IN THE DEEP END: EVALUATION 101

AUSTRALIAN EVALUATION SOCIETY CONFERENCE

LAUNCESTON, WEDNESDAY 19 SEPTEMBER 2018

POLICY PERFORMANCE

THE BUSINESS OF GOVERNMENT

PLANNING | IMPLEMENTATION | EVALUATION
OPENING REMARKS

WELCOME

RATIONALE FOR SESSION

STYLE OF SESSION

TODAY WE WILL COVER:
   A. WHY
   B. WHAT
   C. HOW
   D. WHO
   E. WHEN
   F. WHERE
BUT FIRST, INTRODUCTIONS
INTRODUCTIONS

1. NAME

2. HOME CITY & STATE

3. ORGANISATION & ROLE

4. PAST/CURRENT STUDY HISTORY

5. EVALUATOR SELF-RATING:
   - **IGNORAMUS**: I really don’t know much about evaluation, I don’t know what I don’t know
   - **LEARNING FAST**: I’m on a wild ride, learning by doing and by speaking with others. Hoping it all clicks into place soon
   - **FINALLY GETTING COMFORTABLE**: I’ve got the hang of the basics and feel like I’m finally doing a good job of this evaluation stuff
   - **INFLUENCER**: I spend my days shaping this field

Introductions
QUICK CHECK

GEOGRAPHY - FURTHEST AND CLOSEST

SELF-RATINGS:
- IGNORAMUS
- LEARNING FAST
- FINALLY GETTING COMFORTABLE
- INFLUENCER
A. WHY
WHY IS EVALUATION NEEDED?

- To inform program policy and planning efforts
- To help monitor (and improve) program performance
- For accountability and transparency purposes
- To reduce waste and maximise efficiency (or cost-efficiency)
- To support decision-making about program continuation or adaptation, e.g. to inform program expansion
- To support the development of new or similar programs
- To re-examine program relevance over time

"The key is to make program evaluation integral to managing government programs at all stages, from planning an initial development through start-up, ongoing implementation, appropriations, and reauthorisation."

IN SHORT... EVALUATION AS AN ESSENTIAL MANAGEMENT FUNCTION"

SOURCE: AMERICAN EVALUATION SOCIETY. 2010
B. WHAT IS EVALUATION?
KEY EVALUATION TERMS

EVALUATION
• A FORM OF SYSTEMATIC INQUIRY TO ASSESS IMPLEMENTATION OR IMPACT OF AN INTERVENTION. IT MAY FOCUS ON THE MERIT, WORTH OR SIGNIFICANCE OF AN INTERVENTION AND SEEKS TO JUDGE SUCCESS/VALUE.

OUTCOME(S)
• A VARIABLE, OR VARIABLES, WHICH MEASURE THE IMPACT OF THE INTERVENTION.

IMPACT
• THE EFFECT OF AN INTERVENTION ON OUTCOMES FOR THE BENEFICIARY POPULATION.

ATTRIBUTION
• THE EXTENT TO WHICH THE OBSERVED CHANGE IN OUTCOMES IS THE RESULT OF THE INTERVENTION, HAVING ALLOWED FOR ALL OTHER FACTORS WHICH MAY ALSO AFFECT THE OUTCOME(S) OF INTEREST

IMPACT EVALUATION
• A STUDY OF THE ATTRIBUTION OF CHANGES IN THE OUTCOME TO THE INTERVENTION.
DEDUCTIVE REASONING

• Consider a theory about the topic of interest. Develop specific hypotheses. Narrow down to collect observations to address the hypotheses. Test the hypotheses with specific data -- a confirmation (or not) of the original theories.

INDUCTIVE REASONING

• Start with observations, look for patterns, arrive at tentative hypotheses, then define theories (bottom up)

QUALITATIVE

• Research that uses reflections, opinions, values, views and other sources that are not readily quantifiable

QUANTITATIVE

• Research that uses data that can be numericised/quantified

MIXED METHODS

• The use of both quantitative and qualitative methods in an evaluation design.
IMPACT EVALUATION AND RCTs

COUNTERFACTUAL
- The value of the outcome for the treatment group in the absence of the intervention.

RANDOMISED CONTROLLED TRIAL (EXPERIMENTAL DESIGN)
- A design in which random assignment has been used to allocate the intervention amongst members of the eligible population.

Since there should be no correlation between participant characteristics and the outcome, any differences in outcome between the treatment and control can be fully attributed to the intervention;

QUASI-EXPERIMENTAL DESIGN
- Impact evaluation designs which create a control group using statistical procedures. The intention is to ensure that the characteristics of the treatment and control groups are identical in all respects, other than the intervention.
**SAMPLING FRAME**

- The complete list of the population of interest in the study. This is not necessarily the complete population of the country or area being studied, but is restricted to the eligible population; for example, families with children under five, or female-headed households. The **sample** is the group that is selected from the population.

**RANDOM ASSIGNMENT**

- An intervention design in which members of the eligible population are assigned at random to either the treatment group or the control group (that is, random assignment). That is, whether someone is in the treatment or control group is solely a matter of chance, and not a function of any of their characteristics (either observed or unobserved).
VALIDITY

CONCLUSION VALIDITY
• TESTING WHETHER THERE IS A RELATIONSHIP BETWEEN TWO VARIABLES

INTERNAL VALIDITY
• ASSUMING THAT THERE IS A RELATIONSHIP, IS THE RELATIONSHIP A CAUSAL ONE? CONSIDER VARIOUS THREATS TO VALIDITY – E.G.:
  ▪ SELECTION BIAS
  ▪ HISTORY (EFFECT OF OTHER VARIABLES/DIFFERENCES)
  ▪ MATURATION (PERSONAL GROWTH)
  ▪ INSTRUMENTATION
  ▪ DIFFUSION OF TREATMENT

CONSTRUCT VALIDITY
• DID WE IMPLEMENT THE PROGRAM WE INTENDED TO IMPLEMENT, AND DID WE MEASURE THE OUTCOME WE WANTED TO MEASURE? IN OTHER WORDS, DID WE OPERATIONALIZE THE CAUSE AND THE EFFECT?

EXTERNAL VALIDITY
• THE EXTENT TO WHICH THE RESULTS OF THE EVALUATION APPLY TO OTHER PERSONS, PLACES OR TIMES.
THEORY-BASED EVALUATION

LOGIC MODEL

• DESCRIBES HOW A PROGRAM SHOULD WORK, PRESENTING THE CAUSAL CHAIN FROM INPUTS, THOUGH ACTIVITIES AND OUTPUTS, TO OUTCOMES.

WHILE LOGIC MODELS PRESENT A THEORY ABOUT THE EXPECTED PROGRAM OUTCOME, THEY DO NOT DEMONSTRATE WHETHER THE PROGRAM CAUSED THE OBSERVED OUTCOME.

A THEORY-BASED APPROACH EXAMINES THE ASSUMPTIONS UNDERLYING THE LINKS IN THE LOGIC MODEL.
SYSTEMATIC REVIEW
• A SYNTHESIS OF THE RESEARCH EVIDENCE ON A PARTICULAR TOPIC, OBTAINED THROUGH AN EXHAUSTIVE LITERATURE SEARCH FOR ALL RELEVANT STUDIES USING STRATEGIES TO MINIMISE ERROR.

META-ANALYSIS
• THE SYSTEMATIC ANALYSIS OF A SET OF EXISTING EVALUATIONS OF SIMILAR PROGRAMS IN ORDER TO DRAW GENERAL CONCLUSIONS, DEVELOP SUPPORT FOR HYPOTHESES, AND/OR PRODUCE AN ESTIMATE OF OVERALL PROGRAM EFFECTS.

Sources for definitions in this section:
Trochim, Social Research Methods (web), available at: https://www.socialresearchmethods.net/kb/measure.php
C. HOW TO EVALUATE?
**Phase 1: Framing/Planning**

**Framing**
- Understand purpose and scope of the evaluation
- Build understanding of the evaluand/intervention
- Establish/refine key questions
- Select evaluation approach and methods
- Develop or clarify the program theory
- Stakeholder identification and planning
- Data familiarisation
- Develop primary data collection tools
- Confirm success measures (KPIs, baseline, targets)

**Project Management Activities:** E.g. Plan team, dates, workflow, budget, approvals

**Evaluation Framework**
- and associated data collection tools
PHASE 2: DATA COLLECTION

DELIVER EVALUATION IN-LINE WITH EVALUATION FRAMEWORK
COLLECT DATA TO BUILD UNDERSTANDING OF HISTORY, CONTEXT, BASELINE, ACTIVITIES, OUTPUTS, OUTCOMES & RESULTS
ENGAGE WITH STAKEHOLDERS (E.G. INTERVIEWS, SURVEYS, FOCUS GROUPS) AND DOCUMENT OUTPUTS

PROJECT MANAGEMENT ACTIVITIES: E.G. ENGAGE WITH CLIENTS, MANAGE SCOPE OF WORK, PROVIDE PROJECT UPDATES

COMPLETION OF DATA COLLECTION
PHASE 3: ANALYSIS

REVIEW PRIMARY AND SECONDARY INFORMATION IN RELATION TO EVALUATION FRAMEWORK
THEME DATASETS, WHERE APPROPRIATE
HOLD TEAM MEETING TO DISCUSS FINDINGS AND POSSIBLE RECOMMENDATIONS
PRESENT PRELIMINARY FINDINGS TO TEST/SOCIALISE RESULTS

PROJECT MANAGEMENT ACTIVITIES: E.G. PLAN REPORT STRUCTURE, LAYOUT AND WRITING APPROACH, PREPARE REPORT DATA APPENDIX, PREPARE FOR FINDINGS DISCUSSION.

COMPLETION OF DATA ANALYSIS AND APPENDICES, PRELIMINARY FINDINGS
PHASE 4: REPORTING

PREPARE REPORT INCLUDING: EXECUTIVE SUMMARY, DESCRIPTION OF EVALUAND (PROGRAM THEORY), METHODOLOGY DESCRIPTION, LIST DATA SOURCES, KEY RESEARCH QUESTIONS AND KEY FINDINGS

PROJECT MANAGEMENT ACTIVITIES: E.G. PROJECT COMPLETION AND WRAP-UP, SEEK FEEDBACK (INTERNAL AND EXTERNAL), MEET KEY STAKEHOLDERS, DISCUSS FINDINGS TO PROMOTE UTILISATION

DRAFT AND FINAL REPORTS
LIST OF EVALUATION METHODS

- Surveys / Questionnaires
- Stakeholder Interviews / Forums
- Observational Case Studies / Site Visits
- Desktop Data Analysis (Program and Policy Materials)
- Literature Review
- Benchmarking / Jurisdiction Scans
- Financial Analysis / Cost-Benefit / Cost-Effectiveness Analysis
<table>
<thead>
<tr>
<th>List of Evaluation Approaches</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised Control Trial</td>
<td>Empowerment Evaluation</td>
</tr>
<tr>
<td>Quasi-Experimental Design</td>
<td>Democratic Evaluation</td>
</tr>
<tr>
<td>Realist Evaluation</td>
<td>Research Synthesis/Meta-Analysis</td>
</tr>
<tr>
<td>Theory-Based Evaluation</td>
<td>Strengths-based Evaluation/Appreciative Inquiry</td>
</tr>
<tr>
<td>Goal-Free Evaluation</td>
<td>?</td>
</tr>
<tr>
<td>Developmental/Real-Time Evaluation</td>
<td>?</td>
</tr>
<tr>
<td>Co-Design / Participatory Evaluation</td>
<td>?</td>
</tr>
<tr>
<td>Utilisation-Focused Evaluation</td>
<td>?</td>
</tr>
</tbody>
</table>
D. WHEN SHOULD(N’T) WE EVALUATE?
EVALUATION AND THE POLICY CYCLE

POLICY PHASE

PROGRAM DESIGN

SAMPLE EVALUATION APPROACHES/TECHNIQUES

NEEDS ANALYSIS
LITERATURE REVIEW (PRACTICES ELSEWHERE)
LOGIC MODELLING (CLARIFY PROGRAM THEORY)
MONITORING AND EVALUATION FRAMEWORK DEVELOPMENT
EVALUATION CAPACITY BUILDING
CO-DESIGN PROCESSES

PROGRAM IMPLEMENTATION

DEVELOPMENTAL (OR REAL-TIME) EVALUATION PROCESS/IMPLEMENTATION EVALUATION
PERFORMANCE MEASUREMENT AND MONITORING

PROGRAM COMPLETION

IMPACT EVALUATION
POST-COMPLETION EVALUATION
LAPSING PROGRAM EVALUATION
E. WHO CAN DO EVALUATION?
AES GUIDANCE: COMPETENCIES

AES PROFESSIONAL LEARNING COMPETENCY FRAMEWORK

• EVALUATIVE ATTITUDE AND PROFESSIONAL PRACTICE
• EVALUATION THEORY
• CULTURE, STAKEHOLDERS AND CONTEXT
• RESEARCH METHODS AND SYSTEMATIC INQUIRY
• PROJECT MANAGEMENT
• INTERPERSONAL SKILLS
• EVALUATION ACTIVITIES

HTTPS://WWW.AES.ASN.AU/IMAGES/STORIES/FILES/PROFESSIONAL%20LEARNING/AES_EVALUATORS_COMPETENCY_FRAMEWORK.PDF
AES GUIDANCE: ETHICAL CONDUCT

AES GUIDELINES FOR THE ETHICAL CONDUCT OF EVALUATIONS

- COMMISSIONING AND PREPARING FOR AN EVALUATION
- CONDUCTING AN EVALUATION
- REPORTING THE RESULTS OF AN EVALUATION

HTTPS://WWW.AES.ASN.AU/IMAGES/STORIES/FILES/MEMBERSHIP/AES_GUIDELINES_WEB_V2.PDF
INFORMATION SOURCES

1. RIGHT HERE AT THE CONFERENCE

2. AUSTRALASIAN EVALUATION SOCIETY WEBSITE (AES.ASN.AU)
   - EVALUATION JOURNAL OF AUSTRALASIA
   - LIST OF CONFERENCES
   - LINKS TO EVALUATION RESOURCES
   - THE CONSULTANTS DIRECTORY
   - EVALUATION TENDERS
   - JOB VACANCIES
   - CALENDAR OF EVENTS
   - SPECIAL INTEREST GROUPS
   - EMAILS (E.G. BLOG)

SEE ALSO:
   - AMERICAN EVALUATION SOCIETY (EVAL.ORG)
   - CANADIAN EVALUATION SOCIETY (EVALUATIONCANADA.CA)
3. INTERNATIONAL JOURNALS (FREE FOR AES MEMBERS):
   - EVALUATION
   - EVALUATION REVIEW
   - ALTERNATIVE
   - JOURNAL OF MIXED METHODS RESEARCH
   - QUALITATIVE RESEARCH
   - EVALUATION AND THE HEALTH PROFESSIONS

SEE ALSO:
   - NEW DIRECTIONS FOR EVALUATION
   - AMERICAN JOURNAL OF EVALUATION

4. INTERNATIONAL EVALUATION RESOURCES:
   - BETTER EVALUATION (BETTEREVALUATION.ORG)
   - RESEARCH METHODS KNOWLEDGE BASE - TROCHIM
     (SOCIALRESEARCHMETHODS.NET)
   - WORLD BANK: INDEPENDENT EVALUATION GROUP
     (IEG.WORLDBANKGROUP.ORG/)
INFORMATION SOURCES

5. CENTRE FOR PROGRAM EVALUATION AT THE UNIVERSITY OF MELBOURNE
   - MASTER OF EVALUATION PROGRAM
   - P.H.D IN EVALUATION

6. ATTEND EVALUATION TRAINING SESSIONS
   - AES WORKSHOPS ON SUBJECTS OF INTEREST
   - EVALUATION TRAINING PROGRAMS LED BY CONSULTANCIES OR VISITING PROFESSORS

7. GOOGLE
INFORMATION SOURCES

8. SELF-DIRECTED READING AND RESEARCH

USE (why?)  METHODS (how?)  VALUING (how?)

Source: Alkin (2012), Evaluation Roots (2nd Ed)
CLOSING REMARKS... !
CLOSING REMARKS

1. THANKS FOR ATTENDING THE CONFERENCE AND THIS SESSION

2. HOW TO GET THE MOST FROM THE CONFERENCE

3. TIME FOR QUESTIONS... ?