

Evaluating Research about Education Programs for Students with Special Needs

Ruth Childs

The Purpose of Education Programs

- According to Ontario's *Education Act* , “the purpose of education is to provide students with the opportunity to realize their potential and develop into highly skilled, knowledgeable, caring citizens who contribute to their society (section 0.1(2))
- As educators, we interpret this to mean that the primary purpose of elementary and secondary education is to support the cognitive, social, and psychological development of students and their mastery of knowledge and skills defined in the curriculum (or, for students with IEPs, their mastery of knowledge and skills defined in their IEPs)
- One of the principal ways a school board achieves this purpose is by “deliver[ing] effective and appropriate education programs to its pupils” (section 169.1 (1))

Four Criteria for Choosing Programs

- Effectiveness (in supporting students' cognitive, social, and psychological development and mastery of knowledge and skills defined in the curriculum or in students' IEPs)
- Appropriateness (for an education context)
- Efficiency (resource requirements)
- Equity (in determining which students have access to the program)

Research Designs

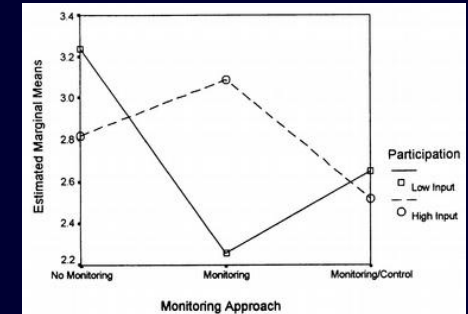
Susan Elgie

Some Research Designs

- Randomized Control Trial (RCT)
- Quasi-Experimental Design (QED)
- Non-Equivalent Groups Design
- Causal Comparative Research
- Correlational Research
- Intervention Research
- Single Case Design
- Mixed Methods Designs
- Grounded Theory
- Ethnology
- Survey
- Meta-analysis
- Interrupted Time Series
- Longitudinal (RTI)
- MORE

We started with the idea that to evaluate research, you need to know the design

Randomized Controlled Trial (RCT)



Sampling:

2 or more groups

Assignment of individuals or clusters of individuals to groups **must be randomized**; may be stratified or matched

Intervention?

Usually yes, for the treatment but not for the comparison group

Setting: Often in a lab setting

Measurement:

Repeated, at least before and after treatment

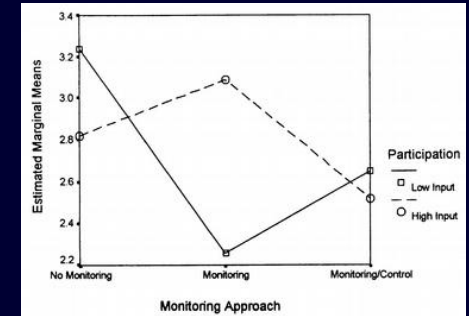
Analysis:

Inferential statistics, often ANOVA

Sample paper:

Given et al. 2008. A randomized, controlled study of computer-based intervention in middle school struggling readers. *Brain and Language*, 106, 83–97

Non-Equivalent Group Design



Sampling:

2 or more groups

Pre-existing groups, non-random assignment, e.g. class—groups should be randomly assigned to treatment

Intervention?

Usually yes, for the treatment but not for the comparison group

Setting: School-based

Measurement:

Repeated

Analysis:

Inferential statistics, often ANOVA/ANCOVA

Sample paper:

Woodward, J. & Brown, C. (2006). Meeting the curricular needs of academically low-achieving students in middle grade mathematics, *Journal of Special Education*. 40, 3.

Single Case Design (SCD)

Sampling:

Study is of single cases, typically 4 or 5.

Cases may be individuals or clusters of individuals

Purposeful selection

Intervention:

Active treatment or intervention

Some designs involve introduction & withdrawal of treatment:

Setting: Often school-based

Measurement:

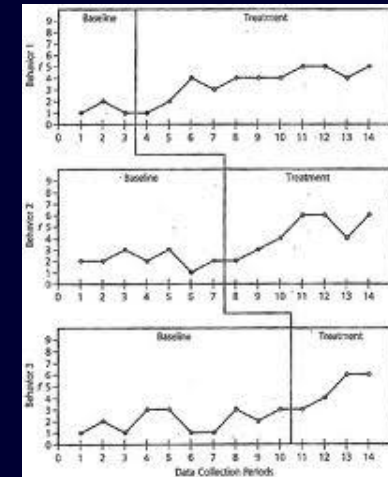
Frequent measurement both before & after treatment

Standard is at least 3 replications of measurement at 3 times or situations

Analysis: Often **graphical**: level, trend, variability

Sample paper: Friel-Patti, S. (2001). Case studies of children using Fast ForWord.

American Journal of Speech-Language Pathology. 10: 203-215.



Correlational Design

Sampling:

- **Random selection** is important
- May be records studies or surveys
- Usually not group-based

Design:

- There is usually not an intervention

Setting: Often distant

Measurement:

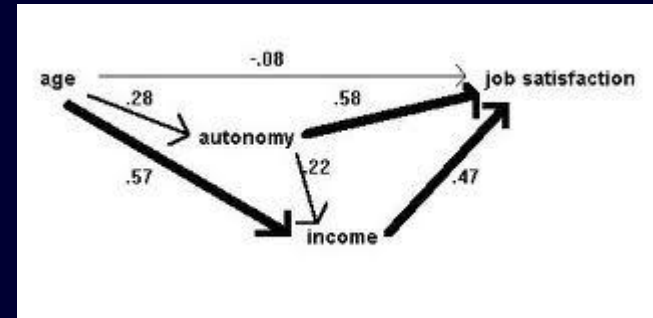
- Quantitative, may be repeated but often not

Analysis:

- **Statistical**, such as multiple regression, path analysis, hierarchical linear modeling

Sample paper:

Jimerson, S. et al. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*. 91, 116-126.



Qualitative Studies

Sampling:

Purposive, theoretical, non-random
Often a rather small number of participants

Setting:

Typically field work, non-obtrusive

Measurement:

Continuous observation, interview, collection of artefacts,

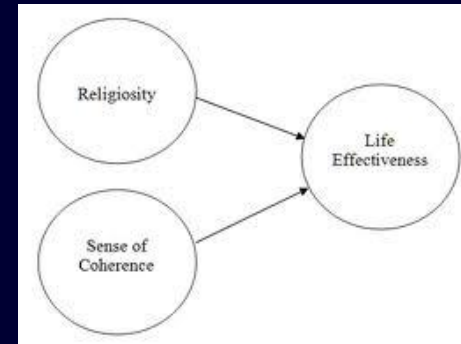
Analysis:

Segmentation, categorization, description

Thematic

Sample paper:

Barga, N.K. 1996. Students with learning disabilities in education: Managing a disability . *Journal Of Learning Disabilities*, 29, 413-421.



???

- Rules for evaluating studies are attached to research designs
- Research designs are complex
- Research papers are complex
- It is very difficult to categorize the studies we read according to the research design
- What to do?

Simplify: Design Building Blocks

- Studies that compare groups
 - Randomly assigned
 - Pre-existing groups
 - Categories, e.g. blind vs. sighted
- Studies with repeated measures
- Correlational studies
- Analysis Methods
 - Inferential statistics
 - Thematic
 - Graphical

We can describe designs using building blocks:

Randomly Assigned Groups + Repeated Measures =
Random Control Trial

Repeated Measures + Visual Analysis =
Single Case Design

People in Categories + Repeated Measures + Correlational=
Longitudinal Study

Qualitative + Correlational =
Mixed Methods

Etc.

And Evaluate Designs Using Rules for Building Blocks —it's easier!

- Judge each Block against its criteria
 - E.g. groups that are to be compared should be equal before intervention or time
 - With repeated measures designs need to check that role of maturity is controlled
 - With correlational designs, random sampling is important
 - Etc.
- & Combine!
 - DOES THIS MAKE SENSE???

Evaluation of Special Education Research

Jayne Herman

When is a practice 'evidence-based'?

- Different synthesis organizations have different standards
- Standards in research evaluation literature may be inappropriate or impossible in special education

What makes Evaluating Special Education Research Unique?

- Small sample sizes
- Heterogeneous populations
- Educational Context

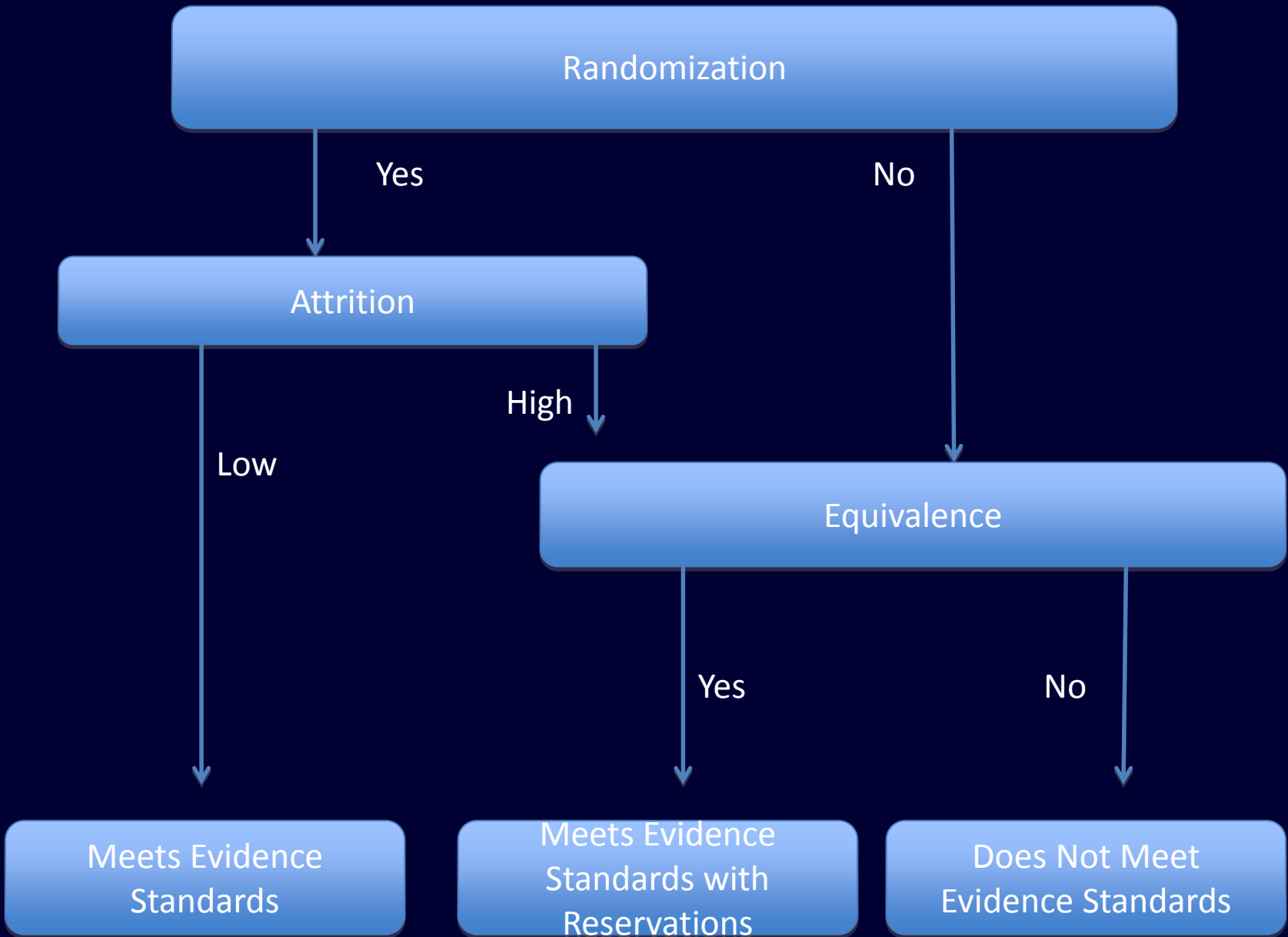
What Works Clearinghouse (WWC) Overview

1. Literature Review
2. Process for evaluating individual studies
3. Synthesize studies

WWC Evaluation Phase

Rate the design of each study into 3 categories

- Meets Evidence Standards
- Meets Evidence Standards with Reservation
- Does Not Meet Evidence Standards



WWC Synthesis Phase

- 1. Positive Effects
- 2. Potentially Positive Effects
- 3. Mixed Effects
- 4. No Discernable Effects
- 5. Potentially Negative Effects
- 6. Negative Effects

WWC- Group Research

Positive Effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Two or more studies showing *statistically significant* positive effects, at least one RCT
- No studies showing *statistically significant* or *substantively important* negative effects.

WWC Advantages and Disadvantages

- Advantages
 - Comprehensive
 - Clear ratings for intervention
- Disadvantages
 - Evidence standards are “rigid”
 - technical/ time consuming

Council For Exceptional Children (CEC) Division for Research

- Published a list of essential and desirable quality indicators to evaluate research in special education
- Designs included are: group, single case, qualitative, regression
- Recommendations for when a practice can be deemed evidenced based

CEC- SCD Research

- Quality indicators for: description of participants/settings, dependent variable, independent variable, baseline, internal validity, external validity, social validity
 - The design provides at least 3 demonstrations of experimental effect at 3 different points in time
 - dependent variables are described with operational precision

WWC/CEC - SCD Synthesis

- Need a minimum of 5 SCD papers that meet evidence standards to be summarized
- Studies must be conducted by at least 3 different research teams at 3 different geographical locations
- The combined number of experiments across the papers totals at least 20

Ontario Context & Small Group Discussions

Christie Fraser

What is unique about Special Education in Ontario?

- policy
- curriculum
- school board responsibilities (SEAC)
- populations
- definitions
- programs
- cultural sensitivities
- IEPs

Fictional Scenario

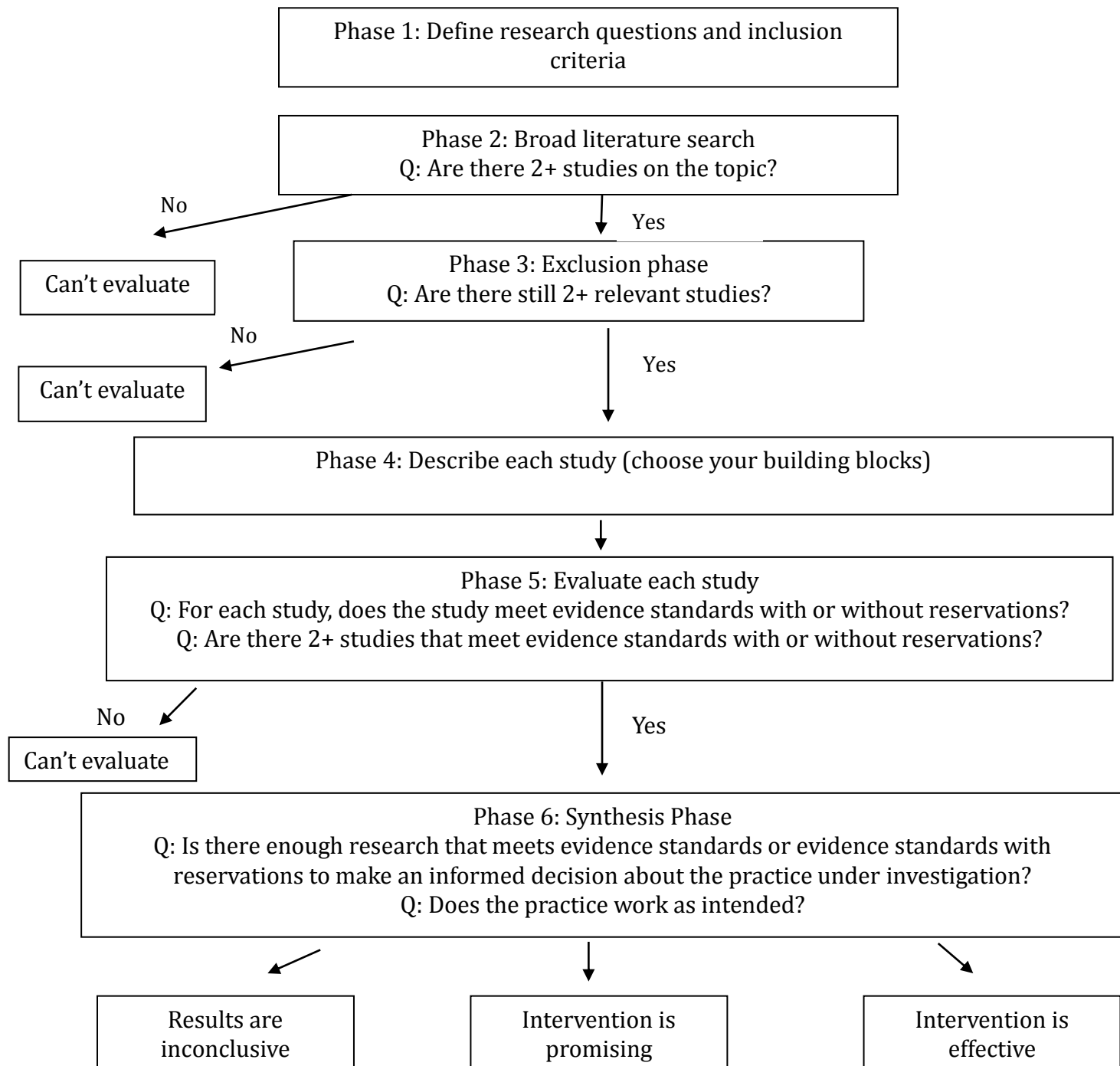
You work in an Ontario school board called Whispering Pines DSB. The board area contains 1 medium to large city, several small cities and rural areas. WPDS is a public board and has 60 elementary and 15 secondary schools under its jurisdiction. Two of the elementary schools are located on reserves and have only aboriginal students. About 10 and 15% of students in the cities need ESL support and a smaller proportion in the rural areas. SES is typical for this rural/urban mix.

Fictional Scenario Con't

Your superintendent comes to you very enthusiastic and really ready to immediately adopt an intervention for special education students, perhaps PALS for junior-level mathematics. The SO asks what you think—what does the research say? You pull up a few studies, read them and find yourself in a dilemma. There are a few studies and they are very difficult to evaluate. What do you do?

Small Group Discussion

- ① Are there demands for this type of research/evaluation in your organization?
- ② How would you currently go about determining whether programs/services for students with special needs are appropriate, effective and equitable?
- ③ What resources are used when evaluating these programs/services?
- ④ What resources do you need to better evaluate these programs/services?



Small Group Discussion Con't

- ⑤ Please go through the decision tree and discuss each phase in terms of its ease and utility.

Question Prompts:

- In what ways must a study be similar to the research question in order to be relevant—e.g. grade, age group, subject area? (inclusion criteria)
- Where should we recommend that people look?
- Do you see issues in this step that we should identify?
- Which building blocks, or research components do you think are needed in this step?
- What do you think are the most important criteria for evaluating research studies?

- ⑥ What obstacles may prevent you from implementing the proposed research/evaluation process in your organization? Conversely, what factors may facilitate?