Using Monitoring and Evaluation to Improve Public Policy

Philip Davies, PhD
American Institutes for Research
Oxford England and Washington, DC
Context of Policy Evaluation in the UK

• Modernising Government
• Better Policy Making
• Evidence-Based Policy and Practice
• Greater Accountability
• Performance Management
• Public Spending and Fiscal Control
• Strategic Development
Why Use M&E For Public Policy?

- Effectiveness - ensure we do more good than harm
- Efficiency - use scarce public resources to maximum effect
- Service Orientation - meet citizen’s needs/expectations
- Accountability - transparency of what is done and why
- Democracy - enhance the democratic process
- Trust - help ensure/restore trust in government and public services
The ‘Experimenting Society’
Donald T. Campbell

“...a society that would use social science methods and evaluation techniques to “vigorously try out possible solutions to recurrent problems and would make hard-headed, multidimensional evaluations of outcomes, and when the evaluation of one reform showed it to have been ineffective or harmful, would move on and try other alternatives” (Campbell, 1999a:9).
What is Evaluation?

A family of research methods which seeks “to systematically investigate the effectiveness of social interventions….in ways that improve social conditions”

(Rossi, Freeman and Lipsey, 1999:20)
Types of Evaluation

• Impact (or *summative*) evaluations
  
  *Does the policy (programme, intervention) work? How large is the likely effect size?*

• Process (or *formative*) evaluation
  
  *How, why, and under what conditions does the policy (programme, intervention) work?*
What is the Policy Question?
Effectiveness of What?

- *Intervention* effectiveness - what works?
- *Resource* effectiveness - at what cost/benefit?
- *Likely diversity* of effectiveness across different groups – what works for whom and when?
- *Implementation* effectiveness - how it works?
- *Experiential* effectiveness - public’s views of policy
How is the policy supposed to work?

Logic Model

Theories of Change

Evidence for Policy
Establishing the Policy Logic/Theory of Change

**Programme Theory**

- Visit to a Prison by Juveniles
- First Hand Experience of Prison Life
- Exposure to Prison Life and Prisoners as Negative Role Models
- Frightens or Scares Juveniles Away from Crime
- Reduces Crime and Offending

**Programme Evidence**

- Visit to a Prison by Juveniles
- First Hand Experience of Prison Life
- Exposure to Prison Life and Prisoners as Positive Role Models
- Stimulates or Attracts Juveniles Towards Crime
- Increases Crime and Offending
How is the policy supposed to work? 

Logic Model

What evidence already exists?

Systematic Reviews

Theories of Change

Harness Existing Evidence

Evidence for Policy
Evidence for Policy

Theories of Change

Harness Existing Evidence

Logic Model

What evidence already exists?

Systematic Reviews

Statistics

Surveys

Qualitative Research

Descriptive Analytical Evidence

What is the nature, size and dynamics of the problem?

How is the policy supposed to work?
Evidence for Policy

Theories of Change

Harness Existing Evidence

Descriptive Analytical Evidence

Attitudinal and Experiential Evidence

How is the policy supposed to work?

Logic Model

What evidence already exists?

Systematic Reviews

What is the nature, size and dynamics of the problem?

Statistics
Surveys
Qualitative Research

How do citizens feel about the policy?

Surveys
Qualitative Research
Observational Studies
How is the policy supposed to work?

Logic Model

What evidence already exists?

Systematic Reviews

What is the nature, size and dynamics of the problem?

Statistics
Surveys
Qualitative Research

How do citizens feel about the policy?

Surveys
Qualitative Research
Observational Studies

Evidence of Effective Interventions (Impact)

Experimental and Quasi-Experimental Studies

Evidence for Policy

Descriptive Analytical Evidence

What Works?
At What Costs?
With What Outcomes?

Evidence of Existing Evidence

Harness Existing Evidence

Experimental and Quasi-Experimental Studies

How is the policy supposed to work?

Logic Model

What evidence already exists?

Systematic Reviews

What is the nature, size and dynamics of the problem?

Statistics
Surveys
Qualitative Research

How do citizens feel about the policy?

Surveys
Qualitative Research
Observational Studies
Evidence for Policy Theories of Change

How is the policy supposed to work?

Logic Model

What evidence already exists?

Systematic Reviews

What is the nature, size and dynamics of the problem?

Statistical Surveys Qualitative Research

Harness Existing Evidence

Descriptive Analytical Evidence

Evidence for Policy

Evidence of Effective Interventions

What Works? At What Costs? With What Outcomes?

Experimental and Quasi-Experimental Studies

Economic and Econometric Evidence

Cost-Benefit Cost-Effectiveness Cust-Utility Analysis

How do citizens feel about the policy?

Surveys Qualitative Research Observational Studies

Evidence of Effective Interventions

Attitudinal and Experiential Evidence

What is the Cost, Benefit and Effectiveness of Interventions?
Evidence for Policy Theories of Change

- How is the policy supposed to work?
- What evidence already exists?
- What is the nature, size and dynamics of the problem?

Harness Existing Evidence

- Systematic Reviews

Descriptive Analytical Evidence

- Statistics
- Surveys
- Qualitative Research

Attitudinal and Experiential Evidence

- Surveys
- Qualitative Research
- Observational Studies

Evidence of Effective Interventions

- Experimental and Quasi-Experimental Studies


Ethical Evidence

- Social Ethics
- Public Consultation

Cost-Benefit Cost-Effectiveness Cust-Utility Analysis

- What are the ethical implications of the policy?

Economic and Econometric Evidence

- Cost-Benefit
- Cost-Effectiveness
- Cust-Utility Analysis

- What is the Cost, Benefit and Effectiveness of Interventions?
Evaluation Evidence in The Policy Process (Linear Model)
Evaluation Evidence in The Policy Process

Evidence-Based Policy

Ideas

Evaluation

Policy Development

Evaluation

Policy Implementation
Impact Evaluations

- Evaluations of Outcome Attainment (have targets been met?)
- Evaluations of Net Effects (against a counterfactual)

- Single Group Pre- and Post- Tests
- Interrupted Time Series Designs
- Matched Comparisons Designs
- Difference of Differences
- Propensity Score Matching
- Regression Discontinuity Designs
- Randomised Controlled Trials

Increasing strength of internal validity and causal inference
Evaluations of Outcome Attainment (Have Targets Been Met?)

**Policy Delivery: trajectories**

- **Low Trajectory**: policy has a lagged impact
- **Mid trajectory**: averaged impact of the policy
- **High Trajectory**: policy has an immediate impact

### Evaluations of Outcome Attainment

- **Intermediate progress indicators or milestones**
- **Historical performance**
- **Project Plan Streams**
- **Long Term Strategic Goal**
- **Mid term Delivery Contract Goal**

Graph showing delivery indicators from 1996 to 2010 with policy steps A, B, and C, and delivery indicators ranging from 50 to 95.
Evaluations of Net Effects (Against a Counterfactual)

• Counterfactual: what would have happened without the program

• Need to *estimate* the counterfactual
  ➢ i.e. find a control or comparison group

• Counterfactual Criteria
  ➢ Intervention & counterfactual groups have identical characteristics on average,
  ➢ Only reason for the difference in outcomes is due to the intervention
Evaluations of Net Effects (Against a Counterfactual)
Quasi-Experimental Methods

Single Group Before and After Studies / Cohort Studies

Effect Size = \( O_2 - O_1 \)

Note: There is no counterfactual
Before and After Examples

- Agricultural assistance program
  - Financial assistance to purchase inputs
  - Compare rice yields before and after
  - Find fall in rice yield
  - Did the program fail?
  - Before is normal rainfall, but after is drought
  - Could not separate (identify) effect of financial assistance program from effect of rainfall
Quasi-Experimental Methods

Interrupted Time Series Design

Road Traffic Deaths, UK, 1950-1972

Road Traffic Act, 1967
Impact Evaluations - Interrupted Time Series Designs

• Evaluation of the impact of the Road Traffic Act 1967
• Many evaluations of medical and public health interventions
• Evaluation of literacy amongst primary school children
• Evaluation of alcohol licensing and crime
• Evaluation of street lighting and crime
• Evaluation of CCTV and crime
Quasi-Experimental Methods

Two Group Before and After-Studies/
Case Control Studies
Matched Comparison Design

Intervention Group (Cases)

Non Intervention Group (Controls)

Matched

Intervention

Outcome = O₁

No Intervention

Outcome = O₂

Effect Size = O₁ - O₂

Note: Counterfactual is O₂
Impact Evaluations - Matched Comparisons Designs

• Also used extensively in UK government policy evaluation
  e.g.
  • Home Office evaluation of Cognitive Therapy for Offenders
  • DWP evaluation of Employment Zones
  • DWP evaluation of Work-Based Learning for Adults (PSM)
  • DfES evaluation of Educational Maintenance Allowance (PSM)
Quasi-Experimental Methods

Regression Discontinuity Design

Regression Discontinuity Trial
With No Treatment Effects

Regression Discontinuity Trial
With an Effective Treatment

Assignment Variable Score

Assignment Variable Score
Indexes Are Common in Targeting of Social Programs

- Anti-poverty programs are targeted to households below a given poverty index
- Pension programs are targeted to population above a certain age
- Scholarships are targeted to students with high scores on standardized test
- CDD Programs are awarded to NGOs that achieve highest scores
Randomised Controlled Trial/Random Allocation Experiment

• The “gold standard” in impact evaluation
• Gives each eligible unit/individual the same chance of receiving the treatment/intervention
• Lottery for who receives benefit
• Lottery for who receives benefit first
• Requires allocation independent of service or policy providers
• Best when ‘blind’ or ‘double blind’ rarely possible in public policy/public service delivery
Randomised Controlled Trial/
Random Allocation Experiment

Eligible population

Baseline

R

Intervention group → Intervention → Outcome = O₁

Control group

No Intervention → Outcome = O₂

Effect estimate = ‘O₁-O₂’-
Counterfactual is O₂
Oportunidades

- National anti-poverty program in Mexico (1997)
- Cash transfers and in-kind benefits conditional on school attendance and health care visits.
- Transfer given preferably to mother of beneficiary children.
- Large program with large transfers:
  - 5 million beneficiary households in 2004
  - Large transfers, capped at:
    - $95 USD for HH with children through junior high
    - $159 USD for HH with children in high school
Oportunidades Evaluation

- Phasing in of intervention
  - 50,000 eligible rural communities
  - Random sample of 506 eligible communities in 7 states - evaluation sample
- Random assignment of benefits by community:
  - 320 treatment communities (14,446 households)
    - First transfers distributed April 1998
  - 186 control communities (9,630 households)
    - First transfers November 1999
ERA Demonstration Project

• What is the most effective and efficient way of:
  • Retaining low paid people in work?
  • Advancing low paid people in the labour market?
ERA Demonstration Project
Multi-Method Evaluation

- Integrated evaluation with policy development and implementation
- Evaluation of existing evidence
- Programme theory evaluation (evaluating logic model)
- Impact evaluation (R.C.T.)
- Implementation evaluation (different models)
- Local context evaluation (Qualitative and Quantitative)
- Qualitative evaluation (clients’ and employers’ perspectives)
- Economic Evaluation (CBA)

Morris, S., Greenberg, D., Riccio, J., Mittra, B., Green, H., Lissenburg, S., and Blundell, R., 2004
Contact

Philip Davies PhD
Executive Director AIR UK
Senior Research Fellow
American Institutes for Research

UK
2 Hill House
Southside
Steeple Aston
Oxfordshire OX25 4SD
Tel: +44 1869 347284
Mobile: +44 7927 186074

USA
1000 Thomas Jefferson Street, NW
Washington DC 20007
Tel: 202 403-5785
Mobile: 202 445-3640