

Michael Bang Petersen, Rune Slothuus, Rune Stubager og Lise Tøgeby

## Eksperimenter: Et redskab i politologens værktøjskasse?

Eksperimenter har ikke tidligere hørt til de metoder, som blev anvendt inden for statskundskaben. I de senere år er eksperimenter imidlertid blevet mere og mere udbredt i amerikansk statskundskab og med dette nummer af *Politica* introduceres eksperimenter også for et dansk publikum. Interessen for eksperimenter er først og fremmest en følge af, at man med eksperimenter får en mulighed for at udtale sig om årsags-virkningsrelationer. Karakteristisk for eksperimenter er, at man som i medicin foretager undersøgelser, hvor egenskaber bliver manipuleret, og deres effekt på andre egenskaber observeret.

Vi kender alle sammen eksperimenter fra skolen. Men det er noget, der hører til i fysik- eller kemilokalet, ikke noget man beskæftiger sig med i de samfundsvidenskabelige eller humanistiske fag. Vi ved også, at eksperimenter er udbredt i medicinsk forskning. Når der er kommet nye behandlingsformer på markedet, skal de afprøves ved blindforsøg, hvor tilfældigt udvalgte patienter får én type behandling, og de resterende en anden. Før ny medicin kan godkendes, må man nemlig sikre sig, at de nye piller har den ønskede effekt og ikke andre. Og det gør man ved hjælp af eksperimenter. Eksperimenter kan bruges til at etablere relationen mellem årsag og virkning.

Der er ingen større tradition for at anvende eksperimenter i statskundskaben – og slet ikke i Danmark. Når vi stræber efter at opnå den samme præcision som i naturvidenskaberne, gør vi det ved hjælp af statistisk manipulation, ved kontrol for tredjevariable. Men herigennem når vi aldrig længere end til at etablere en sammenhæng mellem to variable. Vi kan ikke udtale os om årsag og virkning.

Eksperimenter er imidlertid i disse år ved at blive mere udbredt i amerikansk statskundskab. Og ved eksperimenter forstår vi undersøgelser som i medicin, hvor egenskaber bliver manipuleret, og deres effekt på andre egenskaber observeret (Campbell og Stanley, 1966). Og hvor det derfor bliver muligt at udtale sig om årsag og virkning.

Dette temanummer introducerer eksperimenter som et af flere redskaber, der står til rådighed for politologiske forskere. Her beskriver vi metoden mere i detaljer, mens de følgende fire artikler hver på sin måde illustrerer brugen af

eksperimenter i politologisk forskning. Formålet er at vise, hvordan vi kan blive klogere på statskundskabens klassiske emner ved at inddrage disse nye metoder.

### Hvad er eksperimenter for noget?

Til daglig bruges begrebet *eksperiment* forholdsvis løsagtigt som betegnelsen for, at man prøver noget nyt og ser, hvad resultatet bliver. Også i statskundskab har man tidligere brugt begrebet eksperimenter forholdsvis bredt, fx i forbindelse med institutionelle ændringer eller med simulationer.

I de senere år har man imidlertid reserveret begrebet *eksperiment* til undersøgelser, hvor to eller flere grupper, der er sammensat ved tilfældig udvælgelse, udsættes for forskellige stimuli, med henblik på at måle deres efterfølgende respons. Og det er den tilfældige fordeling, eller *randomiseringen*, der er det afgørende og definerende træk. Nogle gange ser man undersøgelser, hvor forskellige grupper godt nok udsættes for forskellige stimuli, men hvor der ikke er sket en ægte randomisering. I så fald taler man om quasi-eksperimenter. For at karakterisere en undersøgelse som et eksperiment skal der altså være foretaget en ægte randomisering.

Mange eksperimenter foregår i laboratorier, men der foretages også eksperimenter under mere naturlige forhold. For at illustrere, hvorledes et sådan eksperiment kan konstrueres, vil vi kort omtale det første eksperiment, som blev rapporteret i *American Political Science Review* i 1956 af Samuel Eldersveld (jf. Druckman et al., 2006). Det handlede om brugen af forskellige overtalelseteknikker med henblik på at øge valgdeltagelsen. En tilfældig udvalgt del af vælgerne blev udsat for en stimulus i form af en personlig henvendelse med en opfordring til at stemme, en anden tilfældigt udvalgt gruppe fik en stimulus i form af skriftligt materiale, og endelig var der en kontrolgruppe, som ikke blev udsat for nogen form for særlig stimulus. Eksperimentet viste, at valgdeltagelsen var væsentligt højere i den gruppe, der havde fået en personlig henvendelse, end i den gruppe, der havde fået skriftligt materiale, og at valgdeltagelsen i denne igen var en smule højere end i kontrolgruppen. Konklusionen var derfor, at en personlig henvendelse øgede vælgernes valgdeltagelse.

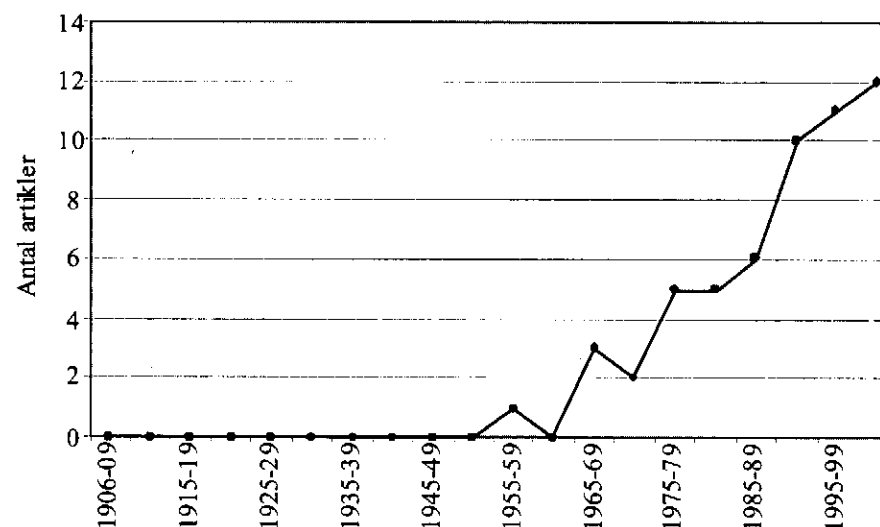
Det nævnte eksempel er også en god illustration af, hvorfor man anvender eksperimenter. Eksperimentet giver for det første mulighed for at undersøge mere dynamiske problemstillinger, eksempelvis hvordan valgdeltagelsen ændrer sig, hvis vælgerne får en personlig opfordring til at stemme. Gennem randomiseringen muliggør eksperimentet for det andet, at man kan kontrollere for andre betydende faktorer. Da deltagerne er blevet fordelt tilfældigt på de forskellige grupper, vil grupperne i princippet være ens sammensat, fx vil der være en lige stor andel kvinder, højtuddannede og meget politisk interesserede. Og endelig for det tredje muliggør eksperimentet, at vi kan bevæge os

fra at undersøge sammenhænge mellem variabler til at påvise kausalrelationer, dvs. identificere årsag og virkning.

### Udbredelsen af eksperimenter i statskundskaben

I sin *Presidential address* til American Political Science Association i 1909 advarede A. Lawrence Lowell politologerne mod at følge i naturvidenskabens fodspor: „Vi er begrænset af, at det er umuligt at gennemføre eksperimenter. Statskundskaben er en observerende, ikke en eksperimenterende videnskab“ (Lowell, 1910: 7). Og hvis man ser på de publicerede artikler i de førende politologiske tidsskrifter blev dette råd omhyggeligt fulgt. Efterhånden aftog imidlertid den principielle modvilje mod eksperimenter. Efter publiceringen af Campbell og Stanleys diskussion af metoden i *Experimental and Quasi-experimental Designs for Research* (1966) blev det efterhånden opfattelsen, at den empiriske forskning ville vinde ved i det mindste at inddrage eksperimentlogikken i sine overvejelser (se fx Lijphart, 1971). Fortsat var det dog meget småt med faglige artikler, der byggede på eksperimenter. Men langsomt skete der noget.

Figur 1. Udviklingen i antallet af artikler i *American Political Science Review*, som rapporterer resultater fra eksperimenter



Figur 1 viser udviklingen i antallet af artikler i *American Political Science Review* (APSR), som rapporterer resultater fra egentlige eksperimenter. Oplysningerne er lånt fra en artikel af James N. Druckman m.fl., publiceret i APSR i 2006. Som man kan se, blev den første artikel publiceret i 1956, og det er den, vi har

refereret ovenfor. Den næste blev først publiceret omkring ti år senere. I de følgende år blev der publiceret i gennemsnit en enkelt artikel om året. Fra 1990 er der publiceret i gennemsnit to om året, og antallet ser ud til fortsat at være svagt stigende.

Nu findes der jo mange andre tidsskrifter end APSR, men dette tidsskrift er formodentlig et godt mål på fagets professionelle normer. Man kan ikke sige, at der nogensinde er blevet publiceret mange artikler, der bygger på eksperimenter, men der er dog siden 1960'erne sket en udvikling fra slet ingen artikler til 2-3 om året i APSR. Man må sige, at siden i hvert fald 1990 har eksperimenter været et anerkendt redskab i politologens værktøjskasse (McGraw og Hoekstra, 1994; Druckman et al., 2006; Gaines et al., 2007).

Men det er karakteristisk, at eksperimenter optræder oftere i nogen subdiscipliner end i andre. De fleste eksperimentbaserede artikler optræder i forbindelse med individuel adfærd, dvs. i artikler om politisk psykologi, valgforskning, medieforskning og parlamentsadfærd. Der er kun få eksperimenter inden for international politik eller forvaltningsforskning.

Væksten i antallet af eksperimentbaserede artikler er en følge af både øget udbud og øget efterspørgsel. Som Druckman et al. (2006) peger på, har der i kvantitativt orienteret forskning været en stigende erkendelse af, at de statistiske redskaber ikke slog til, når man skulle konkludere på det centrale spørgsmål om kausalitet. Der har derfor været en søgen efter nye værktøjer, der kunne hjælpe os, og her fremstod eksperimentet som en frugtbar mulighed.

Men væksten har også været en følge af, at vi har fået adgang til nye indsamlingsteknikker. Langt de fleste af de tidlige eksperimenter, der er publiceret i politologiske tidsskrifter, har været laboratorieeksperimenter og ofte med universitetsstuderende som deltagere. At eksperimenterne forudsatte, at man kunne få folk til at møde op i et laboratorium, begrænsede naturligvis anvendelsen og begrænsede ofte også mulighederne for at generalisere. Dette lagde en naturlig dæmper på interessen for eksperimenter. Med opfindelsen af CATI-teknikken (computer-assisted telephone interviewing) eksploderede imidlertid mulighederne for at foretage eksperimenter. Fordi man kan få computeren til at foretage randomiseringen af stimuli, er det pludselig muligt at indbygge eksperimenter i almindelige survey (se Sniderman og Grob, 1996).

### Styrker og svagheder ved anvendelsen af eksperimenter

Den eksperimentelle logik er den grundlæggende logik, vi benytter os af, hver gang vi foretager videnskabelig inferens (Lijphart, 1971). Det er fx også den logik, der ligger bag statistiske kontrolanalyser. Eksperimenter adskiller sig imidlertid fra de metoder, vi ellers anvender, på i hvert fald to måder (McGraw, 1996).

For det første er der tale om et aktivt indgreb i tingenes tilstand. Forskeren

skaber selv de nødvendige betingelser for observationen i stedet for blot – som Lowell mente, man måtte – at observere naturligt forekommende fænomener. Forskeren har således også fuld kontrol over, hvilke stimuli forsøgspersonerne udsættes for.

For det andet løser man i eksperimentet kontrolproblemet på en anden måde end ellers, nemlig gennem randomisering. Det overordnede princip for testning af hypoteser i empirisk forskning er, at man skal kontrollere for tredjevariable. Vi kan et øjeblik vende tilbage til eksemplet fra Eldersveld om effekten af forskellige metoder til at øge valgdeltagelsen. I det virkelige liv vil der være stor sandsynlighed for, at de mennesker, der får en personlig henvendelse fra partiernes kampagnemedarbejdere, også er dem med længst uddannelse og størst politiske interesse. Hvis man skal undersøge effekten af henvendelsen, må man derfor kontrollere for effekten af uddannelse og interesse. Indsamles oplysningerne om henvendelserne efterfølgende gennem en almindelig surveyundersøgelse, må man løse kontrolproblemet ved statistisk at holde de andre variabler konstant, som kunne tænkes at influere på valgdeltagelsen – forudsat selvfølgelig, at man har målt dem i den pågældende undersøgelse.

Anvender man i stedet for et eksperiment, gennemføres kontrollen ved hjælp af den randomiserede udvælgelse. Man sørger altså på forhånd for, at de grupper, der får en henvendelse, er magen til de grupper, der ikke får henvendelsen. Når undersøgelsespersonerne ved tilfældig udtrækning enten får en personlig henvendelse, en skriftlig henvendelse eller slet ingen henvendelse, vil de tre grupper have samme sammensætning. Ved en perfekt randomisering vil der være den samme andel med høj politisk interesse i alle tre grupper og den samme andel med lav politisk interesse. Uddannelsessammensætningen vil også være den samme i de tre grupper. Disse egenskaber bliver således holdt konstant. Men manipulationen foregår her rent fysisk, ikke statistisk. Når kontrollen foregår gennem randomisering, får man også kontrolleret for alle andre faktorer, også dem, der ikke engang er målt i undersøgelsen. I dette tilfælde sætter fantasien altså ingen grænser. Og man kan derfor være sikker på, at de umiddelbare forskelle, der fremkommer mellem grupper, der har modtaget forskellige stimuli, også kan føres tilbage til disse stimuli. Undersøgelsespersonerne er nemlig ens i alle andre henseender end den modtagne stimulus.

Hermed har vi også løst kausalitetsproblemet. Stimulussen kommer nemlig i alle tilfælde før responsen. Og der kan ikke være tale om nogen selvselektion, der kan vende kausalrelationen på hovedet. På denne måde har vi bevæget os fra sammenhænge til kausalrelationer.

Eldersvelts eksperiment om valgdeltagelsen blev foretaget som led i en faktisk politisk kampagne ved et lokalvalg i USA. Almindeligvis er situationen omkring eksperimenter betydelig mere kunstig, og det gælder både laboratorieforsøg og surveyeksperimenter. Deltagerne er under alle omstæn-

digheder vidende om, at de deltager i en undersøgelse, og dette kan i sig selv påvirke adfærden. Specielt laboratoriesituationen er kunstig, men også telefoninterviewet er det. Derudover har de to teknikker hver deres styrker og svagheder.

Fordelen ved laboratorieeksperimenterne er først og fremmest, at man kan udsætte folk for mere omfattende og komplicerede stimuli, end man kan i forbindelse med surveyeksperimenter. Man kan bede dem læse en avisartikel eller flere, og man kan vise film eller tv-programmer. Man kan også spørge dybere ind til reaktionerne på det set og dermed få bedre hold på det, man kalder *mediatorerne* eller de mellemkommende variabler. Man har altså bedre muligheder for at finde ud af, hvorfor folk reagerer på stimulussen, som de gør.

Laboratorieeksperimenternes svaghed er først og fremmest den *eksterne validitet*. Laboratoriesituationen er kunstig, og det kan være vanskeligt at konstruere realistiske stimuli. Samtidig er det vanskeligt at komme i kontakt med et repræsentativt udsnit af befolkningen. Interviewpersonerne udgør ofte det, man kalder et *convenience sample*. Det er altså folk, der er lette at få fat i, fx fordi de befinder sig på samme sted. Laboratorieforsøg udføres da også ofte med elever eller studerende fra diverse uddannelsesinstitutioner. Det er svært at generalisere resultaterne til hele befolkningen og til den daglige politik. Det betyder ikke, at laboratorieforsøg er nytteløse, de har blot deres begrænsning. Og eksperimenter er stadigvæk den bedste metode til at få viden om kausal-sammenhænge.

Surveyeksperimenternes styrke er først og fremmest, at det er muligt at foretage eksperimentet på et repræsentativt udsnit af befolkningen, og at det er muligt at indsamle oplysninger om et meget stort antal personer. Hermed sikres først og fremmest en større ekstern validitet. Også telefoninterviewet skaber selvfølgelig en kunstig situation, men det gælder alle surveyundersøgelser. Den repræsentative stikprøve gør det imidlertid muligt at generalisere til hele befolkningen.

Det store antal cases betyder samtidig, at man kan undersøge betydningen af de såkaldte *moderatorer*, dvs. faktorer, der har betydning for, om stimulussen har en effekt. Ofte vil forskellige grupper påvirkes forskelligt af forskellige stimuli. Fx kan man forestille sig, at folk med en omfattende politisk viden og interesse påvirkes anderledes end folk med ringe viden eller interesse, måske fordi informationen ikke er ny for dem, eller fordi deres holdninger er mere sammenhængende.

Samtidig er der nogle rent praktiske fordele ved surveyeksperimenterne. Fordi computeren er i stand til at foretage stadig nye randomiseringer, kan man indbygge mange forskellige eksperimenter i den samme undersøgelse. Det er også muligt at forhindre kunstig konsistens imellem spørgsmålene, fordi

man fx i stedet for at spørge alle om deres holdning til forskellige politiske gruppers rettigheder, kan spørge nogle respondenter om én gruppes rettigheder og andre om andre gruppers rettigheder. Og endelig kan man randomisere spørgsmålene, så heller ikke spørgsmålsrækkefølgen spiller nogen rolle (Gaines et al., 2007). Mulighederne for at randomisere, der er indeholdt i computerteknikken, betyder således, at vi kan konstruere hundredvis af forskellige spørgeskemaer, hvor både versioner af det samme spørgsmål og rækkefølgen af spørgsmålene varierer.

Men surveyeksperimenterne har også deres svagheder. Det er stadigvæk vigtigt, at eksperimenterne bliver så realistiske som muligt – og det kan være svært at sikre (Gaines et al., 2007). Også her må man til stadighed bekymre sig over den eksterne validitet: Kan man virkelig generalisere fra undersøgelsen til den politiske hverdag? Dernæst kan man ikke udsætte folk for så omfattende stimuli som ved laboratorieeksperimenter. Man kan ikke sætte dem til at læse en artikel eller se et tv-indslag. Variationerne i stimuli vil som regel være begrænset til udskiftning af ord eller variationer i introduktionen til et spørgsmål. Der er derfor også snævre grænser for, hvad man kan undersøge ved hjælp af surveyeksperimenter – i hvert fald over telefonen. Men i øjeblikket er der stadigvæk mange uudnyttede muligheder. Dertil kommer så, at man via internettet kan foretage undersøgelser, hvor man kan nå ud til en stor og bred deltagergruppe, samtidig med at deltagerne kan vises avancerede stimuli i form af billeder eller videoklip, og man kan foretage målinger af fx svartiden.

Eksperimenter, uanset om de foregår på den ene eller den anden måde, rejser hver gang nogle etiske problemer. I nogle tilfælde vil det være muligt via det gennemførte eksperiment at gribe ind i selve den politiske proces. Det var fx tilfældet, når Eldersveld øgede valgdeltagelsen i et bestemt lokalområde. Dette var der næppe så store etiske problemer forbundet med, men man kan også forestille sig situationer, der er mindre uskyldige.

Man kan desuden indvende, at vi i en vis forstand snyder deltagerne, når vi lader dem indgå i eksperimenter, specielt hvis de ikke bliver fortalt, at de er deltagere i et eksperiment, eller hvad pointen med eksperimentet er. Ved laboratorieforsøg har man mulighed for efterfølgende at foretage en grundig *debriefing*, dvs. fortælle forsøgspersonerne om eksperimentet. Det kan man ikke på samme måde i forbindelse med surveyeksperimenterne, selvom det også er muligt – og nogle gange nødvendigt – at debriefe.

I den undersøgelse, som vi selv har gennemført, og som indgår i to af artiklerne i dette nummer af *Politica*, har vi valgt at løse nogle af de etiske problemer på den måde, at vi aldrig har fortalt folk noget, der ikke var sandt. Vi har fx ikke tillagt partier eller personer synspunkter, som de ikke offentligt har indtaget. Og det var indimellem vanskeligt, fordi man ikke altid kunne finde ud af, hvad partierne mente. Men det ændrer ikke ved, at vi har udsat vore

interviewpersoner for manipulation. Der er altid et etisk problem, som man konkret må overveje.

### Temanummerets artikler

Temanummeret indeholder fire artikler, der hver for sig demonstrerer brugen af eksperimenter inden for statskundskaben. De handler om medier, holdningsdannelse og vælgeradfærd, dvs. de emner, hvor eksperimenter i almindelighed er mest udbredt. Tre af artiklerne bygger på surveyeksperimenter og en enkelt på laboratorieeksperimenter.

Kathrine Adelbert Hansen og Helle Nygaard Jensen har skrevet en artikel om politisk kynisme, som bygger på et laboratorieeksperiment. De undersøger, om det påvirker folk, at de ser en tv-udsendelse, som fokuserer meget på det politiske spil og de strategiske manøvrer omkring et lovforslag og ikke så meget på forslaget indhold. Øger det den politiske kynisme?

De to artikler af Michael Bang Petersen, Rune Slothuus, Rune Stubager og Lise Togeby er skrevet på baggrund af et nyt stort projekt om samspejlet mellem politiske omgivelser og individuelle prædispositioner i holdningsdannelsen. Den første artikel undersøger befolkningens holdninger til velfærdsydelser: Hvem, mener man, er værdige til at modtage hjælp fra det offentlige og hvem ikke? Den anden artikel handler om politiske tolerance, og det centrale spørgsmål er, om tolerancen varierer efter, hvilke grupper der tales om. Men dernæst bliver der gjort forsøg på at ændre folks holdninger ved at præsentere dem for argumenter, der går mod deres oprindelige standpunkt. Hvor bevægelige er folk så i denne situation?

Den fjerde artikel, der er skrevet af Kasper Møller Hansen og Mickael Bech, anvender en særlig teknik, som betegnes *discrete choice*. Spørgsmålet er her, hvordan man prioriterer forskellige goder: lav arbejdsløshed, større løn, en bestemt statsminister?

I alle fire artikler har vi forsøgt at fokusere på det substantielle spørgsmål og ikke på metoden. Artiklerne demonstrerer, hvordan vi kan blive klogere af at anvende disse nye metoder som supplement til de gammelkendte.

### Litteratur

- Campbell, Donald T. and Julian C. Stanley (1966). *Experimental and quasi-experimental designs for research*, Chicago: Rand McNally & Company.
- Druckman, James N., Donald P. Green, James H. Kuklinski and Arthur Lupia (2006). „The Growth and Development of Experimental Research“, *American Political Science Review*, Vol. 100, No. 4, pp. 627-635.

- Eldersveld, Samuel J. (1956). „Experimental Propaganda Techniques and Voting Behavior“, *American Political Science Review*, Vol. 50, No. 1, pp. 154-165.
- Gaines, Brian J., James H. Kuklinski and Paul J. Quirk (2007). „The Logic of the Survey Experiment Reexamined“, *Political Analysis*, Vol. 15, No. 1, pp. 1-20.
- Lijphart, Arend (1971). „Comparative Politics and the Comparative Method“, *American Political Science Review*, Vol. 65, No. 3, pp. 682-693.
- Lowell, A. Lawrence (1910). „The Physiology of Politics“, *American Political Science Review*, Vol. 4, No. 1, pp. 1-15.
- McGraw, Kathleen M. (1996). „Political Methodology: Research Design and Experimental Methods“, pp. 769-786 in Robert Goodin and Hans-Dieter Klingemann (eds.), *A New Handbook of Political Science*, Oxford: Oxford University Press.
- McGraw, Kathleen M. and Valerie Hoekstra (1994). „Experimentation in Political Science: Historical Trends and Future Directions“, *Research in Micropolitics*, Vol. 4, pp. 3-29.
- Sniderman, Paul M. and Douglas B. Grob (1996). „Innovations in Experimental Design in Attitude Surveys“, *Annual Review of Sociology*, Vol. 22, pp. 377-399.