Critical Issues in the Use of Randomized Clinical Trials and Control Groups within Applied Settings: Rationale, Challenges and Benefits

Hill M. Walker, Ph.D.
Center on Human Development
University of Oregon
Eugene, Oregon
Topics to Cover

• Randomized Control Trials (RCT)
  Definition and Key Characteristics
• Advantages and Disadvantages of RCTs
• Examples of Large-Scale RCTs
• Single Subject vs. RCT Research Designs within Schools
• Integrating Single Subject and RCT Designs in a 3-Stage R & D Process
• Criteria Used to Classify Evidence-Based Intervention and Practice Approaches
Changed Landscape of Prevention Science and Public Demand for Reliable Evidence on What Works

• Increasing demand for large-scale prevention studies that demonstrate positive outcomes.
  – Typically, RCTs are the preferred method of demonstrating such effects.
  – Other types of methods and evidence are regarded as less rigorous.

• Prevention approaches and intervention studies that meet this standard are highly regarded and lead to many positive benefits.

• Blueprint Programs Example
Randomized Control Trials and Critical Errors

• RCTs defined--Randomized Controlled Trials are studies that randomly assign individuals to an intervention group or to a control group, in order to measure effects of the intervention.

• RCTs are considered the surest way to avoid making serious errors on health, education and policy issues that can affect the public health and safety.
  – Pre/post only and control-comparison studies often produce erroneous conclusions.
  – Random assignment allows one to evaluate whether the intervention itself, or some other set of factors, causes the observed outcomes.
If Your Goal is a Date on Friday Night,

Sure, I'd love to go see a movie

then Single Subject Designs Will Suffice
However, if you want to get invited to the prom,

I would like you to be my date for the junior-senior prom.

then you need to think about a randomized control trial
Group Versus Individual Approaches to Research as Ways of Knowing and Influencing Behavioral Outcomes

• Extensive Debates in the Applied Behavior Analysis Literature in the 1970s Regarding which Approach was Superior.
  
  Gelfand and Hartman, Kazdin, Michael, Baer, and others

• Issue was not Resolved and Ultimately Ran Its Course with Protagonists and Antagonists Solidifying their Respective Positions.
Continuing Challenges for Advocates of Single Subject vs. Randomized Control Research Designs

• Need to demonstrate positive impact on individuals while also addressing population-based questions relating to public safety, health, education, social policy and so forth.

Sidman (2006) notes that:

• Influencing large groups is not where applied behavior analysis has made its greatest impact and progress.

• Not all problems are subject to the same kinds of solutions and approaches.

• Need to address the challenge of achieving widespread behavior change simultaneously.
Sidman calls for:

• Alternative approaches that can accomplish broad, population-based changes and adoption of new innovations.

• Examples are:
  – Epidemiology
  – Actuarial-based decision-making
  – Scaling up promising interventions to evaluate their efficacy and effectiveness
  – Large-scale studies of policy decisions
Seminal RCT-Derived Medical and Educational Findings

**Medical:**
- vaccines for polio, measles, and hepatitis B
- interventions for hypertension and high cholesterol
- cancer (Hodgkins, leukemia, multiple myeloma)

**Educational:**
- reduced class sizes in grades K-3
- reading improvement and phonics
- high quality early childhood development programs
A Case of Quasi-experimental Studies Leading to Wrong Conclusions Affecting the Public Health

• For years, medical advice and wisdom advised menopausal aged women to accept hormone replacement therapy (HRT).

• Based on a series of quasi-experimental studies, findings suggested a range of benefits associated with HRT (e.g. reduced coronary and cancer risks).

• A large scale RCT of hormone replacement therapy published in 1998 showed that not only was this not true, but the HRT actually increased one's risk for heart disease and cancer.
Use of RCTs in Educational Contexts

• Researchers have had considerable difficulty in selling RCTs to school administrators.

  – As a rule, educational decision-makers will insist on use of wait-list control group designs in which:

    1) usual care controls receive the same intervention within one school year.

    2) long-term followup assessments are thus not possible.

  – When parents and/or teachers discover they are in a control group condition, participant dropout is much more likely.

  – Teachers in control conditions are often motivated to enhance options for target participants to compensate for their not having access to an intervention perceived as beneficial.
Aldous Huxley

• “The Single Greatest Tragedy of Science is the Cold-Blooded Slaying of a Beautiful Theory by an Ugly Fact.”

  – Randomized Control Trials are often the Cruel Means by Which this Tragedy Occurs.
School Context

• Factors that Influence Educator Adoption of New Practices:
  – Fits seamlessly into ongoing school routines
  – Consistent with school and educator values
  – Universal versus targeted interventions
  – Solves a high priority problem or issue
  – Time and effort costs are reasonable
  – Teacher perceives s/he has the skills and resources to apply practice effectively
Albuquerque Public Schools (APS) 
Efficacy Study of the First Step to 
Success Program 
2004 – 2008 

- Four-year large scale study funded by the Institute of Education Sciences (IES).
- 200 participants in grades 1-3 involved in the study.
(APS) Efficacy Study (Continued)

• 101 received First Step (FS) -- 99 were Usual Care Controls (UCC).

• Participants, along with teachers and classrooms, were randomly assigned to FS or UCC.

• Outcome measures were collected at pre, post, and follow-up time points.
First Step Pre Post Change Scores
APS Efficacy Study
Social Skills (Teacher and Parent)

Teacher

Parent

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS,</td>
<td>83</td>
<td>95</td>
<td>89</td>
<td>98</td>
</tr>
<tr>
<td>UCC,</td>
<td>84</td>
<td>86</td>
<td>88</td>
<td>92</td>
</tr>
</tbody>
</table>

N = 101
N = 99
First Step Pre Post Change Scores
APS Efficacy Study
Problem Behavior (Teacher and Parent)

Teacher

Parent

123
121
113
119
112
111
103
110

Pre
Post
Pre
Post

Raw Score Ratings

FS, N = 101
UCC, N = 99
First Step Pre Post Change Scores
Albuquerque Public Schools Efficacy Study

Academic Engaged Time (Classroom Observations)

% Academic Engagement

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Step</td>
<td>43</td>
<td>57</td>
</tr>
<tr>
<td>UCC</td>
<td>41</td>
<td>49</td>
</tr>
</tbody>
</table>

First Step N = 101
UCC N = 99
First Step National Effectiveness Study 2006-2011

Key Features

• Implemented across Five Sites
  – California (San Jose) Year 1
  – Illinois (Chicago Area) Year 1
  – Florida (Tampa) Year 2
  – West Virginia (Huntington Area) Year 2
  – Oregon (Eugene Area) Year 2

• 432 Participants
  – 244 First Step
  – 188 Usual Care Control
## Effect Sizes

### First Step Efficacy vs. Effectiveness Study

<table>
<thead>
<tr>
<th></th>
<th>APS Efficacy Study</th>
<th>SRI Effectiveness Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>San Jose &amp; Chicago</td>
<td></td>
</tr>
<tr>
<td><strong>Social Skills</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td>.87</td>
<td>1.11</td>
</tr>
<tr>
<td>Parent</td>
<td>.54</td>
<td>.24</td>
</tr>
<tr>
<td><strong>Problem Behavior</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td>.73</td>
<td>.63</td>
</tr>
<tr>
<td>Parent</td>
<td>.69</td>
<td>.29</td>
</tr>
<tr>
<td><strong>Academic Engaged Time</strong></td>
<td>.44</td>
<td>.36</td>
</tr>
</tbody>
</table>
Single Case versus Randomized Control Trials in Determining Efficacy-Effectiveness
<table>
<thead>
<tr>
<th>Single Case Methodology</th>
<th>Randomized Control Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Analysis of level and slope changes between phases</td>
<td>• Comparison of group averages and variability</td>
</tr>
<tr>
<td>• Target single individuals</td>
<td>• Random assignment to treatment &amp; control groups</td>
</tr>
<tr>
<td>• Continuous recording of measures</td>
<td>• Pre, Post, Follow-up administration of measures</td>
</tr>
</tbody>
</table>
**Single Case Methodology**

- Allows daily inspection of covariation of Independent and Dependent Variables
- Visual inspection of intervention effects across phases
- Strong internal validity

**Randomized Control Trials**

- Allows subgroup analyses of variations in subject group performance
- Allows calculation of effect size estimates of magnitude of treatment effect
- Strong external validity
Example of the Sensitivity and Value of Using a Single Subject Design in Understanding Teacher FS Implementation

First Step Fidelity Study

(Rob Horner and Colleagues)
Teacher Fidelity of Implementation and Student Problem Behavior during First Step and First Step Plus Coach Feedback

Star shows coach present, feedback after observation

Joe

Edward

Amus

Percent of Intervals with Problem Behavior & Percent Fidelity

Card Fidelity  Problem Behavior
First Step Fidelity Study
Key Features

- Students showed low rates of problem behavior during initial teacher phase with increasing trends over time.

- Teacher fidelity initially moderate, possibly reinforced by low rates of problem behavior.

- Teachers were responsive to the feedback and perceived the additional follow-up from coaches as positive and useful.
“Given the Choice of Studying One Rat for 1,000 Hours or 1,000 Rats for One Hour Each, I Would Choose the Former.”

B. F. Skinner
Laboratory-type intensely studying one rat using a huge magnifying glass
Laboratory-type trying to study a sea of rats using a small magnifying glass
Single Subject Design

Baseline  Intervention
S1
S2
S3
Continuous Recording
RCT Design

Intervention Implemented

Pre | During | Post | Follow-Up

Control Group (N = 50)/Experimental Group (N=50)

Random assignment to Groups
Single Subject Designs

• Are highly cost-efficient

• Allow study of the interactions of treatment and dependent variables on a continuous basis

• Provides for achieving powerful experimental control
RCT Group Intervention Designs

- Are demanding to implement and manage effectively
- Logistics are often complex
- Are unacceptable to many educators
- Can be cost inefficient re: yield
Use of Randomized Control Trials and Single Case Methodology in Researching the First Step to Success Program

- Small scale RCT used to test initial efficacy of First Step
- Single case studies used to test efficacy of First Step with other target populations and settings
- Diverse populations
- Rural settings
Five-Step Research to Practice Process

Step One: Identify high priority areas and critical problem(s) needing solution

Step Two: Engage in fact-finding and initial research

Step Three: Secure external funding to support sustained effort

Step Four: Develop and evaluate a model program or practice to solve problem

Step Five: Provide technical assistance to state and federal agencies, program adopters and professionals
Three-Stage Research and Development Model

• **Research Settings**
  – Experimental Classroom
  – Regular Classroom and Playground Settings
  – Field Test Settings

• **Target Behavior Disorders**
  – Acting Out
  – Peer Aggression
  – Social Withdrawal
  – Academic Survival Skills
Behavior Management Packages

• CLASS (AO-Acting Out)
• PASS (SSK – Academic Survival Skills)
• PEERS (SE-Social Withdrawal)
• RECESS (Aggression)
Each Completed Behavior Management Package Contains:

(1) Program Materials
(2) Consultant Manual
(3) Teacher Manual and/or
(4) Playground Supervisor Manual
Intervention Delivery Process

• **Local Program Coordinator**—Masters package content and procedures, recruits behavioral coaches to implement program, maintains liaison-coordinator role.

• **Behavioral Coach**—Screens and recruits target participants, teachers and classrooms, sets up intervention program, transfers control to teacher or playground supervisor.

• **Classroom Teacher-Playground Supervisor**—Assumes control of program’s operation after behavioral coach transfers it, works with coach to troubleshoot program.

• **Target Student**—Participates in and cooperates with program during implementation.

• **Parents**—Provide support and home rewards throughout program’s operation.
RECESS

- Reprogramming Environmental Contingencies for Effective Social Skills: An Early Intervention for Aggressive Young Children

- Age-Grade Range: K-3

- Key Components:
  - Social Skills Training
  - Praise, Points and Feedback
  - Group Rewards at School – Individual Rewards at Home
  - Response Cost
  - Daily Debriefing and Coaching

- Duration: 3 months

- Target Setting: Playground
Behavioral Responses Discriminating Between Socially Negative-Aggressive and Normal Children

1. Disturbs other children (teases, provokes fights, and interrupts others).
2. Openly strikes back with angry behavior to teasing from other children.
3. Argues and must have the last word in verbal exchanges.
4. Displays physical aggression toward objects.
5. Uses coercive tactics to force the submission of peers (manipulates and threatens).
6. Speaks to others in an impatient or cranky tone of voice.
7. Says uncomplimentary or unpleasant things to other children (engages in name calling, ridicule, and verbal derogation).

Figure 1: Behavioral Responses that Discriminate
FIGURE A.4  Group I: Positive Interactions during Baseline, Intervention, and Follow-Up Phases
FIGURE A.5  Combined Effects of Social Praise, Tokens, and Cost Contingency for a Second Group of Target Subjects
Name: John
Date: 5-24-94

Regular points

Bonus points and description

Morning recess

1. Stayed calm when ball hit him in the face.

Lunch recess

1. Helped Jimmy when he fell.

Afternoon recess

INSTRUCTIONS FOR RECESS SUPERVISORS

A. Praise: Give praise frequently during each recess (at least once every 5 minutes). Keep tally in circle provided. Initial tally at end of each recess. (During days 1-7, the consultant should praise every 2 or 3 minutes.)

B. Bonus points: If child is extra thoughtful or handles a difficult situation admirably, give a bonus point, tell the child, and, if possible, jot down a brief description.

C. Point losing: Cross out one point for every rule broken. Tell child as soon as possible (see rules on clipboard).

D. Timeout: If all the regular points are lost during a given recess, have the child sit out for the duration of that recess. Bonus points are not lost. If child doesn't cooperate with this, inform the teacher or consultant.

FIGURE A.7  Experiment II: Robert and Gregory—Rate of Negative Social Responses to Peers ($\bar{X}$ = average rate of negative social responses, $T$ = treatment period, and $G$ = generalization period)
FIGURE A.6  Experiment I: Paegan and Chuck—Rate of Negative Social Responses to Peers (K = average rate of negative social responses, T = treatment period, and C = generalization period)
FIGURE A.10 Minimum Rate per Minute of Negative Interactions Exhibited by Dennis. Closed dots represent an average of two sessions; open dots represent one session. Follow-up sessions are represented by histograms. Shaded histograms represent 2-day sessions; open histogram represents one-day session. Source: From “The Use of Peer Monitors to Reduce Negative Interactions during Recess,” by B. S. Dougherty, S. A. Fowler, and S. C. Paine. In Journal of Applied Behavior Analysis, 18, 1985, 141–153. Reprinted by permission.
### TABLE 8.8  Negative-Response Rate During Preintervention and Intervention Phases of RECESS for Ten Socially Negative-Aggressive Children During Recess and Classroom Periods

<table>
<thead>
<tr>
<th>Phases</th>
<th>SUBJETS 1</th>
<th>SUBJETS 2</th>
<th>SUBJETS 3</th>
<th>SUBJETS 4</th>
<th>SUBJETS 5</th>
<th>SUBJETS 6</th>
<th>SUBJETS 7</th>
<th>SUBJETS 8</th>
<th>SUBJETS 9</th>
<th>SUBJETS 10</th>
<th>TOTAL</th>
<th>MEAN</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recess</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>.91</td>
<td>.99</td>
<td>.85</td>
<td>.97</td>
<td>1.05</td>
<td>.34</td>
<td>.55</td>
<td>.22</td>
<td>.47</td>
<td>.11</td>
<td>.69</td>
<td>.35</td>
<td></td>
</tr>
<tr>
<td>1a Consultant-operated</td>
<td>.01</td>
<td>.07</td>
<td>.06</td>
<td>.02</td>
<td>.01</td>
<td>.11</td>
<td>.01</td>
<td>.04</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1b Recess-supervisor-operated—consultant present</td>
<td>.17</td>
<td>.04</td>
<td>.03</td>
<td>.05</td>
<td>.06</td>
<td>.02</td>
<td>.02</td>
<td>.01</td>
<td>.00</td>
<td>.05</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Recess-supervisor-operated—extended to classroom</td>
<td>.02</td>
<td>.05</td>
<td>.11</td>
<td>.24</td>
<td>.32</td>
<td>.09</td>
<td>.07</td>
<td>.01</td>
<td>.18</td>
<td>.14</td>
<td>.14</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>3 Fading</td>
<td>.08</td>
<td>.12</td>
<td>.13</td>
<td>—</td>
<td>.38</td>
<td>.13</td>
<td>.10</td>
<td>.02</td>
<td>—</td>
<td>—</td>
<td>.12</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td><strong>Classroom</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1 Prerecess program</td>
<td>.00</td>
<td>—</td>
<td>.28</td>
<td>.11</td>
<td>.22</td>
<td>.25</td>
<td>.00</td>
<td>.46</td>
<td>.41</td>
<td>.10</td>
<td>.26</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>B2 Postrecess program</td>
<td>.02</td>
<td>.19</td>
<td>.15</td>
<td>.25</td>
<td>.07</td>
<td>.28</td>
<td>.00</td>
<td>.15</td>
<td>.26</td>
<td>.07</td>
<td>.15</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>2 Class extension</td>
<td>.00</td>
<td>.06</td>
<td>.04</td>
<td>.18</td>
<td>.09</td>
<td>.08</td>
<td>.00</td>
<td>.05</td>
<td>.06</td>
<td>.05</td>
<td>.07</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>3 Fading</td>
<td>.01</td>
<td>.00</td>
<td>.01</td>
<td>—</td>
<td>.04</td>
<td>.13</td>
<td>.00</td>
<td>.12</td>
<td>—</td>
<td>—</td>
<td>.05</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Experimental Students</th>
<th>Control Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referral/Baseline</td>
<td>Mean = 1.06</td>
<td>SD = .82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recess Only</td>
<td>Mean = .03</td>
<td>SD = .04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom Extension</td>
<td>Mean = .10</td>
<td>SD = .10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fading</td>
<td>Mean = .06</td>
<td>SD = .10</td>
</tr>
</tbody>
</table>

*Table 1: Rate of Negative-Aggressive Behavior Table*
Criteria Used to Classify Evidence-Based Intervention and Practice Approaches
Levels Of Evidence For Evidence-Based Practices

• Level I: Best Support
  
  (1) At least 2 successful RCTs with follow-up
  
  (2) A large series of single case studies (N = > 9) demonstrating efficacy
  
  (3) Intervention is manualized
  
  (4) Client samples clearly specified
  
  (5) Successfully replicated at least twice (by others)
Additional Criteria Often Used For Level I

• 1-year or greater sustainability of treatment effects

• Positive Cost-Benefit Analyses

• High Fidelity of Implementation

• Robust Effect Size Relative to Controls
Level II: Good Support

(1) Two successful RCTs using wait-list control group or

(2) A small series of single case experiments (>3) demonstrating efficacy (specify group, manuals, comparative)
Level III: Moderate Support

(1) One RCT demonstrating efficacy equal to an established treatment or

(2) A small series of single case experiments (>3) demonstrating efficacy (specify group, manuals, comparative)
Level IV: Minimal Support

(1) Intervention does not meet criteria for Level I, II, or III.

Level V: Known Risks

(1) At least one study demonstrating harmful effects
Recommended Resources


ACCESSING THIS POWERPOINT ON THE WEB

• http://ivdb.uoregon.edu

• Click on Hill Walker—ABAI Education Conference Presentation, Reston, Va., Sept. 6, 2008