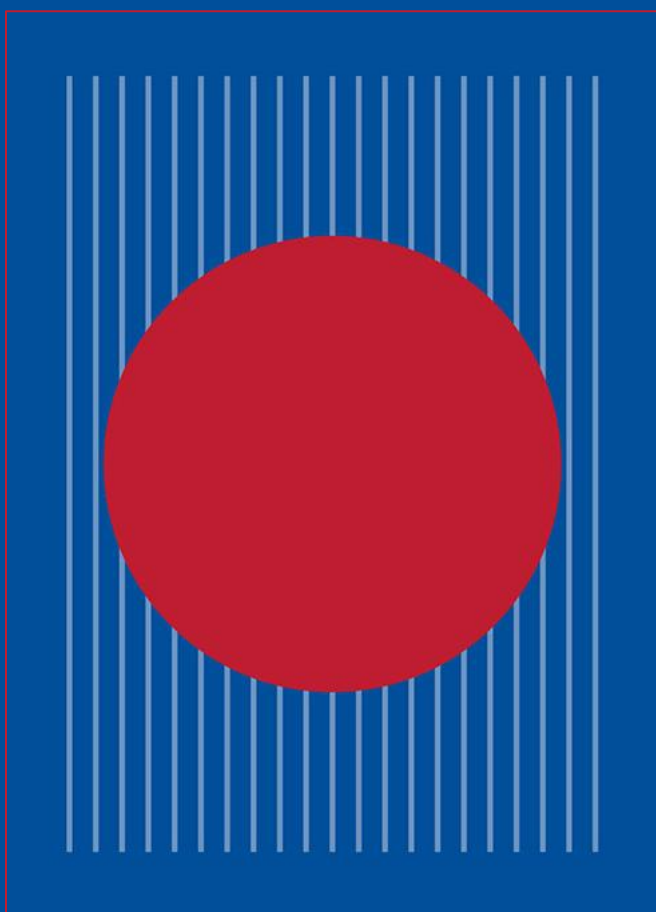


Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet



Rasmus Højbjerg Jacobsen, Asger Graa Andreasen, Maria Keilow, Karl Fritjof Krassel, Rune Vammen Lesner, Niels Peter Mortensen, Beatrice Schindler Rangvid og Mette Verner

VIVE

Afdækning af muligheder for udbredelse af SØM til indsatser på børne- og undervisningsområdet

© VIVE og forfatterne, 2022

e-ISBN: 978-87-7119-991-8

Projekt: 301805

Modelfoto: Ricky John Molloy/VIVE

Finansiering: Børne- og Undervisningsministeriet

VIVE

Det Nationale Forsknings- og Analysecenter for Velfærd

Herluf Trolles Gade 11

1052 København K

www.vive.dk

VIVEs publikationer kan frit citeres med tydelig kildeangivelse.



Forord

Denne rapport indeholder en afdækning af mulighederne for at udstrække principperne bag Den Socialøkonomiske Investeringsmodel (SØM) til seks konkrete indsatser inden for dagtilbuds- og skoleområdet. Rapporten er finansieret af Børne- og Undervisningsministeriet.

Dele af det projekt, der ligger til grund for rapporten, har været udført i samarbejde mellem Børne- og Undervisningsministeriet og VIVE, men det er VIVE, der har det faglige ansvar for rapportens indhold.

Afdækningen indeholder analyser af mulige virkningskæder for sammenhængen mellem en indsats og mulige økonomiske konsekvenser, en undersøgelse af evidens for effekten af de enkelte indsatser, et metodeforslag til beregning af økonomiske konsekvenser samt en opsamling for hver indsats med angivelse af mulighederne for at gennemføre beregninger.

Rapporten er skrevet af projektchef Rasmus Højbjerg Jacobsen (projektleder), analytiker Asger Graa Andreasen, forsker Maria Keilow, senioranalytiker Karl Fritjof Krassel, seniorforsker Rune Vammen Lesner, chefanalytiker Niels Peter Mortensen, seniorforsker Beatrice Schindler Rangvid og professor Mette Verner.

Rapporten har været læst og kommenteret af to eksterne reviewere, ligesom den har gennemgået VIVEs interne kvalitetssikring. Reviewerne takkes for gode og brugbare kommentarer.

Hans Hummelgaard

Forsknings- og analysechef for VIVE Effektmåling



Indholdsfortegnelse

Sammenfatning	6
---------------	---

1	Baggrund og indhold	20
1.1	Baggrund	20
1.2	Undersøgelsens indhold	21

2	Mekanismer for virkning – teoretiske sammenhænge mellem indsatser og effekter	26
2.1	Formulering af forandringsteoriene	28
2.2	Vidensunderbygning af de enkelte trin i folkeskoleforandringsteorien	29
2.3	Vidensunderbygning af de enkelte trin i dagtilbudsområdet	38

3	Tværgående metodeovervejelser	44
3.1	Betragtninger vedrørende modeltype	44
3.2	Modelteknisk beskrivelse	45

4	To-voksen-ordninger	56
4.1	Indsatsens indhold og afgrænsning	56
4.2	Mulige mekanismer for indsatsens virkning	57
4.3	Evidens for indsatsen	58
4.4	Modeludvikling	61
4.5	Ny evidens og ny data	65

5	Normering i dagtilbud	66
5.1	Indsatsens indhold og afgrænsning	66
5.2	Mulige mekanismer for indsatsens virkning	67
5.3	Evidens for indsatsen	68
5.4	Modeludvikling	70
5.5	Ny evidens og ny data	76

6	Udsat skolestart og omgængere	78
6.1	Indsatsens indhold og afgrænsning	78

6.2	Mulige mekanismer for indsatsens virkning	78
6.3	Evidens for indsatsen	80
6.4	Modeludvikling	84
6.5	Ny evidens og ny data	88
<hr/>		
7	Klassekvotient	90
7.1	Indsatsens indhold og afgrænsning	90
7.2	Mulige mekanismer for indsatsens virkning	90
7.3	Evidens for indsatsen	92
7.4	Modeludvikling	95
7.5	Ny evidens og ny data	100
<hr/>		
8	Ordblindeindsatser	101
8.1	Indsatsens indhold og afgrænsning	101
8.2	Mulige mekanismer for indsatsens virkning	101
8.3	Evidens for indsatsen	104
8.4	Modeludvikling	107
8.5	Ny evidens og nyt data	111
<hr/>		
9	Betydningen af uddannet personale i skoler og dagtilbud	113
9.1	Indsatsens indhold og afgrænsning	113
9.2	Mulige mekanismer for indsatsens virkning	114
9.3	Evidens for indsatsen	116
9.4	Modeludvikling	119
9.5	Ny evidens og ny data	126
<hr/>		
	Litteratur	127
<hr/>		
	Bilag 1 Procesbeskrivelse for formulering af forandringsteorierne før vidensunderbygning	135
<hr/>		
	Bilag 2 Dataundersøgelse	139
<hr/>		
	Bilag 3 Litteratursøgningsprotokol	144
<hr/>		

Sammenfatning

Denne rapport indeholder en afdækning af muligheden for at udbrede principperne bag Den Socialøkonomiske Investeringsmodel (SØM) til seks indsatser inden for Børne- og Undervisningsministeriets (BUVM) område. De seks indsatser, der på forhånd var udvalgt af BUVM, er: to-voksen-ordning, normering i dagtilbud, udsat skolestart/omgængere, klassestørrelse, ordblindeindsatser og uddannet personale i skole/dagtilbud.

Afdækningen har været afgrænset til at se på mulige effekter, der kan optræde inden for en 4-årig tidshorisont.

Sammenfatningen er struktureret således, at der først kommer en kort gennemgang af rapportens overordnede konklusioner og derefter følger en mere fyldestgørende sammenfatning af rapporten.

Konklusioner vedrørende indsatser

Afdækningen munder ud i nedenstående anbefalinger for de seks betragtede indsatser. For afklaring af begreber, uddybninger og metodeforslag henvises læseren til den resterende del af sammenfatningen og de enkelte kapitler.

Sammenfatningstabel 1 viser en samlet oversigt over anbefalingerne, og efter tabellen følger en kort beskrivelse af de enkelte indsatser. Tabellens søjler skal læses på følgende måde: "Anbefaling om videre arbejde med model" angiver VIVEs vurdering af, om det er muligt at gennemføre beregninger for indsatsen. "Anbefalede link-outcomes ved analyse med indirekte evidens" angiver, hvilke variable der bør inddrages som forbindelsesled mellem indsatsens direkte effekt og de mulige økonomiske konsekvenser, når der anvendes indirekte evidens. "Population" angiver den population, det er muligt at gennemføre beregninger for på baggrund af den fundne evidens.

Sammenfatningstabel 1 Oversigt over rapportens anbefalinger

Indsats	Anbefaling om videre arbejde mod model	Anbefalede link-outcomes ved analyse med indirekte evidens	Population
To-voksen-ordning	Ja, via indirekte evidens	DNT* læsning	Evt. kun mellemtrin, da evidens er målt i 6. klasse
Normering dagtilbud	Teknisk muligt, men der findes kun svag evidens for virkning, så det tilrådes ikke	Sprogvurdering eller DNT 2. klasse	Ældste børn i dagtilbud (børnehave)
Udsat skolestart/omgængere (opdeles)			
Udsat skolestart	Ja, med direkte evidens for nogle sammenhænge, men ellers via indirekte evidens	Sprogvurdering fra dagtilbud, hvis population er stor nok, ellers DNT læsning 2. klasse, matematik 3. klasse	Børn, som får udsat skolestart; evt. kan anvendes en bred population fra indskoling, hvis populationen af børn, som får udsat skolestart, er for lille til at gennemføre statistiske beregninger
Omgængere	Nej, der er ikke fundet tilstrækkelig evidens		
Klassestørrelse	Muligt, men der er kun begrænset evidens for effekt	DNT læsning (2., 4., 6., 8. klasse)	Hele populationen af elever i grundskolen er relevant, men analysen kan underopdeles på aldersklasser
Ordblindeindsatser (opdeles)			
Lokale	Ja, med indirekte evidens og med forskellige effektstørrelser for forskellige indsats typer	DNT læsning (2., 4., 6., 8. klasse)	Svage læsere, fx afgrænset via resultater fra de nationale test
Tværgående/strukturelle**	Nej, der er ikke fundet evidens		
Uddannet personale (opdeles)			
Skoler	Muligt, men kun svag evidens for virkning	DNT 8. klasse eller 9. klasse afgangseksamen	Evt. kun udskoling, da dansk evidens påviser effekt her
Dagtilbud	Ja, med indirekte evidens, hvis opkvalificering anses for en del af indsatsområdet	Sprogvurdering eller DNT 2. klasse	Ældste børn i dagtilbud (børnehave)

Anm.: "Direkte evidens" betyder, at der er observeret en direkte årsagssammenhæng mellem indsatsen og den økonomiske konsekvens. "Indirekte evidens" betyder, at der er observeret en effekt af indsatsen på områder, som igen kan påvirke de økonomiske konsekvenser. Denne analyse involverer et forbindelsesled (link-outcome) mellem indsatsens effekt og den økonomiske konsekvens.

Note: * DNT er forkortelse for de nationale test. Der tages forbehold for anvendelsen af de nationale test for skoleåret 2019/20 på grund af covid-19-pandemien.

** Tværgående/strukturelle ordblindeindsatser hentyder til de tilfælde, hvor der fx er tale om en omlægning af hele kommunens ordblindetilbud.

Kilde: Egen tilvirkning.

I de tilfælde, hvor det i tabellen anbefales at gå videre med modelarbejdet, skal det også besluttes, hvilke konsekvenser den økonomiske model skal medregne. Svaret på dette spørgsmål afhænger af, hvilket evidensniveau der ønskes for modellen, og af, i hvor høj grad der lægges vægt på de teoretiske sammenhænge mellem indsats og konsekvens. I Sammenfatningstabel 2 findes en oversigt over, hvilke områder der kan medregnes i modellen på baggrund af de forskellige mulige valg. Tabellen er opdelt på skoleindsatser og dagtilbudsindsatser, fordi datakilder og de derfra mulige økonomiske beregninger er ens på tværs af indsatserne på hvert af de to områder. Derudover er de generiske forandringsteorier, der også er et muligt kriterium for udvælgelse af konsekvenser, også ens for de to hovedområder.

Sammenfatningstabel 2 Mulige beregningsområder for økonomiske konsekvenser

Direkte evidens		Indirekte evidens	
		Fra kvalificerede forandringsteorier	Yderligere muligt i registerdata
Skoleindsatser	Mindre kriminalitet og forældres arbejdsudbud for udsat skolestart, ellers ikke, se Sammenfatningstabel 1	Skoler - omgængere - specialundervisning - skoleskift Forældre - arbejdsudbud Elever - mindre kriminalitet	Skoler - personalefravær - personaleomsætning Uddannelse - ungdomsuddannelser Sundhed - primær sundhedssektor - hospitaler Sociale indsatser - forebyggende foranstaltninger - anbringelser - ofre for kriminalitet
Dagtilbudsindsatser	Nej, se Sammenfatningstabel 1	Ansatte - arbejdsudbud Børn - påvirkning af skolegang, jf. ovenfor	Skoler - udsat skolestart Sundhed - primær sundhedssektor - hospitaler Sociale indsatser - forebyggende foranstaltninger - anbringelser - ofre for kriminalitet

Anm.: Datakilderne er nærmere beskrevet i Bilagstabel 2.1 og Bilagstabel 2.2. De generiske kvalificerede forandringsteorier kan ses i kapitel 2.

Kilde: Egen tilvirkning.

Tabellen viser, at hvis der ønskes direkte evidens for sammenhænge, er det meget få områder, der kan medtages.

Med indirekte evidens er der flere muligheder, men inden for dette niveau skal det også besluttes, hvor meget vægt man fx vil lægge på muligheden for sammenhænge, der er understøttet af de kvalificerede forandringsteorier i forhold til en model, der dækker mere bredt og dermed også medtager sammenhænge, der ikke er understøttet af de kvalificerede forandringsteorier. Generelt viser rapportens analyse med kvalificerede forandringsteorier, at de virkningskæder, der kan understøttes, først og fremmest fører til økonomiske konsekvenser inden for områderne forældres arbejdsudbud, børnenes brug af specialundervisning samt kriminalitet, jf. søjlen herom i Sammenfatningstabel 2. Det har ikke været muligt at finde kvalificerede virkningskæder for øvrige potentielle økonomiske konsekvenser.

Hvorvidt en økonomisk konsekvens er understøttet af de kvalificerede forandringsteorier eller ej, betyder ikke noget for de modeltekniske metoder til beregning af en mulig konsekvens, da der i forandringsteorierne alene er tale om teoretiske sammenhænge.

Ved brug af indirekte evidens kan modellen bygges op på to måder. Den ene måde anvender evidens fra litteraturen for sammenhængen mellem link-outcome og økonomisk konsekvens, mens den anden etablerer denne sammenhæng via en statistisk estimering. Evidensen fra litteraturen kan fx bestå af en sammenhæng mellem læring og brug af specialundervisning i skolen eller mellem skolebørns trivsel og deres brug af sundhedsydelse. Fordelen ved denne tilgang er, at de økonomiske konsekvenser vil være båret af viden om evidens, mens ulempen vil være, at der ikke nødvendigvis kan findes evidensgrundlag for alle de mulige områder for de økonomiske konsekvenser. Da det ikke som en del af dette projekt har været undersøgt, hvilken evidens der er på disse områder, fremgår de mulige konsekvenser ved denne metode ikke af Sammenfatningstabel 2.

De områder, der kan dækkes af mulige statistiske beregninger, er angivet i sidste søjle i Sammenfatningstabel 2, da denne mulighed i høj grad afhænger af de tilgængelige data. De statistiske beregninger vil kunne gennemføres ved brug af registerdata og først og fremmest data fra de nationale test og den nationale trivselsmåling. Disse beregninger vil ikke være en årsagssammenhæng, men alene en statistisk korrelation, hvor der er kontrolleret for en række socioøkonomiske baggrundsoplysninger. Til gengæld vil sådanne beregninger, som det fremgår af tabellen, kunne dække et bredt spektrum af mulige økonomiske konsekvenser. Ud over de områder, hvor der er observerbare økonomiske konsekvenser, kan også områder som fx opstart på en ungdomsuddannelse (for de ældste skoleelever) være en mulig konsekvens.

En mulighed for videre modeludvikling kunne derfor være at organisere efter to hovedspor:

- Gennemføre en litteraturafdækning med henblik på at tilvejebringe kvantitative estimater på sammenhængen mellem link-outcomes og de økonomiske konsekvenser
- Foretage en estimering af statistiske sammenhænge mellem link-outcomes og en bred liste af mulige konsekvenser (se Sammenligningstabel 2).

Resultatet af denne udvikling vil være to sæt af estimater, der sammen med estimaterne fra evidensafdækningen i denne rapport vil kunne anvendes til økonomiske beregninger. På baggrund af disse to hovedspor vil det fx være muligt at opbygge en model, hvor en bruger selv vil kunne vælge mellem forskellige evidensniveauer i sin beregning. Det vil eksempelvis kunne munde ud i en model, hvor en bruger ved en beregning for fx en to-voksen-ordning vil kunne vælge, om der kun skal medtages nogle områder, der er evidens for, eller om et bredt spektrum af områder, der er statistisk beregnede, skal medtages. I begge tilfælde vil den grundlæggende evidens for indsatsens virkning stamme fra evidensundersøgelsen her i rapporten.

Nedenfor følger en kortfattet uddybning af indholdet i Sammenfatningstabel 1 og 2.

To-voksen-ordning

Det er ikke muligt at gennemføre modelberegninger baseret på direkte evidens for, at en to-voksen-ordning har direkte økonomiske konsekvenser. Men der kan gennemføres beregninger på basis af evidens for indsatsens virkning på elevers læring, der kan have efterfølgende økonomiske konsekvenser. Det anbefales, at der gennemføres en statistisk beregning af sammenhængen mellem elevernes resultater i læsning i de nationale test og de efterfølgende økonomiske konsekvenser.

Normering i dagtilbud

Det er muligt at gennemføre beregninger med direkte evidens for pædagogers sygefravær (arbejdsmarkedstilknytning) og for forældres arbejdsudbud, men der er tale om evidens af mindre god kvalitet. Det anbefales derfor i stedet at anvende indirekte evidens gennem effekten på børns læring målt ved sprogvurdering ved skolestart eller eventuelt resultater i dansk i de nationale test i 2. klasse. Da der skal anvendes resultater i skolen som link, kan beregningen kun gennemføres for de ældste børn i dagtilbud. De fundne effekter af forbedret normering i dagtilbud er ret små, og derfor må det forventes, at de økonomiske konsekvenser også vil være små.

Udsat skolestart og omgængere

Det anbefales ikke at igangsætte beregninger for omgængere, da der ikke er fundet evidens for, at indsatsen har en effekt. De fleste studier viser en meget begrænset eller ingen effekt.

For udsat skolestart er det muligt at gennemføre beregninger med direkte evidens for forældres arbejdsudbud. Hvis der skal foretages en beregning for brug af specialundervisning, skal der anvendes indirekte evidens igennem elevernes læring, hvor der kan anvendes resultater fra de nationale test i dansk i 2. klasse eller i matematik i 3. klasse til dannelse af de relevante effektstørrelser. Der findes en stor mængde litteratur, der påviser effekter af udsat skolestart.

Klassestørrelse

Det er muligt at gennemføre beregninger på indirekte evidensniveau for en reduktion i klassestørrelse gennem en effekt på elevernes læring. En sådan beregning skal i givet fald anvende resultater fra de nationale test eller fra 9. klasses eksamen som forbindelsesled. Hvis arbejdet igangsættes, skal man være opmærksom på, at der samlet set kun er meget begrænset evidens for, at klassestørrelse har positive effekter, da den samlede litteratur udviser meget forskellige effekter.

Ordblindeindsatser

Det er ikke muligt at gennemføre beregninger på direkte evidensniveau for ordblindeindsatser, da der ikke foreligger evidensundersøgelser hertil. Der er imidlertid evidens for, at flere forskellige indsatser har en effekt på elevernes læseevner, hvorfor det er muligt for disse indsatser at gennemføre beregninger på indirekte evidensniveau. Som forbindelsesled i disse beregninger skal anvendes resultater fra de nationale test i dansk eller resultater fra 9. klasses eksamen i dansk. Der er ikke fundet evidensundersøgelser, der dækker strukturelle ordblindeindsatser, hvorfor det ikke anbefales at gennemføre modelberegninger herfor.

Uddannet personale i dagtilbud og skole

Der kan ikke gennemføres beregninger på direkte evidensniveau, da der ikke foreligger evidensviden hertil. For brug af uddannet personale i skolen er der kun fundet meget små effekter, hvorfor det ikke anbefales at påbegynde denne beregning. For brug af uddannet personale i dagtilbud er det muligt at gennemføre en beregning med indirekte evidensniveau, da der er fundet en positiv effekt på børnenes sproglige kompetencer. Beregningen bør anvende resultater af sprogvurderingen eller resultater fra de nationale test i 2. klasse

som forbindelsesled. Derfor vil beregningen kun kunne gennemføres for de ældste børn i dagtilbud.

Yderligere evidens og data

På en række områder er det muligt at igangsætte et arbejde, der kan munde ud i større viden til økonomiske beregninger. Disse er nævnt i rapporten. Det gælder fx for arbejdet med at udbrede eksisterende danske effektmålinger til at dække også økonomiske effektmål samt et arbejde med at tilvejebringe individoplysninger for fx normering i dagtilbud og to-voksen-ordninger. Ved at igangsætte et sådant arbejde sikres det, at en kommende model gradvist vil kunne opkvalificeres og udvides, efterhånden som ny evidens og ny data kommer til rådighed.

Rapportens indhold

Denne rapport indeholder en afdækning af muligheden for at udbrede principperne bag Den Socialøkonomiske Investeringsmodel (SØM) til seks konkrete indsatser inden for Børne- og Undervisningsministeriets (BUVM) område. De seks indsatser var på forhånd udpeget af BUVM. Hensigten er, at afdækningen kan tjene som beslutningsgrundlag for det videre arbejde med at udvikle en økonomisk model for området.

Det er forventeligt, at indsatser inden for børn og undervisning – ud over den direkte effekt på fx læring eller trivsel – vil have en effekt på efterfølgende brug af offentlige ydelser og tilbud som fx brug af specialundervisning eller tilbøjeligheden til at starte på en ungdomsuddannelse. Der er også mulige effekter hos personalet og forældre.

Afdækningen er afgrænset til at se på følgende seks indsatser, der på forhånd var udvalgt af BUVM:

- Normeringer i dagtilbud
- Klassekvotient
- Betydningen af uddannet personale i dagtilbud og folkeskole
- To-voksenordning
- Ordblindeindsatser
- Udsat skolestart og omgængere.

En anden afgrænsning er, at der som en del af opdraget kun ses på effekter og økonomiske konsekvenser, der realiseres inden for en tidshorisont på ca. 4

år. Dette skyldes, at en model særligt er tænkt som et arbejdsredskab for kommunerne, hvor det forventes, at denne tidshorisont er den relevante.

Disse afgrænsninger har en klar betydning for afdækningens indhold. Først og fremmest betyder afgrænsningen til en 4-årig horisont, at en række potentielle økonomiske konsekvenser ikke er inddraget i afdækningen. Dette gælder fx effekter af bedre og mere uddannelse på arbejdsmarkedstilknytning og indtjening som voksen – en effekt, der først realiseres efter en del år. Et andet eksempel er mulige effekter af indsatser i dagtilbud og tidligt i grundskolen på fx frafald/gennemførelse på ungdomsuddannelse, hvilket også har en økonomisk konsekvens. Dermed betyder afgrænsningen, at det ikke vil være de samlede økonomiske konsekvenser, der indgår.

Indhold af afdækningen

Det grundlæggende undersøgelsesspørgsmål i afdækningen er:

Er det muligt at gennemføre en SØM-lignende beregning for de betragtede indsatser med de tilgængelige data og ved anvendelse af en tilstrækkelig god metode?

I analysen har vi opdelt undersøgelsesspørgsmålet i en række underspørgsmål:

- a) *Kan indsatsen afgrænses/optælles?* Dette spørgsmål er vigtigt, fordi en økonomisk model skal producere et kvantitativt bud på de økonomiske konsekvenser af en indsats.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* For at modellen giver et bud på de relevante konsekvenser af de betragtede indsatser, er det vigtigt at analysere, hvilke der kan være tale om. Typisk vil de seks indsatser, vi betragter her, ikke direkte have økonomiske konsekvenser (idet de omkostninger, der medgår ved driften af indsatsen, ikke regnes som en konsekvens), men derimod påvirker en række mellemtrin i årsagskæden. Disse mellemtrin kan fx bestå i trivsel og læring for deltagerne. Virkningskæderne for indsatserne studeres i rapporten ved hjælp af opstillede forandringsteorier.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Det er naturligvis vigtigt, at indsatsen virker, for at den kan have økonomiske konsekvenser (det er også den grundlæggende præmis for beregninger i SØM). I afdækningen indgår derfor en evidensundersøgelse, hvor det for hver indsats er blevet undersøgt, om (i) indsatsen virker på de områder, hvor den er tænkt til at virke (først og fremmest læring og trivsel), samt om (ii) der er evidens for, at indsatsen har direkte effekt på de områder, hvor der er målelige økonomiske konsekvenser.

- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Med baggrund i svarene på de tre foregående spørgsmål er det muligt for hver indsats at give et svar på, om der kan gennemføres modelberegninger ud fra det nuværende vidensniveau.

Ved besvarelsen af spørgsmål d) arbejder vi med en inddeling i tre beregningstyper, når vi omtaler den mulige beregning af økonomiske konsekvenser.

Det øverste niveau er *direkte evidens-beregning*, som betyder, at der findes kausal evidensviden for en sammenhæng mellem en udført indsats og en økonomisk konsekvens. Dette kunne fx være tilfældet, hvis der er påvist en sammenhæng mellem klassekvotienten og elevernes brug af specialundervisning.

Det næste niveau er *indirekte evidens-beregning*, der betyder, at der ikke findes evidens for den direkte sammenhæng, men at der kan udledes en sammenhæng via et mellemlid. Dette kunne fx være, hvis der er evidens for, at udsat skolestart virker positivt på elevers læring, og der samtidig er evidens for, at forøget læring betyder mindre brug af specialundervisning. Det er også muligt at kombinere evidens fra litteraturen med egne statistiske beregninger. Det endelige resultat af beregningen vil være en økonomisk konsekvens, også selvom det anvendte link-outcome ikke nødvendigvis kan opgøres monetært.

Det nederste niveau er *beregning af statistisk sammenhæng*, som betyder, at der kan påvises en statistisk sammenhæng mellem en indsats og en økonomisk konsekvens, uden at denne sammenhæng nødvendigvis er en årsags-sammenhæng.

Valget af model vil afhænge af, hvilket evidensniveau der ønskes. Generelt vil en model, der kun baserer sig på direkte evidens og/eller indirekte evidens, kunne tilbyde beregninger, der i nogen grad kan tilskrives årsagssammenhænge, fordi den anvender kausale estimater. Det vil give den fagligt stærkeste sammenhæng mellem indsats og resultater af beregningen. Til gengæld vil en sådan model kun kunne dække relativt få områder, fordi der ikke nødvendigvis findes evidensunderbygning for alle områder. Det kan medføre, at de mulige økonomiske konsekvenser ikke kan opgøres særligt præcist. Omvendt vil beregninger fra en model, der alene baserer sig på statistiske sammenhænge, ikke kunne tilskrives årsagssammenhænge og dermed ikke være en kausal model. Derimod vil modellen kunne dække mange flere områder, fordi der ikke vil være samme krav til, at der skal eksistere evidens for sammenhænge. Den eksisterende SØM-model er baseret på en kombination af viden om effekt af indsatser og statistiske sammenhænge mellem et succesfuldt forløb og efterfølgende økonomiske konsekvenser.

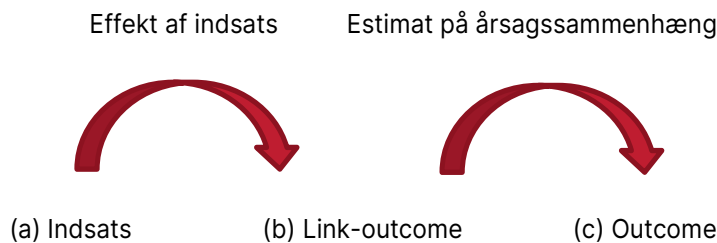
Rapportens kapitel 3 indeholder en grundig metodegennemgang af, hvordan beregninger kan foretages med anvendelse af henholdsvis direkte og indirekte

evidens. Metoden ved direkte evidens består i direkte at anvende det tilgængelige estimat for sammenhængen mellem indsatsen og den efterfølgende økonomiske konsekvens.

Ved indirekte evidens er metoden illustreret i Sammenfatningsfigur 1. I tilfælde, hvor en direkte sammenhæng mellem indsatsen (a) og det relevante outcome (c) ikke kan etableres, vil det være fordelagtigt at benytte et link-outcome (b). Denne situation kan opstå, hvis fx indsatsen sænket klassekvotient evalueres ved testresultater, men det relevante outcome-mål er brug af specialundervisning. Testresultatet bliver i dette tilfælde benyttet som link-outcome ved at etablere en årsagssammenhæng mellem testresultater og det ønskede outcome.

Sammenfatningsfigur 1

Beregning af effekt ved brug af link-outcome



Estimater fra statistiske beregninger på data kan udnyttes på to måder til modellen. For det første kan de anvendes til at kombinere med metoden i Sammenfatningsfigur 1, hvor sammenhængen mellem (b) og (c) kan estimeres. Dette vil ikke nødvendigvis resultere i et kausalt estimat, men vil kunne supplere i de tilfælde, hvor evidens ikke kan findes. For det andet kan der gennemføres direkte beregning af korrelationen mellem at have modtaget en indsats og efterfølgende økonomiske konsekvenser. Dette vil ikke være en årsagssammenhæng, og en sådan beregning vil kun kunne gennemføres, hvis det er muligt at identificere de børn, der har modtaget indsatsen, i data.

Gennemgang af de seks indsatser

Nedenfor gennemgås kort svarene på spørgsmålene a-d) for hver af de seks indsatser. For mere detaljerede svar henvises til rapportens kapitler 4-9.

To-voksen-ordning

- a) *Kan indsatsen afgrænses/optælles?* Indsatsen kan fx afgrænses ved at optælle de børn, der har været omfattet af to-voksen-ordning og omregne til fuldtidspersoner, dvs. svarende til, at barnet har været omfattet af ordningen hele skoleåret.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formodentlig går gennem børnenes læring og trivsel og kan have økonomiske konsekvenser på forældres arbejdsudbud, brugen af specialundervisning og begået kriminalitet.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Der er en meget begrænset litteratur, men et dansk studie har fundet en signifikant virkning på resultater i dansk ved et forsøg i 6. klasse. Samme studier viser også nul effekt på andre effektmål.
- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan gennemføres beregninger med indirekte evidens ved hjælp af link gennem resultatet i de nationale test i læsning. Beregninger baseret på direkte evidens eller en ren statistisk sammenhæng kan ikke gennemføres, da evidensgrundlag og datatilgængelighed ikke tillader dette.

Normering i dagtilbud

- a) *Kan indsatsen afgrænses/optælles?* Indsatsen afgrænses ved at optælle de børn, der har været omfattet af ændret normering, og omregne til fuldtidsækvivalenter (ved eksempelvis en indsats, hvor normeringen, altså hvor meget personale der er til stede pr. barn, ændres halvdelen af dagen for 12 børn; det svarer til, at 6 model-børn er omfattet af ændringen hele dagen, når der regnes i fuldtidspersoner).
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formodentlig går gennem bl.a. børnenes skoleparticipation og trivsel, større forældretilfredshed, mindre sygefravær hos personalet, der alle kan have økonomiske konsekvenser for dagtilbudets udgifter, forældres arbejdsudbud samt børnenes brug af ydelser i skolen.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Litteraturen tyder på, at der er en mindre effekt på børns ordforråd, men generelt er evidensen på området sparsom og blandet.

- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan gennemføres beregninger med indirekte evidens ved hjælp af link gennem resultatet sprogvurderinger eller eventuelt de nationale test i læsning i 2. klasse. Derfor kan beregningen kun gennemføres for de ældste børn i dagtilbud. Da de fundne effekter af forbedret normering i dagtilbud er ret små, må det dog forventes, at de økonomiske konsekvenser også vil være små målt pr. person. Det kan dog godt blive til betydelige effekter på aggregeret niveau, fordi den betragtede population er stor. Det er kun muligt at gennemføre en beregning baseret på statistisk sammenhæng, såfremt data for normering i dagtilbud indgår på aggregeret niveau og ikke på individniveau, fordi data på individniveau ikke foreligger. Dette vil alt andet lige gøre resultaterne mere usikre.

Udsat skolestart/omgængere

- a) *Kan indsatsen afgrænses/optælles?* Udsat skolestart betyder, at et barn starter et år senere i skolen, hvorved barnet bliver et ekstra år i børnehaven. For omgængere er den mest relevante indsatsdefinition, at eleven modtager undervisningen på et klassetrin to gange. Optællingsenheden kan her være, at barnet modtager et ekstra års undervisning i det samme pensum.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formodentlig især går gennem børnenes skoleparathed, mindre sygefravær hos personalet samt større forældretilfredshed og bedre børnetrivsel (på grund af bedre match mellem forudsætninger og skolegang) og ad denne vej kan have økonomiske konsekvenser for dagtilbuddets udgifter, forældres arbejdsudbud samt børnenes brug af ydelser i skolen, hvor virkningskæden går gennem børnenes læring.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Der er ikke meget evidens for virkningen af at gå en klasse om, men en relativt stor litteratur tyder på, at der er positive effekter af udsat skolestart.
- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan gennemføres beregninger med direkte evidens på forældres arbejdsudbud samt kriminalitet. Hvis der ønskes beregninger for brug af specialundervisning, skal der anvendes indirekte evidens ved hjælp af link gennem resultater fra de nationale test i læsning i 2. klasse eller matematik i 3. klasse.

Klassekvotient

- a) *Kan indsatsen afgrænses/optælles?* I forskningslitteraturen skelnes mellem klassekvotient og klassestørrelse, men i denne modelsammenhæng anser vi de to for at være ækvivalente. Ændringer i klassestørrelse kan optælles ved at se på, hvor mange børn der er dækket af ændringen i hvor stor en del af skoletiden.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formodentlig går gennem børnenes læring og trivsel og kan have økonomiske konsekvenser på forældres arbejdsudbud, brugen af specialundervisning, færre omgængere og skoleskift samt mindre begået kriminalitet.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Der er en stor mængde litteratur om betydningen af klassestørrelser, men resultaterne er blandede. En del studier finder nuleffekter, og mange finder positive effekter af mindre klassestørrelse. Effekterne er ofte små, men enkelte studier finder også effekter af betydelig størrelse. Et nyere dansk studie finder kun meget små effekter.
- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan ikke gennemføres beregninger baseret på direkte evidens, men der kan gennemføres beregninger med indirekte evidens ved anvendelse af link for læring som fx resultater i de nationale test eller eksamen i 9. klasse. De målte effekter er meget små og oftest insignifikante, men da populationen potentielt alle er grundskoleelever, kan den samlede effekt godt være betydelig.

Ordblindeindsatser

- a) *Kan indsatsen afgrænses/optælles?* I analysen har vi kvalitativt skelnet mellem lokale indsatser, typisk på en enkelt skole, og mere strukturelle indsatser, typisk for en hel kommune. Optællingen og afgrænsningen af indsatsen er dog ret ens, idet det blot drejer sig om at optælle det antal børn, som er dækket af forandringen, hvad enten den er lokal eller mere strukturel.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formodentlig går gennem børnenes læring og trivsel og kan have økonomiske konsekvenser på forældres arbejdsudbud, brugen af specialundervisning, færre omgængere og skoleskift samt mindre begået kriminalitet.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Der er ikke fundet litteratur, der påviser virkning af strukturelle ændringer. Til gengæld er der evidens for, at bestemte undervisningsmetoder har effekt på børnenes læseevner. Der er en ret stor litteratur, der evaluerer

indsatser over for ordblinde, og generelt er der en del positive effekter af de målte metoder.

- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan ikke gennemføres beregninger baseret på direkte evidens, men der kan godt gennemføres beregninger med indirekte evidens ved anvendelse af link for læsning som fx resultater i de nationale test i dansk eller eksamen i dansk i 9. klasse. Disse beregninger bør gennemføres, sådan at der kan tillades forskellige effektstørrelser, alt efter hvilken specifik indsats der regnes på.

Uddannet personale i skole og dagtilbud

- a) *Kan indsatsen afgrænses/optælles?* Uddannet personale kan enten bestå i, at det er personale med formelle uddannelseskompetencer, som fx undervisningskompetence i fag i skolen – eller i, at der er tale om opkvalificeringsforløb af personalet. I begge tilfælde gælder det, at optælling består i at medregne de børn, der har været påvirket af ændringen. For lærere med undervisningskompetence i faget, de underviser i, kan man fx beregne indsatsen til antal børn, der fuldtid har modtaget undervisning af en lærer med undervisningskompetence i det fag, der undervises i. For dagtilbud kan man fx beregne ændringen i den tid, det gennemsnitlige barn har sammen med uddannet personale.
- b) *Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?* Forandringsteorier, der indeholder de kvalificerede virkningskæder, viser, at mekanismerne formentlig går gennem børnenes læring og trivsel og kan have økonomiske konsekvenser på forældres arbejdsudbud, brugen af specialundervisning, færre omgængere og skoleskift samt mindre begået kriminalitet. I dagtilbuddet går mekanismen bl.a. gennem børnenes trivsel og læring til effekter på skoleparathed og anvendelse af skoleydelse.
- c) *Er der evidens for, at indsatsen virker, og i givet fald på hvad?* Der er evidens for betydningen af uddannet personale. At lærerne har undervisningskompetence i det fag, der undervises i, har en mindre, men positiv effekt på læring, og i dagtilbud har danske lodtrækningsforsøg påvist en effekt på børnenes sproglige kompetencer.
- d) *Kan der gennemføres modelberegninger, og i givet fald hvordan?* Der kan ikke gennemføres beregninger baseret på direkte evidens, da der ikke foreligger evidens herfor. Der kan godt gennemføres beregninger med indirekte evidens. For skoleindsats kan anvendes link i form af resultater i de nationale test i 8. klasse eller eksamen i 9. klasse, da et dansk studie har vist evidens for virkningen her. For dagtilbud kan gennemføres beregning, hvor der anvendes link i form af resultater af sprogvurdering ved overgang til skole eller eventuelt resultatet i de nationale test i 2. klasse. Derfor kan beregningen kun gennemføres for de ældste børn i dagtilbud.

1 Baggrund og indhold

1.1 Baggrund

Den Socialøkonomiske Investeringsmodel (SØM) blev i 2016-2017 udviklet af VIVE og Incentive for Socialstyrelsen på baggrund af en bevilling på satspuljeaftalen i efteråret 2015. Den første modelversion blev offentliggjort i januar 2018, og i øjeblikket er version 2.3 tilgængelig til download på Socialstyrelsens hjemmeside.

Modellens formål er give et bud på de økonomiske konsekvenser over tid af sociale indsatser over for afgrænsede målgrupper af socialt udsatte voksne og børn og unge. Det har først og fremmest været tanken, at modellen kunne anvendes som supplement og vidensbase til at understøtte beslutninger om iværksættelse af specifikke indsatser over for specifikke lokale målgrupper. Det kan fx ske ved, at en kommune gennemfører en SØM-beregning for flere forskellige foreslåede indsatser og inddrager resultaterne herfra i den kommunale beslutningsproces sammen med de øvrige input, der er relevante for beslutningen.

Det er dog rimeligt at antage, at det ikke kun er velfærdsindsatser på det specialiserede socialområde, der medfører afledte økonomiske konsekvenser. Således må det også formodes, at indsatser inden for børn og undervisning – ud over den direkte effekt på fx læring eller trivsel – vil have en effekt på efterfølgende brug af offentlige ydelser og tilbud som fx specialundervisning, tilbøjeligheden til at starte på en ungdomsuddannelse og eventuelt også sundheds-tilbud. Derudover er der en klar påvist sammenhæng mellem et succesfuldt skole- og uddannelsesforløb og individets senere tilknytning til og fastholdelse på arbejdsmarkedet. Derfor giver det god mening at undersøge, om SØM's principper om estimering af økonomiske konsekvenser kan udstrækkes til børne- og undervisningsområdet.

På baggrund af ovenstående er det formålet med denne rapport at afdække, om og i givet fald hvordan principperne bag SØM kan udstrækkes til at dække indsatser på Børne- og Undervisningsministeriets (BUVM) område.

Afdækningen er på opdrag fra BUVM afgrænset til at se på følgende seks indsatser:

- Normeringer i dagtilbud
- Klassekvotient

- Betydningen af uddannet personale i dagtilbud og folkeskole
- To-voksenordning
- Ordblindeindsatser
- Udsat skolestart og omgængere.

En anden afgrænsning er, at der ses på effekter og økonomiske konsekvenser, der realiseres inden for en tidshorizont på ca. 4 år. Dette skyldes, at den påtænkte model særligt er tænkt som et arbejdsredskab for kommunerne, hvor det forventes, at denne tidshorizont er den relevante.

Disse afgrænsninger har en klar betydning for afdækningens indhold. Først og fremmest betyder afgrænsningen til en 4-årig horisont, at en række potentielle økonomiske konsekvenser ikke kan inddrages i afdækningen. Dette gælder fx arbejdsmarkedstilknytning og -indtjening som voksen – en effekt, der først realiseres efter en del år. Et andet muligt eksempel er effekten af indsatser i dagtilbud og tidligt i grundskolen på fx frafald/gennemførelse på ungdomsuddannelse, hvilket også medfører en økonomisk konsekvens.

1.2 Undersøgelsens indhold

En model til mulig beregning af økonomiske konsekvenser er en kompleks størrelse. Derfor belyser denne afdækning en række områder, der alle er med til at besvare det grundlæggende undersøgelsesspørgsmål:

Er det muligt at gennemføre en SØM-lignende beregning for de betragtede indsatser med de tilgængelige data og ved anvendelse af en tilstrækkelig god metode?

I analysen har vi opdelt undersøgelsesspørgsmålet i en række underspørgsmål, som vi præsenterer her.

Kan indsatsen afgrænses/optælles?

Dette spørgsmål er vigtigt, fordi en økonomisk model skal producere et kvantitativt bud på de økonomiske konsekvenser af en indsats. Det må alt andet lige forventes, at en stor indsats (dvs. en, der berører mange personer) har større samlede økonomiske konsekvenser end en lille indsats, hvorfor det er vigtigt, at denne forskel kan kvantificeres, hvis modellen skal regne rigtigt. Det kan kun lade sig gøre, hvis det er muligt at optælle de individer, beregningen skal foretages for.

Hvilke mekanismer/virkninger kan drive de økonomiske konsekvenser?

For at modellen giver et bud på de relevante konsekvenser af de betragtede indsatser, er det vigtigt at analysere, hvilke der kan være tale om. Typisk vil de seks indsatser, vi betragter her, medføre økonomiske konsekvenser en række mellemtrin i årsagskæden, hvor indsatserne først påvirker fx trivsel og/eller læring for deltagerne. Disse mellemtrin vil kun i begrænset omfang medføre egentlige økonomiske konsekvenser, men de kan efterfølgende fx påvirke behovet for specialundervisning eller sygefraværet hos personale eller forældre og derigennem give anledning til en økonomisk gevinst.

Afdækningen indeholder en analyse, hvor VIVE i samarbejde med BUVM har opstillet teoretiske modeller (forandrings teorier) for virningskæderne for de enkelte indsatser. Disse modeller er efterfølgende blevet kvalificeret af VIVE, så kun mekanismer med et litteraturgrundlag eller eventuelt andet grundlag er tilbage i de endelige modeller. Modellerne bidrager med et grundlag for at udvælge de økonomiske konsekvenser, der er relevante for hver indsats.

Er der evidens for, at indsatsen virker, og i givet fald på hvad?

Det naturligvis vigtigt, at indsatsen virker, for at den kan have økonomiske konsekvenser. På børne- og undervisningsområdet findes en stor volumen af litteratur, der har set på effekter af de enkelte indsatsområder. Det betyder, at det i visse tilfælde er muligt at finde estimater på effekten af en indsats – ikke bare på fx børns læring og trivsel, men også på områder som fx forældres arbejdsudbud eller børnenes senere kriminalitet. Disse direkte effekter kan altså etablere en kausal sammenhæng mellem indsatsen og den efterfølgende økonomiske konsekvens.

I afdækningen indgår der derfor en evidensundersøgelse, hvor det for hver indsats er blevet undersøgt, om indsatsen har en virkning ud fra gennemførte effektmålingsstudier. De målte effekter er inden for en række områder relateret til både ikke-økonomiske som fx elevers læring og trivsel og økonomiske som fx personalets sygefravær. Inden for projektets tidsramme har det ikke været muligt at gennemføre en omfattende systematisk litteratursøgning med screening af fundne studier. Derfor har vi i stedet gennemført en litteratursøgning, der først og fremmest har haft til hensigt at finde eventuelle allerede gennemførte systematiske litteraturoversigter. Hvis det ikke har kunnet lade sig gøre, har vi i stedet søgt efter anden litteratur om emnet, jf. nærmere nedenfor.

VIVE har gennemført litteraturgennemgangen for indsatserne to-voksen-ordning, normering i dagtilbud, klassekvotient og ordblindeindsatser, mens BUVM har gennemført litteraturgennemgangen for indsatserne omgængere og udsat skolestart samt betydningen af uddannet personale i skole og dagtilbud. BUVM har fulgt en

protokol udarbejdet af VIVE, og der har været løbende dialog og mulighed for afklaring af eventuelle spørgsmål.

VIVE har på baggrund af den fundne litteratur sammenfattet evidensen på de enkelte områder.

Ud over at give et generelt indblik i, om der er evidens for indsatsens virkning, udgør de centrale fundne evidensstudier også en selvstændigt input til de økonomiske beregninger i de tilfælde, hvor der er fundet evidens for indsatsernes virkning på økonomiske outcomes. Ved udvælgelse af studier fra evidensundersøgelsen til det mulige videre arbejde med modellen har VIVE lagt vægt på følgende kriterier:

- Studier af høj evidensmæssig kvalitet, dvs. studier med lodtrækningsdesign og generelt kontrolgruppedesign, vægter højest
- Studier udført på baggrund af data fra Danmark (eller eventuelt Norden) skal i videst mulig omfang indgå
- Der skeles ikke til studiernes resultater ved udvælgelsen.

Dette fører ikke nødvendigvis til en entydig udvælgelse, så der vil i nogle tilfælde også være tale om en skønsmæssig vurdering. Dette skyldes bl.a., at der er meget stor forskel på mængden af evidens, når man ser på tværs af de seks indsatser. For nogle kan man godt anvendes stringente regler for udvælgelse, mens det for andre ikke er muligt, fordi der er meget få studier. I disse tilfælde kan det være en fordel at medtage et studie, også selvom det ikke helt opfylder kriterierne.¹

Kan der gennemføres modelberegninger, og i givet fald hvordan?

Med baggrund i svarene på de tre foregående spørgsmål er det muligt for hver indsats at give et svar på, om der kan gennemføres modelberegninger ud fra det nuværende vidensniveau.

Ved besvarelsen af dette spørgsmål arbejder vi med en inddeling i tre beregningstyper, når vi omtaler den mulige beregning af økonomiske konsekvenser.

Det øverste niveau er *direkte evidens-beregning*, som betyder, at der findes kausal evidensviden for en sammenhæng mellem en udført indsats og en økonomisk konsekvens. Dette kunne fx være tilfældet, hvis der er påvist en sammenhæng mellem klassekvotienten og elevernes brug af specialundervisning.

¹ Fra evidensundersøgelsen leveres også et Excel-bilag med en database over alle de fundne effektstørrelser, der kan indgå i arbejdet med en kommende model. Indholdet i dette bilag vil fx kunne indgå som baggrund for modellen til udvælgelse af de konkrete effektantagelser til beregninger.

Det næste niveau er *indirekte evidens-beregning*, der betyder, at der ikke findes evidens for den direkte sammenhæng, men kan udledes en sammenhæng via et mellemlid. Dette kan fx være, hvis der er evidens for, at udsat skolestart virker positivt på elevers læring, og der samtidig er evidens for, at forøget læring betyder mindre anvendelse af specialundervisning.

Det nederste niveau er *beregning med statistisk sammenhæng*, som betyder, at der kan påvises en sammenhæng mellem en indsats og en økonomisk konsekvens, uden at denne sammenhæng nødvendigvis er en årsagssammenhæng.

Af de tre niveauer er de to øverste egentlige årsagssammenhænge, hvis der ligger veludførte effektmålinger til grund for den påviste evidens, mens det nederste niveau som nævnt ikke nødvendigvis er en årsagssammenhæng – man kan altså ikke sige, om det er indsatsen, der har afstedkommet den observerede forskel.

Det vil også være muligt at kombinere evidens med en beregning med statistisk sammenhæng. Det kan fx gennemføres ved at udnytte viden om effekten af en indsats på et link-outcome og herefter estimere en sammenhæng mellem niveauet for et link-outcome og de forskellige områder, hvor der kan være økonomiske konsekvenser.

Det kan ikke på forhånd afgøres, om en type model er bedre end en anden – det afhænger helt af, hvilket evidensniveau der ønskes. Generelt vil en model, der kun baserer sig på direkte evidens og/eller indirekte evidens, kunne tilbyde beregninger, der i nogen grad kan tilskrives årsagssammenhænge, fordi den anvender kausale estimater. Til gengæld vil en sådan model kun kunne dække relativt få områder, fordi der ikke nødvendigvis findes evidensunderbygning for alle områder. Det kan medføre, at de mulige økonomiske konsekvenser undervurderes. Omvendt vil beregninger fra en model, der alene baserer sig på statistiske sammenhænge, ikke kunne tilskrives årsagssammenhænge og dermed ikke være en kausal model. Derimod vil modellen kunne dække mange flere områder, fordi der ikke vil være samme krav til, at der skal eksistere evidens for sammenhængene. Den eksisterende SØM-model er alene baseret på statistiske sammenhænge. Kombinationsmodellen med evidens og beregning af statistisk sammenhæng kan have et element af kausalitet, fordi der er evidens for en af de to dele. Den efterfølgende statistiske sammenhæng mellem link-outcome og den økonomiske konsekvens er derimod ikke nødvendigvis en årsagssammenhæng.

Som en del af svaret på spørgsmålet om mulige modelberegninger belyses også, om der er tilstrækkeligt datagrundlag for at udføre beregningerne.

Som følge af, at den modelberegning, der kan gennemføres for en given indsats, afhænger af, hvilket evidensniveau der vælges, er anbefalingerne for de

enkelte indsatser inddelt i, hvad der er muligt med de tre niveauer direkte evidens, indirekte evidens og statistisk sammenhæng. For hver indsats indeholder rapporten konkrete anbefalinger til, om en model kan gennemføres – og i så fald på hvilket evidensniveau – samt anbefalinger til nye datakilder og analyser, der vil kunne forbedre grundlaget for fremtidige modelberegninger.

1.2.1 Kapitlernes indhold

Resten af rapporten er opbygget på følgende måde: Kapitel 2 indeholder en gennemgang af analysen, hvor de forskellige virkningskæder for de seks indsatser er undersøgt. I kapitlet opstilles forandringsteorier for indsatserne, og der inddrages litteratur for at kvalificere de mulige sammenhænge.

Kapitel 3 indeholder tværgående metodebetragtninger og -forslag til det videre arbejde frem mod en økonomisk model. Den foreslåede metode lægger sig op ad den metode, der anvendes i en økonomisk model udviklet af Washington State Institute for Public Policy (WSIPP).

Kapitlerne 4-9 indeholder gennemgang af hver af de seks indsatser. For hver indsats gennemgås de fire underspørgsmål, der blev præsenteret ovenfor, og der samles op på svaret på hovedspørgsmålet om, hvorvidt det er muligt at gennemføre beregninger. Kapitlerne indeholder passager med gentagelser, idet det er meningen, at hvert kapitel skal kunne læses selvstændigt.

Rapportens bilag indeholder en gennemgang af processen omkring arbejdet med forandringsteorierne (Bilag 1) og en separat gennemgang af dataundersøgelsen (Bilag 2).

2 Mekanismer for virkning – teoretiske sammenhænge mellem indsatser og effekter

I denne afdækning har indgangen til arbejdet med at undersøge mulighederne for en økonomisk model været at undersøge, hvorvidt det er muligt at formulere sammenhængende og meningsfulde forklaringskæder (forandringsteorier), der kan føre til en sammenhæng mellem de konkrete indsatsområder og de mulige økonomiske konsekvenser.

Formuleringen af forandringsteorierne har både haft et processuelt og et analytisk sigte. Forandringsteorierne er blevet formuleret efter et forløb med afholdelse af workshops med deltagere fra VIVE, styrelser og ministerie, jf. afsnit 2.1 nedenfor. Disse workshops har fungeret som en samlende proces for deltagerne, hvor viden fra deltagerne kunne samles i de forskellige forandringsteorier. Metoden med workshops blev netop valgt for bedst muligt at få inddraget den samlede viden fra de tre aktører.

Formålet med den forandringsteoretiske tilgang som en del af løsningen af denne opgave har været:

- at samle sammen på den tilgængelige viden om de forskellige områder
- at få formuleret forbindelser, der eventuelt ikke normalt ses for at få dem undersøgt
- at sikre, at alle vinkler kan afdækkes
- at forsøge at forbinde indsatser med økonomisk målbare indikatorer
- at præsentere et visuelt og let tilgængeligt billede af hvordan indsatserne tænkes at virke.

Vi anvender en forandringsteoretisk tilgang til at give indsigt i, hvilke virkingskæder indsatserne kan virke gennem for at udløse en økonomisk effekt. Forandringsteorier er oftest brugt som en del af virkningsevalueringer (Dahler-Larsen, 2013), som også er kendt som den teoribaserede evaluering (Chen, 1990; Pawson & Tilly, 1997; Weiss, 1997; Donaldson, 2007). Virkningsevalueringen griber typisk fat i den overordnede teori, som et program eller indsats bygger på. Hovedspørgsmålet er, om de forestillinger om forholdet mellem program og resultat, som et program bygger på, kan bekræftes, afkræftes eller eventuelt udvikles. I denne rapport og i processen frem mod rapporten er forandringsteorier brugt som en metode til at formulere virkingskæderne, der udgør de teoretiske forbindelser mellem indsats og potentielle effekter som

målt ved hjælp af de økonomiske indikatorer. Virkningskæderne består af en række forbindelser, der forbinder indsats med en potentiel økonomisk konsekvens, hvor antallet af led i kæden afhænger af argumentationen for virkningskæden. Vi arbejder i dette kapitel med at undersøge, om de formodede virkningskæder kan bekræftes via fund i litteraturen.

Forandringsteoriene i denne rapport skal ikke ses som bestandige billeder på indsatsernes forventede virkninger, men derimod som produkter af den konkrete proces og den konkrete fremsøgte viden til vidensunderbygning af forandringsteoriene. Forandringsteoriene skal ikke ses som det endelige resultatet af denne afdækning, men derimod som det analytiske bindeled mellem indsatserne og de potentielle økonomiske indikatorer.

Formålet med forandringsteorier i forbindelse med denne afdækning er at vise, hvordan indsats og potentiel økonomisk indikator kan forbindes teoretisk ud fra mulige virkningskæder. Forandringsteoriene kan dermed ses som et redskab til at sandsynliggøre eksistensen af en mulig økonomisk konsekvens af indsatserne.

Forandringsteorier er teoretiske konstruktioner, og derfor vil det heller ikke altid være muligt at underbygge forandringsteoriernes empirisk. I dette afsnit underbygges de enkelte dele af forandringsteoriene med empiriske undersøgelser, hvor det er muligt. Ud fra muligheden for at underbygge med viden fra forskellige kilder vurderes det, hvorvidt en given forbindelse skal fastholdes i forandringsteorien – og hvis den fastholdes, hvorvidt det er en velunderbygget (stærk) forbindelse eller en mindre underbygget (svag) forbindelse. Denne vurdering bygger på en række forhold omkring den empiriske underbygning, herunder studiernes omfang, mængden af studier, evidensniveau, relevans, overførbarhed til kontekst mv.

Den endelige opbygning af forandringsteoriene bygger dermed også på de tilgængelige kilder og på, at de skal være operationelle. En del af forbindelserne falder bort i forløbet, og dette skyldes ikke, at konstruktionen af forandringsteoriene er fejlagtig eller mangelfuld. Den mindre volumen i de endelige forandringsteorier skyldes i høj grad, at der mangler mulighed for empirisk underbygning af de enkelte sammenhænge. Dette gør sig mere gældende på dagtilbudsområdet, end det gør på folkeskoleområdet. Det lægger dermed også op til, at der mangler mere viden – herunder forskningsmæssig viden – om sammenhængene på dagtilbudsområdet end på skoleområdet.

I dette afsnit arbejdes der med to generiske forandringsteorier, der dækker indsatser på henholdsvis skole- og dagtilbudsområdet, og som blev resultatet af de afholdte forandringsteoriworkshops, jf. næste afsnit.

De to generiske modeller anvendes i de efterfølgende indsatskapitler til at koble indsatserne på de relevante forandringskæder i de generiske forandringsteorier, hvorved forandringsteorierne går fra at være generiske til at være forandrings-teorier for de enkelte indsats. På indsatsniveau undersøges det herefter, om der kan findes evidens for dele af forandringskæderne. For eksempel kan det undersøges, om der er evidens for, at en større andel af uddannet personale har betydning elevernes trivsel eller for personalets sygefravær.

2.1 Formulering af forandringsteorierne

I forbindelse med dette projekt blev formuleringen af forandringsteorierne lagt processuelt an. Afsættet for formulering for alle seks indsats var en workshop, hvor relevante medarbejdere fra henholdsvis Børne- og Undervisningsministeriet, Styrelsen for Uddannelse og Kvalitet og VIVE var inviteret til at deltage og give input til formuleringen af forandringsteorierne. Deltagerne var udvalgt ud fra, at de skulle have faglig indsigt i den pågældende indsats. Deltagerne på workshopperne leverede input til forandringsteorierne og de virkende mekanismer. Afsættet i workshops gør også, at det er workshopperne, der i første omgang er styrende for indholdet af forandringsteorierne. VIVE har derefter kvalificeret og gennemarbejdet forandringsteorierne – både for at sikre konsistens og for at undersøge muligheden for vidensunderbygningen. Det overordnede ansvar for formuleringen af de enkelte forandringsteorier er placeret hos VIVE. Det samlede overblik over proces og de generiske forandringsteorier, som er formuleret på baggrund af processen, kan ses i Bilag 1.

VIVE har på baggrund af de formulerede bruttoforandringsteorier dannet to generiske forandringsteorier for henholdsvis folkeskole- og dagtilbudsområdet. Disse bruges i det videre arbejde til at underbygge de enkelte forbindelser med viden.

2.1.1 VIVEs vidensunderbygning

VIVE har på baggrund af de to generiske forandringsteorier opdelt forandringsteorierne til de enkeltstående forbindelser, der hver er søgt vidensunderbygget. Denne vidensunderbygning er sket både gennem den tilhørende evidensundersøgelse og ved indhentning af yderligere viden. Den yderligere viden er fx indhentet gennem afsøgning af relevante videnskabelige databaser eller andre kilder til forskning på området, viden fra evaluering samt viden præsenteret i rapporter og andre formidlingskilder. VIVE forholder sig i det følgende løbende til kvaliteten og tyngden af den empiriske viden til de enkelte forbindelser. I vidensunderbygningen af forandringsteorierne anses en sammenhæng mellem to størrelser, fx trivsel og læring, for underbygget, hvis der

kan påvises en sammenhæng mellem de to størrelser. Kravet er ikke, om det kan påvises, om sammenhængen går den ene eller den anden vej.

På baggrund af vidensunderbygning har VIVE formuleret to endelige generiske forandringsteorier, der kan anvendes til det videre arbejde frem mod en økonomisk model.

2.1.2 De endelige forandringsteorier

Slutteligt i processen med forandringsteorierne har VIVE sorteret i den grafiske fremstilling af forandringsteorien, således at der fremgår to nye versioner af de generiske forandringsteorier, hvor det kun er de underbyggede forbindelser, der er fastholdt. Derudover er de stærkt underbyggede forbindelser markeret med fuldt optrukne streger, mens de svagt underbyggede er markeret med stiplede streger.

Disse to generiske forandringsteorier er derefter ført tilbage til de enkelte indsatser, der i hver deres forandringsteori er bundet op på de relevante virkningskæder i den relevante forandringsteori.

2.2 Vidensunderbygning af de enkelte trin i folkeskoleforandringsteorien

I det følgende præsenteres den relevante fremsøgte viden til de enkelte trin af forandringsteorien for folkeskoleområdet. Den fremsøgte viden præsenteres under de enkelte forbindelser. Dele af vidensgrundlaget peger imidlertid ikke særskilt på en enkelt forbindelse, men kan bruges til at underbygge flere forbindelser.

2.2.1 Øget elevtrivsel fører til øget læring

Denne forbindelse, som den blev formuleret på den afholdte workshop, forholder sig til sammenhængen mellem trivsel og læring. Hypotesen er, at hvis den enkelte elevs trivsel øges, er det muligt at øge læringsmængden, fordi der er mere overskud til den faglige del af skolegangen. En diskussion fra workshoppen, som også er relevant her, er, om det er muligt at isolere trivsel fra læring i skolen, eller om det i højere grad er to emner, der er så tæt sammenknyttede for eleverne, at det ikke er muligt at forholde sig til det ene uden også at forholde sig til det andet.

I denne undersøgelse ses elevtrivsel som bestående af en række forskellige forhold omkring eleverne. Det dækker fx både over elevernes sociale og mentale trivsel og socio-emotionelle kompetencer for eleverne, men det er fælles for dem, at det er den del af trivslen, der udspiller sig i en skole-setting. Hvor social og mental trivsel oftest måles ved surveys, der spørger til forskellige dele af humør/holdninger, så er socio-emotionelle kompetencer mere rettede på at måle interaktionen med andre.

Der er i VIVEs afsøgning af området fundet studier, der både fokuserer på den direkte sammenhæng mellem trivsel og læring, og studier, der undersøger dette igennem andre forbindelser, fx gennem elevernes helbredstilstand. En del af de fremsøgte referencer forholder sig også til det omvendte perspektiv: hvordan øget læring kan medføre øget trivsel i skolen. For viden om enkelte indsatsers sammenhæng med læring specifikt findes også en del studier i den tilhørende evidensundersøgelse til denne rapport.

I en rapport fra "Departement for Education" i Storbritannien vises en række forskellige forbindelser mellem trivsel og læring (Gutman & Vorhaus, 2012). Fokus er her elever i alderen 7-13 år og trivsels indflydelse på senere uddannelsesmæssige outcomes i alderen 11-16 år. En anden rapport fra Storbritannien, der forholder sig til emnet, ser ud over det også på, hvordan elevernes (fysiske/somatiske) helbred kan indvirke på deres opnåede uddannelsesmæssige outcomes (Brooks, 2014). I sidstnævnte undersøgelse finder man, at elever, der har bedre helbred og trivsel, har en større sandsynlighed for at præstere bedre rent fagligt, og at gode sociale og emotionelle kompetencer har en sammenhæng med både bedre helbred og trivsel *og* bedre resultater.

I forhold til sammenhængen mellem trivsel, forstået som elevens socio-emotionelle styrker eller udfordringer, og deres præstationer i faglige test, kan nævnes undersøgelsen af sammenhængen mellem elevens trivsel og disses nationale tests (Ministeriet for Børn, Undervisning og Ligestilling, 2016), hvor der ses en positiv sammenhæng mellem gode testresultater og elevernes egenoplevede trivsel. Hvor det i førnævnte er den nationale trivselsundersøgelse, der anvendes i analyserne, anvender andre analyser et SDQ-spørgeskema (the Strengths and Difficulties Questionnaire) til at operationalisere (mis-)trivsel blandt elever, og finder ligeledes en positiv sammenhæng: jo større udfordringer eleverne har rent socio-emotionelt, desto lavere scorer de i faglige tests (Keilow et al., 2019).

En række andre undersøgelser forholder sig også til denne sammenhæng, i andre lande, aldersgrupper og operationaliseret på forskellige måder (fx Lewis et al., 2011; Lyons & Huebner, 2016; Ng, Huebner & Hills, 2015; Oishi, Diener & Lucas, 2007).

Ikke alle studier finder de samme sammenhænge, fx finder en undersøgelse på baggrund af surveydata i dansk kontekst ikke nogen selvstændig sammenhæng mellem *emotional engagement*, hvor skoletrivsel er et komponent, og læring, når der også tages højde for komponenter som elevernes socioøkonomiske baggrund, motivation og deltagelse (Rangvid, 2018).

2.2.1.1 Studier om sammenhæng mellem mentalt helbred (diagnoser) og læring

Nogle undergrupper af elever kan opleve andre effekter end gruppen af elever set under ét. En relevant undergruppe at forholde sig til i denne sammenhæng er de elever, der har forskellige udfordringer i forhold til deres mentale helbred. Dette afsnit drejer sig om modifikationer til forbindelser, der går gennem forbedret læring. Eksempelvis finder en undersøgelse af børn og unge med funktionsnedsættelser, at de ikke præsterer lige så højt fagligt (målt på karaktergennemsnit), som de resterende elever. Ligeledes finder studiet, at børn og unge med funktionsnedsættelser vurderer deres egen trivsel lavere end den resterende gruppe elever (Mortensen, 2020). Studiet har ikke undersøgt, om der er en sammenhæng imellem, hvordan eleverne klarer sig på de to mål.

For gruppen af børn mellem 11 og 15 år finder et studie (DeSocio & Hootman, 2004) en sammenhæng mellem børnenes mentale sundhed og dårligere læring og øget fravær. Andre studier viser, at dårligt mentalt helbred i barndommen, fx ADHD, hænger sammen med forskellige skole-outcomes, fx opnået uddannelse, faglige testscorer og at gå klasser op (Currie & Stabile, 2004; Currie, 2008). Et andet studie viser, at især udadreagerende problemer (sammenlignet med indadreagerende problemer) hænger sammen med dårligere lærings-outcomes (Evensen et al., 2016).

Samlet set er der flere forskellige studier, der underbygger, at der er en sammenhæng mellem trivsel og læring, men det er ikke entydigt, hvilken retning sammenhængen går, eller om de to begreber i stedet skal ses som et samlet hele. Samlet anses sammenhængen at være underbygget. Dermed vil en indsats, som enten påvirker trivsel og/eller læring, aktivere virkningskæden, der kan føre videre frem mod de potentielle økonomiske indikatorer. Herudover bemærkes det også, at der for gruppen af børn og unge med diagnoser relateret til mentalt helbred findes en særlig forbindelse vedrørende typen af psykiske udfordringer i forhold til læring.

2.2.2 Øget elevtrivsel fører til færre omgængere

Udgangspunktet for denne forbindelse i forandringsteorien var en hypotese om, at hvis elevernes trivsel var god, var de i højere grad i stand til at modtage læring. Den forventede forbindelse videre fra læring var, at der gennem øget

faglig læring også kunne opnås, at færre elever blev nødt til at gå en klasse om, fordi de ikke havde nået tilstrækkeligt fagligt niveau til at starte på næste klassetrin. Til en vis grad går hypotesen også direkte gennem den sociale trivsel, at elever der havde bedre trivsel, også i højere grad var i stand til at udvikle sig i socialt og dermed var en stærkere del af fællesskabet og kunne følge klassekammeraternes sociale udvikling til næste klassetrin.

VIVE har i afsøgningen af denne forbindelse ikke fundet noget videnskabelig litteratur, der kan underbygge forbindelsen. Tænketaenken DEA har lavet en analyse af sammenhængen mellem barnets trivsel i første skoleår (0. klasse), og hvorvidt barnet går 0. klasse om (DEA, 2020). Studiet bygger på registerdata og finder, at omgængere i 0. klasse trives i lavere grad end den resterende gruppe elever.

Forbindelsen underbygges af studiet fra DEA og fastholdes dermed i den endelige forandringsteori.

2.2.3 Øget elevtrivsel medfører øget inklusion

I workshoppen blev denne forbindelse formuleret som en forbindelse, der går gennem lærernes praksis, hvor en øget elevtrivsel i klassen kan understøtte, at der i klasserumssituationer er mere overskud til at rumme elever med andre behov og dermed at øge inklusionen. Altså at man for de samme lærerressourcer kan rumme flere elever – særligt elever med andre behov, der måske tidligere har været i ekskluderede tilbud.

VIVE har i afsøgningen af viden ikke kunnet finde viden, der understøtter denne forbindelse direkte eller indirekte, hvorfor den ikke er med i den endelige forandringsteori for folkeskoleområdet.

2.2.4 Bedre indeklime giver bedre lærertrivsel, bedre læring og mindre fravær

I workshoppen tog denne hypotese afsæt i en rapport fra Incentive og DTU, der samler op på de samfundsøkonomiske gevinster ved forbedret indeklime (Madsen et al., 2020). Rapporten samler op på forskningen om påvirkning af elevers læring og læreres trivsel set på baggrund af en række forskellige indikatorer. I forhold til elevernes læring ses bl.a. på karakterer og valg af uddannelse på længere sigt. Desuden ses også på læringen som resultat af, at eleverne har mindre fravær. Det lavere fravær skyldes ifølge rapporten, at eleverne er udsat for mindre fugt, mindre risiko for skimmelsvamp og mere hensigtsmæssig temperatur. Sammenhængen er stærkere for eleverne end for

lærerne, da eleverne er i klasseværelserne flere timer end lærerne. Lærernes trivsel ses i denne forbindelse som værende fravær af sygdom.

På baggrund af rapporten fastholdes forbindelserne i forandringsteorien.

2.2.5 Bedre elevtrivsel fører til mindre kriminalitet

På workshopen blev denne sammenhæng formuleret som flere forskellige forbindelser mellem elevernes trivsel og fællesskabet i klasse frem mod elevernes potentielle kriminalitet. Da der i denne sammenhæng kun ses på forandringer i en 4-årsperiode, vil der for langt de færreste elever være tale om, at man kan definere det som en kriminel løbebane. Derfor er forbindelsen udelukkende, hvorvidt eleverne begår kriminalitet.

Som underbygning af denne forbindelse anvendes en udgivelse fra Det Kriminalpræventive Råd (Balvig, 2011), der på baggrund af gentagne spørgeskemaundersøgelser undersøger, hvad der gør, at unge er henholdsvis lovlydige og ikke-lovlydige. Her findes det på baggrund af regressionsanalyser af paneldata, at der er sammenhæng med flere forskellige vurderinger fra eleverne af deres skolegang og det, at de falder i den ene eller anden gruppe. Det gør sig fx gældende for, hvorvidt eleverne kan lide at gå i skole, hvilket kan ses som en selv vurderet trivselsmåling.

På baggrund af ovenstående analyser fastholdes forbindelserne, men de vurderes at være svage, da det udelukkende er dele af elevernes trivsel, der selv vurderes.

2.2.6 Bedre elevtrivsel fører til mindre fravær

Elevtrivsels sammenhæng med fravær blev på workshopen formuleres som en relativt simpel forbindelse, hvor det, at eleverne har det godt i klassen, gør, at de har mindre lyst til at være fraværende. Hvis denne hypotese gør sig gældende, vil man dermed også kunne minimere fravær gennem en bedre elevtrivsel.

Det har ikke været muligt at vidensunderbygge denne sammenhæng direkte, men i en spørgeskemaundersøgelse fra 2018 (Rambøll, 2018) er kommunale konsulenter, der arbejder med fraværsområdet, blevet bedt om at vurdere, hvad årsagerne til elevers fravær er med særligt fokus på langtidsfravær. Her vurderer 40 %, at elevernes fravær skyldes personlige udfordringer som fx skoletræthed og manglende motivation, mens 27 % vurderer, at de skyldes sociale udfordringer i skolen som fx mobning.

På baggrund af ovenstående spørgeskemaundersøgelse fastholdes forbindelsen, men den vurderes at være svagt underbygget, da det er den samlede vurdering af de unge, der anvendes til at underbygge forbindelsen.

2.2.7 Mindre fravær fører til øget arbejdsudbud blandt forældre

På workshoppen blev denne hypotese baseret på antagelsen om, at forældre er nødt til at blive hjemme og passe børn i tilfælde af sygdom og fravær. De dage, hvor forældrene passer deres børn, er de ikke til rådighed for arbejdsmarkedet, og derfor er forældrenes arbejdsudbud også mindre.

Sammenhængen er umiddelbar logisk og kan også funderes på andre kilder (Madsen et al., 2020). Sammenhængen bliver dog i nævnte rapport nuanceret, og derfor skal det heller ikke ses som en 1:1 sammenhæng, hvor en dags mindre fravær giver en dags mere arbejdsudbud for forældrene. Dette skyldes ifølge rapporten tre forhold: en del af forældrene er i forvejen ikke en del af arbejdsstyrken; ikke alt elevfravær kræver forældrepassning, og at en del af eleverne passes af andre end forældrene, fx bedsteforældre (Madsen et al., 2020).

På baggrund af ovenstående rapport fastholdes sammenhængen i den endelige forandringsteori.

2.2.8 Øget elevtrivsel fører til færre skoleskift

I workshoppen blev denne hypotese formuleret på baggrund af to mekanismer i klasserne: dels en fastholdelsesmekanisme, dels en fraskydningsmekanisme. Den første forståelse er, at eleverne i højere grad bliver fastholdt i en klasse, hvis de har god trivsel, da trivsel hermed også ses som at være en del af fællesskabet. I fraskydningsmekanismen er forståelsen, at elever, der ikke trives eller er del af fællesskabet, i højere grad bliver adskilt fra fællesskabet. Samlet set er der tæt sammenhæng mellem den enkelte elevs trivsel og trivslen generelt i klassen i denne hypotese, som den blev formuleret på workshoppens.

Denne forbindelse underbygges af flere forskellige studier; fx viser en rapport på baggrund af analyser af den nationale panelundersøgelse "Børn og unge i Danmark" (Ottosen et al., 2018), at flere, der har skiftet skole, oplyser, at skoleskiftet er sket netop, fordi de trives dårligt i deres gamle klasse – det gør sig især gældende for de ældre piger. Herudover viser analyserne, at 41 % af de børn, der i panelet havde skiftet skole, angav, at mindst et af deres skoleskift skyldes, at de ikke trivedes i den gamle skole.

Herudover kan nævnes en række internationale studier, der også beskæftiger sig med skoleskift og sammenhængen med elevernes mentale sundhed og

trivsel. I forlængelse af dette ses i dette studie, at socio-emotionelle udfordringer, som fx ADHD, har en sammenhæng med risikoen for at gå om (Currie, 2008).

Denne sammenhæng findes til en vis grad underbygget i det materiale, som VIVE har kunnet fremfinde i denne forbindelse. Sammenhængen ses dog som stærkest, hvis eleverne har udfordringer med den mentale sundhed.

2.2.9 Mere undervisningsdifferentiering fører til øget elevtrivsel og læring

I workshoppen blev denne forbindelse beskrevet som muligheden for, at eleverne i højere grad trives og dermed lærer mere, hvis undervisningen er differentieret og dermed tilpasset til den enkelte elevs behov.

Der er i afsøgningen af relevant viden udelukkende fundet en række mindre studier, der beskæftiger sig med denne forbindelse. Som eksempler på dette kan nævnes et mindre kvalitativt studie af, hvordan undervisningsdifferentiering påvirker læring og trivsel. De samme elevers læring og trivsel sammenlignes i en almindelig undervisnings-setting (første del af kurset) og et alternativt setting med undervisningsdifferentiering (kursets anden del). Studiet finder en positiv virkning på begge resultatmål omkring læring og trivsel (Avci et al., 2009) Studiet bygger på et begrænset antal individer, nemlig 22, og er udelukkende kvalitativt.

Et andet mindre studie (Baumgartner, Lipowski & Rush, 2003) viser på en før-/eftertest en positiv virkning på bl.a. elevernes læsefærdigheder. Studiet bygger på en mindre intervention i tre klasser i USA.

Et lidt større studie, hvor sammenhængen også er undersøgt over længere tid, nemlig 8 år, er foretaget på en grundskole i USA (Beecher & Sweeny, 2008). Studiet finder en positiv sammenhæng mellem undervisningsdifferentiering og elevernes læse-/skrivefærdigheder og matematikkundskaber. Sammenhængen er særligt stærk for elever fra ressourcetsvage hjem.

Forbindelsen fastholdes på baggrund af den fremsøgte viden, men den ses ikke at være stærk og bygger udelukkende på mindre og ikke-danske undersøgelser.

2.2.10 Mere undervisningsdifferentiering fører til mindre brug af specialundervisning

På workshoppene blev denne forbindelse set som mulighed for, at elever i mindre grad har brug for specialundervisning eller mindre segregering i specialklasse, hvis undervisningen bliver tilrettelagt i forhold til den enkelte elevs behov.

Det har ikke været muligt for VIVE at finde viden, der direkte underbygger denne forbindelse. Indirekte kan forbindelsen underbygges af en international publikation, der omhandler, hvordan undervisningsdifferentiering bruges til at undervise elever med særlige behov i almenundervisningen, men der er ikke set på, om det kan forhindre, at elever kommer i segregerede specialtilbud (Strogilos, 2018).

Yderligere underbygges denne forbindelse også indirekte af to studier, som viser, at undervisningsdifferentiering gavner inkluderede elever (Kern et al., 2001; Lee et al., 2010).

Da det ikke har været muligt for VIVE at finde viden, der underbygger denne forbindelse direkte og for hele populationen, fastholdes den ikke i den endelige model.

2.2.11 God lærertrivsel medfører bedre læring for eleverne

Perspektivet her er, at når lærerne trives, så er undervisningen i højere grad af god kvalitet, hvorved læringen blandt eleverne understøttes.

Sammenhængen underbygges fx af et studie, der finder en positiv sammenhæng mellem netop lærertrivsel og læring (Turner & Theilking, 2019). Derudover finder et andet studie en positiv sammenhæng mellem lærernes generelle velbefindende ved starten af skoleåret og elevernes læring i slutning af skoleåret (Duckworth, Quinn & Seligman, 2009). Et italiensk studie blandt 2000 lærere viste, at trivsel blandt lærerne målt ved bl.a. tilfredshed med jobbet og selvtillid førte til, at eleverne fik højere karakterer ved deres afsluttende eksaminer (Caprara et al., 2006).

Sammenhængen ses dermed som understøttet og fastholdes i forandringsteorien, dog ses den at være svag, dels fordi studierne ikke er gennemført i dansk kontekst, dels fordi den kvantitative mængde af studier er begrænset.

2.2.12 God lærertrivsel medfører lavere læreromsætning

Hypotesen på workshoppen var her, at hvis lærerne trives, så fastholdes de i højere grad i deres stilling.

Sammenhængen ses underbygget i et kvantitativt studie fra Taiwan blandt 250 lærere, hvor det ses, at lærertrivsel mindsker intentionen fra lærerne om at forlade skolen (Chang, Chiu & Liu (2017)). I modsætning til dette finder et større dansk kvantitativt studie ingen sammenhæng mellem personaleomsætning og sygefravær og lærertrivsel i et spørgeskemastudie blandt 440 lærere (Anderson & Winther, 2011).

På baggrund af ovenstående og grundet den begrænsede mængde om emnet, det er lykkedes VIVE at finde, untlades forbindelsen i forandringsteorien.

2.2.13 Øget elevtrivsel medfører mindre brug af specialundervisning

Hypotesen her ligger i formuleringen fra workshoppen tæt på forbindelsen mellem elevtrivsel og øget inklusion. Dog omhandler denne hypotese mere specifikt elever, der ikke er i fare for at blive ekskluderet, men eventuelt har behov for mindre indsats i form af specialundervisning i begrænset form.

Det har i VIVEs afsøgning af viden ikke været muligt at underbygge denne sammenhæng direkte. Et bud på en indirekte underbygning er en forskningskortlægning, der fokuserer på trivsel blandt elever med særlige behov i henholdsvis almenklassen og specialtilbud. Denne kortlægning belyser dog ikke direkte, om der er en sammenhæng i forhold til, om øget trivsel kan nedsætte brugen af specialundervisning (Dyssegaard et al., 2013).

Denne forbindelse har dermed ikke været mulig at underbygge med viden i VIVEs afsøgning, hvorfor den untlades i den endelige forandringsteori.

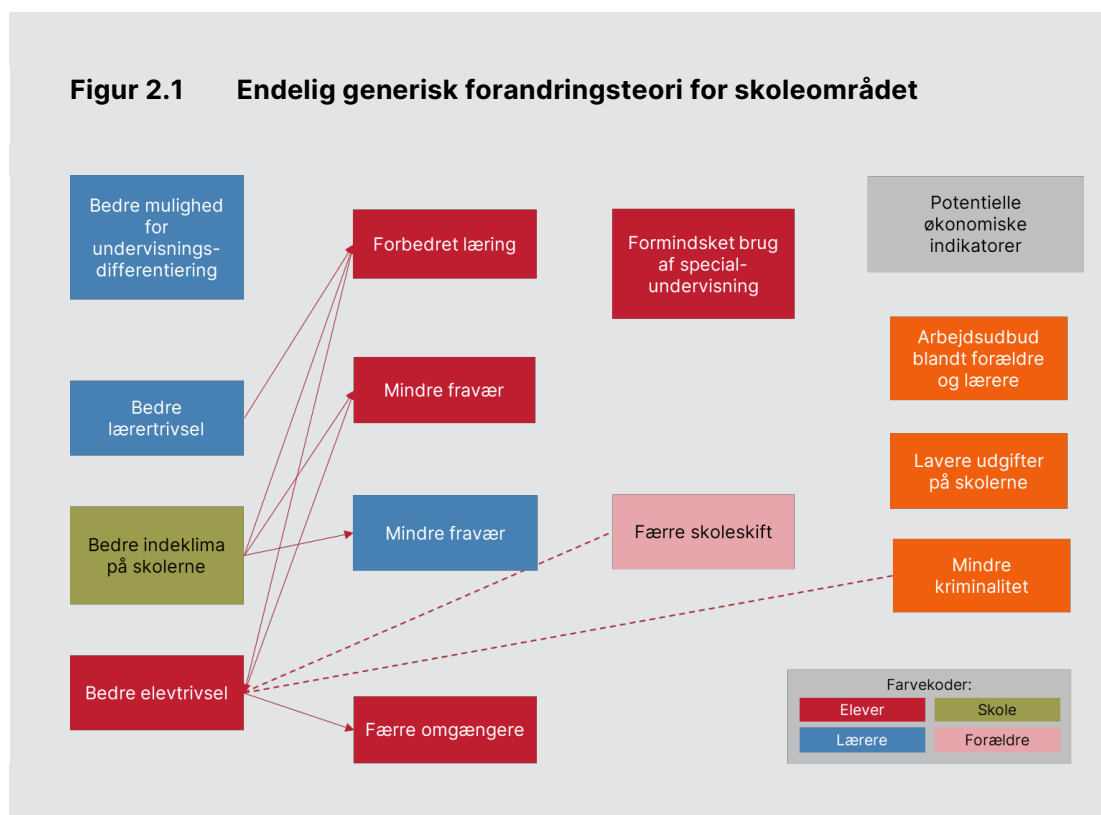
2.2.14 Øget elevtrivsel medfører lavere udgifter til specialundervisning

I workshoppen blev denne forbindelse formuleret som forbindelsen om, at hvis eleverne trives, så kan de i højere grad være en del af normalundervisningen. Dermed er der færre elever, der ekskluderes til specialtilbud eller har brug for særlige indsatser, hvorved der er en besparelse på dette område.

Det har ikke været muligt for VIVE at finde studier, der underbygger denne forbindelse, hvorfor den ikke fastholdes i den endelige model.

2.2.15 Endelig forandringsteori for skoleområdet

På baggrund af gennemgangen af de enkelte trin i forandringsteorien er de trin, det ikke har været muligt at finde belæg for, blevet udeladt af den endelige model. Den endelige generiske forandringsteori for skoleområdet kan ses i Figur 2.1.



2.3 Vidensunderbygning af de enkelte trin i dagtilbudsområdet

VIVE har på samme måde som på skoleområdet søgt efter viden til at underbygge de enkelte forbindelser i den generiske forandringsteori for dagtilbudsområdet. Det har generelt været sværere at finde relevant viden til at underbygge disse forbindelser, hvorfor nogle af underbygningerne også bygger på et bredere vidensperspektiv end i skoledelen af forandringsteoriene.

Dette skal også ses i lyset af, at dagtilbudsområdet er forskningsmæssigt mindre belyst end skoleområdet. Det kan derfor også udledes, at der er brug for mere viden om sammenhænge på dagtilbudsområdet set ud fra et mere kvantitativt perspektiv for at underbygge brugen af en økonomisk model på dagtilbudsområdet. En mulighed her er også at undersøge allerede indsamlet

data for sammenhænge mellem den leverede indsats på dagtilbudsområdet og de ønskede resultatmål.

I det følgende gennemgås de enkelte forbindelser fra forandringsteorien, ligesom vidensunderbygningen vurderes i forhold til, hvorvidt de enkelte forbindelser fastholdes i forandringsteorien for dagtilbudsområdet.

2.3.1 Bedre trivsel blandt medarbejderne medfører lavere sygefravær

I workshoppen blev denne hypotese formuleret som den – på papiret – simple sammenhæng mellem, at hvis medarbejderne trives, så vil der også være et lavere sygefravær blandt medarbejderne. Gennem VIVEs afsøgning og forsøg på at vidensunderbygge denne forbindelse viser det sig dog, at denne forbindelse ikke kan stilles op så simpelt. Et eksempel på dette er forskningsprojektet Pionerprojektet (Framke, 2016), hvor arbejdsmiljøet i daginstitutionerne forsøges påvirket gennem en række interventioner. Der deltog 78 institutioner i undersøgelsen, hvoraf de 44 modtog en intervention. Det lykkedes ikke for projektet at påvirke trivslen, men det lykkedes derimod at påvirke sygefraværet positivt. En konklusion fra den tilhørende ph.d.-afhandling er, at forbindelsen mellem sygefravær og trivsel bør undersøges nærmere generelt på arbejdsmarkedet og specifikt på dagtilbudsområdet.

Et andet studie (Gørtz & Andersson, 2014) bruger *child-to-teacher ratio* som et mål for arbejdspress, som anses som en dimension af trivsel. De bruger *large scale register data*, og metoden har høj evidenskvalitet. De finder, at flere børn pr. medarbejder øger korttidssygefraværet i vuggestuer (0-3 år), men ikke i børnehaver.

Det har ikke været muligt at underbygge forbindelsen direkte, men den fastholdes som en svag forbindelse i forandringsteorien, der med fordel kan undersøges nærmere. Desuden fastholdes de videre forbindelser frem i forandringsteorien mod udgifter på daginstitutionernes budget og medarbejdernes arbejdsudbud, til trods for at de ikke er videnskæssigt underbygget, men anses som at være logiske følger af et lavere sygefravær.

2.3.2 Bedre trivsel blandt medarbejderne fører til mindre medarbejderomsætning

Det har ikke været muligt at fremsøge kilder, der direkte understøtter denne forbindelse. Nogen studier viser en anderledes forbindelse mellem de to forhold, nemlig at dårlig trivsel og andre arbejdsmiljørelaterede forhold er en

medvirkende årsag til jobskifte (fx Hale-Jinks & Knopf, 2006; Albertson & Kagan, 1987; Borg, Riding, & Falzon, 1991; Fletcher & Payne, 1982; Galinsky, 1988; O'Connor & Clarke, 1990).

Da der ses en forbindelse mellem trivsel og medarbejderomsætning, fastholdes forbindelsen i forandringsteorien, men den markeres til at være svag, da den ifølge de fremsøgte kilder har en anden forbindelse, end den blev formuleret på workshoppen. Forbindelsen videre frem mod udgifter på daginstitutionens budget fastholdes også som udgifter til rekruttering af medarbejdere og oplæring af medarbejdere. Den sidstnævnte forbindelse markeres som svag, da det ikke har været muligt at fremsøge viden, der underbygger denne forbindelse. Fastholdelsen skal ses i lyset af ræsonnementet om, at rekruttering ikke er gratis, men det er en forbindelse, der med fordel kan undersøges videre, hvis den økonomiske indikator ønskes som en del af den økonomiske model på dagtilbudsområdet.

2.3.3 Tættere relation mellem børn og voksen giver bedre overlevering til skolestart

I workshoppen blev denne hypotese formuleret og udviklet med et blik på, at hvis medarbejderne har eller opnår en god relation til børnene i dagtilbuddene, kender de i højere grad også børnene og deres styrker og udfordringer. Dette kendskab til børnene kan så, hvis man følger argumentationen fra workshoppen, anvendes i overleveringen i forbindelse med skolestart, hvor de oplysninger, som medarbejderne overleverer, er tættere på barnet og barnets behov og derfor også i højere grad kan anvendes af skolen til at arbejde med børnene i forbindelse med skolestart.

Denne forbindelse underbygges af en række australske anbefalinger omhandlende overgangen mellem dagtilbud og skole (Dockett & Perry, 2014). Anbefalingerne bygger på en omfattende mængde forskning. De fokuserer på, at den gode overlevering sker, når den afgivende medarbejder har kendskab til barnet, og at kendskabet overleveres til skolen og bruges til at opbygge en relation til barnet, hvori kendskabet er afsæt for videre læring og arbejde med det enkelte barns trivsel.

Dermed findes forbindelsen videnskæssigt underbygget, og den fastholdes i forandringsteorien for dagtilbudsområdet.

2.3.4 God eller øget børnetrivsel medfører mere skoleparate børn

Sammenhængen under denne overskrift blev på workshoppen formuleret med afsæt i, at hvis børnene i dagtilbuddene har god trivsel eller opnår bedre trivsel, har de mere overskud til at indgå i dagligdagen i dagtilbuddet. Dermed underbygges også børnenes deltagelse i de pædagogiske indsatser, hvoraf nogle peger frem mod skolen, både set som socialt udviklende og i forhold til mere læringsorienterede aktiviteter såsom lege med ord, bogstaver eller tal. Endelig kan en del af dette også være børnenes sproglige udvikling, der også, hvis man følger denne hypotese, kan ses at være en del af en aktivitet, som god trivsel kan underbygge udbyttet af for børnene, og som også er afsæt for skoleparathed og læring i skolen.

Det har ikke været muligt for VIVE at fremsøge litteratur, der underbygger denne sammenhæng direkte, men flere forskellige dele kan anvendes til at underbygge forskellige dele af denne sammenhæng.

Danmarks Evalueringsinstituts undersøgelse af kvalitet i de danske dagtilbud (EVA, 2020) med anvendelse af ECERS-3 til vurdering af kvaliteten på en række forskellige områder viser, at kvaliteten varierer mellem de forskellige dagtilbud, både på tværs af kommuner og i de enkelte kommuner. I rapporten anvendes sammenhængen mellem kvalitet og læring og videre frem til skoleparathed for børnene, og her bruges ECERS til at vurdere kvaliteten af dagtilbuddet. Hvorvidt det så i sidste ende gør sig udslag i mere skoleparathed for børnene, der har været i de dagtilbud, som EVA vurderer som at være af høj kvalitet, kan med fordel belyses gennem videre analyse af det indsamlede data.

I forskningskortlægningen "Daginstitutionens betydning for børns udvikling, en forskningsoversigt" (Christoffersen, Højen-Sørensen & Laugesen, 2014) gennemgås en stor mængde forskning om betydningen af dagtilbuddene på børnenes udvikling, men forbindelsen mellem trivsel og skoleparathed berøres kun indirekte.

I forskningsoversigten fra Clearinghouse (Dyssegaard et al., 2013) ses der på forskellige faktoreres indflydelse på skoleparathed, herunder også på dagtilbuddenes kvalitet. Her ses ikke direkte på sammenhængen mellem børns trivsel og deres skoleparathed, men der ses positive forbindelser mellem skoleparathed og det, at børn har gået i et dagtilbud, samt mellem skoleparathed og det, at børn er en del af et dagtilbud af høj kvalitet. Det er altså ikke trivslen isoleret, der undersøges, men trivsel er her set som en del af den leverede kvalitet.

Det har ikke været muligt at fremsøge viden, der undersøger denne sammenhæng direkte, men derimod ses trivsel at være en del af den leverede kvalitet,

som igen ses som en del af det, der er med til at understøtte børnenes skoleparathed. Forbindelsen fastholdes derfor, men samtidig bemærkes det, at det er en forbindelse, der med fordel kan udforskes nærmere.

2.3.5 God trivsel for børnene giver mindre behov for andre sociale indsatser

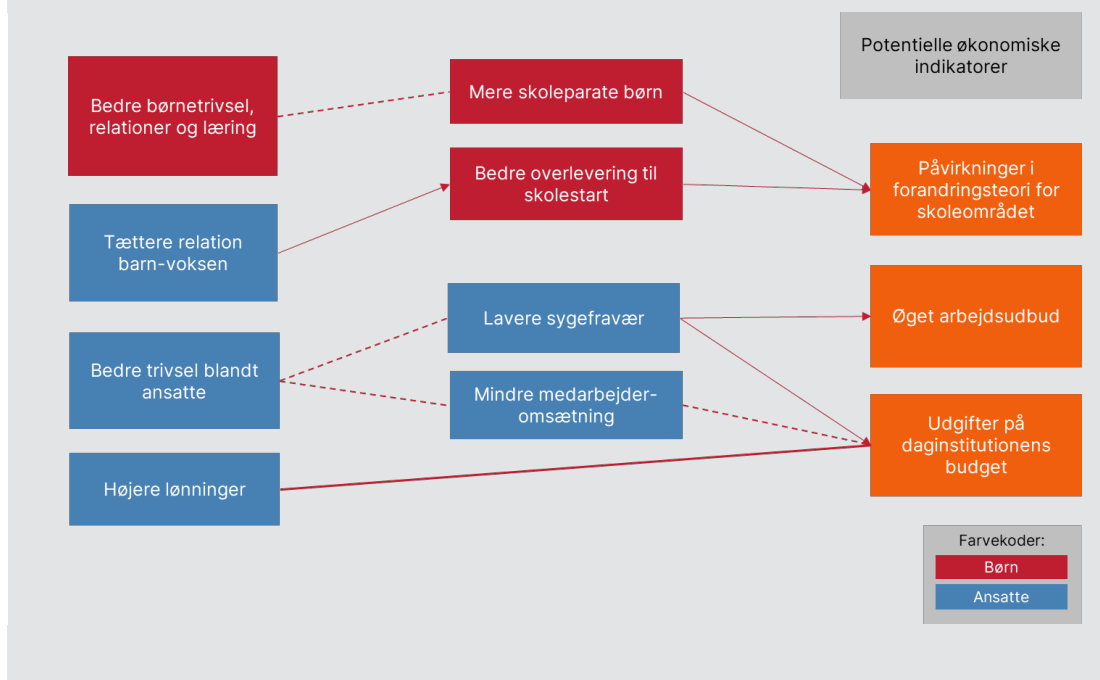
Den sidste forbindelse fra den generiske forandringsteori, der søges underbygget i denne del, blev på workshoppen formuleret som en hypotese om, at hvis børnene trives godt i dagtilbuddet, kan dagtilbuddet tage sig af nogle af de udfordringer, som ellers måske kunne have udviklet sig således, at der var behov for andre sociale indsatser. Argumentationen for hypotesen kan eksemplificeres på følgende måde; Hvis et barn trives godt i dagtilbuddet og har tillid til medarbejderne, vil det måske i højere grad være muligt for dagtilbuddet at tage hånd om udfordringer, inden de giver sig udslag i foranstaltninger eller undersøgelser fra kommunen.

Det har ikke været muligt for VIVE at underbygge denne hypotese videnskabsmæssigt, hvorfor den ikke medtages i den endelige forandringsteori for dagtilbudsområdet. Den omvendte sammenhæng er blevet belyst i forskellige undersøgelser (fx Bengtsson, Knudsen & Nielsen, 2008), men ikke sammenhængen fra trivsel og frem mod sociale indsatser.

2.3.6 Endelig forandringsteori for dagtilbudsområdet

Efter at de forbindelser, der ikke er underbygget, er undladt i forandringsteorien, giver det følgende grafiske fremstilling af forandringsteorien.

Figur 2.2 Endelig generisk forandringsteori for dagtilbudsområdet



De følgende kapitler 4-9 indeholder resultaterne af afdækningen fordelt på de seks indsats, altså indsatsniveauet. I denne gennemgang tjener de kvalificerede forandringsteorier som et input til udvælgelsen af de områder af økonomiske konsekvenser, der kan indgå i økonomiske beregninger i en model. Forandringsteorien er dog kun ét blandt flere input, når den endelige udvælgelse skal foretages.

3 Tværgående metodeovervejelser

Dette kapitel indeholder VIVEs generelle metodeovervejelser i forhold til udviklingen af en økonomisk model for de seks indsatsområder, der indgår i rapporten. En række af disse metodeovervejelser er generelle og kan ikke knyttes til en enkelt indsats, hvorfor de i stedet gennemgås samlet her.

3.1 Betragtninger vedrørende modeltype

Afdækningen i denne rapport besvarer spørgsmålet om, hvorvidt det er muligt at udvikle en model, der tager afsæt i den eksisterende SØM-model. Selvom vi i både den modeltekniske beskrivelse og i rapporten generelt tillader, at metoden afviger en del fra den, der anvendes i SØM, er der dog i selve valget af SØM som grundlag valgt en given modeltype.

Modellen, der beskrives her og i rapporten generelt, er således designet til alene at se på budgetøkonomiske konsekvenser, dvs. konsekvenser for de offentlige aktører. Der er altså ikke tale om en fuld samfundsøkonomisk model, hvor også de privatøkonomiske aktører (private husholdninger og virksomheder) indgår. Derudover er modellen også på forhånd udvalgt til at alene at måle på økonomiske outcomes – altså det, der kan måles i kroner og ører. Det betyder, at modellen ikke inddrager ikke-monetære effekter på fx trivsel eller psykisk velbefindende (undtagen som eventuel mellemregning) og ikke forsøger at værdisætte disse områder.

Ovenstående betragtninger er vigtige at holde sig for øje, hvis resultater fra modellen skal sammenlignes med resultater fra andre modeltyper, som eventuelt inddrager andre områder. Dertil er det vigtigt, at det gøres klart for brugeren, at modellen alene regner på de budgetøkonomiske konsekvenser, og at væsentlige mål som læring og trivsel ikke direkte værdisættes i modellen. Da disse mål ofte er centrale i vurderingen af indsatser på børne- og uddannelsesområdet, kan modellen ikke stå alene som beslutningsgrundlag for, hvorvidt en given indsats er formålstjenlig.

3.2 Modelteknisk beskrivelse

Den økonomiske model har til formål at give et estimat på den monetære nutidsværdi af udvalgte indsatser på BUVM's område inden for et 4-årigt perspektiv. VIVE foreslår, at modellen – ligesom SØM – udvikles med inspiration i Washington State Institute for Public Policy (WSIPP) cost-benefit-modellen². En af styrkerne ved modellen er, at den tilbyder cost-benefit-beregninger af en række indsatser inden for den samme ramme. Det giver derved et rimeligt sammenligningsgrundlag til budgetlægning/konsekvensvurdering på tværs af indsatser.

Udgangspunktet for modellen er, at en børne- og undervisningsindsats betragtes som en investering, der giver et løbende afkast over tid. Det teoretiske afsæt for denne tilgang er kompetenceudviklingsmodeller udviklet af bl.a. Nobelprismodtager James Heckman. På kort sigt forventes der en meromkostning forbundet med investeringen, som helt eller delvist kan blive opvejet af et afkast på lang sigt. Modellen sammenregner omkostninger og afkast. Da modellen inden for rammerne af denne afdækning kun har et 4-årigt perspektiv, vil der være en del mulige mellem- til langsigtede økonomiske konsekvenser, der ikke kommer med i beregningen. Et af de primære mål for investeringer – især inden for skoleområdet – er udviklingen af akademiske kompetencer (human kapital). Det samfundsmæssige afkast af dette sker først for alvor, når individerne indtræder på arbejdsmarkedet. Dette vil for langt de fleste ikke ske inden for den 4-årige tidshorisont, der er valgt som grundlag for denne rapport.

Resultatet fra modelberegningerne skal derfor tolkes som de forventede omkostninger ved en konkret børne- og undervisningsinvestering, hvor der er taget højde for konkrete kortsigtede afkast af investeringen.

Modellen er partiel i den forstand, at afkastet af investeringer vurderes lokalt, og eventuelle generelle ligevægtseffekter ikke inkluderes i beregningen. Beregningen vil eksempelvis ikke medtage betydningen af den øgede efterspørgsel efter lærere, herunder effekten på deres lønninger af indsatser med reducerede normeringer. Desuden medtages eventuelle dødvægtstab ved skatteindkrævning ikke i modellen, dvs. potentielle adfærdseffekter af højere skatter. Disse antagelser i modellen er rimelige, da modellen skal benyttes lokalt og ikke til at træffe beslutninger, der dækker hele landet. Det fremhæves desuden, at modelberegningen er af ren monetær karakter. Det vil sige, at andre relevante konsekvenser af indsatser såsom trivsel og i særdeleshed kompetenceudvikling ikke er direkte værdisat i modelberegningen, jf. ovenfor.

Det vurderes, at et af følgende to estimater (NPV_{CB} og NPV_B) skal benyttes som output fra modellen.

² Den seneste dokumentation af modellen er WSIPP (2019).

$$NPV_{CB} = \sum_{i=1}^4 \frac{\sum_{j=1}^{n_i} (Q_{ij} * P_{ij}) - C_i}{(1+r)^i}$$

$$NPV_B = \sum_{i=1}^4 \frac{\sum_{j=1}^{n_i} Q_{ij} * P_{ij}}{(1+r)^i}$$

NPV_{CB} betegner nutidsværdien af løbende omkostninger og afkast fra en konkret investering over en 4-årig periode. NPV_B foretager samme beregning, men indbefatter kun afkast fra investeringen. Begge estimater vurderes at være brugbare i en kommunal budgetlægningskontekst. I ligningen betegner i antal år efter indsatsen, og r betegner raten for afskrivning af den fremtidige værdi. Denne parameter betegner derved den vægt, som fremtidige afkast har på nuværende tidspunkt. j betegner et givet outcome-mål (fx specialundervisning), n_i er antallet af benyttede outcome-mål for indsats i , og C betegner indsatsens omkostning i et givet år. P er prissætningen af et givet outcome-mål, og Q betegner størrelsen af ændringen i outcome-målet, der opstår på baggrund af indførslen af indsatsen. Q_{ij} er derved størrelsen af ændringen i outcome-målet j , der sker i år efter indsatsen. P_{ij} er den monetære værdi pr. enhed af outcome-mål j på tidspunkt i . En beregning opdelt på kommune, region og stat kan også opnås ud fra de to formler ved at opdele i tre beregninger, hvor den eneste forskel er, at prisen justeres til at være henholdsvis den andel, der vedrører kommune, region og stat.

3.2.1 Effektstørrelse

Den centrale parameter i udregningen af nutidsværdien af løbende omkostninger og afkast fra en konkret investering over en 4-årig periode (NPV_{CB}) er størrelsen af ændringen i outcome-målet, der sker på baggrund af indsatsen (Q). For hvert relevant outcome-mål (j) består denne af et estimat på indsatsens kausale effekt på dette outcome-mål β samt et mål på det lokale niveau af outcome-målet uden indsatsen ($Base_i$). Bemærk, at brugen af lokale niveauer er en væsentlig afvigelse fra WSIPP-modellen. Årsagen er, at WSIPP-modellen er designet til brug for centraladministrationen i Washington State, hvor modelovervejelserne i denne afdækning har et kommunalt udgangspunkt. Det er derfor væsentligt, at modeloutputtet er retvisende lokalt, hvorimod WSIPP-modellen kan tage udgangspunkt i et gennemsnit på tværs af lokaliteter, da denne benyttes til at underbygge beslutningsprocessen i centraladministrationen.

Den konkrete beregning af β og $Base$ følger nedenfor. Hvis eksempelvis forskningslitteraturen har vist, at et fald i klassekvotienten fra 25 til 22 børn medfører et fald i brug af specialundervisning på 0,2 standardafvigelse (β), og standardafvigelsen på specialundervisning lokalt er 0,1 ($Base_i$), vil effektstørrelsen for ændringen i klassekvotient være en funktion af disse to tal.

$$Q_i = f(Base_i, \beta_i)$$

Denne beregning sikrer, at effektstørrelsen er baseret på forskningsmæssig evidens (β) og samtidig er repræsentativ for den lokale kontekst ($Base_i$). Den centrale antagelse her er, at der kan findes evidens for estimater på β , der afspejler effekten i en lokal kontekst. Det vil eksempelvis sige, at hvis der kun findes et β i forskningslitteraturen, og dette β er beregnet på en kontekst med lokalt specialundervisningsniveau på 0,3, så kræver antagelsen i modellen, at effekten (målt i standardafvigelser) vil være den samme i brugerens lokale kontekst. Hvis forskningslitteraturen til gengæld tilbyder en række gode estimater på β fra forskellige kontekster, vil denne antagelse være af mindre betydning. Da evidens for langt de fleste af de seks BUVIM-indsatsområder er begrænset, jf. de efterfølgende kapitler, vil der ofte være behov for en generaliseringsantagelse. Det anbefales, at der foretages en konkret vurdering af, hvorvidt det skal være muligt at foretage modelberegningen i tilfælde, hvor den lokale kontekst er langt fra det niveau, hvorfra der er forskningsevidens. Det vil eksempelvis sige i tilfælde, hvor kommunens specialundervisningsniveau er ekstraordinært højt. I dette eksempel vil man have to muligheder. Enten informerer man brugeren om den øgede usikkerhed ved estimatet eller undlader at gøre beregninger mulige for ekstraordinært høje værdier. Det vurderes, at begge tilgange vil være fagligt forsvarlige.

Det er desuden værd at fremhæve, at både $Base$ og β har subscript i , hvilket indikerer, at de tillades at ændre sig over den 4-årige periode efter indsatsen. Det må eksempelvis antages, at sandsynligheden for at deltage i specialundervisning ændrer sig i kraft med, at barnet bliver ældre, og at effekten på specialundervisning ikke nødvendigvis er den samme året efter indsatsen og 4 år senere. Det er væsentligt at holde sig for øje, at $Base$ beregnes uden indsatsen for at undgå udfordringer med, at indsatsen påvirker niveauet af $Base$ (endogenitetsudfordringer). Hvordan persistens i effekterne bør behandles, følger i afsnit 3.2.6.

3.2.2 Beregning af effektstørrelser ved direkte evidens

Den konkrete beregning af effektstørrelsen Q afhænger af, hvorvidt outcome-målet er kontinuert eller binært. Den tilgang, der benyttes, er at omregne begge til kontinuerte variable.

For outcome-mål, der i udgangspunktet er kontinuerte såsom fx testresultater, udregnes effektstørrelsen som:

$$Q_i = Base_i * \beta_i$$

I denne beregning er β omregnet som effekt-estimatet fra forskningslitteraturen (den gennemsnitlige forskel mellem indsats og kontrolgruppe) divideret med estimatets standardafvigelse. Dette omregnede estimat betegnes ofte som Cohen's d . $Base$ er i denne sammenhæng standardafvigelsen i outcome-målet i den lokale kontekst.

For outcome-mål, der i udgangspunktet er binære, såsom fx hvorvidt en elev har modtaget specialundervisning eller ej, benyttes følgende udregning af effektstørrelsen (Cox-transformation):

$$Q_i = \frac{Base_i * e^{\beta_i * 1.65}}{1 - Base_i + Base_i * e^{\beta_i * 1.65}} - Base_i$$

$Base$ er i denne sammenhæng målt i procent, og β betegner den procentvise ændring forårsaget af indsatsen. Q bliver for binære outcome-mål dermed ændringen i procentpoint.

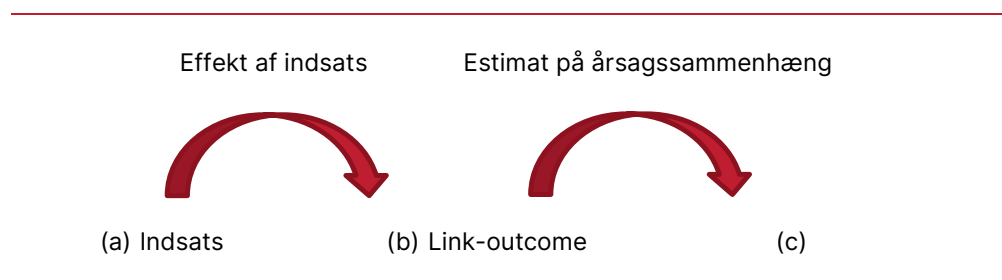
3.2.3 Beregning af effektstørrelser ved indirekte evidens

Modellen beskrevet ovenfor indbefatter følgende to trin:

1. Find i forskningslitteraturen estimater på effekten af hver indsats målt ved de relevante outcome-mål
2. Find enhedspriser for outcome-mål.

Det er imidlertid langt fra givet, at forskningslitteraturen tilbyder alle estimater på effekten af indsatser på alle relevante outcome-mål. For at imødekomme denne udfordring tilføjes et ekstra trin til modellen. Figur 3.1 illustrerer modellen ved brug af dette ekstra trin. I tilfælde, hvor en direkte sammenhæng mellem indsatsen (a) og det relevante outcome (c) ikke kan etableres, vil det være fordelagtigt at benytte et link-outcome (b). Denne situation kan opstå, hvis fx indsatsen "reduceret klassekvotient" evalueres ved testresultater, men det relevante outcome-mål i modellen er brug af specialundervisning. Testresultatet bliver i dette tilfælde benyttet som link-outcome ved at etablere en årsags-sammenhæng mellem testresultater og det ønskede outcome (specialundervisning).

Figur 3.1 Beregning af effektstørrelse ved brug af links



Proceduren ved brug af links bliver derfor som følger:

1. Find i forskningslitteraturen estimater på effekten af hver indsats på link-outcome
2. Find estimater på forholdet mellem link-outcome og de relevante outcome-mål
3. Find enhedspriser for outcome-mål.

Under punkt 1 afsøges forskningslitteraturen for studier, der evaluerer de ønskede børne- og undervisningsindsatser. Disse vil typisk være evalueret med en kort tidshorisont og med troværdige estimater på indsatsens forventede primære mål. I skolesammenhænge kan dette fx være et mål for udviklingen af human kapital som resultatændringer i faglige test i skolen. Bemærk, at uanset hvilken enhed effekten er opgjort i, så kan effekten omregnes til en effektstørrelse (effect size). Under punkt 2 afsøges derefter, hvorledes dette konkrete effektmål er relateret til de outcome-mål, der er relevante i forhold til en beregning af den monetære værdi af indsatsen i et 4-årigt perspektiv. For en indsats i skolen vil dette eksempelvis betyde, at man finder et estimat på sammenhængen mellem en skolekarakterændring og den resulterende ændring i forbruget af specialundervisning. Disse "links" mellem effekten af indsatsen og det ønskede outcome-mål fremskaffes som udgangspunkt ved afsøgning af forskningslitteraturen. Hvis forskningslitteraturen ikke tilbyder estimater på denne sammenhæng af en rimelig kvalitet (dvs. kausalitetsstudier), fremskaffes estimaterne i stedet ved egne statistiske beregninger på data.

Estimaterne på effektstørrelser af links-effekter behandles tilsvarende til effektstørrelserne på effekterne af indsatser som beskrevet i afsnit 3.2.2. Tabel

3.1 illustrerer forskellen ved beregning med direkte og indirekte evidens i beregningen af den monetære effekt af en indsats på et givet outcome-mål. Første søjle i tabellen angiver etiketter, søjle 2 viser beregningen af effektstørrelsen ved direkte evidens, søjle 3 viser den ekstra beregning, der skal til, hvis det er nødvendigt at benytte links, og man derved kun har indirekte evidens, og søjle 4 angiver, i hvor høj grad informationen er i modellen eller gives af brugeren.

I søjle fire ses det, at brugeren angiver populationsgennemsnit for både links-outcome og det endelige outcome. Disse vil, afhængig af outcome, være et mål på standardafvigelsen i den lokale kontekst. Hvis brugeren ikke forventes at kende disse tal, kan det være hensigtsmæssigt at tilbyde valgmuligheder, fx baseret på oplysninger fra registerdata.

Tabel 3.1 Modelstruktur ved direkte og indirekte evidens

	Effektstørrelse ved direkte evidens	Effektstørrelse ved indirekte evidens	Bruger input
Effekttestimat	β^{ac}	β^{ab} og β^{bc}	Nej
Populationsgennemsnit	$Base^{ac}$	$Base^{ab}$ og $Base^{bc}$	Delvis
Effektstørrelse	$f(\beta^{ac}, Base^{ac})$	$g(\beta^{ab}, Base^{ab})$ og $h(\beta^{bc}, Base^{bc})$	Nej
Total effektstørrelse (Q)	$f(\beta^{ac}, Base^{ac})$	$g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc})$	Nej
Monetær effektstørrelse	$Q(\beta^{ac}, Base^{ac}) * P$	$Q(\beta^{ab}, \beta^{bc}, Base^{ab}, Base^{bc}) * P$	Nej

En model baseret på direkte evidens vurderes at indeholde mindre bias end en model baseret på indirekte evidens. I det omfang, at forskningsstudier har påvist kausale estimater på den direkte sammenhæng mellem indsatsen og de relevante outcome-mål, bør en model baseret på direkte evidens derfor benyttes.

3.2.4 Valg mellem flere mulige link-outcomes

En given indsats kan have effekt på mange områder, fx kan en uddannelsesindsats have effekt på både trivsels- og læringsmål (samt på andre områder). Når man skal finde de økonomiske konsekvenser på afledte (SØM-)variable som fx brug af specialundervisning, sundhed m.m., skal det besluttes, hvordan de tilfælde, hvor der er flere mulige link-outcomes, skal behandles. Enten kan man tillade, at der vælges ét – og kun ét – link-outcome pr. beregning, eller

også skal man finde en metode til at sammenveje resultater af beregninger fra flere forskellige link-outcomes.

Som ovenfor kan der findes inspiration i WSIPP-modellen, der har fire regler i den forbindelse:

1. Direkte effekt tæller før indirekte effekt
2. Største effekt tælles, når der kun er positive effekter
3. Største tab og største gevinst når der er både positive og negative effekter
4. Største tab, når der kun er negative effekter.

Som en lettere omskrevet version af denne foreslår VIVE en valgmekanisme, der kan gennemføres på følgende måde: Hvis der er flere mulige link-outcomes, udregnes de samlede økonomiske konsekvenser for alle de outcomes, der er effektstørrelser for (eventuelt efter en selektering af effektstørrelser ud fra fx studiekvalitet og kontekst), og der tilbagediskonteres over de 4 år.

I praksis vil vi foreslå, at der kun anvendes et maksimalt antal link-outcomes pr. indsats, fx to lærings-outcomes og et trivsels-outcome. Ud over det rent praktiske argument, så skal der også findes relevante data for de mulige link-outcomes, før de kan inddrages.

Der anvendes kun statistisk signifikante estimater af effektstørrelser og de økonomiske konsekvenser.

Herefter kan modellen give mulighed for at vælge mellem de tre følgende:

1. Den mest positive effekt. For hver konsekvens tages den største tilbagediskonterede værdi.
2. Den mest konservative effekt. For hver konsekvens tages den mindste tilbagediskonterede værdi.
3. Gennemsnit af effekter. For hver konsekvens udregnes det simple gennemsnit af alle værdier.

Hvis der er både positive og negative beta-estimater på et område, hvor vi ville forvente positive estimater, så anser vi det for at være det mest konservative at vælge den mindste, altså den negative og mest ikke-favorable værdi, da dette giver de mindste samlede økonomiske konsekvenser.

Som i WSIPP-modellen burde direkte effekter som udgangspunkt have forrang i forhold til indirekte effekter, men det vil forventeligt ikke få megen betydning

i praksis, da der kun findes meget få direkte effekt-estimer af tilstrækkelig god kvalitet, jf. rapportens efterfølgende kapitler.

3.2.5 Beregning af effektstørrelser ved statistisk beregnet sammenhæng

I det omfang hverken indirekte eller direkte evidens er tilgængelig, kan modellen baseres på statistisk beregnede sammenhænge i data. Dette er dog kun muligt, i det omfang data på individer, der har modtaget indsatsen, og de relevante outcomes på disse individer er tilgængelige. Ved en model alene baseret på statistiske beregninger på data vil modellen følge sølje 2 i Tabel 3.1, blot ved at effekt-estimerne er beregnet på data og ikke baseret på eksisterende evidens. En model baseret alene på beregnede statistiske sammenhænge vil som udgangspunkt ikke kunne betegnes som en kausal model, og resultater fra en sådan model vil derfor ikke kunne benyttes til at konkludere på årsagssammenhænge. Dette vurderes at være en væsentlig begrænsning ved denne type model sammenholdt med en model baseret på direkte eller indirekte evidens.

3.2.6 Persistens i effekter

Da børne- og undervisningsindsatserne skal evalueres op til 4 år, efter at indsatsen påbegyndes, skal modellen kunne håndtere ændringer i effekten af indsatsen over tid. Man kan eksempelvis forestille sig, at en indsats har en stor positiv effekt på testresultater det første år, men at effekten derefter er aftagende. For andre effekter kan man forvente det modsatte: en stigning over tid.

I tilfælde, hvor direkte evidens benyttes, vil en af tre tilgange fra nedenstående prioriterede liste kunne benyttes til at tage hånd om ændringen i effekten over tid:

1. Den tilgang, der forventes at forårsage den mindste bias, er at afsøge forskningslitteraturen for estimer på effekter af indsatsen målt flere år efter indsatsen. Baseret på disse punkt-estimer vil en funktionel form kunne estimeres. Denne funktionelle form vil give et estimat på den årlige effekt op til 4 år efter indsatsen.
2. Hvis indsatsen ikke er evalueret flere år, efter den er implementeret, beholdes evalueringsestimatet fra den kortvarige evaluering, men litteraturen afsøges for estimer på persistens i outcome-målet. Det vil eksempelvis sige, at man kender effekten af en indsats på specialundervisning i år 1, men ikke i år 2, 3 og 4. Der findes til gengæld evidens for, i hvor høj grad specialundervisning i år 1 har betydning for

sandsynligheden for at deltage i specialundervisning de følgende år. Hvis denne sammenhæng er kendt for bare et par år, vil en funktionel form kunne estimeres, og dermed vil den ønskede sammenhæng kunne beregnes.

3. Hvis der kun findes kortvarige evalueringer af indsatsen, og der ikke eksisterer evidens for persistens i outcome-målet, da vil denne persistens kunne estimeres ved brug af danske registerdata.

Hvis indirekte evidens benyttes, vil de samme tre rangerede tilgange kunne benyttes, blot på link-outcome i stedet for det endelige outcome-mål.

3.2.7 Subgrupperinger

Modellen beskrevet i de tidligere afsnit er en generaliseret model, hvor effekten af indsatser måles for et gennemsnitligt individ. Det anbefales, at man i udviklingsfasen begynder med at beregne en sådan generaliseret model. Denne tilgang må dog i mange tilfælde forventes at være for aggregeret for brugeren. Man vil derfor med fordel på sigt kunne udvide modellen med subgrupperinger. For nogle brugere kunne det eksempelvis være interessant at opdele effekten af en ændring i klassekvotienten på, hvorvidt indsatsen skal ske i indskolingen, mellemtrinnet eller udskolingen. En sådan opdeling vil være mulig inden for rammerne af modellen, og antagelserne, der skal til, vil afhænge af den viden, der er tilgængelig i forskningslitteraturen. Konkret vil en opdeling i subgrupper indbefatte to elementer:

- Brugeren skal opgive *Base*-niveauet for den konkrete subgruppe, hvorpå man ønsker et mål for nutidsværdien af indsatsen, eksempelvis testresultat i mellemskolingen. Kommunale gennemsnit kan beregnes, så brugeren kun behøver at vælge kommune.
- Forskningslitteraturen skal afsøges for effektmål for alle subgrupper, man ønsker at tillade i modellen. Hvis effektmål ikke eksisterer for alle ønskede subgrupper, bør der foretages en vurdering af, hvorvidt de ikke-eksisterende effektmål kan imputeres på baggrund af de effektmål, man har for andre subgrupper. Eksempelvis bør en faglig vurdering danne grundlag for, om man kan imputere effekten på en ændring af klassekvotienten i mellemskolen, hvis denne ikke findes i forskningslitteraturen, men man har troværdige mål for indskolingen og udskolingen.

3.2.8 Usikkerhed

Det kan på sigt være hensigtsmæssigt at tilbyde brugeren information om præcisionen i estimatet af nutidsværdien af en given indsats. Metodisk vil et

estimat på usikkerheden i modellen kunne opnås ved at benytte en Monte Carlo-procedure. Denne tilgang benyttes i WSIPP og vil i princippet uden videre kunne overføres til modellen. I WSIPP præsenteres et estimat på sandsynligheden for, at en indsats samlet set giver et positivt afkast. Givet tidshorisonten i modellen er dette dog ikke et meningsfuldt mål, da børne- og undervisningsindsatser i en del tilfælde ikke må forventes at give overskud inden for den korte tidshorisont (op til 4 år). Det foreslås derfor, at man mere generelt i stedet estimerer 95 % konfidensintervaller og præsenterer brugerne for et estimat, der fortæller, at man med 95 % sikkerhed må forvente et estimat på et givet minimum. Det anbefales dog, at man venter med at indføre estimater på usikkerhed i modellen, til man har udviklet en funktionel model.

Før man beslutter at indføre mål for usikkerhed i modellen, bør man gøre sig følgende overvejelser:

- Det mål på usikkerhed, som modellen kan tilbyde, er et internt usikkerhedsmål. Det betyder, at målet giver et estimat på usikkerheden, givet at antagelserne i modellen grundlæggende er korrekte. Usikkerhedsmålet vil ikke kunne hjælpe brugeren med en forståelse af den usikkerhed, der ligger i, at modellen ikke nødvendigvis er repræsentativ i forhold til den lokale kontekst, altså den eksterne usikkerhed. Man bør derfor foretage en afvejning af, hvorvidt et sådant internt usikkerhedsmål er relevant for brugeren, og dernæst i hvor høj grad at forståelsen af usikkerhedsmålet kan kommunikeres tilfredsstillende til brugeren, så brugeren ikke misforstår målet som værende repræsentativt for al usikkerhed, der kan opstå ved forventningen til indsatsens nutidsværdi.
- For rent praktisk at kunne udføre en beregning af usikkerheden i modellen er det nødvendigt at foretage en systematisk litteraturgennemgang og derefter metaanalyser, så man opnår et ikke-selekeret billede af fordelingen af estimater i forskningslitteraturen. Det er i sidste ende disse fordelinger, der danner grundlag for estimering af usikkerheden. Herunder bør der også foretages en faglig vurdering af, hvorvidt den eksisterende empiriske forskning er tilstrækkelig bred til, at et mål på usikkerheden baseret på denne vil være retvisende.
- En anden praktisk problemstilling ved at benytte Monte Carlo-tilgangen til at opnå estimater på usikkerheden er, at det gør modeludviklingen og potentielt også modellen langsom. Det skyldes, at estimationen kræver, at alle kombinationer af input i modellen skal køres minimum 999 gange.

3.2.9 Aggregering af beregninger

Modelberegning foretages med udgangspunkt i et enkelt individ, der modtager en konkret indsats. For brugeren vil den relevante indsats ofte blive implementeret på et højere niveau, fx på klasse-, skole- eller kommuneniveau. En af

modellens grundlæggende antagelser er, at der ikke medregnes generelle ligevægtseffekter. Denne antagelse er central i omregningen af effekter fra individniveau til et højere niveau, da det gør det muligt at aggregere resultater lineært. Det vil sige, hvis der er 23 elever i en klasse, og de alle modtager en given indsats, så vil den samlede monetære effekt være 23 gange den gennemsnitlige individ-effekt.

Et andet eksempel på, at aggregeringer skal håndteres i modellen, er en indsats, der implementeres i indskolingen, hvor den ældste årgang i indskolingen modtager indsatsen i et år, mens de yngste vil modtage indsatsen i 4 år. Tilgangen vil her igen være at beregne antallet af individ-skoleår og derefter aggregere lineært. Det understreges, at denne linearitetsantagelse på nuværende tidspunkt er nødvendig for modelimplementeringen, men at den i nogen grad har potentiale til at tilføje bias til estimatet. Antagelsen vil over tid kunne løses, hvis evidens for effekten af variation i indsatsdosering forefindes i forskningslitteraturen. Da modellen har til formål at understøtte den lokale beslutningsproces, vurderes antagelsen om ikke at medregne generelle ligevægtseffekter at være rimelig, men det bemærkes, at generelle ligevægtseffekter vil være af stigende betydning i takt med, at aggregeringsniveauet hæves. Eksempelvis vil udbuddet af skolelærere ikke blive væsentligt påvirket af, at en enkelt eller to skoler ansætter flere lærere for at kunne reducere normeringerne i klasserne, men hvis denne indsats implementeres på landsplan, vil udbuddet af ledige lærere blive en væsentlig faktor, der vil skulle indgå i beregningerne.

4 To-voksen-ordninger

4.1 Indsatsens indhold og afgrænsning

To-voksen-ordninger i grundskolen indebærer, at der er to voksne til stede i klassen i den almindelige undervisning. Således er denne form for indsats relateret til normeringer og klassestørrelse. Dog er der her yderligere den mulighed, at kombinationen af de to voksne kan varieres. De to voksne kan have samme kompetencer og vil dermed kunne ses som substitutter i undervisningen (fx begge med læreruddannelse), eller de kan have forskellige kompetencer og derved ses som komplementer i undervisningen (fx en lærer og en pædagog). Den praktiske udformning af to-voksen-ordningen kan således variere meget, og disse forskellige udformninger kan principielt have forskellige effekter på de relevante outcomes. I forbindelse med arbejdet med en økonomisk model er det dog nødvendigt at betragte spørgsmålet generelt, så to-voksen-ordninger ses under ét, selvom den lokale implementering kan være forskellig. Det betyder, at de forskellige modeller af to-voksen-ordninger indgår i såvel forandringsteorier som evidensundersøgelser. Forskellige udformninger vil i nogen grad kunne håndteres i den praktiske modelopbygning ved at indregne forskellige effektstørrelser i beregningerne.

Når to-voksen-ordninger skal afgrænses til en mulig modelberegning, vil vi anbefale, at det bliver gjort individbaseret, jf. de generelle metodeovervejelser i afsnit 3.2.9. Helt konkret kan det gennemføres ved at optælle, hvor mange børn der er dækket af en to-voksen-ordning i hvor stor del af skoletiden. Det vil altså sige, at en klasse på 26 børn, hvor der er to-voksen-ordning i halvdelen af tiden, svarer til, at 13 model-børn er omfattet af ordningen hele skoleåret. Som for øvrige beregninger forudsættes det, at der er tale om en indsats, der varer hele året. Dette setup bygger på en implicit antagelse om, at indsatsens effekt er lineær i antallet af børn, der modtager indsatsen, samt i hvor intensiv (dvs. andel af timer) indsatsen er.

En problemstilling er, at man finder den ovenstående definition for generel, og man kunne ønske forskellige beregninger for underinddelinger på fx indskolingen, mellemtrinnet og udskolingen eller på forskellig uddannelsessammensætning af de to voksne i klassen. Dette vil i praksis være svært at gennemføre, da der kun delvist er vidensgrundlag til sådanne opdelinger.

4.2 Mulige mekanismer for indsatsens virkning

Figur 4.1 viser den kvalificerede forandringsteori for to-voksen-ordning. Forandringsteorien viser en teoretisk modellering af de mulige mekanismer mellem indsatsen og den økonomiske forandring på de områder, jf. kapitel 2. Mekanismemodellen siger ikke i sig selv noget om, hvordan de økonomiske beregninger kan foretages, men illustrerer blot, hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

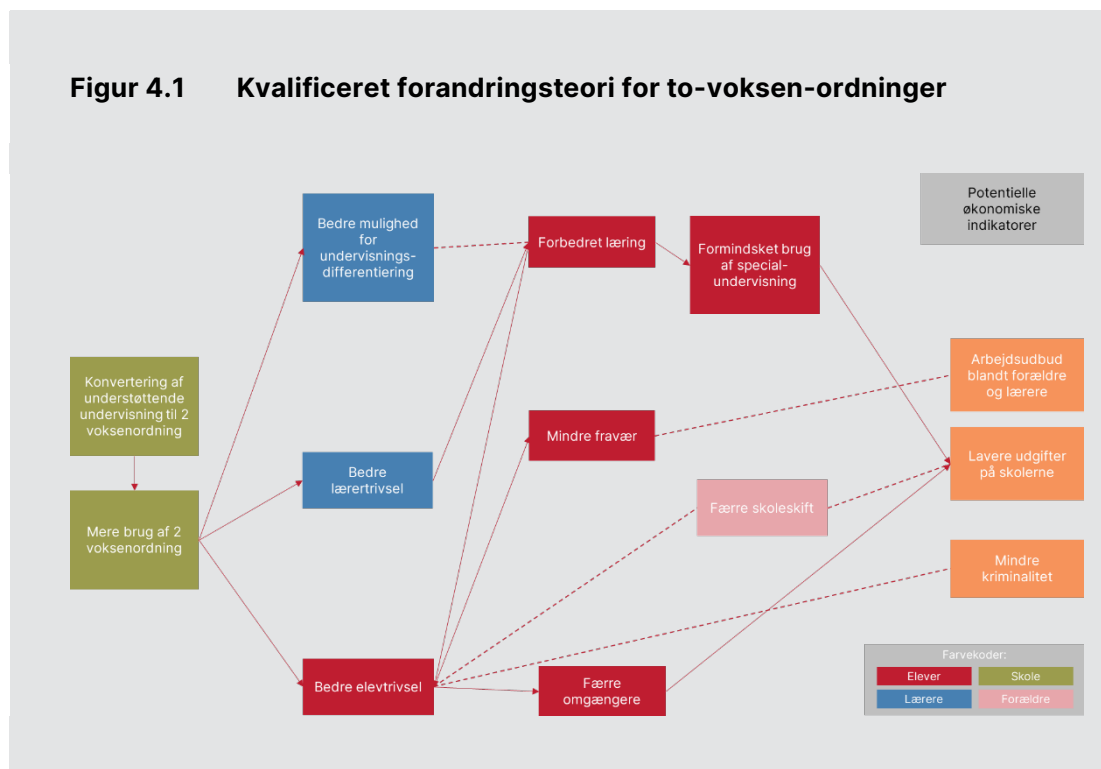
Figuren illustrerer, at vi har underbygget virkningskæder, der munder ud i tre områder for de økonomiske konsekvenser.

For det første kan der være afledte lavere udgifter på skolerne. I den kvalificerede teoretiske model kommer denne effekt som følge af henholdsvis elevernes mindre brug af specialundervisning, færre omgængere og færre skoleskift. I den teoretiske model er det henholdsvis elevernes trivsel og læring, der indvirker på disse størrelser.

For det andet viser den teoretiske model en mulig effekt på forældres arbejdsudbud, der går igennem elevens forøgede trivsel og deraf lavere fravær fra skolen.

For det tredje er det en mulig økonomisk effekt, at der vil blive begået mindre kriminalitet som følge af elevernes mulige forbedrede trivsel.

Figur 4.1 Kvalificeret forandringsteori for to-voksen-ordninger



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet. "Konvertering af understøttende undervisning til to-voksen-ordning" henviser til muligheden herfor i folkeskoleloven.

Fuldt optrukne linjer angiver stærke sammenhænge, mens stiplede linjer angiver svage sammenhænge, jf. kapitel 2.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2 og Bilag 1.

I afsnit 4.4 anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.

4.3 Evidens for indsatsen

Dette afsnit ser på den kvantitative evidens for virkningen af indsatsen, der er fundet efter en litteratursøgning.

Som nævnt i afsnit 4.1 er der store overlap mellem to-voksen-ordningen og indsatsområderne klassekvotient og normering i skolen. Der findes kun en meget begrænset litteratur, der direkte vedrører to-voksen-ordninger. Der er således kun identificeret to relevante studier af to-voksenordninger i litteraturen, hvor indsatsen passer ind under ovenstående definition af to-voksen-ordninger, og hvor kvaliteten i effektmålingen anses for tilstrækkelig høj³.

³ Epinion (2020) undersøger virkningen af, at skoler kan fravige § 14b og § 16a i folkeskoleloven om undervisningens samlede varighed og har for de frigivne ressourcer afprøvet en række indsatser,

Der findes derudover en større litteratur om brugen af undervisningsassistenter ("Teaching Aides" og "Teaching Assistants" (TA's)). En stor del af disse studier er foretaget i Storbritannien, hvor der i de seneste år har været øget fokus på brugen af disse i skolerne. Brugen af undervisningsassistenter i Storbritannien er dog ikke organiseret som to-voksen-ordninger, hvor de to voksne er i klassen, men oftest bruges undervisningsassistenter til at give specielt udfordrede elever 1:1-undervisning eller undervisning i helt små grupper, fx 1:3. Et enkelt studie (Sibieta, 2020) undersøger en indsats med undervisningsassistenter, hvor der både indgår 1:1, 1:3 og undervisningsassistenter i klassen, men effekten af de enkelte elementer er ikke undersøgt. Andre indsatser, hvor der benyttes undervisningsassistenter, er indsatser, hvor undervisningsassistenter indgår i klasseundervisningen som støtteperson for et enkelt barn, der har en form for handicap. Desuden er der identificeret en række studier, der på mere overordnet studerer sammenhængen mellem antal af undervisningsassistenter på skoler og karaktergennemsnit på skoler. Overordnet vurderer vi dog ikke, at disse studier omhandler ordninger, der er sammenlignelige med en dansk to-voksen-ordning.

Fra litteraturen er der derfor kun to studier, der er vurderet relevante i denne sammenhæng.

Penney (2018) undersøger effekten af undervisningsassistenter i projektet STAR – et projekt, der foregik i Tennessee i midten af 1980'erne og involverede mere end 11.000 elever på 80 skoler. Brugen af fuldtids undervisningsassistenter blev randomiseret på børnehaveklasseniveauet og 1. klasse. Ifølge Gerber et al. (2001) brugte undervisningsassistenterne ca. 30 % af deres tid på egentlig undervisning og de resterende 70 % på andre opgaver (fx retning af opgaver, papirarbejde mv). Effekterne estimeres som "*dynamic average treatment effects*" for forskellige kombinationer af timing i brugen af undervisningsassistenter. Det vil sige, at nogle klasser havde undervisningsassistenter tilknyttet i børnehaveklassen og ikke i 1. klasse; for andre var det omvendt, og nogle havde undervisningsassistenter i begge år. Studiet finder, at brugen af fuldtids undervisningsassistenter har positiv effekt på elevernes testscore i læsning i slutningen af 1. klasse (også for dem, der kun havde undervisningsassistenter i børnehaveklassen). Størrelsen af de estimerede effekter er 0,08-

herunder to-lærer-ordning, og finder ingen signifikante effekter på elevernes faglige trivsel eller udvikling. Da undervisningens samlede varighed ændres, når to-voksen-ordningen indføres, er der ikke tale om en effektmåling af to-voksen-ordningen, men en samlet vurdering af betydningen af disse forandringer. Resultatet er ikke udgivet i peer-reviewet tidsskrift.

Jensen et al. (2020) evaluerer folkeskolereformen og herunder muligheden for at konvertere understøttende undervisningstimer til to-voksen-undervisning (§ 16a) og finder ingen signifikante effekter på elevernes faglige resultater. Da den understøttende undervisning reduceres, når to-voksen-ordningen indføres, er der ikke tale om en effektmåling af to-voksen-ordningen, men en samlet vurdering af betydningen af disse forandringer. Resultaterne er ikke udgivet i peer-reviewet tidsskrift.

0,15 SD (standardafvigelse). Heterogene effekter viser, at disse effekter hovedsagelig skyldes positive effekter for elever fra familier med høj socioøkonomisk status og hvide studerende.

Andersen et al. (2018) evaluerer to-voksen-indsatser, som blev implementeret i Danmark i 2012-2013. Indsatsen blev rullet ud som et lodtrækningsforsøg, hvor 6. klasser på 105 skoler fik en to-voksen-indsats. Der er blev benyttet to typer af indsatser: Den ene indsats bestod i, at klassen havde to læreruddannede i klassen gennemsnitligt 11,6 lektioner om ugen i 85 % af skoleåret, mens den anden indsats bestod i, at klassen havde en lærer-uddannet og en ikke-læreruddannet i klassen gennemsnitligt 20,7 lektioner om ugen i 85 % af skoleåret. Studiet identificerer statistisk signifikante effekter på resultater af læsetests, men ingen signifikante resultater for matematiktests. De signifikante effektstørrelser ligger i størrelsesordenen 0,06-0,17 SD og er altså beskedne. Der er en tendens til, at effekterne er størst, når den ekstra voksne ikke er uddannet som lærer, men effekterne er dog ikke signifikant forskellige. Estimationer af heterogene effekter viser, at effekterne er størst for elever med lavtuddannede forældre, hvor effekten estimeres til 0,17 SD.

I lodtrækningsforsøget bag Andersen et al. (2018) deltager udelukkende 6. klasser, og det er derfor ikke muligt at forudsige, om der vil findes tilsvarende effekter af at indføre to-voksen-indsatser på andre klassetrin. Analysen i Andersen et al. (2018) foretages for kortsigtede outcomes, og der er for nuværende ikke offentliggjort eventuelle resultater for langsigtede outcomes.

Fundene fra de to studier er sammenfattet i nedenstående tabel.

Tabel 4.1 Sammenfatning af evidensviden for to-voksen-ordninger

Evidensområde	Evidens	Indgår nordiske?	Noter
Børn			
Test score	+	Nej	Tendens til størst effekt for elever med lavtuddannede forældre.
Læsetest	+	Ja	
Matematiktest	0	Ja	

Anm.: Et plus indikerer, at en *forbedring* i normeringen (flere elever dækket af en to-voksen-ordning) har en positiv effekt på udfald; omvendt for et minus.

De to studier ser ikke på effekter, der ligger ud over de her præsenterede, og eventuelle afledte økonomiske effekter på personale og/eller forældre er således ikke analyseret. Der blev dog fundet evidens, der kan bruges som indirekte-(link)outcome. Baseret på evidensundersøgelsen anbefaler vi, at man tager udgangspunkt i resultaterne fra studiet Andersen et al. (2018), hvori der findes en effekt på læsning på 0,09 SD. Effekt-estimatet omregnes til en effektstørrelse, der kan anvendes i en beregning.

4.4 Modeludvikling

Vi gennemgår mulighederne for modeludvikling med basis i de tre evidensniveauer direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Det er på nuværende tidspunkt ikke muligt at gennemføre beregninger for to-voksen-ordninger baseret på direkte evidens, da en sådan evidens ikke foreligger.

Indirekte evidens

Outcomes og link-outcome

Figur 4.2 illustrerer modellen ved et indirekte evidensniveau. Indsatsen er i afsnit 4.1 modelteknisk afgrænset til et fuldtidsindivid udsat for én ekstra voksen i klassen. Estimatet på effekten på læsning fra studiet Andersen et al. (2018) kan benyttes til at etablere evidens for indsatsens effekt. I studiet fremgår en gennemsnitlig effektstørrelse på 0,09 SD⁴ for 20,7 lektioner om ugen i 85 % af skoleåret. Dette estimat omregnes til modellens enhed ved en lineær justering. Det vil sige, at effekt-estimatet, som anbefales brugt i modellen, er $0,09 * (33/20,7) / 0,85 = 0,17$, når der omregnes til en helårseffekt.⁵

Som link-outcomes til beregningen bør anvendes resultater fra de nationale test i læsning. Disse gennemføres i 2., 4., 6. og 8. klasse. Ønsker man særskilt at se på de ældste børn, kan resultatet fra afgangsprøven i dansk benyttes. Valget af link er i dette tilfælde i høj grad betinget af den begrænsede tilgængelige evidens, men der foretages samtidig en vurdering af, at disse link-out-

⁴ Det er almindeligt at måle effekt i antal standardafvigelsers forbedring af det anvendte effektmål. Denne metode sikrer, at effektstørrelser, der måles på forskellige effektmål, kan sammenlignes. Målet, der også kaldes Cohen's d, er også omtalt i kapitel 3.

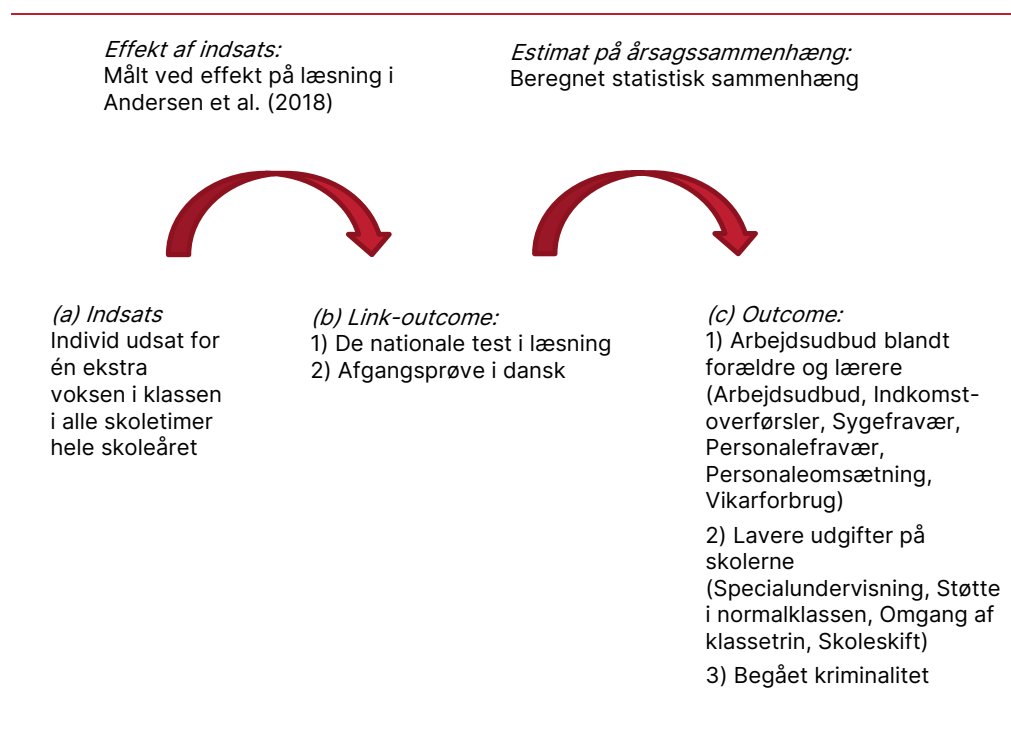
⁵ De 33 timer er det antal timer, der er i 6. klasse.

comes forventes at fange en væsentlig del af den effekt, der er anført i forandringsteorien. Det forventes således, at testresultatet i væsentlig grad er korreleret med forbedring i læring, formindsket fravær og højere trivsel i klassen.

Det foreslås, at relationen mellem link-outcome og de økonomiske konsekvenser bliver beregnet baseret på statistiske sammenhænge i data. I tilfælde, hvor der forefindes evidens for en relation mellem link-outcome og et konkret økonomisk outcome, vil dette dog med fordel kunne erstatte den statistisk beregnede relation. En beregning af relationen mellem link-outcome og outcome åbner rent teknisk op for muligheden for at inkludere alle outcomes, der er tilgængelige data på (jf. Bilagstabel 2.1). Hvis man tager den kvalificerede forandringsteori som udgangspunkt i udvælgelsen af relevante outcomes, vil man kunne inkludere følgende outcomes: 1) Arbejdsudbud blandt forældre og lærere (Arbejdsudbud, Indkomst-overførsler, Sygefravær, Personalefravær, Personaleomsætning, Vikarforbrug), 2) Lavere udgifter på skolerne (Specialundervisning, Støtte i normalklassen, Omgang af klassetrin, Skoleskift) og 3) Begået kriminalitet. Det vil også være muligt at se på overgangen til ungdomsuddannelse for de ældste elever.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL (Styrelsen for IT og Læring) og DST (Danmarks Statistik), hvor den betingede korrelation mellem outcome og link-outcomes beregnes, fx den betingede korrelation mellem brug af specialundervisning og de nationale test i læsning. I regressionerne betinges på standard socio-demografiske baggrundsvARIABLE fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset og baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret og tilgængeligheden af de nationale tests, der samlet kun dækker folkeskolen med informationer fra 2010 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for problemer i forbindelse med multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Figur 4.2 Model for indirekte evidens for to-voksen-ordninger



Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen. Dette baseline-niveau kan baseres enten på en lokal værdi for den kommune, beregningen skal dække, eller på en landsdækkende gennemsnitlig værdi. Det kan ikke forventes, at brugeren ligger inde med disse parametre. Derfor skal modellen tilbyde opdaterede tal fra alle landets kommuner på disse parametre. Disse tal beregnes på baggrund af danske registerdata som årlige, kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerte outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer for det valgte link-outcome. Det vil i denne sammenhæng sige, at der skal anvendes standardafvigelse på testresultater i læsning fra de nationale test for hver kommune samt landsgennemsnittet.

Beregn effektstørrelser for hvert relevant outcome

Som angivet i afsnit 3.2.3 består estimatet af den monetære effekt på et outcome i et givet år efter indsatsen af følgende beregning:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat fra Andersen et al. (2018), $Base^{ab}$ er det lokale baselineniveau på testresultatet i læsning, β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome, og $Base^{bc}$ er baseline-niveauet for det relevante outcome. Da testresultatet er en kontinuert variable, reduceres udtrykket for kontinuerte outcomes til:

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

Og for binære outcomes til:

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right)$$

Beregningen foretages på individniveau for hvert indsatsår. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsatsår ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimaterne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode, er der behov for estimater på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette for to-voksen-ordninger, da indsatsen kun er evalueret, umiddelbart efter indsatsen er udført. Ved brug af link-outcome fra testresultater i læsning kan persistensen estimeres. Populationen er igen begrænset af tilgængeligheden af de nationale tests til folkeskoleelever fra 2010 og frem. De nationale test i læsning gennemføres på hvert andet klassetrin. Dette benyttes til at beregne den betingede korrelation i testresultater mellem to efterfølgende test-år. Estimatet benyttes til at foretage en lineær ekstrapolation over den 4-årige periode, hvorved der opnås et mål for den betingede korrelation mellem testresultater i år 1 og testresultater i de følgende 3 år.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi pr. individ, der modtager den fulde indsats som anført i afsnit 3.2. Hvis brugeren eksempelvis ønsker at beregne effekten af en indsats med en ekstra voksen i klassen i 10 klasser, hvor den ekstra voksne er til stede 10 % af skoleåret i 2 år, så omregnes dette først til antal fuldtidseksponerede individer. I dette tilfælde vil det være 1 klasse (23 individer) i 2 år, dvs. 46 modelindivider. Dernæst aggregeres estimatet pr. individ op med dette antal individer.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistisk sammenhæng, er det nødvendigt at kunne knytte viden om, hvilke børn der har været omfattet af to-voksen-ordninger til de efterfølgende realiserede outcomes. Dette er i øjeblikket ikke muligt, da der ikke findes data på individniveau om anvendelse af to-voksen-ordninger af tilstrækkelig god kvalitet.

4.5 Ny evidens og ny data

Som nævnt ovenfor mangler der viden om, hvilke børn, klasser og skoler der har haft eller er i gang med to-voksen-ordninger. Data herom vil være et godt udgangspunkt for både fremtidige beregninger og for fremtidige effektmålinger.

Derudover påviser evidensundersøgelsen også generelt, at der er ret få studier, der har beskæftiget sig med to-voksen-ordninger. Derfor vil det være tilrådeligt at gennemføre et studie, der ud over effekterne på testresultater også ser på andre mål som fx elevtrivsel, brug af specialundervisning, omgang og påvirkning af forældres arbejdsudbud. Man kunne eksempelvis tage udgangspunkt i de elever, der deltog i lodtrækningsforsøget evalueret i Andersen et al. (2018) og måle på flere effekter for disse elever. Dog vil der måske være et problem med manglende statistisk power i analyser på områder med små forventede effekter, da en stor del af populationen i studiet ikke er i målgruppen for fx specialundervisning.

5 Normering i dagtilbud

5.1 Indsatsens indhold og afgrænsning

Normering i dagtilbud beskriver en relation mellem antallet børn og voksne i et dagtilbud. Med normering menes typisk antal børn pr. voksen, hvilket betyder, at en større ratio indebærer flere børn pr. voksen, mens en lavere ratio omvendt indebærer færre børn pr. voksen. En indsats rettet mod normering i dagtilbud vil typisk indbefatte en form for grænseværdi for normering, der kan opnås ved enten at justere tæller (antallet af børn) eller nævner (antallet af voksne) i ratioen.

Normering i dagtilbud betegnes som en strukturel kvalitetsparameter i dagtilbud, som umiddelbart lader sig afkode og måle. Dette er dog mere kompliceret end som så. Som angivet i Christoffersen & Højen-Sørensen (2013) spiller en lang række forhold ind herunder: sygefravær (både børnenes og de voksnes), afspadsering, ferie, børnenes uanmeldte fridage, børnenes aktiviteter, administrative opgaver samt observationstidspunkt.

Kvalitet i dagtilbud er mere end normering og strukturel kvalitet. Der sondres i litteraturen mellem strukturel kvalitet, proceskvalitet og resultat-kvalitet (EVA, 2017). I en SØM-sammenhæng vil både strukturelle og processuelle kvalitetsdimensioner kunne indgå som indsatser i det omfang, de lader sig kvantificere og manipulere. Man kunne eksempelvis forestille sig SØM-beregninger i relation til indsatser, som forbedrer den processuelle ledelseskvalitet. Resultat-kvalitetsdimensionen er også relevant, da data på børnenes kompetencer og færdigheder vil kunne fungere som link-outcomes (i det omfang der findes data her på). I denne afdækning belyses alene normering i dagtilbud, men det sker med fuld anderkendelse af, at kvalitet opererer i mange dimensioner i samspil med normering.

Når normering i dagtilbud skal afgrænses til en mulig modelberegning, vil vi anbefale, at det bliver gjort individbaseret, jf. de generelle metodeovervejelser i kapitel 3. Ved en indsats, hvor normeringen eksempelvis ændres halvdelen af dagen for 12 børn, svarer det til, at 6 modelbørn er omfattet af ændringen hele dagen. Som for de øvrige beregninger forudsættes det, at der er tale om en indsats, der varer hele året. Dette setup indebærer en implicit antagelse om, at indsatsens effekt er lineær i gruppestørrelse og antal voksne, samt hvor intensiv indsatsen er.

Som begreb dækker dagtilbud over så forskellige institutioner som vuggestuer, dagplejen, børnehaver og aldersintegrerede institutioner. Modellen antager i udgangspunktet, at betydningen af normering er ens i alle typer institutioner. Som det udfoldes i det følgende, bliver denne antagelse udfordret af evidensen på området, hvorfor en model i bedste fald som minimum må sondre mellem de forskellige typer af institutioner. Der er eventuelt mulighed for at slå dels vuggestuer og vuggestuelignende dele af aldersintegrerede institutioner, dels børnehaver og børnehavelignende dele af aldersintegrerede institutioner sammen. Det forventes ikke, at dagplejen vil kunne indgå i modellen, da aldersspændet er for stort.

5.2 Mulige mekanismer for indsatsens virkning

Figur 5.1 viser den kvalificerede forandringsteori for normering i dagtilbud. Den kvalificerede forandringsteori viser en teoretisk modellering af de mulige økonomiske effekter af ændret normering på de områder, hvor vi har haft mulighed for at underbygge de opstillede mekanismer, jf. kapitel 2. Den kvalificerede model siger ikke i sig selv noget om, hvordan de økonomiske beregninger kan foretages, men illustrerer blot, hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

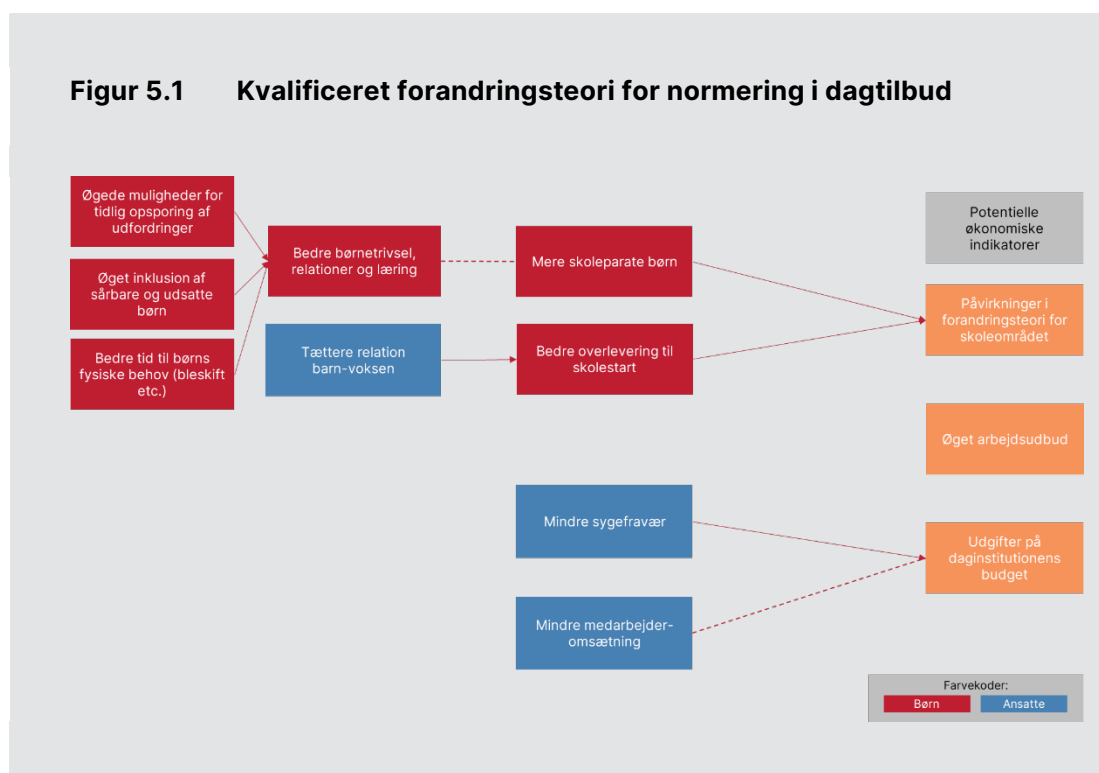
Figuren illustrerer, at vi har fundet underbyggede økonomiske outcomes på tre områder.

For det første kan forbedret normering medføre bedre børnetrivsel, bedre relationer og læring, hvilket kan give mere skoleparate børn og deraf positive påvirkninger på skolen. I den kvalificerede forandringsteori kommer denne effekt som følge af henholdsvis elevernes mindre brug af specialundervisning, færre omgængere, mindre brug af udsat skolestart og færre skoleskift.

For det andet viser den teoretiske model en mulig effekt på øget arbejdsudbud. Som det beskrives i næste afsnit, er der fundet et centralt studie, som netop undersøger sammenhængen mellem normeringer i dagtilbud og forældres arbejdsudbud. Denne mekanisme blev ikke identificeret under brainstormingen ved udarbejdelsen af forandringsteoriene, hvorfor den ikke er inkluderet i forandringsteorien i afsnit 2.3. Studiet forekommer dog yderst relevant i denne sammenhæng.

For det tredje er det en mulig økonomisk effekt, som relaterer sig til diverse strukturelle forhold omkring institutionen og personalets adfærd.

Nedenfor anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet. Kassen "Påvirkninger i forandringsteori for skoleområdet" henviser til, at virkingskæderne her kan fortsætte, når barnet træder ind i skolealderen, og derigennem iværksætte nogle af forandringkæderne for indsatser på skoleområdet.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

5.3 Evidens for indsatsen

Evidensen på effekten af normering i dagtilbud er forholdsvis begrænset. Årsagen hertil kan være, at universelle dagtilbud, som det kendes fra Danmark, det øvrige Norden og Holland, kun i mindre grad benyttes andre lande.

Der er fundet to systematiske reviews af betydningen af normering: Perlman et al. (2017) samt Christoffersen, Højen-Sørensen & Laugesen (2014)⁶. Perlman et al. (2017) fokuserer på *Early Childhood Education and Care (ECEC)* for børn mellem 30 og 72 måneder. Ud fra en stikprøve på 6.988 studier opfylder blot 29 studier inklusionskriterierne til en kvalitativ syntese, mens blot 3 studier overlever frem til metaanalysen. I alt rapporteres der på i studierne på

⁶ Der er ligeledes et Campbell-review fra VIVE. Arbejdet er ved afdækningens afslutning endnu ikke så langt, at det har kastet resultater af sig.

83 forskellige udfald i kategorierne kognitive kompetencer (10 studier), sproglige kompetencer (21 studier), matematiske kompetencer (12 studier) samt sociale og emotionelle kompetencer (18 studier). I den kvalitative syntese fremhæver Perlman et al. (2017), at 2 ud af 10 studier finder en signifikant sammenhæng mellem forbedret normering og kognitive kompetencer. 5 ud af 21 studier finder en signifikant sammenhæng mellem forbedret normering og børnenes sproglige kompetencer, mens 2 studier finder den modsatte sammenhæng. 2 ud af 12 studier rapporterer signifikante resultater mellem normering og matematik pegende i hver sin retning. Slutteligt viser 4 ud af 18 studier en signifikant sammenhæng mellem forbedret normering og børnenes adfærd. I metaanalysen finder Perlman et al. (2017) en samlet korrelation mellem normering og en standardiseret ordforrådstest på 0,03 (*PPVT, Peabody Picture Vocabulary Test*) med et 95 % konfidensinterval på 0,00-0,05, hvilket betyder, at korrelationen kun er på grænsen til at være signifikant.

I et større review af daginstitutionernes betydning for børns udvikling berører Christoffersen, Højen-Sørensen & Laugesen (2014) også normeringer. I den kvalitative syntese sammenfatter de, at bedre normering har betydning for trykgere tilknytning til pædagoger, bedre voksen-barn-interaktioner samt kognitiv og sproglig udvikling. Ligeledes giver bedre normering bedre social tilpasning og bedre adfærd. Endelig peger bedre normering også i retning af lavere personaleomsætning, da bedre normering giver lavere sygefravær blandt personalet. Det studie, der henvises til vedrørende personalets sygefravær, er et arbejdspapir, som sidenhen er blevet udgivet (Gørtz & Andersson, 2014) som en artikel, hvori det findes, at normering har betydning for sygefraværet i vuggestuer, mens det er insignifikant for sygefraværet i børnehaver.

Sammenlignet med Perlman et al. (2017) er Christoffersen, Højen-Sørensen & Laugesen (2014) ikke så systematisk i sin afrapportering. Der er ikke foretaget et egentligt metareview, ligesom Christoffersen, Højen-Sørensen & Laugesen (2014) heller ikke dokumenterer, hvor stor en andel af de præsenterede studier der underbygger en afrapportering af et signifikant fund. Reviewet er en central rapport i dansk litteratur om normeringer i daginstitutioner, men i en operationel SØM-sammenhæng er Christoffersen, Højen-Sørensen & Laugesen (2014) mindre relevant, da der er behov for at kunne afløfte *effect sizes* (eller oplysninger, der kan omregnes hertil) som input i modellen.

Kjeldsen (2018) undersøger sammenhængen mellem normering og forældrenes arbejdsudbud på både den ekstensive og intensive margin. Ud af 27 forskellige specifikationer af udfaldsmål finder Kjeldsen (2018) 6 signifikante estimater, som alle peger i retning af øget arbejdsudbud – særligt for mødre – ved forbedret normering. Der anvendes simple OLS-regressioner, hvorfor kvaliteten af evidensen relativ lav.

Tabel 5.1 opsummerer den samlede evidensviden på normeringer i dagtilbud for både reviews og enkeltstudier. Der er fundet studier af normeringers effekt på børns kognitive kompetencer, sproglige kompetencer, matematiske kompetencer, sociale kompetencer samt eksamensresultater. Ligeledes er der afdækket studier af normerings betydning for sygefravær blandt personalet samt arbejdsmarkedsdeltagelse for forældrene. Helt overordnet viser de inkluderede studier et blandet billede af effekter af normering. For alle outcomes findes eksempler på både positive og 0-effekter af forbedret normering. For matematiske kompetencer findes tilmed, at de matematiske kompetencer reduceres ved forbedret normering. Det skal indvendes, at Tabel 5.1 skal tages med det forbehold, at antallet af inkluderede studier i alle tilfælde er relativt lavt.

Tabel 5.1 Sammenfatning af evidensviden for normering i dagtilbud

Evidensområde	Evidens	Indgår nordiske?	Noter
Børn			
Kognitive kompetencer	0/+	Nej	Overvejende insignifikante resultater
Sproglige kompetencer	0/+	Nej	
Matematiske kompetencer	-/0/+	Nej	
Sociale kompetencer	0/+	Nej	
Eksamensresultater	0/+	Ja	
Pædagoger			
Sygefravær	0/+	Ja	Kun signifikante effekter for vuggestuepædagoger
Forældre			
Arbejdsmarkedsdeltagelse	0/+	Ja	Lavt evidensniveau

Note: Et plus indikerer, at en *forbedring* i normeringen har en gunstig effekt på udfald. Omvendt for et minus.

5.4 Modeludvikling

Vi gennemgår mulighederne for modeludvikling med basis i de tre evidensniveauer direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Inspireret af forandringsteorien og på baggrund af den fundne evidens kan der principielt foretages direkte beregninger af normeringer i dagtilbud for to outcomes:

- Vuggestuepædagogers sygefravær baseret på Gørtz & Andersson (2014)
- Forældres arbejdsmarkedsdeltagelse baseret på Kjeldsen (2018).

Der er dog et væsentlig forbehold at tage her, da beregningerne i begge tilfælde vil basere sig på kun et studie, som i begge tilfælde tilmed har et relativt lavt evidensniveau. På den baggrund kan VIVE derfor ikke anbefale en model, som bygger på direkte evidens.

Indirekte evidens

Ved brug af indirekte evidens ser mulighederne lidt bedre ud. Som beskrevet i afsnit 3.2.3 består proceduren af tre trin, når der anvendes indirekte evidens:

1. Find i forskningslitteraturen estimerer på effekten af hver indsats på link-outcome
2. Find estimerer på forholdet mellem link-outcome og de relevante outcome-mål
3. Find enhedspriser for outcome-mål.

Link-outcomes og outcomes

Perlman et al. (2017) fandt i et metareview en korrelation mellem normering og en standardiseret test for ordforråd på 0,03. Korrelationen blev estimeret til at være på grænsen til det signifikante. Hvis det forventes, at denne sammenhæng også gælder en dansk kontekst, og estimatets usikkerhed kan accepteres, vil det være muligt at benytte indirekte evidens og link-outcomes som grundlag for beregningerne, idet estimatet fra Perlman et al. (2017) kan anvendes i første trin af proceduren.

Den standardiserede test for ordforråd anvendt i metaanalysen af Perlman et al. (2017) anvendes ikke bredt og systematisk i Danmark. Der er derfor nødvendigt at afsøge andre, lignende datakilder. Det, der kommer nærmest, er:

1. Sprogvurderinger i forbindelse med skolestart (i begyndelsen af 0. klasse)
2. Nationale test i læsning fra 2. klasse.

Der ses bort fra trivselsmål i skolen som link-outcome, da der ikke er tilstrækkelig evidens herfor. Da der anvendes link-outcomes, som ligger tidsmæssigt umiddelbart efter afslutningen af dagtilbuddet eller efter dagtilbuddet, så giver det kun mening at fortolke resultaterne af beregningen for de ældste børn i dagtilbud.

Næste trin i proceduren er at komme fra link-outcomes til outcomes. Her findes to metoder. Første metode er at afsøge forskningslitteraturen for relevante estimater, mens anden metode er at afdække statistiske sammenhænge mellem link-outcome og outcome. En beregning af relationen mellem link-outcome og outcome åbner rent teknisk for muligheden for at inkludere alle outcomes, der er tilgængelige data på. Hvis man har den kvalificerede forandringsteori som udgangspunkt i udvælgelsen af relevante outcomes, vil man kunne inkludere følgende outcomes:

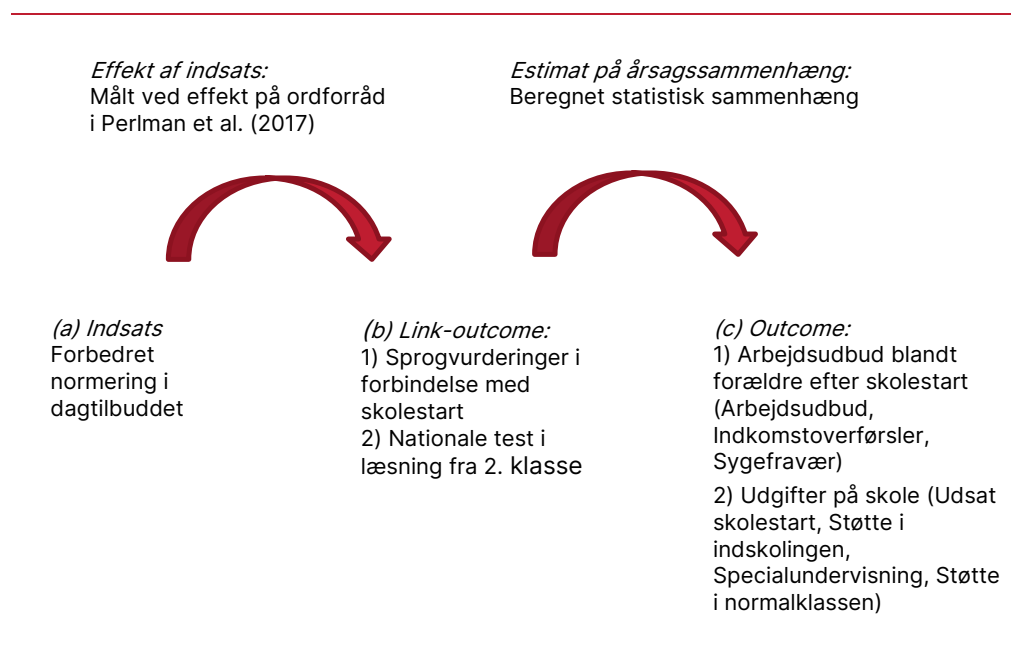
1. Arbejdsudbud blandt forældre efter skolestart (Arbejdsudbud, Indkomstoverførsler, Sygefravær)
2. Udgifter på skole (Udsat skolestart, Støtte i indskolingen, Specialundervisning, Støtte i normalklassen).

Det er vigtigt at forholde sig til den metodemæssige distinktion mellem indirekte evidens og ren statistisk sammenhæng. Når det gælder indirekte evidens, kan der med afsæt i den kvalificerede forandringsteori anvendes statistiske sammenhænge mellem link-outcome og outcome, hvis forskningslitteraturen ikke tilbyder estimater mellem de to. Når det gælder rene statistiske sammenhænge, menes statistiske sammenhænge mellem indsats og outcomes. Det vil sige, at link-outcomes i denne situation slet ikke benyttes. Der vendes tilbage til rene statistiske sammenhænge sidst i dette afsnit.

Hvis litteraturen ikke tilbyder estimater mellem link-outcome og outcome, må der foretages en estimation. Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og link-outcomes beregnes, fx den betingede korrelation mellem brug af specialundervisning og de nationale test i læsning i 2. klasse. I regressionerne betinges på standard socio-demografiske baggrundsvARIABLE fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret og tilgængeligheden af de nationale tests, der samlet kun dækker folkeskolen med informationer fra 2010 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Nedenstående figur viser de forskellige trin, hvor estimatet fra Pearlman et al. (2017) anvendes som bro mellem indsats og link-outcomes.

Figur 5.2 Model for indirekte evidens for normering i dagtilbud



Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen. Dette baseline-niveau kan enten baseres på en lokal værdi for den kommune, beregningen skal dække, eller på en gennemsnitlig national værdi. Det kan ikke forventes, at brugeren ligger inde med disse parametre. Derfor skal modellen tilbyde opdaterede tal fra alle landets kommuner på disse parametre. Disse tal beregnes på baggrund danske registerdata som årlige, kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerte outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer på det valgte link-outcome. I dette eksempel betyder det, at der skal findes baseline-niveauer for hver af de to link-outcome kandidater:

1. Sprogvurderinger i forbindelse med skolestart
2. Nationale test i læsning fra 2. klasse.

Disse baseline-niveauer skal findes på kommunalt og nationalt niveau. Derved vil brugeren efter valg af kommune blive præsenteret for en række tal, som eventuelt kan rettes til efter behov.

Beregning af effektstørrelser for hvert relevant outcome

Beregningen af effektstørrelser for normeringer i daginstitutioner vil være ækvivalent til metoderne, der anvendes ved andre indsatser. Som angivet i afsnit 3.2.3 består estimatet af den monetære effekt på et outcome i et givet år efter indsatsen af følgende beregning:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat fra Perlman et al. (2017), $Base^{ab}$ er det lokale baselineniveau på testresultatet (de nationale test i 2. klasse eller sprogvurderinger), β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome (fx udsat skolestart), og $Base^{bc}$ er baseline-niveauet for det relevante outcome (fx procentsats af børn, der har udsat skolestart). Da testresultatet er en kontinuert variabel, reduceres udtrykket for kontinuerte outcomes til:

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

og for binære outcomes til:

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right).$$

Beregning foretages på individniveau for hvert indsatsår. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsatsår ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimerne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode, er der behov for estimer på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette, da indsatserne i den relevante litteratur kun er evalueret, umiddelbart efter indsatsen er udført. Til gengæld er det muligt at bringe link-outcomes i spil igen og ved hjælp heraf estimere persistensen.

Udfordringen er dog, at populationen alt efter valg af link-outcome er begrænset af tilgængeligheden af henholdsvis de nationale tests og sprogvurderinger i forbindelse med skolestart for en flerårig periode. For de nationale tests vedkommende vil matematik-testen i 3. klasse kunne bringes i spil, mens det er sværere at finde yderligere målepunkter for sprogvurderinger ved skolestart. I tillæg er data fra sprogvurderinger i forbindelse med skolestart først og kun delvist tilgængelige fra 2016.

Persistens findes ved at beregne den betingede korrelation i testresultater mellem to efterfølgende testår. Dette estimat benyttes derpå til at foretage en lineær ekstrapolation over den 4-årige periode. Nedenstående tabel angiver en oversigt over datapunkter i persistensberegningen, hvis nationale test i 2. klasse anvendes som link-outcome. I eksemplet er indsatsen foretaget i år 0, hvilket vil sige sidste år i dagtilbud. Det er reelt muligt at ekstrapolere over en længere periode, om end usikkerheden også stiger.

Tabel 5.2 Eksempel på datapunkter i beregningen af persistens

År	Institutionstrin	Datapunkt
0	Daginstitution – indsats	–
1	0. klasse	Ekstrapoleret
2	1. klasse	Ekstrapoleret
3	2. klasse	DNT dansk
4	3. klasse	DNT matematik

En alternativ tilgang vil være at beregne persistens i outcomes. Dette sker implicit ved at estimere sammenhængen mellem link-outcome og outcome direkte. Denne metode lader sig derfor relativt let implementere, og det er muligt at tage højde for persistens direkte ved estimation af et persistenskorrigeret β^{bc} benævnt $\tilde{\beta}^{bc}$.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi pr. modelbarn, der modtager den fulde indsats som anført i afsnit 3.2.5. Dernæst aggregeres estimatet pr. modelbarn op med antallet af børn.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistiske sammenhænge, er det nødvendigt at kunne knytte viden om, hvilke børn der har været

omfattet af en indsats til de efterfølgende realiserede outcomes. Disse data haves som udgangspunkt ikke på individniveau for normering i dagtilbud, men kan konstrueres på baggrund af aggregerede normeringsdata på kommune- eller områdeniveau, hvis man er villig til at acceptere den approksimation af den reelle normering, dette indebærer. Statistikbanken fra DST har en opgørelse over normering på kommuneniveau opdelt på institutionstype, der vil kunne anvendes som kilde⁷. I skrivende stund dækker data årene 2015-2019. Der vil være en del usikkerhed i forbindelse med en sådan approksimation, som alt andet lige vil have en afledt negativ effekt på modellens præcision.

Metoden til at beregne den statistiske sammenhæng er, at der for hvert inkluderet outcome gennemføres en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og den normering, der gældende for de pågældende børn, beregnes, fx den betingede korrelation mellem brug af specialundervisning og indsatsen. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST som eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret, dvs. fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Med dataudfordringerne in mente vurderer VIVE samlet set, at der ikke er basis for at konstruere en model alene baseret på statistiske sammenhænge.

5.5 Ny evidens og ny data

VIVE vurderer, at der på nuværende tidspunkt ikke er basis for at lave en model på normeringer i dagtilbudsområdet. Det er teknisk muligt, men den forskningsmæssige evidens er sparsom, både når det gælder direkte og indirekte evidens, og begrænsningerne i data gør en model baseret alene på statistiske sammenhænge for usikker.

Med afsæt i den kvalificerede forandringsteori blev der identificeret to studier, som belyste den direkte sammenhæng mellem normeringer i dagtilbud og relevante outcomes. Desværre havde begge studier et relativt lavt evidensniveau, hvorfor det ikke var tilrådeligt at basere en model alene på dem.

⁷ www.statistikbanken.dk/BOERN3

En model baseret på indirekte evidens vil kræve et solidt estimat mellem indsats og et eller flere link-outcomes. Som et bud herpå blev et estimat på normeringens betydning for resultater i en ordforrådstest fra et meta-review diskuteret. Estimatet baserede sig på tre studier og var på grænsen til det signifikante. Det vil derfor klart kunne anbefales at gennemføre velspecificerede effektmålinger af effekten af normeringer i dagtilbud.

Yderligere vil en model på baseret på indirekte evidens være begrænset af mangel på gode datakilder på kompetenceniveau for børn i dagtilbud, der kan bruges som link-outcome. På dagtilbudsområdet vil ny data med kompetencemål (både mestring, trivsel og læring) være meget brugbare til at beregne indirekte evidens af indsatser i dagtilbud.

Et andet område, hvor datakvaliteten kan forbedres, er ved at oprette et detaljeret register, der er i stand til at koble børn og det personale i dagtilbud, de har daglig kontakt med. Dette vil gøre det muligt at udvikle en model baseret på statistiske sammenhænge og på sigt danne grundlaget for forskningsevvidens, der kan bruges til en model baseret på direkte evidens.

6 Udsat skolestart og omgængere

6.1 Indsatsens indhold og afgrænsning

Udsat skolestart og omgængere beskriver begge indsatser, hvor barnet får et ekstra år med det samme læringsindhold – enten i børnehaven eller i skolen. I forhold til udsat skolestart vil en indsats indebære, at et barn starter et år senere i skolen end det år, hvor det fylder 6 år. Barnet modtager herved et ekstra års undervisning/udvikling i børnehaven. I forhold til omgængere vil den mest relevante indsatsdefinition være, at barnet modtager undervisningen på et klassetrin to gange og dermed får 1 års undervisning mere med det samme pensum.

Modelberegninger foretages på individniveau. Den monetære effekt pr. individ-år skal efterfølgende aggregeres i forhold til omfanget af den lokale indsats, dvs. antallet af modelbørn (individ-år), som indsatsen berører.

I tilfældet med børn, der får udsat deres skolestart eller går et klassetrin om, er begge indsatser på individniveau og berører derfor kun det enkelte barn. Antallet af modelbørn vil derfor være lig antallet børn, der får udsat deres skolestart eller går et klassetrin om.

6.2 Mulige mekanismer for indsatsens virkning

Figur 6.1 viser den kvalificerede forandringsteori for udsat skolestart/omgængere. Forandringsteorien viser en teoretisk modellering af de mulige mekanismer mellem indsatsen og den økonomiske forandring på disse områder, jf. kapitel 2. Den kvalificerede forandringsteori siger ikke i sig selv noget om, hvordan de økonomiske beregninger kan foretages, men illustrerer blot, hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

Figuren illustrerer, at vi på skoleområdet har fundet evidens-underbyggede økonomiske outcomes på tre områder: arbejdsudbud blandt forældre og lærere, lavere udgifter og mindre kriminalitet. Når den 4-årige tidshorisont tages i betragtning, er kriminalitet dog kun et relevant outcome for de ældste elever

i skolen.⁸ For omgængere kan det eventuelt også være relevant at inddrage overgang til ungdomsuddannelse, om end de færreste omgængere er i de ældre skoleklasser.

På skoleområdet kan udsat skolestart og omgang give afledte lavere udgifter på skolerne. I den kvalificerede forandringsteori kommer denne effekt som følge af elevernes mindre brug af specialundervisning, færre omgængere (se-
nere hen) og færre skoleskift. I den teoretiske model er det henholdsvis elevernes trivsel og læring, der indvirker på disse størrelser. Vedrørende omgængere kan der dog være en modsatrettet effekt i øgede udgifter på grund af øget kapacitetsbehov samt behov for flere ansatte i skolen. Der kan også på dagtilbudsområdet være afledte øgede udgifter på grund af øget kapacitetsbehov og behov for flere ansatte.

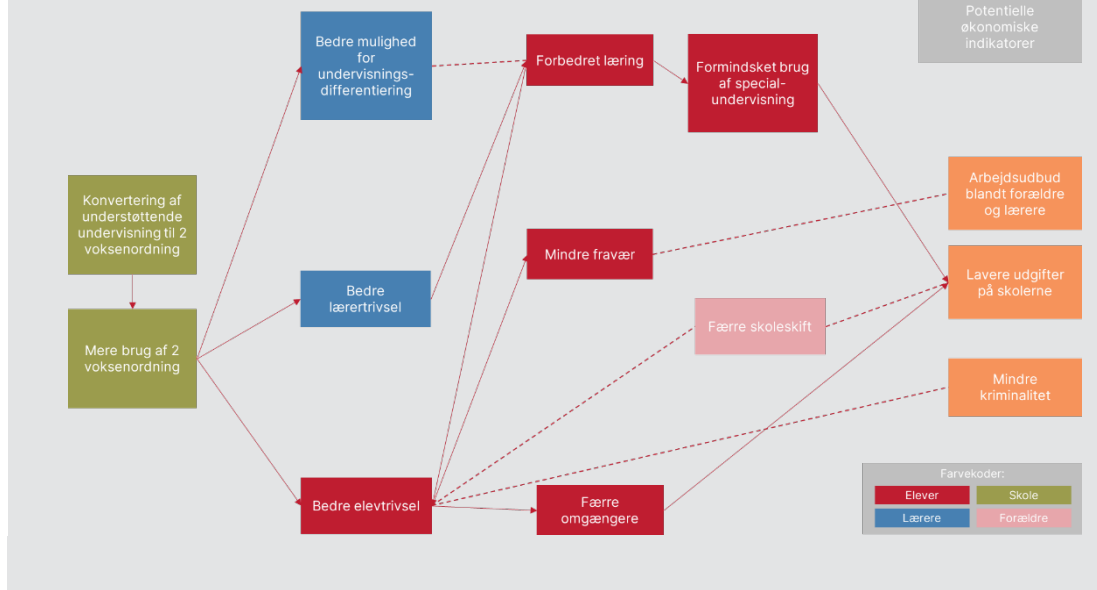
For det andet viser den teoretiske model en mulig effekt på forældres arbejdsudbud, der går igennem elevens forøgede trivsel og deraf lavere fravær fra skolen.

For det tredje er det en mulig økonomisk effekt, at der vil blive begået mindre kriminalitet som følge af elevernes mulige forbedrede trivsel.

I afsnit 6.4 anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.

⁸ Grundet den 4-årige horisont indgår andre ellers relevante outcomes som beskæftigelse ikke.

Figur 6.1 Kvalificeret forandringsmodel for udsat skolestart/omgængere



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet. De fuldt optrukne linjer angiver stærkt evidensunderbyggede forbindelser, mens de svagt underbyggede er markeret med stiplede streger.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

6.3 Evidens for indsatsen

Dette afsnit beskriver den fundne evidens om udsat skolestart og omgængere. For denne indsats er det BUVM, der har været ansvarlige for indsamlingen af evidens, og fremstillingen er baseret på det materiale, ministeriet har leveret til VIVE (VIVE har ikke foretaget en vurdering af, om dette materiale er dækkende, idet denne vurdering forudsættes foretaget af BUVM). Præsentationen af den fundne evidensviden er opdelt i områderne "udsat skolestart" og "omgængere", da der er tale om to forskellige indsats. Der er fundet evidens for en række områder, herunder mål vedrørende elevens skoletid som læring, ikke-kognitive mål, elevadfærd, omgang og skolegang, men også mål for kriminalitet, helbred, beskæftigelse og indkomst som voksen.⁹ Der er ikke fundet evidens for effekten på specialundervisning.

Det skal bemærkes, at mange studier kun belyser effekten af udsat skolestart indirekte, da 'indsatsen', der evalueres, er alder ved skolestart. Denne indsats

⁹ I evidensundersøgelsen er medtaget alle outcomes, også de, som ikke er med i den kvalificerede forandringsteori.

dækker over naturlig aldersvariation blandt børn, som indskoles til tiden i forhold til de officielle regler – og altså ikke får udsat deres skolestart. Den naturlige aldersvariation fremkommer ved at børn, der er født langt før 'cut'-datoen for indskoling (i Danmark er det tidligt på året/foråret), vil være ældre ved indskoling end børn født tættere på cut-datoen. Den naturlige aldersvariation ved almindelig indskoling vil være på op til 1 år (hvor nogle elever indskoles som 5,5-årige og andre som 6,5-årige). Denne 'indsats' har det til fælles med indsatsen 'udsat skolestart', at eleven vil være ældre ved indskoling (typisk vil elever, der skulle indskoles som 5,5-årige, få udsat deres skolestart, til de er 6,5 år). De to indsatses adskiller sig dog ved, at der kan være stigma forbundet med at få udsat sin skolestart, hvilket må forventes at trække effekten af udsat skolestart i negativ retning sammenlignet med effekten af at være født først på året. Det må desuden forventes, at børn, der har fået udsat skolestarten, adskiller sig fra andre skolebørn på en række parametre, der er relateret til årsagen til den udsatte skolestart. Denne forskel skal tages med i vurderingen af relevansen af den pågældende evidens.

De fundne studier er generelt af høj kvalitet, idet de alle anvender metoder med statistisk kontrolgruppe. De mest brugte metoder i den fundne evidens er forskellige varianter af regression discontinuity-designet (RDD), instrumentvariabel-metoder og matching. Der er ikke fundet studier, der er baseret på randomiserede forsøg.

Meget af evidensen er baseret på amerikanske data, men der er også studier fra andre lande (Canada, Italien, Tyskland), heriblandt også studier, der belyser problematikken med danske data og data for andre nordiske lande, som dermed er af særlig relevans i en dansk kontekst. For del-indsatsen *omgængere* er der dog alene fundet evidens fra USA.

For indsatsen *udsat skolestart* finder mange studier nul-effekter, men en del finder også mellemstore til store positive resultater. Dog findes for nogle evidensområder som kriminalitet, egen beskæftigelse og indkomst/løn, også negative resultater. Nogle resultater er dog meget kontekst-afhængige. For eksempel viser et studie, at effekten af en højere alder ved indskoling medfører højere kriminalitet (Cook & Kang, 2016). Studiet er gennemført i amerikansk sammenhæng, hvor undervisningspligten ophører ved en fast aldersgrænse (mellem 16-18 år, forskelligt fra stat til stat) og ikke som i Danmark efter et fast antal skoleår. Det betyder, at jo senere elever begynder i skole, jo kortere er undervisningspligten i praksis. For elever med en højere alder ved indskoling vil der gå længere tid mellem, at undervisningspligten ophører for dem og den formelle afslutning af skolegangen med et high school-diplom. Det medfører i praksis, at flere af de relativt sent-indskolede vælger at droppe ud uden diplom, hvilket igen relateres til højere kriminalitet. I et skolesystem som det danske, hvor længden af undervisningspligten ikke afhænger af alderen ved indskoling, vil denne mekanisme ikke være relevant.

Med hensyn til studier af særlig relevans er der fundet tre danske studier, som alle er af høj kvalitet (benytter alle regression discontinuity-design). Ligesom for de internationale studier finder en stor del af evidensen nul-effekter, men også positive effekter, som i de danske studier gennemgående er store.

For indsatsen *omgængere* er resultaterne meget blandede: her findes en stor del nul-effekter og ellers effekter både i positiv¹⁰ og negativ retning. Alle fundne studier er fra USA, og der er dermed ikke fundet evidens for effekter af at gå 1 år om i en dansk eller nordisk sammenhæng.

Fundene fra evidensundersøgelsen er sammenfattet i Tabel 6.1.

Tabel 6.1 Sammenfatning af evidensviden for udsat skolestart og omgængere

Evidensområde	Evidens	Indgår nordiske?	Noter
Udsat skolestart/alder ved skolestart			
Læring	0/+	Ja	Måles først og fremmest som test scores
Ikke-kognitive mål	+	Nej	Fx selvkontrol, tilgang til læring m.m.; baseret på ét studie
Gå en klasse om	+	Nej	
Mentalt helbred	0/+	Ja	Fx adfærdsvanskeligheder, ADHD, emotionelle vanskeligheder, opmærksomhed
Kriminalitet	-/0/+	Ja	Fx sigtet, dømt; se bemærkningen ovenfor vedrørende den negative effekt
Forældres beskæftigelse	0/+	Ja	Kun målt for mødre; baseret på ét studie
Egen beskæftigelse	-/0/+	Ja	Fx (fuldtids-)beskæftigelse
Indkomst/løn	-/0/(+)	Ja	Løn, indkomst
Fysisk helbred	0	Nej	baseret på ét studie
Skolegang	+	Ja	Fx opnået uddannelse, længde af skolegang, indskrivning i avancerede kurser/uddannelser, fastholdelse i uddannelse
Modtager af sociale ydelser som voksen	0/+	Ja	Baseret på ét studie
Omgængere			
Elevadfærd	0/(+)	Nej	Fx suspendering, fravær, engagement
Kriminalitet	0	Nej	Fx forbrydelse, forseelse; baseret på ét studie
Læring	-/0/+	Nej	Bl.a. testscorer
Motivation	+	Nej	Fx motivation til videre uddannelse, værdisættelse af uddannelse; baseret på ét studie
Omgænger senere	+	Nej	Omgang på senere klassetrin; baseret på ét studie

¹⁰ Bemærk, at resultatet for de to outcomes kun med positivt fortegn i tabellen (motivation og omgang i senere år) må tages med et forbehold, idet det i begge tilfælde er baseret på et enkelt studie.

Evidensområde	Evidens	Indgår nordiske?	Noter
Skolegang	-/0/+	Nej	Fx højeste gennemførte klassetrin, alder når eleven forlader skolen, high school-påbegyndelse og gennemførelse
Trivsel	0	Nej	Følelsen at høre til i skolen; baseret på ét studie

Anm.: Plustegnet indikerer, at senere skolestart eller omgang har en gavnlig/ønskelig effekt på outcomes. For læring betyder det fx, at testscorer er højere/bedre, mens det for andre outcomes fx indikerer lavere kriminalitet, *ikke* at gå en klasse om eller ikke at modtage sociale ydelser. Parenteser i søjlen 'Evidens' indikerer, at den angivne effektretning er mindre stærkt evidensunderbygget.

Studier med danske data er særligt relevante for eventuel anvendelse i en økonomisk model. Der findes en række danske studier, som bruger kvasi-eksperimentelle metoder og er dermed på et højt evidensniveau (der er ikke gennemført RCT'er vedrørende udsat skolestart eller omgang).

Når vi tager udgangspunkt i de tre outcomes på skoleområdet, findes der direkte dansk evidens for to samt indirekte evidens for det tredje. Med hensyn til kriminalitet finder Landersø, Nielsen & Simonsen (2015), at børn, der startede et år senere i skole, fordi de blev født i starten af januar i stedet for i slutningen af december, ikke oplevede en lavere sandsynlighed for at begå kriminalitet end børn, der startede skole rettidigt. Et andet studie af de samme forfattere (Landersø, Nielsen & Simonsen, 2019) finder effekter for mødrenes arbejdsudbud. Helt konkret finder de, at hvis et barn er ældre ved skolestart, forbedres moderens beskæftigelse i gennemsnit i årene omkring skolestart.

Der er ikke fundet direkte evidens for det sidste outcome (udgifter i skolen), men indirekte evidens via læring.¹¹ Det vurderes, at de mest relevante studier for modeludviklingen er i Landersø, Nielsen & Simonsen (2015) og Sievertsen (2015). Denne vurdering er foretaget på baggrund af følgende kriterier: 1) Studierne er danske, 2) der er tale om studier med et godt metodedesign, og 3) effekterne gælder for en bred elevgruppe.

Sievertsen (2015) finder, at børn, der er ældre ved skolestart, klarer sig bedre i dansk i testene i 2., 4. og 6. klasse. Der er ingen effekt på testresultatet i 8. klasse. I matematik er der en positiv effekt af en senere skolestart på præstationer i 3. klasse, men ikke i 6. klasse. Effekttørrelserne for de signifikante effekter er alle omkring 0,5 SD og må betegnes som store.

Landersø, Nielsen & Simonsen (2015) viser til gengæld, at sen skolestart medfører højere karakterer i matematik ved folkeskolens afgangsprøve (men ikke i

¹¹ Der er endvidere evidens på senere omgang, men ikke i nordiske studier.

dansk). Effektstørrelsen er omtrent $0,4 SD^{12}$ og dermed i samme størrelsesorden eller lidt mindre end Sievertsens resultater for de mindre klasser.

Generelt er mønstret hos Sievertsen (2015), at der er effekter i de mindre klasser, men ikke i de store, mens Landersø, Nielsen og Simonsen (2015) finder en positiv effekt i 9. klasse. Konklusionen med hensyn til de større klasser i de to studier stemmer derfor ikke overens. Det kan dermed være relevant at begrænse modelberegningen til de mindre klasser.

6.4 Modeludvikling

Nedenfor gennemgår vi mulighederne for modeludvikling med basis i de tre evidensniveauer direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Det ville være muligt at gennemføre beregninger for udsat skolestart baseret på direkte evidens, da der foreligger evidens på højt kvalitetsniveau for Danmark vedrørende kriminalitet og forældrenes arbejdsudbud. Det vurderes dog, at kriminalitet som outcome i praksis ikke er centralt i forhold til indsatsen udsat skolestart, da det ligger uden for den 4-årige beregningshorisont, som betragtes i denne afdækning. Hvis det alligevel har interesse, vil relevante estimater kunne hentes hos Landersø, Nielsen & Simonsen (2015).

Vedrørende forældrenes arbejdsudbud vil estimaterne af effekten af udsat skolestart fra Landersø, Nielsen & Simonsen (2020) kunne anvendes. I forhold til en 4-årig horisont er det estimaterne for moderens beskæftigelse, når barnet går i 1.-3. klasse, der er relevante (der er ingen estimater for 0. klasse). Effektstørrelserne ligger i alle 3 år omkring 0,05, dvs. udsat skolestart øger sandsynligheden for, at moderen er i beskæftigelse med 5 procentpoint.

Indirekte evidens

Outcomes og link-outcomes

For et sidste outcome – skoleudgifter – er der ikke direkte evidens, men der vil kunne frembringes indirekte evidens via link-outcome. Det vurderes på nuværende tidspunkt ikke at være meningsfuldt at gennemføre beregninger for omgængere, da der for nuværende ikke findes relevant evidens. I det følgende beskrives derfor modellen for udsat skolestart. Hvis der senere frembringes

¹² Beregnet som gennemsnit af 0,3 for piger og 0,5 for drenge.

relevant evidens for omgængere, vil en tilsvarende model for omgang kunne anvendes.

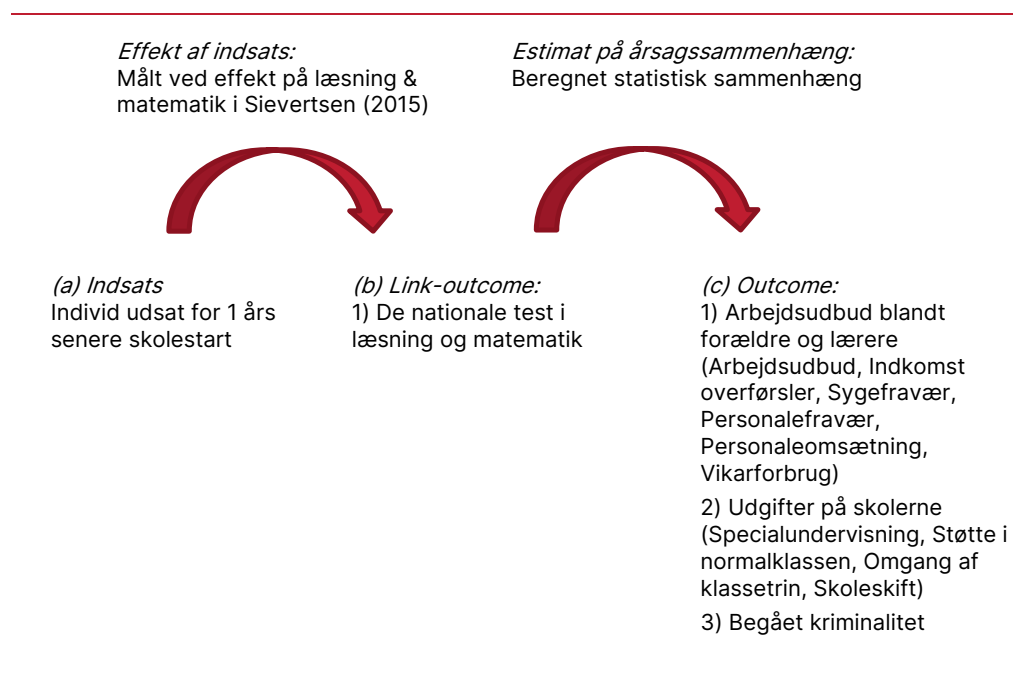
Figur 6.2 illustrerer modellen ved et indirekte evidens-niveau. Indsatsen er i afsnit 6.1 modelteknisk afgrænset til et barn, der starter 1 år senere i skolen, end den overordnede regel for skolestart tilsiger. Hos Landersø, Nielsen & Simonsen (2015) er der kun læringseffekter i slutningen af 9. klasse, hvilket ligger uden for den 4-årige horisont. Det vil derved være studiet af Sievertsen (2015), der skal benyttes som empirisk grundlag. Hos Sievertsen findes store effekter på elevernes læse- og regnekompetencer i 2. og 3. klasse på 0,56 SD¹³. Med en standardafvigelse på omtrent 25 point svarer det til 14 point på 100-point skalaen.

Som link-outcome til beregningen for skoleudgifterne bør anvendes resultater fra de nationale test i læsning og matematik for 2. og 3. klasse. Valget af link er i dette tilfælde i høj grad betinget af den begrænsede tilgængelige evidens, men der foretages samtidig en vurdering af, at disse link-outcomes forventes at fange en væsentlig del af effekt baseret på forandringsteorien. Det forventes således, at testresultatet i væsentlig grad er korreleret med forbedring i læring, formindsket fravær og højere trivsel i klassen.

Det vil også være muligt at tilføje andre skole-outcomes, hvis dette skønnes at være relevant (jf. Bilagstabel 2.1). Der foretages for hvert inkluderet outcome en regression baseret på data fra DST, hvor den betingede korrelation mellem outcome og link-outcome beregnes, fx den betingede korrelation mellem brug af specialundervisning og resultat fra de nationale tests. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængeligheden af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for test af multiple hypoteser ved brug af Bonferroni-korrektion eller lignende.

¹³ SD = standardafvigelse. Det er almindelig kutyme i forskningslitteraturen at angive effekter i standardafvigelser, som gør, at størrelsen af effekter kan sammenlignes på tværs af forskellige outcome-mål.

Figur 6.2 Model for indirekte evidens for udsat skolestart



Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen. Dette baseline-niveau kan baseres enten på en lokal værdi for den kommune, beregningen skal dække, eller på en landsdækkende gennemsnitlig værdi. Det kan ikke forventes, at brugeren ligger inde med disse parametre. Derfor skal modellen tilbyde opdaterede tal fra alle landets kommuner på disse parametre. Disse tal beregnes på baggrund af danske registerdata som årlige kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerede outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer for det valgte link-outcome, dvs. for standardafvigelsen på testresultater i læsning og matematik fra de nationale test for hver kommune samt landsgennemsnittet.

Dette resulterer i en model, hvor brugeren blot behøver at vælge kommune. Brugeren vil derefter blive præsenteret for en række tal, som kan rettes til efter behov.

Beregn effektstørrelser for hvert relevant outcome

Som angivet i afsnit 3.2 skal der for hvert outcome foretages en beregning af den monetære effekt af indsatsen. Dette skal gøres for hvert af de 4 år, efter indsatsen blev implementeret. Når alle monetære effekter er regnet ud, skal de lægges sammen som beskrevet i afsnit 3.2.9. Den monetære effekt af en indsats på et givet outcome beregnes på følgende måde:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat på testresultater fra Sievertsen (2015), $Base^{ab}$ er det lokale baseline-niveau på testresultatet i læsning og matematik, β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome, og $Base^{bc}$ er baseline-niveauet for det relevante outcome. Da testresultatet benyttes som link-outcome for både dagtilbud og skole, og da testresultater er en kontinuert variable, reduceres udtrykket for kontinuerte outcomes til:

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

og for binære outcomes til:

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right)$$

Beregningen foretages på individniveau for hvert indsats-år. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsats-år ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimaterne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode, er der behov for estimater på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette for udsat skolestart, da indsatsen kun er evalueret i henholdsvis år 3 og 4, efter indsatsen er udført (dvs. i 2. klasse for læsning og 3. klasse for matematik). Ved brug af link-outcomes fra testresultater i læsning og matematik kan persistensen (dvs. den vedvarende effekt) estimeres. Populationen er begrænset af tilgængeligheden af de nationale tests til folkeskoleelever fra 2010 og frem. Dette benyttes til at beregne den betingede korrelation i testresultater mellem to efterfølgende test-år. Man har derved opnået et mål for, hvor persistente

testresultater er. Dette estimat benyttes til at foretage en lineær ekstrapolation over den 4-årige periode. Man opnår derved et mål for den betingede korrelation mellem testresultater i år 1 og testresultater i de følgende 3 år.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi pr. individ, der modtager den fulde indsats som anført i afsnit 3.2. Dernæst aggregeres estimatet pr. individ op med dette antal individer.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistisk sammenhæng, er det nødvendigt at kunne knytte viden om, hvilke børn der har været omfattet af en indsats til de efterfølgende realiserede outcomes.

For udsat skolestart og omgængere er dette muligt ved brug af DST's elevregistre, der har observationer fra 2008 og frem. Der foretages for hvert inkluderet outcome en regression baseret på data fra DST, hvor den betingede korrelation mellem outcome, og hvorvidt individet har fået udsat deres skolestart, henholdsvis er gået en klasse om, bliver beregnet, fx den betingede korrelation mellem brug af specialundervisning og indsatsen. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret, dvs. fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

6.5 Ny evidens og ny data

Hvis man ønsker en model for udsat skolestart baseret på direkte evidens af andre mål for forældrenes arbejdsudbud end beskæftigelse (fx Indkomstoverførsler, Sygefravær) eller for outcomes, der vedrører udgifter på skolerne (Specialundervisning, Omgang af klassetrin, Skoleskift), kan man med fordel tage udgangspunkt i den empiriske strategi i Sievertsen (2015) og Landersø, Nielsen & Simonsen (2015), men benytte de relevante outcomes som afhængige variable i stedet for testresultater. Det vurderes, at der p.t. ikke er brug for yderligere data vedrørende udsat skolestart.

For at kunne udvikle en model for omgængere, skal der tilvejebringes dansk (eller nordisk) evidens på højt kvalitetsniveau for outcomes (eventuelt link-outcomes). Det vurderes, at der p.t. ikke er brug for yderligere data vedrørende omgængere for at opnå dette, men det kræver gennemførelse af et større eksperiment. Eksperimentet kunne med fordel være et naturligt eksperiment, mens et lodtrækningsforsøg vurderes som mindre oplagt. Et naturligt eksperiment kunne enten designes som en gradvis implementering i kommuner af øget brug af omgang, hvor kommuner, der implementerer, i første bølge sammenlignes med kommuner, der implementerer senere. En anden tilgang kunne være at anvende testresultater sammen med en fast grænse for at fastlægge, hvilke elever skal gå en klasse om. Her ville elever, der lige netop ikke når op til grænsen og derfor må gå klassetrinet om, sammenlignes med elever, der lige netop opnår et testresultat, der gør, at de får lov at fortsætte til næste klassetrin. Før man overvejer at sætte et eksperiment i gang, bør det dog selvsagt vurderes nøje, om det er realistisk/ønskeligt at gennemføre et eksperiment omkring omgængeri.

7 Klassekvotient

7.1 Indsatsens indhold og afgrænsning

I forskningslitteraturen skelnes mellem klassekvotient og klassestørrelse. I en forskningsoversigt (Filges, Sonne-Schmidt & Nielsen, 2018) argumenteres, at reduktion af henholdsvis klassestørrelse og klassekvotient (elev-lærer-ratio) er to forskellige – om end relaterede – indsatser, og der fokuseres snævert på klassestørrelse. Klassekvotienten – den gennemsnitlige klassestørrelse på årgangs- eller skoleniveau – benyttes ofte som approksimation for klassestørrelse i anvendelser, hvor data om antallet af elever i den enkelte klasse ikke er til rådighed. Den gennemsnitlige klassestørrelse anvendes også som instrument for den faktiske klassestørrelse (Jakubowski & Sakowski, 2006). En anden fordel ved at benytte den gennemsnitlige klassestørrelse på årgangs- eller skoleniveau er, at bias (inden for årgangen, mellem klasserne) i estimaterne kan være mindre. I denne rapport indgår både studier baseret på klassekvotient og klassestørrelse.

I modelsammenhæng vil en indsats indebære, at klassestørrelsen reduceres med en elev. Skaleringsenheden vil her være, at barnet modtager undervisning i en klasse med én elev færre et helt skoleår. Når indsatser vedrørende reduktion af klassestørrelse skal afgrænses til en mulig modelberegning, vil vi anbefale, at det bliver gjort individbaseret, jf. de generelle metodeovervejelser i afsnit 3.2.9. Helt konkret kan det gennemføres ved at optælle, hvor mange børn der er dækket af ændringen i hvor stor en del af skoletiden. Ændringer i klassestørrelsen gælder typisk alle – eller langt de fleste – undervisningstimer,¹⁴ dvs. at det i en klasse på 26 børn svarer til, at 26 modelbørn er omfattet af ordningen hele skoleåret. Som for øvrige beregninger forudsættes det, at der er tale om en indsats, der varer hele året. Dette setup indebærer en implicit antagelse om, at indsatsens effekt er lineær i antallet af børn, der modtager indsatsen, samt i hvor intensiv (dvs. andel af timer) indsatsen er.

7.2 Mulige mekanismer for indsatsens virkning

Figur 7.1 viser den kvalificerede forandringsteori for klassestørrelse. Forandringsteorien viser en teoretisk modellering af de mulige mekanismer mellem indsatsen og den økonomiske forandring på disse områder, jf. kapitel 2. Den

¹⁴ I de større klasser kan undervisningen i enkelte fag, fx tysk/fransk, finde sted i andre grupper end basisklassen.

kvalificerede forandringsteori siger ikke i sig selv noget om, hvordan de økonomiske beregninger kan foretages, men illustrerer blot, hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

Figuren illustrerer, at vi har fundet økonomiske outcomes på tre områder, hvor der også er fundet en evidens-underbygning af mekanismerne: arbejdsudbud blandt forældre og lærere, udgifter på skolerne og mindre kriminalitet. Når den 4-årige tidshorisont tages i betragtning, er kriminalitet dog kun et relevant outcome for de ældste elever i skolen.¹⁵

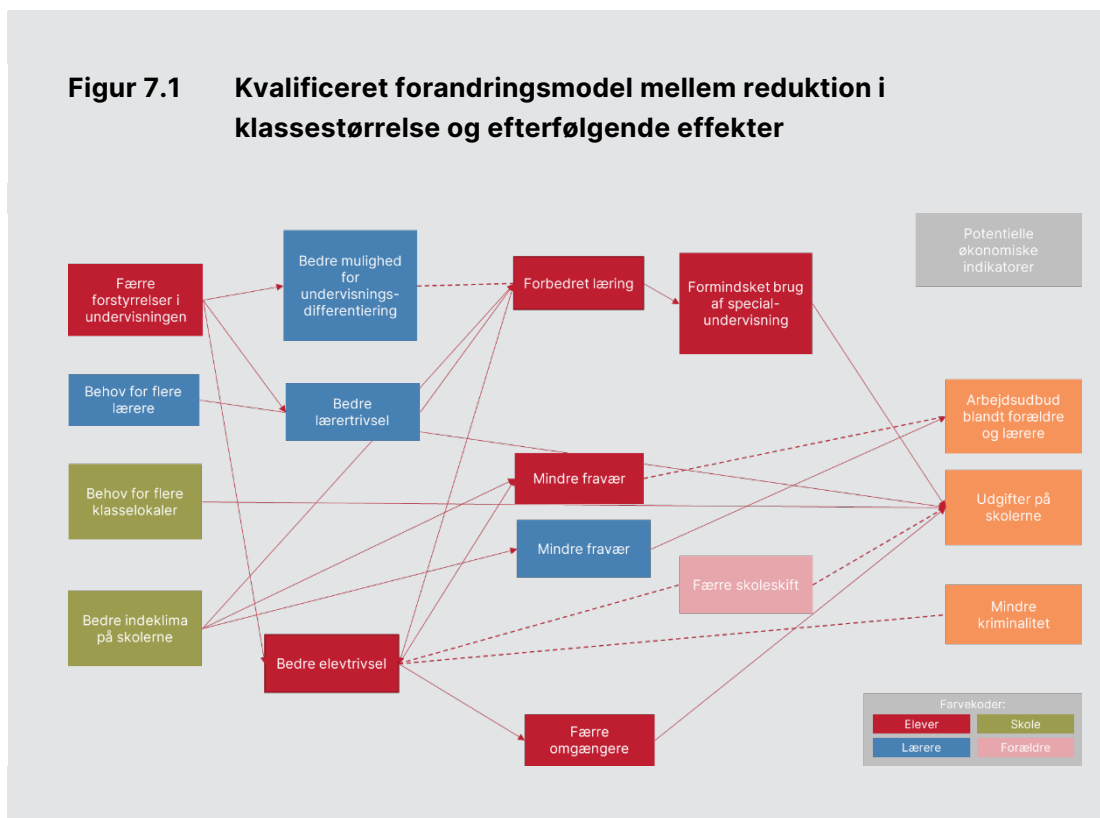
For det første kan der være afledte ændringer i udgifterne på skolerne. I den kvalificerede forandringsmodel kommer denne effekt som følge af elevernes mindre brug af specialundervisning, færre omgængere og færre skoleskift. I den teoretiske model er det henholdsvis elevernes trivsel og læring, der indvirker på disse størrelser.

For det andet viser den teoretiske model en mulig effekt på forældres arbejdsudbud, der går igennem elevens forøgede trivsel og bedre indeklima på skolerne og deraf lavere fravær fra skolen blandt elever og lærere.

For det tredje er det en mulig økonomisk effekt, at der vil blive begået mindre kriminalitet som følge af elevernes mulige forbedrede trivsel.

¹⁵ Grundet den 4-årige tidshorisont indgår andre ellers relevante outcomes som beskæftigelse ikke.

Figur 7.1 Kvalificeret forandringsmodel mellem reduktion i klassestørrelse og efterfølgende effekter



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet. De fuldt optrukne linjer angiver stærkt evidensunderbyggede forbindelser, mens de svagt underbyggede er markeret med stiplede streger.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

I afsnit 7.3 anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.

7.3 Evidens for indsatsen

Dette afsnit ser på den kvantitative evidens for virkningen af indsatsen, der er fundet efter en litteratursøgning.

Langt de fleste fundne studier fokuserer på klassestørrelsens effekt på læring, oftest målt ved testresultater eller karakterer. Der er dog også studier, der belyser ikke-kognitive resultatmål som fx selvkontrol, engagement, opmærksomhed. Endvidere er der studier med længere-sigtede resultatmål som opnået uddannelsesnivea, løn mv., men den er betydeligt mindre. Der er ikke fundet studier, der undersøger effekter af klassestørrelse på andre evidensområder,

som kan være relevante, såsom om mindre klasser gør, at elever i mindre grad udskilles til segregeret specialundervisning.¹⁶

Der er en del litteratur om effekter af klassestørrelse, og mange af studierne er af høj til meget høj kvalitet, idet de anvender metoder på højt evidensniveau, fx kontrollerede forsøg og kvasi-eksperimenter. Meget af evidensen er baseret på amerikanske data, men der er også en del studier fra andre lande, heriblandt også studier, der belyser problematikken med danske data og data fra andre nordiske lande, som dermed er af særlig relevans i en dansk kontekst.

Der er to meget kendte studier i klassestørrelseslitteraturen. Det ene er et stort anlagt eksperiment i USA fra 1985 omkring effekten af reduceret klassestørrelse på elevers faglige resultater (Student-Teacher Achievement Ratio – STAR), hvor både lærere og elever inden for de samme skoler blev fordelt i små (13-17 elever) og store (22-25 elever) klasser. Det andet studie er udført i israelsk kontekst og er baseret på en kvasi-eksperimentel metode, som benytter regler for maksimal klassestørrelse (Angrist & Lavy, 1999). Ideen er, at reglen for maksimal klassestørrelse skaber en variation i klassers størrelse, hvor nogle skoler må oprette flere, men mindre klasser på grund af reglen, end skolen måske ellers ville have gjort. For begge studier fandtes store positive effekter af mindre klasser på elevernes faglige resultater af omtrent samme størrelse (Krueger, 1999; Mosteller, 1995; Angrist & Lavy, 1999).

Der er dog senere blevet stillet spørgsmålstegn ved de fundne positive effekter (Hanushek, 1999; Angrist et al., 2019). Den væsentligste kritik af STAR-eksperimentet er, at den oprindelige lodtrækning ikke holder, fordi elever i løbet af de første år blev flyttet mellem store og små klasser. Og en opfølgende analyse for Angrist & Lavy (1999)-studiet, hvor de ser på en senere tidsperiode med andre skoleforhold (Angrist et al., 2019), finder ingen signifikante effekter af klassestørrelse.

Der er fundet ét metastudie for effekter af klassestørrelse (Filges, Sonne-Schmidt & Nielsen, 2018).¹⁷ Studiet fokuserer snævert på effekten af antal elever i klassen på elevernes læring i grundskolen. Det betyder i praksis, at studier, der anvender den gennemsnitlige klassestørrelse på årgangs- eller skoleniveau – i stedet for det faktiske antal elever i klassen – ikke indgår i studiet. For at indgå i metaanalysen kræves det desuden, at studiet angiver den for en metaanalyse nødvendige information. Kun 10 studier ender med at blive inklu-

¹⁶ Sturm (1997) viser, at henvisningsprocenten er en tredjedel lavere efter indfasningen af klassestørrelsesreduktionen. Det er dog en rent beskrivende sammenhæng og kan ikke tolkes kausalt.

¹⁷ Der pågår i VIVE Campbell et arbejde med en ny forskningsoversigt om effekter af klassestørrelse på specialområdet:
(<https://www.vive.dk/da/nyheder/2020/vive-campbell-undersoeger-betydningen-af-klassestoerrelser-paa-specialomraadet/>)

deret i metaanalysen (4 STAR-studier og 6 andre). Metaanalysen viser, at antallet af elever i en klasse i bedste fald har meget lille effekt på elevernes faglige resultater.

Evidensgrundlaget for Tabel 7.1 er baseret på nyere studier (fra omtrent år 2000 og senere) på højt evidensniveau suppleret med studier fra de nordiske lande. En fuldstændig afdækning af litteraturen på klassestørrelse-området ville være meget omfattende.

Det overordnede billede, der tegner sig, er, at der er blandede resultater. En del studier finder nul-effekter, men mange finder positive effekter af klassestørrelse. Effekterne er ofte små, men nogle studier finder også effekter af betydelig størrelse.¹⁸

Studier med danske data er særligt relevante for en eventuel anvendelse i en dansk økonomisk model. Der findes en række danske studier, som bruger kvasi-eksperimentelle metoder og er dermed på et højt evidensniveau (der er ikke gennemført RCT'er vedrørende klassestørrelse i Danmark). Nogle af dem benytter ligesom Angrist & Lavy (1999) regler for maksimal klassestørrelse (Browning & Heinesen, 2007; Nandrup, 2016), mens de øvrige benytter andre kvasi-eksperimentelle strategier (Heinesen, 2009; Krassel & Heinesen, 2014). På nær Heinesen (2009), som finder en mellemstor positiv effekt af en reduktion af klassestørrelse,¹⁹ finder de øvrige studier alle nul- eller kun små positive effekter. Det overordnede billede af den danske evidens vedrørende effekter af klassestørrelse er derfor ingen eller kun små positive effekter af reduceret klassestørrelse.

Fundene fra evidensundersøgelsen er sammenfattet i Tabel 7.1.

¹⁸ Ved sammenligningen af effektstørrelserne bør man være opmærksom på, at intensiteten af indsatsen – målt ved størrelsen af reduktionen af klassestørrelse – varierer mellem studierne og bør omregnes for at opnå sammenlignelighed. I STAR-eksperimentet (RCT) er reduktionen på 7-8 elever, mens studier, der bruger regler for maksimal klassestørrelse, typisk afreporterer effektstørrelser for en reduktion med én elev.

¹⁹ Resultatet er dog fundet i en undervisningssituation, som måske ikke er så repræsentativ både med hensyn til og den gennemsnitlige klassestørrelse (fransk, 13 elever).

Tabel 7.1 Sammenfatning af evidensviden for reduktion af klassestørrelse

Evidensområde	Evidens	Indgår nordiske?	Noter
Læring	0/+/>++	Ja	Måles først og fremmest som test-scores/karakterer
Omgang	+	Nej	
Ikke-kognitive mål	0/+	Ja	Fx trivsel, selvkontrol, samvittighedsfuldhed, troen på sig selv (self-efficacy); elev-engagement, forstyrrelse, opmærksomhed; litterær interesse)
Helbred (mentalt)	0	Ja	
Afsluttet uddannelse	+	Ja	
Løn/indkomst	+	Ja	
Kriminalitet	0	Nej	

Anm.: Plustegnet indikerer, at en *reduktion* af klassestørrelsen har en positiv effekt på outcomes, fx øget læring, mindre omgang. Parenteser i søjlen 'Evidens' indikerer, at den angivne effektretning er mindre stærkt evidensunderbygget.

Det vurderes, at det mest relevante studie for modeludviklingen er Nandrup (2016). Denne vurdering er foretaget på baggrund af følgende kriterier: 1) Studiet er dansk, 2) der er tale om et studie med et godt metodedesign, og 3) effekterne gælder for en bred elevgruppe. En særlig fordel ved at bruge Nandrup (2016) til modeludviklingen er, at analyserne nemt kan gennemføres og udvides med nye analyser, fx for andre outcomes, i en dansk kontekst.

Nandrup (2016) benytter maksimum klassestørrelse-regler til at estimere effekter af en reduktion i klassestørrelsen. Effekten bliver evalueret ved testresultater i de nationale test i 2., 3., 6. og 8. klasse. Studiet finder blandede effekter af en reduktion i klassestørrelsen. Der findes ingen signifikante effekter på testresultater i 3. og 8. klasse. I 2. klasse findes positive effekter på testresultater i læsning, og i 6. klasse i både læsning og matematik. De signifikante effektstørrelser ligger på omkring 0,010-0,015 SD, svarende til 0,3-0,4 point på en 100 point-skala. Dette vurderes at være meget små effektstørrelser.

7.4 Modeludvikling

Vi gennemgår mulighederne for modeludvikling med basis i de tre evidensniveauer direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Det vurderes på nuværende tidspunkt ikke muligt at gennemføre beregninger for klassestørrelsesreduktion baseret på direkte evidens, da en sådan evidens kun foreligger i meget begrænset omfang.

Indirekte evidens

Studiet af Nandrup (2016) giver mulighed for rent teknisk at udvikle indirekte evidens for den monetære effekt af klassestørrelse, men da estimerne i studiet (og i den øvrige forskning) er af blandet fortegn og generelt med små estimater, må det forventes, at en sådan model vil have meget små afledte økonomiske konsekvenser som output.

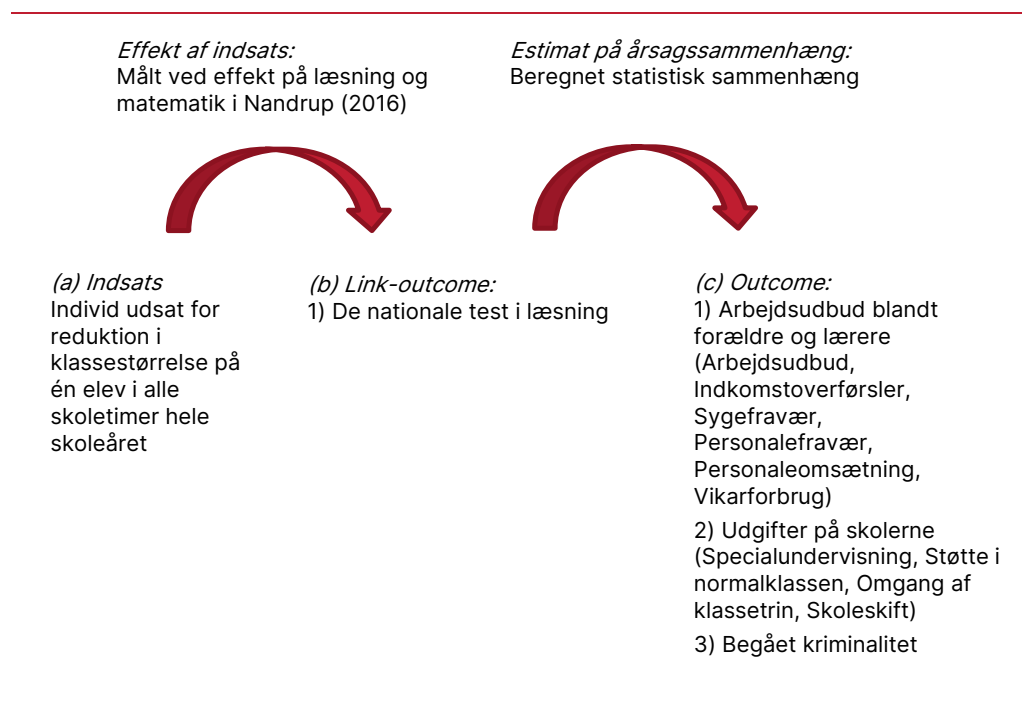
Den anbefalede metode til en modelberegning, hvis det besluttes at gennemføre en beregning, er gennemgået i det efterfølgende.

Figur 7.2 illustrerer modellen ved et indirekte evidens-niveau. Indsatsen er i afsnit 7.1 ovenfor modelteknisk afgrænset til et individ udsat for en reduktion i klassestørrelsen på én elev i alle skoletimer hele skoleåret. Estimerne af effekten på læsning og matematik fra studiet Nandrup (2016) benyttes som evidensgrundlag. Her findes ingen signifikante effekter på 3. og 8. klassetrin, men på 2. og 6. klassetrin. Det anbefales derfor, at modellen afgrænses til mellemtrinnet. Effektstørrelserne i studiet er mellem 0,013-0,015 SD i læsning og matematik. Hvis effekten antages at være lineær, svarer det til en effektstørrelse på omtrent 0,04 SD for en reduktion i klassestørrelsen på 3 elever. Valget af link er bestemt af den tilgængelige evidens, men samtidig vurderes det, at disse link-outcomes kan fange en betydelig del af de effekter, der er beskrevet i forandringsteorien. Det forventes således, at testresultaterne i væsentlig grad er korreleret med forbedring i læring, formindsket fravær og højere trivsel i klassen.

Vi foreslår, at relationen mellem link-outcome og de økonomiske konsekvenser bliver beregnet baseret på statistiske sammenhænge i data. I tilfælde, hvor der forefindes evidens for en relation mellem link-outcome og et konkret økonomisk outcome, vil dette dog med fordel kunne erstatte den statistisk beregnede relation. En beregning af relationen mellem link-outcome og outcome åbner rent teknisk op for muligheden for at inkludere alle outcomes, der er tilgængelige data på (jf. Bilagstabel 2.1). Hvis man har den kvalificerede forandringsteori som udgangspunkt i udvælgelsen af relevante outcomes, vil man kunne inkludere følgende outcomes: 1) Arbejdsudbud blandt forældre og lærere (Arbejdsudbud, Indkomstoverførsler, Sygefravær, Personalefravær, Personaleomsætning, Vikarforbrug), 2) Udgifter på skolerne (Specialundervisning, Støtte i normalklassen, Omgang af klassetrin, Skoleskift) og 3) Begået kriminalitet.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og link-outcomes beregnes. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset og baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret og tilgængeligheden af de nationale tests, der samlet kun dækker folkeskolen med informationer fra 2010 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Figur 7.2 Model for indirekte evidens for klassestørrelse



Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen, dvs. for den population, som beregningen skal dække. Dette baseline-niveau kan enten baseres på en lokal værdi for den kommune, beregningen skal dække, eller på en landsdækkende gennemsnitlig værdi. Det kan ikke forventes, at brugeren ligger inde med disse

parametre. Derfor skal modellen tilbyde opdaterede tal på disse parametre fra alle landets kommuner. Disse tal beregnes på baggrund af danske registerdata som årlige kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerte outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer op det valgte link-outcome, dvs. for standardafvigelsen på testresultater i læsning fra de nationale test for hver kommune samt landsgennemsnittet.

Dette resulterer i en model, hvor brugeren blot behøver at vælge kommune. Brugeren vil derefter blive præsenteret for en række tal, som kan rettes til efter behov.

Beregn effektstørrelser for hvert relevant outcome

Som angivet i afsnit 3.2 skal der for hvert outcome foretages en beregning af den monetære effekt af indsatsen. Dette skal gøres for hver af de 4 år, efter indsatsen blev implementeret. Når alle monetære effekter er regnet ud, skal de lægges sammen som beskrevet i kapitel 3. Den monetære effekt af en indsats på et givet outcome beregnes på følgende måde:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat fra Nandrup (2016), $Base^{ab}$ er det lokale baseline-niveau på testresultatet i læsning og matematik, β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome, og $Base^{bc}$ er baseline-niveauet for det relevante outcome. Da testresultatet er en kontinuert variable, reduceres udtrykket for kontinuerte outcomes til:

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

Og for binære outcomes til:

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right)$$

Beregningen foretages på individniveau for hvert indsats-år. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsats-år ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimaterne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode, er der behov for estimater på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette for klassestørrelse, da indsatsen kun er evalueret umiddelbart, efter indsatsen er udført. Ved brug af link-outcome fra testresultater i læsning kan persistensen (dvs. den vedvarende effekt) estimeres. Populationen er begrænset af tilgængeligheden af de nationale tests til folkeskoleelever fra 2010 og frem. Dette benyttes til at beregne den betingede korrelation i testresultater mellem to efterfølgende test-år. Man har derved opnået et mål for, hvor persistent testresultater er. Dette estimat benyttes til at foretage en lineær ekstrapolation over den 4-årige periode. Man opnår derved et mål for den betingede korrelation mellem testresultater i år 1 og testresultater i de følgende 3 år.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi pr. individ, der modtager den fulde indsats som anført i afsnit 3.2. Dernæst aggregeres estimatet pr. individ op med dette antal individer.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistisk sammenhæng, er det nødvendigt at kunne knytte viden om, hvilke børn der har været omfattet af en indsats til de efterfølgende realiserede outcomes.

For klassestørrelse er dette muligt ved brug af DST's elevregistre, der har observationer fra 2008 og frem. Der foretages for hvert inkluderet outcome en regression baseret på data fra DST, hvor den betingede korrelation mellem outcome, og hvorvidt individet går i en mindre klasse, bliver beregnet, fx den betingede korrelation mellem brug af specialundervisning og indsatsen. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST, eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset og baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregistret, dvs. fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

7.5 Ny evidens og ny data

Det vurderes på nuværende tidspunkt ikke at være meningsfuldt at gennemføre beregninger for en klassestørrelsesreduktion på direkte eller indirekte evidensniveau, da forskningsmæssig evidens på skoleområdet peger på små effekter med hensyn til testscorer.

Hvis man ønsker en model baseret på direkte evidens, kan man med fordel tage udgangspunkt i den empiriske strategi i Nandrup (2016), men benytte de relevante outcomes som afhængige variable i stedet for testresultater. Fordi studiet kun finder små effekter på testscorer, er der dog også rimelig sandsynlighed for, at der findes små effekter på outcomes, der kan monetariseres.

Endvidere har VIVE i januar 2021 udarbejdet et Notat om Forsøg med Alternativ Ressourceanvendelse for BUVM (Beuchert, Dietrichson & Jensen, 2021), som skitserer et lodtrækningsforsøg (RCT) for at evaluere effekten af en ned-sættelse af klassestørrelse. Projektet vil kunne levere RCT-estimer af en klassestørrelsesreduktion for Danmark, og resultaterne herfra vil med fordel kunne bruges til modelberegninger.

Det vurderes, at der p.t. ikke er brug for yderligere data vedrørende klassestørrelse.

8 Ordblindeindsatser

8.1 Indsatsens indhold og afgrænsning

Indsatser for ordblinde kan – som det også skete i analysen i kapitel 2 – inddeles i to kvalitativt forskellige indsatsstyper:

- En afgrænset, oftest lokal indsats over for en gruppe af ordblinde børn; typisk en mindre gruppe på en enkelt skole. Indsatsen vil oftest have en start- og en slutdato og vil fx kunne bestå i, at der afprøves en ny metode til undervisning af ordblinde børn.
- En strukturel ændring af ordblindetilbuddet i kommunen eller i et andet større område. Dette vil påvirke alle ordblinde børn, der går i skole i det pågældende område, og kan fx være, hvis man vælger at omlægge undervisningen til et centralt tilbud i kommunen, eller hvis man på anden måde ændrer rammevilkårene for ordblindeindsatsen i kommunen.

Selvom der således er en indholdsmæssig forskel på indsatsen, alt efter om den er lokal eller strukturel, så vil det i modelmæssig sammenhæng være de samme relevante mål, der skal anvendes. Derfor vil vi foreslå, at ordblindeindsatser alene afgrænses ved at optælle antallet af børn, der har været omfattet af indsatsen. Såfremt fx en indsats indføres gradvist, så vil dette kunne håndteres ved beregningsteknisk at antage, at eksempelvis halvdelen af eleverne modtager indsatsen første år, men det fulde antal elever modtager indsatsen andet år. På samme måde kan der tages hånd om, hvis elever på forskellige årgange modtager indsatsen.

Modelberegninger foretages på individniveau, jf. afsnit 3.2.9. Den monetære effekt pr. individ-år skal efterfølgende aggregeres i forhold til optællingen af den lokale indsats.

8.2 Mulige mekanismer for indsatsens virkning

Figur 8.1 og Figur 8.2 viser kvalificerede forandringsteorier for virkningskæderne for ordblindeindsatser af henholdsvis den lokale og den strukturelle type. Forandringsteorierne viser en teoretisk modellering af de mulige mekanismer mellem indsatsen og den økonomiske forandring på disse områder, jf. kapitel 2. Mekanismemodellerne siger ikke i sig selv noget om, hvordan de

økonomiske beregninger kan foretages, men illustrerer blot, hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

Figurerne illustrerer, at vi har fundet økonomiske outcomes på tre områder for ordblindeindsatserne – og illustrerer samtidig, at selvom virkningskæderne er lidt forskellige i de to figurer, så er det de samme. I begge tilfælde er de økonomiske konsekvenser, der står tilbage i forandringsteorierne efter kvalificering, forældrenes arbejdsudbud, udgifter på skolerne og mindre kriminalitet, hvilket også understøtter at betragte ordblindeindsatser samlet i en modelmæssig kontekst. Det er også muligt at inddrage overgang til ungdomsuddannelser for indsatser rettet mod elever i udkolingen.

Effekten på forældres arbejdsudbud opstår i begge figurer gennem en virkningskæde via elevernes forventede forbedrede trivsel, der igen fører til mindre skolefravær og derigennem påvirker muligheden for, at forældrene kan være på arbejde. I forandringsteorien for lokale indsatser indgår endvidere, at tidligere opsporing af problemer med ordblindhed kan føre til mindre fravær, som kan føre til forøget arbejdsudbud.

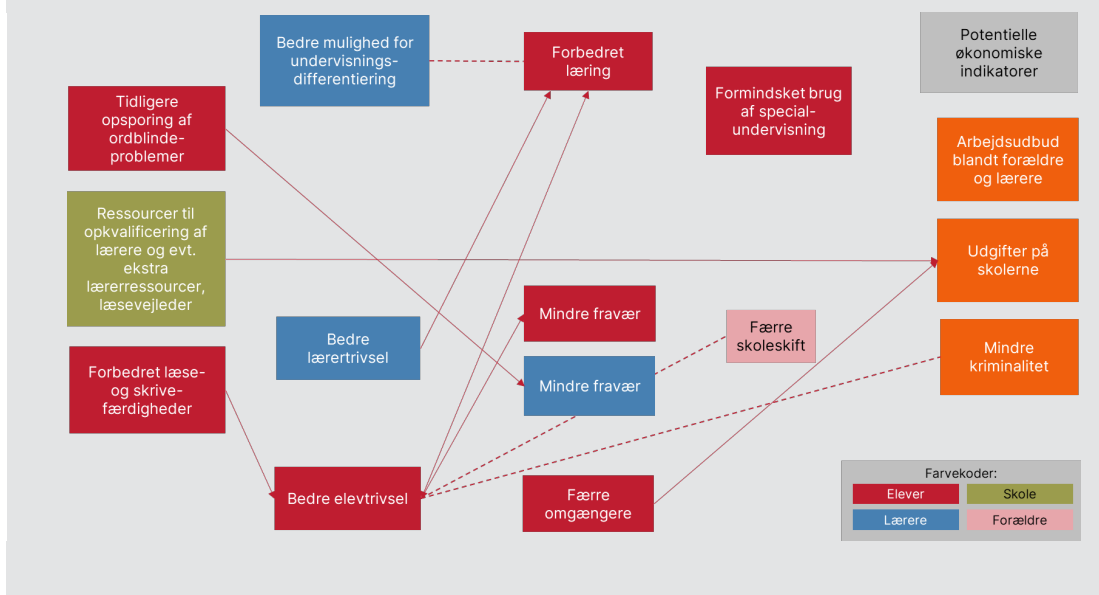
Virkningskæderne, der medfører lavere udgifter på skolerne, går igennem henholdsvis forbedrede læse- og skrivefærdigheder, forbedret læring og formindsket brug af specialundervisning for en kæde. En anden virkningskæde går for de strukturelle tiltags vedkommende igennem fortsat deltagelse i fagligt fællesskab, bedre elevtrivsel og færre omgængere og skoleskift.

Et tredje område, hvor der ifølge begge de kvalificerede forandringsteorier er mulighed for en økonomisk konsekvens, er, at der kan blive begået mindre kriminalitet som følge af elevernes mulige forbedrede trivsel, om end dette outcome ikke anses for at være centralt for ordblindeindsatser.

Ud over de outcomes, der indgår i de kvalificerede forandringsteorier, vil det også være relevant at se på fx overgang til ungdomsuddannelse som et outcome for målgruppen.

I afsnit 8.4 anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.

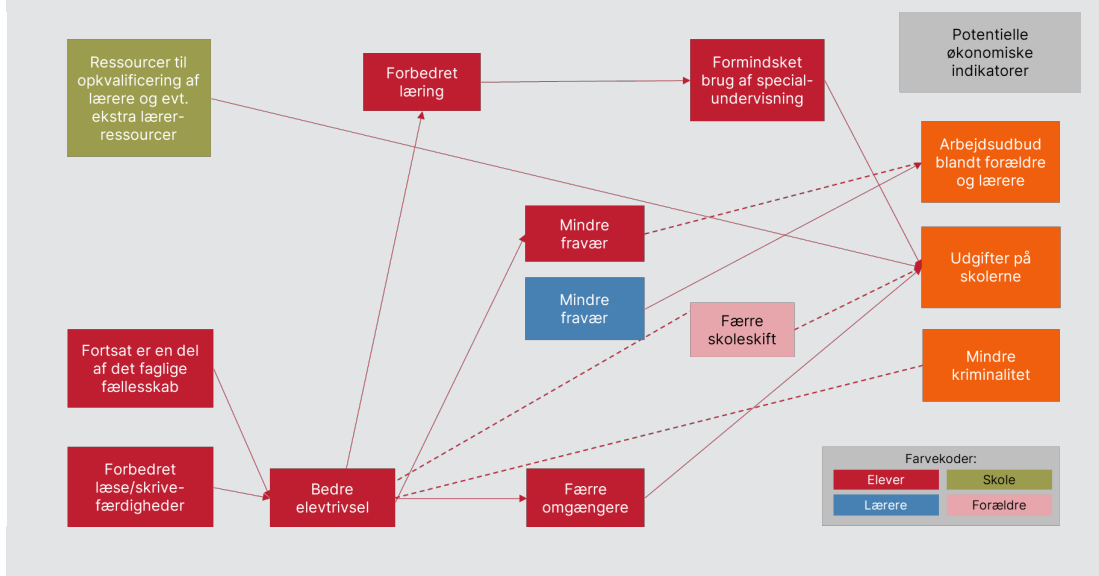
Figur 8.1 Kvalificeret forandringsteori for sammenhængen mellem ordblindeindsatser (lokale) og efterfølgende effekter



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens der er angivet i feltet.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

Figur 8.2 Teoretiske sammenhænge mellem ordblindeindsatser (strukturelle) og efterfølgende effekter



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens der er angivet i feltet.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

8.3 Evidens for indsatsen

Dette afsnit ser på den kvantitative evidens for virkningen ordblindeindsatser, der er fundet efter en litteratursøgning.

I denne gennemgang medtager vi indsatser og metaanalyser, som inkluderer børn og unge, der i studierne defineres som enten ordblinde eller som børn og unge med læse- eller stavehandicap²⁰, uanset hvilken definition studierne anvender²¹. Målgrupper, som deler fællestræk med børn og unge med ordblindhed, fx børn i *risiko* for ordblindhed (*at risk of dyslexia*), unge med generelle indlæringsvanskeligheder (*learning disabilities*) eller unge med faglige vanskeligheder (*academic difficulties*), inkluderes ikke i denne kortlægning. Se fx Dietrichson et al. (2020) for et systematisk review af indsatser for unge med – eller i risiko for – faglige vanskeligheder.

Ordblindeindsatserne anvender forskellige tilgange til at opnå en positiv effekt, og ofte afrapporterer metaanalyserne effektstørrelser særskilt for hver type af indsats. Det gør de, fordi indsatserne – og deres effektstørrelser – kan være ret forskellige. På den ene side komplicerer det evidensbilledet, men det giver på den anden side mulighed for at afrapportere, hvilke typer af indsatser der kan dokumentere de største og sikreste effekter, og hvilke indsatser der ikke lader til at være effektfulde. Figur 8.2 giver et overblik over mest udbredte indsatstyper på området.

På ordblindeområdet findes flere metaanalyser, der undersøger effekter af indsatser for børn og unge med ordblindhed (Galuschka et al., 2020; Galuschka et al., 2014; SBU, 2014). Metaanalyserne medtager hovedsageligt studier med randomiserede kontrollerede forsøg, hvor ordblinde enten tildeles en specifik indsats eller indgår i kontrolgruppen. Studierne anvender dermed metoder på højt evidensniveau og har generelt høj kvalitet. Dog er det ikke alle studier i metaanalyserne, som anvender fuld randomisering i tildelingen af enten indsats eller kontrolgruppe, og ikke alle studier gennemfører en undersøgelse af *treatment fidelity*, som er de strategier, der undersøger, overvåger og sikrer, at indsatsen implementeres helt efter planen. Den nyeste metaanalyse

²⁰ På engelsk betegnes ordblindhed *dyslexia*, og læse- og stavehandicap betegnes henholdsvis *reading disability* og *spelling disability*.

²¹ Ordblindhed er vanskeligt at definere og vanskeligt at afgrænse entydigt. Overordnet set kan ordblindhed defineres som en dysfunktion i visse sproglige funktioner – især sprogets lydæssige form – som er vigtige for at kunne udnytte skriftens lydprincip. Ordblindhed er et spørgsmål om grænse-dragning, og i gennemsnit vil 1-2 elever i hver skoleklasse kunne defineres som ordblinde (SBU, 2014). Men hos de børn og unge, der defineres som ordblinde, eksisterer mange individuelle variationer i forhold til graden af læse- og stavevanskeligheder. På grund af vanskelighederne ved at afgrænse ordblindhed anvender studierne ikke helt samme afgrænsning. De fleste studiers inklusionskriterie er børn og unge, der afviger med en bestemt grænse (fx 1 eller 1,5 standardafvigelse) fra gennemsnittet på en læsetest, samtidig med at den unge scorer inden for normalområdet på kognitive test.

af Galuschka et al. konkluderer dog, at disse forhold ikke influerer på de overordnede resultater af analyserne (Galuschka et al., 2020).

Evidensen er både baseret på studier fra amerikanske, engelske og nordiske lande, og evidensen herfra kan meningsfuldt overføres til en dansk kontekst. Det skyldes, at de skandinaviske sprog og det engelske sprog på mange måder er sammenlignelige, fordi de alle er germanske sprog og har en såkaldt dyb ortografi.²² Herudover skyldes det, at kontrolgruppen i indsatserne i langt de fleste tilfælde modtager almindelig undervisning som vanligt. Effekterne måles derfor op imod den fremgang, ordblinde ville opnå, hvis de fortsat blev tildelt almindelig undervisning.

Langt de fleste studier kredser om ordblindeindsatsernes effekt på læring – og her især ordblindes læse- eller stavekompetencer. Det er derfor oftest læse- og stavefærdigheder, der udgør evidensområdet for indsatserne. Selvom disse primære effekter ikke er tilstrækkelige i en måling af økonomiske konsekvenser, er de højst relevante at få kortlagt, netop fordi det er den primære effekt på læse- og staveudfald, som kan afføde eventuelle afledte effekter.

Der findes generelt få eksempler på undersøgelser, der evaluerer effekter af indsatser på andre udfald end de rent læringsbaserede og kan betegnes relevante økonomiske konsekvenser. I dansk kontekst findes imidlertid indsatsen *Plan T*, et intensivt læringsforløb for unge med ordblindhed i Odense Kommune, som evalueres på udfald som fravær, trivsel, skoleparathedsvurdering etc. (Rosholm, Nielsen & Hvidman, 2020). Internationalt betragtet har få dog forsøgt at udvikle og evaluere trivselsindsatser for børn og unge med ordblindhed (Boyes et al., 2021), selvom børn og unge med ordblindhed kan have større risiko for at have emotionelle og adfærdsrelaterede udfordringer. En ny pilotindsats – *Clever Kids* – viser dog lovende resultater, men der er behov for en indsats i større skala for at kunne sikre reliabiliteten af resultaterne fra denne indsats (Boyes et al., 2021). Der savnes således evidens for ordblindeindsatser effekt på trivsel, og der savnes evidens for indsatsernes afledte effekter.

Tabel 8.1 sammenfatter evidensen på ordblindeområdet. Helt overordnet finder metaanalyser både positive effekter og nul-effekter af ordblindeindsatser målt på både læse- og staveudfaldsmål. De varierende effektstørrelser skyldes dels, at nogle indsatser dokumenterer positive såvel som nul-effekter, men især, at det ikke er alle ordblindeindsatser, som alle metaanalyser kan dokumentere sikre effekter for. Litteraturen finder de største effekter af indsatser, der involverer én-til-én-undervisning og lavere effektstørrelser for indsatser gennemført i hele klassen.

²² Dyb ortografi betyder, at der er forholdsvis stor afstand mellem ordlyde og de bogstaver, som lydene repræsenterer (Elbro, 2006: 31). Derfor kan evidensen fra engelsktalende lande meningsfuldt overføres til dansk kontekst.

Generelt finder alle metaanalyser de sikreste effekter for af indsatser af fonem-grafem-typen (dvs. vedrørende lyde og bogstaver), og det er da også denne indsatstype, der er mest undersøgt i litteraturen. Det systematiske review af det svenske SBU fra 2014 konkluderer, at man kun kan udtale sig sikkert om effekten af denne indsatstype, fordi evidensgrundlaget for andre indsatstyper er for spinkelt (SBU, 2014).

Metaanalysen af Galuschka et al. konkluderer imidlertid, at flere ordblindeindsatser er effektfulde, når vi ser på læse- og staveudfaldsmål (Galuschka et al. 2020). Foruden fonem-grafem-indsatser gælder det indsatser, der træner fonologisk opmærksomhed, morfologiske (ordets opbygning) indsatser og ortografiske (stavning) indsatser. Der kan kun dokumenteres positive effekter for morfologiske indsatser på staveudfaldsmål, og litteraturen finder forskellige effektstørrelser for denne indsats. De ortografiske indsatser evalueres kun for staveudfald, hvor der findes middelstore effekter.

Der kan ikke dokumenteres nogen effekter af hukommelsesindsatser eller understøttende software, og generelt er der få undersøgelser af effekten disse indsatser med i metaanalyserne. Studier finder dog, at almindelig individuel undervisning ikke kan erstattes af computerbaserede tilgange. Dette konkluderes også af en metaanalyse, der undersøger effekten af understøttende software for alle elever og ikke blot ordblinde (Williams & Beam, 2019). I en dansk undersøgelse har man undersøgt effekten af it-støtte på ordblinde elevers skriftsproglige udvikling (Arnbak & Petersen, 2017). Undersøgelsen viser, at de ordblinde elever har gavn af redskaberne, når de skal løse læse- og skriveopgaver, men hovedparten anvendte ikke redskaberne i hverdagen. Derfor har undersøgelsen ikke mulighed for at undersøge, hvilken effekt brug af it har på udviklingen af ordblindes skriftsproglige færdigheder.

Tabel 8.1 Sammenfatning af evidensviden for ordblindeindsatser

Evidensområde	Evidens	Indgår nordiske studier?	Noter
Læsning	0/+	Ja	Måles først og fremmest som testscores (standardiserede eller ikke-standardiserede tests)
Stavning	0/+	Ja	Måles først og fremmest som testscores (standardiserede eller ikke-standardiserede tests)
Ikke-kognitive mål	0	Ja	For eksempel fravær, trivsel, skoleparathedsvurdering

8.4 Modeludvikling

Vi gennemgår mulighederne for modeludvikling med basis i de tre evidensniveauer direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Det vurderes på nuværende tidspunkt ikke muligt at gennemføre beregninger for ordblindeindsatser, da en sådan evidens ikke foreligger.

Indirekte evidens

Det vurderes på nuværende tidspunkt ikke at være meningsfuldt at gennemføre beregninger for strukturelle indsatser, da det ikke har været muligt at finde evidensunderbygning for denne type indsats.

Til gengæld viser evidensundersøgelsen, at der er en del positive effektmålinger for konkrete lokale indsattyper, jf. ovenfor, som er blevet undersøgt i litteraturen, så vi anbefaler at tage udgangspunkt i disse til en modelberegning for ordblindeindsatser.

Figur 8.3 illustrerer modellen ved et indirekte evidensniveau for lokale ordblindeindsatser. Indsatsen er i afsnit 8.1 ovenfor modelteknisk afgrænset til en optælling af de individer, der deltager i en konkret lokal indsats som fx afprøvelse af en ny undervisningsmetode. Da der i den fundne litteratur er afprøvet adskillige indsatsformer, er det svært at udvælge en enkelt relevant effektstørrelse til beregningen. Hvis man ser på den samlede mængde effektstørrelser for effekt på læsning, tyder det på en effektstørrelse i størrelsesorden 0,35-0,4 SD. Vi vil dog i stedet i dette tilfælde anbefale, at beregningen indrettes sådan, at der kan anvendes forskellige effektstørrelser fra de forskellige indsattyper, der findes i evidensundersøgelsen.

Som link-outcomes til beregningen bør resultater fra de nationale test i læsning eller eventuelt resultatet af afgangsprøve i 9. klasse i dansk anvendes. Valget af link er i bestemt af den tilgængelige evidens, der i høj grad måler på læseevne. Samtidig må man dog også forvente, at disse link-outcomes kan opfange en væsentlig del af effekt baseret på forandringsteorien. Det forventes således, at testresultatet i væsentlig grad er korreleret med forbedring i læring, formindsket fravær og højere trivsel.

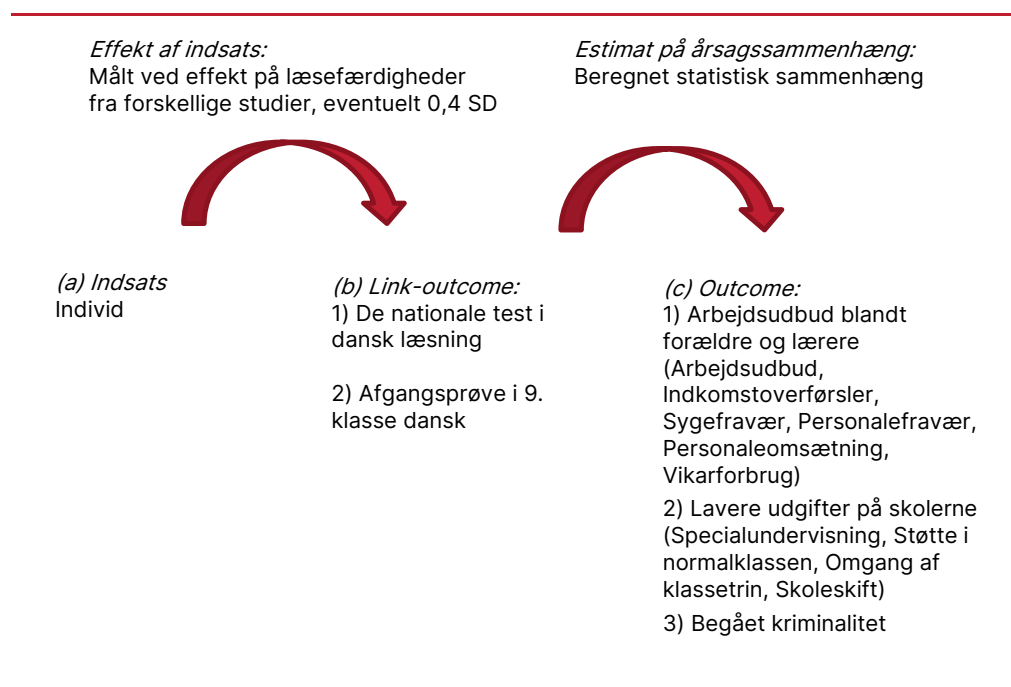
Det foreslås, at relationen mellem link-outcome og de økonomiske konsekvenser bliver beregnet baseret på statistiske sammenhænge i data. Såfremt der

kan findes evidens for en relation mellem link-outcome og et konkret økonomisk outcome, vil dette dog med fordel kunne erstatte den statistisk beregnede relation. En beregning af relationen mellem link-outcome og outcome åbner rent teknisk op for muligheden for at inkludere alle outcomes, der er tilgængelige data på (jf. Bilagstabel 2.1). Hvis man benytter den kvalificerede forandringsteori som udgangspunkt i udvælgelsen af relevante outcomes, vil man kunne inkludere følgende outcomes: 1) Arbejdsudbud blandt forældre og lærere (Arbejdsudbud, Indkomstoverførsler, Sygefravær, Personalefravær, Personaleomsætning, Vikarforbrug), 2) Lavere udgifter på skolerne (Specialundervisning, Støtte i normalklassen, Omgang af klassetrin, Skoleskift) og 3) Begået kriminalitet.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og link-outcomes beregnes, fx den betingede korrelation mellem brug af specialundervisning og resultat fra de nationale tests. I regressionerne betinges på standard socio-demografiske baggrundsvARIABLE fra DST som eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. I dette tilfælde, hvor der er tale om en indsats, der kun tilbydes en delpopulation af alle skoleelever, bør populationen til beregningen også afspejle dette. Dette kan fx gøres ved kun at inkludere de elever, der scorer lavt i de nationale test i læsning.

Hvis man vælger at inkludere specialundervisning som et outcome, vil analysen være afgrænset af indholdet i specialundervisningsregisteret og tilgængeligheden af data fra de nationale tests, der samlet kun dækker folkeskolen med informationer fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Figur 8.3 Model for indirekte evidens for ordblindeindsatser (lokale)



Kilde: Egen tilvirkning.

Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen. Dette baseline-niveau kan enten baseres på en lokal værdi for den kommune, beregningen skal dække, eller på en landsdækkende gennemsnitlig værdi. Såfremt der forventeligt er meget forskellige baseline-niveauer for fx forskellige aldersklasser – og såfremt der ønskes en model, der kan foretage beregninger opdelt på alder – skal der være alderopdelte baseline-niveauer til rådighed. Det kan ikke forventes, at brugeren ligger inde med disse parametre. Derfor skal modellen tilbyde opdaterede tal fra alle landets kommuner på disse parametre. Disse tal beregnes på baggrund danske registerdata som årlige, kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerte outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer for det valgte link-outcome, dvs. standardafvigelsen på testresultater fra de nationale test for hver kommune samt landsgennemsnittet.

Brugeren har derfor blot behov for at vælge kommune og vil derefter blive præsenteret for en række tal, som kan rettes til efter behov.

Beregn effektstørrelser for hvert relevant outcome

Som angivet i kapitel 3 består estimatet af den monetære effekt på et outcome i et givet år efter indsatsen af følgende beregning:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat på testresultater for henholdsvis skole (Kristensen & Skov, 2019) og dagtilbud (Jensen et al., 2009 og Bleses et al., 2020), $Base^{ab}$ er det lokale baseliniveau på testresultatet i læsning, β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome, og $Base^{bc}$ er baseliniveauet for det relevante outcome. Da der benyttes test- eller eksamensresultat som link-outcome, og da testresultater er en kontinuert variable, reduceres udtrykket for kontinuerte outcomes til

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

Og for binære outcomes til

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right)$$

Beregning foretages på individniveau for hvert indsatsår. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsatsår ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimaterne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode er der behov for estimater på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette, da indsatserne kun er evalueret umiddelbart efter, at de er udført. Ved brug af link-outcome fra testresultater, kan persistensen estimeres. Populationen er igen begrænset af tilgængeligheden af de national tests i folkeskoleelever fra 2010 og frem. Dette benyttes til at beregne den betingede korrelation i resultater mellem to efterfølgende testår. Dette estimat benyttes til at foretage en lineær ekstrapolation over den 4-årige periode. Man opnår derved et mål for den betingede korrelation mellem testresultater i år 1 og testresultater i de følgende 3 år.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi per individ, der har deltaget i en indsats som anført i kapitel 3. Dernæst aggregeres estimatet per individ op med dette antal individer.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistisk sammenhæng, er det nødvendigt at kunne knytte viden om, hvilke børn der har været omfattet af en indsats til de efterfølgende realiserede outcomes.

Dette er for øjeblikket ikke muligt, da der ikke findes individdata for, hvilke børn der er ordblinde, og som derfor ville være i målgruppen for en indsats.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og hvorvidt individet har modtaget undervisning af en lærer med relevante kompetencer beregnes. F.eks. den betingede korrelation mellem brug af specialundervisning og indsatsen. I regressionerne betinges på standard sociodemografiske baggrundsvariable fra DST. Eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet, og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome vil populationen være afgrænset af specialundervisningsregistret, dvs. fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen anbefales det at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

8.5 Ny evidens og nyt data

Som det fremgik af evidensundersøgelsen er der generelt ikke målt på økonomiske outcomes i undersøgelsen af effekten af ordblindeindsatser. Det vil derfor være ønskeligt, hvis sådanne målinger kunne gennemføres, så et bedre evidensniveau tilvejebringes. Dette kunne fx gøres ved en opfølgning af analysen gennemført af Rosholm, Nielsen & Hvidman (2020).

Data for ordblinde i Danmark kan delvist skaffes gennem adgang til Notas²³ databaser. Imidlertid er der ikke tale om en systematisk opgørelse af ordblinde, hvilket ellers kunne være ønskeligt til analyse- og forskningsformål. Som nævnt ovenfor kan populationen approksimeres ved at se på de børn, der

²³ Nota er Kulturministeriets bibliotekstilbud til mennesker med læsehandicap.

scorer lavt på læsetest, men der er alene tale om en approksimation, da der ikke er viden om baggrunden for den lave score.

Endelig er der data for screeninger for ordblindhed, der foretages fra 3. klasse i grundskolen, som vil kunne anvendes fremadrettet, når der kommer flere år-gange til rådighed.

9 Betydningen af uddannet personale i skoler og dagtilbud

9.1 Indsatsens indhold og afgrænsning

På skoleområdet vil en indsats indebære, at en lærer uden relevante kompetencer udskiftes med en lærer med de relevante kompetencer. Relevante kompetencer er i dette kapitel i første omgang tænkt bredt og omfattende både *generel* undervisningskompetence (dvs. at have gennemført en uddannelse til lærer) og undervisningskompetence *i faget*. Fsva. skoleområdet indsnævres det i kapitel 9.3 til alene at fokusere på sidstnævnte.

Skaleringsenheden vil være, at barnet modtager undervisning af en lærer med de relevante faglige kompetencer et helt skoleår. Afhængig af den konkrete specifikation kan man vælge at specificere "relevante kompetencer" forskelligt. Det kan naturligt afgrænses enten ved, at der blot kræves, at læreren skal have en formel læreruddannelse eller tilsvarende kompetencer, eller ved at kræve, at læreren skal have formel undervisningskompetence i de enkelte timer.

På dagtilbudsområdet vil den mest relevante indsatsdefinition være at en ansat, der har kontakt med børn i det daglige, men ikke har relevante formelle kompetencer, udskiftes med en person med de relevante kompetencer. Det har dog i evidensafsøgningen ikke vist sig mulighed for at foretage en sådan sammenligning, jf. nedenfor. Denne indsatsdefinition vil derfor kun kunne implementeres ved et lavere evidensniveau og i det omfang at data på kompetencer hos ansatte i dagtilbud kan fremskaffes.

En alternativ indsatsdefinition på dagtilbudsområdet er, at alle medarbejdere i et dagtilbud modtager et kortere opkvalificeringsforløb. Den modeltekniske indsatsdefinition vil derved være at barnet modtager et dagtilbud med ændret indhold. Skaleringsenheden kan her være at barnet modtager denne indsats i et helt år.

Modelberegninger foretages på individniveau. Den monetære effekt per individ-år skal efterfølgende aggregeres i forhold til den omfanget af den lokale indsats, dvs. antallet af model-børn (individ-år) indsatsen berør. I tilfældet med et kortere opkvalificeringsforløb for alle ansatte i et dagtilbud vil model-

børnene være alle børn i dagtilbuddet i de år, de har tilbage i dagtilbuddet eller op til 4 år. Effekten aggregeres lineært op på baggrund af antallet af model-børn.

9.2 Mulige mekanismer for indsatsens virkning

Figur 9.1 og Figur 9.2 viser kvalificerede forandringsteorier for hhv. uddannet personale i skoler og i dagtilbud. Forandringsteorierne viser en teoretisk modellering af de mulige mekanismer mellem indsatsen og den økonomiske forandring på de områder, jf. kapitel 2. Mekanismemodellerne siger ikke i sig selv noget om, hvordan de økonomiske beregninger kan foretages, men illustrerer blot hvilke mekanismer eventuelle økonomiske konsekvenser kan opstå igennem. Ved alene at betragte den kvalificerede forandringsteori sikrer vi, at der er tale om mekanismer, der enten er påvist direkte som en effekt, eller hvor kvalitative studier eller udsagn har sandsynliggjort sammenhængen.

Figurerne illustrerer, at vi har fundet økonomiske outcomes på tre områder for hhv. dagtilbud og skole, hvor der også er fundet en underbygning af mekanismerne.

På skoleområdet kan der være afledte lavere udgifter på skolerne. I den kvalificerede forandringsteori kommer denne effekt som følge af hhv. elevernes mindre brug af specialundervisning, færre omgængere og færre skoleskift. I den teoretiske model er der hhv. elevernes trivsel og læring, der indvirker på disse størrelser.

Den kvalificerede forandringsteori på skoleområdet viser også en mulig effekt på forældre og læreres arbejdsudbud, der går igennem elever og lærers øgede trivsel og deraf lavere fravær fra skolen.

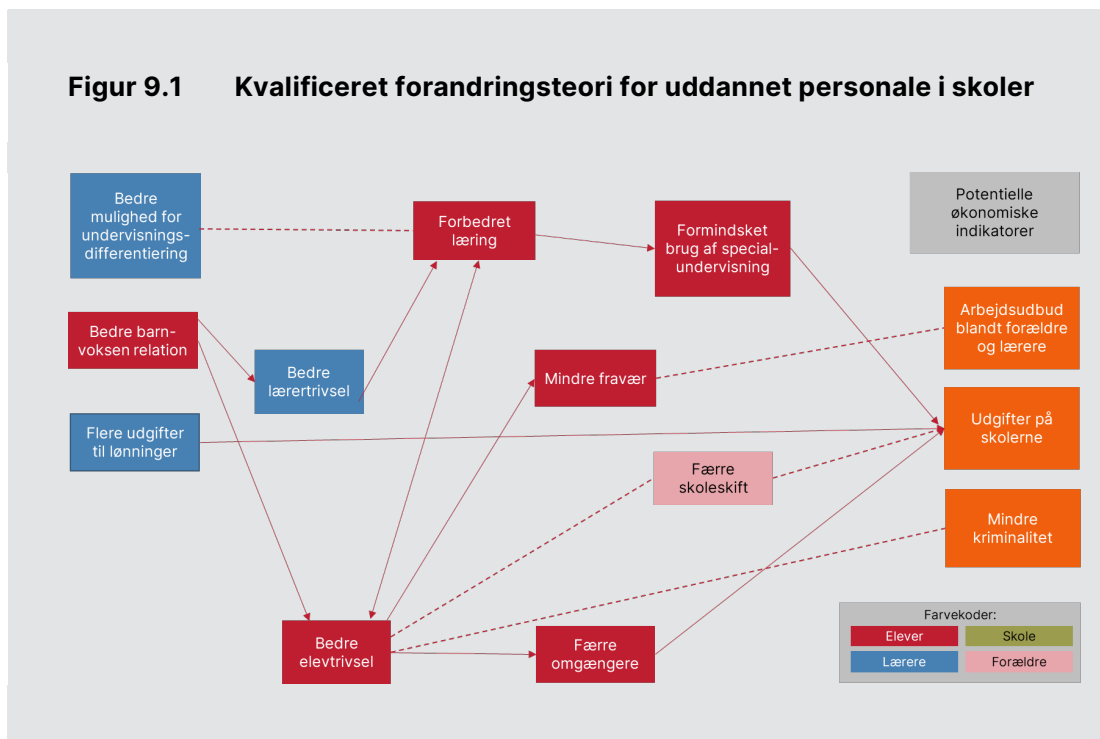
Et tredje område, hvor der inden for skoleområdet er en mulig økonomisk effekt, er, at der kan blive begået mindre kriminalitet som følge af elevernes mulige forbedrede trivsel.

På dagtilbudsområdet kan der være afledte lavere udgifter på dagtilbuddet. I den kvalificerede forandringsteori kommer denne effekt som følge af forbedret trivsel blandt de ansatte og heraf lavere sygefravær og mindre medarbejderomsætning. Et lavere sygefravær vil desuden have en selvstændig betydning i form af øget arbejdsindsats blandt de ansatte.

Hertil kommer at børns trivsel, relationer og læring og voksen-barn relationerne forventes at kunne blive påvirket positivt. Dette kan have en indvirkning

på barnets skoleparathed og på overgangen til skolestart. Det kan samlet set forventes at påvirke barnet adfærd i skolen i bred forstand (se Figur 9.1).

I afsnit 9.4 nedenfor anvender vi inputtet fra den kvalificerede forandringsteori som et muligt forslag til de områder, der kan medtages i en økonomisk beregning, hvis der vælges en evidensunderbygget model.

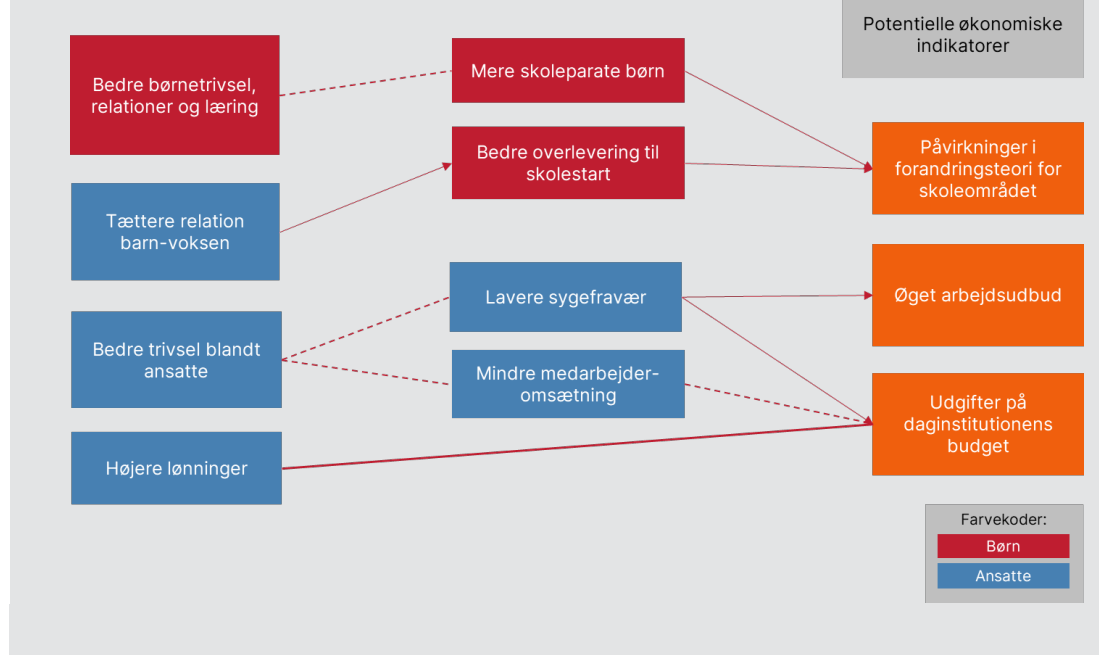


Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet.

Note: Fuldt optrukne linjer angiver stærke sammenhænge, mens stiplede linjer angiver svage sammenhænge, jf. kapitel 2.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

Figur 9.2 Kvalificeret forandringsteori for uddannet personale i dagtilbud



Anm.: Farvekoderne i figuren angiver, hvem der oplever den konsekvens, der er angivet i feltet.

Note: Fuldt optrukne linjer angiver stærke sammenhænge, mens stiplede linjer angiver svage sammenhænge, jf. kapitel 2.

Kilde: Egen tilvirkning på baggrund af workshops og efterfølgende kvalificering, jf. kapitel 2.

9.3 Evidens for indsatsen

Dette afsnit ser på den kvantitative evidens for virkningen af uddannet personale i skole og dagtilbud, der er fundet efter en litteratursøgning.

For denne indsats er det Børne- og Undervisningsministeriet, der har været ansvarlige for indsamlingen af evidens. Det nedenstående er baseret på det materiale, ministeriet har leveret til VIVE, og der er ikke foretaget en vurdering af, om dette materiale er dækkende, idet denne vurdering forudsættes foretaget af BUVM.

For dette indsatsområde undersøges i litteraturen mange forskellige lærerkvalifikationer, som sammenlignes med en eller flere andre typer af lærerkvalifikationer. For eksempel i Darling-Hammond et al. (2005) sammenlignes Teach For America (TFA) certificering med alle andre slags lærere, hvoraf nogle er certificerede, andre ikke. Man kan derfor for dette indsatsområde ikke nøjes

med blot at se på signifikansværdi og effekttestimatets fortegn for at danne sig et generelt billede af, i hvilken retning evidensen peger. Det er derimod nødvendigt, at se på selve indsætterne i hvert enkelt studie og - ligeså vigtigt - se på, hvad indsatsen sammenlignes med.

Metoderne, som er anvendt gentagne gange i den fundne evidens, er simple OLS eller fixed effects modeller, samt randomiserede kontrollerede forsøg.²⁴ Der er fundet evidens for en række områder, herunder ikke-kognitive mål og mental helbred, men hovedparten er baseret på læring som resultatmål.

Generelt viser litteraturen et blandet billede for indsatsen *uddannet personale*. En betydelig del af studierne finder nuleffekter. Ellers peger resultaterne i begge retninger, dog mest i positiv retning.

Tabel 9.1 viser en sammenfatning af den fundne evidens.

Tabel 9.1 Sammenfatning af evidensviden for uddannet personale i skoler og dagtilbud

Evidensområde	Evidens	Indgår nordiske?	Noter
Læring	(-)/0/+	Ja	Fx sproglig udvikling, kommunikation, ordforråd; matematisk forståelse, læringsparathed, kende bogstaver, tal, farver
Ikke-kognitive mål	0/+	Ja	Fx koncentration, samarbejde, selvregulering, sociabilitet
Helbred (mental)	0/+	Ja	Fx emotionelle problemer, adfærdsproblemer, pro-social adfærd, hyperaktivitet (SDQ-scorer)
Undervisningskvalitet	0/(+)	Nej	Fx undervisning & interaktion, emotionel klima, læringsklima, fysiske forhold & udstyr
Skoleskift	0	Nej	Stadig i off. skole (i samme by), skiftet skole; baseret på et enkelt studie
Indkomst	+	Nej	Livsindkomst, baseret på et enkelt studie
Organisering	0/+	Nej	Emotionel støtte, klasseledelse, læringsstøtte; baseret på et enkelt studie

Anm.: Hvor parenteser betyder meget små effektstørrelser.

Der er fundet evidens fra de nordiske lande for både læring, ikke-kognitive mål og mental helbred. Det drejer sig om seks danske studier (fem på dagtilbudsområdet og et på skoleområdet), heraf tre udført med lodtrækningsforsøg, og dermed på højeste kvalitetsniveau med hensyn til evidens. Ligesom

²⁴ Andre metoder der anvendes inkluderer ANCOVA/variansanalyse, hierarkiske lineære modeller, metastudie, ikke-parametrisk vækstmodel, difference-in-differences, samt regression discontinuity design med instrumentvariable.

for den internationale litteratur, er der også for de danske studier blandede resultater, hvor der findes både en del nuleffekter, men også mange mellemstore til store positive effekter.

Det vurderes at de mest relevante studier for modeludviklingen er Kristensen & Skov (2019) på skoleområdet og Jensen et al. (2009), Jensen et al. (2017) og Bleses et al. (2020) på dagtilbudsområdet. Denne vurdering er foretaget på baggrund af følgende kriterier: 1) studierne er alle danske og 2) der er tale om studier med gode metodedesigns.

Kristensen & Skov (2019) sammenligner lærere med kompetencer i det fag de underviser med lærer uden de relevante faglige kompetencer. Metoden er at benytte variation i kompetencedækning oplevet af den samme elev, men på tværs af fag. Den primære analyse i studiet sammenligner effekten af lærer med undervisningskompetencer i faget med lærer uden det. Der foretages desuden sammenligninger af lærer med ordinær læreruddannelse, merituddannelse og anden uddannelse. Effekten bliver evalueret ved testresultater ved de nationale test i 2., 4., 6. og 8. klasse samt ved afgangsprøven i 9. klasse. Studiet finder blandede effekter af undervisningskompetencer i faget hos lærerne. Der findes ingen signifikante effekter på testresultater i 2., 4. og 6. klasse. I 8. klasse findes positive effekter på testresultater i biologi, geografi, fysik/kemi, men en negative effekt på testresultatet i dansk. I 9. klasse findes der negative effekter i nogle fag, men der er dog overvejende positive effekter blandt de statistisk signifikante parameterestimer, herunder en positive effekt på karakteren i faget dansk. Der er stor variation i effektstørrelser, forfatterne vurderer dog at de fleste effektstørrelser ligger på omkring 0,1-0,15 karakterpoint. Det vurderes til at være meget små effektstørrelser. Studiet giver mulighed for rent teknisk at udvikle indirekte evidens for den mone-tære effekt af kompetencedækning i faget på skoleområdet, men da estimerne i studiet (og i den øvrige forskning) er af blandet fortegn og generelt med meget små estimer må det forventes at en sådan model vil have meget små afledte økonomiske konsekvens som output. Samlet kan der med en stor analysepopulation dog godt være tale om større aggregerede effekter.

Børne- og Undervisningsministeriet har i evidensafsøgningen identificeret tre gode danske studier der ser på effekten af uddannet personale på dagtilbudsområdet. Alle tre studier benytter lodtrækningsforsøg til at identificere effekten af personalekompetencer. Der er dermed tale om evidens af høj kvalitet. Alle tre studier evaluerer effekten af at opkvalificere personalet i dagtilbud, så der er ikke tale om effekten af, at børnene har uddannet personale kontra ikke-uddannet personale, og resultaterne skal ses i dette lys. Begrebet "uddannet personale" kan dog godt siges også at inkludere personale, der er blevet opkvalificeret. Studierne behandler forskellige opkvalificering programmer, så effekterne kan ikke direkte sidestilles. Jensen et al. (2009) og Jensen et al. (2017) ser på effekter for 3-5-årige, mens Bleses et al. (2020) ser på effekter

for 2-3-årige. Studierne estimerer effekter på læring, mental helbred og på andre typer af kompetencer 1 til 2 år efter indsatsens start. Det vurderes ikke muligt at benytte studierne til at skabe direkte evidens for monetære effekter, men studierne kan benyttes som indirekte evidens for monetære effekter. Jensen et al. (2009) og Bleses et al. (2020) finder effekter på sproglige kompetencer med effektstørrelser fra 0.19 til 0.36 (Jensen et al. (2009), Sproglig kompetence: effekt på 0.36, Bleses et al. (2020), ordforråd: effekt på 0.27 og Bleses et al. (2020), sprogbrug: effekt på 0.19). Jensen et al. (2017) finder både signifikante og insignifikante effekter. Det vurderes at disse estimater er særligt egnede til at udvikle en model baseret på indirekte evidens.

9.4 Modeludvikling

Vi gennemgår mulighederne for modeludvikling med basis i de tre evidensniveauer, direkte evidens, indirekte evidens og statistisk sammenhæng.

Direkte evidens

Det vurderes på nuværende tidspunkt ikke muligt at gennemføre beregninger for uddannet personale i skoler og dagtilbud baseret på direkte evidens, da en sådan evidens kun foreligger i meget begrænset omfang.

Indirekte evidens

Indirekte evidens i skolen

Den evidens, der kan bruges i et indirekte evidens-setup, peger på små effekter med skiftende fortegn. Rent teknisk kan man dog godt, på baggrund af den nuværende evidens, gennemføre en beregning, og der findes dansk evidens, der understøtter positive, men små effekter i udskolingen. Metoden hertil beskrives i det følgende.

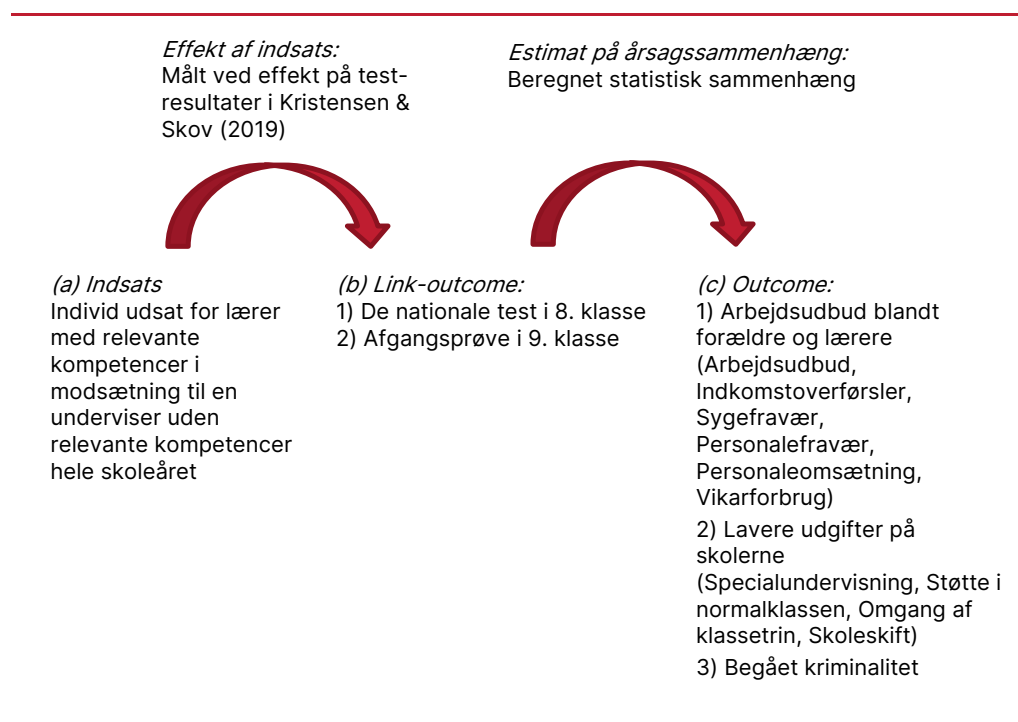
Figur 9.3 illustrerer modellen ved et indirekte evidensniveau for en indsats i skolen. Indsatsen er i afsnit 9.1 ovenfor modelteknisk afgrænset til, at eleverne undervises af lærer med relevante kompetencer i modsætning til en underviser uden relevante kompetencer hele skoleåret. Studiet af Kristensen & Skov (2019) benyttes som evidensgrundlag. Her findes ingen signifikante effekter på testresultater i 2., 4., og 6. klasse, men for 8. klasse og 9. klasse. Det anbefales derfor, at modellen for nærværende afgrænses til udskolingen. Forfatterne vurderer, at effekterne ligger på omkring 0,1-0,15 karakterpoint. Dvs. en effektstørrelse på testresultater på 0.03-0.05 SD for en lærer med undervisningskompetence i modsætning til en underviser uden undervisningskompetence.

Som link-outcomes til beregningen bør anvendes resultater fra de nationale test i 8. klasse eller afgangsprøveresultatet i 9. klasse. Valget af link er i dette tilfælde i høj grad betinget af den begrænsede tilgængelige evidens, men der foretages samtidig en vurdering af, at disse link-outcomes forventes at fange en væsentlig del af effekten baseret på den kvalificerede forandringsteori. Det forventes således, at testresultatet i væsentlig grad er korreleret med forbedring i læring, formindsket fravær og højere trivsel i klassen.

Det foreslås, at relationen mellem link-outcome og de økonomiske konsekvenser bliver beregnet baseret på statistiske sammenhænge i data. I tilfælde, hvor der kan findes evidens for en relation mellem link-outcome og et konkret økonomisk outcome, vil dette dog med fordel kunne erstatte den statistisk beregnede relation. En beregning af relationen mellem link-outcome og outcome åbner rent teknisk op for muligheden for at inkludere alle outcomes, der er tilgængelige data på (jf. Bilagstabel 2.1). Hvis man benytter den kvalificerede forandringsteori som udgangspunkt i udvælgelsen af relevante outcomes, vil man kunne inkludere følgende outcomes: 1) Arbejdsudbud blandt forældre og lærere (Arbejdsudbud, Indkomstoverførsler, Sygefravær, Personalefravær, Personaleomsætning, Vikarforbrug), 2) Lavere udgifter på skolerne (Specialundervisning, Støtte i normalklassen, Omgang af klassetrin, Skoleskift) og 3) Begået kriminalitet.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og link-outcomes beregnes, fx den betingede korrelation mellem brug af specialundervisning og resultat fra de nationale tests. I regressionerne betinges på standard socio-demografiske baggrundsvariable fra DST. Eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret og tilgængeligheden af de nationale tests, der samlet kun dækker folkeskolen med informationer fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multiple hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Figur 9.3 Model for indirekte evidens for uddannet personale i skolen



Indirekte evidens i dagtilbud

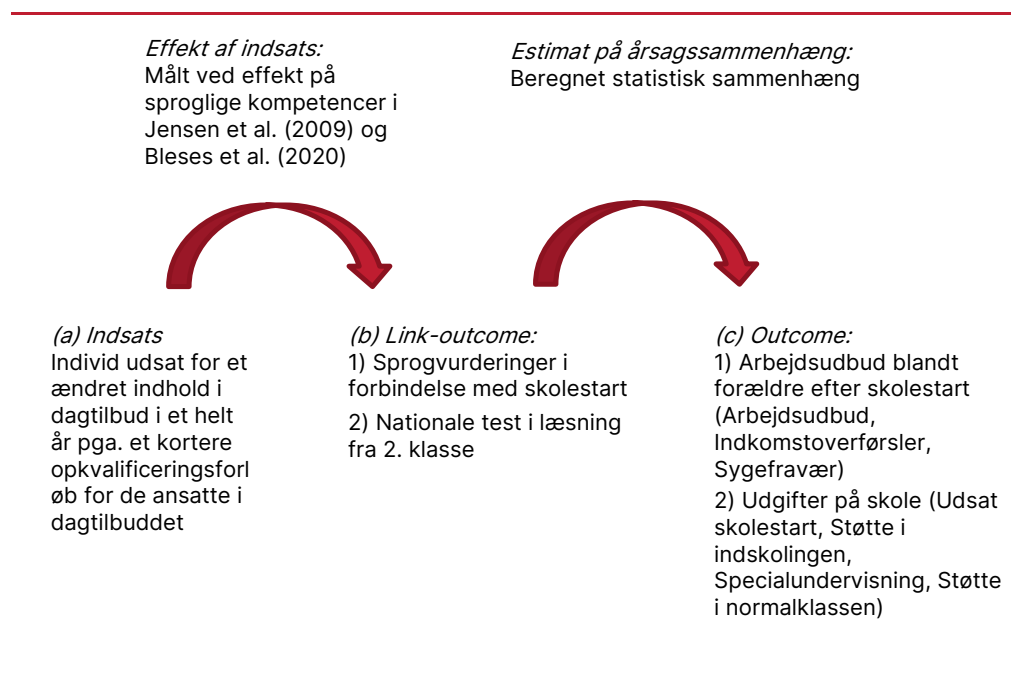
Figur 9.4 illustrerer modellen ved et indirekte evidens-niveau ved en indsats i dagtilbud. Indsatsen er i afsnit 9.1 defineret til at være, at alle medarbejdere i et dagtilbud modtager et kortere opkvalificeringsforløb, og dermed at barnet medtager en indsats, der lyder på en forbedret kvalitet i dagtilbuddet, hvor enheden er et år. Studierne af Jensen et al. (2009) og Bleses et al. (2020) benyttes som evidensgrundlag. Her finder effekter på sproglige kompetencer med effektstørrelser fra 0.19 til 0.36 SD.

Som link-outcomes til beregningen bør anvendes resultater fra sprogvurderinger i forbindelse med skolestart eller sprogvurderingen fra dagtilbud og eventuelt de nationale test i læsning fra 2. klasse. Det vil være fordelagtigt at benytte det tidligste af de to link-outcomes, dvs. sprogvurderinger i forbindelse med skolestart eller i dagtilbud. Valget af link er i dette tilfælde udfordret af den generelle mangel på registerdata om kompetenceudvikling og trivsel blandt børn i dagtilbud. Det er således nødvendigt at benytte registerdata på skolebørn for at opnå et rimeligt link-outcome. Disse link-outcomes vurderes

til at være fagligt brugbare, da de er relateret til outcome fra de empiriske studier, og i nogen grad fanger forventede effekter fra ændring i trivsel og udvikling hos børn i dagtilbud. Begrænsningen ved brug af disse link-outcomes er dog, at man ikke vil kunne se på alle effekter i dagtilbuddet, der kan forventes på baggrund af den kvalificerede forandringsteori, såsom brug af privat dagtilbud, brug af specialdagtilbud, personalefravær, personaleomsætning og vikarforbrug. Man vil ligeledes ikke kunne bruge disse link-outcomes til at måle på effekter på forældre før skolestart. På grund af disse begrænsninger giver det kun mening at gennemføre beregninger for de ældste børn i dagtilbud. De relevante outcomes vil derved blive reduceret til 1) Arbejdsudbud blandt forældre efter skolestart (Arbejdsudbud, Indkomstoverførsler, Sygefravær) og 2) Udgifter på skole (Udsat skolestart, Støtte i indskolingen, Specialundervisning, Støtte i normalklassen). Det vil hertil være muligt at tilføje andre skole-outcomes, hvis dette skønnes relevant (jf. Bilagstabel 2.1). Det fremhæves, at brug af disse skolestarts link-outcomes giver bedst mening for de ældste børn i dagtilbud på grund af den 4-årige beregningshorisont, da man kun vil kunne måle effekter efter link-outcome resultatet. Eksempelvis vil der for et barn på sit sidste år i dagtilbud måles link-outcome i år 1, hvorimod det for børn der har mere end 4 år tilbage i dagtilbud ikke vil være meningsfuldt at måle link-outcomes. Hvis man på denne baggrund vælger at fokusere på de ældste børn i dagtilbud, bør man undlade at benytte Bleses et al. (2020) som evidensgrundlag, da dette studie ser på børn i 2 til 3 års alderen. Det vil derved være studiet Jensen et al. (2009), der skal benyttes som empirisk grundlag.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome og link-outcome beregnes. F.eks. den betingede korrelation mellem brug af specialundervisning og sprogligvurdering i forbindelse med skolestart. I regressionerne betinges på standard sociodemografiske baggrundsvariable fra DST, som eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet, og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome vil populationen være afgrænset af specialundervisningsregistret og tilgængeligheden af sprogligvurdering i forbindelse med skolestart, der samlet kun dækker en undergruppe af børn der starter i folkeskolen. Hvis et stort antal outcomes inkluderes i modellen, anbefales det at man korrigerer for multipel hypotesetest ved brug af Bonferroni korrektion eller lignende.

Figur 9.4 Model for indirekte evidens for uddannet personale i dagtilbud



Baseline

For at anvende den indirekte evidens skal der findes estimater for det relevante baseline-niveau for beregningen. Dette baseline-niveau kan baseres enten på en lokal værdi for den kommune, beregningen skal dække, eller på en landsdækkende gennemsnitlig værdi. Det kan ikke forventes, at brugeren ligger inde med disse parametre. Derfor skal modellen tilbyde opdaterede tal på disse parametre fra alle landets kommuner. Disse tal beregnes på baggrund danske registerdata som årlige, kommunale gennemsnit for den pågældende målgruppe. For binære outcomes er baseline-niveauet en procentsats, og for kontinuerede outcomes er baseline-niveauet standardafvigelsen. For at kunne beregne indirekte evidens er der desuden behov for baseline-niveauer på det valgte link-outcome, dvs. standardafvigelsen på testresultater fra de nationale test for hver kommune samt landsgennemsnittet. Dette resulterer i en model, hvor brugeren blot behøver at vælge kommune. Brugeren vil derefter blive præsenteret for en række tal, som kan rettes til efter behov.

Beregn effektstørrelser for hvert relevant outcome

Som angivet i kapitel 3 skal der for hvert outcome foretages en beregning af den monetære effekt af indsatsen. Dette skal gøres for hver af de 4 år, efter indsatsen blev implementeret. Når alle monetære effekter er regnet ud, skal de lægges sammen som beskrevet i kapitel 3. Den monetære effekt af en indsats på et givet outcome beregnes på følgende måde:

$$P * g(\beta^{ab}, Base^{ab}) * h(\beta^{bc}, Base^{bc}),$$

hvor P er enhedsprisen på det betragtede outcome, β^{ab} er det omregnede estimat på testresultater for henholdsvis skole (Kristensen & Skov, 2019) og dagtilbud (Jensen et al., 2009 og Bleses et al., 2020), $Base^{ab}$ er det lokale baselineniveau på testresultatet i læsning, β^{bc} er det beregnede estimat på sammenhængen mellem testresultatet og det relevante outcome, og $Base^{bc}$ er baselineniveauet for det relevante outcome. Da der benyttes testresultat som link-outcome for både dagtilbud og skole, og da testresultater er en kontinuert variable, reduceres udtrykket for kontinuerte outcomes til:

$$P * \beta^{ab} * Base^{ab} * \beta^{bc} * Base^{bc}$$

Og for binære outcomes til:

$$P * \beta^{ab} * Base^{ab} * \left(\frac{Base^{bc} * e^{\beta^{bc} * 1.65}}{1 - Base^{bc} + Base^{bc} * e^{\beta^{bc} * 1.65}} - Base^{bc} \right)$$

Beregningen foretages på individniveau for hvert indsatsår. Det vil sige, at beregningen kan justeres for varierende indsatsdosis for forskellige indsatsår ved at foretage separate beregninger. Det vil ligeledes være muligt at tage højde for persistens over tid som beskrevet nedenfor. Den konkrete implementering heraf afhænger af valgt metode til at opgøre persistens.

Beregning af persistens i estimerne

For at modellen skal kunne give et bud på effekten af en indsats over en 4-årig periode, er der behov for estimer på effekten af indsatsen på alle outcomes for alle 4 år. Der er ikke forskningsmæssig evidens for dette, da indsatsen kun er evalueret, umiddelbart efter indsatsen er udført. Ved brug af link-outcome fra testresultater, kan persistensen (dvs. den vedvarende effekt) estimeres. Populationen er igen begrænset af tilgængeligheden af de nationale tests i folkeskoleelever fra 2010 og frem og for dagtilbud af tilgængeligheden af sprogvurderinger i forbindelse med skolestart (med mindre man vælger at benytte resultat fra nationale test i 2. klasse som link-outcome). Dette benyttes til at beregne den betingede korrelation i testresultater mellem to efterfølgende testår. Man har derved

opnået et mål for, hvor persistent testresultater er. Dette estimat benyttes til at foretage en lineær ekstrapolation over den 4-årige periode. Man opnår derved et mål for den betingede korrelation mellem testresultater i år 1 og testresultater i de følgende 3 år.

Effekten af indsatsen beregnes som ovenfor beskrevet for hvert outcome for hvert af de 4 år efter indsatsen. Disse tal omregnes til en nutidsværdi per individ, der modtager den fulde indsats som anført i kapitel 3. Dernæst aggregeres estimatet per individ op med dette antal individer.

Statistisk sammenhæng

For at kunne gennemføre beregninger baseret alene på statistisk sammenhæng er det nødvendigt at kunne knytte viden om, hvilke børn der har været omfattet af en indsats, til de efterfølgende realiserede outcomes.

På skoleområdet er dette muligt ved brug af register for lærernes kompetencer fra STIL, der har observationer fra 2013 og frem. Alternativt kan benyttes det nye lærer-elev-register, der forventes at være tilgængeligt hos DST i den nærmeste fremtid.

Der foretages for hvert inkluderet outcome en regression baseret på data fra STIL og DST, hvor den betingede korrelation mellem outcome, og hvorvidt individet har modtaget undervisning af en lærer med relevante kompetencer, bliver beregnet, fx den betingede korrelation mellem brug af specialundervisning og indsatsen. I regressionerne betinges på standard socio-demografiske baggrundsvARIABLE fra DST. Eksempelvis forældrenes uddannelsesniveau, indkomst, arbejdsmarkedsstatus og familieforhold samt barnets køn, etnicitet og tidlige mål for faglige kompetencer. Populationen er begrænset baseret på tilgængelighed af outcomes i registrene. Hvis man vælger at inkludere specialundervisning som et outcome, vil populationen være afgrænset af specialundervisningsregisteret, dvs. fra 2011 og frem. Hvis et stort antal outcomes inkluderes i modellen, anbefales det, at man korrigerer for multipel hypotesetest ved brug af Bonferroni-korrektion eller lignende.

Det vurderes ikke på nuværende tidspunkt at være meningsfuldt at foretage en beregning af indsatser med uddannet personel i dagtilbud af en rimelig kvalitet. Dette skyldes, at det på nuværende tidspunkt ikke med et rimeligt detaljeringsniveau er muligt at koble børn og det personale i dagtilbud, de har daglig kontakt med.

9.5 Ny evidens og ny data

Det vurderes på nuværende tidspunkt ikke at være meningsfuldt at gennemføre beregninger for uddannet personale i skole på direkte eller indirekte evidensniveau, da forskningsmæssig evidens på skoleområdet peger på små effekter med skiftende fortegn. På dagtilbudsområdet er der forskningsmæssig evidens for positive effekter, så her vil det være mere meningsfuldt at foretage en modeludvikling. Dette vil kunne gøres med et indirekte evidensniveau. En sådan beregning vil dog være begrænset af mangel på gode datakilder på kompetenceniveau for børn i dagtilbud, der kan bruges som link-outcome, idet data fra sprogtest stadig kun findes i en relativ kort tidsperiode. Hertil kommer, at der ikke er gode datakilder, der kan koble børn og personale. Rent teknisk kan man dog godt gennemføre en beregning af indirekte evidens på både skole- og dagtilbudsområdet på baggrund af den nuværende evidens. På skoleområdet er det ligeledes muligt at foretage beregninger alene baseret på statistiske sammenhænge.

På dagtilbudsområdet vil ny data med kompetencemål være meget brugbart til at beregne den indirekte evidens af indsatser i dagtilbud. Et alternativ er at oprette et detaljeret register, der er i stand til at koble børn og det personale i dagtilbud, de har daglig kontakt med. Dette vil gøre det muligt at udvikle en model baseret på statistisk beregninger og på sigt danne grundlaget for forskningsevidens, der kan bruges til en model baseret på direkte evidens.

På skoleområdet er der etableret gode datakilder, der kobler elever med lærere. Disse benyttes bl.a. i Kristensen & Skov (2019). Hvis man ønsker en model baseret på direkte evidens, kan man med fordel tage udgangspunkt i den empiriske strategi i denne forskningsartikel, men benytte de relevante outcomes, som afhængige variable i stedet for testresultater. Før man begiver sig ud i disse beregninger bør man dog have in mente at Kristensen & Skov (2019) finder effekter på testresultater med skiftende fortegn. Der er på den baggrund også en rimelig sandsynlighed for, at der findes små effekter med skiftende fortegn på outcomes, der kan monetariseres.

Litteratur

- Albertson, LM & Kagan, DM (1987), Occupational stress among teachers. *Journal of Research & Development in Education*, 21(1), 69-75.
- Andersen SC, Beuchert L, Nielsen HS & Thomsen MK (2018), The Effect of Teacher's aides in the Classroom: Evidence from a Randomized Trial, *Journal of the European Economic Association*, Volume 18, Issue 1, February 2020, pp. 469-505, <https://doi.org/10.1093/jeea/jvy048>
- Anderson SC & Winter S (red.) (2011), Ledelse, læring og trivsel i folkeskolerne. SFI Rapport 11:47. København https://pure.vive.dk/ws/files/254638/1147_Ledelse_laering_trivsel.pdf
- Angrist, JD & Lavy V (1999), Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2), 533-575.
- Angrist JD, Lavy V, Leder-Luis J & Shany A (2019). Maimonides' Rule Redux. *American Economic Review: Insights*, 1(3), 309-24.
- Arnbak E & Petersen DK (2017). Projekt It og Ordblindhed: En undersøgelse af it-støtte til ordblinde elever på mellemtrinnet. DPU Aarhus Universitet. Rapport. København.
- Avci S, Yüksel A, Soyer M, Balikcioglu S (2009), The Cognitive and Affective Changes Caused by the Differentiated Classroom Environment Designed for the Subject of Poetry. *Educational Sciences: Theory & Practice* 9 (3) • Summer 2009 • 1069-1084
<https://files.eric.ed.gov/fulltext/EJ858920.pdf>
- Balvig F (2011), Lovlydig ungdom. Det Kriminalpræventive Råd. København.
- Baumgartner T, Lipowski M & Rush C (2003), Increasing reading achievement of primary and middle school students through differentiated instruction. (ERIC Document Reproduction Service No. ED479203).
<https://files.eric.ed.gov/fulltext/ED479203.pdf>
- Beecher M & Sweeny SM (2008), Closing the achievement gap with curriculum enrichment and differentiation: One school's story. *Journal of Advanced Academics*, 19(3), 502-530. <https://files.eric.ed.gov/fulltext/EJ810785.pdf>

- Bengtsson TT, Knudsen L & Nielsen VL (2008), Kortlægning af kommunernes foranstaltninger til udsatte unge - Udviklingen efter anbringelsesreformen. SFI Rapport 08:30. København.
- Beuchert L, Dietrichson J & Jensen VM (2021), Forsøg med alternativ ressourceanvendelse: Baggrundsnotat med forslag til modeller. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.
- Bleses D, Jensen P, Slot P & Justice L (2020), Low-cost teacher-implemented intervention improves toddlers' language and math skills. *Early Childhood Research Quarterly* 53(4), 64-76
- Borg MG, Riding RJ & Falzon JM (1991), Stress in Teaching: a study of occupational stress and its determinants, job satisfaction and career commitment among primary schoolteachers, *Educational Psychology*, 11:1, 59-75, DOI: 10.1080/0144341910110104
- Boyes M, Leitao S, Claessen M, Dzidic P, Badcock NA & Nayton M (2021), Piloting "Clever Kids": A randomized-controlled trial assessing feasibility, efficacy, and acceptability of a socioemotional wellbeing program for children with dyslexia. *British Journal of Educational Psychology* 91(3), pp. 950-971
- Brooks F (2014), The link between pupil health and wellbeing and attainment (Issue November). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/370686/HT_briefing_layoutvFINAL.vii.pdf
- Browning M & Heinesen E (2007), Class size, teacher hours and educational attainment. *The Scandinavian Journal of Economics*, 109(2), 415-438
- Caprara GV, Barbaranelli C, Steca P & Malone PS (2006), Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44(6), 473-490. <https://doi.org/10.1016/j.jsp.2006.09.001>
- Chang C, Chiu L & Liu J (2017), A Study on the Relationship between Well-Being and Turnover Intentions among Rural School Teachers: School Organizational Climate as a Moderating Variable. *Creative Education*, 8, 523-538. <https://m.scirp.org/papers/75699>
- Chen H-C (1990), Theory-Driven Evaluations, London: SAGE Publications.
- Christoffersen MN & Højen-Sørensen A-K (2013), Børnehavens normeringer – En forskningsoversigt over opgørelsesmetoder. SFI Rapport 13:39. København

- Christoffersen MN, Højen-Sørensen A-K & Laugesen L (2014), Daginstitutionens betydning for børns udvikling – En forskningsoversigt. SFI Rapport 14:23. København
- Cook PJ & Kang S (2016), Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout and crime initiation. *American Economic Journal: Applied Economics*, 8(1), 33-64
- Currie J (2008), Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47(1), 87.
- Currie J & Stabile MB (2004), Child Mental Health and Human Capital Accumulation: The Case of ADHD. NBER Working paper series, 10435, 1–42.
- Dahler-Larsen P (2013), Evaluering af projekter - og andre ting, som ikke er ting, Syddansk Universitetsforlag. Odense
- Darling-Hammond L, Holtzman DJ, Gatlin SJ & Heilig JV (2005), Does Teacher Preparation Matter? Evidence about Teacher Certification, Teach for America, and Teacher Effectiveness. *Education Policy Analysis Archives*, 13(42)
- DEA (2020), Omgængere er udfordret fagligt og social. DEA Rapport. København. <https://dea.nu/i-farver/publikationer/omgaengere-er-udfordret-fagligt-og-socialt/>
- DeSocio J & Hootman J (2004), Children's mental health and school success. *J Sch Nurs*; 20: 189-96
- Dietrichson J, Filges T, Klokke RH, Viinholt BC, Bøg M & Jensen UH (2020), Targeted school-based interventions for improving reading and mathematics for students with, or at risk of, academic difficulties in Grades 7–12: A systematic review. *Campbell Systematic Reviews*, 16(2), e1081
- Dockett S & Perry B (2014), Continuity of Learning: A resource to support effective transition to school and school age care. Canberra, ACT: Australian Government Department of Education
- Donaldson SI (2007), Program theory-driven evaluation science: Strategies and applications, Lawrence Erlbaum Associates Publishers. New Jersey
- Duckworth AL, Quinn PD & Seligman MEP (2009), Positive predictors of teacher effectiveness. *The Journal of Positive Psychology*, 4(6), 540-547. <https://doi.org/10.1080/17439760903157232>

- Dyssegaard CB, Egeberg JdH & Steenberg K (2013), Skoleparathed – Systematisk forskningskortlægning. Dansk Clearinghouse for Uddannelsesforskning IUP, Aarhus Universitet
- Dyssegaard CB, Larsen MS & Hald BM (2013), Elever med særlige behov i almen skolen i forhold til inklusion, trivsel og selvværd. EMU Forskningskortlægning.
- Elbro C (2006), Literacy acquisition in Danish: A deep orthography in cross-linguistic light. *Handbook of orthography and literacy*, 31-45
- Epinion (2020), Fleksibel tilrettelæggelse af skoledagen. Rapport. København.
- EVA (2017), Forskningsopsamling: Kvalitet i dagtilbud – Pointer fra forskning. Rapport Danmarks Evalueringsinstitut (EVA). København
- EVA (2020), Læringsmiljø i kommunale børnehaver, Rapport Danmarks Evalueringsinstitut. København (<https://www.eva.dk/dagtilbud-boern/laeringsmiljoe-kommunale-boernehaver>)
- Evensen, M, Lyngstad, TH, Melkevik, O & Mykletun, A (2016), The Role of Internalizing and Externalizing Problems in Adolescence for Adult Educational Attainment: Evidence from Sibling Comparisons using Data from the Young HUNT Study. *European Sociological Review*, 32(5), 552–566
- Filges T, Sonne-Schmidt CS & Nielsen BCV (2018), Small class sizes for improving student achievement in primary and secondary schools: a systematic review. *Campbell Systematic Reviews*, 14, 1-107
- Fletcher, B & Payne, RL (1982), Levels of Reported Stressors and Strains Amongst Schoolteachers: some UK data, *Educational Review*, 34:3, 267-278, DOI: 10.1080/0013191820340308
- Framke E (2016), Effect and process evaluation of a participatory organizational intervention focusing on core job tasks. Ph.d.-afhandling. Aalborg Universitet
- Galinsky, E (1988), Parents and Teacher-Caregivers: Sources of Tension, Sources of Support, *Young Children*, 43 n3 pp. 4-12
- Galuschka K, Görge R, Kalmar J, Haberstroh S, Schmalz X & Schulte-Körne G (2020), Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist*, 55(1), 1-20

- Galuschka K, Ise E, Krick K & Schulte-Körne G (2014), Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLOS ONE*, 9(2), e89900.
- Gerber SB, Finn JD, Achilles CM & Boyd-Zaharias J (2001), Teacher Aides and Students' Academic Achievement. *Educational Evaluation and Policy Analysis*, 23(2), 123–143. <https://doi.org/10.3102/01623737023002123>
- Gutman LM & Vorhaus J (2012), The impact of pupil behaviour and wellbeing on educational outcomes. Report for the Department of Education, 2012, pp. 3–42.
- Gørtz M & Andersson E (2014), Child-to-teacher ratio and day care teacher sickness absenteeism, *Health Economics*, 23: pp. 1430-1442 (<https://onlinelibrary.wiley.com/doi/epdf/10.1002/hec.2994>)
- Hale-Jinks, C, Knopf, H & Knopf, H (2006), Tackling Teacher Turnover in Child Care: Understanding Causes and Consequences, Identifying Solutions, *Childhood Education*, 82:4, 219-226, DOI: 10.1080/00094056.2006.10522826
- Hanushek EA (1999), Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects. *Educational Evaluation and Policy Analysis*, 21(2), 143-163
- Heinesen E (2009), Estimating Class-size Effects using Within-school Variation in Subject- specific Classes. *The Economic Journal*, 120(545), 737-760
- Jakubowski M & Sakowski P (2006), Quasi-experimental estimates of class size effect in primary schools in Poland. *International Journal of Educational Research*, 45(3), 202-215
- Jensen B, Holm A, Allerup P, Kragh A (2009), Effekter af indsatser for socialt udsatte børn i daginstitutioner. HPA-projektet. Danmarks Pædagogiske Universitet. København
- Jensen B, Jensen P & Rasmussen AW (2017), Does professional development of preschool teachers improve children's socio-emotional outcomes? *Labour economics*, 45, 26-39
- Jensen VM, Bjørnholt B, Mikkelsen MF, Nielsen CP & Ladekjær E (2020), Den længere og mere varierede skoledag – En analyse af reformens elementer. VIVE Rapport. København.

- Keilow, M, Sievertsen, HH, Niclasen, J & Obel, C (2019), The Strengths and Difficulties Questionnaire and standardized academic tests: Reliability across respondent type and age. *PLOS ONE* 14(7), 1–12.
<https://doi.org/https://doi.org/10.1371/journal.pone.0220193>
- Kern L, Delaney B, Clarke S, Dunlap G & Childs K (2001), Improving the classroom behavior of students with emotional and behavioral disorders using individualized curricular modifications. *Journal of Emotional and Behavioral Disorders*, 27, 615-631
- Kjeldsen MM (2018), Effekten af børnepasningens pris og kvalitet på forældrenes arbejdsudbud. De Økonomiske Råds Sekretariat, Arbejdsrapport 2018:1
- Krassel KF & Heinesen E (2014), Class-size effects in secondary school. *Education Economics*, 22(4), 412-426
- Kristensen N & Skov PR (2019), Betydningen af kompetencedækning og læreruddannelsesbaggrund. VIVE Rapport. København
- Krueger, AB (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2), 497-532
- Landersø R, Nielsen HS & Simonsen M. (2015), School Starting Age and The Crime-Age Profile. *The Economic Journal*, 127(602), 1096–1118.
- Landersø R, Nielsen HS & Simonsen M (2020), Effects of School Starting Age on the Family. *Journal of Human Resources*, 55(4), 1258-1286
- Lee S, Wehmeyer ML, Soukup JH & Palmer SB (2010), Impact of curriculum modifications on access to the general education curriculum for students with disabilities. *Exceptional Children*, 76 (2), 213-233
- Lewis AD, Huebner ES, Malone PS & Valois R (2011), Life satisfaction and student engagement in adolescents. *J Youth Adolesc*; 40:249–62
- Lyons MD & Huebner ES. (2016), Academic characteristics of early adolescents with higher levels of life satisfaction. *Appl Res Qual Life*; 11:757–71.
- Madsen M, Hauberg DS, Kolstrup K & Toftum J (2020), Samfundsøkonomiske gevinster ved forbedret indeklime. Rapport Incentive og DTU. København
- Ministeriet for Børn, Undervisning og Ligestilling (2016). Sammenhængen mellem elevernes trivsel og elevernes nationale testresultater

- Mortensen, NP (2020), Uddannelsesresultater og -mønstre for børn og unge med funktionsnedsættelser, VIVE-rapport. København.
- Mosteller, F (1995). The Tennessee study of class size in the early school grades. *The Future of Children*, 5(2), 113-127
- Nandrup AB (2016). Do class size effects differ across grades? *Education Economics*, 24(1), 83-95
- Ng ZJ, Huebner ES & Hills JK (2015), Life satisfaction and academic performance in early adolescents: Evidence for reciprocal association. *J Sch Psychol*, 53:479-91.
- O'Connor PR & Clarke VA (1990), Determinants of Teacher Stress. *Australian Journal of Education*; 34(1):41-51. doi:10.1177/000494419003400103
- Oishi S, Diener E & Lucas RE. The optimum level of well-being: Can people be too happy? *Perspect Psychol Sci* 2007; 2:346-60
- Ottosen MH, Andreasen AG, Dahl KM, Hestbæk A-D, Lausten M & Rayce SLB (2018), Børn og Unge i Danmark - Velfærd og trivsel 2018. København: VIVE – Det Nationale Forsknings- og Analysecenter for Velfærd.
- Pawson R & Tilley N (1997), Realistic evaluation, London: SAGE Publications
- Penney J (2018), Dynamic Treatment Effects of Teacher's Aides in an Experiment with Multiple Randomizations, *Economic Inquiry*, Volume 56, 2, 1244-1260
- Perlman M, Fletcher B, Falenchuk O, Brunsek A, McMullen E & Shah PS (2017), Child-Staff ratios in early childhood education and care settings and child outcomes: A systematic review and metaanalysis. *PLOS ONE*, 12(1):e0170256
- Rambøll (2018), Undersøgelse af hjemmeundervisning, fravær og børn uden for undervisningstilbud, Rapport til Undervisningsministeriet. København
- Rangvid, BS (2018), Student engagement in inclusive classrooms, *Education Economics*, 26:3, 266-284.
- Rosholm M, Nielsen SA & Hvidman C (2020), *Effektevaluering af tre intensive læringsforløb*. TrygFondens Børneforskningscenter. Rapport. Aarhus

- SBU (2014), Dyslexi hos barn och ungdomar - tester och insatser. En systematisk litteraturöversikt. Statens beredning för medicinsk utvärdering (SBU) Rapport nr. 225. Stockholm
- Sibieta L (2020), Effect of an area-wide campaign to improve the use of teaching assistants: an application of synthetic control methods, in PhD-thesis: Essays in the Economics of Education: School resource Decisions and Student Achievement, UCL PhD thesis
- Sievertsen HH (2015), En god start - Betydningen af alder ved skolestart for barnets udvikling. SFI Rapport 15:38. København
- Strogilos V (2018), The value of differentiated instruction in the inclusion of students with special needs/ disabilities in mainstream schools; https://www.shs-conferences.org/articles/shsconf/pdf/2018/03/shsconf_gctale2018_00003.pdf
- Sturm, P (1997), Nevada's Class-Size Reduction Program, Carson City, NV: Senate Committee on Human Resources. www.leg.state.nv.us/lcb/research/bkground/97-07.HTM.
- Turner K & Theilking M (2019), Teacher Wellbeing: Its Effects on Teaching Practice and Student Learning, *Issues in Educational Research*, v29 n3 p938-960
- Weiss CH (1997). How Can Theory-Based Evaluation Make Greater Headway? *Evaluation Review* 21(4):501-524
- Williams C & Beam S (2019), Technology and writing: Review of research. *Computers & education*, 128, 227-242
- WSIPP (2019). Benefit-Cost Technical Documentation – Washington State Institute for Public Policy Benefit – Cost Model (<http://www.wsipp.wa.gov/TechnicalDocumentation/WsippBenefitCostTechnicalDocumentation.pdf>)

Bilag 1 Procesbeskrivelse for formulering af forandringsteoriene før vidensunderbygning

Workshops

Der blev i forbindelse med formuleringen af forandringsteoriene afholdt 8 workshops af halvanden times varighed med følgende indsatser og fokusområder:

- Klassekvotient (5. november 2020)
- Normering i dagtilbud (1. december 2020)
- Udsat skolestart (14. december 2020)
- To-voksen-ordninger (14. december 2020)
- Større andel uddannet personale i dagtilbud (14. december 2020)
- Større andel uddannet personale i skolen (14. december 2020)
- Ordblindeindsatser (14. december 2020 og opfølgende 6. januar 2021).

På hver workshop introducerede VIVE først tilgangen til arbejdsgangen og processen for workshoppen og redegjorde for begreberne i forandringsteoretisk arbejde. Herefter brainstormede deltagerne på indhold og elementer til forandringsteoriene. Et vigtigt fokuspunkt for udarbejdelsen af virkningskæderne var, at den formulerede virkning skulle kunne forventes at indtræffe inden for den 4-årsperiode, som er afgrænsningen af de økonomiske effekter, der indgår i projektet. De økonomiske indikatorer er også formuleret på workshops og kan derfor stamme fra forskellige kilder. VIVEs deltagere på de respektive workshops havde indsigt i den for hver indsats gennemførte evidensundersøgelse, hvorfor formulerede indikatorer også kunne stamme fra denne. Økonomiske indikatorer fra en workshop blev i nogle tilfælde taget med videre til de næste workshops, således at forandringsteoriene informerede og underbyggede hinanden.

Bruttoforandringsteorier

På baggrund af input fra de afholdte workshops og med inddragelse af VIVEs faglige viden på områderne formulerede VIVE overordnede bruttoforandringsteorier for de enkelte indsatser. Bruttoforandringsteoriene er som udgangspunkt ikke underbyggede af litteratur og skal derfor også udelukkende ses som det umiddelbare output af de afholdte workshops.²⁵ Dermed indgik de i arbejdet som grundlag for den videre udformning af forandringsteoriene. På

²⁵ Bruttoforandringsteoriene kan downloades som selvstændige Powerpoint-filer fra VIVEs hjemmeside.

dette stadie er en del af de specifikke virkningskæder således udviklet på baggrund af, hvordan virkningskæden hypotetisk set kan formuleres, men uden at tage højde for, om der er empirisk grundlag for de foreslåede sammenhænge og uden at kende kvaliteten af et eventuelt empirisk vidensgrundlag. Forandringsteoriene i bruttoversion skal derfor også ses som en del af processen og ikke som noget, der kan bruges til at sige noget direkte om sammenhænge eller mangel på samme. Derimod udgør de en vigtig inspiration for og grundlag for det videre arbejde.

Kvalificering og klargøring til vidensunderbygning

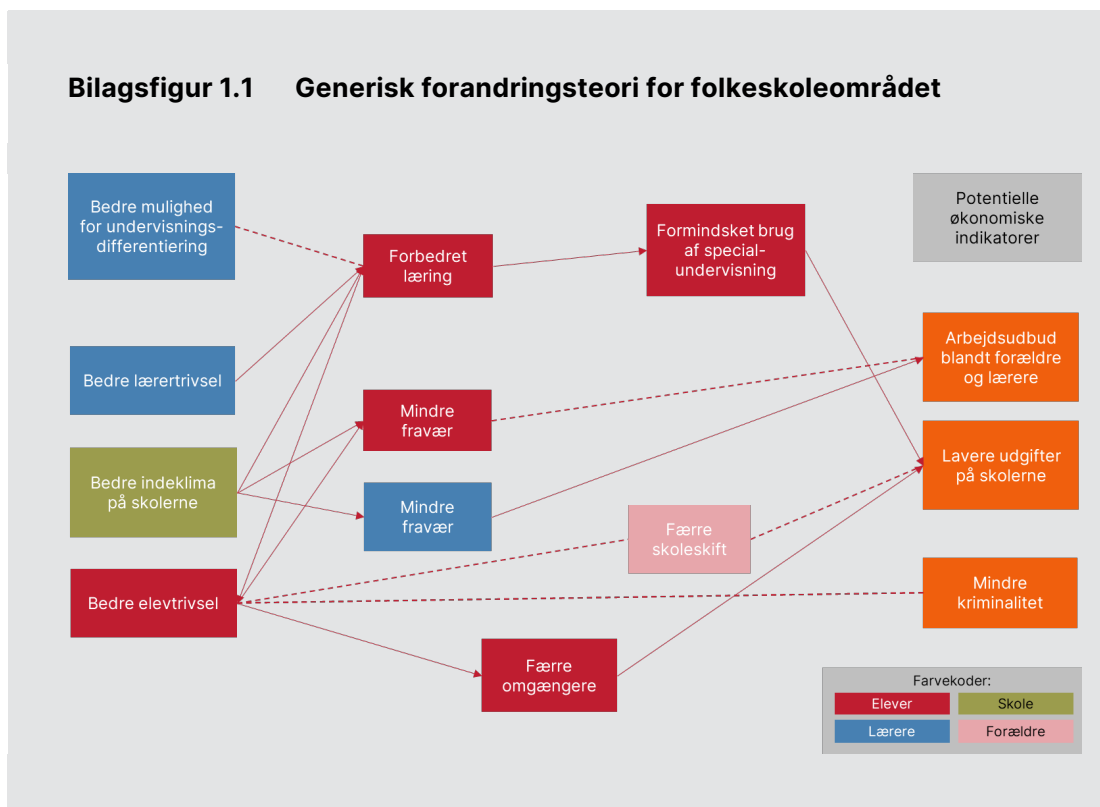
På baggrund af bruttoforandringsteoriene har VIVE formuleret to generiske forandringsteorier for henholdsvis dagtilbud og skoleindsatser. Alle seks indsatsers virkningskæder vil teoretisk set aktivere én eller flere virkningskæder i den tilhørende generiske forandringsteori, fx kan en indsats inden for skoleområdet aktivere den virkningskæde, der begynder med "bedre elevtrivsel". Dermed får indsatserne en fælles vej frem mod de potentielle økonomiske konsekvenser gennem forandringsteoriernes virkningskæder. De generiske forandringsteorier er stadig teoretiske forståelser af virkningskæderne.

VIVE har i forbindelse med arbejdet med formuleringen af de to generiske forandringsteorier gennemarbejdet de formulerede bruttoforandringsteorier. Fra disse er kun medtaget de virkningskæder, der forventes at medføre en økonomisk virkning som udtrykt gennem de potentielle økonomiske konsekvenser, ligesom virkningskæder, der ikke umiddelbart er meningsfulde i det videre arbejde, er udeladt, fx er administrative udgifter i skolerne udeladt. Dette arbejde har VIVE udført på baggrund af faglig indsigt i indsatserne og områderne. Den faglige indsigt underbygges desuden af følgende kriterier for forbindelserne i de generiske forandringsteorier:

- Virkningskæden skal kunne forbindes med en økonomisk indikatorer. Hvis virkningskæden ender blindt eller i en indikator, der ikke er økonomisk målbar, undlades den.
- Virkningskæderne skal have potentiale til at nå at påvirke den formulerede økonomiske konsekvens inden for den 4-årsperiode, som dette projekt omhandler.

Den generiske forandringsteori for folkeskoleområdets indsatser kan ses i Bilagsfigur 1.1.

Bilagsfigur 1.1 Generisk forandringsteori for folkeskoleområdet

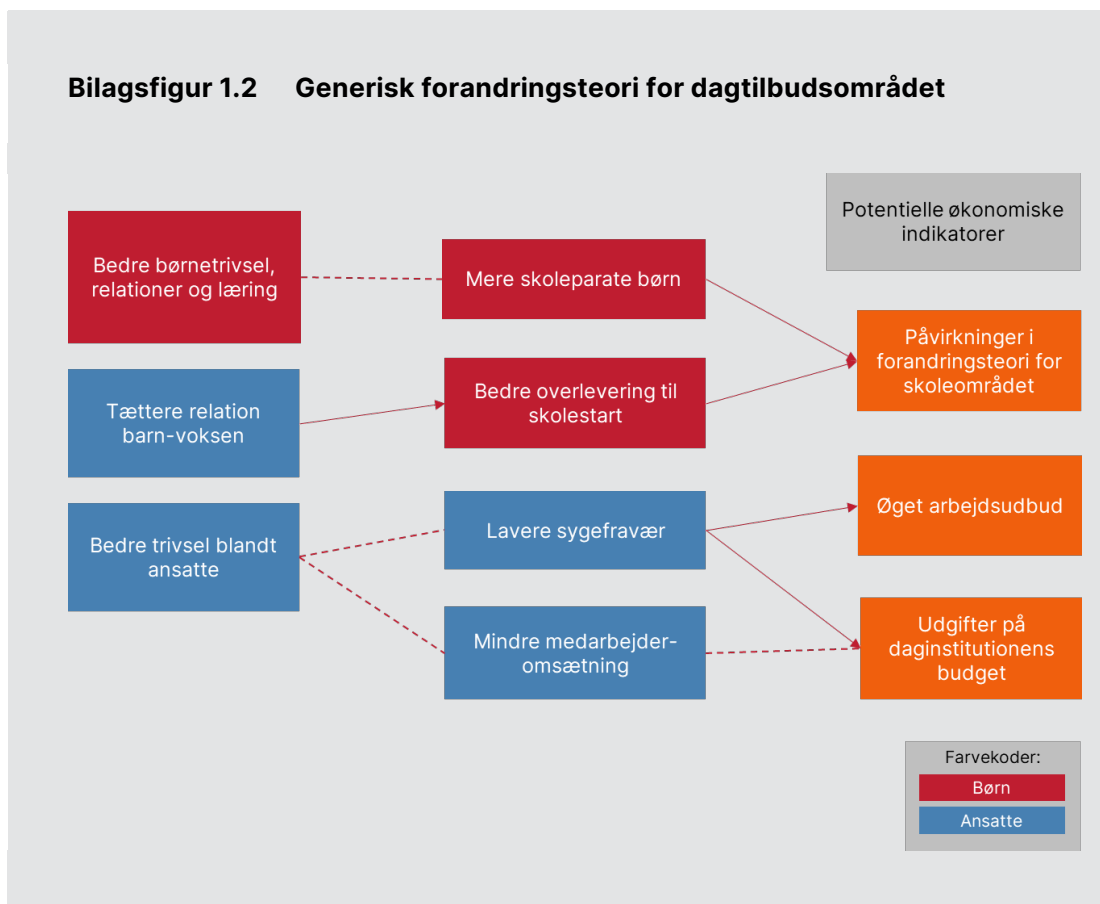


De generiske forandringsteorier kan bruges til at belyse de forskellige indsatser, som de er formuleret i bruttoforandringsteorierne på baggrund af workshopperne. Som eksempel på, hvordan de generiske forandringsteorier skal anvendes, kan bruttoforandringsteorien om ændret klassestørrelse bruges. I den forandringsteori, som den blev formuleret på baggrund af workshoppen, er der en række forskellige umiddelbare virkninger såsom "Bedre indeklima", "Bedre lærertrivsel" og "Bedre mulighed for undervisningsdifferentiering". Disse forbindelser kan genfindes i den generiske forandringsteori, og dermed kan de formulerede virkningskæder anvendes til at forbinde indsats med potentiel økonomisk indikator. Dette gør sig vel at mærke kun gældende, hvis disse kæder også vidensunderbygges og fastholdes i de endelige forandringsteorier.

I denne del af rapporten holdes de økonomiske indikatorer på et meget overordnet niveau, men de vil senere i rapporten blive uddybet og konkretiseret. Ligeledes skal de økonomiske indikatorer ikke tolkes som en bestemt post i den kommunale økonomi, men kan godt dække over en række forskellige udgiftsposter. Et eksempel på dette er den meget overordnede post "Udgifter på skoleområdet", der dækker over en række forskellige poster i forskellige regnskaber.

Den generiske forandringsteori for dagtilbudsområdet kan ses i Bilagsfigur 1.2.

Bilagsfigur 1.2 Generisk forandringsteori for dagtilbudsområdet



Efter formuleringen af de generiske forandringsteorier var næste skridt i processen frem mod de endelige forandringsteorier, der anvendes i indsatskapitlerne i denne bog, at VIVE foretog en vidensunderbygning. Denne sidste del af processen fremgår af kapitel 2.

Bilag 2 Dataundersøgelse

Dataundersøgelsen har afdækket mulige datakilder for beregninger i en kommende økonomisk model. I dette kapitel opsummeres dataundersøgelsen helt overordnet, mens de enkelte outcomes vil blive nærmere beskrevet under hver indsats.

Dataundersøgelsen er gennemført som et samarbejde imellem VIVE og BUVM. BUVM har gennemført afdækningen for skole-/dagtilbudsområdet, mens VIVE har gennemført afdækningen for de tre øvrige områder. Det er ligeledes VIVE, der har stået for sammenfatningen af dataundersøgelsen. Ud over beskrivelserne i rapporten findes også et separat Excel-bilag, som kan rekvireres ved henvendelse til VIVE, med en samlet oversigt over de undersøgte outcomes og områder.

Metode i dataundersøgelsen

Dataundersøgelsen blev udført i tre trin. Første trin var at opstille lister for, hvilke områder det kunne være relevant at have data for, samt hvilken detaljeringsgrad der var nødvendig. Dette arbejde mandede ud i, at der blev arbejdet inden for følgende opdeling:

- Relevante outcomes dagtilbud
- Relevante outcomes 0.-5. klasse
- Relevante outcomes 6.-9. klasse.

Næste trin var at yderligere underopdele på relevante områder inden for hvert af de tre aldersinddelinger:

- Skole-/dagtilbudsområdet (fx læring, trivsel, fravær)
- Sundhedsudgifter for individet (fx brug af sundhedsvæsenet)
- Adfærdsudgifter for individet (fx kriminalitet og sociale foranstaltninger)
- Adfærdsudgifter for familiemedlemmer (fx forældres arbejdsmarkedstilknøytning).

Sidste trin var herefter at undersøge den præcise datadækning på disse områder. Dette er udfoldet i næste afsnit og det tilknyttede Excel-bilag. Herudover er der også knyttet kommentarer til data i gennemgangen af de enkelte indsatser i de kommende kapitler.

Resultat

Generelt viser dataafdækningen, at der er en god datatilgængelighed på skoleområdet, men ikke helt så detaljerede data på dagtilbudsområdet.

Data på dagtilbudsområdet

På dagtilbudsområdet findes data for bl.a. personalefravær, den statslige sociale normering og sprogvurderinger. Imidlertid er kun personalefravær registreret på en måde, der opfylder kriterierne for så vidt angår længde og dækning. Det skyldes, at den statslige sociale normering først er tilgængelig fra og med 2019, mens sprogvurderinger kun er tilgængelige i for 2016-2018 og kun for et udvalg af kommuner. Dette betyder, at børnenes trivsel, som må anses for at være en helt central parameter i målingen af effekter af indsatser i dagtilbud, findes der ikke data for, ligesom der også kun er meget begrænset data om anvendelse af støtte i dagtilbud og specialdagtilbud på individniveau.

Data på skoleområdet

På skoleområdet er der betydeligt bedre data, hvilket først og fremmest skyldes muligheden for at anvende data fra de nationale test og den nationale trivselsmåling. Data fra de nationale test er til rådighed for 2010-2019, mens data fra trivselsmålingen er til rådighed for 2015-2019 – hvilket altså i begge tilfælde opfylder kriterierne for længde og dækning. Herudover er der også på skoleområdet data for både elevfravær og personalefravær samt for anvendelse af specialklasser og specialskole. Disse data opfylder i alle tilfælde kriteriet for, hvor længe der findes data for. Andre relevante data, der findes på skoleområdet, er afgangskarakterer ved 9. og 10. klasses eksamen samt påbegyndelse af ungdomsuddannelse.

Det bemærkes, at data fra de nationale test ikke findes for alle individer for alle dæknings-år, da de nationale test kun afholdes i bestemte fag på bestemte klassetrin. Det betyder, at anvendelsen af denne datakilde eventuelt skal gøre brug af sammenkobling af data fra flere individer for at danne "tids-serier".

Det bemærkes ligeledes, at data for skolefravær, de nationale test og den nationale trivselsmåling kun er tilgængelige for folkeskoler, men ikke for fri- og efterskoler. Dette bør dog ikke betyde så meget i denne sammenhæng, da den økonomiske model er tænkt som et redskab til indsatser på folkeskoleområdet.

Data relateret til andre samfundsområder

Ud over de områder, der direkte er relateret til skole og dagtilbud, skal en kommende model også kunne måle andre områder. I den Socialøkonomiske Investeringsmodel er der således mulighed for at inddrage områder som arbejdsmarked, indkomstoverførsler, socialområdet, kriminalitet og sundhed, og det vil være naturligt også at se på disse områder til en model på BUVM's område.

På disse områder findes der bredt dækkede registerdata med individoplysninger, der stilles til rådighed fra Danmarks Statistisk og Sundhedsdatastyrelsen. Disse datakilder vil bl.a. kunne anvendes til at analysere:

- Forældres arbejdsudbud. Dette område vil kunne belyses både igennem optjent arbejdsindkomst og gennem modtagelse af indkomsterstattende ydelser fra det offentlige. Et område, der dog ikke findes særlig god data om, er sygefravær på arbejdsmarkedet. Det skyldes, at der for private virksomheder kun indsamles data for sygefravær hos en stikprøve af virksomhederne.
- Sundhedsudgifter. På dette område vil det fx være muligt at se på besøg hos lægen, indlæggelser og ambulante behandlinger på hospital. Der findes detaljerede data herfor igennem en længere årrække.
- Sociale foranstaltninger. Data på dette område dækker sociale tilbud iværksat af kommunerne over for socialt udsatte børn og unge. På dette område findes detaljerede data på individniveau.
- Kriminalitet. På dette område findes der individdata både for sigtelser for unge under 15 år og for domme for personer fra 15 år. Disse vil kunne inddrages i en økonomisk model. Der findes også data for anmeldte forbrydelser og for ofre for forbrydelser, men disse vurderes det ikke i første omgang muligt at knytte prisbeløb på.

Et område, som der ikke findes data på, er generelt trivsel blandt familiemedlemmer. Dette område kunne godt være et relevant outcome for indsatser på området, men det vil ikke kunne inddrages her.

Bilagstabel 2.1 viser en oversigt over de potentielle økonomiske konsekvenser for en skoleindsats og de mulige datakilder, mens Bilagstabel 2.2 viser en tilsvarende oversigt for dagtilbudsområdet.

Bilagstabel 2.1 Potentielle outcomes for skoleindsatser

Outcome-domæne	Outcome	Mulige link-outcomes	Datakilde
Barnet	Specialundervisning	Læring Trivsel	DST
	Støtte i normalklassen	Læring Trivsel	DST
	Omgang af klassetrin	Trivsel Læring	DST
	Skoleskift	Trivsel Læring	DST
Institutionen	Personalefravær	—	DST
	Personaleomsætning	—	DST
	Vikarforbrug	—	
Uddannelse	Gennemførelse af uddannelse	Trivsel Læring	DST
Sundhedsområdet	Mental sundhed	Trivsel Ungdomsuddannelse	DST/SDS
	Fysisk sundhed	Trivsel Ungdomsuddannelse	DST
Socialområde	Ofre for kriminalitet	—	DST
	Forebyggende foranstaltninger	Trivsel	DST
	Anbringelser	Trivsel	DST
Adfærd (barnet)	Begået kriminalitet	Trivsel Læring Ungdomsuddannelse	DST
	Arbejdsudbud (fritidsjob)	Trivsel Læring Ungdomsuddannelse	DST
Adfærd (familien)	Arbejdsudbud	Trivsel	DST
	Indkomstoverførsler	—	DST
	Sygefravær	Trivsel	DST

Kilde: Egen tilvirkning.

Bilagstabel 2.2 Potentielle outcomes for dagtilbudsindsatser

Outcome-domæne	Outcome – inspireret af i forandringsteori og konkretiseret af tilgængelige data	Mulige link-outcomes	Datakilde
Barnet	Udsat skolestart	Trivsel Læring	DST
	Støtte i indskolingen	Trivsel Læring	DST
	Omgang af klassetrin	Trivsel Læring	DST
	Brug af privat dagtilbud	—	DST
	Brug af specialdagtilbud	—	—
Institutionen	Sygefravær blandt personalet	—	DST
	Medarbejderomsætning	—	DST
	Kapacitetsbehov	—	—
	Antal ansatte	—	—
	Modtager af puljemidler til sociale normeringer	—	BUVM
Sundhedsområdet	Mental sundhed	Trivsel	DST/SDS
	Fysisk sundhed	Trivsel	DST
Socialområde	Ofre for kriminalitet	—	DST
	Forebyggende foranstaltninger	Trivsel	DST
	Anbringelser	Trivsel	DST
Adfærd (familien)	Arbejdsudbud	—	DST
	Indkomstoverførsler	—	DST
	Sygefravær	—	DST

Bilag 3 Litteratursøgningsprotokol

Det grundlæggende spørgsmål, der søges besvaret med evidensundersøgelsen, er, hvorvidt der i den internationale litteratur er evidens for, at hver af de seks indsatser virker – altså medfører et forbedret outcome på det område, hvor det var tiltænkt. På undervisningsområdet vil de outcomes, der er relevante som effektmål, ofte vedrøre enten faglige mål som forbedring i faglige test eller prøvekarakterer, eller det vil være mål for elevernes trivsel i skolen. Endvidere kan elevernes fravær i skolen være relevant som et mellemliggende resultatmål. Endelig kan der også være tale om mål, der ser på, hvor mange der gennemfører skolen eller kommer videre i uddannelsessystemet.

For hver af de seks indsatser vil det således blive undersøgt, om der er evidens for indsatsens effekt. Inden for projektets tidsramme er det ikke muligt at gennemføre en omfattende systematisk litteratursøgning med screening af fundne studier. Derfor tages udgangspunkt i eksisterende (systematiske) litteraturoversigter til formålet. Det svarer i nogen grad til det, der også kaldes en *rapid evidence assessment*, der adskiller sig fra en mere grundig litteraturnemgang med hensyn til bredden og dybden af gennemgangen. Den generelle litteratursøgning suppleres med ekstra søgninger med henblik på at øge antallet af skandinaviske studier, der vurderes at have størst relevans i en dansk kontekst.

Indsatser

- Udsat skolestart, omgængere
- Betydningen af uddannet personale i skoler og dagtilbud
- Normeringer i dagtilbud
- Klassekvotient (grundskolen)
- Tovoksenordning (grundskolen)
- Ordblindeindsatser.

Outcomes

Følgende outcomes kan fx være relevante som effektmål på undervisningsområdet:

- faglige mål/test (fx andelen, der har et bestemt niveau)
- mål for elevernes trivsel i skolen
- mål for fravær
- gennemførelse af skole/ungdomsuddannelse
- overgang til ungdomsuddannelse
- normeringer i dagtilbud (trin 2 outcome)
- forældres arbejdsudbud (trin 2 outcome).

International litteratursøgning

Kilder:

- - Web of Science
- - Eric
- - Psycinfo
- - Campbell
- - evt. Cochrane.

Søgestrategi

Strategien er at finde frem til en eller flere centrale publikationer, som giver referencer til centrale enkeltstudier.

For at identificere centrale publikationer søges i første omgang efter systematiske reviews og meta-analyser. Kun hvis det ikke findes, ledes videre efter øvrige oversigtsartikler og evt. efter artikler i højtrangerende videnskabelige tidsskrifter.

De centrale publikationer er tænkt som en hurtigere genvej til et antal relevante enkeltstudier. De således fundne enkeltstudier screenes, og relevante studier optages i vores evidensdatabase (se nedenfor).

Søgningerne i de ovennævnte ressourcer suppleres med:

- Referencer identificeret via direkte søgninger i de nævnte databaser
- 'Referencer fra referencer' forstået som referencer identificeret gennem eksisterende systematiske reviews inden for området samt referencer fra referencer identificeret i databasesøgningerne (snowballing).
- Citation tracking, dvs. der identificeres artikler, som henviser til det centrale review for at finde litteratur, der er nyere end de(n) centrale publikation(er) (forward snowballing).

Søgetermer

Generelle søgetermer:

School, primary school, lower secondary school, elementary school, middle school, junior high school; daycare, pre-school, kindergarten
(Systematic) reviews, meta-analysis

1. *Udsat skolestart, omgængere*
Delayed school enrollment, school starting age, school entrance age, school readiness; grade repetition, grade retention
2. *Uddannet personale i skoler og dagtilbud*
Qualified teachers/staff, teacher/staff qualifications/certification/competencies/education/specialization
3. *Normeringer i dagtilbud*
Child-to-adult ratio in daycare/childcare, child-staff ratio, staff-child ratio, teacher-child ratio, staff-to-child ratio, educator to child ratio, childcare ratio; class size, group size
4. *Klassekvotient*
Class size; pupil-teacher ratio, student-teacher ratio
5. *Tovoksenordning*
Teacher's aide(s), teaching assistant, education assistant, team teacher; co-teaching
6. *Ordblindeindsatser*
Dyslexia treatment/cure/intervention/program/therapy/tutoring/support; reading disability/disorder; dyslexic learners.

Inklusion- og eksklusionskriterier:

- Der medtages kun kvantitative studier:
 - Højt evidensniveau, dvs. metastudier, lodtrækningsforsøg eller effektmåling med statistisk udvalgt kontrolgruppe (herunder regressionsstudier, hvor nogle individer ikke har modtaget nogen indsats eller har modtaget forskellige niveauer/doser af indsats)
 - Som udgangspunkt medtages ikke studier med før-efter-måling
- Der medtages:
 - Publicering efter 2000
- Der medtages ikke:
 - Studier, som ikke omhandler dagtilbud eller grundskole
 - Meget korte studier (mindre end 10 sider)
 - Bachelor- og masteropgaver.

Etablering af database til opsamling af oplysninger fra evidenssøgningen:

Der etableres en Excel-database over de oplysninger, der indsamles fra hvert studie.

Nordisk litteratursøgning

Kilder:

- Forskningsdatabasen
- Evalueringsportalen, Norge
- TrygFondens Børneforskningscenter
- ROCKWOOL Fonden
- Helsedirektoratet
- VIVE/SFI/KORA
- Kunnskapssenteret, Norge
- Bufdir – Nettbiblioteket, Norge
- Washington State Institute for Public Policy (WSIPP)
- Evidensdatabasen (DPU)
- Campbell
- Det svenske Skoleforskningsinstitut
- Education Endowment Foundation
- What Works
- European Expert Network on Economics of Education
- EVA
- NB-ECEC (Nordic Base of Early Childhood Education and Care).

VIVÉ