Challenges in Utilizing Randomized Control Trials in Social Program Evaluation: Lessons Learned

A presentation by Robert Malatest
R.A. Malatest & Associates Ltd.
to the 2011 Canadian Evaluation Society Conference
Edmonton, Alberta
May 2011

About Our Firm

• Long-standing history of completing program evaluation since 1985
• Firm has 4 offices (Ottawa, Toronto, Edmonton, Victoria) and specializes in the evaluation of public sector programs/services at the federal, provincial, local and/or Pan-Canadian level
• The firm has more than 60 full-time staff, plus another 100 – 150 staff working in three call centres
• We routinely conduct 12 to 15 evaluations/ program reviews per year (value ranging from $50,000 to $2 million)

Presentation Overview

• Why is random assignment a “better” evaluation approach?
• Our experience in random assignment.
• Barriers & challenges to utilizing random assignment for social program evaluation.
• Best Practices to Support RCT in Program Evaluation.
• Summary.
Why Random Assignment?

- Evaluators constantly strive to answer the question "What is the net impact of the program?"
- To do this, we typically construct a comparison group.
- While we can often adjust/correct for a number of factors (demographics, experience, etc.), we are never confident that the comparison group is a good match for program participants – particularly with respect to program participant "motivation".

The Comparison Group Quandry

- Comparison groups are usually selected on the basis of:
  - prioritization - those with highest needs assigned to intervention, those with less need assigned to the comparison group
  - "first in" basis - intervention clients are selected on the basis of "first come", later applicants are assigned to the comparison group
  - those in a program are compared to those not in a program (first and second stage matching to develop an appropriate comparison group)
- Evaluators are constantly asked whether the comparison group is in fact a good population for the counterfactual position

Random Assignment Benefits

- Evaluations that utilize random assignment offer a better way of ensuring comparability of the intervention group with a very similar group.
- A key similarity is both the treatment and control group were eligible participants in the study - the only difference is the random assignment to the intervention.
- This method then controls for "motivational" issues that cannot be easily accounted for in most other comparison group constructs.
Random Assignment Projects - Case Study Examples

• Evaluation of a program to measure the effectiveness of case management for “at risk” students in three Ontario colleges (Foundations for Success)
• Evaluation of an Aboriginal training program (BC)
• Evaluation of a youth gang prevention initiatives (Ontario, AB, SK)

Challenges of Using Random Assignment

• Administrative and logistical challenges
• Program design issues
• Sample issues
• Challenges on implementing random assignment in a social policy context
  ➢ ethical considerations
  ➢ other

Administrative and Logistical Challenges

• Random assignment typically involves more time and effort at the research design stage than is the case for evaluations that utilize administrative data as the comparison. Examples include:
  ➢ need to obtain informed consent from potential project participants (intervention and control)
  ➢ more resources need to be allocated to evaluation design (control for cross-contamination, complexity of design, other)
  ➢ recruitment issues - some individuals may not want to participate if there is a possibility that they will not receive services/benefits
Program Design Issues

- Need to examine the extent to which random assignment can truly isolate participants from control (i.e., random assignment of different family members, students within a classroom, etc.)
- Difficulty in measuring impacts should the program design and/or evaluation be changed due to implementation issues (not unique to random assignment)
- In order to accurately measure participation impacts, it is often required that additional data be collected from delivery agencies

Sample Issues

- Sufficient sample of the various cohorts must exist to enable the random sampling method to be applied
- Programs being evaluated with ongoing intake must have sufficient continual enrolment to make effective use of this process

Ethical Considerations

- This presents the greatest challenge to implementing random assignment
- Service providers/contract authorities may initially agree to random assignment, but may withdraw support when they realize that some “at risk” clients will not receive services
Example - Youth Gang 
Prevention Pilot Project

- Client was originally supportive of random assignment methodology
- When advised that in order to demonstrate impact, some control group youth could become more at risk of gang affiliation due to “non-treatment”, support for random assignment methodology was withdrawn
- Evaluation methodology utilized voluntary participation (into the program) targeting high risk youth that reduced the comparability/extrapolation to the “at risk” population

Example – Aboriginal 
Youth Training Evaluation

- Project/evaluation design envisioned random assignment
- Limited ability to recruit participants resulted in a small control group
- Service provider desired to include “most appropriate” clients as part of the intervention group – counter to the random assignment model

Example - Youth Leadership 
Pilot Project

- Funder envisioned a RCT approach to the evaluation of a youth leadership program. Service provider had only limited understanding of RCT.
- Service provider selected clients with the (mis)understanding that control group clients could be randomly assigned at a later date.
- Community events occurred that likely affected the perceptions of youth that were later identified as potential RA control participants which will weaken the evaluation.
Other Considerations

- Limited understanding and support for random assignment – often viewed as a “black box” due to limited exposure to random assignment by contract authorities and service providers
- Level of experience and capacity in Canada is limited
- Random assignment studies may require more “processes” when compared to other evaluations (i.e., ethics review, informed consent, participation agreements, other)
- Random assignment studies are not “perfect” – they have their own biases (see Stafford, 2002, Cambridge University)

Conditions for Successful Implementation of Random Assignment

- Requires sufficient support from service providers and funders for the notion that the program design will not include the provision of services to all clients
- Particularly well-suited for social program evaluations that are:
  I. in the pilot stage or is a demonstration project
  II. where it is anticipated that the number of clients will exceed the ability to serve all clients (rationing would have to occur in any extent)
  III. minimal risks to contamination between intervention and control group

Other Implementation Issues

- Pilot projects rarely result in optimal delivery in the first year of operations. From an evaluation standpoint, evaluations that utilize random assignment should provide funding for two year/two phase evaluation (Year 1/Phase 1 data may not reflect the true operation[s] of the pilot)
- Year 1/Phase 1 should sometimes be viewed as a “pilot” of the random assignment methodology/program delivery. Corrections after Year 1 can contribute to improved evaluation in Year 2
Best Practices to Support the Use of RA in Program Evaluation

- Longer than average planning phase to facilitate pre-selection activities (ethics review, service provider “education”, other)
- Large number of potential clients for both the intervention and control group (many RA projects are shelved due to limited numbers of clients)
- Incorporate training for service providers as to the “mechanics” of RA evaluations

Evaluation “Benefits”

- Match of participant and comparison clients typically better than through the use of other matching processes
- “Motivation” for both groups are equivalent
- Very easy to attribute impacts to the intervention rather than having to explain possible differences in samples/environment

With Large Samples – Random Assignment Typically Generates Very Close Match Between Control and Intervention Groups (FFS)

Source: Foundations for Success (n=1,740)
Random Assignment Resulted in Groups with Similar "Motivations"

Source: Foundations for Success (n=1,740)

Random Assignment Methodology Allowed for Greater Confidence in Estimating Program Impacts

Source: Foundations for Success (n=1,740)

Lessons Learned - Practitioner

- Random assignment typically requires more planning and implementation time relative to other evaluation methodologies
- Expect challenges in terms of recruitment – even when given a high probability of acceptance, some individuals will refuse to be part of the study – OVERSAMPLE!
- Easier to demonstrate outcomes without having to resort to complex regression models (bootstrapping, propensity score matching, etc., don’t appear in the evaluation report)
- Need to build in process evaluation elements to explain why results may/may not have occurred
Lessons Learned - Funders

- See bullet #1 previous – planning resources!
- Contract Authority reluctance to engage in random assignment usually revolves around "ethical" issues; however, to make informed decisions is it not in the interest of good public policy to use best available tools to determine program effects?
- Random assignment projects lend themselves to:
  - evaluation of pilot projects (a new "intervention")
  - projects which will involve a rationing of services (some allocation methodology would have been developed anyway)
  - projects that do not involve significant ramifications for excluded participants

Hopes for the Future

- Random assignment can work to demonstrate social program impacts – funders need to support random assignment projects in Canada
- An important element of any random assignment study is the dissemination of both process (how it was done) and outcomes to the research community – capacity building is important – but it comes at a cost
- May require Treasury Board or other oversight agencies to demand more rigourous evaluation models to encourage evaluation authorities to "move out of their comfort zone" to adopt random assignment evaluation models

For More Information About Foundations for Success:

Final Impacts Report: