

# **Implementing Randomized Field Trials in Social Settings**

**Judy Gueron  
MDRC**

**September 24, 2003**

**National Research Council  
Center for Education  
Committee on Research in Education**

# Why Random Assignment?

- With RFTs you can more confidently separate fact from advocacy. With alternatives, this is much less true
- RFTs involve up-front challenges, but the political and financial costs are high for evaluations that end in methodological disputes
- But RFTs are not always the right approach

# Today

- Summary comments
- Eight key challenges
- Comparative costs

# What It Takes

- Our experience suggests that implementing a successful RFT is an *art* that involves the continuing interaction between creative and flexible research skills and operational and political savvy

# Four Stages Of A Study

## 1. Start up

- Design: theory, questions, scale, duration, internal vs. external validity
- Selling it
- Recalibration
- Resist starting RA too soon

## 2. Operational

- Monitoring Es and Cs
- Equivalent, high quality impact data
- Process and context data

# Stages 3 and 4

## 3. Analysis

- Estimating impacts and possibly costs and benefits
- Explaining impacts: what causes success?

## 4. Diffusion

- Interpret and promote findings
- Impact policy and practice

# Key Lessons

- Stage 1 is make it or break it
- Anticipate: this will be very hard
- Success requires very senior people who know the “art”
- Cannot be done on the cheap: resources and time
- Resist trying to answer all questions in one study
- Initially, give priority to internal vs. external validity
- Knowledge accumulates: many studies, ultimately syntheses

# Challenge 1: Address the Important Questions

1. Does the treatment achieve more than services available in the community?
2. Is the treatment *per se* effective?
3. Is one treatment more effective than another?

## Lessons

- Implement the treatment: a fair test
- Keep your eye on the controls
- Assure a meaningful E/C service differential

Not unique to RFTs

## Challenge 2: Meet Ethical and Legal Standards

- Do not deny entitlements
- Do not reduce service levels
- Address important, unanswered questions
- Inform students and parents
- Assure data confidentiality
- Use RFTs only if there is no less intrusive way
- Have high probability that results will be used

Most are not unique to RFTs

## Challenge 3: Convince People That RFTs Are Essential and → Policy

- Support from prestigious academics and panels
- Some “stick” – the welfare waiver model
- Obstacles
  - Burden, ethics, politics
  - Impacts are smaller than outcomes
- Modest changes are enough to bring visibility and affect policy
- Counter the hype

# Challenge 4: Balance Research Ambition and Operational Reality

- Rare opportunities: resist the urge to answer all questions
- Decisions that → site burden
  1. Location/duration of random assignment
  2. Handling of controls
  3. Intrusiveness of data collection,
  4. Unit of random assignment: students, schools, teachers
  5. Complexity of the design
  6. Emphasis on external validity
  7. Time before start RA
- Some are not unique to RFTs

# Internal versus External Validity

- What question do you care about
  - internal validity: can/does it work?
  - external validity: a bottom-line number?
  - what affects effects; continuous improvement?
- Having samples in many places is preferable *if* you can assure randomness, programs, and process data
- At this stage in education, emphasize internal validity vs. externally valid samples that jeopardize internal validity
- Stages in knowledge development
- A research agenda that builds toward synthesis

## Challenge 5: Assure Randomness & Increase Participation

- Random assignment is an all-or-nothing process. It doesn't help to be a little bit random
- Researchers set or control the process
- Explain to administrators/teachers up front
  - Real value for young people
  - What is involved
  - What will be learned
  - Why uniquely reliable
  - Impact on future funding
- Can build a constituency

## Challenge 6: Have Enough Sample, Locations, Time

- Detecting policy relevant impacts
  - Sample of students or schools large enough
  - Long enough follow-up to yield reliable results
  - Multiple and diverse locations
- Not unique to RFTs

# Challenge 7: Measure Multiple Outcomes; Link Treatment To Results

- Researcher and stakeholder theories
- A range of outcomes
- Determining what causes success or failure: getting inside the black box
  - Descriptive comparisons
  - Differential impact studies (multiple treatment groups)
  - Various forms of synthesis
  - Building a statistical model from multiple experiments
- Not unique to RFTs

# Challenge 8: Enforce Research Status Over Time

- More difficult if
  - Multiple services
  - Extends over many years/sites
  - Controls and experimentals in the same school
  - Two or more treatments provided by the same teachers in the same schools
- Most not unique to RFTs

# Costs of RFTs Versus Large Quasi-Experiments

- Large-scale field research with primary data collection is very expensive
- True for RFTs and alternatives
- RFTs cost more up front but less later
- Well-implemented RFTs are more cost effective, in terms of impact on policy

# Comparative Cost: Stages 1 and 2

- |                            |   |
|----------------------------|---|
| 1. Start up                | + |
| ➤ Design                   | ≈ |
| ➤ Selling it               | + |
| • site recruitment         | + |
| • site incentives          | + |
| • participant incentives   | ≈ |
| • control services         | ? |
| ➤ Recalibration            | + |
| ➤ Wait (fair chance)       | ≈ |
| 2. On-going                | + |
| ➤ Monitoring Es and Cs     | + |
| ➤ High quality data        | ≈ |
| ➤ Process and context data | ≈ |

# Comparative Cost: Stages 3 and 4

- 3. Analysis —
  - Estimating impacts —
  - Explaining impacts —
- 4. Diffusion —
  - Interpret and promote findings —
  - Impact policy and practice —