Evaluation Designs with Small Samples

December 2, 2010
Webinar overview

- This webinar will be recorded, so that it can be available on the Centre’s website as an educational resource
- The slides and any supplementary materials will be sent to participants at the end of the session
- The session will take approximately 45 minutes including time for questions
- You may ask questions or make comments at the end of the session by:
  - Typing your question or comment
What is the Centre?

Bringing people and knowledge together to promote the best mental health and well-being for every child and youth

Knowledge | Capacity | Partnerships

Full continuum of effective and accessible mental health services for children and youth
Evaluation designs with small samples

- Karen Milligan from Integra and
- Kim Harris from the Centre for Children and Families in the Justice System of the London Family Court Clinic
Roadmap to Establishing YourTreatment as Evidence-Based

- Treatment results in observed benefits
- Must be shown in controlled research
- Randomized Clinical Trial (RCT) is gold standard
- Must be replicated in 2 controlled studies by independent research

⇒ EXPERIMENTAL CONTROL IS ESSENTIAL

But an RCT is not always feasible in clinical settings…

Chambless & Hollon, 1998
Efficacy vs. Effectiveness

**Efficacy** – the extent to which a treatment brings about a beneficial outcome under *optimally controlled conditions*.

**Effectiveness** – the extent to which a treatment brings about beneficial outcomes under *usual conditions*.
Challenges Faced in Applied Research

- Ethical considerations
  - All clients need intervention (can’t randomize to a no treatment condition)
  - Clients need intervention immediately and cannot wait until end of study
- Cost and time required to provide treatment and control groups with the treatment
- Number of clients that need the same treatment at the same time
- Small number of clients in a treatment
- Group Design – no standard treatment to compare to
Pre-Post (Cohort) Evaluation

- Tells us if there is a change in an outcome from beginning to end of treatment.
- Can help us in identifying key outcomes.

However....

- Cannot prove that the treatment causes the impact.
- Outcome could be due to other factors (e.g., time, maturation).
- Only based on two time points in time (cannot look at pattern or trend).
How can we address these challenges?

Multiple Baseline Designs

(Time Series, Single Subject Designs)
Time Series Designs: AB Design

- Essential foundation to understanding multiple baseline designs but not a research design in and of itself.
- A = current level of behaviour/outcome prior to intervening.
- B = current level of behaviour/outcome after intervention is introduced.

| A = Baseline Phase | B = Intervention Phase |
Multiple Baseline

- At least 2 directly observable and well defined phenomena are tracked

- Phenomena can be:
  - Different *clients*
  - Different *behaviours*
  - Different *settings*

- Control over possible confounding variables (e.g., maturation) is established by staggering the introduction of treatment.
Multiple Baseline Designs: The Basics

- Baseline data are gathered on two or more participants.
- Each baseline is tracked for an extended period of time.
- After stable baselines are established, intervention is introduced to only one of the baselines (i.e., one participant).
- After a stable pattern of behaviour is established, intervention is applied to a second baseline (i.e., the next participant).
- Repeat until intervention has been applied to all baselines.
Multiple Baseline Designs: The Basics

Essentially…

CHANGE THE OUTCOME AFTER INTRODUCTION OF THE INTERVENTION

+ 

ABSENCE OF CHANGE IN OUTCOME FOR BASELINE PARTICIPANTS

= 

CHANGE IS DUE TO INTERVENTION
FIGURE 2. Percentage of off-task behavior for all baseline and treatment sessions. Shaded areas represent the proportion of each session during which children worked independently.
Percentage of On-Task Behaviour

Sessions

Child 1

Child 2

Child 3
Multiple Baseline

How do you know the treatment worked?

Visual analysis of graphed data
- Behaviour change occurs with the introduction of the treatment and remains unchanged in other baselines
- Pattern = height, slope, and shape

Data are most compelling when:
- Behaviour change occurs immediately with the introduction of the treatment
- Data in baseline and treatment do not overlap

Statistical Analysis software programs are also available.
How do I select my outcome variable?

- Keystone variable
- Must be well-defined/operationalized
- Must be able to observe or administer across a number of time points (e.g., no practice effects).
- Pattern of high frequency (e.g., impacts on daily functioning and seen in specific setting)
- Ideally, change is readily seen when treatment is introduced
- Pattern of change is predictable
- Pattern of change is replicated
How many participants and observations are enough?

Participants:
- 3 or more
- More participants = more experimental control

Observations:
- At least 3 points in baseline recommended BUT most important to establish stability or predictable pattern
- Do not want to start treatment with an increase in baseline
- More time points establishes the pattern with more certainty
Challenges

- Similarity in participants and interventions to ensure that they would be impacted by external variables in a similar manner
- When stable baselines cannot be established
- When independent baselines cannot be established
- If a delay in treatment would have significant adverse effects
- Multiple measurement time points – time intensive
Non concurrent multiple baseline

Watson and Workman (1981)

- Concurrent nature of MB can be a problem in applied settings

- Clients with the same target behaviour may rarely present themselves at the same point in time

- It may be unethical to delay treatment until several clients can be engaged concurrently
Non Concurrent MB: The Basics

- Determine the lengths of several baseline phases in advance of treatment evaluation
- Randomly assign the client to one of the pre-determined baseline lengths
- Carry out baseline observations and implement treatment at the pre-determined point in time
- NOTE: a stable baseline must also have been achieved
Sequential Evaluation of Reinforced Compliance and Graduated Request Delivery

**Fig. 1.** Percentage of compliance for all four children across all phases. Shaded areas represent mean baseline compliance levels to requests from the probability level being trained in that phase.
What can a Multiple Baseline Design tell you?

- Does the outcome of interest change from baseline to intervention (and follow up)?
- What is the nature of the change (timing, level, variability)?
- That the intervention is effective (at least 3 participants, 2 replications by independent researchers).
- Can only generalize findings to the participant, settings examined.
A Few Helpful Resources...


Questions?
For more information

Karen Milligan, PhD
Integra
Director of Research and Psychology, Integra
kmilligan@integra.on.ca

Kim Harris, PhD
Centre for Children and Families in the Justice System
of the London Family Court Clinic
kim.harris@lfcc.on.ca

Evangeline Danseco, PhD
Director, Evaluation & Research, Centre of Excellence
613.737.7600 Ext. 3319
edanseco@cheo.on.ca