

BEYOND OUTCOMES FOCUSED EVALUATION: INVESTIGATING AND PREDICTING IMPLEMENTATION QUALITY

RYAN J. GAGNON, M.A.

BARRY A. GARST, Ph.D.



WHAT IS IMPLEMENTATION RESEARCH?

- The science and study of taking a program to effect, that is identifying the factors that may contribute to the **design**, **delivery**, and/or **selection** of a program

Eccles & Mittman, 2006; Sloboda, Dusenbury, & Petras, 2014

GUIDING QUESTIONS

- Why study implementation and its relevant components
 - Why study “it” and What is “it”
- Measurement of implementation quality
 - How do we measure “it”
- Linking implementation quality to program outcomes
 - Does “it” matter

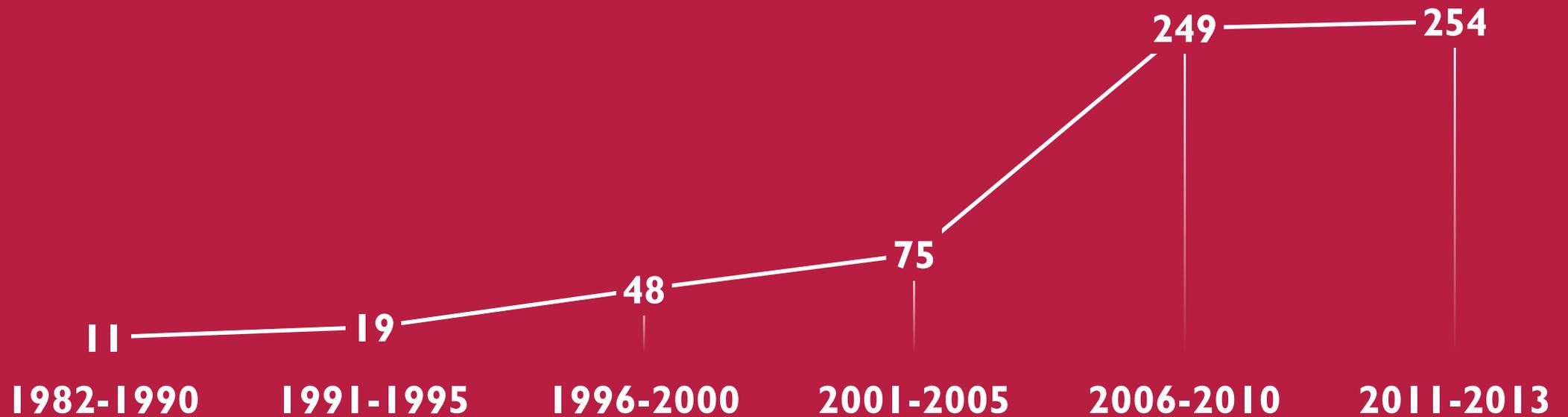
WHY DO I STUDY IMPLEMENTATION?

- Because I was doing “it” poorly
- My performance was lagging behind my less experienced and less trained staff and peers
- I am an academic so I over think



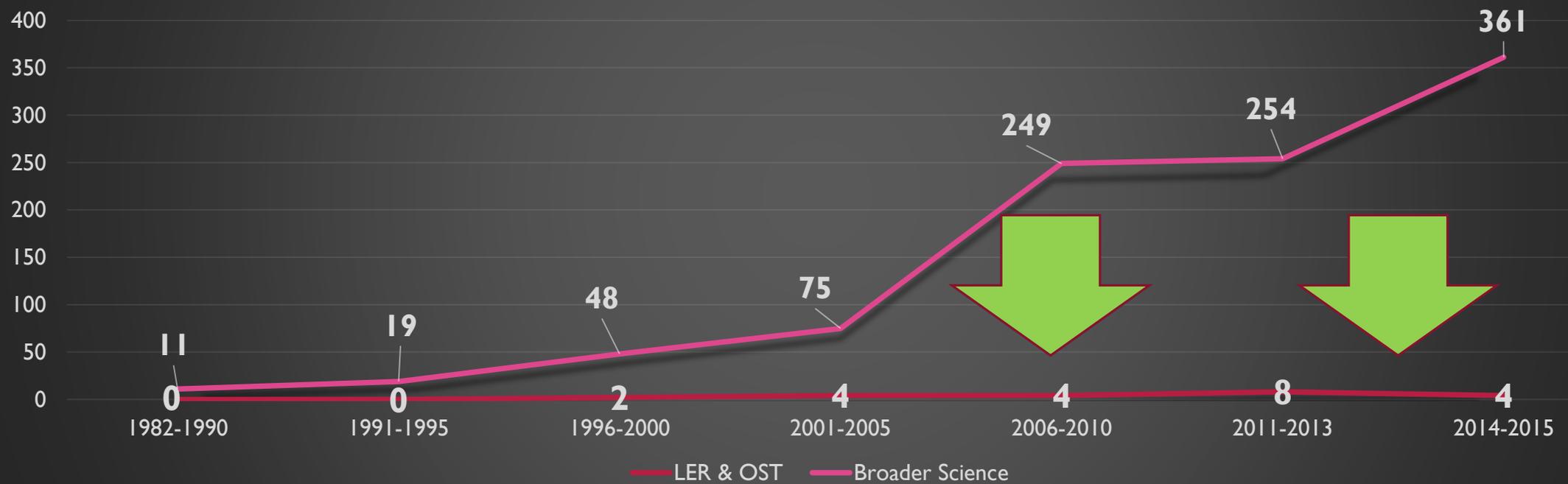
THE STUDY OF IMPLEMENTATION IS GROWING!

IMPLEMENTATION SCIENCE REFERENCES IN BROADER MEDICAL AND SOCIAL SCIENCES



THIS TREND IS NOT REFLECTED IN OUT-OF-SCHOOL TIME, LEISURE, EXPERIENTIAL EDUCATION, AND RECREATION PROGRAMS

Comparing LER & OST Implementation Assessment to Broader Social & Medical Science



WHY NOT?

- Unstructured nature of programs?
- Lack of resources?
- Lack of evidence based programming?
- Resistance to external standards and structure?

Tucker & Rheingold, 2010



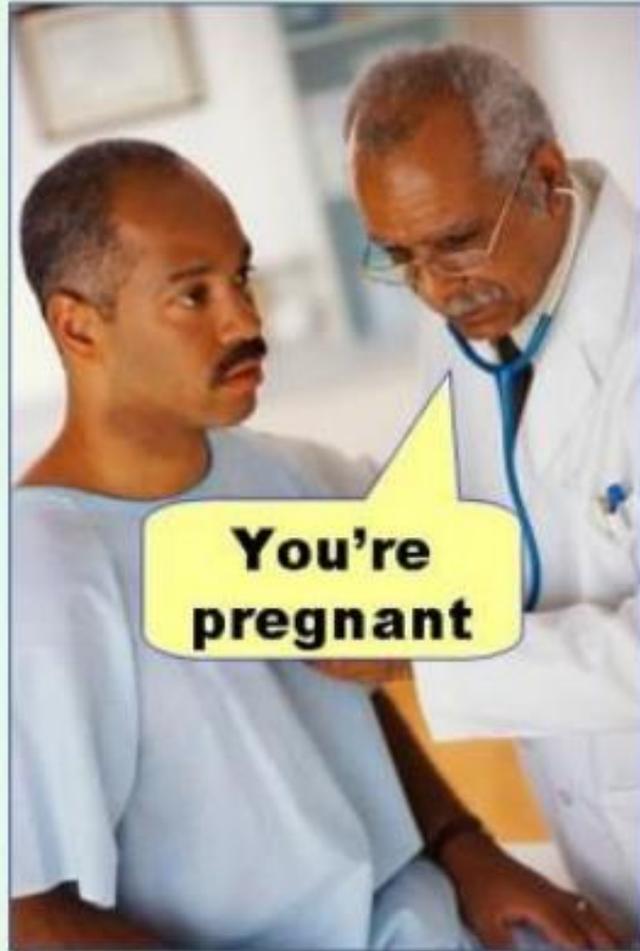
-
- What other barriers have you (your organization) experienced re: the assessment of implementation?

IMPLEMENTATION ASSESSMENT PROVIDES CLUES TO PROGRAM SUCCESS OR FAILURE

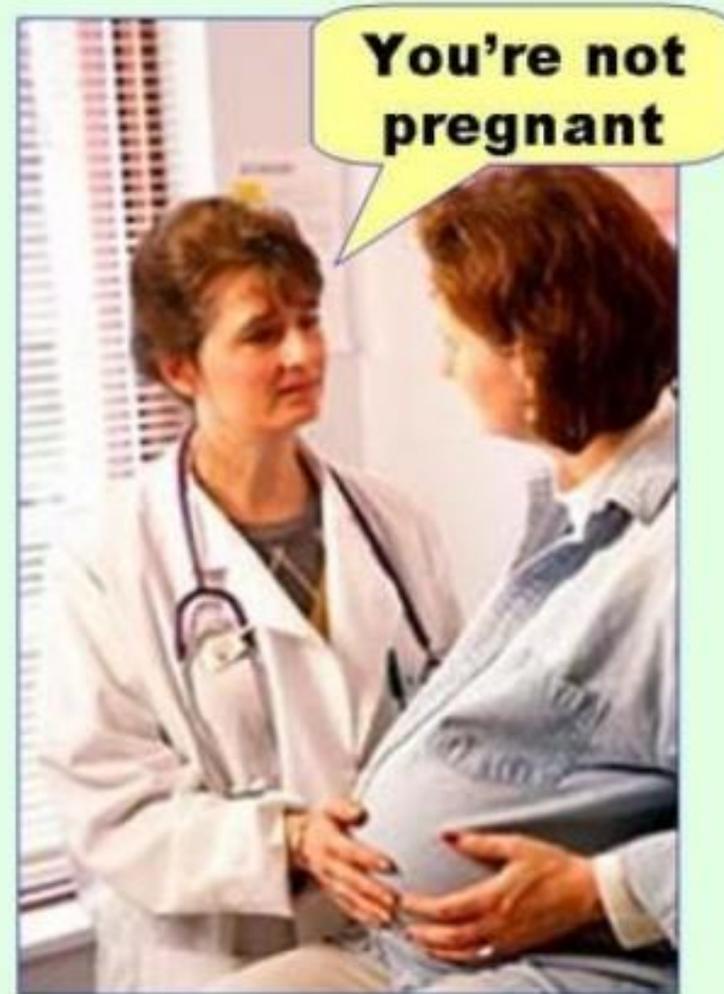
- “Small effect sizes or findings inconsistent with well-reasoned hypotheses may not be related to the efficacy of the program as it was designed, but rather be related to failure to implement the program as intended”

Caldwell et al., 2008

Type I error
(false positive)



Type II error
(false negative)



IMPLEMENTATION ASSESSMENT TELLS US ABOUT THE REAL WORLD

- Determines if research-based programs are **practical** in real-world settings
- Ensures that programs are **delivered consistently** across sites
- Highlights potential explanations for **omissions or modifications** to a program

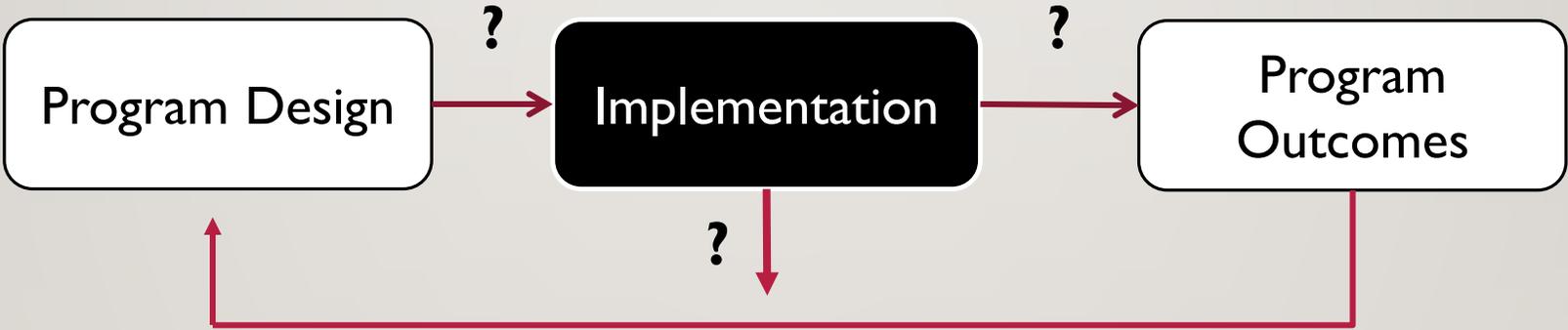
Duerden & Witt, 2012; Fixsen, Blase, Naoom, & Wallace, 2009; Johnson, Mellard, Fuchs, & McKnight, 2006

IMPLEMENTATION ASSESSMENT ALLOWS US TO IMPROVE

- If a program is not able to be delivered as designed this may highlight **a need to alter** the program to better suit the needs of the community it is intended to serve
- Documented intentional or unintentional modifications to a program can lead to “**better**” program outcomes

Dane & Schneider, 1998

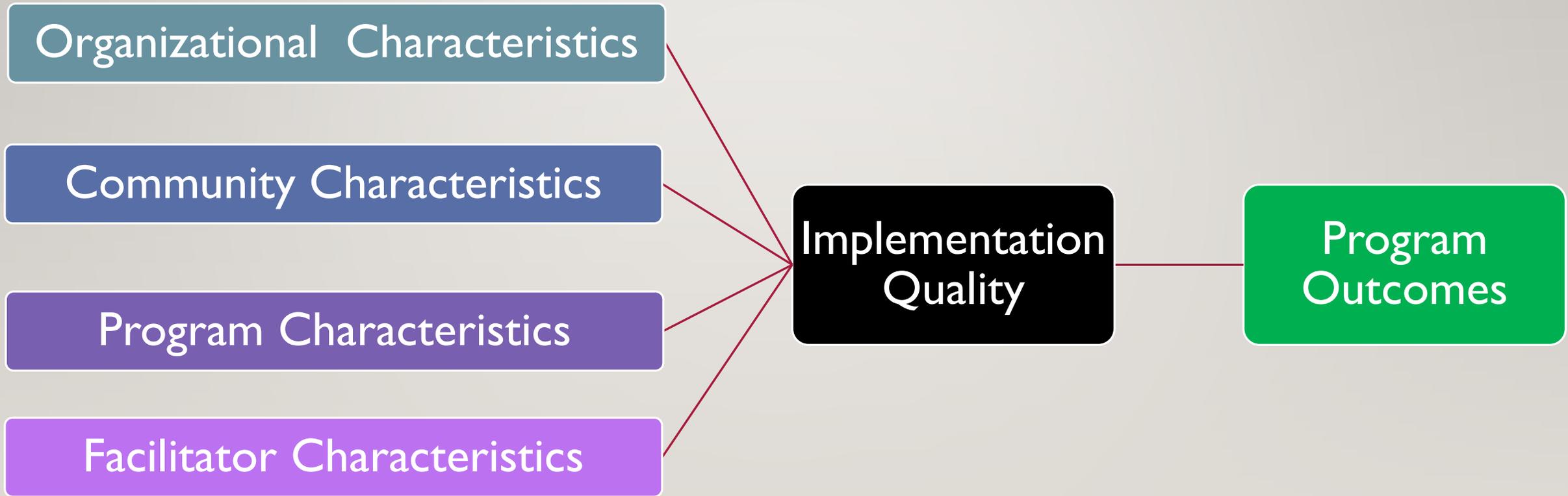
IMPLEMENTATION ASSESSEMENT TELLS US MORE ABOUT PROGRAM SUCCESS AND FAILURE



PAPER ONE: A SYSTEMATIC REVIEW OF FACTORS THAT INFLUENCE FACILITATOR IMPLEMENTATION QUALITY AND RELATED CONCEPTS



IMPLEMENTATION QUALITY IS INFLUENCED AT MULTIPLE LEVELS



Organizational Characteristics

```
graph TD; A[Organizational Characteristics] --> B[Administrative Support]; A --> C[Leadership Quality]; A --> D[Technical Support]; B --> E[Logistical, Financial, & Personnel]; C --> F[Buy-In & Vision]; D --> G[Training & Technical Assistance];
```

Administrative Support

Logistical, Financial, & Personnel

Leadership Quality

Buy-In & Vision

Technical Support

Training & Technical Assistance

Community Characteristics

```
graph TD; A[Community Characteristics] --> B[Program Saturation]; A --> C[Program Fit]; A --> D[Cultural Fit]; B --> E[Too many programs]; C --> F[Developmental Fit & Need]; D --> G[Language & Ethnicity of Developer];
```

Program Saturation

Too many programs

Program Fit

Developmental Fit & Need

Cultural Fit

Language & Ethnicity of Developer

Program Characteristics

```
graph TD; A[Program Characteristics] --> B[Program Complexity]; A --> C[Program Parsimony]; A --> D[Participant Responsiveness]; B --> E[Rigor & Depth]; C --> F[Clear Input & Outputs]; D --> G[Engagement & Usefulness];
```

Program Complexity

Rigor & Depth

Program Parsimony

Clear Input & Outputs

Participant Responsiveness

Engagement & Usefulness

Facilitator Characteristics

```
graph TD; A[Facilitator Characteristics] --> B[Experience]; A --> C[Training]; A --> D[Program Support]; A --> E[Competency]; B --> B1[Program Specific & General]; C --> C1[Program Specific & Broad]; D --> D1[Motivation & Buy-in]; E --> E1[Development & Ability];
```

Experience

Program Specific
& General

Training

Program Specific
& Broad

Program Support

Motivation &
Buy-in

Competency

Development
& Ability



-
- Who has an implementation assessment process in place currently that addresses **all or some** of these four dimensions?

WE KNOW WHAT IMPLEMENTATION IS, BUT WE HAVE AN ISSUE OF **DEFINITIONAL CONGRUENCE**

- **Program adaptation** “changes made to the program, particularly material that is added to the program”
- **Program deviation** “the addition or removal of material from the original program design by the program facilitator”

Berkel et al., 2011; Century, Freeman, & Rudnick, 2008



SOLUTION?

SYSTEMATIC REVIEW AND CONTENT ANALYSIS

- Systematic review in its broadest sense is to summarize the best available research to answer **a question or set of questions**
- Exclusion and/or inclusion within a systematic review utilize **clearly defined criteria** to such a degree that an outside team could easily replicate the review and achieve near identical findings

OUR QUESTION

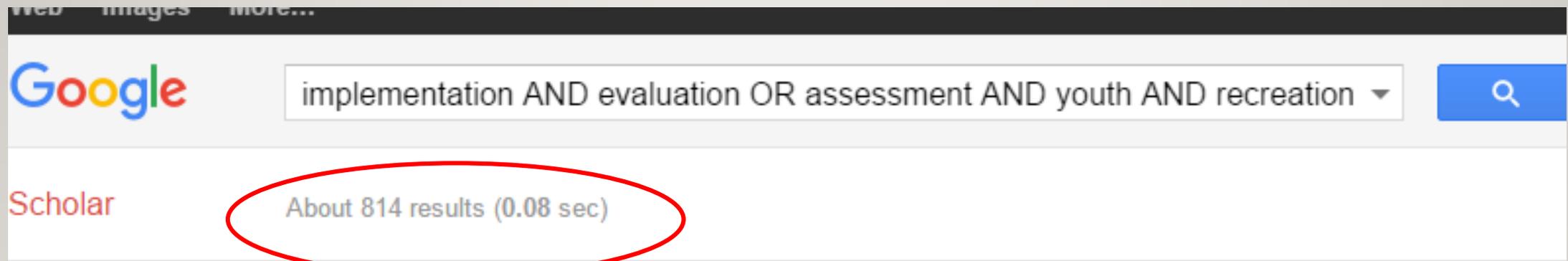
- What facilitator characteristics contribute to program implementation quality?
 - How are they defined and operationalized?

OUR CRITERIA

- 2005 and Forward
- Published Journal Articles
- Indexed on Social Sciences Citation Index
- Non-Medical/Treatment
- Non-Correctional or Justice
- Non-Academic
- Non-Individual-Centered Programs

BOOLEAN OPERATORS

- We utilized Boolean operators to limit our search



- Google Scholar Indexes:

Social Sciences Citation Index

Overcome information overload and focus on essential data from the world's leading journals

FACTORS THAT INFLUENCE FACILITATOR IMPLEMENTATION QUALITY (**BOOLEAN OPERATORS**)

- **Structural**

- Training, Program Design, Resources

- **Process**

- Buy-in, Quality of Delivery, Participant Responsiveness, Competency, Experience

- **Related Concepts**

- Fidelity
- Adherence
- Dosage
- Program Differentiation
- Adaptation
- Deviation

MANUAL REFINEMENT THROUGH ABSTRACTS

First Poll	814
Medical/Treatment	-161
Correctional/Justice	-67
Academic	-209
Individual Centered	-31
Duplicate	-93
	253

Second Poll	253
Grey	-115
Book Chapter	-8
Citation	-17
Non-Empirical	-21
	92

CONTENT ANALYSIS OF TERMS

- A research method for **the subjective interpretation** of the content of text data through the **systematic classification process** of coding and identifying themes or patterns
- In our review the **text data** are the **operational definitions** of the terms of interest

Hsieh & Shannon, 2005

FIDELITY SYNONYMS AND/OR EQUIVALENT TERMS

- Treatment Integrity
- Program Integrity
- Program Fidelity
- Implementation Fidelity
- Treatment Adherence
- Degree of Implementation

FIDELITY DEFINITIONS

- **Fidelity**, also referred to as adherence, integrity, and quality of implementation, is the extent to which the delivery of an intervention adheres to the protocol or program model **as intended by the developers of the intervention**
 - **Fidelity**: Whether prescribed program components were **delivered as instructed** in program protocol
 - **Fidelity**, the degree to which treatment is **delivered as intended**
 - **Fidelity**, the extent to which the innovation **corresponds to the originally intended program**
- 

CONSTRUCTION OF “UNIVERSAL” DEFINITION

- ...as intended by the developers of the intervention...
- ...delivered as instructed...
- ...corresponds to the originally intended program...
- ...delivered as intended...
- **Fidelity:** the degree to which a program is delivered as intended by the program developers

SO WHAT

- The goal of this process was to uncover the latent and observed definitions of concepts relating to implementation quality
- The findings indicate that definitional “crossover” and “averaging” is occurring

NOW WHAT

- We have operationalized 15 terms which will be used for measurement and analysis in further investigations
- We have provided researchers with a clear path and “clean” definitions

QUESTIONS, COMMENTS, OR WITTY ANECDOTES



PAPER TWO: ASSESSING FACILITATOR INFLUENCES ON PROGRAM IMPLEMENTATION



REMINDER: IMPLEMENTATION QUALITY IS INFLUENCED AT MULTIPLE LEVELS



WHY ARE YOU ASSESSING IMPLEMENTATION QUALITY?

- Uncover alternative methods of program delivery?
- Staff evaluation?
- Because the funder said you had to?
- To ensure program is being delivered?



DIRECT OBSERVATION

Observation Checklist

Coder: _____ Date: _____ Time: _____

Activity: _____ Number of Participants: _____

CRITERION: Quality of facilitator delivery

How **ENTHUSIASTIC** was the facilitator?

	Not at all enthusiastic	Somewhat enthusiastic	Enthusiastic	Very enthusiastic
1 st observation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2 nd observation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3 rd observation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How **CLEAR** were the facilitator's directions?

	No directions are clear	Some directions are clear	Most directions are clear	All directions are clear
1 st observation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2 nd observation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Mainieri & Anderson, 2015



BREAKOUT GUIDING QUESTIONS

- How did you rate the quality of the **facilitator's delivery**? Strengths and weaknesses?
- How did you rate the participants' **engagement and responsiveness**? Strengths and weaknesses?
- Was this observational exercise easy? How does this example differ from other experiences you've had with direct observation-based assessment?

INDIRECT

- Self-Report
- Adherence
- The FCPC
- Limitations

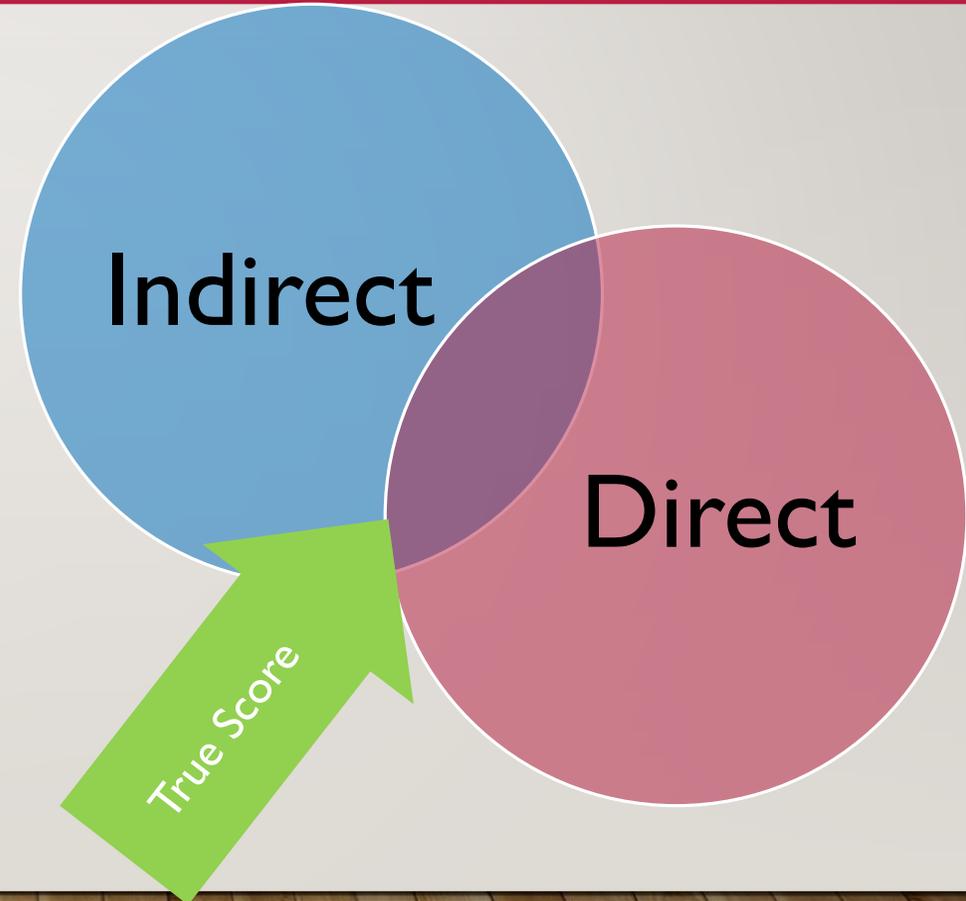
Gagnon, Stone, & Garst, 2016



HYBRID

- Both direct and indirect strategies are combined and then a “true-score” is created
- Self-reporters over-estimate
- Observers under-estimate
- Limitations

Durlak & DuPre, 2008



ALTERNATIVE METHOD ONE

PEER REPORT

- Peers complete a modified version of the self-report of their co-facilitator
- Limitations



Mainieri, 2016

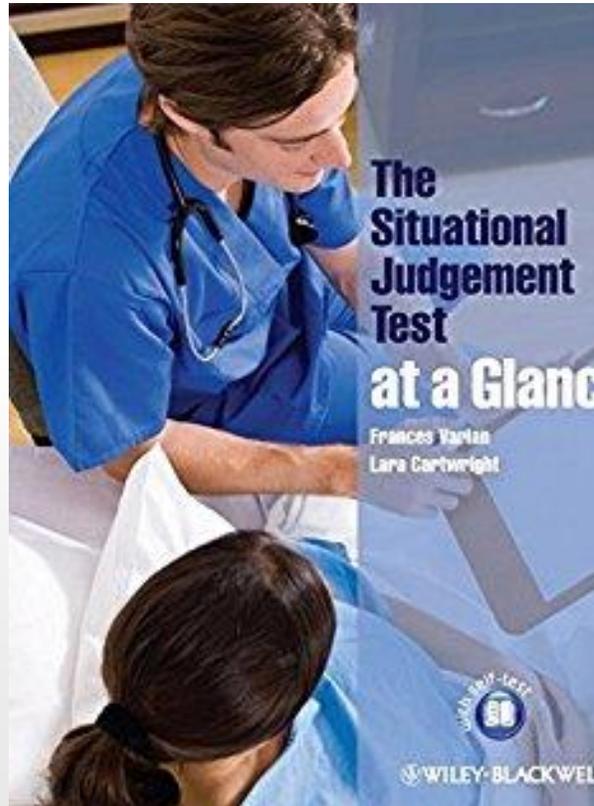
ALTERNATIVE METHOD TWO MICRO-OBSERVATION

- Highly scheduled programs have brief time-periods selected (5-10 minutes) for observation via remote video or in person
- Limitations

Pettigrew & Gagnon, 2016



ALTERNATIVE METHOD THREE SITUATIONAL JUDGEMENT TESTING



- SJTs present respondents with a brief scenario and then ask him or her to select the best choice or indicate what he or she would do

Barrett, Doverspike, & Young, 2010

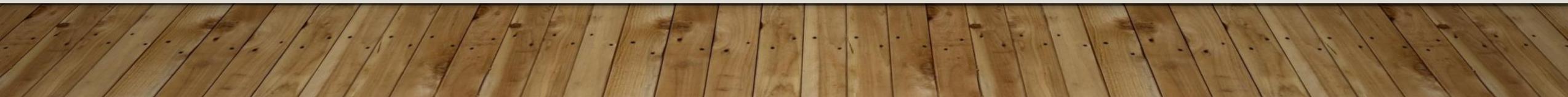
DEVELOPMENT OF SITUATIONAL JUDGEMENT TESTS

(1) Review of literature to develop constructs of interest.

(4) SMEs review and provide feedback to proposed scenarios.

(2) Development of SJT scenarios.

(3) Identification and recruitment of SMEs.



PRODUCT

Scenario 5 Program Differentiation (This information is hidden from respondents)

You are delivering a five-hour program to develop self-regulative behaviors in a group of nine 12-14 year olds. About an hour into the program, several of the participant's mention that they "played the same games" last week in their afterschool program you are not affiliated with. You know that it is important to stick with the program plan as evidenced in your own training with program. Please respond with what you believe will be the most effective strategy to achieve the program goals and maintain the program plan?

Please enter the strategy that you believe **would best achieve** the program goals

Please enter your best response here and if applicable what additional information or questions you have about the scenario.

SJT DEVELOPMENT STEPS 5-8

(5) Scenarios refined based upon SME feedback.

(6) SMEs review scenarios and provide response choices on how they would respond to given scenario based upon question(s) of interest.

(8) Finalized SJTs are implemented with cohort of interest.

(7) Responses are refined/reduced based upon alternative SME team.

PRODUCT #2

In this portion you will be presented with scenarios commonly encountered during the facilitation of programs. The scenario details may not be an exact fit to your own programs, but we'd still like to know how you would respond in the given scenario. **Read the scenario and then circle the action you would most likely take.**

	<i>You are delivering a 90-minute resilience program to a group of eight 13-15 year olds. Your co-facilitator is not adhering to the program plan and mentioned to you beforehand that she knew the teens would not be engaged. Her poor attitude towards the program is now causing the teens to disengage. Please select the strategy you believe would be the most effective to reengage your partner while maintaining program quality?</i>
Circle One	
1	Remove co-facilitator from program
2	Address co-facilitator's issue(s) in private while group participates in alternative activity
3	Take a group break and visit with co-facilitator in private
4	Introduce unrelated activity
5	It is not correctable, cancel program
6	Revisit & Refocus on program goals with co-facilitator
7	Move forward without addressing co-facilitator
8	Acknowledging and addressing co-facilitator issue in front of participants
9	It is not correctable, keep going.
10	Revisit & Refocus on program goals with entire group including co-facilitator.
11	Working with that particular co-facilitator

SJT DEVELOPMENT STEPS 9-10

(9) SJT responses are analyzed for study purpose.

(10) SJTs refined for re-implementation

PRODUCT 3

Circle One	<i>You are delivering a 90-minute resilience program to a group of eight 13-15 year olds. Your co-facilitator is not adhering to the program plan and mentioned to you beforehand that she knew the teens would not be engaged. Her poor attitude towards the program is now causing the teens to disengage. Please select the strategy you believe would be the most effective to reengage your partner while maintaining program quality?</i>
1	Remove co-facilitator from program
2	Address co-facilitator's issue(s) in private while group participates in alternative activity
3	Take a group break and visit with co-facilitator in private
5	It is not correctable, cancel program
6	Revisit & Refocus on program goals with co-facilitator
7	Move forward without addressing co-facilitator
8	Acknowledging and addressing co-facilitator issue in front of participants
10	Revisit & Refocus on program goals with entire group including co-facilitator.

SO WHAT?

- We developed an alternative method to predict implementation quality
- Utilized practices germane to corporate hiring screening practices

NOW WHAT?

- Need to see if the SJT's meaningfully and as theorized predict implementation quality and program outcomes

PAPER THREE: MEASURING THE EFFECT
OF IMPLEMENTATION QUALITY
PREDICTION ON OUTCOMES
*PRELIMINARY ANALYSES



GUIDING RESEARCH GOAL

- The goal in this portion of the paper is to examine the predictive qualities of SJTs, traditional self-reports, and facilitator characteristics on parent reported program outcomes
- What implementation qualities and factors best predict program outcomes?

OUR SAMPLES

- The samples in this portion of the study are camp counselors from Residential Summer Camps within the southeastern United States ($N = 63$) and parents of youth attending one of the camps ($N = 671$).

COUNSELOR PROCEDURES & DESCRIPTIVES

- Completed three identical measures, Pre-Training, End-of-Training, End-of-Season
- 63 Counselors completed all three measurement occasions (74 completed at least two)
- Primarily female (59.4%), currently college student (89.8%), white (61.3%), and in 2nd season (72.1%).

PARENT PROCEDURES & DESCRIPTIVES

- Completed an online survey one week after child's camp experience
- Reminder sent two weeks after camp experience
- 671 Responses after data cleaning
- 36.3% Response Rate
- Respondents were primarily female (68.3%), well-educated (74.1% with Bachelors or above), White (83.1%), and affluent (81.7% of sample at or \$60,000 per year).

PARENT OBSERVATION MEASURE, CAMP SKILL DEVELOPMENT

- Measures 9 common dimensions of the camp experience
- Both reliable and valid as indicated by psychometric testing with differing groups (e.g., CFA, Rho, Correlation)

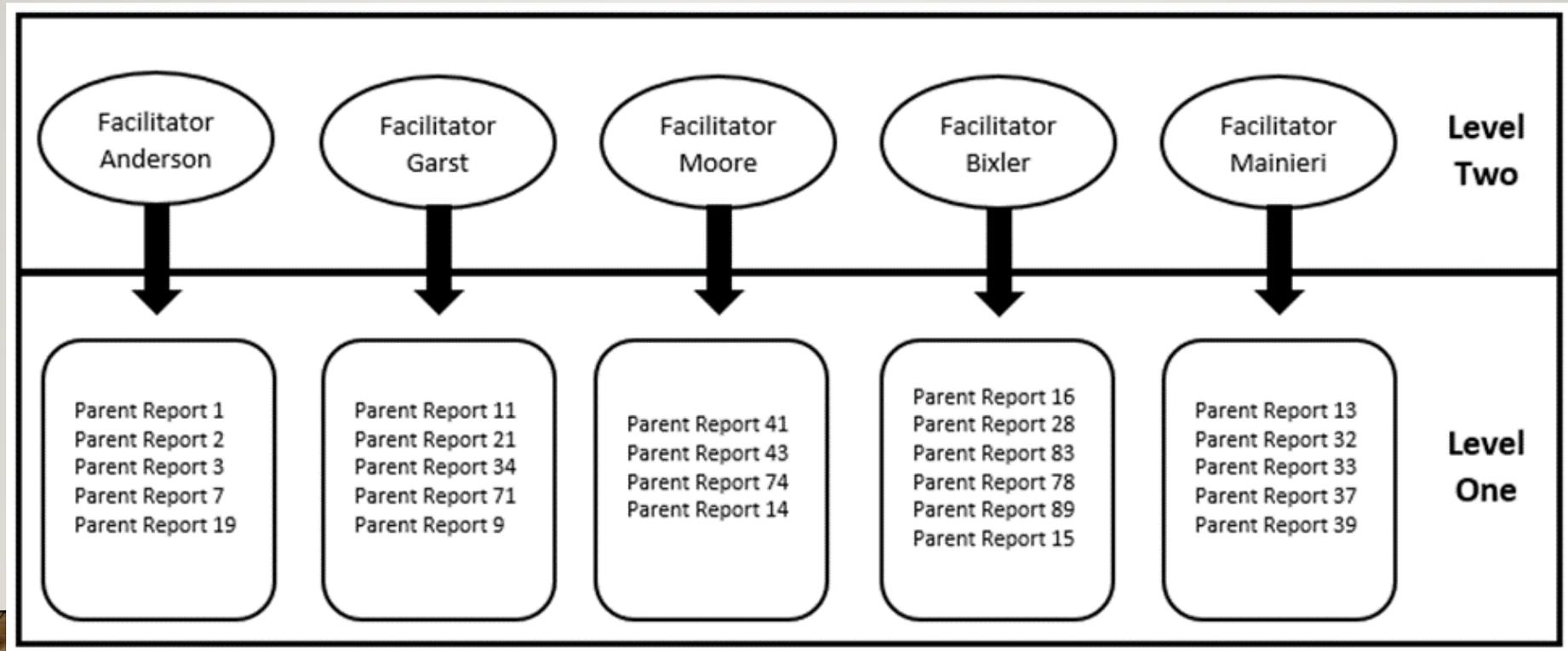
Garst & Gagnon, 2016

Primary Outcomes Associated with Camp Experience	
Outcome	Definition
Independence	The ability to start and accomplish tasks without external motivation ¹
Resilience	The ability to adapt through stress and adversity ²
Attitude	The ability to monitor and regulate behaviors ³
Cooperation	The ability to work with others towards a shared goal ⁴
Communication	The ability to articulate thoughts and feelings in a meaningful way ³
Social Competency	The ability to listen to, bond, and empathize with others ³
Critical Thinking	The ability to thoughtfully, thoroughly, logically, and creatively consider a problem ⁵
Decision Making	The ability to consider and choose the correct course of action ⁶
Exploration	The degree of desire and/or motivation to learn new topics or capability to try new/different experiences ⁵

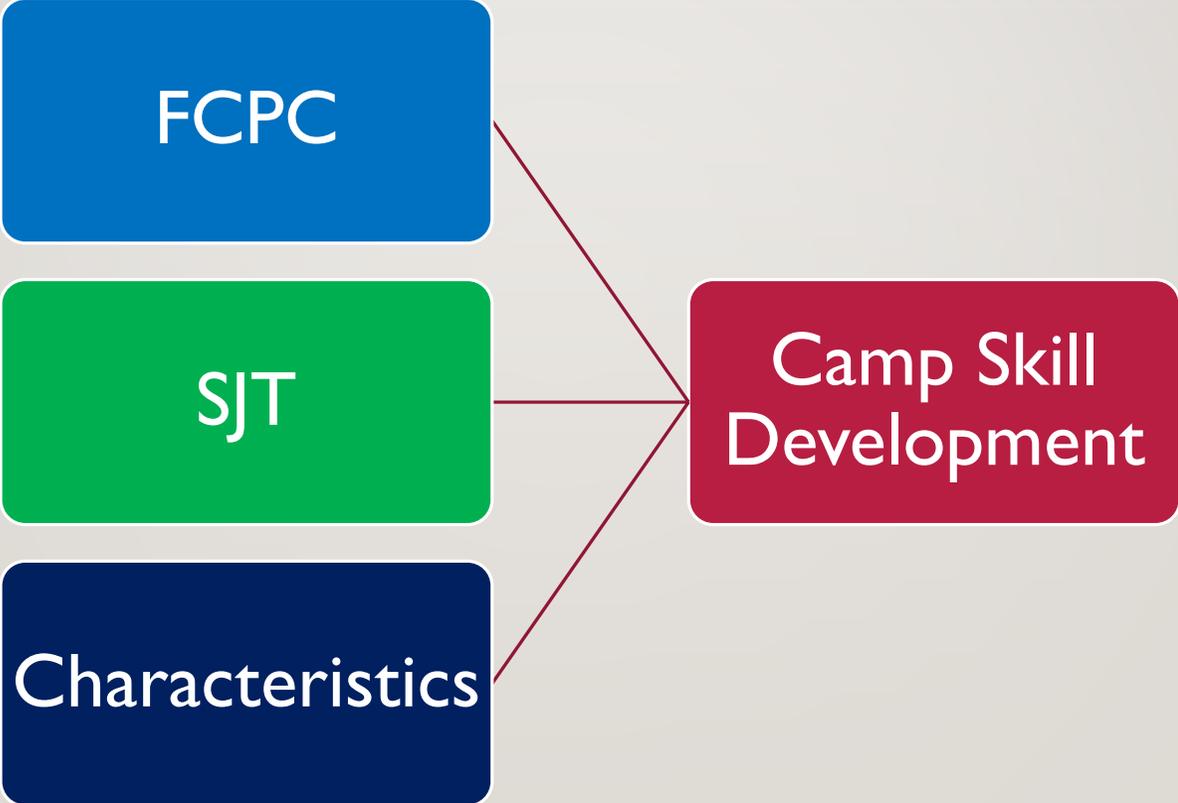
ANALYSIS

*PRELIMINARY

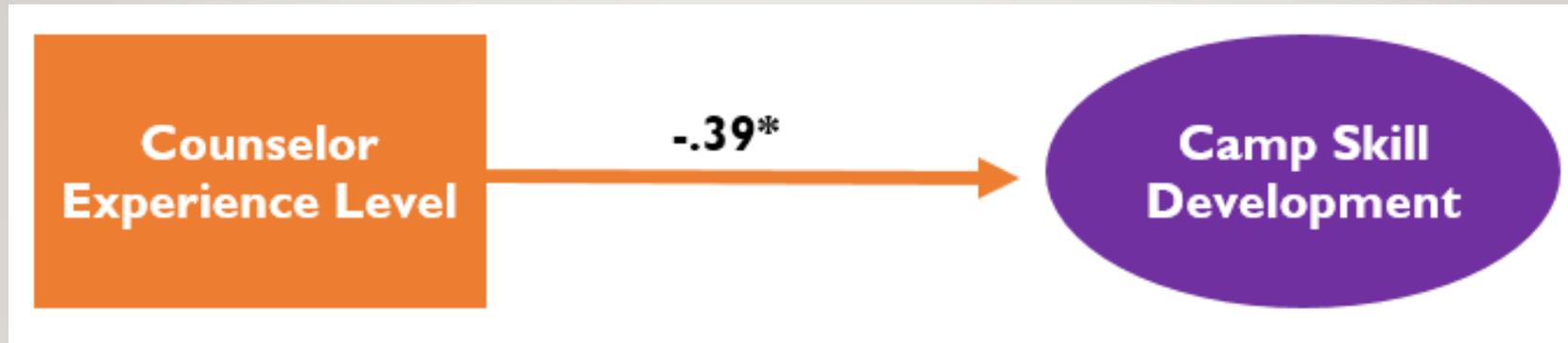
- Multilevel structural equation modeling (ML-SEM)



SIMPLIFIED LATENT HLM (AKA ML-SEM)



SIMPLE...BUT INTERESTING PRELIMINARY RESULTS

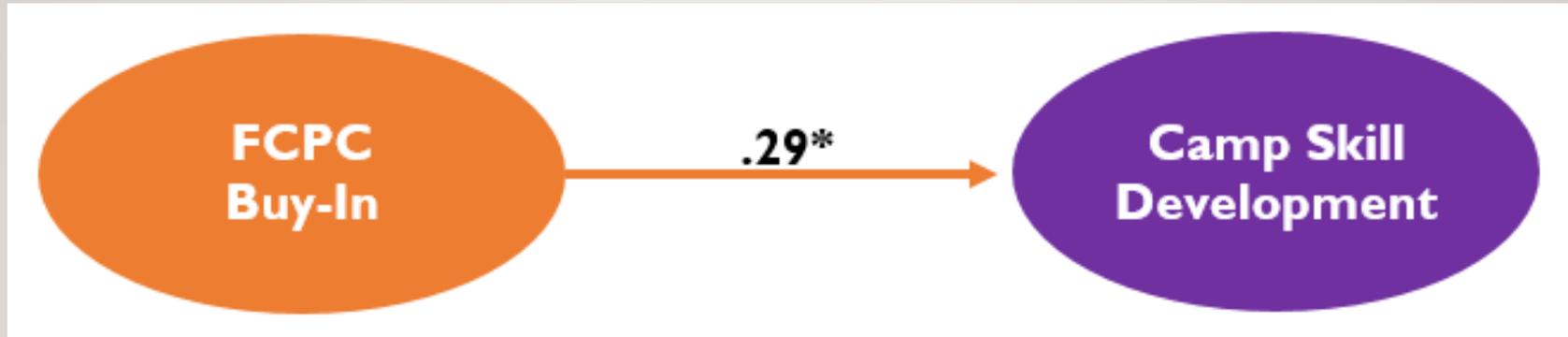


Summary of Fit Indices for Measurement Model ($N = 671$).

	χ^2	df	RMSEA	SRMR	CFI
Measurement Model	853.409	384*	.057 (90% CI, .052, .062)	.057	.936

* indicates $p \leq .001$.

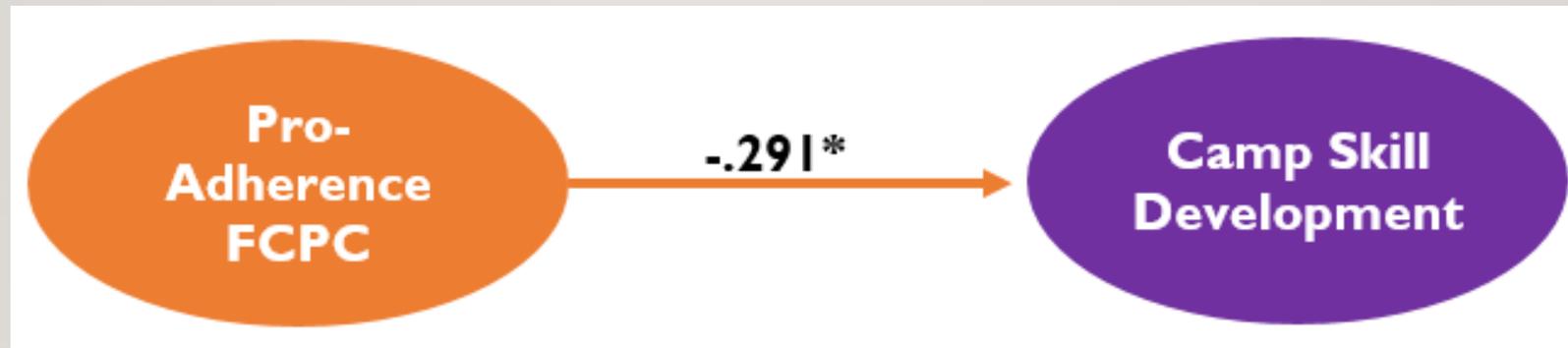
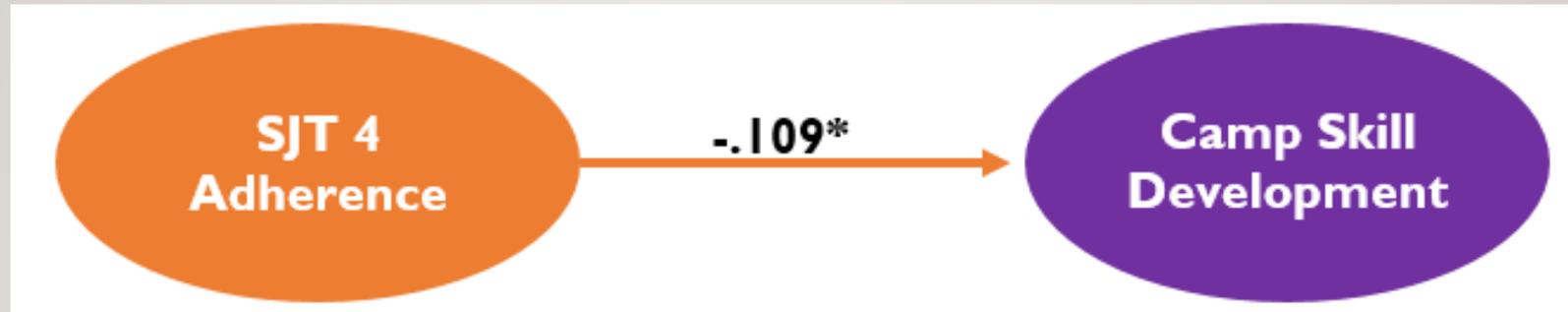
DIFFERENCES IN BUY-IN EFFECT



DIFFERENCES IN REPORTED TRAINING VERSUS TRAINING CONFIDENCE



DIFFERENCES IN PRO-ADHERENCE BELIEF AND SJT ADHERENCE



SO WHAT

- Preliminary Analyses did partly confirm some implementation theory
- This style of assessment could be used to influence hiring decisions, training, and program design
- Unique approach and less resource intensive than other styles of assessment

NOW WHAT

- Need to do comparative analyses with direct observations?
- How do SJTs perform in less or more structured programs?

QUESTIONS, THOUGHTS, WITTY ANECDOTES?

- All slides and references are available now at ryangagnon.com/presentations
- rjgagno@Clemson.edu
- bgarst@Clemson.edu