, der blev fortolket som snyd. Generelt blev forståelsen af de specifikke initiativer italesat inden for rammerne af en grundlæggende forståelse af, at børnene samt deres læring og udvikling var i centrum. Ved at handlingerne førte til positive resultater, blev forståelsen af de nye tiltags nødvendighed i lige så høj grad skabt i handlingen som ud fra en forudgående nødvendighed. Legitimiteten af de igangsatte initiativer var altså ikke forudbestemt, men den var resultatet af at bemærke de positive konsekvenser. Ved både at indfortolke bestående og nye initiativer i den pædagogiske praksis blev de bestående rammer både bekræftet og udviklet.

Selvom snyderi gennemgående blev betragtet som uetisk og upassende, så anvendte alle fire skoler på forskellig vis praksisser, der (jf. Popham 2001; Amrein-Beardsley 2009) kunne have været opfattet som *teaching-to-the-test*. Dette skete eksempelvis ved at øve eksamensformen og ved at lære eleverne eksamensstrategier til at optimere elevernes eksamenskarakterer. Det var også markant, at selvom initiativerne blev introduceret sammenfaldende med incitamentsmodellen og som en konsekvens heraf, blev de nye praksisser oftest fortolket ind i den eksisterende praksis fremfor at være en konsekvens af det øgede økonomiske



incitament. Meningen med handlingerne blev dermed skabt ind i forudgående handlingsmønstre. I modsætning til de overvejende negative holdninger til forøget accountability i litteraturen (Hardy *et al.* 2019; Jacobsen & Rothstein 2015) var skolelederernes vurderinger af konsekvenserne dermed overvejende positive.

### 3. Forståelsen af den resultatbaserede finansiering

Skolerne var opmærksomme på, at der kunne være forskelle skoler imellem i mulighederne for at realisere resultatmålene og dermed opnå bonussen. Der blev især lagt mærke til, hvordan forskelle i de konkrete afgangsklassers elevsammensætning kunne have indflydelse på de samlede resultater og derfor ville påvirke både, hvilke indsatser der var nødvendige, og hvilke resultater der kunne opnås. I forlængelse heraf gav alle fire skoler udtryk for, at de startede med at vurdere deres specifikke muligheder, herunder de økonomiske risici samt muligheden for at opnå resultatmålet ved at vurdere hver enkelt elevs faglige kompetencer og forventede karakterer. Med udgangspunkt i vurderinger af eleverne blev den umiddelbare forståelse løbende diskuteret, tal blev opdateret og tiltag blev vurderet regelmæssigt. Lederne involverede sig på denne måde i de projekter, der udsprang af skolepuljen.

Vores overordnede forventning, inden vi gennemførte dataindsamlingen, var, at deltagelsen i skolepuljen ville forhøje *stakes*, og at dette ville få tilsvarende konsekvenser som ved andre styringstiltag, der er baseret på anvendelsen af tests med forøgede *stakes* - og dermed potentielt kunne have dysfunktionelle konsekvenser. Ledernes forståelser indikerede imidlertid, at *stakes* ikke blot var fastlagt ud fra incitamentsmodellens præmisser eller skolernes karakteristika, idet skolernes initiativer i sig selv skabte ændringer i *stakes*. Selvom skolepuljen gav skolerne mulighed for yderligere finansiering uden at ændre noget og uden at løbe en risiko, hvis årgangen betragtedes som "sikker", igangsattes nye initiativer.

Alle interviewpersoner forsøgte at nå resultatmålene og dermed opnå bonus. Blot ved at deltage i skolepuljen fik karakterne derfor større betydning. Men *stakes* blev oplevet forskelligt skolerne imellem, da deres mening blev skabt i en interaktion mellem skolernes indsats og klassernes karakteristika. En viceskoleleder beskrev, at opnåelsen af bonussen

dels [har] noget at gøre med det arbejde, de præstationer, [vi] har leveret. Og så har det at gøre med de elever, der er nede i klassen. Det tænker man nu. Og så har det jo hele tiden at gøre med børnenes formåen, socioøkonomisk, intelligens osv. Det spiller ind. (B1)

Eksempelvis oplevede skole A det første år stort set ikke noget pres ved forøgede *stakes*, fordi forventningen var, at resultatmålet ville blive opnået relativt let. Lederen lagde således vægt på, at "den årgang, vi havde det første år, det var en ret dygtig årgang ... det var stort set

gratis" (A1), og han forventede ikke større ændringer i undervisningspraksis. Skolelederen opsummerede det således:

Vi har nok lavet flere tests, været mere fokuseret på enkelte elever og på enkelte fag, end vi gjorde [før] ... men når jeg sådan tænker tilbage og bliver spurgt: Har vi egentligt lavet en hel masse? Nej. Men vi har fokuseret på at nå det her mål, og det tror jeg måske har været en del af det. (A1)

I modsætning til skole A opfattede skoleleder C *stakes* højere. De besluttede at investere i projektet ved at bruge flere penge på ekstra personale, end der var afsat i det oprindelige budget. Investeringerne forøgede presset, hvorfor handlingen betød, at forståelsen af karaktererne og deres konsekvenser blev ændret. At investere blev fremhævet som risikabelt både af skolelederen selv og andre, da det kunne påvirke skolen negativt, eksempelvis ved at føre til fyringer, hvis skolen ikke opnåede målet. Den forøgede risiko blev forsøgt håndteret ved at opretholde en sikkerhedsmargin på mindst en elev ekstra i forhold til resultatmålet, så man ikke kun lige planlagde at reducere antallet af svage elever nok til at opnå bonussen.

I løbet af det andet år ændrede skole C's økonomiske forhold sig imidlertid pga. kommunens generelle budgetreduktioner. Skolelederen gennemførte besparelserne ved at opgive nogle af de initiativer, der var blevet igangsat som følge af skolens deltagelse i skolepuljen. I forlængelse heraf forklarede skolelederen, at projekterne ikke var blevet afsluttet af hensyn til lærerne. Efter de særlige initiativer var blevet opgivet, fandt han ud af, at presset var blevet forstået anderledes af lærerne, end han selv havde opfattet det. Lederen havde forventet, at projektet kun ville lægge et pres på ham, fordi det var ham, der skulle stå til ansvar over for den større budgetmæssige risiko. Selvom lederen udtrykte, at han havde tilstræbt, at lærerne ikke skulle mærke et forøget pres, havde lærerne – særligt 9. klasses lærerne – følt det. Først efter at de særlige initiativer var opgivet, gav lærerne udtryk for dette. Han beskrev:

Og så får de bare givet udtryk for: "Puh, det er meget rarere i år. Vi ved, at vi bare kan koncentrere os om eleverne og gøre dem så dygtige. Vi vil selvfølgelig gerne i mål og have pengene... og at de skal få 4 i dansk og matematik. Men vi ved, at der ikke er det her pres, og vi ved, at vores job ikke er afhængige af, at vi når i mål" (C2).

Handlingsmønsteret baseret på puljen havde dermed fået en anden mening for lærerne, end det var tiltænkt fra skolelederen. Dette viser, at meningskabelsen ikke blot sker mellem skolelederne (jf. Balogun & Johnson 2004), men også mellem ledelse og lærere. Efter den nye regering havde fjernet muligheden for at opnå bonus, blev indsatsen forstået som "en mere oprigtig hjælp" (C2), fordi fokus ændrede sig til "at eleven kan blive så dygtig som han/hun kan fremfor at have fokus på at nå 4-tallet" (C2). Retrospektivt fortolkede lederen, at hvis projektet var fortsat på samme vis, "så ville der have været flere sygdomsmeldinger med stress.

Og vi har haft nogen sidste år også, hvor det efterfølgende helt klart kommer frem: Det var det, der var med til at presse dem” (C2). Efter denne oplevelse ændrede skolelederen sit syn på incitamentsmodellen og blev mere kritisk i forhold til incitamentsstrukturen og i særdeleshed i forhold til dens fokus på elever med et karaktergennemsnit omkring 4. Dermed blev meningen med skolepuljen ikke alene skabt af dens karakteristika og de resultater, der blev opnået, men også af de handlinger, den førte til.

#### **4. Inspirationsprogrammet**

Alle skoler opfattede, at de havde gavn af inspirationsprogrammet, og mange af de initiativer, der blev igangsat, udsprang af inspirationsprogrammets kurser og workshops. Selvom der blev lagt vægt på skolepuljen som en initierende eller motiverende faktor, blev inspirationsprogrammet i lige så høj grad nævnt, når vi spurgte til skolepuljen, idet inspirationsprogrammet formede indsatserne. En skoleleder uddybede det således:

Puljen har været en katalysator [til] at komme i gang, tænker jeg. Men jeg tænker, at havde det bare været puljen i sig selv, havde den ikke gjort det, for havde vi ikke fået alt den viden på workshops ... så var vi ikke kommet dertil, hvor vi [er]. Ikke på nuværende tidspunkt. (C1).

Det indikeres således, at ledernes opfattelse af skolepuljen blev påvirket af, at de knyttede den sammen med inspirationsprogrammet. Edwards & Mindrila (2019) har pointeret, at de praksisser, som anvendes af højt præsterende skoler, ofte indebærer “slow, incremental growth and require exceptional effort by educators” (*ibid*, s. 4). Incitamentsmodellen krævede imidlertid øjeblikkelige forbedringer, og skolerne lagde derfor vægt på, at de var nødt til at vælge “de her lavthængende frugter, altså, hvad ville vi med størst sandsynlighed kunne løfte mest muligt i løbet af seks måneder” (B1). Reaktion på resultatbaseret finansiering argumenteres ofte også for at ske på ”the quickest, cheapest and simplest manner possible” (Li & Ortagus 2019, s. 325). Selvom reaktionerne på incitamentsmodellen kunne opfattes som en nedprioritering af langsigtede og mere grundlæggende læringsindsatser, var fokus på den umiddelbare gavn for eleverne. F.eks. blev udlevering af gratis morgenmad til eleverne forklaret som en nem måde at sikre, at eleverne fik en god start på eksamenerne. Ved at tillægge initiativerne betydning for eleverne blev både de nye praksisser og skolepuljens præmisser inkorporeret i skolernes forståelsesrammer, fremfor at der blev skabt nye betydninger og meninger med undervisningen.

#### **5. Vurdering af nye initiativer**

Adspurgt til, hvilke nye tiltag der var sat i værk, kom flere ledere selv ind på, hvorvidt der var tale om *teaching-to-the-test*. De gav udtryk for, at det kunne være negative konnotationer forbundet med en sådan praksis, men lederne lagde vægt på, at den konkrete praksis omkring

eksamensafviklingen var legitim, fordi det gav eleverne den største chance for at få succes. F.eks. nævnte skole A, at den øgede testfrekvens og anvendelsen af testresultater til at vurdere elevernes forventede karakterer gav mulighed for at igangsætte intensive læringsforløb for udvalgte elever, når det blev vurderet, at de ville have gavn af det.

Skolepuljen blev etableret for at fremme de svage elevers resultater. Svaghed var i princippet klart defineret ved karaktermålsætningerne fremfor særlige karakteristika ved eleven. Da definitionen retrospektivt var baseret på, hvilke karakterer eleverne faktisk opnåede, indebar den et paradoks: Hvis en 'svag' elev opnåede 4 ved afgangseksamenerne, ville skolen have gjort en forskel og netop løftet elevens præstation, men i så fald ville eleven ikke længere indgå som en svag elev. Derudover var de elever, der ved resultatopgørelsen talte som 'svage', ikke nødvendigvis dem, der ud fra andre betragtninger var svage, men blot dem, der på eksamensdagen ikke opnåede karakterkravet. 'Svaghed' og deraf incitamentsmodellens målsætning var dermed på trods af den klare definition tvetydig i praksis. En forbedring af indsatserne krævede derfor en fortolkning af målgruppen.

Hvad enten det var, fordi bonussen blev opgjort på skoleniveau, eller det var, fordi målgruppen ikke kunne afgrænses på forhånd, så indebar målgruppefortolkningen, at skolerne havde fokus på de elever, der blev opfattet som afgørende for at opnå bonussen. Det nødvendiggjorde en løbende fortolkning af elevernes resultater ved interne prøver og afleveringsopgaver, såvel som vurderinger af de nye indsatsers effekt. Det indebar f.eks., at elever, der på et givet tidspunkt blev opfattet som svage og dermed udvalgt til intensive læringsforløb, blev drøftet på ledelsesniveau for at vurdere deres fremskridt og sandsynlighed for succes ved afgangseksamenerne. Booher-Jennings (2005) har med begrebet "bubble kids" som beskrivelse af de marginale elever (Deming *et al.* 2016) pointeret, at det kan være problematisk udelukkende at fokusere på elever, der opfattes som afgørende for opnåelsen af fastlagte målsætninger. Da skolepuljens bonus var knyttet til afgangsklasserne, kunne disse på tilsvarende vis fremstå som marginalklasser, som kunne få fokus på bekostning af de andre klassers elever. Denne afgrænsning var imidlertid heller ikke helt entydig: Selvom eleverne i indskolingen ikke på noget tidspunkt ville kunne påvirke resultaterne i forhold til incitamentsmodellen, så ville elever i 8. klasse dog indgå i tildelingsgrundlaget det følgende år.

På alle skolerne blev målgruppen løbende afgrænset, udvidet og fortolket. Selv på de skoler, der i mindre grad igangsatte nye initiativer, var der opmærksomhed på, hvordan muligheden for at opnå bonus udviklede sig i lyset af elevgruppens læringsfremskridt. Afgrænsningen af de enkelte elevers svaghed var i afgangsklasserne baseret på en løbende vurdering af børnenes behov, og fokusering på enkeltelever kunne i praksis derfor altid fortolkes som

hensigtsmæssig, fordi netop disse elever ville have udbytte af det særlige fokus. Til forskel herfra var afgrænsningen mellem klassetrinene fastlagt, og de valg, der blev truffet, når indsatser for hele årgange blev besluttet, var langt mere eksplicite. Disse indsatser fik også mere vægt, når skolelederne blev spurgt om, hvilke initiativer der var blevet igangsat.

Generelt forstod skolelederne det ikke som problematisk, at der blev fokuseret på de afgangselever, der blev opfattet som marginale i forhold til opnåelsen af bonussen. Dette var, fordi mange initiativer i praksis blev forstået som klassebaserede, og fordi initiativerne ofte også omfattede andre klasser. Det var f.eks. tilfældet, da skole B forøgede antallet af undervisningstimer i dansk og matematik samt igangsatte et intensivt læringsforløb for udvalgte elever fra flere årgange, fordi læsning, stavning og matematik blev opfattet som fagområder, hvor mange elever havde forbedringspotentialer. Derfor blev det besluttet, at alle skolens dansk- og matematiklærere skulle deltage i inspirationsprogrammets aktiviteter, så alle klasserne på skolen kunne få gavn af indsatserne. I erkendelse af at det brede fokus indebar væsentlige omkostninger, blev initiativet dog senere afgrænset til klasselærerne for afgangsklasserne.

Selvom det ikke var skoleledernes intention, blev det oftest afgangsklasserne, der blev fokuseret på. De var definerende for svaghedsbegrebet, og dermed var de afgørende for bonussen. Selvom skolernes indsatser ofte blev forstået som gavnlige for alle eleverne ved at gennemføre initiativer, der omfattede alle klassetrin, så var hovedparten af de initiativer, der blev nævnt i praksis, dog mere snævert rettede mod afgangsklasserne. Skole C besluttede f.eks. at ansætte mere personale, end der var budgetteret med for at kunne have tre fremfor to afgangsklasser.

Skole C indførte også læringskonferencer med deltagelse af alle afgangsklassernes lærere, elevernes skriftlige opgaver blev mere individuelt tilpasset, og der blev fokuseret mere på skole-hjem-samarbejdet, f.eks. ved at skolelederen deltog i forældremøder. Tilsvarende besluttede Skole D at forbedre tilbagemeldinger i dansk og matematik, lave studiegrupper, hvor elever med forskellige kompetenceniveauer blev blandet og have tre lærere til at undervise to klasser i dansk og matematik fremfor to, som man plejede. Det gav mulighed for, at elever kunne inddeles i tre hold på baggrund af faglige kompetencer. Disse nye aktiviteter blev gennemført ved at øge skolens samlede udgifter fremfor at reducere ressourcer tildelt andre klasser. Selvom de forøgede udgifter ville blive finansieret, hvis skolen opnåede bonussen, var resultatet stadigvæk, at de marginale afgangsklasseelever fik ekstra opmærksomhed. Lederne var dog opmærksomme på problemstillingen og understregede, at de havde et ansvar for at håndtere dilemmaerne mellem modellen, der gav incitament til at fokusere på de marginale elever, og skolens hensigt med at gøre alle elever så dygtige som

muligt. Ligeledes lagde flere ledere også vægt på det generelle ansvar, de havde for at undgå utilsigtede negative konsekvenser af incitamentsmodellen og initiativerne på skolerne.

Generelt havde skolerne primært fokus på dansk og matematik, der var de fag, der var afgørende i incitamentsmodellen. Men fagene blev forstået som mere generelt nyttige for eleverne, fordi disse færdigheder også er nødvendige for at få succes i andre fag. Dog bemærkede lederne fra flere skoler, at de måske havde for stort fokus på eksamen. Eksempelvis besluttede lærerne på skole C sammen med hver elev i 9. klasse, hvilke typer af matematikopgaver den enkelte elev burde fokusere på at løse til eksamen for at opnå flest mulige point og derved få den bedst mulige karakter, hvorefter det valgte opgavefokus styrede skolearbejdet i matematik resten af skoleåret. Denne praksis blev opfattet som hensigtsmæssig, fordi den blev fortolket som en del af en indsats, hvor eleverne skulle lære at håndtere eksamenssituationen. Skolelederen for skole C forklarede, at elever "skal vide rent færdighedsmæssigt, hvad er det, der kræves af mig, når jeg skal til [eksamen]. ... man skal vide, hvad det er for en eksamensform. Og det er der, det er acceptabelt." (C1).

Jennings og Bearak (2014) pointerer, at *teaching-to-the-test* kan have både positive og negative konsekvenser afhængigt af, hvordan de praktiseres. Dette gav de interviewede ledere ligeledes udtryk for. Selvom de generelt forstod egne praksisser som legitime og hensigtsmæssige for eleverne, var de opmærksomme på, at der var grænser for det acceptable, og flere eksempler på potentielt uhensigtsmæssige aktiviteter blev nævnt. Det var imidlertid karakteristisk, at forståelsen for hensigtsmæssighed oftest blev etableret ved omtaler af hypotetiske praksisser, som ikke var acceptable, og som måske kunne finde sted på andre skoler, men som de ikke selv ville involvere sig i. Skabelsen af mening i egne handlinger blev deraf legitimeret gennem henvisning til andres problematiske men hypotetiske handlinger.

## 6. Varierende opfattelser

Det var imidlertid interessant, at der ikke var fuld enighed skolerne imellem om, hvor grænserne gik, da de i høj grad forstod dem ud fra rammerne af den praksis, de selv havde valgt. Når først handlingerne var valgt skete der dermed en retfærdiggørelse af dem, og dermed en fastholdelse af praksissen som irreversibel. Lederen for skole C sagde eksempelvis, at hvis testforberedelse foregik som "en multiple choice opgave, og ... så kører vi lige 25 [gamle eksamensopgaver] - det ville ikke være i orden" (C1). Skole B praktiserede netop denne type intensiv eksamenstræning. Lederen for skole C påpegede endvidere, at det var problematisk at udpege elever til intensive læringsforløb, mens B og D praktiserede denne metode uden at bemærke noget forkert ved det. Lederen fra skole D fremhævede, at *teaching-to-the-test*-baserede initiativer var uacceptable, hvis det kun havde effekt på det første års

resultater, men det var dog stadig tilladeligt, hvis det blev gjort for at opnå bonussen for at øge niveauet efterfølgende. Eksempelvis øgede skolen fokus på emner, hvor det var opfattelsen, at eleverne på 9. årgang havde særlige vanskeligheder, idet det blev vurderet, at "hvis vi skulle nå det i år, så er det sted, hvor de er svagest kørende, det er matematik. Så det var en rent strategisk prioritering fra vores side" (D2).

Da konsekvenserne af enkeltstående indsætter kun vanskeligt kan adskilles fra den undervisningspraksis, der uden skolepuljens eksistens ville have fundet sted, kan vi ikke med sikkerhed afgøre, om den valgte praksis på skolerne faktisk var gavnlige for afgangseleverne, eller hvorvidt den førte til, at andre elevers udbytte blev forringet eller forbedret. Skoledernes forskellige forståelser indikerer imidlertid, at det er et spørgsmål om, hvilken mening initiativerne gives, og at der er brede rammer for, hvordan eksisterende forståelsesrammer og handlingsmønstre kan inkorporere nye praksisser, og hvordan initiativer, der i nogle sammenhænge kunne forstås negativt, kan blive opfattet positive ved at fokusere på bestemte facetter eller konsekvenser af teknikkerne.

Ingen af de deltagende skoler havde på forhånd fået tilført flere ressourcer fra forvaltningen, fordi de deltog i skolepuljen. Der var derfor grænser for, hvilke nye initiativer der kunne igangsættes uden at nedprioritere andre aktiviteter på skolerne. Ligeledes ville en forøgelse af tidsforbruget på et fagområde i praksis indebære en reduktion af andre fagområder, hvis ikke det samlede timetal skulle øges. På nogle af skolerne var der derfor en erkendelse af, at pensum blev indsnævret ved at øge fokus på dansk og matematik, og det blev forstået som en reduktion af tiden til andre fag. Denne prioritering blev dog på et par af skolerne dog også knyttet sammen med kommunernes overordnede strategier og mål. Skoleleder B så det f.eks. som en naturlig beslutning at styrke dansk og matematik, både fordi "de er grundlæggende fag", og fordi det matchede "den kommunale [prioritering] om læsning" (B2). På tilsvarende vis knyttede andre skoler de nye initiativer sammen med ting, der allerede blev gjort, eller som de opfattede som initiativer, der alligevel ville være blevet taget. De prioriteringer, som blev foretaget som en konsekvens af skolepuljen, blev ved disse fortolkninger forstået som noget, der alligevel ville være blevet besluttet.

Selvom det forøgede fokus på testede emner som pointeret af både Au (2007) og Brill *et al.* (2018) indebærer mindre tid til andre ikke-testede emner, blev de nye praksisser på skolerne forstået som hensigtsmæssige, da de gavnede eleverne, og fordi de var meningsfulde i forhold til de etablerede rammer for, hvordan skolerne altid har arbejdet. Selvom mange af de nye initiativer ikke blev opfattet som indført kun pga. incitamentsmodellen og opnåelse af dens mål, funderede lederen fra skole C over, hvad den meningsskabelse, der fulgte af skolepuljen og det øgede eksamensfokus, havde betydet. Både elever og lærere havde skabt en ny mening

med fagene, som var påvirket af, hvad der gav point til eksamenerne. Men også det forøgede fokus på eksamen og på de specifikke dele af pensum, der var vigtigst for eksamensresultaterne, kunne gives mening inden for den bestående praksis. Lederen fra skole A konstaterede, at selvom fokus på eksamen muligvis var blevet øget pga. incitamentsmodellen, så var det noget, "du altid træner" (A1). Adspurgt, om det forøgede fokus på afgangsklasserne og især de svageste i disse klasser kunne have negative konsekvenser for andre elevgrupper, forklarede lederen for B:

Det lette svar vil være 'ja', og hvad er det så, det er... det her år, der har vi gjort sådan her og sagt: ja, nu er det de her otte stykker ... det bliver dem, vi har fokus på. Det betyder jo ikke, og det her er vigtigt for mig at sige, at vi er gået på kompromis med, hvordan læringssituationen er for [andre elevgrupper]. Men det betyder, der har været en ressourcefordeling, [der prioriterer de otte elever] (B1).

Viceskoleleder fra skole D udtrykte bekymret, at nogle skoler kunne "levere nogle fantastiske resultater et, to år, og så leverer de ikke mere" (D2), fordi de fokuserede på "nogle hacks, man kan lave, hvis man er kynisk nok" (D2) for at nå målene uden faktisk at forbedre sig. Tilsvarende bemærkede lederen for skole C, at man godt kunne forvente, at nogle skoler kun tog initiativerne for at opnå bonus, men han udtrykte også klart, at dette ikke var hans hensigt:

Der skal simpelthen ikke være nogen, der kan komme at sige til mig, at jeg kun gjorde det her for pengenes skyld. ... Jeg vil have det, så jeg kan kigge mig selv i øjnene også om 10 år, når jeg kigger tilbage på det her projekt. Ja, [de specifikke elever] hjalp vi, men vi hjalp også dem [over og under] (C1).

## **G. Diskussion og konklusion**

Mens der i udlandet er omfattende erfaringer med finansieringsmodeller, hvor der tildeles ressourcer til offentlige virksomheder på baggrund af resultat- eller effektmål, er det en forholdsvis ny model i Danmark. Enkelte kommuner og regioner har anvendt effektbaserede tildelingsmodeller (Bukh & Christensen 2020b), ligesom den statslige model for nærhedsfinansiering<sup>11</sup> af regionerne og statens finansiering af den kommunale beskæftigelsesindsats også kan siges at være en resultatbaseret tildelingsmodel<sup>12</sup>. Bortset fra sådanne enkeltstående eksempler var danske erfaringer sparsomme, da Folketinget i 2017 besluttede at indføre skolepuljen (Børne- og Undervisningsministeriet 2017a).

Offentlige virksomheder i Danmark finansieres ofte ved, at finansieringen enten er fastlagt ved budgetårets begyndelse, eller ved at finansieringen er aktivitetsbaseret, f.eks. ved at være afhængig af antal borgere, der visiteres til en indsats. Sådanne finansieringsmodeller tildeler de offentlige virksomheder ressourcer nogenlunde i forhold til de aktiviteter, der udføres, og



dermed de udgifter, der afholdes. Til forskel herfra øges budgetrisikoen, når der er en resultatbaseret komponent i bevillingsmodellen, fordi der ikke nødvendigvis er sammenhæng mellem de resultater, der opnås, og de udgifter, der afholdes. Dette princip er ikke ukendt på uddannelsesområdet, idet videregående uddannelsesinstitutioners finansiering i Danmark er knyttet til, om de studerende består eksamenerne. På tilsvarende vis indebar incitamentsmodellen, at eksamenskaraktererne fik en relativt stor betydning for skolernes ressourcetildeling, og at budgetusikkerheden blev forøget. Der var i det hele taget markant mere på spil ved afgangseksamenerne end sædvanligt. Dog skal det ved vurderingen af resultaterne tages i betragtning, at skolepuljen indebar en mulighed for at opnå en ekstra finansiering, og ikke at den generelle tildeling blev gjort resultatafhængig.

Det kan synes indlysende, at der er tilbageløbseffekter fra tests og eksamener, således at undervisningen samt elevernes adfærd og læringsudbytte påvirkes heraf. I praksis er der et kompliceret samspil, hvor eksamener og tests på den ene side netop anvendes for at forøge læringsudbyttet, og på den anden side har som formål at evaluere elevernes udbytte af undervisningen; ideelt set uafhængig af eksamenssituationen. Forskning har vist, at tilbageløbseffekter forøges, jo mere der er på spil for skoler, lærere og elever (Alderson & Wall 1993; Kelley & Protsik 1999; Lee & Medina 2018). Derfor forventede vi, at tilbageløbseffekterne ved afgangseksamenerne ville blive forøget som følge af incitamentsmodellen. Vi studerede, hvordan lederne fra fire folkeskoler, der var udvalgt til at deltage i skolepuljen, opfattede incitamentsmodellen samt de ændringer i undervisningspraksissen, der blev foretaget på skolerne som konsekvens af skolernes deltagelse. Vi havde i særdeleshed fokus på, om tilbageløbseffekter blev forstået som positive, dvs. førte til en undervisningspraksis, der blev opfattet som hensigtsmæssig, eller negative, hvis praksissen førte til uønskede resultater. Vores udgangspunkt var dermed, at hensigtsmæssigheden blev vurderet ud fra de konsekvenser, der var forbundet med praksissen, samt at vi lod vurderingen være baseret på lederne opfattelser fremfor en foruddefineret karakteristik af, hvad der er god eller dårlig undervisningspraksis.

Tidligere forskning (Gunnulfsen & Roe 2018; Jongbloed & Vossensteyn, 2016; Paletta *et al.* 2019) har vist, hvordan lederne på uddannelsesinstitutioner har en central rolle, når nye tiltag skal implementeres. Det gælder ikke mindst, når skoler skal leve op til krav og forventninger, der er introduceret udefra, og når disse krav skal integreres med den lokale praksis (Lowenhaupt *et al.* 2016; Rutledge *et al.* 2010). I sådanne situationer må ledere fortolke initiativer og krav og skabe nye meninger ud fra den eksisterende praksis (f.eks. Bukh *et al.* 2020a; 2020b; Spillane *et al.* 2002). Selvom de ledere, der blev interviewede, var kritiske i forhold til budgetusikkerheden, som modellen skabte, så fokuserede de først og fremmest på mulighederne for skabe forandringer samt på, hvordan de kunne øge muligheden for at opnå bonussen. Generelt handlede ledelserne bevidst i forhold til incitamentsmodellen, og de lagde

ikke skjul på, at mange initiativer blev igangsat på grund af skolepuljen og mhp. at opnå bonussen. Men det var også gennemgående for de eksempler, som lederne fremhævede, at de blev opfattet som hensigtsmæssige, fordi de gavne eleverne. Meningen med handlingsmønstrene afhang deraf både af forsøget på opnå bonussen og muligheden for at gavne eleverne, hvilket blev forstået som forenelige mål.

Oftest var det niende klasserne, der var i fokus. Dette stemte overens med intentionerne bag skolepuljen, da det var disse elevers eksamensresultater, der skulle forbedres for, at skolen fik mulighed for en bonus. Nogle gange omfattede de igangsatte initiativer dog også andre klasser, selvom det kunne være vanskeligt at fastholde det bredere fokus, når der opstod behov for at prioritere indsatserne. Ved at lægge vægt på at initiativerne skulle være gavnlige, dannedes imidlertid en opfattelse af initiativernes berettigelse, og generelt blev alle de eksempler på nye initiativer, som vi blev præsenteret for, forstået som hensigtsmæssige. Selv handlinger taget på skolen, hvor lærere oplevede et forhøjet pres, blev anset som gode for eleverne. Problemet med stress og potentielle sygemeldinger omhandlede, at lærerne oplevede en øget budgetusikkerhed, og hvordan dette påvirkede meningen med deres arbejde ift. at opnå de nødvendige resultater. Sammenlignet med litteraturen, der overvejende er kritisk over for anvendelsen af tests og eksamensresultater, når skolers ansvar for elevernes læring forøges (Hardy *et al.* 2019; Jacobsen og Rothstein 2015), var ledernes holdninger til skolepuljens intention, indhold og konsekvenser markant positive. Der blev lagt vægt på, hvordan incitamentsmodellen fungerede som en katalysator for ændringer, samtidig med at den hidtidige undervisningspraksis ofte blev fortolket, så de nye handlinger kunne forstås som en naturlig forlængelse af den hidtidige praksis.

Ved studier af tilbageløbs effekter er der ofte fokus på de potentielt negative konsekvenser i form af blandet andet *teaching-to-the-test* (Madau 1988; Popham 2001), indsnævring af pensum (Amrein-Beardsley 2009; David 2011) og fokusering på de marginale elever (Burgess *et al.* 2005; Deming *et al.* 2016; Neal & Schanzenbach 2010). Denne artikel indikerer, at skolerne reagerede strategisk på incitamentsmodellen ved at igangsætte initiativer, der på mange måder svarer til disse tilbageløbs effekter. Men modsat den litteratur, der ofte understreger dysfunktionaliteten af de utilsigtede konsekvenser inden for grundskoleområdet (f.eks. Amrein-Beardsley 2009; Paletta *et al.* 2019) og resultatbaseret finansiering (Blankenberg & Phillips 2016; Li & Ortagus 2019; Umbricht *et al.* 2017), viser dette studie flere eksempler på, at selvom initiativerne potentielt ikke var i overensstemmelse med skolepuljens intentioner, blev der skabt en lokal mening, hvor initiativerne var i elevernes interesse. Dermed viser studiet også, at man skal være varsom med at afgøre, om en undervisningspraksis er god eller dårlig, alene fordi den udgør *teaching-to-the-test*, indsnævring af pensum eller fokusering på marginale elever (Amrein-Beardsley 2009; Paletta

et al. 2019). Der er mange specifikke detaljer vedrørende implementeringen og forståelsen af initiativerne og deres konsekvenser, der også har betydning for vurderingen.

Det var markant, at u hensigtsmæssige initiativer typisk kun blev omtalt hypotetisk og som noget, der måske fandt sted på andre skoler, men som de ikke selv ville gøre. Kun en af lederne opfattede det som en mulighed, at hans egen skole ville kunne snyde, og selv da blev det ikke overvejet som noget reelt, da situationen, hvor det ville kunne finde sted, blev opfattet som usandsynlig. Derimod blev hypotesen brugt som et udtryk for, at puljen kunne have problematiske konsekvenser i praksis. Der var forskellige forståelser mellem lederne i forhold til, hvilke handlinger de anså for problematiske. Som i eksemplet ovenfor, blev disse forståelser konstrueret gennem hypotetiske historier, der spillede en markant rolle ved lederne forståelse af egne initiativer som hensigtsmæssige. Meningsdannelsen udspillede sig således ved, at lederne primært bemærkede de positive og gavnlige elementer ved deres egne indsatser og de negative aspekter af hypotetiske handlinger. Ved at iværksætte initiativer, der blev opfattet som gavnlige, og ved at forstå initiativerne som en fortsættelse af den eksisterende praksis, blev de hypotetiske eksempler på snyd og dysfunktionalitet rammesættende for gavnlige handlinger ved at pointere, hvad der ville være brud på ordentlig praksis. Skolepuljen og dens præmisser blev dermed oftest integreret i de mentale modeller i stedet for at forstyrre de bestående forståelser.

Lederne var opmærksomme på, hvordan initiativerne kunne forøge præstationspresset og påvirke lærerne. Der var forskelle, ikke kun skolerne imellem, men også mellem de to årgange af afgangsklasser i interviewperioden, hvilket betød, at der var forskelle i muligheden for at opnå bonus i de to år. Resultaterne indikerer, at lederne fortolkning af, hvor krævende det ville være at opnå forbedringsmålsætningerne, havde betydning for, hvordan der blev handlet på den enkelte skole. Hvis forbedringsmålene blev opfattet som lette, f.eks. fordi eleverne i et specifikt år var mere kvalificerede end i baseline-året, fik de nye initiativer mindre ledelsesopmærksomhed, og ændringerne var begrænsede. Hvis målet til gengæld blev opfattet som vanskeligt at opnå, f.eks. på grund af økonomiske nedskæringer, opfattede skolerne et større pres for at opnå forbedringer (og dermed forhåbentlig en bonus), og de nye aktiviteter fik større opmærksomhed. Udover at presset for at opnå bonus blev bestemt af, hvordan lederne opfattede skolens situation, er det også værd at bemærke, at dette pres kunne udvikle sig over tid og var ikke givet alene ved incitamentsmodellen, men blev i lige så høj grad skabt af de igangsatte initiativer og den måde, disse blev forstået på.

Da undersøgelsen kun omfatter fire skoler, bør resultaterne fortolkes med en vis varsomhed. Eksempelvis var forskellige versioner af *teaching-to-the-test* udbredte, men det var fordelene ved at gøre eleverne fortrolige med eksamensformen og betydningen af, at de var

velforbredte, motiverede (og mætte), der blev tillagt vægt ved ledernes fortolkninger. Når forholdene omkring eksamenssituationen blev prioriteret højt, blev det forstået som en styrkelse af elevernes udbytte. Denne opfattelse er formodentlig påvirket af, at de danske eksamener faktisk afspejler målsætninger og generelle færdigheder, der værdsættes af lærere (Beuchert og Nandrup 2018) og er vigtige for elevers læringsudbytte. De potentielt negative sider af *teaching-to-the-test*, der kendes fra litteraturen (f.eks. De Wolf og Janssens 2007), kunne ikke genfindes i ledernes forståelse af deres egen praksis. Internationale studier har vist, at et øget fokus på tests og eksamener kan føre til, at flere elever ikke deltager i testene (De Wolf & Janssens 2007, s. 390). Vi oplevede ikke indikationer på lignende utilsigtede bivirkninger, men det skyldes muligvis, at forbedringsmålene ifm. skolepuljen blev målt i forhold til alle elever, der var indskrevet på skolen og ikke kun i forhold til de elever, der deltog i afgangseksamenerne. Selvom vi ikke kan validere denne betingelse, viser resultatet vigtigheden af at overveje de specifikke detaljer, når man udvikler resultatbaserede finansieringsmodeller.

## Referencer

- Akpınar, K.D. & B. Cakildere. (2013). Washback Effects of High-stakes Language Tests of Turkey (KPDS and ÜDS) on Productive and Receptive Skills of Academic Personnel, *Journal of Language and Linguistic Studies*, 9(2), 81-94.
- Alderson, J. C. & D. Wall. (1993). Does Washback Exist?, *Applied Linguistics*, 14(2), 115-129. doi: 10.1093/applin/14.2.115. <https://doi.org/10.1093/applin/14.2.115>
- Al-Samarrai, S., U. Srestha, A. Hasan, N. Nakajima, S. Santoso & W. H. A. Wijoyo. (2018). Introducing a Performance-Based Component into Jakarta's School Grants: What Do We Know About Its Impact After Three Years?, *Economics of Education Review*, 67, 110-136. <https://doi.org/10.1016/j.econedurev.2018.10.005>
- Amrein-Beardsley, A. (2009). The Unintended, Pernicious Consequences of "Staying the Course on the United States' No Child Left Behind Policy", *International Journal of Education Policy and Leadership*, 4(6), 1-13. doi: 10.22230/ijepl.2009v4n6a199.
- Andersen, S. C., L. M. Andreassen, G. S. Harrits, U. Hvidman, M. V. Kristensen & S. Munkedal. (2019) Evaluering af skolepuljen: 1. midtvejsrapport, Aarhus, Trygfondens Børneforskningscenter.
- Andersen, S. C., L. M. Andreassen, G. S. Harrits, C. G. Jensen, S. Y. Junge & S. Munkedal. (2020) Evaluering af skolepuljen: 2. midtvejsrapport, Aarhus, Trygfondens Børneforskningscenter.
- Andersen, M. M., R. Dilling-Hansen & A. V. Hansen. (2020). Expanding the Concept of Social Impact Bonds, bringses i *Journal of Social Entrepreneurship*, doi: 10.1080/19420676.2020.1806100
- Andersen, L. B., C. B. Jacobsen, M. L. F. Jakobsen, T. Pallesen & S. Serritzslew. (2017). *Økonomiske incitamenten i den offentlige sektor*, 2. udgave. København: Hans Reitzels Forlag.
- Andersen, S. C. & H. S. Nielsen. (2020). Learning From Performance Information, *Journal of Public Administration Research and Theory*, 30(3), 415-431. doi: 10.1093/jopart/muz036.
- Ankestyrelsen. (2020). Tilsynsudtalelse om kommuners brug af private konsulenter på det sociale område. J.nr. 19-62648, J.nr. 19-62648, 24. juni 2020.
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis, *Educational Researcher*, 36(5), 258-267. doi: 10.3102/0013189x07306523
- Balogun, J. and Johnson, G. 2004. "Organizational restructuring and middle manager sensemaking", *Academy of Management Journal*, 47(4), 523-549. <https://doi.org/10.5465/20159600>
- Beuchert, L. V. & A. B. Nandrup. (2018). The Danish national tests at a glance, *Danish Journal of Economics*, (1), 1-37. <https://doi.org/10.2139/ssrn.2996431>

Biggs, J. B. (1995). Assumptions Underlying New Approaches to Educational Assessment, *Curriculum Forum*, 4(2), 1-22.

Bjerrild, S. (2017). Dumpekarakter til Løkkes præmiepulje, *Folkeskolen*, No. 20, pp. 14-19.  
Blankenberg, B. & A. Phillips. (2016). Performance Funding in Illinois Higher Education: The Roles of Politics, Budget Environment, and Individual Actors in the Process, *Educational Policy*, 30(6), 884-915. doi:10.1177/0895904814556748.

Booher-Jennings, J. (2005). Below the Bubble: 'Educational Triage' and the Texas Accountability System, *American Educational Research Journal*, 42(2), 231-268. doi: 10.3102/00028312042002231.

Brill, F. J., K. Grayson, L. Kuhn & S. O'Donel. (2018). *What Impact Does Accountability have on Curriculum, Standards and Engagement in Education? A Literature Review*. Berkshire: National Foundation for Educational Research.

Bukh, P. N. & K. S. Christensen. (2020a). Tildelingsmodeller og effektbaserede principper i den offentlige sektor, *Samfundslederskab i Skandinavien*, 35(4), 207-249.  
<https://doi.org/10.22439/sis.v35i4.6019>

Bukh, P. N. & K. S. Christensen. (2020b). Effektbaserede ressourcetildelingsmodeller for forløbs- og periodebaserede indsatser, *Samfundslederskab i Skandinavien*, 35(6). <https://doi.org/10.22439/sis.v35i6.6111>

Bukh, P. N., K. S. Christensen & A. K. Svanholt. (2020a). Omkostningsbevidsthed i socialt arbejde: fra spændingsfelt til professionalisme, *Tidsskrift for Professionsstudier*. 31.

Bukh, P. N., K. S. Christensen & A. K. Svanholt. (2020b). Making sense of cost-consciousness in social work, *Qualitative Research in Accounting & Management*.  
<https://doi.org/10.2139/ssrn.3721062>

Bukh, P. N. & A. K. Svanholt. (2019). Styring i spændingsfeltet mellem vilde problemer og stramme budgetter, *Samfundslederskab i Skandinavien*, 34(3), 229-266.  
<https://doi.org/10.22439/sis.v34i3.5812>

Bukh, P. N. & A. K. Svanholt. (2020). Empowering Middle Managers in Social Services Using Management Control Systems, *Journal of Public Budgeting, Accounting & Financial Management*, 32(2), 267-289. doi:10.1108/JPBAFM-06-2019-0096.  
<https://doi.org/10.1108/JPBAFM-06-2019-0096>

Burgess, S. M., C. Propper, H. Slater & D. Wilson. (2005). Who Wins and who Loses from School Accountability? The Distribution of Educational Gain in English Secondary Schools, Centre for Economic Policy Research, Discussion Paper No. 5248.

Børne- og Undervisningsministeriet (2017a). Pulje til løft af fagligt svage elever i folkeskolen, København, Danmark. <https://www.uvm.dk/puljer-udbud-og->

prisuddelinger/puljer/puljeoversigt/tidligere-udmeldte-puljer/grundskole/pulje-til-loeft-af-fagligt-svage-elever-i-folkeskolen-skolepuljen. (tilgået 20. juli 2020)

Børne- og Undervisningsministeriet (2017b). Program for elevløft. <https://www.uvm.dk/folkeskolen/laeringskonsulenterne/pulje-til-elevloeft/program-for-elevloeft>. (tilgået 20. juli 2020)

Cheng, L. (1999). Changing Assessment: Washback on Teacher Perceptions and Actions, *Teaching and Teacher Education*, 15(3), 253-271. doi: 10.1016/S0742-051X(98)00046-8. [https://doi.org/10.1016/S0742-051X\(98\)00046-8](https://doi.org/10.1016/S0742-051X(98)00046-8)

David, J. (2011). What students need to learn: High-stakes testing narrows the curriculum, *Educational Leadership*, 68(6), 78-80.

Deming, D. J., S. Cohodes, J. Jennings, & C. Jencks. (2016). School Accountability, Postsecondary Attainment, and Earnings, *Review of Economics and Statistics*, 98(5): 848-862. doi: 10.1162/REST\_a\_00598. [https://doi.org/10.1162/REST\\_a\\_00598](https://doi.org/10.1162/REST_a_00598)

Donaldson, M. L. & J. L. Papay. (2015). Teacher evaluation for accountability and development. I *Handbook of Research in Education Finance and Policy*, H. F. Ladd & M. E. Goertz (red.), 174-193. New York: Routledge

Dougherty, K.J., S. M. Jones, H. Lahr, R. S. Natow, L. Pheatt & V. Reddy, V. (2016), *Performance Funding for Higher Education*. Baltimore: John Hopkins University.  
Edminston, D. & A. Nicholls. (2018). Social Impact Bonds: The Role of Private Capital in Outcome-Based Commissioning, *Journal of Social Policy*, 47(1), 57-76. doi: 10.1017/S0047279417000125

Edwards, N. R. & D. L. Mindrila. (2019). Improving Graduation Rates: Legitimate Practices and Gaming Strategies. *Education Policy Analysis Archives* 27 (41), 1-25. doi: 10.14507/epaa.27.4222.

Figlio, D. N., & H. F. Ladd. (2015). School Accountability and Student Achievement. In *Handbook of Research in Education Finance and Policy*, edited by H. F. Ladd and M. E. Goertz, 194-210. New York, NY: Routledge

Flick, U. (2009), *An Introduction to Qualitative Research*, 4th edition, London, Sage Publications.

Ganon-Shilon, A. & C. Schechter. (2018). School Principals' Sensemaking of Their Leadership Role during Reform Implementation. *International Journal of Leadership in Education* 22(3), 279-300. doi: 10.1080/13603124.2018.1450996.

Gioia, D. A. & K. Chittipeddi. (1991), Sensemaking and sensegiving in strategic change initiation, *Strategic Management Journal*, 12(6), 433-448. <https://doi.org/10.1002/smj.4250120604>

Gioia, D. A. & J. B. Thomas. (1996), Identity, Image, and Issue Interpretation: Sensemaking During Strategic Change in Academia, *Administrative Science Quarterly*, 41(3), 370-403. <https://doi.org/10.2307/2393936>

Gunnulfsen, A. E. & A. Roe. (2018). Investigating teachers' and school principals' enactments of national testing policies: A Norwegian stud. *Journal of Educational Administration*, 56(3), 332-349. doi: 10.1108/JEA-04-2017-0035. <https://doi.org/10.1108/JEA-04-2017-0035>

Hammer, S. & J. Høpner. (2019). *Meningsskabelse, organisering og ledelse: En introduktion til Weicks univers*, 2. udgave, København: Samfundslitteratur.

Hardy, I., V. Reyes, & M. O. Hamid. (2019). Performative Practices and 'Authentic Accountabilities': Targeting Students, Targeting Learning? *International Education Journal: Comparative Perspectives* 18(1), 20-33.

Herbst, M. (2008). *Financing Public Universities: The Case of Performance Funding*. Dordrecht: Springer.

Holm, L. & K. B. Kousholt. (2019). Beyond Washback Effect: A Multi-disciplinary Approach Exploring How Testing Becomes Part of Everyday School Life Focused on the Construction of Pupils' Cleverness, *Annual Review of Critical Psychology*, 16, 917-952.

Jacobsen, R., & R. Rothstein. (2015). Educational Goals: A Public Perspective. I *Handbook of Research in Education Finance and Policy*, H. F. Ladd and M. E. Goertz (eds.). New York: Routledge.

Jennings, J. L. & J. M. Bearak (2014). 'Teaching to the Test' in the NCLB Era: How Test Predictability Affects our Understanding of Student Performance. *Educational Researcher* 43(8), 381-389. doi: 10.3102/0013189X14554449.

Jongbloed, B., & Vossensteyn, H (2016). University Funding and Student Funding: International Comparisons. *Oxford Review of Economic Policy*, 32(4), 576-595. doi: 10.1093/oxrep/grw029

Kelley, C. & J. Protsik. (1999). Risk and Reward: Perspectives on the Implementation of Kentucky's School-Based Performance Award Program, *Educational Administration Quarterly*, 33(4): 474-505. doi: 10.1177/0013161X97033004004.

Koyama, J. (2014). Principals as Bricoleurs: Making Sense and Making Do in an Era of Accountability, *Educational Administration Quarterly* 50(2), 279-304. doi: 10.1177/0013161X13492796.

Kvale, S. (1996). *InterViews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks: Sage.



- Ladd, H. F. & R. P. Walsh. (2002). Implementing Value-Added Measures of School Effectiveness: Getting the Incentives Right, *Economics of Education Review*, 21(1), 1-17. doi: 10.1016/S0272-7757(00)00039-X.
- Lambersky, J. (2016). Understanding the Human Side of School Leadership: Principals' Impact on Teachers' Moral, Self-Efficacy, Stress, and Commitment, *Leadership and Policy in Schools*, 15(4), 379-405. doi: 10.1080/15700763.2016.1181188.
- Lavy, V. (2007). Using Performance-Based Pay to Improve the Quality of Teachers, *The Future of Children* 17(1), 87-109. doi: 10.1353/foc.2007.0007.
- Lee, J. D. & O. Medina. (2018). Results-Based Financing in Education: Financing Results to Strengthen Systems. Results in Education for All Children (REACH), World Bank Group. <https://doi.org/10.1596/31250>
- Lowenhaupt, R., J. P. Spillane & T. Hallett. (2016). Education Policy in Leadership Practice: "Accountability Talk" in Schools, *Journal of School Leadership*, 26(5), 783-810. doi: 10.1177/105268461602600503.
- Li, A. Y., & Ortagus, J. C. (2019), Raising the Stakes: Impacts of the Complete College Tennessee Act on Underserved Student Enrollment and Sub-baccalaureate Credentials, *The Review of Higher Education*, 43(1), 295-333. doi: 10.1353/rhe.2019.0097
- Lukka, K. & E. Vinnari. (2014), Domain theory and method theory in management accounting research, *Accounting, Auditing & Accountability Journal*, 27(8), 1308-1338. <https://doi.org/10.1108/AAAJ-03-2013-1265>
- Madau, G. (1988). The influence of testing on the curriculum. I L. Tanner, red. *The politics of reforming school administration*. London: The Falmer Press, 83-121.
- Maitlis, S. (2005). The social processes of organizational sensemaking, *Academy of Management Journal*, 48(1), 21-49. <https://doi.org/10.5465/amj.2005.15993111>
- Maitlis, S. & M. Christianson. (2014). Sensemaking in organizations: taking stock and moving forward, *The Academy of Management Annals*, 8(1), 57-125. Maitlis, S. & M. Christianson. (2014). Sensemaking in organizations: taking stock and moving forward, *The Academy of Management Annals*, 8(1), 57-125. <https://doi.org/10.5465/19416520.2014.873177>
- Maitlis, S. & T. B. Lawrence. (2007). Triggers and Enablers of sensegiving in Organizations, *The Academy of Management Journal*, 50(1), 57-84. <https://doi.org/10.5465/amj.2007.24160971>
- Martin-Rios, C. (2016), Innovative management control systems in knowledge work: a middle manager perspective, *Journal of Management Control*, 27, 181-204. <https://doi.org/10.1007/s00187-015-0228-8>

- Miles, M. B., A. M. Huberman & J. Saldaña. (2014). *Qualitative Data Analysis: A Methods Sourcebook*, 3. udgave. Thousand Oaks, Sage.
- Neal, D. & D. W. Schanzenbach. (2010). Left Behind by Design: Proficiency Counts and Test-Based Accountability, *The Review of Economics and Statistics*, 92(2), 263-283. doi: 10.1162/rest.2010.12318.
- Paletta, A., E. Ferrari & G. Alimehmeti. (2019). How Principals use a new Accountability System to promote Change in Teacher Practices: Evidence from Italy. *Educational Administration Quarterly*, 46, 25-48. doi: 10.1177/0013161X19840398.
- Popham, W. J. (2001). Teaching to the Test?, *Educational Leadership*, 58(6), 16-21.
- Regeringen. (2016). Regeringens politik A til Å: Folkekolen, <https://www.regeringen.dk/regeringens-politik-a-å/folkeskolen/> (tilgået 20. juli 2020)
- Rutledge, S. A., D. N. Harris & W. K. Ingle. (2010). How principals 'bridge and buffer' the new demands of teacher quality and accountability: A mixed-methods analysis of teacher hiring, *American Journal of Education*, 116(2), 211-242. doi: 10.1086/649492
- Saldaña, J. (2016). *The Coding Manual for Qualitative Researchers*, 3. udgave. Los Angeles: Sage.
- Sandberg, J. & H. Tsoukas. (2015), Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development, *Journal of Organizational Behavior*, 36(2S). 6-32. DOI: 10.1002/job.1937
- Shirrell, M. (2016). New principals, accountability, and commitment in low-performing schools, *Journal of Educational Administration*, 54(5), 558-574. doi 10.1108/JEA-08-2015-0069
- Siverbo, S., M. Caker & J. Åkesson. (2019). Conceptualizing dysfunctional consequences of performance measurement in the public sector, *Public Management Review*, 21(12), 1801-1823 doi: 10.1080/14719037.2019.1577906
- Smith, P. C. (2007). *Formula Funding of Public Services*. Oxon: Routledge. <https://doi.org/10.4324/9780203013021>
- Sonenshein, S. (2010). We´re changing-or are we? untangling the role of progressive, regressive, and stability narratives during strategic change implementation, *Academy of Management Journal*, 53(3), 477-512. <https://doi.org/10.5465/amj.2010.51467638>
- Spillane, J. P., J. B. Diamond, P. Burch, T. Hallett, L. Jita & J. Zoltners. (2002). Managing in the Middle: School Leaders and the Enactment of Accountability Policy. *Educational Policy* 16(5), 731-762. doi: 10.1177/089590402237311.

Spillane, J. P. & A. W. Kenney. (2012). School Administration in a Changing Education Sector: The US Experience. *Journal of Educational Administration* 50, 541–561. doi: 10.1108/09578231211249817.

STIL. (2019). De socioøkonomiske referencer for grundskolekarakterer, Styrelsen for IT og læring, Børne- og Undervisningsministeriet. Tilgået fra <https://www.uvm.dk/-/media/filer/uvm/stat/pdf19/191025-metodenotat-bag-om-de-sociookonomiske-referencer-for-grundskolekarakterer.pdf> (tilgået 23. juli 2020)

Sundheds- og Ældreministeriet. (2019). Cirkulære om statsligt, resultatafhængigt tilskud i 2020 til regionernes sundhedsvæsen (Nærhedsfinansiering), CIR1H nr 10186 af 18/12/2019.

Undervisningsministeriet. (2013). Aftale om et fagligt løft af folkeskolen mellem regeringen (Socialdemokraterne, Radikale Venstre og Socialistisk Folkeparti), Venstre og Dansk Folkeparti. Tilgået fra <https://www.uvm.dk/folkeskolen/folkeskolens-maal-love-og-regler/politiske-oplaeg-og-aftaler> (tilgået 23. juli 2020)

Undervisningsministeriet. (2014). Improving the Public School: Overview of Reform of Standards in the Danish Public School (Primary and Lower Secondary Education), Copenhagen, Denmark. Tilgået fra <https://www.uvm.dk/-/media/filer/uvm/publikationer/engelsksprogede/2014-improving-the-public-schools.pdf>. (tilgået 20. juli 2020)

Undervisningsministeriet. (2018). Bekendtgørelse om pulje til løft af fagligt svage elever i folkeskolen, BEK no. 117 of 21/02/2018, København, Danmark.

Umbricht, M. R., F. Fernez & J. C. Ortagus. (2017). An Examination of the (Un)intended Consequences of Performance Funding in Higher Education. *Educational Policy* 31(5), 643–673. doi: 10.1177/0895904815614398.

Warner, M.E. (2013). Private finance for public goods: social impact bonds. *Journal of Economic Policy Reform*, 16(4), 303-319, doi: 10.1080/17487870.2013.835727

Weick, K.E. (1979). *The Social Psychology of Organizing*, 2. udgave, New York: McGraw-Hill.

Weick, K.E. (1995). *Sensemaking in Organizations*, Thousand Oaks: Sage.

Weick, K.E. (2001). *Making Sense of the Organization*. Oxford: Blackwell Publishing.

Weick, K.E. & K. M. Sutcliffe. (2003). Hospitals as cultures of entrapment: a re-analysis of the Bristol Royal Infirmary, *California Management Review*, 45(2), 73-84. <https://doi.org/10.2307/41166166>

de Wolf, I. F. & F. J. G. Janssens. (2007). Effects and Side Effects of Inspections and Accountability in Education: An Overview of Empirical Studies. *Oxford Review of Education*, 33(3), 379–396. doi: 10.1080/03054980701366207.

Yuan, K., V. N. Le, D. F. McCaffrey, J. A. Marsh, L. S. Hamilton, B. M. Stecher & M. G. Springer. (2013). Incentive Pay Programs do not Affect Teacher Motivation or Reported Practices: Results from three Randomized Studies. *Educational Evaluation and Policy Analysis*, 35(1), 3-22. doi: 10.3102/0162373712462625.

---

<sup>1</sup>Tak til James Høpner og Sverri Hammer for konstruktive og brugbare kommentarer til en tidligere version af artiklen. Desuden tak til Betina Sørensen for assistance med manuskriptet.

<sup>2</sup> PISA er en forkortelse for Programme for International Student Assessment, som er en OECD-undersøgelse, der sammenligner skoleelevers kompetencer på tværs af OECD-lande. Siden år 2000 har hensigten med PISA-undersøgelserne været at teste 9. klasses elevers kunnen inden for læsning, matematik og naturfag.

<sup>3</sup> Vi anvender de engelske begreber: *Low-stakes* hhv. *high-stakes* til at beskrive i hvor høj grad, der er noget på spil ved den anvendte styring. *Low-stakes* nationale tests indebærer, at resultaterne af de nationale tests har meget lille indvirkning på skole, lærere og elever, mens de eksempelvis ville havde været *high-stakes* tests, hvis lærerne enten opnåede en personlig belønning eller risikerede at blive fyret, hvis eleverne klarede sig henholdsvis godt eller dårligt.

<sup>4</sup> De nationale mål er: 1. Folkeskolen skal udfordre alle elever, så de bliver så dygtige, de kan, 2. Folkeskolen skal mindske betydningen af social baggrund i forhold til faglige resultater, og 3. Tilliden til og trivslen i folkeskolen skal styrkes blandt andet gennem respekt for professionel viden og praksis. Der er fire operative resultatmål for de nationale mål: "1. Mindst 80 pct. af eleverne skal være gode til at læse og regne i de nationale tests, 2. Andelen af de allerdygtigste elever i dansk og matematik skal stige år for år, 3. Andelen af elever med dårlige resultater i de nationale tests for læsning og matematik skal reduceres år for år, og 4. Elevernes trivsel skal øges." (Undervisningsministeriet 2013, s. 23)

<sup>5</sup> Teknisk set er der tale om en direkte effektbaseret tildeling med anvendelse af en effekttakst (præmiebeløbet) knyttet til opnåelsen af en minimumseffekt (jf. Bukh & Christensen 2020b). I artiklen anvender vi generelt betegnelsen 'bonus' fremfor præmiebeløb, da denne betegnelse har været udbredt i medierne såvel som blandt de skoler, vi interviewede.

<sup>6</sup> Washback effekt er et bredt anvendt begreb inden for skoleområdet. Det refererer til, hvordan tests påvirker, hvordan læseplaner designes mhp., hvad eleverne måles på, undervisningspraksis og læringsadfærd. Det er ikke indlysende, om washback-effekter er positive eller negative.

<sup>7</sup> Begrebet *sensemaking* er vanskeligt at oversætte præcist til dansk (jf. Hammer & Høpner 2019), men det kan oversættes til meningsskabelse. Vi anvender i et vist omfang begreber som f.eks. enactment uden at oversætte dem til dansk, mens vi andre steder har anvendt danske ord som "ledetråde" (clues) i stedet for det engelske ord. Ligeledes anvender vi ligesom Maitlis & Christianson (2014, p. 109) ordene "mening" og "forståelse" synonymt for at markere den mening, som sensemaking-processerne fører til, da mening ikke betyder, at det er "meningsfuldt", dvs. har værdi eller indebærer et formål.

<sup>8</sup> Skolepuljen, der formelt betegnes "Pulje til løft af fagligt svage elever i folkeskolen" (Børne- og Undervisningsministeriet 2017a; Undervisningsministeriet 2018) omtales nogle gange også som "elevløftspuljen". Ministeriet anvender i denne bekendtgørelse (Undervisningsministeriet 2018) betegnelsen 'præmiebeløb', hvilket ofte også er den betegnelse, der anvendes i medierne, se f.eks. Folkeskolens temanummer om skolepuljen d. 23. november 2017 (Bjerrild 2017).

---

<sup>9</sup> Karakteren beregnes på baggrund af de seks bundne prøver i dansk og matematik i 9. klasse, således at gennemsnittet beregnes for karaktererne i dansk (mundtlig, skriftlig fremstilling, retskrivning og læsning) og matematik (uden hhv. med hjælpemidler) hver for sig. Hvis eleven har over 4 i begge gennemsnit, er skolepuljens succeskriterium opfyldt.

<sup>10</sup> 'Socioøkonomisk' refererer til elevernes sociale og økonomiske baggrund, mens 'reference' indebærer, at tallet kan bruges som et sammenligningsgrundlag for skolens faktisk opnåede karakterer. STIL (2019) beskriver de tekniske detaljer vedrørende beregningen. Denne type mål, der internationalt går under betegnelsen "value-added measures" (Ladd & Walsh 2002), anvendes ofte (Donaldson & Papay, 2015, pp 180-182) som mål for kvaliteten af undervisning.

<sup>11</sup> Puljen til nærhedsfinansiering udgør samlet for regionerne 1,5 mia. kr. i 2020 og er baseret på den enkelte regions opfyldelse af 4 kriterier: (1) reduktion i antal sygehusforløb pr. borger, (2) reduktion i sygehusaktivitet pr. kroniker med KOL og/eller diabetes, (3) reduktion i andelen af somatiske indlæggelser, der fører til en akut genindlæggelse inden for 30 dage samt (4) stigning i andelen af virtuelle sygehusforløb. Ved at opfylde et af kriterierne modtager regionen 1/3 af finansieringen, ved to kriterier 2/3, og hvis tre eller alle fire er opfyldt, modtager regionen 100% af finansieringen (Sundheds- og Ældreministeriet 2019).

<sup>12</sup> Udover erfaringer med resultatbaseret finansiering inden for den offentlige sektor findes enkelte eksempler på, at danske kommuner anvender resultatbaserede betalingsmodeller i forhold til private leverandører. Erfaringerne er dog også her begrænsede, og der er lovgivningsmæssige komplikationer ifm. anvendelsen af resultatbaseret betaling på de klassiske velfærdsområder; se f.eks. Ankestyrelsen (2020). Herudover er der i Danmark en vis interesse for de såkaldte "Social Impact Bonds" (Warner 2013; Edminston & Nicholls 2018), og der arbejdes i regi af Den Sociale Investeringsfond på at udvikle danske erfaringer hermed. Selvom Social Impact Bonds er baseret på en resultatafhængig betaling, har de hidtidige erfaringer herhjemme været, at den resultatafhængige del enten udgør en meget beskedne del eller reelt set slet ikke indgår. Se også Andersen *et al.* (2020).