

Til
Undervisningsministeriet

Dokumenttype
Bilag

Dato
August 2014

BILAG 2

DESIGN OG METODE-

BILAG

BILAG 2 DESIGN OG METODEBILAG

INDHOLD

Design- og metodebilag	Error! Bookmark not defined.
1.1 Forskningsdesign	Error! Bookmark not defined.
1.2 Analysemetoder	2
1.3 Sammenfatning	3

1. Design- og metodebilag

I forskningskortlægningen er studierne blevet vurderet på baggrund af deres forskningsmæssige kvalitet og forsknings- og emnemæssige relevans i relation til indeværende reviewspørgsmål. Dette bilag giver et overblik over de forskningsdesign og analysemetoder, som anvendes i de inkluderede studier. Bilaget har dermed karakter af en læsevejledning til de metodiske begreber, som bliver brugt i abstracts og syntese.

1.1 Forskningsdesign

De inkluderede studier i forskningskortlægningen er baseret på en række forskellige forskningsdesign. I nedenstående tabel forklares de forskellige forskningsdesign, som er anvendt i studierne.

Tabel 1: Oversigt over forskningsdesign

<p>Systematisk review</p> <p>Et systematisk review er et review af den samlede forskning inden for et givent område. Systematiske reviews er baseret på en systematisk strategi for søgningen efter studier (fx hvilke databaser der søges i, hvilke keywords der søges efter, hvilke tidsskrifter der er blevet gennemgået ved manuelle søgninger etc.) og indeholder kriterier for at inkludere og ekskludere studier i reviewet samt metoder til at vurdere og sammenfatte resultaterne af de inkluderede studier.</p>
<p>Randomiseret, kontrolleret forsøg (RCT)</p> <p>Randomiserede, kontrollerede forsøg indebærer, at forsøgspersonerne inddeles i mindst én indsats- og kontrolgruppe. Indsatsgruppen modtager den indsats, som studiet undersøger effekten af. Kontrolgruppen modtager ingen indsats og fortsætter typisk med den gængse praksis på området, fx et bestemt undervisningsforløb. Kontrolgruppen fungerer således som et sammenligningsgrundlag, som resultaterne i indsatsgruppen kan holdes op imod. Herved kan effekten af en indsats vurderes.</p> <p>For at der er tale om et RCT, er det afgørende, at fordelingen mellem indsats- og kontrolgruppen er tilfældig. Den tilfældige tildeling til grupper sikrer mod selektionsbias og medfører, at indsats- og kontrolgruppen har de samme karakteristika inden for statistisk tilfældige grænser. Målet er dermed at sikre, at påvirkningen, som indsatsgruppen modtager, er den eneste forskel mellem de to grupper, således at eventuelle effekter for indsatsgruppen kan tilskrives indsatsen.</p>
<p>Kvasi-eksperiment</p> <p>I kortlægningen anlægges en bred forståelse af et kvasi-eksperimentelle forskningsdesigns. Kvasi-eksperimentelle forskningsdesigns dækker således over flere typer design, såsom naturlige eksperimenter, matching og propensity score matching.</p> <p>Kvasi-eksperimenter indebærer ligesom RCT-studier typisk en indsats- og kontrolgruppe (i nogle tilfælde kaldt en sammenligningsgruppe). Indsatsgruppen modtager en indsats, mens kontrolgruppen ikke gør. Alternativt kan et kvasi-eksperiment bestå af flere indsatsgrupper og ingen kontrolgruppe, hvilket muliggør vurdering af relative effekter. Til forskel fra RCT-studier er allokeringen af deltagerne i grupper ikke tilfældig (randomiseret). Dette kan eksempelvis skyldes, at tilfældig udvælgelse ikke er mulig, eller at man har interesse i en særlig målgruppe.</p> <p>I naturlige eksperimenter anvendes en naturlig variation i den/de uafhængige variable, som er opstået eksogent, dvs. uden for et eksperiment, til at undersøge forskelle mellem en indsats- og en sammenligningsgruppe. I naturlige eksperimenter vil der sjældent være tale om en randomiseret fordeling af enhederne i indsats- og kontrolgrupperne.</p> <p>Matching kan anvendes til fordelingen af grupper, hvor randomisering ikke er mulig. I matching er henholdsvis indsats- og sammenligningsgrupperne sammensat, så de ligner hinanden på variable, som antages at være centrale i forhold til det specifikke forskningsspørgsmål, som undersøges.</p> <p>Propensity score matching er en form for statistisk matching, hvor man forsøger at estimere effekten af en indsats ved at tage højde for forhold, som kan påvirke sandsynligheden for, at en enhed modtager den givne indsats.</p>
<p>Før- og eftermålinger</p> <p>Design med før- og eftermålinger indebærer, at en gruppe forsøgspersoner testes på et outcome af interesse, før og efter de gives en indsats. Formålet er at undersøge, om der er sket en forandring på</p>

de områder, som indsatsen søger at påvirke. Der kan eksempelvis være tale om en indsats, der sigter mod at forbedre elevers motivation. Elevernes motivation måles før og efter indsatsen for at belyse, om indsatsen har en virkning.

Longitudinelt studie

I longitudinelle studier følges de samme individer over tid. Det er forskelligt, om gruppen af individer, der følges over tid, modtager en indsats eller ej. Typisk indhentes måleresultater for deltagerne på flere tidspunkter i løbet af undersøgelsens forløb. Der findes forskellige typer longitudinelle studier, hvoraf kohortestudier er et af de mest udbredte longitudinelle forskningsdesign. I kohortestudier følges individer, som er kendetegnet ved at have oplevet den samme begivenhed inden for en given tidsperiode. Et typisk eksempel på kohortestudier er at følge børn, der er født i samme år.

Tværsnitsstudie

I tværsnitsstudier indsamles data på tværs af en gruppe individer eller organisationer på et givent tidspunkt. I et tværsnitsstudie vil data have form af kvantitative eller kvantificerbare data om mere end to variable, idet formålet er at belyse variation på tværs. Tværsnitsstudier kan afdække samvariation mellem variable, men kan ikke fastslå retningen af en sammenhæng, idet der kun indgår data fra ét tidspunkt.

Casestudie

Casestudier er kvalitative studier af en eller få enheder, fx skoler, som sigter mod dybdegående viden om de undersøgte cases. Dataindsamlingen i casestudier kan foregå på mange forskellige måder, eksempelvis via interviews, fokusgruppeinterviews, observation og tekstanalyse. Casestudier er velegnede til at undersøge komplekse sammenhænge, men kan kun i begrænset omfang belyse effekten af en indsats.

Et casekontrolstudie er imidlertid et eksempel på et casestudiedesign, der kan bidrage til at identificere kausale sammenhænge. I et casekontrolstudie udvælger man forsøgspersonerne på baggrund af, om de er kendetegnet ved det karakteristika eller den egenskab, man ønsker at undersøge (casegruppen) eller ikke (sammenligningsgruppen). Potentielle kausale sammenhænge undersøges herefter ved at sammenligne relevante outcomes i gruppen med (casegruppen) eller uden (sammenligningsgruppen) den relevante egenskab. Hvis der er en systematisk forskel i outcomes mellem case- og sammenligningsgruppen, kan forskellen henføres til tilstedeværelsen og/eller fraværet af den egenskab, som kun kendetegner den ene gruppe, såfremt det alene er denne egenskab, der varierer mellem grupperne. Casekontrolstudier kombineres ofte med longitudinelle undersøgelser.

Kilder: Rieper, Olaf & Hanne Foss Hansen (2007): *Metodedebatten om evidens*. København: AKF Forlaget. Andersen, Lotte Bøgh, Kasper Møller Hansen & Robert Klemmensen (Red.) (2012): *Metoder i Statskundskab*. København: Hans Reitzels Forlag.

1.2 Analysemetoder

Studierne i forskningskortlægningen anvender mange forskellige analysemetoder til at belyse forskellige indsatser, praksisser og redskabers betydning for elevers alsidige udvikling og sociale kompetencer. I abstracts anvendes en række samlebetegnelser for analysemetoderne, som uddybes nedenfor.

Tabel 2: Tabel over analysemetoder

Variansanalyse

Variansanalyse bruges som en fællesbetegnelse, der dækker over en række statistiske analysemetoder, der analyserer forskelle i flere gruppers gennemsnit. De forskellige typer variansanalyser er af varierende statistisk styrke.

Den mest simple form for variansanalyse er en statistisk test af, om flere gruppers gennemsnit er forskelligt (ANOVA). Der er flere forskellige varianter af denne type variansanalyse. Eksempelvis anvendes envejs-ANOVA, når der er én uafhængig variabel, mens tovejs-ANOVA anvendes, når der er mere end én uafhængig variabel. Multivariat variansanalyse (MANOVA) benyttes, når der er mere end én afhængig variabel.

Analyse af samvarians (ANCOVA) er en mere avanceret analysemetode, som belyser, om gennemsnittene for en afhængig variabel er ens på tværs af forskellige niveauer af en kategorisk, uafhængig variabel. Multivariat analyse af samvarians (MANCOVA) anvendes, når der er mere end én afhængig variabel.

Regressionsanalyse

Regressionsanalyse er en bred betegnelse, som dækker over forskellige typer statistiske analyser, der undersøger sammenhængen mellem en afhængig variabel og flere uafhængige variable. Analysemetoden estimerer effekten af en given uafhængig variabel på den afhængige variabel, mens der kontrolleres for betydningen af øvrige uafhængige variable (kontrolvariable) i analysen.

Der findes mange forskellige typer regressionsanalyser, som kan anvendes, alt efter hvilke typer variable der indgår i modellen og deres statistiske sammenhæng. Hvis den afhængige variabel eksempelvis antages at være intervallskaleret, kan multivariat, lineær regression anvendes, mens logistisk regression anvendes, hvis den afhængige variabel er binær. Ikke-parametriske statistiske regressionsmodeller anvendes, når de sædvanlige antagelser om populationens fordeling (normalitetsantagelsen) ikke er opfyldt.

Økonometriske vækstmodeller

En analysemodel inden for gruppen af økonometriske vækstmodeller er structural equation modelling (SEM), hvor statistiske data og kvalitative kausalantagelser bruges til at teste og estimere kausalsammenhænge.

Der skelnes grundlæggende mellem to hovedkomponenter inden for SEM-modeller: Henholdsvis *strukturelle modeller*, der anvendes til at vise afhængighed mellem endogene og eksogene variable, og *målemodeller*, der anvendes til at påvise sammenhænge mellem latente variable og deres indikatorer. SEM kan både bygge på variansanalyse, regressionsanalyse, stianalyser og hierarkiske regressionsmodeller, fx multilevel analyse.

Statistisk modellering (structural path analysis) er en specifik form for SEM, som ikke indeholder målemodellen i SEM, men alene bruger den strukturelle model af SEM til at afdække direkte og indirekte afhængighed mellem udvalgte variable.

Faktoranalyse indeholder omvendt kun målemodellen fra SEM og anvendes til at påvise underliggende faktorer mellem korrelerede variable.

Hierarkiske modeller

Hierarkiske modeller er en samlebetegnelse for forskellige statistiske analysemodeller, som er kendetegnet ved variation på flere niveauer, fx både på skoleniveau og på elevniveau. Det er centralt i forhold til hierarkiske modeller, at forskelle mellem populationen på forskellige niveauer er relateret til en teoretisk forventning om variation på tværs af de forskellige niveauer.

Der er flere forskellige typer analysemetoder inden for de hierarkiske modeller, herunder både lineære og ikke-lineære modeller. Brugen af fixed effects i multilevel analyser giver således mulighed for at fjerne uobserverede effekter, eksempelvis effekten på eleverne af at være i en bestemt klasse på en bestemt skole. Random effects kan bruges, når man antager, at den uobserverede effekt ikke er korreleret med nogen af de uafhængige variable.

Kvasi-eksperimentelle analysetilgange

Kvasi-eksperimentelle analysetilgange bruges som en samlebetegnelse for analysemetoder, hvor det ikke er muligt at skabe en faktisk kontrolgruppe. Kvasi-eksperimentelle analysemetoder dækker blandt andet over difference-in-difference og regression discontinuity modellering.

Difference-in-difference bruges til at estimere effekten af en indsats ved at udregne forskellen mellem før- og eftermåling for både indsats- og kontrolgruppe, og herefter trækkes ændringen for kontrolgruppen fra ændringen for indsatsgruppen, hvilket giver et estimat for effekten af indsatsen. Regression discontinuity udnytter tærskelværdier til at estimere effekt ved at sammenligne observationer, der ligger tæt på hver side af tærskelværdien.

Kilder: Agresti, Alan & Barbara Finlay (2009): *Statistical methods in the social sciences* (4. udgave). Pearson Prentice Hall.
Wooldridge, Jeffrey M. (2009): *Introductory econometrics – a modern approach* (4. udgave). South-Western Cengage Learning.

1.1 Sammenfatning

Ovenstående giver en kort indføring i de forskningsdesigns og analysemetoder, som er anvendt på tværs af de studier, der indgår i forskningskortlægningen om alsidig udvikling og sociale kompetencer. Det er centralt at have for øje, at kombinationen af forskningsdesign og analysemetode har betydning for styrken af et studies resultater.

Eksempelvis kan en simpel variansanalyse af data indsamlet via et randomiseret, kontrolleret forsøg medføre robuste resultater, idet den tilfældige udvælgelse gør, at indsats- og kontrolgruppe har samme karakteristika (også uobserverede faktorer). Hvis

kontrol- og indsatsgruppe derimod ikke er tilfældigt udvalgt – som i et kvasi-eksperimentelt design – vil en variansanalyse ikke give helt lige så robuste resultater. Anvendelsen af en regressionsanalyse i kvasi-eksperimentelle designs vil dog kunne kompensere for svagheder i undersøgelser, hvor indsats- og kontrolgruppe ikke er tilfældigt udvalgt, idet analysen kan kontrollere for relevante baggrundsfaktorer, eksempelvis køn, alder og socioøkonomisk baggrund.

Størrelsen af et studies stikprøve har ligeledes betydning i sammenhæng med forskningsdesign og analysemetode. Eksempelvis kræver et randomiseret, kontrolleret forsøg en langt mindre stikprøve end et tværsnitsstudie. Selv for randomiserede, kontrollerede forsøg er der dog en nedre grænse for størrelsen af stikprøven, hvis resultaterne skal være robuste. Analysemetoden er også relevant i den sammenhæng, idet brug af eksempelvis hierarkiske modeller medfører et behov for en større stikprøve.

Forskningsdesign, analysemetode og stikprøvestørrelse bør altså ikke ses som uafhængige faktorer i vurderingen af et studies kvalitet, men skal sammenholdes for at danne et retvisende billede af styrken af et studies resultater.