

LØSNINGSBESKRIVELSE

FØLGEFORSKNING TIL FOR- SØGSPROGRAM OM MO- DERSMÅLSUNDERVISNING

- FORSØGSINDSATSER RETTET MOD 1. KLASSETRIN

This document presents the protocol for the randomized controlled trial that is reported in Andersen, Guul, & Humlum (2021): "How First-Language Instruction Transfers to Majority-Language Skills", *Nature Human Behaviour*, forthcoming. The protocol was an appendix to the contract between the Ministry of Education and Aarhus University signed before the beginning of the trial. Sensitive information (such as the budget) has been removed or redacted.

INDHOLD

1	Indledning.....	3
2	Projektet.....	4
2.1	Overblik	4
2.2	Forsøgsdesign.....	5
2.3	dataindsamling	5
2.4	Efterundersøgelse og effektanalyse	14
2.5	Longitudinal opfølgning	15
3	Praktiske forhold	15
3.1	Administration og organisation af deltagelse	15
3.2	Endelig fastlæggelse af forsøgsdesign	16
3.3	Sammenligningsskolerne.....	16
3.4	Afreportering af resultater.....	16
3.5	Formidling	17
3.6	Etik og datasikkerhed	17
3.7	Tidsplan	18
3.8	Bemanding	18
3.9	Budget	20
	Referencer	23
B	Foreløbige styrkeberegninger.....	24
B.1	Beslutningsvariable og parametre	24

Det er veldokumenteret, at tosprogede elever i gennemsnit klarer sig betydeligt dårligere end etnisk danske elever i det danske uddannelsessystem, Egelund & Rangvid (2005), Egelund et al. (2011) og Andersen & Rambøll (2010). Dette mønster ses også i mange andre af de lande, som vi normalt sammenligner os med, Kirsch et al. (2002). Selv andengenerationsindvandrere klarer sig betydeligt dårligere end etnisk danske elever, Nielsen & Rangvid (kommende). På trods af dette er der stadig mangel på viden om, hvordan vi bedst hjælper de tosprogede elever med at klare sig bedre i uddannelsessystemet. Nyere uddannelsesøkonomisk forskning peger på, at det generelt bedre kan betale sig at sætte ind med tidlige indsatser (fx i grundskolen) end med senere indsatser (fx ved overgang fra grundskole til ungdomsuddannelse), Heckman (2008).

For nylig har SFI og EVA lavet en kortlægning og analyse af eksisterende indsatser for tosprogede elever, Andersen et al. (2012). Denne analyse påpeger nogle interessante sammenhænge, fx at tosprogede elever på skoler med relativt få tosprogede, der tages ud af deres egen klasse for at modtage særlig undervisning, klarer sig dårligere end tosprogede, der ikke tages ud af klassen. Endvidere påpeges det i analysen at supplerende dansk som andetsprog for klasser med mere end 10 % tosprogede kan have positive effekter. Analysen kan dog ikke besvare de dybereliggende spørgsmål, som fx hvorvidt det er gavnligt for de tosprogede børn at modtage særlig undervisning udenfor klassen eller hvorvidt det er gavnligt at modtage ekstra timer i dansk som andetsprog, da selektionsproblematikken ikke håndteres eksplicit. I tråd hermed viser en rapport fra AKF, der bl.a. sammenligner erfaringerne med modersmålsundervisning i Danmark og Sverige, at der ikke kan drages entydige konklusioner om effekten af sådanne indsatser, Mehlbye et al. (2011). Dette skyldes især, at deltagelse i modersmålsundervisning er frivillig i de afdækkede indsatser.

Selv ved brug af avancerede statistiske metoder har det vist sig umådeligt svært at besvare denne type spørgsmål. Ved at implementere et randomiseret kontrolleret forsøg, som Undervisningsministeriet (UVM) foreslår, kan man dog komme uden om mange af de statistiske problemer, der er forbundet med en sådan evaluering. Når indsatserne, man ønsker at evaluere, er allokeret tilfældigt på skoler, kan man nemlig i princippet blot sammenligne resultaterne for eleverne i de skoler, der har modtaget indsatsen, med resultaterne for eleverne i de skoler, der ikke har modtaget indsatsen.

I dette projekt vil alle skoler med tosprogede elever i 1. klasse (i skoleåret 2014/15) have mulighed for at deltage. Blandt de tilmeldte vil nogle blive udtrukket tilfældigt til at modtage modersmålsundervisning rettet mod elevgrupper med tyrkisk, arabisk eller et tredje modersmål. De øvrige skoler vil fungere som sammenligningsskoler i skoleåret 2014/2015 og modtage modersmålsundervisning for deres elevgrupper i 1. klasse året efter (i skoleåret 2015/2016). Der er særligt fokus på modersmålsbaseret undervisning i tyrkisk og arabisk, men skoler, der selv kan stille en lærer til rådighed og udvikle deres egne materialer på et tredje sprog, kan også deltage i forsøget.

Dette forsøg vil således give skoler og kommuner langt bedre viden end de har i dag i forhold til, hvordan de kan hjælpe en af de elevgrupper, som i øjeblikket har sværest ved at klare sig i folkeskolen og få en ungdomsuddannelse. Dette forsøg vil give svar på, hvordan modersmålsundervisning påvirker elevernes faglige progression og socioemotionelle velfærd. Vi vil også foretage beregninger af de økonomiske gevinster og omkostninger ved forsøget, så indsatsernes effekt så vidt muligt kan sættes i forhold til deres forskellige omkostninger.

Det er en vanskelig opgave at designe, udføre og evaluere et så stort randomiseret kontrolleret forsøg. For at sikre at denne opgave løses bedst muligt har vi samlet en tværfaglig projektgruppe, der har erfaring

med design og evaluering af randomiserede kontrollerede forsøg på uddannelsesområdet, fx blandt andet i forbindelse med Forsøg med tolærerordninger og forsøg med modersmålsbaseret undervisning i 4. klasse. Projektgruppen består af følgende medlemmer (for nærmere oplysninger inklusive cv'er, se bilag B).

- [REDACTED]
- [REDACTED]
- [REDACTED]

Denne projektbeskrivelse redegør for, hvordan projektet foreslås implementeret, og hvordan de forskellige opgaver i denne forbindelse vil blive løst. Til sidst gennemgås en række emner af mere praktisk karakter. Eventuelle ændringer i forhold til denne projektbeskrivelse vil ske efter aftale med UVM.

2 PROJEKTET

Der er nedsat en arbejdsgruppe, der har til opgave at konkretisere og beskrive de enkelte forsøgsindsatser. Følgforskningen bistår Undervisningsministeriet med følgende primære aktiviteter: design af effekt- og implementeringsevaluering, dataindsamling, dataanalyse og afrapportering af resultater. Før vi opridses de forskellige delelementer af projektet, og hvordan de forskellige opgaver påtænkes håndteret, vil vi give et kort overblik over projektet, som vi ser det.

2.1 OVERBLIK

Forsøgsprogram om modersmålsundervisning er et ambitiøst randomiseret kontrolleret forsøg med flere forskellige indsatsgrupper. Dette betyder alt andet lige, at vi har behov for et stort antal deltagende skoler for at kunne gennemføre og evaluere forsøget. Alle folkeskoler, der på 1. klassetrin (i skoleåret 2014/15) har mindst en elev, der har et andet modersmål end dansk, men som har samme modersmål (fx tyrkisk) har mulighed for at deltage i forsøget.¹

Som en del af evalueringen af forsøgsprogrammet vil vi anvende registeroplysninger omkring de deltagende skoler og deres elever. Ved også at anvende disse registerdata i designfasen, kan vi få mulighed for at foretage et stratificeret forsøg. På baggrund af data fra tilmeldingssurvey samt fra registre om skolerne matches skoler/elevgrupperne to og to indenfor hver af de tre indsatsgrupper (tyrkisk, arabisk eller andre modersmålsgruppe). For at sikre, at sammenlignings- og indsatsgrupper ligner hinanden mest muligt i udgangspunktet og dermed øge den statistiske styrke i forsøget, vil vi anvende *optimal nonbipartite matching* til at matche skolerne. Den metode tager højde for flest mulige baggrundsfaktorer i udgangspunktet. Dette øger præcisionen af de beregnede effekter af forsøget. Efter stratificering trækkes der lod om, hvilken af de matchede skoler, der skal være i henholdsvis sammenlignings- og indsatsgruppe i skoleåret 2014/2015. De skoler, der skal være i sammenligningsgruppe skoleåret 2014/2015, vil skulle fungere som indsatsgruppe i skoleåret 2015/2016 og omvendt

¹ Det kan være problematisk at inddrage privatskoler og frie grundskoler i forsøget, idet fx privatskolerne har en tendens til at være meget forskellige, Andersen & Serritzlew (2007) og Rangvid (2008), hvor folkeskolerne trods alt udgør en mere homogen gruppe. Det kan gøre det svært at finde sammenlignelige privatskoler i indsats- og kontrolgrupperne.

Som nævnt indledningsvist vil dette forsøg kunne give svar på nogle helt centrale spørgsmål om, hvordan vi kan løfte det faglige niveau i den danske folkeskole generelt – og specielt hvordan man kan lukke præstationsgabets mellem elever med dansk som modersmål og tosprogede elever. Forsøget kan enten betragtes som ét samlet forsøg eller som et forsøg med tre forskellige indsatser. I sidstnævnte tilfælde, hvor målet er at undersøge effekten af de tre indsatser adskilt, er der – vurderet ud fra styrkeberegninger – behov for deltagelse af et meget stort antal skoler (mellem 70 og 100 i hver indsatsgruppe afhængig af elevgruppernes størrelse). Derfor foreslås det i denne løsningsbeskrivelse, at der gøres en stor indsats for at få flest mulig skoler til at tilmelde sig forsøget. Hvis der ikke er deltagende skoler nok, kan vi risikere, at der faktisk er positive effekter af de implementerede indsatser, men at vi ikke har skoler og elever nok til at kunne påvise disse effekter. Derfor giver vi her et overblik over nogle af vores løsningsforslag ift. antallet af skoler og diskuterer i de følgende afsnit, hvordan denne udfordring kan håndteres:

- Anvende gode førmålinger, der kan øge forklaringsgraden
- Anvendelse af avancerede matching strategier på basis af relevante variable fra spørgeskemaundersøgelser og registre, fx tidligere elevers faglige resultater og antal tosprogede elever
- Danne indsats- og sammenligningsgrupper på samme skole (dvs. fx en gruppe tyrkisk-sprogede elever modtager modersmålsundervisning mens en gruppe arabisk-sprogede elever fungerer som sammenligningsgruppe (for elevgrupper på andre skoler)

2.2 FORSØGSDESIGN

Inden vi går i nærmere detaljer omkring følgeforskningens resterende opgaver, vil vi kort diskutere relevante problemstillinger relateret til forsøgets design ift. at afklare indsatsens effektstørrelse. I den forbindelse er to faktorer centrale: 1) om indsatserne behandles separat eller samlet og 2) hvor mange elever, der er i hver indsatsgruppe. Hvis elevgrupperne er små (4 elever) vil det mindst kræve 100 elevgrupper at kunne udtale sig fornuftigt om effekterne af indsatsen jf. styrkeberegningerne i bilag 1. Dette er vel at mærke, hvis forsøget betragtes som én indsats. Skal indsatserne behandles separat vil det kræve ligeså mange elevgrupper i hver indsatsgruppe. Er elevgrupperne store (8 elever) vil det kræve mindst 70 elevgrupper.

Der er flere måder, hvorpå vi kan mindske behovet for skoler i forsøget, men de er ofte forbundet med mere eller mindre uheldige følgevirkninger. Det foretrukne scenarie er, at vi ikke har indsats- og kontrolgrupper på samme skole af hensyn til forsøgets validitet. På baggrund af antallet af tilmeldte skoler, vil vi dog tage stilling til, om det kan blive nødvendigt at tillade enten flere indsatser på samme skole, eller at en skole indgår som både sammenlignings- og som indsatsgruppe med forskellige elevgrupper. Eftersom indsatsen kun er målrettet afgrænsede elevgrupper, mindskes problemerne hermed, idet elevgrupper, der modtager modersmålsundervisning, vil have svært ved at påvirke andre elevgrupper, der har et andet modersmål. Således kan man overveje om flere indsatser kan implementeres på samme skole, fx at en elevgruppe af arabisk herkomst og en elevgruppe af tyrkisk herkomst fra den samme skole modtager modersmålsundervisning, eller den ene elevgruppe fungerer som sammenligningsgruppe for en anden skole.

2.3 DATAINDSAMLING

Førundersøgelsen, efterundersøgelsen og alle andre surveys relateret til forsøgsprogrammet foretages i dataindsamlingsredskaberne SurveyXact, Rambøll Results og Rambøll Sprog, der er specialudviklet af Rambøll.

For at mindske behovet for survey-oplysninger og dermed belastningen af skoler, lærere og elever vil de indhentede survey-data blive koblet med data fra Danmarks Statistiks registre, der fx indeholder oplysninger omkring elevernes forældre og dermed deres sociale baggrund. Med henblik på at koble survey-data til registerdata vil vi indsamle cpr-numre for alle deltagende elever og lærere og institutionsnumre og kommunekoder for alle deltagende skoler.

De indsamlede data fra surveys til skoleledere og lærere færdigbehandles af Rambøll, således at de kan indgå i en kvantitativ analyse af effekterne af de enkelte forsøgsindsatser på skolernes pædagogiske indsatser og undervisning.

Efter udtrækning af skoler, der skal deltage i forsøget, gennemføres i august 2014 en omfattende førundersøgelse af de pågældende skoler og de elever, som skal indgå i forsøget. En tilsvarende førundersøgelse gennemføres i august 2015. I forhold til design og analyse af forsøget tjener førundersøgelsen flere formål. Først og fremmest giver førundersøgelsen mulighed for at undersøge forholdene på skolerne og elevernes brug af sprog i hverdagen, færdighedsniveau, trivsel mm., før forsøget starter. I det omfang en førundersøgelse inkluderer de dimensioner, man er interesseret i at måle effekter på, er førundersøgelsen også nyttig i den forstand, at den øger den statistiske styrke og dermed reducerer behovet for skoler i forsøget betydeligt. Førundersøgelsen vil dels bestå af surveys til relevante lærere samt af sprogvurderingstests af de relevante elever på de udtrukne skoler. Generelt vil efterundersøgelsen indsamle data på de samme parametre som førundersøgelsen.

ELEVERNE

Vi vil i forbindelse med undersøgelsen af eleverne fokusere på følgende områder:

- Faglige kompetencer målt med sprogvurderingsværktøj
- Socioemotional velfærd (inklusive trivsel i skolen, selvværd, selvtillid)
- Motivation og arbejdsindsats

Jævnfør kravspecifikationen lægger vi vægt på at finde et mål af faglige kompetencer, der både måler sproglig refleksion og kommunikative kompetencer i overensstemmelse med faghæfte 46 om fælles mål for modersmålsundervisning. Vi skal dog gøre opmærksom på, at fagbeskrivelsen i høj grad lægger vægt på elevernes kompetencer på modersmålet, mens vi i vores målinger fokuserer på elevernes danskfaglige kompetencer.

I forhold til at måle elevernes faglige kompetencer er det en udfordring, at elever, der starter i 1. klasse ikke med sikkerhed kan læse. Det er derfor ikke muligt at anvende gængse læsetests til at måle elevernes kompetencer. De eksisterende sprogtests er dog i vid udstrækning baseret på at screene børn med sprogvanskeligheder. Det kan derfor være, at mange børn i den bedre halvdel scorer maksimumpoint i testen. Sådanne "lofteffekter" betyder, at det kan være umuligt at måle, hvis disse børn har haft en positiv effekt af forsøget. Problemet bliver dog mindre af, at tosprogede børn og særligt dem med ikke-vestlige forældre typisk klarer sig dårligere i disse tests end børn med danske forældre. Udover at de faglige kompetencer skal måles sprogligt og læserelateret, vil det være en stor fordel rent praktisk, hvis testen kan aflægges elektronisk, ligesom det er centralt, at testen er valideret i en dansk kontekst. Ud fra disse kriterier er sprogvurderingsværktøjet udviklet af Dorthe Bleses det bedst mulige redskab og samtidig et af de mest udbredte på området, selvom testen er målrettet elever i børnehaveklassen.² Der vil således være

² Af andre sprogvurderingstests kan nævnes "Sprogvurdering" af Heinze, 2012, der imidlertid er væsentligt mere tidskrævende. Derudover er "Bogstaveprøve" af Møller & Juul, 2009 og "Children, Communication Check List" af Bishop, 2012 andre mulige alternativer, der dog begge også vil være plaget af lofteffekter.

risiko for lofteffekter i forhold til at differentiere mellem eleverne - især i slutningen af 1. klasse, men det faktum, at vi primært skal måle på effekterne for tosprogede gør formentlig problemet mindre. Således viser en tidligere anvendelse af sprogvurderingsværktøjet på børnehaveklasebørn, at knap 60 % af de ikke-vestlige tosprogede børn har så store sprogproblemer, at de skal have en særlig indsats, Højen & Bleses, 2012: 10. Vi har ikke kunne få tal for, hvor mange af de resterende godt 40 %, der rammer loftet i testen, men det vil formentlig kun være en mindre del af dem.

Med henblik på at teste elevernes evner på modersmålet vil VIA University College udvikle en dialogisk sprogtest, der kan kortlægge modersmålet hos minoritetsbørn i forsøget, herunder elevens samtalekompetence, hverdagslige ordforråd og skolerelevante ordforråd. Sprogtesten kan enten nyudvikles eller kan være en bearbejdning af eksisterende tests til formålet, under hensyntagen til rettigheder. VIA vil endvidere gennemføre sprogtesten med brug af lokale konsulenter

Udover sprogvurderingsværktøjet vil vi foreslå, at de deltagende elever på skolerne skal tage testen i dansk, læsning for 2. klasse som en frivillig test i december i 1. klasse. Det kan ikke fungere som førmåling, da indsatsen på det tidspunkt har været i gang. Men det kan bruges som en måling af, om indsatsen har haft effekt allerede efter nogle måneders forløb. Derudover er det obligatorisk, at eleverne skal tage testen i foråret i 2. klasse. I senere analyser vil vi kunne bruge den test til at måle eventuelle langtidseffekter af indsatserne.³ På den baggrund foreslår vi følgende plan for måling af dansksproglige kompetencer:

E2014	E2014	F2015	E2015	F2016
Sprogvurdering	Læsning, 2. klasse (frivillig test)	Sprogvurdering	Læsning, 2. klasse (frivillig test)	Læsning 2. klasse

Det skal bemærkes, at det kræves ekstra af de deltagende skoler, at de forpligter sig til at tage to test i løbet af de frivillige testperioder i efteråret 2014 og efteråret 2015.⁴ Desuden skal de samme målinger foretages for de elevgrupper, der starter i 1. klasse i efteråret 2015.

I forhold til at måle elevernes socioemotionelle velfærd har en række kriterier været afgørende for valget af målemetode. Igen har målgruppens alder haft betydning for valget af test, idet det ikke er realistisk at forvente, at elever i 1. klasse meningsfuldt kan vurdere deres egen socioemotionelle situation. Det har derfor for det første været centralt at finde et instrument, der kan udfyldes af elevernes lærere. Det betyder at vi kan forvente at få mere valide svar. For det andet har det været centralt at anvende et mål, der er valideret i en dansk kontekst således at spørgsmål, svar kategorier og skalaer er forskningsfunderede. På baggrund af disse kriterier er CBCL (Child Behavior Checklist) og SDQ (Strengths and Difficulties Questionnaire) relevante tests. I valget mellem disse to er SDQ kortere, og har i tidligere studier vist sig at være ligeså effektiv som CBCL, Goodman & Scott (1999); Klasen et al.(2000).

³ Det vil være ideelt, hvis de deltagende elever kan få mulighed for at aflægge den nationale test i dansk, læsning for 2. klasse i slutningen af eller lige efter indsatsperioden, dvs. i juni 2015 i slutningen af 1. klasse, eller begyndelsen af 2. klasse i august/september 2015. Vi anbefaler, at UVM undersøger muligheden for dette.

⁴ De præcise tidspunkter for testperioderne i skoleåret 2014/15 er endnu ikke offentliggjort, så dette bygger på en antagelse om, at testperioderne er mindst lige så lange som i 2013/14. Det anbefales, at den frivillige testperiode i efteråret 2014 forlænges frem til den sidste uge inden jul

Da både sprog vurderingen og vurderingen af socioemotionel velfærd uanset måleinstrument skal udføres af en lærer, er det vigtigt for at kunne måle effekterne af indsatsen, at der afsættes ressourcer til at gennemføre disse målinger på både indsats- og sammenligningsskoler.

Det er endnu ikke fastlagt hvilke spørgsmål, der skal bruges til at måle elevernes motivation og arbejdsindsats. Vi foreslår at anvende en kombination af spørgsmål anvendt i andre danske undersøgelser på skoleområdet omskrevet til lærervurderinger.

I overgangen fra børnehaveklassen (0. klasse) til 1. klasse skifter eleverne som udgangspunkt lærer. Ved før-målingens gennemførelse i august vil den nye lærer forventeligt have et ringere udgangspunkt for at vurdere elevernes sproglige kompetencer samt socioemotionelle velfærd end den gamle lærer. Vi foreslår således, at børnehaveklasselæreren besvarer førmålingen i august, mens eftermålingen besvares af den nye 1. klasselærer.

Rent praktisk vil dataindsamlingen foregå ved, at der sendes samlede forsendelser per skole med kuverter adresseret til børnehaveklasselæreren ved førmålingen samt 1. klasselæreren ved eftermålingen. Forsendelsen vil indeholde logins for de elever, der enten modtager modersmålsundervisning eller er i sammenligningsgruppen. Login vil give adgang til både at udføre sprogvurdering samt trivselsmåling for eleven. Således minimerer vi besværet for den enkelte lærer mest muligt. Derudover vil der være vedlagt et følgebrev, der fortæller om formålet med dataindsamlingen samt anfører en udførlig vejledning i forhold til hvordan læreren besvarer surveyene.

LÆRERNE

Det foreslås at gennemføre en spørgeskemaundersøgelse blandt lærere på 1. klassetrin både før og efter indsatsens implementering. Formålet med at lave en spørgeskemaundersøgelse til lærerne er todelt: For det første vil det undersøges om det implementeringsmiljø, som indsatsen gennemføres i, har betydning for indsatsens effekter. For det andet vil det undersøges, hvorvidt gennemførelsen af indsatsen med modersmålsundervisning har betydning for skolens bredere praksis.

Således vil spørgeskemaundersøgelsen til lærerne komme ind på den pædagogiske indsats i praksis (udvikling, planlægning, tilrettelæggelse og evaluering af undervisning, teamsamarbejde på klassetrin, fagfelt eller lignende), pædagogiske indsatser for tosprogede elever (dansk som andetsprog, andre tilgange til inklusion af tosprogede, forældreinddragelse) samt lærernes egne kompetencer i forhold til undervisning af tosprogede.

Spørgeskemaet skal besvares af lærere både på indsats- og sammenligningsskoler. Vi foreslår, at skemaet skal besvares af 1 dansk- og 1 engelsklærer per klasse på 1. klassetrin. Det er vores vurdering, at det især er relevant at stille spørgsmålene til sproglærerne, og at det vil være disse lærere, der især vil samarbejde med modersmålsunderviseren.

Ud fra erfaringer fra tidligere randomiserede forsøg foreslår vi at anvende lærerne som primær kilde til at afdække skolens tilgang til undervisning af tosprogede elever. Det er ikke vores indtryk, at vi får tilstrækkelig viden ved at gennemføre en selvstændig spørgeskemaundersøgelse til skolelederne. Yderligere viser andre studie på skoleområdet, at lærernes vurdering af ledelsen har større sammenhæng med elevernes resultater, end ledelsens egne vurderinger af samme. Vi vil dog inddrage ledelsen, når vi gennemfører casestudier (se afsnit X).

For at lette skolernes regnskabsafklæggelse i forbindelse med de midler, de modtager fra UVM til gennemførelse af forsøget, foreslås det, at lærerne i spørgeskemaet efter indsatsens afslutning skal afkrydse, om

de har brugt midlerne som forudsat i UVM's bevillingsbrev, og eventuelt hvordan de har brugt dem. Den præcise formulering leveres af UVM, så det sikres, at det lever op til regnskabskravene. Dog forudsættes det, at det er spørgsmålsformuleringer, som kan besvares med enkelte afkrydsninger. Da lærerne skal krydse afkrydse ressourceanvendelsesspørgsmål på skolens vegne foreslås det, at UVM allerede i bevillingsbrevet gør skolerne opmærksomme på dette krav til regnskabsaflægningen.

IMPLEMENTERING

Det er selvsagt ikke kun designet af et randomiseret kontrolleret forsøg, der har betydning for, hvorvidt forsøget bliver en succes i sidste ende. En forudsætning for et succesfuldt forsøg er, at skolerne gennemfører den indsats de er udtrykt til samt at alle skoler deltager i den dataindsamling, der gør, at vi kan undersøge om forsøget har haft en effekt. En vigtig brik i dette er en effektiv og målrettet kommunikation med de deltagende skoler. Vi vil beskrive dette aspekt senere. Derudover anser vi det som afgørende viden at blive klogere på om *måden* indsatsen implementeres på har betydning for indsatsens effekter samt mere grundlæggende at komme nærmere på *hvorfor* en given indsats har effekt. Det er derfor følgeforskningen tilbyder at gennemføre både fidelitetsspørgeskemaundersøgelse samt casestudier, som beskrives i de kommende afsnit.

Støtte til skoler og kommuner med implementering

I det følgende beskrives vores aktiviteter i forhold til information, som har det overordnede formål at klæde skolerne på til at gennemføre indsatsen med størst grad af fidelitet samt til at gennemføre dataindsamlingsaktiviteter med størst mulig kravopfyldelse. Det anses som ekstremt vigtigt, at aktørerne i forsøget gives den information og vejledning, de har behov for. Hvis skolerne ikke bliver klædt ordentligt på til at gennemføre de aktiviteter, der er foreskrevet i forsøgsprogrammet, vil effektmålingerne ikke reflektere effekten af de enkelte indsatser, men effekten af de justerede indsatser, som skolerne praktiserer. Det er altså vigtigt, at de relevante aktører implementerer forsøgsprogrammet med *fidelitet* for at opnå brugbar viden omkring hvilke indsatser, der har hvilke effekter.

På baggrund af blandt andet erfaringer fra forsøget med tolærerordninger samt de tidligere forsøg med modersmålsbaseret undervisning finder vi det centralt at sikre metodisk robust implementering og dermed for at opnå valide og brugbare forskningsresultater at holde fokus på følgende faktorer:

Først og fremmest er den indledende information til skoler omkring deres roller og opgaver i forsøget vigtig. Det er herigennem, at vi sikrer os, at skolerne er gearet til at implementere den pågældende indsats som ønsket. Indsatserne vil generelt kræve, at skolerne prioriterer de nødvendige ressourcer og ledelseskræfter. Det er vigtigt, at skolerne er opmærksomme på dette på et tidligt tidspunkt, så de har tid til at træffe de nødvendige foranstaltninger. Erfaringen fra tidligere forsøg er, at det er centralt at kommuner og skoler i god tid modtager skriftlig orientering om deres deltagelse i forsøget. Den skriftlige kommunikation skal på den ene side være kort og præcis i sit indhold. På den anden side skal den sørge for at skoler og kommuner får fyldestgørende information om kravene til deres deltagelse, tidsplan og hvor de kan få svar på eventuelle udestående spørgsmål. Dernæst kan det være formålstjenstligt at afholde informationsmøder med personlig kontakt til skolerne, således at der skabes tillid gennem dialog, ligesom afklaring af mange spørgsmål kan være nemmere her end via udelukkende skriftlig kommunikation. Der vil blive holdt informationsmøder med deltagende kommuner og skoler.

Udover den indledende information er det vigtigt at bibeholde kontakten med skolerne under selve forsøget for at sikre, at de fortsat har fokus på, hvordan den pågældende indsats ønskes implementeret. Det

kan således ellers være et problem, at skolerne mister interessen for forsøget undervejs og derfor slækker på implementeringen. Der kan også opstå tvivlsspørgsmål på de enkelte skoler undervejs, hvilket i de nuværende og foregående forsøg med modersmålsbaseret undervisning har vist sig ved, at den af Rambøll oprettede hotline har modtaget en stor mængde henvendelser fra skoler, der har haft et behov for at spørge til specifikke forhold i forsøget og deres egen rolle i forsøget. På baggrund af henvendelserne har det været muligt at yde ekstra støtte til skolerne i en solid implementering. En sådan hotline forventes derfor også at være yderst anvendelig i nærværende forsøg.

Det er stadig relativt ny praksis at gennemføre forsøgs- og udviklingsarbejde på skoleområdet i Danmark ved hjælp af randomiserede kontrollerede forsøg. Derfor er det vigtigt at indhente viden omkring den praktiske implementering af forsøgsindsatserne, som vil finde sted på skolerne, for at få et robust vidensgrundlag om den gennemførte indsats. I det følgende beskrives hvordan en del af denne viden blive indhentet via spørgeskemaundersøgelserne om fidelitet, samt hvordan casestudierne kan bidrage hertil.

Fidelitetsspørgeskemaundersøgelse

Indsamling af viden omkring den praktiske implementering af de forskellige elementer af forsøget er afgørende af to årsager. For det første er det interessant at få et indblik i hvordan skolerne og underviserne har grebet modersmålsundervisningen an. Denne viden gør det nemlig muligt at sige, om skolerne lever op til kravene til indsatsene og med hvilken grad af efterlevelse. For det andet er det centralt at få undersøgt om måden undervisningen gennemføres på, har betydning for effekterne for elevernes faglige og trivselsmæssige udbytte. Således anbefales det at gennemføre en fidelitetssurvey blandt modersmålsunderviserne, hvor de bedes registrere aktiviteter og output i et elektronisk skema samt besvare en række uddybende spørgsmål. Spørgeskemaundersøgelsen vil gennemføres to gange i løbet af forsøget.

Som udgangspunkt mener vi at følgende elementer bør indgå i spørgeskemaundersøgelsen:

- Modersmålsunderviserens formelle kompetencer (uddannelse, linjefag, ect.) samt mere uformelle kompetencer (undervisnings- og/eller pædagogisk erfaring)
- Deltager modersmålsunderviserne i det intenderede kursus? Vurderer de, at deres didaktiske kompetencer er forøget med kurset?
- Er underviserne motiverede for at applicere de tilegnede kompetencer i deres undervisning og aktiverer de deres tilegnede kompetencer i undervisningen?
- Læser og forstår underviserne vejledningsmaterialet og anvender de det?
- Varighed/intensitet af indsats (antal lektioner, lærer/elev ratio i indsatsen, elevs fravær)
- Gennemfører underviserne undervisningen som intenderet (anvendelse af vejledning, undervisningsmateriale, anvendelse af modersmål)
- Samarbejder underviserne med elevernes andre lærere på 1. klassetrin, således at den viden der opnås om elevernes sprogkompetencer i modersmålsundervisningen såvel i de andre fag kan udnyttes til at styrke elevens samlede sprogkompetencer.
- Inddrager underviserne elevernes forældre i gennemførelsen eller forberedelsen af undervisningen og i så fald hvordan?
- Deltager eleverne aktivt som intenderet (fremmøde, aktiv deltagelse i undervisningen, evt. laver lektier)

Undersøgelserne af skolernes implementering af de enkelte indsats vil være vigtige i forhold til at fortolke resultaterne af effektanalysen. Vi søger således muligheden for, at forsøgsprogrammet vil kunne bidra-

ge med viden om, *hvorfor* de enkelte indsatser virker, og ikke kun *hvilke* indsatser der virker. Endelig er implementeringsforskningen vigtig for at opnå erfaringer, der kan bruges ved design og implementering af fremtidige randomiserede kontrollerede forsøg. Den opnåede erfaring og viden i forbindelse med implementeringen af forsøgene vil blive beskrevet og analyseret i forbindelse med afrapporteringen af forsøgene.

Casestudier

Vi foreslår, at der udover den beskrevne fidelitetssurvey gennemføres ni casestudier på udvalgte skoler.

Formålet med casestudierne er primært at understøtte effektanalyserne ved at formulere hypoteser om mekanismerne hvorved modersmålsundervisning fremmer børns læring og trivsel. Et sekundært bidrag af casestudierne er at beskrive og tolke skolers og kommuners erfaringer med at implementere forsøget.

Fokus for casestudierne
<ul style="list-style-type: none">- Mekanismer: Hvordan styrker modersmålsundervisning elevernes faglighed og trivsel?- Kontekst: Hvilke øvrige faktorer påvirker udmøntningen af indsatserne?- Skalering: Hvordan kan kommuner og skoler bedst organisere modersmålsundervisning?

Mekanismer

Det randomiserede forsøgsdesign vil primært give svar på, hvorvidt modersmålsundervisning har en positiv påvirkning på elevernes sproglige kompetencer og deres socioemotionelle velfærd. Dette er et væsentligt bidrag, idet eksisterende forskning ikke har entydige svar på netop dette. Effektevalueringen vil således give svar på om og i givet fald hvor meget effekt indsatserne har. Uden supplerende analyser, vil vi dog stadig mangle svar på, *hvorfor* modersmålsundervisning har en eventuel effekt. Således foreslår vi at lave en teoribaseret implementeringsevaluering med særlig fokus på de mekanismer, der skaber effekten. Ved at gennemgå eksisterende forskning og undersøgelser formuleres hypoteser om de mekanismer, der gør, at modersmålsundervisning kan styrke elevernes faglige progression samt trivsel og motivation. Disse hypoteser oversættes igen til observerbare implikationer som inspirerer udarbejdelsen af interviewguides og observationsguide. Hypoteserne vil eksempelvis kunne fx målrettes modersmålsunderviserens didaktiske strategier (herunder undervisningsdifferentiering), hvordan og på hvilken måde forældrene involveres, og hvordan temaer og begreber fra de øvrige fag integreres i modersmålsundervisningen. Vi vil også i forbindelse med mekanismerne have blik for, hvorvidt undervisningen er gennemført i overensstemmelse med inspirations- og vejledningsmateriale (fidelitet). Resultaterne af casestudierne vil blive anvendt til at blive klogere på, hvilke hypoteser der er mest relevante at forfølge i den endelige spørgeskemaundersøgelse til både elevernes eksisterende lærere samt modersmålsunderviseren. Således vil casestudierne medvirke til at kvalificere de endelige effektanalyser, hvor surveydata inddrages til at undersøge betydningen af implementeringsfaktorer.

Kontekst

Derudover vil casestudierne fokusere på den kontekst, som indsætterne gennemføres i. En nyere undersøgelse fra AKF (2012) viser, at tosprogede elever i Danmark kan forbedre deres læseforståelse ved at deltage i modermålsundervisning, men at effekten af modermålsundervisning ikke er entydig. Det tolker forskerne således, at rammerne og vilkårene for modermålsundervisningen har stor betydning for effekten. Casestudierne vil derfor fokusere på, udvikle hypoteser om og undersøge, hvordan rammerne påvirker udmøntningen af indsatsen – og kan derigennem tilvejebringe nyttig viden, der kan bidrage til forklaringer af effekten. Undersøgelsen af skolens kontekst vil derfor fokusere på ledelsens prioritering af eksempelvis dansk som andetsprog, og hvordan modermålsunderviserne organisatorisk er forankret på skolen. Det vil også være en relevant kontekstfaktor at undersøge, hvilke rationaler der ligger bag placeringen af den supplerende undervisning – om det udelukkende er praktiske hensyn til eksempelvis elevernes skema, eller om også pædagogiske hensyn ift. undervisningens tilrettelæggelse og gennemførelse spiller en rolle.

Skalering

Som lovgivningen er i dag, skal kommunerne tilbyde børn fra EU lande (samt Færøerne og Grønland) undervisning i deres modersmål. Kommunerne er ikke forpligtet til at tilbyde modermålsundervisning til børn fra andre lande, men kan hvis de ønsker det. Undervisningen skal så vidt muligt foregå på elevens skole, men kan pga. elevtal, lærerkapacitet e.l. foregå på en anden skole i kommunen. Undervisningen gives i særskilte timer og så vidt muligt på aldersinddelte hold i tre til fem ugentlige timer, og i fortsættelse af elevernes almindelige skoletid. Med forsøget for 1. klassetrin vil en række kommuner få ny erfaring med at gennemføre modermålsundervisning. Disse erfaringer er værdifulde at opsamle for de deltagende kommuner samt andre kommuner, der måtte ønske at gennemføre undervisning for deres elever, der har et andet modersmål end dansk. I den forbindelse bør det afdækkes, hvordan kommunen og skolerne har rekrutteret en modermålsunderviser, hvordan de rent praktisk har håndteret at give undervisningen som supplerende undervisning udenfor normal skoletid osv.

Hvilke kriterier der skal være for udvælgelse af skoler afhænger af det endelige formål og fokus for casestudierne. Skal casestudierne belyse hvordan kommuner bedst kan organisere og gennemføre modermålsundervisning, kan det være hensigtsmæssigt at sørge for at besøge kommuner og skoler, der adskiller sig mest muligt fra hinanden for på den vis at kunne tilpasse rådene forskellige kommuners vilkår. Skal casestudierne anvendes til at teste hypoteser om mekanismerne for modermålsundervisning, kunne det være hensigtsmæssigt at holde en række faktorer konstante og således besøge skoler, der ligner hinanden på mange parametre. Formålet med casestudierne og således udvælgelseskriterierne fastlægges dog endeligt i dialog med kunden.

Rent praktisk foreslår vi, at casestudierne afvikles fx i starten af december 2014 eller i slutningen af januar 2015. Således vil kommuner og skoler på dette tidspunkt have gjort sig nogle erfaringer med at gennemføre indsætterne. Ligeledes er der mulighed for at fundene kan styrke forsøgets resterende løbetid bl.a. ved at kvalificere efter-surveys til lærere. Casestudierne kan enten fokusere på flere udvalgte skoler eller alternativt udvalgte kommuner. Vælges sidstnævnte fremgangsmåde gennemføres interviews med lærere og ledere fra flere af kommunens deltagende skoler. Dette kan give et bredere indblik i praksis og erfaringer mens casestudier på enkeltskoler, vil give et mere dybdegående indblik i enkelte skoler. Uanset om

en skole eller kommune tilgang udvælges vil observation og elevinterview kun finde sted på én udvalgt skole (en per skole eller en per kommune).

Nedenstående tabel illustrerer, hvilke aktører vi foreslår at interviewe.

Estimeret varighed	Aktivitet
8.00-9.00	Interview med kommunal kontaktperson
9.00 - 10.00	Evt. interview med deltagende elever (3-6 elever afhængigt af gruppens størrelse)
10.00 - 11.30	Interview med dansk- og engelsklærere til de deltagende elever (1 lærer per involverede 1. klasser. Evt. mulighed for fokusgruppeinterviews med lærere fra flere af kommunens deltagende skoler)
12.00 - 13.00	Interview med repræsentanter fra skoleledelsen med kendskab til implementeringen af indsatsen (mulighed for interviews med flere repræsentanter for ledere på kommunens deltagende skoler)
13.00 - 14.30	Interview med underviser i Almen Sprogforståelse (Evt. mulighed for fokusgruppeinterviews med undervisere fra flere af kommunens deltagende skoler)
14.30 - 15.30	Observation af modersmålsundervisning (1 lektion)
15.30-16.30 senere)	(eller Fokusgruppeinterview med forældre til elever, der deltager i modersmålsundervisning med deltagelse af tolk (Evt. mulighed for fokusgruppeinterviews med forældre fra flere af kommunens deltagende skoler)

2.4 EFTERUNDERSØGELSE OG EFFEKTANALYSE

Efter forsøget er gennemført, og der er blevet indsamlet data fra førundersøgelsen, undersøgelser i forsøgsperioden og efterundersøgelsen, vil effekterne af de forskellige indsatser blive estimeret. Analysen vil omfatte effekter på samtlige af de målte udfald, herunder faglige resultater, trivsel, motivation og arbejdsindsats. Det vil endvidere blive undersøgt, hvorvidt elever med forskellig social baggrund påvirkes forskelligt af indsatserne. Påvirkes elever fra en svag social baggrund fx mere af indsatserne end elever med en stærk social baggrund? Fra Danmarks Statistiks registerdata har vi detaljerede oplysninger omkring forældrenes uddannelse, arbejdsmarkedsstatus osv. I det omfang det er muligt, vil vi også undersøge, om effekterne af indsatserne er heterogene i andre dimensioner, fx om indsatserne påvirker drenge og piger forskelligt, eller om indsatserne påvirker fagligt svage elever anderledes end fagligt stærke elever. Vi vil også undersøge, om karakteristika ved elevernes lærere i andre relevante fag har betydning for, hvor stort

udbytte de enkelte elever får af indsatserne. Det vil også være relevant at se på, om effekterne af indsatserne varierer i forhold til andelen af tosprogede i klassen. En del tidligere danske undersøgelser peger på, at andelen af tosprogede på en skole har en betydning for elevernes faglige niveau, og hvordan indsatserne for tosprogede virker, se henholdsvis Jensen & Rasmussen (2011) og Andersen et al. (2012)..

UVM ønsker at få afdækket, om den måde, hvorpå skolerne implementerer indsatserne, har betydning for resultaterne. Eftersom det ikke er tilfældigt, hvordan skolerne *vælger* at implementere en given indsats, vil det dog være svært at sige noget endegyldigt på dette punkt. Dette understreger vigtigheden af undervejs i forsøget at forsøge at sikre en høj grad af fidelitet i implementeringen i forhold til at producere generaliserbare resultater. Det vil være en vigtig del af den endelige analyse at se på forskelle i implementering på tværs af skoler og karakteristika ved forskellige typer skoler, der vælger forskellige implementeringsformer (fx ledelsesopmærksomhed, lærerteamsamarbejde, lærerkvalifikationsstrategier mv.), da dette er vigtigt i forhold til fortolkning og generaliserbarhed af undersøgelsens resultater.

Effektanalyserne kan afrapporteres i flere trin for at få præsenteret nogle hovedresultater hurtigst muligt efter indsatsens ophør, samtidig med at nogle data først vil blive tilgængelige senere ligesom nogle analyser vil tage længere tid at gennemføre. Effektanalyserne vil således først og fremmest undersøge effekter på elevernes læsning, trivsel og motivation. I næste trin vil det også blive søgt afdækket, om og eventuelt hvordan indsatserne har påvirket skolernes pædagogiske indsatser og undervisning generelt, og om indsatsernes effekt afhænger af forhold på skolerne. På længere sigt kan langtidseffekter på nationale test i 2. klasse undersøges.

2.5 LONGITUDINAL OPFØLGNING

Med henblik på at se på de langsigtede effekter af indsatserne vil det være interessant at følge de medvirkende skoler og elever over tid. Dette kan gøres relativt nemt og billigt ved brug af Danmarks Statistiks registerdata. Vi vil blandt andet koble data fra forsøget med data omkring senere uddannelsesforløb (senere nationale test, deltagelse i specialundervisning, karakterer, gymnasial uddannelse, erhvervsfaglig uddannelse osv.) og beskæftigelsesforløb (løn, erfaring, sociale ydelser med mere). På den måde kan vi se, om indsatserne påvirker eleverne senere i livet, fx deres færd gennem uddannelsessystemet, og hvordan de klarer sig på arbejdsmarkedet. Det er selvfølgelig interessant i sig selv, om indsatserne har umiddelbar effekt på elevernes faglige formåen som målt ved de nationale test, men i sidste ende er det væsentligt, om en eventuel positiv effekt af indsatserne på de faglige resultater også materialiserer sig i en større sandsynlighed for fx at få en uddannelse ud over folkeskolen (jf. 95-procent-målsætningen) og at komme i beskæftigelse. Der går dog selvsagt en del tid, før vi vil være i stand til at besvare sådanne spørgsmål, så i første omgang vil den longitudinale opfølgning primært bestå i at undersøge, om indsatserne påvirker resultaterne i senere nationale test, fx de obligatoriske test i læsning på 2. klasses trin, altså ca. 1 år efter forsøgets afslutning. Da vi allerede har nogen information omkring mere generelle sammenhænge mellem tidlige mål og senere fagligt niveau, ungdomsuddannelse osv., kan dette også være med til at give os en ide om indsatsernes langsigtede effekter.

3 PRAKTISKE FORHOLD

Der er en række forhold af mere praktisk karakter, som vi kort gør rede for i det følgende.

3.1 ADMINISTRATION OG ORGANISATION AF DELTAGELSE

UVM står for at indgå samarbejdsaftaler med de deltagende kommuner. Som tidligere nævnt vil det sandsynligvis være nødvendigt at relativt mange skoler deltager i forsøget. Det er derfor rigtig vigtigt, at forsøget har stor opbakning. Samtidig er det vigtigt, at de, som deltager i forsøget, deltager i de dataindsamlinger, som er krævet. Erfaringer fra Forsøg med tolærerordninger samt Forsøgsprogrammet for modersmålsundervisning på 4. klassetrin viser, at det er en stor fordel, at aftaler indgås mellem ministeriet og kommuner, så kommunerne kan hjælpe med at sikre at skolerne overholder aftalerne.

Rambøll står efterfølgende for den primære kommunikation med de deltagende kommuner, for selve administrationen af kommunernes og skolernes tilmelding til forsøget. Dette har vi erfaring med fra de nævnte forsøg. AU og Rambøll gennemfører en kommunikationsstrategi overfor deltagende kommuner og skoler, således at de udviklede indsatser kommunikeres effektivt. Desuden sikres det, at alle kommuner og skoler forstår forsøgets metode, og hvad der forventes af dem i forbindelse med implementering for at sikre fidelitet.

Endvidere understøttes kommunernes og skolernes implementering af forsøgsindsatserne af en hotline, der etableres af Rambøll. Dette giver mulighed for, at kommuner og skoler både per telefon og e-mail dagligt kan henvende sig med spørgsmål. Derudover vil vi følge op på skoler, hvor deltagelsen i surveys ikke er tilstrækkelig høj, eller hvor monitorering af skolernes registrering af aktiviteter signalerer problemer.

Endelig gennemfører Rambøll en ny tilmeldings- og registreringsrunde i foråret 2015, når forsøget skal køre i anden runde for skoleåret 2015/16.

3.2 ENDELIG FASTLÆGGELSE AF FORSØGSDESIGN

Som tidligere beskrevet er det en reel mulighed, at det ikke vil være praktisk muligt at gennemføre samtlige indsatser beskrevet i udbudsmaterialet. Kortlægningen vil afdække dette i detaljer. Derudover er det potentielt ikke – ud fra det opstillede forsøgsprogram – muligt at besvare samtlige spørgsmål som formuleret i udbudsmaterialet. Vi lægger derfor op til, at endelig stillingtagen til forsøgets design fastlægges på baggrund af de tilmeldinger, der kommer.

3.3 SAMMENLIGNINGSSKOLERNE

For at kunne afdække indsatsernes faktiske effekter, er det centralt, at de samme målinger gennemføres for såvel indsats- som sammenligningsskoler. Det er derfor centralt, at der sættes midler af til at foretage de beskrevne tests både for indsats- og sammenligningsskoler. Særligt skal der afsættes ressourcer til at gennemføre test af sprogvurderinger og socioemotional velfærd, da disse tests skal foretages af en lærer på både indsats- og sammenligningsskoler.

3.4 AFRAPPORTERING AF RESULTATER

Der vil i efteråret 2016 blive udfærdiget en samlet afrapportering af resultaterne af indsatserne. Derudover vil vi holde UVM orienteret omkring relevante resultater fra den longitudinale opfølgning. Rapporter vil indeholde en grundig beskrivelse og analyse af det implementerede forsøg. Dette dækker over blandt andet metode (herunder design og randomisering), teori, dataindsamling, implementering, empiri, analyse (herunder den faktiske effektanalyse, hvor vi betragter forskellige udfaldsmål: dels faglige resultater men også socioemotional velfærd og trivsel mm.) og følgeforskernes konklusioner. Det vil blive grundigt beskrevet, hvordan de enkelte indsatser rent faktisk er blevet implementeret blandt andet ved hjælp af indsamlede data fra monitorering af skoler og casebesøg på udvalgte skoler.

Som nævnt vil afrapporteringen ske i flere trin, så hovedresultater kan blive afrapporteret primo september, mens yderligere implementeringsanalyser samt analyser af mere langsigtede analyser vil følge senere.

3.5 FORMIDLING

Såfremt UVM ønsker at afholde en konference eller på anden vis formidle resultaterne af forsøget, står vi naturligvis til rådighed. Værdien af et storstilet forsøgsprogram som dette består i den viden, der genereres, og det er derfor vigtigt at videreformidle denne viden til et bredere publikum. Specielt vil det være relevant både for de involverede skoler og andre skoler at blive bekendt med resultaterne fra forsøget, således at skoleledere og lærere kan lære af erfaringerne fra forsøget. Hvis de enkelte indsatser viser sig at være forbundet med positive effekter vil det være centralt at få beskrevet indsatserne i detaljer (eventuelle kurser, opkvalificeringsforløb mv.), således at alle skoler i princippet har mulighed for at implementere dem.

Vores formidling af projektets resultater består primært af analyserapporten. Derudover planlægger vi at formidle resultaterne af projektet i førende, internationale tidsskrifter inden for vore fagområder. Denne formidling vil selvfølgelig finde sted *efter* offentliggørelsen af de relevante projektrapporter.

3.6 ETIK OG DATASIKKERHED

Forskningsprojekter og specielt projekter, der involverer randomisering af enheder, er pålagt nogle etiske hensyn. Dette projekt vil følge de "Vejledende Retningslinier for Forskningsetik i Samfundsvidenskaberne" som formuleret af det tidligere Statens Samfundsvidenskabelige Forskningsråd. Dette indebærer blandt andet, at der tages hensyn til at de deltagende skoler og elever ikke belastes unødigt i forbindelse med forsøget. Eftersom der er tale om et randomiseret forsøg, kan vi imidlertid ikke behandle alle deltagende skoler og elever ligeligt. Derudover er der en række hensyn, der skal tages, når man arbejder med følsomme oplysninger på individniveau. I overensstemmelse med gældende regler i persondataloven er projektet anmeldt til Datatilsynet.

3.7 TIDSPLAN

For at give et overblik over projektets forløb fra de indledende undersøgelser til den longitudinale opfølgning viser vi her vores foreløbige tidsplan for projektet.

2014

Uge 15-18	Tilmeldingsperiode for kommuner
Uge 19	Randomisering
Uge 20	Besked til deltagende skoler
Juni	Skoler opretter elever og lærere i Rambøll Results
August	Førundersøgelse lærere, elever
Uge 39	Indsats igangsættes
Oktober-november	Fidelitetssurvey 1 blandt modersmålsundervisere
November-december	Evt. frivillig 2. klasses læsetest
December	Casestudier på udvalgte skoler

2015

Februar-marts	Fidelitetssurvey 2 blandt modersmålsundervisere
Marts-april	Tilmelding og registrering af nye 1. klasser
Maj-juni	Efterundersøgelse herunder evt. frivillig 2. klasses læsetest
August	Førundersøgelse lærere, elever for nye 1. klasser
Uge 39	Indsats igangsættes
Oktober-november	Fidelitetssurvey 1 blandt modersmålsundervisere
November-december	Evt. frivillig 2. klasses læsetest

2016

Februar-marts	Fidelitetssurvey 2 blandt modersmålsundervisere
Maj-juni	Efterundersøgelse herunder evt. frivillig 2. klasses læsetest
Oktober	Af rapportering af hovedresultater

Fremover Longitudinal opfølgning

3.8 BEMANDING

Her listes projektgruppens medlemmer og kompetencer. For mere detaljerede oplysninger omkring de enkelte medlemmer, se Bilag B.

- **Institut for Økonomi og Institut for Statskundskab, AU:** [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- **Rambøll Management Consulting:** [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED]

De to parter har stor erfaring med at samarbejde om randomiserede kontrollerede forsøg, hvilket er en garanti for, at samarbejdet vil være effektivt og fagligt fokuseret, idet vi ikke skal bruge tid på at forstå hinandens tilgange, styrker og svagheder. Vi har etableret samarbejdet i respekt for hinandens styrker og den værdi, vores samarbejde giver UVM ved at kunne tilrettelægge og gennemføre et forsøg med de højeste metodiske standarder og en effektiv ressourceudnyttelse og projektafvikling, der følger de opstillede tidsplaner.

ARBEJDSDELING

[REDACTED] AU, er projektledere og står som udgangspunkt for al kontakt til UVM i forbindelse med følgeforskningen. Hvis det på et senere tidspunkt anses for hensigtsmæssigt, kan der gøres yderligere undtagelser af denne karakter. Følgeforskningen vil i så fald informere UVM om dette.

AU tager sig primært af randomisering, design af spørgeskemaundersøgelser på baggrund af input fra Rambøll, databehandling, dataanalyse og afrapportering til UVM. Rambøll tager sig primært af fidelitets-survey og implementeringsanalyse, administration af tilmeldinger, udarbejdelse og gennemførelse af spørgeskemaundersøgelser og casestudier, information til samt vejledning af skoler. AU og Rambøll yder dog hinanden gensidig support i forhold til de fleste af disse opgaver.

REFERENCER

- Andersen, D., Jakobsen, V., Jensen, V. M., Nielsen, S. S., Pedersen, K. C. Z., Petersen, D. S. & K. M. Thorsen. 2012. *Indsatser for tosprogede elever. Kortlægning og analyse*. SFI: København.
- Andersen, S. C. & Rambøll Management Consulting. 2010. *Mehmet og modkulturen. En undersøgelse af drenge med etnisk minoritetsbaggrund*. Rambøll. Rapport.
- Andersen, S. C. & S. Serritzlew. 2007. "Vad händer när offentliga skolor utsätts för konkurrens från friskolor? Erfarenheter från Danmark". I: Lindbom, Anders (red.): *Friskolorna och framtiden – segregation, kostnader och effektivitet*, Institutet för Framtidsstudier: Stockholm.
- Egelund, N. & B. S. Rangvid. 2005. *PISA-København 2004*. Akf forlaget: København.
- Egelund, N., Nielsen, C. P. & B. S. Rangvid. 2011. *PISA Etnisk 2009. Etniske og danske unges resultater i PISA 2009*. AKF Rapport. AKF: København.
- Goodman, R., & S. Scott (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful?. *Journal of abnormal child psychology*, 27(1), 17-24.
- Hanushek, E. A., Kain, J. F., Markman, J. M. & S. G. Rivkin. 2003. "Does peer ability affect student achievement?". *Journal of Applied Econometrics* 18(5): 527-544.
- Heckman, J. J. 2008. "Schools, Skills, and Synapses". *Economic Inquiry* 46(3): 289-324.
- Jensen, P. & A. W. Rasmussen. 2011. "The effect of immigrant concentration in schools on native and immigrant children's reading and math skills". *Economics of Education Review* 30: 1503-1515.
- Kirsch, I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J. & C. Monseur. 2002. *Reading for Change. Performance and Engagement Across Countries. Results From PISA 2000*. OECD: Paris.
- Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., Rothenberger, A. & Goodman, R. (2000). Comparing the German versions of the strengths and difficulties questionnaire (SDQ-Deu) and the child behavior checklist. *European child & adolescent psychiatry*, 9(4), 271-276.
- Lavy, V. & A. Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School". *American Economic Journal: Applied Economics* 3(2):1-33.
- Mehlbye, J., Rangvid, B. S., Larsen, B. Ø., Fredriksson, A., & Nielsen, K. S. (2011). *Tosprogede elever undervisning i Danmark og Sverige*. AKF, Anvendt Kommunal Forskning
- Nielsen, H. S. & B. S. Rangvid. "The impact of parents' years since migration on children's academic achievement". Kommande i *IZA Journal of Migration*.
- Rangvid, B. S. 2008. "Private School Diversity in Denmark's National Voucher System". *Scandinavian Journal of Educational Research* 52(4): 331-354.
- Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A. & S. Raudenbusch. 2011. "Optimal Design Plus Empirical Evidence: Documentation for the "Optimal Design" Software". Arbejdsrapport.
- Schochet, P. Z. 2008. "Statistical Power for Random Assignment Evaluations of Education Programs". *Journal of Educational and Behavioral Statistics* 33(1): 62-87.

Rothstein, J. 2009. "Student Sorting and Bias in Value Added Estimation: Selection on Observables and Unobservables". *Education Finance and Policy* 4(4): 537-571.

B FORELØBIGE STYRKEBEREGNINGER

I forsøget skal randomiseringen som udgangspunkt foregå på skoleniveau. Der er således tale om en såkaldt "two-level cluster randomized trial", se Spybrook et al. (2011). Niveau 1 er eleverne og niveau 2, hvor randomiseringen sker, er skolerne. Der vil kun være to niveauer i designet, da kun én elevgruppe udtrækkes til deltagelse i forsøget på hver skole.

I forhold til beregninger af forsøgets styrke er der en række beslutningsvariable og parametre, der er nødvendige at fastsætte. Da vi ikke nødvendigvis kender den sande værdi af disse parametre, må vi i en del tilfælde lade det bero på kvalificerede gæt ud fra den eksisterende litteratur på området. I andre tilfælde kan det afhænge af præferencer, fx hvor store effekter på test-scores vi gerne vil være i stand til at fange. I det følgende beskrives valget af disse beslutningsvariable og parametre.

B.1 BESLUTNINGSVARIABLE OG PARAMETRE

Statistisk styrke er det centrale begreb og angiver sandsynligheden med hvilken, vi vil være i stand til at finde en sand effekt af en given størrelse givet et vist signifikansniveau. Vi følger konventionen og opererer med en styrke på 80 %.

Statistisk signifikans angiver sandsynligheden med hvilken vi fejlagtigt vil finde en effekt, selv om den sande størrelse af effekten er nul. Her opererer man som regel med et signifikansniveau på 0,05, altså er der kun 5 procent chance for at finde en effekt, hvis den sande effekt er nul.

Intra-class correlation angiver hvor stor en del af den samlede varians der stammer fra variation mellem skoler. Som udgangspunkt vil en del af variansen i test-scores stamme fra variation indenfor skoler og en del fra variation mellem skoler. Baseret på amerikanske studier fastsætter vi denne parameter til 0,15.

Forklaringsgraden af baggrundsvARIABLE er anvendelig i forhold til at vurdere, hvor meget usikkerheden af de estimerede effekter kan mindskes ved at inkludere disse variable i analyserne. Dette kan både være socioøkonomiske karakteristika og deslige, men specielt adgang til en "før"-test-score øger erfaringsmæssigt forklaringsgraden betragteligt og dermed den andel af variansen af testresultaterne, der kan forklares af baggrundskarakteristika. I beregningerne antages en forklaringsgrad på 0,5. Analyser på amerikanske data tyder på, at forklaringsgraden vil være mindst 0,5, når man har adgang til et tidligere test score mål, se fx Schochet (2008) og Rothstein (2009).

Minimum Detectable Effect Size (MDES) angiver den mindst mulige effekt, som forsøget skal være i stand til at kunne finde. For at dette er sammenligneligt på tværs af studier angives denne som regel i standardafvigelsesenheder. For eksempel betyder en MDES på 0,2, at forsøget vil være i stand til at finde effekter på testresultater, der er større end 0,2 standardafvigelser. Man anvender ofte en cost-benefit baseret MDES, som angiver at MDES skal være den effekt, som præcis er stor nok til at give et samfundsøkonomisk overskud. En mere pragmatisk tilgang er at se på, hvilken størrelse effekter andre tilsvarende indsatser har givet anledning til. Cost-benefit-tilgangen fører ofte til en MDES på 0,1, mens den anden tilgang fører til en MDES på omkring 0,2-0,3. Da der er tale om en relativt intensiv indsats, vælger vi at sætte den parameter til mellem 0,2 og 0,3. Er effekterne mindre er det formentlig ikke interessant at fortsætte i samme spor i nye projekter.

Antallet af elever i klassen/elevgruppen er antallet af individer i en klynge. Da forsøget består af gruppebaserede indsatser, laves beregninger for et par forskellige gruppestørrelser, da det inden kortlægningen må siges at være svært at forudse størrelsen af disse grupper.

Hvis en skole udtrækkes til at modtage en af indsatserne er det som udgangspunkt alle elever af arabisk, tyrkisk eller anden herkomst på 1. klassestrin, der modtager indsatsen. Såfremt en elevgruppe på en given skole viser sig at være for stor, kan dette løses ved at trække lod mellem de pågældende elever. I forhold til forsøgets styrke er udfordringen imidlertid snarere, at elevgrupperne er relativt små. De nedenstående beregninger gælder både hvis indsatserne betragtes som tre separate indsatser, og hvis de betragtes som en samlet indsats. I førstnævnte tilfælde der dog gøres opmærksom på, at det påkrævede antal elevgrupper i så fald vil gælde for hver af de tre indsatsgrupper.

Eftersom vi ikke kender elevgruppernes størrelse har vi foretaget beregninger både for en konservativ gruppestørrelse på 4 elever og på en mere optimistisk gruppestørrelse på 8 elever. Med en gruppestørrelse på 4 vil det kræve omkring 103 skoler at opnå en MDES på 0,3 og langt over 200 skoler at opnå en MDES på 0,2. Med en gruppestørrelse på 8 elever vil det kræve mere end 70 skoler at opnå en MDES på 0,3 og omkring 150 skoler at opnå en MDES på 0,2. For at vi kan udtale os om, hvorvidt forsøg med modersmålsbaseret undervisning har positive effekter, er det således nødvendigt at have mindst 100 deltagende skoler med 4 elever på hver skole eller lidt færre skoler, men med flere deltagende elever på hver skole.