



*Summer Institute on Program Evaluation
May 30 - June 3, 2016*

Evaluation Designs

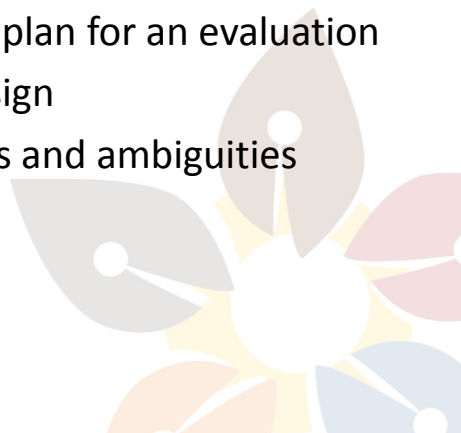
Javier Mignone
Department of Community Health Sciences
The University of Manitoba
Javier.mignone@umanitoba.ca



Design and Measurement Decisions

There are no rigid rules for making methods decisions. Corollaries to this lack of rules are:

- There is no single best plan for an evaluation
- There is no perfect design
- There are always errors and ambiguities





Design and Measurement Decisions (cont'd)

Different methods are appropriate for different situations (paradigm of choices).

Evaluation methods are to be judged on the basis of appropriateness, utility, practicality, credibility, and relevance. These criteria are necessarily situational and context-bound. One cannot judge the adequacy of methods used in a specific evaluation without knowing the purpose of the evaluation, the intended uses of the findings, the resources available, and the trade-offs negotiated.



Evaluation Designs

- **After only (post-program)**
 - Evaluation is done after the program is completed
- **Retrospective (post-program)**
 - Participants are asked to recall or reflect on their situation, knowledge, attitude, behaviour, etc., prior to the program
- **Before-after (before and after the program)**
 - Program participants or situations are looked at before the program and then again after the program



Evaluation Designs (cont'd)

- During (additional data “during” the program)
 - Collecting information at multiple times during the course of the program
- Time series (multiple points before and after program)
 - Involves a series of measurements at intervals before the program begins and after it ends
- Case study
 - Uses multiple sources of information and multiple methods to provide an in-depth and comprehensive understanding of the program



Qualitative Data Approaches (mostly)

- Open-ended or semi-structured interviews
- Focus groups
- Direct observation
- Participatory observation
- Journal
- Document analysis
- Combination of approaches (e.g., photo-voice, life story board)



Quantitative Data Approaches (mostly)

- Data collected from existing records
- Development of codes for existing records
- Development of questionnaires (surveys)
 - Self-administered
 - Administered by interviewer in-person or by telephone
 - Questions can be open-ended or close-ended
 - Other options: e.g., polling booth surveys



Mixed Methods

Is an approach to inquiry that combines or associates both qualitative and quantitative forms.

Sequential

Quan → Qual

Qual → Quan

Concurrent

Quan

Qual




Inductive and Deductive Strategies

- *Inductive*
 - Evaluator attempts to make sense of a program without imposing preexisting expectations on the program setting
 - Inductive designs begin with specific observations and build toward general patterns
- *Hypothetico-deductive*
 - Requires the specification of main variables and the hypothesis before data collection begins
 - Hypothesis based on explicit theoretical framework (Patton, 2008)




Inductive and Deductive Strategies

Qualitative/Naturalistic Paradigm	Quantitative/Experimental Paradigm
Qualitative data (narratives, descriptions)	Quantitative data (numbers, statistics)
Naturalistic inquiry	Experimental designs
Case studies	Treatment and control groups
Inductive analysis	Deductive hypothesis testing
Subjective perspective	Objective perspective




Inductive and Deductive Strategies

Qualitative/Naturalistic Paradigm	Quantitative/Experimental Paradigm
Close to the program	Aloof from the program
Holistic contextual portrayal	Independent and dependent variables
Systems perspective focused on interdependencies	Linear, sequential modeling
Dynamic, ongoing view of change	Pre-post focus on change
Purposeful sampling of relevant cases	Probabilistic, random sampling



Inductive and Deductive Strategies

Qualitative/Naturalistic Paradigm	Quantitative/Experimental Paradigm
Focus on uniqueness and diversity	Standardized, uniform procedures
Emergent, flexible designs	Fixed, controlled designs
Thematic content analysis	Statistical analysis
Extrapolations	Generalizations




Some Basic Designs

1)	I	X	O			
2)	I	X	O	O		
3)	I	O	X	O		
4)	I	O	X	O	O	O
5)	I	O	X	O	X	O
6)	I	O	X	O	O	O
	C1	O	Y	O	O	O
	C2	O	O	O	O	O

X = Intervention; O = Observation; I = Intervention group; C1 = Comparison group 1; C2 = Comparison group 2; Y = Intervention in comparison group

(Shadish, et al., 2002)



Some Basic Designs (cont'd)

Basic designs

- True experimental designs (random assignment)
- Quasi-experimental designs (non-random assignment)
 - Non-equivalent control group with pre- and posttest
 - Time-series with non-equivalent control group
- Non-experimental designs
 - Single group pretest-posttest
 - Single group time-series



Case Study Designs

Case study is an ideal methodology when a holistic, in-depth investigation is needed. Whether the study is experimental or quasi-experimental, the data collection and analysis methods are known to hide some details. Case studies, on the other hand, are designed to bring out the details from the viewpoint of the participants by using multiple sources of data. (Tellis, 1997)



Case Study Designs (cont'd)

According to Yin (1994) there are at least four applications for a case study model:

- To explain complex causal links in real-life interventions
- To describe the real-life context in which the intervention has occurred
- To describe the intervention itself
- To explore those situations in which the intervention being evaluated has no clear set of outcomes.



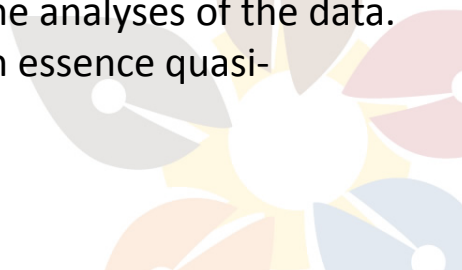
Adequacy Assessments

Seek to determine if the expected changes occurred. Inferences about the adequacy of program outcomes depend on the comparison of the performance or impact of the project with previously established adequacy criteria.



Plausibility Assessments

Seek to determine if a program had an effect above and beyond other external influences. These evaluations attempt to control for the influence of confounding factors by choosing control groups before an evaluation is begun, or afterwards during the analyses of the data. This type of design is in essence quasi-experimental





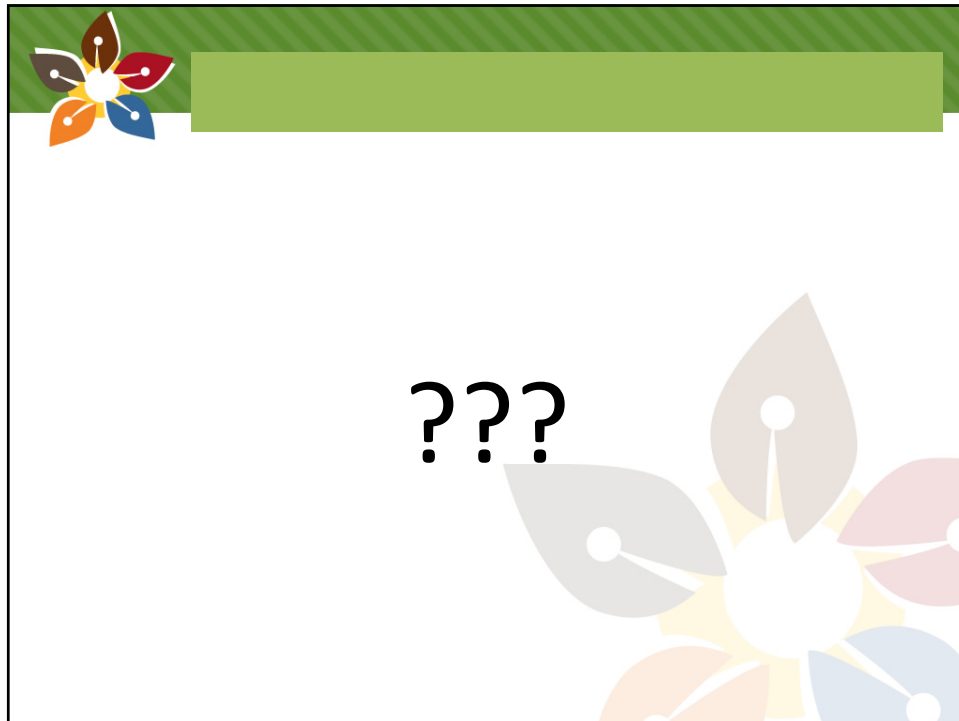
Probability Assessments

Aim at ensuring that there is only a small known probability that the difference between program and control areas were due to confounding, bias, or to chance. These evaluations require randomization of treatment and control activities to the comparison groups, being the “gold standard” of academic *efficacy* research (experimental studies). (Habicht, et al., 1999)



The 3 E's

- Efficacy: the ability of an intervention to produce the desired beneficial effect in expert hands and under ideal circumstances.
- Effectiveness: the degree to which action(s) achieves the intended results under normal or usual (real) circumstances.
- Efficiency: The production of the desired effects or results with minimum use of time, resources and efforts.



References/Bibliography

- Cresswell JW (2003) *Research design: Quantitative, qualitative, and mixed methods approaches*. Thousand Oaks: Sage Publications.
- Habicht JP, Victora CG, Vaughan JP (1999) Evaluation designs for adequacy, plausibility, and probability of public health programme performance and impact. *International Journal of Epidemiology*, 28, 10-18.
- Patton MQ (1997) *Utilization-focused evaluation: The new century text*. 3rd edition. Sage: Thousand Oaks.
- Patton MQ (2008) *Utilization-focused evaluation*, 4th Edition, Thousand Oaks: Sage Publications.
- Shadish WR, Cook TD, Campbell DT (2002) *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin: Boston.
- Tellis, W (1997) Application of a case study methodology. *The Qualitative Report*, 3(3).
- Yin RK (2009) *Case study research: Design and methods*. 4th edition.