Evaluating Social Programs

Maytree – Wellesley Public Policy Training Institute

“If you want to know what water is, don’t ask the fish”

-- Marshall McLuhan

John Stapleton, April 6th and 7th 2006
Why do we do evaluation and performance assessment?

• To see if our intentions and objectives translate into outcomes
  We can have the best of intentions but is the result the one we expected?

• To see if we are organized properly to be able to know whether things are working
  Would we know if things weren’t working?
Are there challenges? (YES!)

Evaluation can divert resources from programs

An exam isn’t always, “a happy recollection of work well done.”
Are there challenges? (YES!)

There are no simple measures of performance
“One size does not fit all”
Have you ever heard anyone say?

• “I have no idea what I do for this place”
• “They have no idea what really goes on here”
• Media: “Another interminable report” vs. “Amazingly, there was no report”
Have you ever heard anyone say?

“They went out and got a bunch of expensive consultants to tell them what any one of us could have told them for free!”

(My apologies to Saul Alinsky)
Have you ever heard anyone say?

• “I’m too busy doing what I’m supposed to be doing to have time to sit around and naval gaze – I’ll participate but don’t expect me to get that project done by the end of the month!”

• “They whitewashed the whole thing – now they’ll think that everything’s great around here – I can’t believe it”
Ultimately, Ben lost his job as a shepherd because he fell asleep every time he tried to count his sheep.
Part I: What is Evaluation?
Program evaluation is essentially a set of philosophies and techniques to determine if a program 'works'. It is a practice field that has emerged as a disciplined way of assessing the merit, value, and worth of projects and programs.

Extraordinary sums have been invested in social programs, but the means of knowing what happened, and why, have not always been available.
Behind the obvious: a host of other more complex questions. . .
What is a program supposed to do?

• It is often difficult to define what a program is supposed to do
• Indirect indicators may be used instead.
• For example schools are supposed to 'educate' people.
  – What does 'educate' mean? If the exact goal cannot be defined well, it is difficult to indicate whether the program 'works'.
There may be unintended or unforeseen consequences of a program.
• These unintended consequences may be as important as the intended consequences.

• Evaluations should measure not just whether the program does what it should be doing, but what else it may be doing.
• Perhaps the most difficult part of evaluation is determining whether it is the program itself that is doing something.

• There may be other processes that are really causing the outcome, or preventing the hoped-for outcome.
Many evaluations cannot determine whether it is the program itself, or something else, that is the 'cause'.
• One main reason that evaluations cannot determine causation involves self selection.
  – For example, in a jobs training program, some people decide to participate, and others do not participate

• Even without random assignment, evaluations can provide information.
  – "People who participate in program xyz were more likely to find a job, while people who did not participate were less likely to find a job."
When there is enough data, statistical analysis can be used to make a 'reasonable' case by showing that other causes are unlikely.
The evaluation can be used to analyze the program process. So instead of focusing on the outcome focus on what the program was doing. This information could help how the program was operating.
Program evaluation is often divided into types of evaluation and types of language.

Formative Evaluation

– occurs early in the program.
– results are used to decide how the program is delivered, or what form the program will take.
Formative Evaluation

Example:

An exercise program for elderly adults would seek to learn what activities are motivating and interesting to this group. These activities would then be included in the program.
Process Evaluation

• Concerned with how the program is delivered.
• Deals with things such as when, and where the program activities occur, and who delivers them.
• Is the program being delivered as intended?
  • An effective program may not yield desired results if it is not delivered properly.
Outcome Evaluation

- Addresses the question of what are the results.
- It is common to speak of short-term outcomes and long-term outcomes.

Example:
In an exercise program, a short-term outcome could be to increase participant’s knowledge of the health effects of exercise. A long-term outcome could be a lesser likelihood of dying from heart disease.
Descriptive studies (Formative):

The researcher describes the goals, objectives, start-up procedures, implementation processes, and anticipated outcomes of a program, presenting the details of each.
Descriptive Studies:

• Are the goals articulated?
• Are the goals communicated throughout the organization?
• Are the objectives formulated in a clear and precise manner?
• Is responsibility for objectives assigned to specific individuals?

“I got a blank memo too.
Management must have added another unwritten rule.”
Normative studies (Process):

The researcher evaluates the program's goals and objectives by multiple values. Types of questions include the following.

• Are the goals minimalist, challenging, or unrealistically high?
• Are goals short or long term?
• Is the articulation of goals continuous or sporadic?
• Are objectives behavioural or attitudinal?
Impact studies (Outcome):

The researcher evaluates program goals and objectives in terms of outcomes.

• What are the short and long term, direct impacts clients or stakeholders?
• What are the indirect costs and benefits?
Centre for Disease Control framework

1999: The Centres for Disease Control and Prevention (CDC) published a six-step framework for conducting evaluation of public health programs with an increased emphasis government programs in the US.
CDC 6 steps

- Engage stakeholders
- Describe the program
- Focus the evaluation
- Gather credible evidence
- Justify conclusions
- Ensure use and share lessons learned
Research design
Four Major Types

1. Survey research:
   • descriptive and normative questions
   • random samples are far superior
   • cross-sectional, panel, and cohort

2. Case study:
   • descriptive and normative questions
   • can be quantitative, but are usually qualitative in support of findings gathered elsewhere
Research design
Four Major Types

3. Field experiment:
• impact questions
• requires a control group
• True, non-equivalent and before and after

4. Secondary data analysis:
• Useful for descriptive, normative, and impact
Evaluation researchers ask the following questions:

- What is the **nature and scope** of the problem?
- Where is it located, whom does it affect, and how does it affect them?
- What is it about the problem or its effects that justifies new, expanded, or modified social programs?
- What feasible interventions are likely to significantly ameliorate the problem?
Evaluation researchers ask the following questions:

- What are the appropriate target populations for intervention?
- Is a particular intervention reaching its target population?
- Is the intervention being implemented well?
- Are the intended services being provided?
Evaluation researchers ask the following questions:

• Is the intervention effective in attaining the desired goals or benefits?
• How much does the program cost?
• Is the program cost reasonable in relation to its effectiveness and benefits?
GET ALL THE INFORMATION YOU CAN, WE'LL THINK OF A USE FOR IT LATER.
Seven main areas of study within evaluation research are:

1. Product evaluation
2. Personnel evaluation
3. Performance evaluation
4. Program evaluation
5. Proposal evaluation
6. Policy evaluation
7. Empowerment evaluation “Assessment of Education”
“Evaluation studies involve a blend of creative thinking and in-depth knowledge of measurement techniques. This is so partly because they must combine knowledge of complicated labour market and economic theories, intricate models of program delivery and learning, elaborate sets of expected and unexpected potential impacts, and highly structured statistical models to estimate outcomes and impacts from partial lists of concepts that often cannot be measured directly.”
Evaluation often differs by level of Government and Agency:

"Pretend" patients are to visit 'walk-in' centres as part of an independent evaluation.
Provincial Governments: e.g. Ontario Health Units

Program evaluation

• systematic collection, analysis and reporting of information

• assists in decision-making.
Provincial Governments: e.g. Ontario Health Units

Stakeholders:

- Individuals and groups (both internal and external) who have an interest in the evaluation;
- Includes program staff or volunteers, participants, other community members, decision-makers, and funding agencies.
Provincial Governments: e.g. Ontario Health Units

Program evaluation

• systematic collection, analysis and reporting of information

• assists in decision-making.
Guiding Principles

- Integrated Program Planning and Evaluation
- Clear Description of the Program
- Explicit Purpose for Identified Need
- Specific Evaluation Questions
- Ethical Conduct
- Systematic Methods
- Clear and Accurate Reporting
Guiding Principles

• Timely and Widespread Dissemination
• Multidisciplinary Team Approach
• Stakeholder Involvement
• Utilization of Evaluation Findings
Evaluation in the Non-Profit Sector: A Clash of Cultures?
Evaluation is usually:

- An afterthought
- The last item on the agenda
- Has negative connotations
- Associated with grades and scores.
We know what these experiences can be like:

• They assume wrong-doing
• They are only qualitative
• They are mainly about control
• They produce anxiety about how we measure up
When evaluation is used only for rating, classification and sorting, client agencies are motivated only to prove to others that they fit or qualify.
Evaluation done only as a fitness test will not lead to innovative and adaptive initiatives with respect to such important societal challenges.
Ten Lessons:

1. Be clear about what you are trying to accomplish.
2. Set the stage properly. Be Candid
3. Pay attention to stakeholders.
4. Link the theory of the intervention to outcomes.
5. Integrate evaluation into the program.
Ten Lessons:

6. Integrate evaluation into daily work.
7. Identify just a few things to evaluate
8. Coordinate evaluation reports with internal decision-making.
9. Use evaluation as a process not simply as a report.
10. Do evaluation only when an organization is ready
The Participatory Model - Opening the "black box"
The Participatory Model:

• Gives primacy to the program’s actual service delivery system

• Understanding of the experiences of program participants is crucial

• Participants are followed from contact, selection, and enrollment in the program
SURE, WE'RE STILL LIVING IN POVERTY, BUT WE GOT GOOD SELF ESTEEM.
The Participatory Model:

- Trajectories of participants with different attributes are charted
- Key junctures when decisions were made are identified
- An attempt is made to understand how and why such decisions are made
- Modal patterns of staff behavior and participant responses to each service component at each key juncture are recorded
An approach that is NOT evaluation:
Program Effectiveness Organizational Development (PEOD)

• United Way Approach
• Outcome measurement – not ‘evaluation’
• Measure against agreed upon goals
P.E.O.D.

Cycle:
• Establish Program Need
• Plan Program
• Implement Interventions
• Assess Program Results
• Act on Findings
PEOD *Logic Model:*

Logical Relationship between program activities and the changes expected to occur from these activities.

"I think you should be more explicit here in step two."
Part II: Case Study in Evaluation: The National Child Benefit (NCB)
Governments agreed:

• Children need best possible start
• Poverty has long term consequences
• Families better off when supported to employment
• Accountability is important
Governments Agreed:

- Evaluation part of transparent accountability
- Findings to identify extent to which governments meeting objectives
- Employs multiple lines of evidence.
NCB

- Collaboration under SUFA
- Combines trends toward:
  - Shared initiatives
  - Delivery of social programs through income tested benefits
NCB approach to Evaluation followed a distinct path:

• Set Objectives
• Conduct Evaluability Assessment
• Report and Monitor annually
• Conduct Evaluation
• Engage in program assessment
What is the NCB initiative?

- NCBS income tested special benefit to low income families with children
- Provincial territorial offset of the NCBS under social assistance programs
- Reinvestment of NCBS offset savings in programs for children
NCB End game:

• New programs for children in all jurisdictions
• A common income ‘platform’ for all low income families whether on social assistance or working poor – harmonized delivery
Set Objectives: (Ministers of Social Services 1997)

- Prevent and reduce the depth of child poverty
- Promote attachment to the workforce by ensuring that families will always be better off as a result of working
- Reduce overlap and duplication with closer harmonization of program objectives and simplified administration
MY RESPONSIBILITIES INCLUDED BUDGETING FOR EDUCATION, HOUSING, TRANSPORTATION AND HEALTH CARE...

AS WELL AS NEGOTIATING WITH STATE, LOCAL AND FEDERAL AGENCIES.

THAT'S IMPRESSIVE! WHERE DID YOU GET THAT KIND OF EXPERIENCE?

I WAS ON WELFARE AFTER MY HUSBAND DIED.
Evaluability Assessment

• Can these objectives be evaluated?
• Can the intervention of the NCBS be independently separated from all else? – the economy, other trends
Evaluability Assessment

- Can the independent effect of the NCB ensure that low-income parents always better off working- can it be separated from other elements in the environment?
- How can we measure reductions in overlap and duplication? – delivery and reinvestments
What was done?

- Consultation with Experts
- External Peer Review
- Multiple lines of evidence
- External contracts with Expert consultants
- Support of the Auditor General
Monitoring and Reporting

NCB has had an annual federal/provincial/territorial report (since the inception)

- Early childhood services and children-at-risk services: $50.4 million (10%)
- Supplementary health benefits: $16.7 million (3%)
- Child benefits and earned income supplements: $152.5 million (32%)
- Child / day care: $167.0 million (34%)
- Other (includes First Nations): $99.0 million (21%)
Strengths

• Common definitions and common reporting
• Time series
• Reports changes in reporting
• Updates on reinvestments
• Changing environments monitored
The Evaluation

- Time series analyses of NCB impact on social assistance (SA) caseloads
- Survival analyses of NCB impact on SA caseloads
- Gross impact on poverty reduction
- Gross impact on reducing the welfare wall
The Evaluation

• Net impact analyses of the NCB on poverty reduction and labour supply

• Survey of NCB recipients views on the NCB

• Focus groups with NCB recipients
The Evaluation

• Surveys of Managers of the NCB programs
• Case studies of NCB programs
• Literature reviews
• Cost effectiveness analysis of the NCB
Constraints

• No control group
• Extraneous factors
Findings

Child Poverty

- NCB had positive impacts
- Reinvestment impact could not be established but have high potential
- Need for attention to child poverty
Findings

Work Incentives

• Work is more attractive

• Success in addressing problems of welfare wall

• Positive impact on labour market attachment but reduced levels of employment for those already in the labour force.
Findings

Harmonization

• Progress in reducing overlap and duplication
• Cost effectiveness can’t be concluded due to data limits but promising
Findings

Assessment

• More to be done
• Largely successful in meeting stated objectives
• No definitive picture
• Multiple lines of evidence corroborated findings of success in meeting stated objectives
"Heads, we pour massive amounts of cash into research and development, revolutionize our industry and become bigger than Microsoft, G.E. and Ford combined. Tails, we grab a six pack, leave early and go fishing."
Post-Module #5: Ideas to Action: Implementation, Monitoring & Evaluation

• What type of role do you envisage for your organization in implementing or influencing implementation in respect of your public policy issue?
• What would be a workable implementation strategy?
• What performance indicators would show whether the policy objectives are being achieved?
• What role could you or your organization play in the evaluation process?