

RAY PAWSONS "REALIST EVALUATION"

Seniorforsker Håkon Finne, SINTEF Teknologi og samfunn

Evalueringskonferansen 2016-09-30, Trondheim



Planen

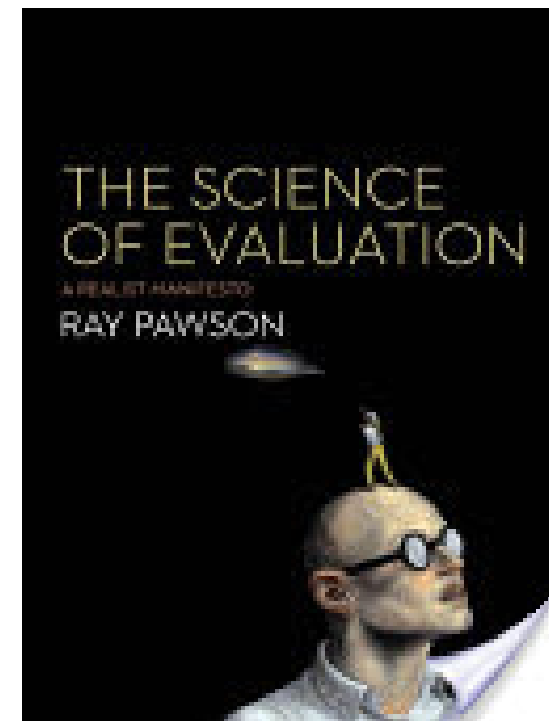
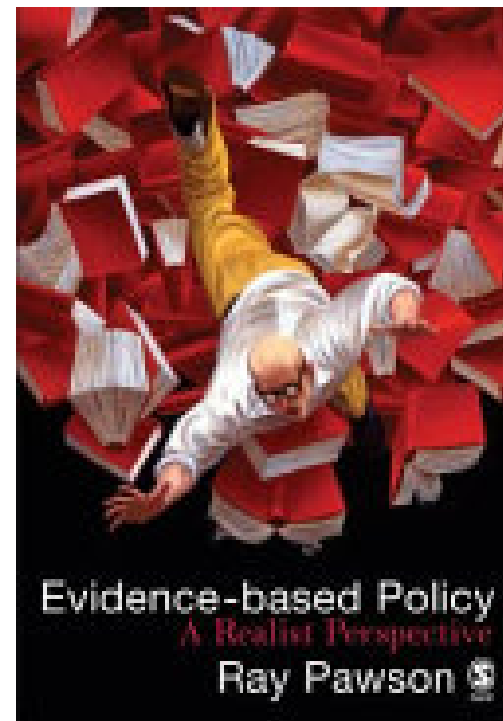
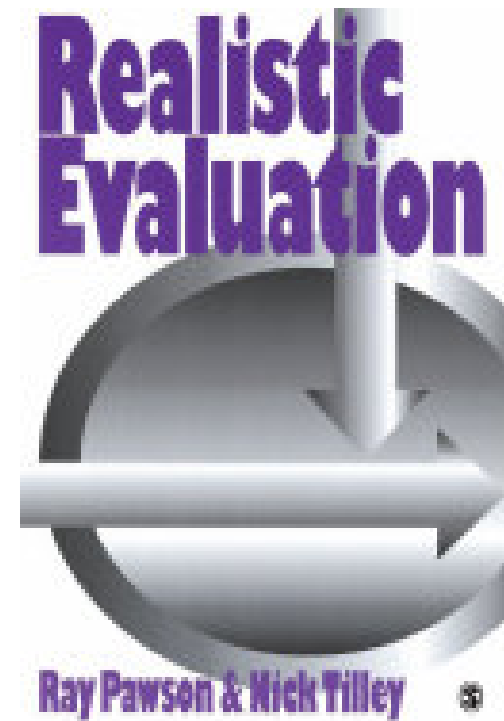
1. Prosjektet: Om evaluering, Ray Pawson og meg selv
2. En sørgelig tilstand
3. Problemet er ikke der vi har trodd
4. Et real-ist-isk alternativ
5. Metodologiske komplekser
6. Fra evaluering til evidens for politikkutforming
7. Jeg har en plan! – eller?



1. Prosjektet

- Evaluering: av effekter av tiltak/programmer/reformer i en sosial virkelighet
- Ray Pawson: Professor i samfunnsvitenskapelig forskningsmetode, University of Leeds
- Håkon Finne: Seniorforsker i SINTEF Teknologi og samfunn, praktiserende evalueringsforsker
- Metode: HF undersøker hva RPs "real-ist-ic evaluation" er

Datamateriale





2. En sørgelig tilstand

- Programeffekter: Små, inkonsistente, lite overførbare: Vi vet ikke hva som virker
- Programevalueringer: Like sprikende, metodiske svakheter, akkumulerer ikke kunnskap, status som annenrangs forskning
- Selv omfattende eksperimentelle design gir inkonsistente resultater
- Problematisk overgang fra enkeltstudier til policyformulering
- Det burde ikke være slik! ?



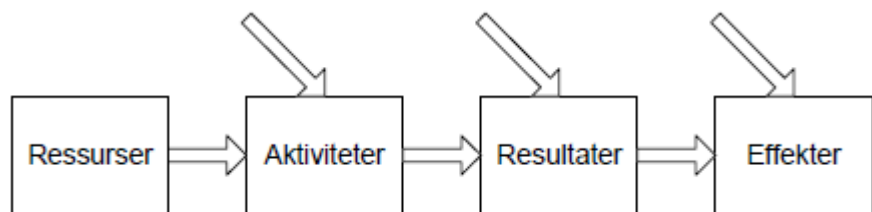
Effektevaluering og programteori i vekst

- Planlegge effekter, underveisevaluering for å se om man er på rett spor, måle effekter i etterkant og når de har satt seg
- Vellykkethet avhengig av målstruktur, programlogikk, resonnementkjeder, programteori
- Økende fokus på "evidens" for "virker tiltaket som tiltenkt"?

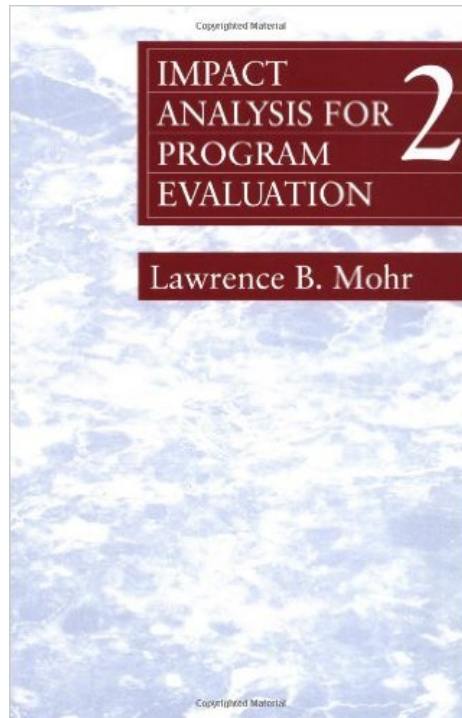
Metoder kommer stadig nærmere gullstandarden

$$\frac{O}{O} - \frac{X}{O} - \frac{O}{O}$$

- Randomisert, kontrollert eksperiment, med kontrollgruppe, helst dobbelt blindt, måler utfallsvariabelen før og etter intervensjonen i begge grupper
- Donald T Campbell: Mange sofistikerte, alternative design og metoder for å kompensere for validitetstrusler pga kompleksitet i sosiale systemer



Number	Chapter	Name	Diagram
1	4	one-shot case study	$T \ Y$
2	10	ex post facto	assorted
Elementary quasi-experimental designs			
3	4	before-after	$A/C: Y_1 \ T Y_2$
4	4	comparative posttest	$(A): \frac{T \ Y_E}{A: \ Y_C}$
Quasi-experimental designs			
5	7	comparative change	$(A)/C: \frac{X_{1E} \ T \ Y_E}{A/C: \ X_{1C} \ Y_C}$
6	9	interrupted time series	$A/C: Y_1 \ \dots \ T Y_{T+1}$
7	9	comparative time series	$(A)/C: \frac{Y_{1E} \ T \ Y_{(T+1)E}}{A/C: \ Y_{1C} \ Y_{(T+1)C}}$
Fortified designs			
8	8	criterion population	$(A)/C: \frac{X_{1E} \ T \ Y_E}{A/C: \ X_{1CP} \ Y_{CP}}$
9	11	subobjective	assorted, based on $T \ S \ Y$
Experimental designs			
10	6	regression discontinuity	$A/C: \frac{A_E \ T \ Y_E}{A/C: \ A_C \ Y_C}$
11	4	R-comparative posttest	$R: \frac{T \ Y_E}{R: \ Y_C}$
12	4	R-comparative change	$R: \frac{X_{1E} \ T \ Y_E}{R: \ X_{1C} \ Y_C}$



3. Problemet er ikke der vi har trodd

- Problemet er ikke at metodene ikke kommer nær nok gullstandarden.
- Det er gullstandarden som er problemet.
- Fordi samfunnsvitere har misforstått hvordan eksperimenter fungerer i naturvitenskapene.

Tre kausalitetsmodeller

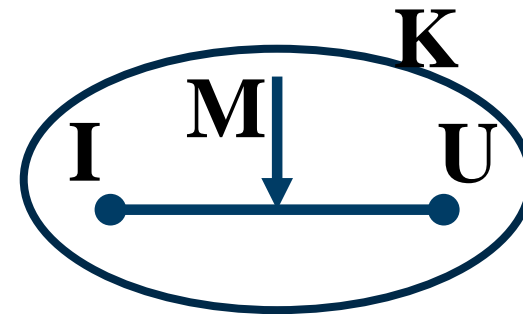
Successionist



Configurational



Generative



Kompleksitet i sosiale systemer

- Choice is the very condition of social and individual change and not some sort of practical hindrance to understanding that change

Derfor forskyves evalueringsspørsmålet

- Fra:
- Virker det?

- Til:
- Hva virker, for hvem, under hvilke betingelser, og hvordan?

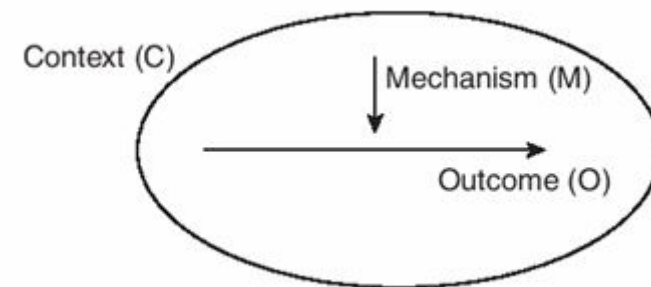


Fører iskrem til mer spinat?

- A certain mother habitually rewards her young son with ice cream after he eats his spinach. What additional information would you need to be able to predict whether the child will: a. Come to love or hate spinach, b. Come to love or hate ice cream, or c. Come to love or hate mother? (Bateson, 1972, p. xvii)

4. Et real-ist-isk alternativ

- Real: orientert mot virkeligheten, reelle problemer, ikke grunnforskning
- Realist: vitenskapsteoretisk posisjon mellom positivisme og konstruktivisme
- Realistic: realistisk, gjennomførbart
- Analytisk skjema: Mechanism, Context, Outcome



Overvåkningskamera (CCTV) utløser **mekanismer** mot tyveri på parkeringsplasser

- Tyven kan tas på fersk gjerning
- Potensielle tyver er redde for å bli videofilmet
- Bilførere foretrekker parkeringshus med CCTV
- Raskere politiutrykning der mistenkelig aktivitet filmes
- Potensielle tyver antar at her tar politiet tyveri alvorlig
- Forskyvning mot raskere tyveri (vanskeligere å oppdage)
- Minner bilførere på tyverifaren og tar forholdsregler
- Tyveribevisste bilførere prioriterer parkeringshuset; de mindre bevisste skyves ut til gateparkering

Kontekst påvirker hvilke mekanismer som utløses

- Enkeltskyver vs gjenger
- Bilførernes bruksmønster
- Biler i blindfelt
- Tilgang på alternative tyverimål
- Tilgang på oppfølgingsressurser
- CCTV nyhet eller innarbeidet

5. Metodologiske komplekser

- Mange spørsmål – mange metoder – refortolking viktig
- Ikke kast ut dine metode-babyer med Pawsons badevann, men vær tydeligere på hva de forutsetter og hva de hjelper til å forstå
- De implisertes egenforståelse er særdeles viktige for å forstå utfall
- Aksjonsforskningsbasert følgeforskning så godt som noe i formativ evaluering fordi det regisserer lokal læring og konsensusbygging
- Men ikke start fra scratch hver gang



6. Fra evaluering til evidens for policyutforming

- Metaevalueringer à la Campbell/Cochrane legger input/output-studier oppå hverandre og veker for avstand til den eksperimentelle gullstandarden
- Pawson: Ikke akkumuler programmer, men programteorier; identiske programteorier hentes fra mange domener
- Særlig synlig i review-tilnærming



7. Jeg har en plan! – eller?

- Oppgradering av evaluering som fag
- Bygg programteorier som CMO, ikke evalueringsdesign som OXO

- Mange skritt fra nå: fra evidens til politikkutforming



Teknologi for et bedre samfunn