

Hvorfor regeringer og velgørende organisationer skal eksperimentere med os

Af Jonathan Breckon, leder af The Alliance for Useful Evidence, Nesta



DANSK LYKKE

Ekstrabladet 27.11.1910

Regeringer og velgørende organisationer har en moralsk pligt til at eksperimentere. Uden at prøve ting af lærer vi aldrig noget. Vi er nødt til at eksperimentere med nye måder at arbejde på ved at starte i det små, før vi lancerer det på landsplan – og selv på internationalt niveau.

Uanset om det drejer sig om de bedste måder at lære vores børn at læse på, om at udarbejde programmer til at få arbejdsløse tilbage i arbejde eller om at fremme organ donation, har vi brug for at eksperimentere. Hvis de gamle måder ikke fungerer, må vi teste nye løsninger. Og disse eksperimenter kan ikke altid udføres på papiret af udvalg af tjenestemænd eller på et universitetslaboratorium, men må i stedet finde sted ude i den rodede, virkelige verden: i vores boliger, på hospitaler og på skoler.

Alt for ofte bliver nye offentlige politikker lanceret nationalt med for lidt ordentlig afprøvning eller evaluering. Egentlig eksperimenterer regeringer på hele befolkninger på samme tid – men uden at lære noget.¹⁾ I Storbritannien har man offentlige projekter til en værdi af mere end 685 milliarder kroner, men man har ingen planer om at evaluere deres reelle effekt.²⁾ Det er uetisk at eksperimentere med os på denne vilkårlige måde, hvor vi ikke finder ud af, om det skader mere, end det gavner.³⁾ Vi er uvidende forsøgskaniner. Ville det ikke være bedre med et mere formelt eksperiment, hvor nye politikker blev afprøvet og evalueret, inden de blev sluppet løs på os allesammen?

Vi kan kun vide, om vi fejler eller ej, hvis vi gennemprøver regeringstiltagene grundigt. Meget af det, vi laver i øjeblikket, ender faktisk med ikke at leve op til det, det bør. Eller ender endda med at forvolde skade. For eksempel virkede det i 1970'erne som en genial plan for den amerikanske "krig mod narko" at forsøge at styre teenagere væk fra kriminalitet ved at lade dem besøge fængsler og møde dømte kriminelle. Ideen var at chokere børnene og skræmme dem væk fra et kriminelt liv. Trods støtte fra amerikanske politistyrker og Det Hvide Hus gav programmer som D.A.R.E og Scared Straight bagslag.⁴⁾ I stedet for at skræmme børnene fik det kriminalitet til at ligne noget, der rent faktisk kunne betale sig, og

lovovertrædelserne steg i visse områder. Men det var først efter solide evalueringer – dvs. systematisk bedømmelse af eksperimenterende forskning – at vi var i stand til at finde ud af, at det ikke fungerede. Vi har brug for samfundsvidenskabens bedste redskaber, hvis vi vil komme til en forståelse af, hvorvidt – og hvordan – politikkerne virker, dvs. alt fra randomiserede, kontrollerede undersøgelser til andre forskningsmetoder, hvilket er beskrevet i en ny rapport fra The Alliance for Useful Evidence.⁵⁾

Hvad vil det sige at eksperimentere?

Der er en vis mængde forvirring i sprogbruget med hensyn til de *metoder*, der bruges ved undersøgelser af socialpolitikken. Interviews med britiske embedsmænd og ministre tilbage i 2003 afslørede en "tåge i forhold til nomenklaturen" mellem de forskellige mekanismer til test af politikker: "trailblazers" (banebrydere), "policy trials" (politikforsøg), "pioneers" (pionerer), "prototypes" (prototyper) eller "benchmarks" (referencemærker), "randomised control trials" (randomiserede, kontrollerede undersøgelser) og andre udtryk blev slynget rundt, uden at det var klart, om de allesammen var rigtige eksperimenter eller ej:

*"Det er ikke kun afdelingerne af de britiske administrationer, der bruger forskellige etiketter; det gør underafdelinger af den samme afdeling eller administration også, hvilket øger forvirringen. De fantasifulde betegnelser for de tidligste evalueringsmetoder til politiske initiativer er mangedoblet – faktisk i en sådan grad at en af de ministre, vi har interviewet, kunne berette, at han eller hun engang fik frit valg til at vælge den betegnelse, som han eller hun bedst kunne lide til et pilotforsøg ud fra en vifte af rivaliserende – men alle lige upassende – muligheder."*⁶⁾

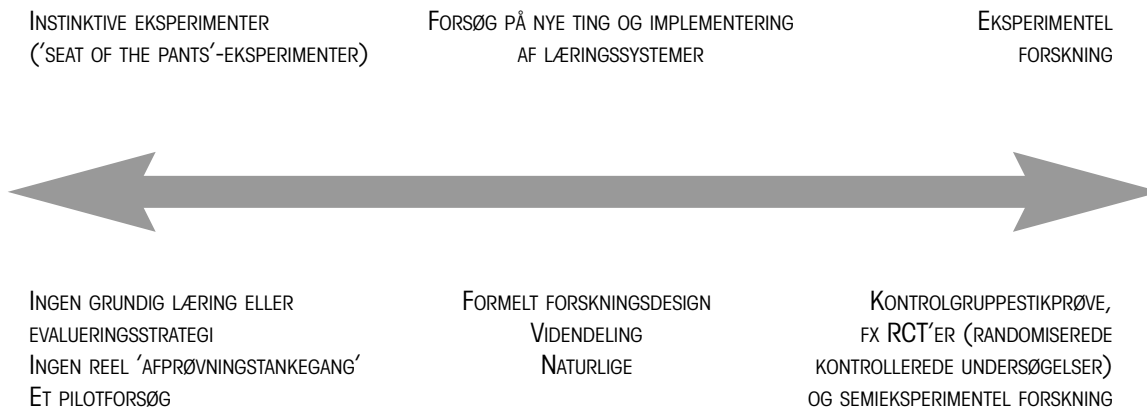
En ordbogsdefinition af "experiment" er et godt udgangspunkt:

"A test or investigation, especially one planned to provide evidence for or against a hypothesis" (dansk oversættelse: "En test eller undersøgelse; især en, der har til formål at tilvejebringe beviser for eller imod en hypotese.")⁷⁾

Ifølge vores definition af eksperimentel regeringsledelse betragtes politisk udvikling som en række bevidste hypotesetest, hvor resultaterne bruges til at gøre de efterfølgende beslutninger mere kvalificerede.

Lad os gøre det klart, hvad vi ikke mener med "at eksperimentere".⁸⁾ Vi er ikke interesserede i "intuitive/instinktive tiltag i den offentlige politik"; dvs. at afprøve nye politiske tiltag og programmer uden at have nogen meningsfuld metode til at vurdere deres virkninger. Det er ikke "løssluppen prøven sig frem eller spontan politikudbredelse" – en definition, der undertiden anvendes i slappere eksperimenterende praksis.⁹⁾

Solide og velegnede forskningsteknikker til at lære af eksperimenterne er det afgørende. Det er bedst at betragte eksperimenterne, som om de befinder sig på et kontinuum, der går fra den uordentlighed, der kendetegner alle de forsøg, som regeringerne og de velegnede organisationer gør på at afprøve nye ideer, og helt over til de eksperimenter, der udføres ved hjælp af de bedste tilgængelige forskningsteknikker, såsom randomiserede, kontrollerede undersøgelser eller andre teknikker. Denne artikel taler for, at man flytter sig til højre på dette kontinuum.



Flytning mod kontinuumets højre side ved at bruge solide metoder til at evaluere virkningen af eksperimenterne – især dem, der bruger kontrolgrupper, såsom (hvis det er hensigtsmæssigt) RCT'er eller semi-eksperimentelle forskningsmetoder.¹⁰⁾

Men eksperimentelle forskningsmetoder, såsom RCT'er, kan være den forkerte måde at måle virkningen på. Vi slår *ikke* til lyd for at bruge en "metodestyret" tilgang, hvor man kun starter med én enkelt metode, som fx randomiserede, kontrollerede undersøgelser, og så anvender den for at forstå konsekvenserne af tiltaget. Det er meget bedre at tænke på, hvor hensigtsmæssige ens metoder egentlig er. Det er ikke nok at have et forskningsdesign af fremragende teknisk kvalitet. Du kunne fx have en solid, kontrol-

leret undersøgelse, som bare ikke fortæller dig noget om, hvorvidt det samme indgreb kunne fungere andre steder. Som den tidligere formand for The European Evaluation Society, Professor Elliot Stern, udtrykte det: "Vi må tage udgangspunkt i det, vi ønsker at vide om programmerne, snarere end i en bestemt værktøjskasse, uanset hvor moderigtig den end måtte være".¹¹⁾ Andre designs som fx dem, der anvender statistiske, kvalitative, teori-baserede og andre tilgange, kan være meget mere anvendelige til at opnå en forståelse af virkningerne.¹²⁾

Det er også vigtigt at gøre opmærksom på, at vi ikke altid ønsker at springe direkte til højre side af kontinuumet; visse nye politiske initiativer er måske ikke egnede til eksperimenter, hvis der endnu ikke er

nogen klar hypotese eller sammenhængende aktivitet. Vi skal undgå faren for at have alt for høje bevistærskler for nye aktiviteter i forhold til mere etableret socialpolitisk arbejde, der er blevet forfinet over flere år. I nogle tilfælde er det simpelthen bare for tidligt at afgøre, hvad der må testes, og det kan derfor være nødvendigt at lade det vokse, før man designer en robust test.

En måde, hvorpå man undgår at generalisere, når man skal evaluere eksperimenter, er at bruge dokumentationsstandarder – en metode, der anvendes af Nesta (se figur 1).¹³⁾ Disse dokumentationsstandarder betyder, at man begynder på et grundlæggende niveau, hvor man kan beskrive, hvad man gør, og hvorfor det har betydning – på en logisk, sammenhængende og overbevisende måde. Man bevæger sig derefter videre op på skalaen til de områder, der befinder sig oppe ved niveau 3, hvor man kan have brug for ekspe-

rimentel forskning og større sikkerhed mht. kausalitet, og derefter videre til dokumentation for replikation og skalering. Bevisniveauerne er således afpasset de forskellige stadier, der findes i udviklingen af et tiltag.

Figur 1. Nestas "dokumentationsstandarder"



At være mere ydmyg i socialpolitikken

Vi vil også gerne se en tilgang, der accepterer en vis ydmyghed. Vi behøver ikke færdige svar, og det kan være nødvendigt at lære lidt efter lidt. Vores kollega Hasan Bakhshi har argumenteret overbevisende for en "eksperimentel tilstand", dvs. en forvandling af den erhvervsmæssige innovationspolitik, hvor regeringen ser bevidst hypotesetestning og udveksling af data som et af sine nøgleindgreb:

"Ethvert eksperiment er en hypotesetest, som nødvendigvis må foregå i en kontekst af usikkerhed, omkring hvad der er eller ikke er sandt. Det starter med en formodning om uvidenhed, og man forsøger at reducere denne ved at bringe in-

formationer for dagen (eksperimentet), og – med hensyn til offentligt motiverede eksperimenter – offentliggøre resultaterne. Pointen er ikke at bevise, at noget er korrekt – en kritik, der også er blevet rettet mod de modeller, der fordrer udpegning af vindere – men snarere at opdage det, man ikke tidligere vidste, og derefter bruge denne nye viden til videre udforskning.¹⁴⁾

Det er vigtigt, at vi bruger solide metoder til at evaluere eksperimenterne, som fx randomiserede, kontrollerede undersøgelser eller semi-eksperimentelle forskningsmetoder.¹⁵⁾ Stærke evalueringsteknikker betyder, at regeringerne kan se, om eksperimentet fungerede eller ej, snarere end at skulle stole på mavefølelser, selvom politikere og regeringer altid vil være nødt til at afbalancere beviserne med andre politiske hensyn i deres beslutningstagning.

At tænke småt

En vigtig del af det at eksperimentere er at starte i det små. Dvs. at undgå de politiske tiltag, der er ”for store til at kunne mislykkes”. Det er en tilgang, der tillader os at være ”innovative og forsigtige på samme tid [...] at afprøve tingene på en overdrevent forsigtig måde”, ifølge en amerikansk holdningserklæring om socialpolitiske eksperimenter, der var finansieret af The MacArthur Foundation.¹⁶⁾ Du kan teste din nye idé, og hvis den fejler, er det kun i begrænset omfang, hvilket jo gør skaden mindre. Men hvis innovationen lykkes, kan du udvide det til den større verden.

The Government Digital Service (hører under det britiske Cabinet Office, som svarer til det danske statsministerium) og The Behavioural Insights Team (som nu er uafhængigt af den britiske regering) har eksperimenteret med det formål at forbedre alt fra opfordringer til organdonation til udfyldelsen af de britiske skatteblanketter. I disse tilfælde begyndte de ”småt”, hvilket betød relativt lave omkostninger – både for det offentlige, men også med hensyn til den potentielle politiske skade, hvis forsøget var mislykkedes. Der er håndgribelige fordele. The Behavioural Insights Team, der hører til Nesta, har arbejdet sammen med den britiske regering for at be-

kæmpe forsinkede indkomstskattebetalinger. De lavede en enkelt lille ændring i regeringens standardbrev efter at have eksperimenteret online med en vifte af forskellige kommunikationsstile. De fandt ud af, at den bedste måde at tilskynde folk til at betale skat på var et simpelt tiltag, der bestod af denne sætning: ”9 ud af 10 mennesker har allerede betalt deres skat”. Fremhævelsen af det, som socialpsykologer kalder normativ adfærd – at gøre det, som andre gør – førte altså til en markant stigning i antallet af de hurtige skattebetalinger. Ændringen af skatteblanketterne fremskyndede, hvad der svarer til over 2 milliarder kroner i ekstra skatteindtægter til den britiske regering.¹⁷⁾

I USA er det almindeligt, at delstatsregeringerne afprøver nye politiske initiativer. Den forbundsstruktur gør det muligt for de enkelte delstater at opstille deres egne forsøg i ret stor skala, så man kan teste de sandsynlige virkninger af en foreslået ny politik eller leveringsmekanisme eller begge dele. En oversigt i 1997 fandt 240 igangværende forsøg i de amerikanske delstater på tværs af alle områder af social politik, herunder velfærd, skolegang og misbrug/afhængighed.¹⁸⁾

Mange eksperimenter skal begynde i det små for at begrænse de politiske og økonomiske omkostninger, der kan være for regeringen, hvis et forsøg mislykkes. Men størrelse er et relativt begreb. ”Småt” for USA og Kina kan fx forekomme meget stort for andre. På trods af Kinas stærkt centraliserede, bureaukratiske og hierarkiske regering, har man kunnet eksperimentere på en måde, der har været relativt ”bottom-up”, men på store økonomiske politikker:

”Selvom ambitiøs planlægning fra centralstaten, prægtige teknokratiske moderniseringsordninger og storprojekter aldrig har været forsvundet fra Kinas politiske dagsorden, så har en etableret proces med eksperimenter, som går forud for vedtagelsen af mange nationale politikker, fungeret som en stærk korrektionsmekanisme.”¹⁹⁾

Det vigtigste ved at starte småt er ikke den fysiske størrelse, men derimod en etik, der går ud på at begynde med at afprøve

idéerne, og som giver plads til tilpasning og læring.

Konklusion: Vi er nødt til at give politikerne noget spillerum. Det er afgørende, at vi afprøver nye idéer. Lad lovende innovationer blomstre – også selvom de er uprøvede og dermed udokumenterede. Faren ved at nedskyde fremsatte ideer, fordi ”der ingen beviser er”, er, at vi kvæler enhver, der ønsker at finde en vej ud af status quo.

Det nuværende politiske klima er modent til eksperimenter

At starte småt kan være uundgåeligt, hvis der ikke er tilstrækkelige midler til at introducere en politik på nationalt niveau. Faktisk byder ressourceknaphed også på muligheden for at eksperimentere med de sociale politikker. Vi kan prøve ting af på visse steder for at se, om de fungerer eller ej.

I det mexicanske socialbistandsprogram Progresas eksperimenterede man kun i visse lokalsamfund – delvist på grund af budgetrestriktioner. Programmet gjorde brug af betingede kontantoverførsler: Man tilbød betalinger til fattige familier, hvis de opfordrede deres børn til at gå i skole regelmæssigt eller til at besøge sundhedsklinikker. Progresas bygger på en simpel idé: I stedet for bare at yde overførselsindkomst til de fattige gennem forskellige tilskud til mad, transport, elektricitet og lignende er det bedre at yde overførselsindkomsten direkte, efter pengemæssige begreber. Ud over kontanter sørgede programmet for naturalydelser i form af sundhedsmæssige fordele og kosttilskud. Det var underlagt en streng evaluering: 320 lokalsamfund blev randomiseret som behandlingsgruppe, mens 186 lokalsamfund fungerede som kontrolgruppe. Programmet er vokset, og det er blevet opretholdt under skiftende politiske regimer: Det påbegyndtes under præsident Zedillo, derefter overtog Fox, og det anvendes stadig i dag. I 2005 hjalp det 5 millioner familier, dvs. næsten alle de mexicanske husholdninger, der lever i ekstrem fattigdom.²⁰⁾

Decentralisering af de offentlige tjenester giver flere muligheder for at eksperimentere, ifølge en erfaren eksperimentel forsker, Professor Gerry Stoker fra Southampton Universitet i Storbritannien.²¹⁾ Overdragelse af

magt fra centralregeringen til skoler, lokale myndigheder og førende fagfolk skaber mulighed for at afprøve forskellige metoder.

Eksperimenter kan faktisk bedst udføres på lokalt niveau. Centralregeringen finder det nemlig næsten umuligt at gøre det ordentligt, ifølge en evaluering fra 2003, som den britiske regering selv foretog af sine pilotforsøg.²²⁾ For mange nationale politikker er simpelthen mejset i sten. Manifestforpligtelser, som er blevet fulgt nøje gennem medierne, og som er blevet drøftet i parlamenterne, kan betyde, at der er investeret for meget politisk kapital til, at der rigtigt er mulighed for forandring.

Er eksperimenter kun relevante i deres specifikke sammenhænge?

Faren ved kun at eksperimentere i et enkelt lokalområde er naturligtvis, at det kan være svært at generalisere med hensyn til resten af verden. Det, som virker i København eller London, virker muligvis ikke i Madrid.

Det kunne jo godt tænkes, at de fordele, der opnås i de lokale eksperimenter, blot skyldes konteksten og humangeografien i disse områder. I deres bog *Evidence-Based Policy: A Practical Guide to Doing It Better*, argumenterer filosofen Professor Nancy Cartwright og økonomen Jeremie Hardie for, at dette er en grundlæggende fejl ved de randomiserede, kontrollerede forsøg og lignende redskaber. Det er i praksis meget svært at være sikker på, at det, der var ”effektivt der, også vil fungere her”.

I lyset af den åbenbare succes med Progresas blev et lignende indgreb afprøvet i USA af borgmester Bloomberg: Opportunity NYC (der finansieredes af private kilder). Kunne de positive effekter, der oplevedes i landdistrikterne i et udviklingsland (Mexico), overføres til et byområde i den rigeste nation i verden? Indtil videre lader svaret til at være nej. Eksperimenterne med de betingede kontantoverførsler er slet ikke begyndt endnu i New York, fordi det stadig anses for uklart, hvorvidt den nye version af modellen er effektiv; bevismaterialet til fordel for den indledende model kunne ikke leve nok op til tidligere løfter til at retfærdiggøre politisk vedtagelse.

Den bedste reaktion på ovenstående er at

gentage forsøgene flere steder (og endda på samme sted for at se, om det ikke bare var en engangsforestilling). Gentagelse er en vigtig del af den eksperimentelle videnskab og samfundsvidenskaben.²³⁾ Det bør derfor være en vigtig del af den eksperimentelle regering.

En anden måde at omgå problemet med den lokale kontekst på er at foretage meget større eksperimenter – og bruge internettet. Behavioural Insights Teamets test til at fremme or-gandonation udførtes på 1 million mennesker, der brugte webstedet GOV.UK. Altså ikke én bestemt placering i Storbritannien, men ved hjælp af landet som helhed. Den stigende anvendelse af internettet og de sociale medier i den offentlige service giver mulighed for flere eksperimenter – i stor skala og både billige og hurtige – som fx den randomiserede, kontrollerede undersøgelse af vælgermobilisering, der nåede 61 millioner brugere under valget til den amerikanske kongres i 2010. Virksomheder som fx Facebook, Capital One og Google foretager eksperimenter af global karakter.

Den lokale kontekst vil altid være en afgørende faktor, når vi skal lære af eksperimenterne. Men den er også en problemstilling ved enhver form for evaluering eller enhver implementering af en politik (en politik, der er effektiv i et område, kan være fuldstændig upassende i en anden). At eksperimenterer er også en proces, der aldrig stopper. Vi bør altid blive ved med at teste og afprøve nye politikker – selv efter de er blevet indført på nationalt niveau.

Konklusion

Det er lige nu, der er behov for de eksperimenterende regeringer. Komplekse samfundsudfordringer som følge af klimaforandringer, det demografiske pres eller store folkevandringer betyder, at det ”som vi plejer” højst sandsynligt ikke vil være godt nok til at tage de udfordringer op, som det enogtyvende århundrede byder os. Ifølge én efterlysning af flere eksperimenter vil det for magthaverne være en realistisk fremgangsmåde, eftersom den ”muliggør eventyrlyst og forsigtighed på én og samme tid”.²⁵⁾ I stedet for ubesindigt at forfølge nye papirbase-rede idéer, der ender med politiske problemer, har vi nu metoden til at teste dem.

Overdreven løfterigdom dominerer desværre den politiske debat og dem, der søger ny finansiering til deres velgørende organisationer. Det

kan være svært at forestille sig politikere, der åbent støtter eksperimenterne på de politiske talerstole. Men det er ikke umuligt. Tre af de vigtigste politiske partier i Finland, som også går et valg i møde, har formelt forpligtet sig til eksperimenter i deres programmerklæringer.²⁶⁾ Et af disse manifestløfter kommer fra Centerpartiet – som på nuværende tidspunkt er foran i meningsmålingerne – og de har endda planer om at ændre forfatningen for at tillade eksperimenterne.

I en anden periode med enormt socialt og økonomisk pres, nemlig 1930’erne, efterlyste USA’s præsident Franklin D. Roosevelt ”dristig, vedholdende eksperimenteren” i The New Deal. Han sagde:

*”Landet har brug for og – medmindre jeg fejlbedømmer dets temperament – kræver landet dristig, vedholdende eksperimenteren. Det er sund fornuft at tage en metode og afprøve den: Hvis den slår fejl, så indrøm det ærligt, og prøv med en anden. Men frem for alt så forsøg noget. De millioner af mennesker, der er i nød, vil ikke stå og se stiltiende på i al evighed, mens de ting, der kan tilfredsstille deres behov, er inden for let overkommelig rækkevidde.”*²⁷⁾

Franklin D. Roosevelt udpegede en vej med ambitiøst og tempofyldt lovgivningsmæssigt arbejde, som sendte 15 store lovforslag til Kongressen i hans første 100 dage, og som var blandet med en åbenhed for at afprøve tingene gennem solid eksperimentel forskning. I 1940 arbejdede 8.000 samfundsforskere for den føderale regering. Og Roosevelt kaldte en af de førende fortalere for at anvende kontrollerede undersøgelser på regeringsprogrammerne, Professor Charles Meriam, for sin ”onkel Charley”.²⁸⁾

Men vi kan ikke forlade os på vores ledere alene. Det kan være op til os vælgere at acceptere, at vi ikke har alle svarene. Og det kan være nødvendigt, at vi tilmelder os eksperimenter på skoler, på hospitaler og i vores kvarterer. Det er ikke en nem ting at bede om, men det er bedre end at pløje videre, som vi plejer, uden hensyn til følgerne.²⁹⁾

Flere institutioner er også begyndt at støtte den eksperimentelle forskning, i form af What Works-Centrene, som er medfinan-

sieret af det britiske Economic and Social Research Council samt de britiske, walisiske og skotske regeringer, der dækker alt fra udannelse til nedbringelse af kriminalitet. The What Works Centre for Local Economic Growth (et samarbejde mellem London School of Economics and Political Science, Centre for Cities og Arup) gennemfører deres egne eksperimenter – de såkaldte “demonstrationsprojekter”. Det britiske Cabinet Office har også oprettet en forsøgsrådgivningsservice til at hjælpe de offentlige ansatte med at designe gode eksperimenter.

Disse praktiske foranstaltninger vil alle være en hjælp, men vi har brug for en politisk kulturforandring. I sidste ende er det, vi gerne vil opnå, mere ydmyghed og færre falske løfter i manifesterne. Hvis dataene og dokumentationen ikke retfærdiggør de politiske beslutninger, bør vi gøre opmærksom på det. Og hvis vi ikke har nogen dokumentation, bør vi forpligte os til at eksperimentere i fremtiden.

Artiklen er oversat fra engelsk til dansk af oversættelsesbureauet Diction.

Noter:

1. Making Evidence Useful: The case for new institutions, Nesta d. 28. februar 2013
2. Evaluering i regering, National Audit Office, december 2013
3. Se for eksempel: “*Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations*”, Annals of American Academy for Political and Social Science, vol. 589, s. 22-40
4. Se for eksempel denne systematiske gennemgang af dokumentationen mht. ”Scared Straight” og andre ungdomsoplysningsprogrammer til forhindring af ungdomskriminalitet. <http://www.Cochrane.org/features/Scared-straight-and-other-Juvenile-Awareness-programmes-Preventing-Juvenile-Delinquency>
5. Alliance for Useful Evidence: “Better Public Services Through Experimental Government”. Marts 2015. <http://www.alliance4usefulevidence.org/publication/Better-Public-Services-through-Experimental-Government/>
6. *Trying it out; The Role of ‘Pilots’ in Policy-Making. Report of a Review of Government Pilots (Cabinet Office, 2003)*
7. Collins English Dictionary, 3. udgave
8. *Social Policy Experimentation; A Position Paper*. Berk et al. s. 390. Evaluation Review Vol. 9(4) August 1985
9. *Policy Experimentation in China’s Economic Rise*. Sebastian Heilmann, Studies in Comparative International Development (2008) 43:1–26. s. 3
10. Mere om kvasi-eksperimentelle metoder kan findes i denne rapport til Europa-Kommissionen, som er udarbejdet af The Alliance for Useful Evidence sammen med JPAL-Europe og London School of Economics: *Guide to Testing Social Policy Innovation*, Europa-Kommissionen, juni 2014 <http://www.alliance4usefulevidence.org/publication/spark-guide-to-testing-social-policy-innovation/#sthash.SbrjWEoC.dpuf>. Som ikke-tekniske vejledninger til politiske beslutningstagere i de eksperimentelle forskningsteknikker anbefales: *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials (Behavioural Insights Team, 2012)*
11. Stern, E. (udgives snart) *Impact Evaluation; A Design Guide for Commissioners and Managers of International Development Evaluations In the Voluntary and Community Sector*. Udarbejdet for Bond, Comic Relief, Big Lottery Fund og DFID s. 4
12. Som en glimrende oversigt over de forskellige tilgange anbefales: Medical Research Council (2008), *Developing and evaluating complex interventions: new guidance*. <http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>
13. Mulgan, G., (2014) “Innovation in the Public Sector; How can public organisations better create, improve and adapt?” London: Nesta. s. 11
14. Hasan Bakhshi, Alan Freeman & Jason Potts, Nesta: State of Uncertainty; Innovation policy through experimentation.
15. Som ikke-tekniske vejledninger til politiske beslutningstagere om de eksperimentelle forskningsteknikker anbefales: *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials (Behavioural Insights Team, 2012, eller Nestas og Ben Goldacres gør-det-selv-hjemmeside (D.I.Y. website)*. Mere om kvasi-eksperimentelle metoder kan findes i: *Guide to Testing Social Policy Innovation*, Europa-Kommissionen, juni 2014 <http://www.alliance4usefulevidence.org/publication/spark-guide-to-testing-social-policy-innovation/#sthash.SbrjWEoC.dpuf>
16. Berk RA, Boruch RF, Chambers DL, Rossi PH & White AD (1985) “Social Policy Experimentation: A Position Paper”, *Evaluation Review*, 9(4), s. 387-429
17. I-teams: The teams and funds making innovation happen in governments around the world. Nesta & Bloomberg Philanthropies, 2014 s. 18
18. *The Digest of Social Experiments* Greenberg & Shroder, 1997. Tredje udgave.
19. *Policy Experimentation in China’s Economic Rise* Sebastian Heilmann, Studies in Comparative International Development (2008) 43:1–26
20. Progress Against Poverty. Sustaining Mexico’s Progresá-Oportunidades Program, Santiago Levy, Brookings Institution, 2006
21. <http://Ann.sagepub.com/content/628/1/47.Short>
22. *Trying it Out; The Role of ‘Pilots’ in Policy Making; Report of a Review of Government Pilots*, Cabinet Office, 2003
23. *Evidence-Based Policy: A Practical Guide to Doing It Better*, Nancy Cartwright & Jeremy Hardie (2012)
24. *Social science and replication*, Den 11. december 2013. Blog af Alex Sutherland & Nicole Janz, University of Cambridge <http://www.alliance4usefulevidence.org/replication/>
25. Berk RA, Boruch RF, Chambers DL, Rossi PH & White AD (1985) “Social Policy Experimentation: A Position Paper”, *Evaluation Review*, 9(4), s. 414
26. <http://www.Nesta.org.uk/blog/Finlands-Political-parties-Embrace-experimentalism>
27. Roosevelt FD (1932) “Address at Oglethorpe University in Atlanta, Georgia”, [Online] The American Presidency Project. Tilgængelig på: <http://www.presidency.ucsb.edu/ws/?pid=88410>
28. The Origins of American Social Science, Ross, 1991, s. 455
29. <https://Quarterly.blog.gov.uk/2015/01/27/What-Works-The-Rise-of-Experimental-Government/>