

Collaborative Dynamics, Change Activities, & Collective Outcomes in Multi-Organizational Reform: A Conceptual Evaluation Framework

Lucas B. Hill, University of Wisconsin-Madison

Julia N. Savoy, University of Wisconsin-Madison

Judy Milton, University of Georgia

Mark Graham, Yale University

Framing

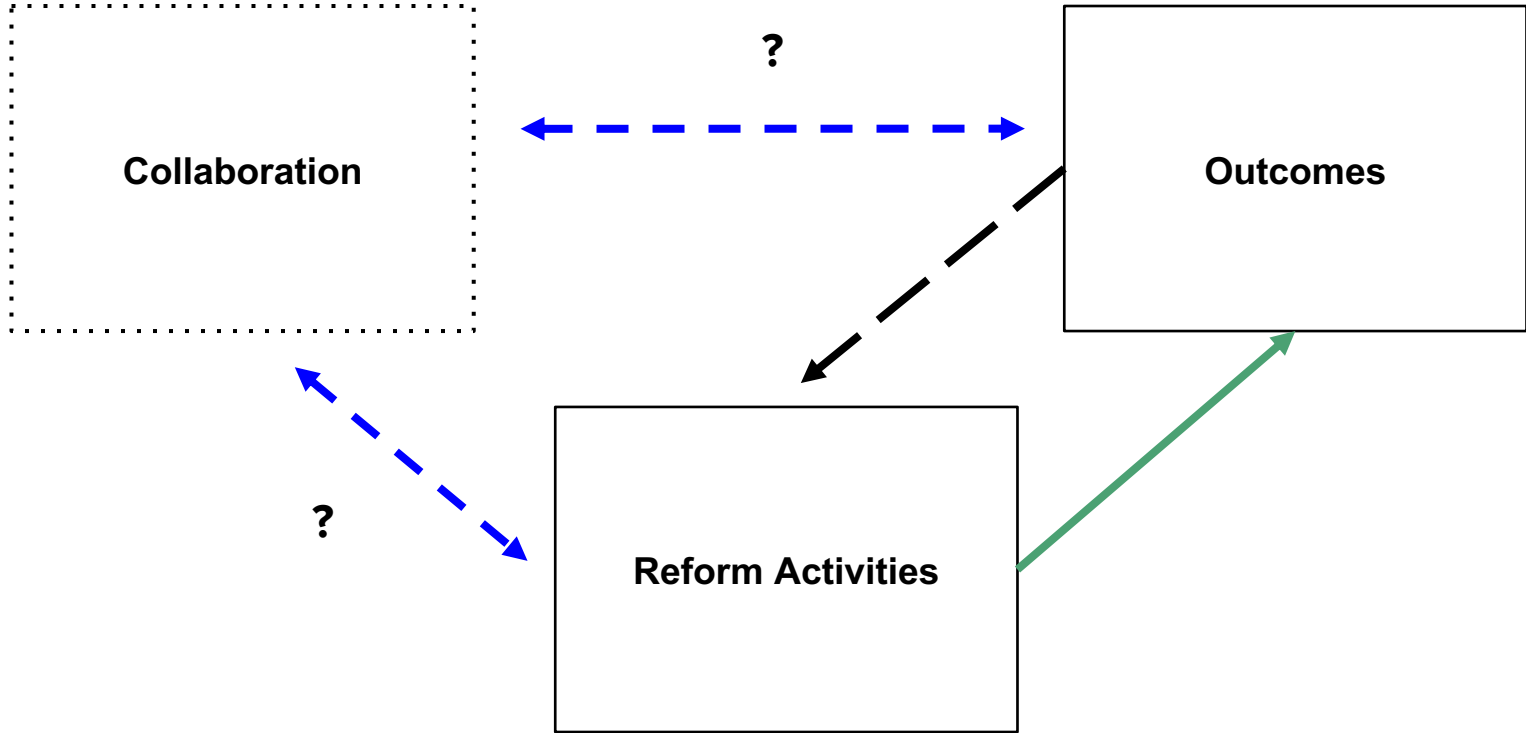
- Higher education reform increasingly relies on large-scale, multi-institutional collaborative initiatives
 - **HUGE** focus within STEM
- Projects are complex and “messy”
- Difficult to determine what impact a collaborative/collective approach has on desired outcomes
- “Classical” evaluation methods unable to address such complexity
- The **purpose** of this paper is to advance a **conceptual, literature-based evaluation framework** that demonstrates the interrelationship and importance of evaluating:
 - how a multi-organizational reform initiative functions (i.e., collaborative dynamics),
 - what the initiative does (i.e., change activities),
 - and collective outcomes.

“Classic” Program Evaluation

- Assumption that programs are stable and predictable (or even controllable...)
 - Did X cause A, B, C?
 - Focus on and measurement of summative effect
- Linear progression of outcomes/impact
 - Demonstrated through left-to-right logic models & often theories of change
- STEM education reform has utilized this “classic approach” for decades
 - Hire an evaluator (often summative/external), determine if project goals were accomplished
- Yet, “wicked” problems, like undergraduate STEM reform, are extremely **COMPLEX**

Move to a Systems Approach

- Movement towards collaborative reform initiatives to replace siloed/ineffective efforts
 - Underlying assumption: more change occurs through collaboration than working alone
- **However...**
 - “Even though evaluators may acknowledge the dynamic and interconnected nature of the complex systems that evaluated programs function within, many evaluators are unsure how to take these dynamics and interconnections into account and thus continue to use evaluation strategies and practices based on a static, mechanistic, and isolated view of the program.” (Moore, Parsons, & Jessup, 2019, p. 78).
 - The assumed change mechanism (collaboration) is often ignored or under resourced in its development/evaluation
- In response...A push towards developmental, formative, and systems-based evaluation
- To evaluate the success of collaborative change, it is critical to evaluate the system elements/linkages of collective reform:
 - Collaborative dynamics
 - (Aligned) activities
 - Collective outcomes

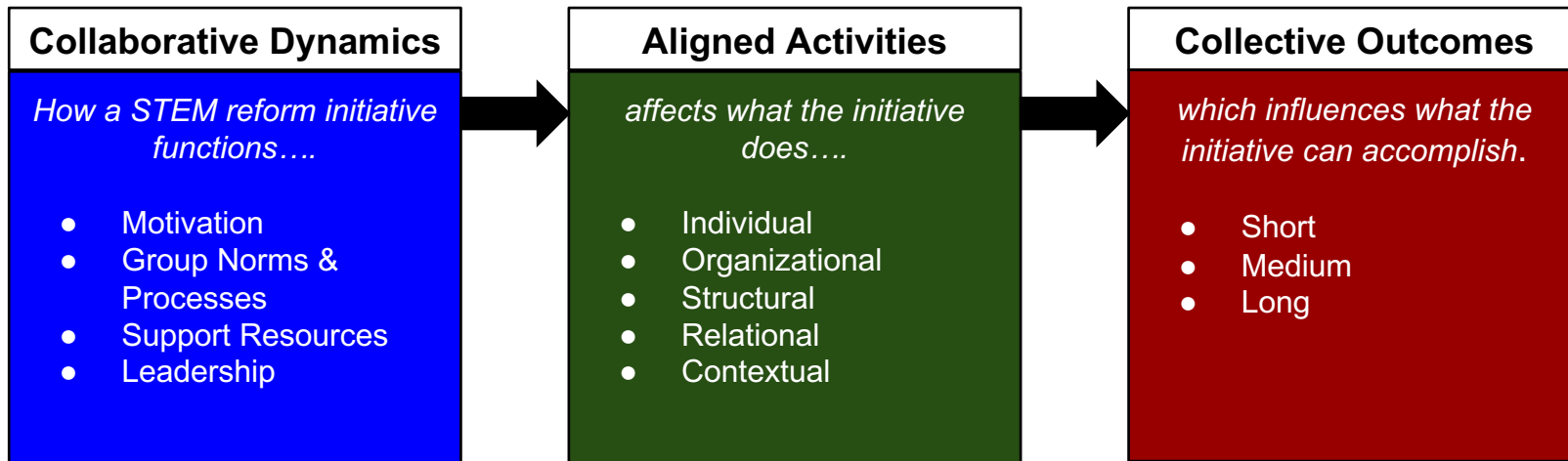


Tiered Evaluation Framework



This material is based on work supported by the National Science Foundation under Grant no. 1649199; Grant No. 1834518, 1834522, 1834510, 1834513, 1834526, 1834521; and Grant No. DUE 1231286. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Systems Context



Collaborative Dynamics

● Genesis

- Development of measures of collective impact in a NSF-funded project
- Literature review of collaboration in multi-sector and multi-institutional initiatives

● Key Parameters (Dimensions of Collaboration Framework)

- Motivation
- Group Norms and Processes
- Support Resources
- Leadership

Reform Activities

- **Genesis**

- Prior work investigating the impact of a large STEM Network (CIRTL)
- Hill et al., 2019; Hill, 2019

- **Key Elements**

- Individual
- Organizational
- Structural (and Logistical)
- Relational
- Contextual

Initiative Outcomes

● Genesis

- Long-time member of the research & evaluation team of CIRTL
- Applying the systems evaluation protocol (SEP) within CIRTL, “messy” pathway modeling

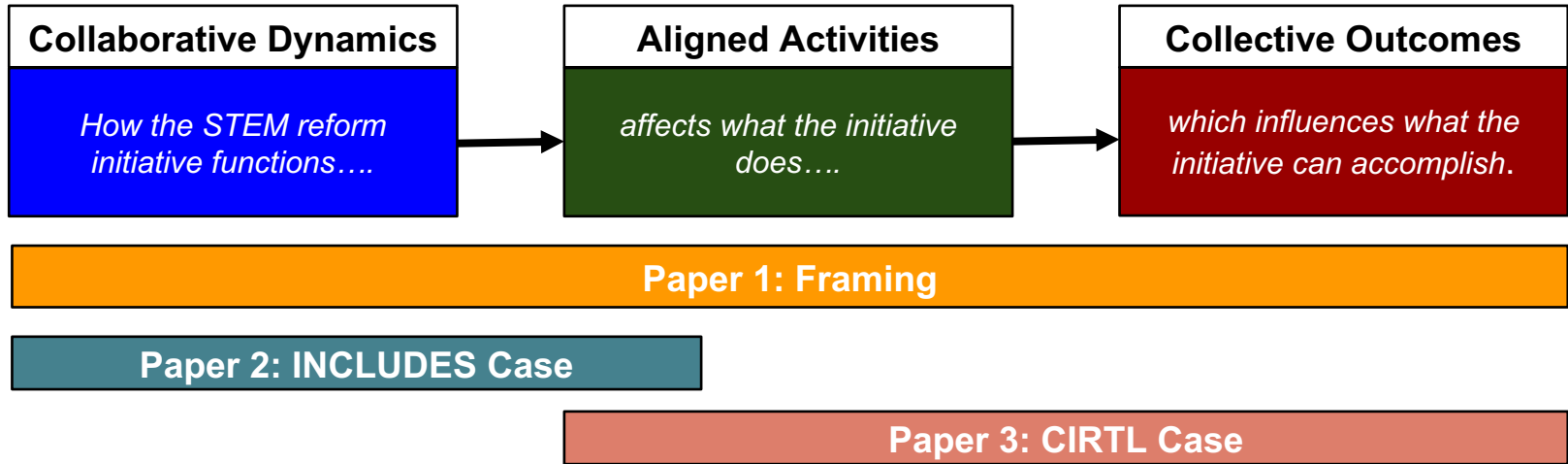
● Key Components

- Short
- Medium
- Long
- Relationships between outcomes
- Feedback loops to activities/collaboration

Implications

- Chance to slow down and examine the **holistic levels of evaluation**...not just programmatic impact
- Collaborative dynamics can **contextualize** and frame activities and outcomes
- Establish **collaborative effect** on reform initiatives
- **Diagnose** areas and ways to improve initiatives
- **Future Work**
 - Examine the evaluation framework elements within multiple reform initiatives
 - Explore statistical models to examine the interrelationships of the three elements

Transitioning to papers #2 and #3....



Evaluating Collective Impact in the CIRTL INCLUDES Pilot Project

Judy Milton, University of Georgia

Lucas B. Hill, University of Wisconsin-Madison

Julia N. Savoy, University of Wisconsin-Madison



This material is based on work supported by the National Science Foundation under Grant no. 1649199 and Grant No. 1834518, 1834522, 1834510, 1834513, 1834526, 1834521. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Collective Impact as a Model for Collaboration

Evaluation of Collaboration

- Intentional learning about CI processes
- Continuous learning within project through regular data feedback cycles
- Assessment of progress as well as outcomes

Evaluation as Collaboration

- Evaluators embedded within each project team
- Close coordination with teams to understand their data needs and build awareness of value of evaluation
- Co-creation of needs assessments and surveys to advance teams' work



Key Elements of Collective Impact (CI)

Collective Impact provides a framework for organizations to align and coordinate as a new system to address complex problems in a way that encourage emergence of innovative solutions and dissemination of effective practices. Key elements for the framework are:

- Common agenda
- Mutually reinforcing activities
- Shared measurements
- Continuous communication
- Backbone infrastructure



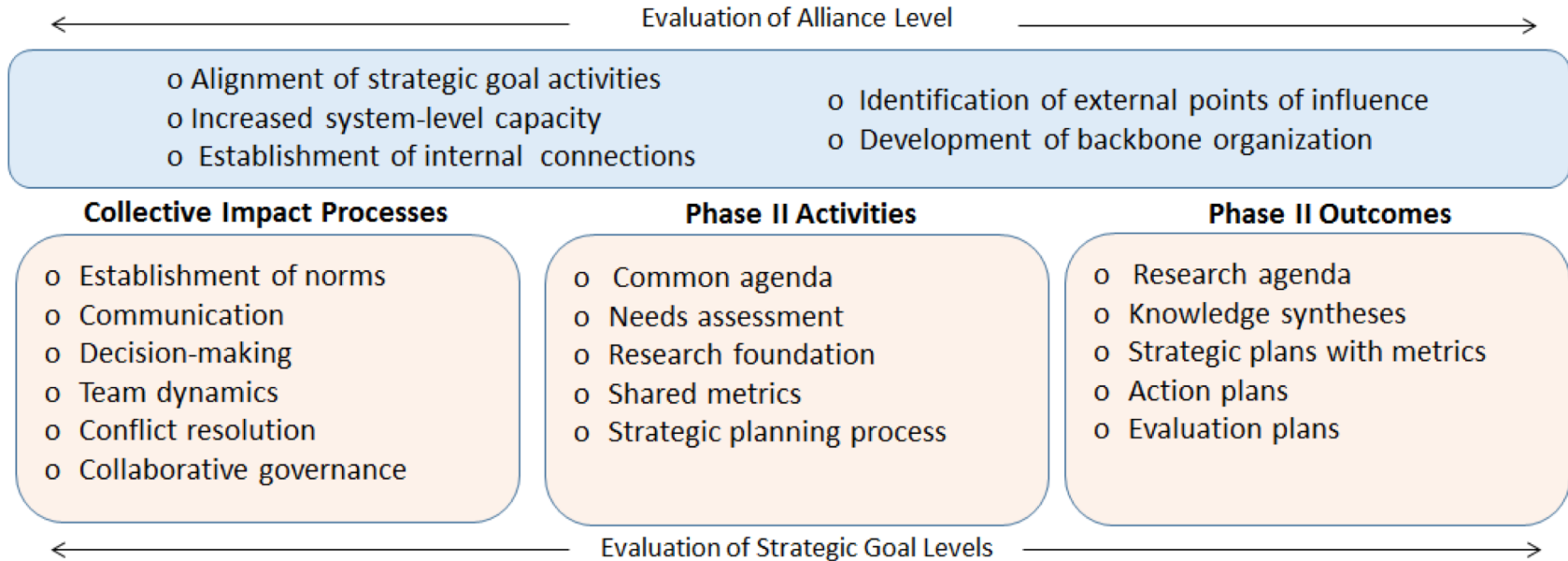
What do we want to learn from evaluation of CI?

Using a developmental approach, we have asked 3 primary questions to evaluate Collective Impact processes, activities, and outcomes:

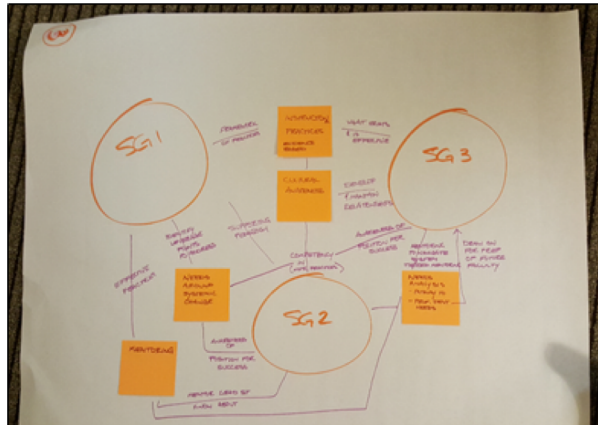
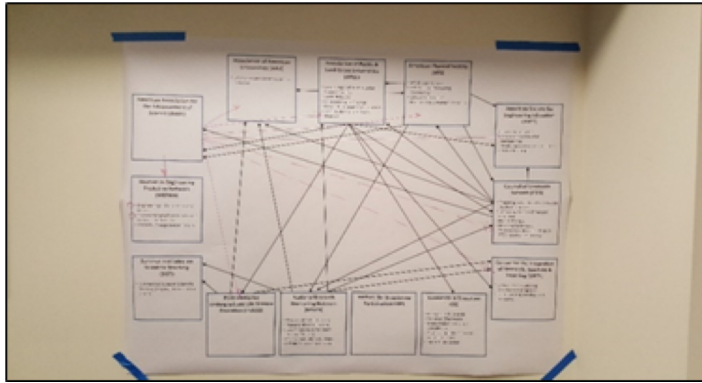
1. Did it occur?
2. How did it occur?
3. What was its quality?
 - a. Was it aligned with the common agenda?
 - b. To what extent were the project members engaged and committed to the work?
 - c. Did it mutually reinforce other activities in the project?
 - d. Was it evidence-based, actionable, and feasible?



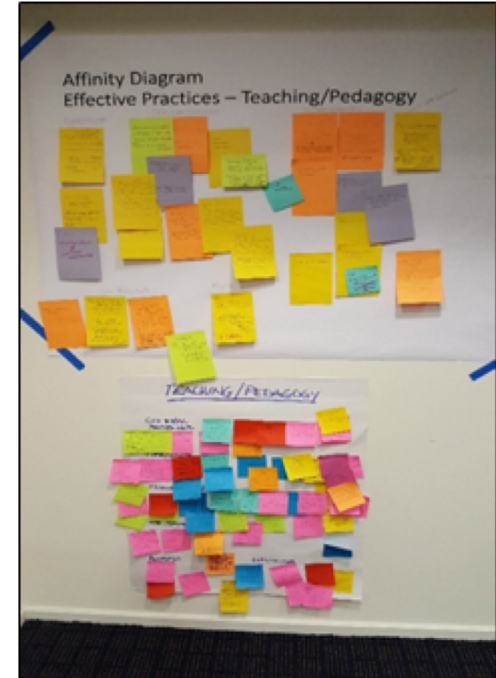
Our CI Evaluation Model in CIRTl INCLUDES Pilot



The Practice of Collective Impact in INCLUDES Collective Impact?



- Rotation Model - GS for CC courses
 - Job shadowing
 - Teach labs
 - Teach courses
 - "matchmaking"
- Fellowships (HawkeyeTech)
- Pedagogical Workshops
 - Current faculty with GS
 - use case studies/scenarios of CC life
 - CIRTL - improved, tailored
- Panels - career prep / expectations
- Videos / virtual Grand Rounds (Health Clinical)
- Educ. about bias - implicit biases, microaggressions, Inc. awareness, immigration
- Regional Internship - includes Aquinets in pedagogical - diversity journals
 - Scaffolded - dup a tree in 1/2
 - 2-3 hours in 1/2
 - 2-3 hours in 1/2
- Center for excellence - on demand - a LEAN? - 50% / 1/2



How do we measure Collective Impact?

Alliance-Level Items	Team-Level Items
The Aspire Alliance agrees on a common problem.	We agree on a common problem.
The Aspire Alliance has shared goals.	We have shared goals.
Aspire Alliance members communicate effectively with each other.	We communicate effectively with each other.
Aspire Alliance members trust each other.	We trust each other.
The Aspire Alliance provides ways for all voices to be heard.	We provide ways for all team member and partner voices to be heard.
The Aspire Alliance make decisions collectively.	We make decisions collectively.
Leadership in the Aspire Alliance is effective.	Leadership in [team] is effective.
Partners (i.e., individuals and organizations) essential to the success of the Aspire Alliance are involved.	We involve essential partners (i.e., individuals and organizations) to advance the work of the [team].
The Aspire Alliance has identified strategies to achieve Alliance-wide goals.	We have identified strategies to accomplish our goals.
Aspire Alliance teams coordinate with one another to implement project activities.	We know how the goals and activities of our team fit within the Aspire Alliance as a whole.
The Aspire Alliance has identified metrics to measure our success.	We have identified metrics to measure our success.
The Aspire Alliance has a clear plan to collect and use data to drive strategic improvement.	We have a clear plan to collect and use data to drive strategic improvement.

Alignment of Collective Impact Metrics

Our measurement of CI in the NSF INCLUDES Alliance lies within multiple frameworks: Collective Impact, the CIRTl INCLUDES pilot, the NSF Collaborative Framework, and the theoretical Collaborative Dynamics framework

Alliance-Level Items	Team-Level Items	Framework Constructs
The Aspire Alliance agrees on a common problem.	We agree on a common problem.	<u>CI</u> : Common agenda <u>CIRTl CI</u> : Alignment <u>NSF</u> : Vision <u>CD</u> : Congruity
The Aspire Alliance has shared goals.	We have shared goals.	<u>CI</u> : Common agenda <u>CIRTl CI</u> : Alignment <u>NSF</u> : Vision <u>CD</u> : Congruity
Aspire Alliance members communicate effectively with each other.	We communicate effectively with each other.	<u>CI</u> : Continuous communication <u>CIRTl CI</u> : Communication <u>NSF</u> : Leadership & communication <u>CD</u> : Interaction



What have we learned about the implementation of CI?

- Decision to use CI needs to be a collective decision
- Implementation takes time, commitment, and resources
- “Shared leadership” needs to be defined and practiced proactively
- Practice of CI sits at the intersection of multiple organizational cultures
- Equity and inclusion should be explicitly stated in order to confront systems of oppression
- Implementation is dynamic and iterative



What are our challenges?

- How should the project balance the work to produce program outcomes with the work to develop collaborative processes?
- How does the project develop and practice inclusive norms, including ones to address dynamics, positionality, and voice?
- How should equity and inclusion be defined by the project? How can these constructs be explicitly measured within the CI framework?
- How can we best use the organizational research on institutional mergers, large-scale collaborations, and shaping of new organizational cultures to enhance this project and inform other funded efforts to bring change initiative to a national scale



Evaluation Capacity Building Through Stakeholder Pathway Modeling of Local Programs in a Multi-Institutional STEM Future Faculty Development Initiative

American Evaluation Association, 2019

Mark Graham & Melanie Bauer
Yale University

Lucas Hill,
University of Wisconsin-Madison



This material is based on work supported by the National Science Foundation under Grant no. 1726625. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

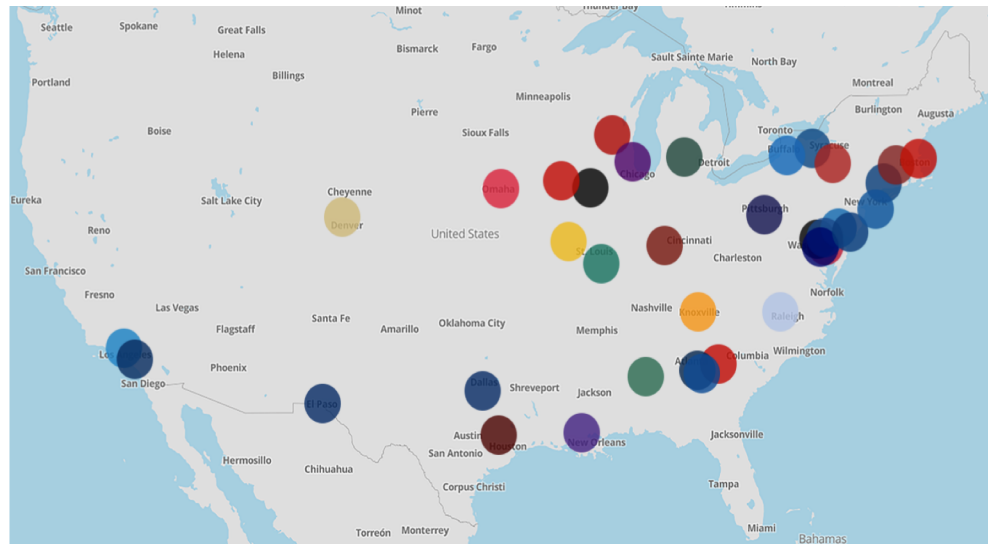
The CIRTLL Network of 38 Institutions

Center for the Integration of Research, Teaching, & Learning (CIRTLL)

University of Wisconsin - Madison

Founded in 2003, now 38 member institutions across the U.S. (NSF funding)

Impacts >5,000 “future faculty” participants annually through >100 Professional Development (PD) activities (e.g., workshops, courses, events) with goal to improve undergraduate STEM education.



Infrastructure. CIRTLL Central at the University of Wisconsin. Each of the 38 institutions has a “local” institutional/campus leader. In-person network meeting every 6 months. Conference call once a month. Active leadership team.

State of Evaluation for the CIRTL Network

Evaluation of CIRTL needs to move towards assessing network-wide impacts

Currently:

- Collect annual reports from institutions (# and types of activities, # of participants, levels of participation)
- Most institutions conducting immediate, post-program surveys
- Some individual institutions studying their program impacts over the longer-term

The CIRTL Network desires a set of common measures to examine broader (cross-institutional) Professional Development training impacts over the short-, medium-, and long-term.

Network-wide information would be very useful to local programs as well, but they are limited in their capacity (time, resources, expertise) to collect this data

Common Measures Across a Network Initiative?

Potential First Steps towards the Solution

The CIRTl Network: a case study for using systems-based evaluation across a program network

Work at institution level toward creating network-level evaluation infrastructure

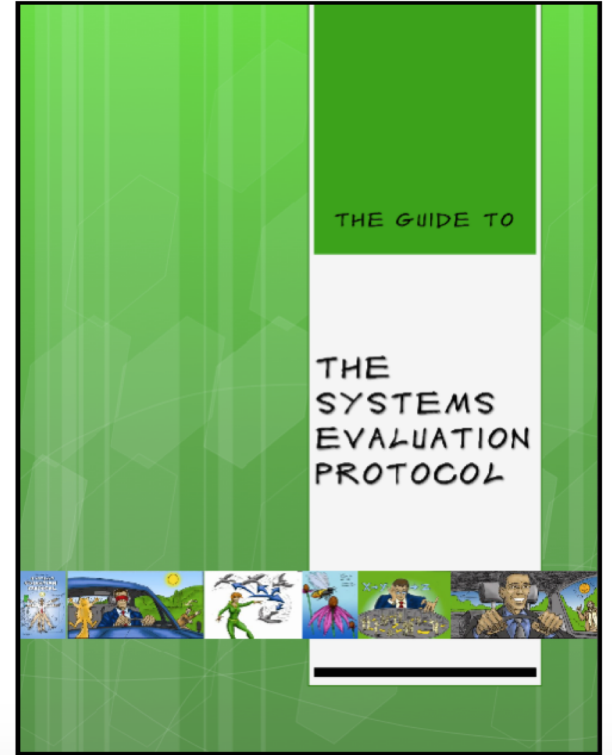
Evaluation capacity building intervention with two cohorts of local CIRTl leaders – a ground up, grass roots approach to common measures

Systems Evaluation Protocol

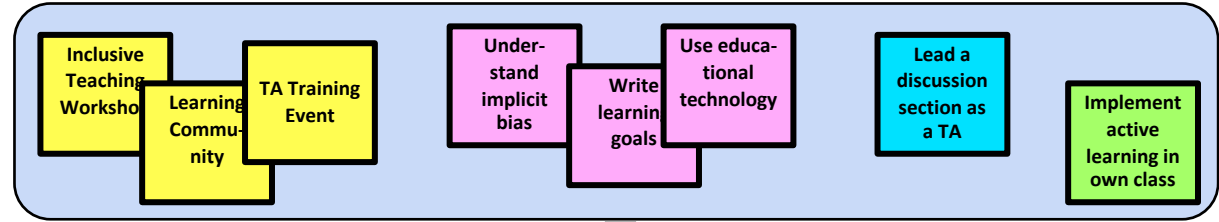
a.k.a. “the SEP”



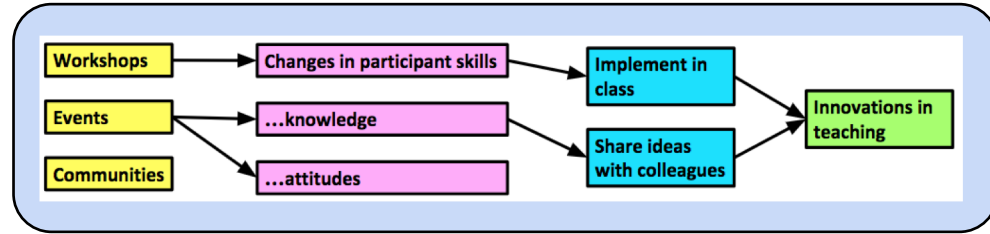
Cornell University
Cornell Office for Research on Evaluation



Outcomes Analysis



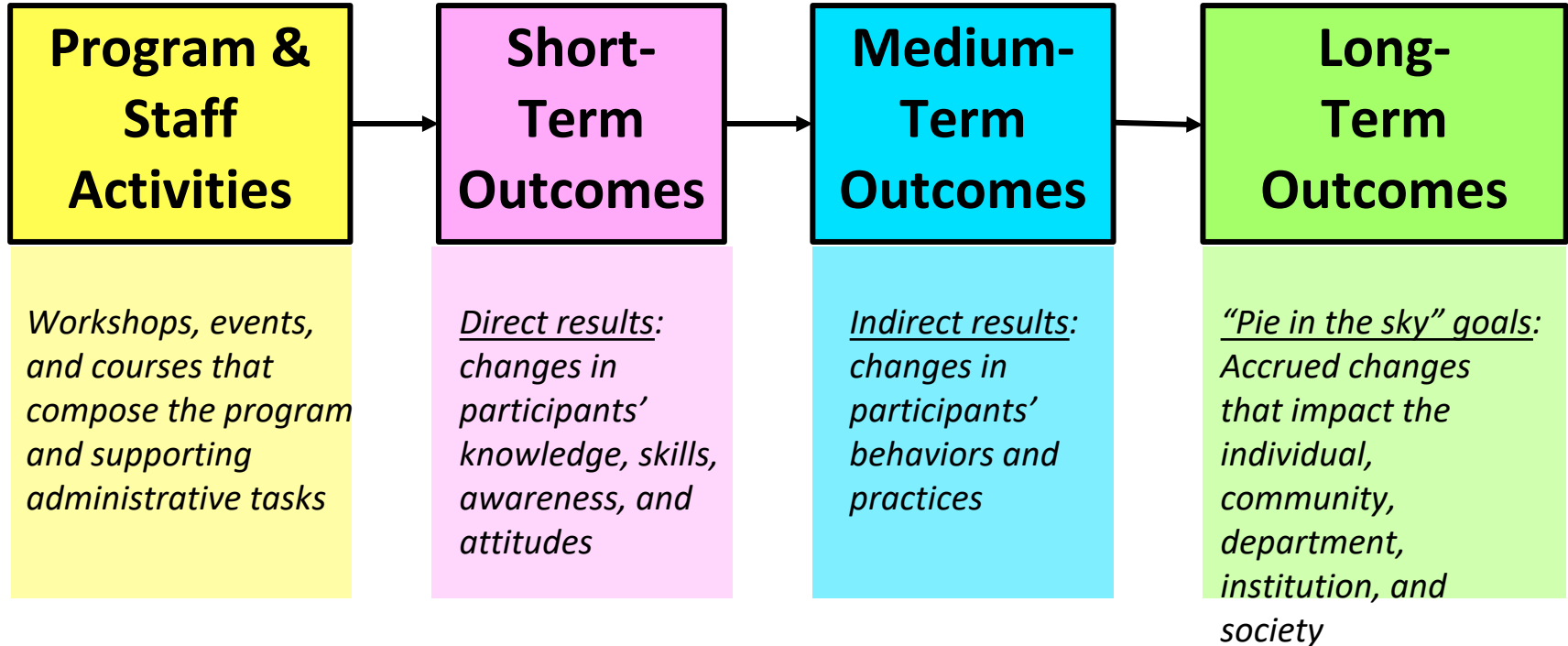
Pathway Model (“Map”)



Evaluation Plan

Evaluation Question	
Measurement	
Sample/Participants	
Design/Procedures (e.g., pre/post)	
Data Management	
Analysis	
Reporting/Utilization	
Timeline	

CIRTL Program Mapping with the SEP



Our Questions

What are the **expectations and experiences** of institution-level CIRTl program leaders as they participate in our SEP-driven evaluation capacity building intervention? **Like it? Not like it? Helpful? Useful?**

Are we improving evaluation knowledge, skills, and attitudes?

What are the **expected program outcomes** (spanning the short-, medium-, and long-term) shared across their institutions? **Commonalities or unique aspects (or both)?**

Are local leaders better prepared to develop plans to measure mid- to longer-term outcomes?

Over two years we worked as evaluation consultants with institution-level program leaders across two cohorts of CIRTl institutions ($N=10$):

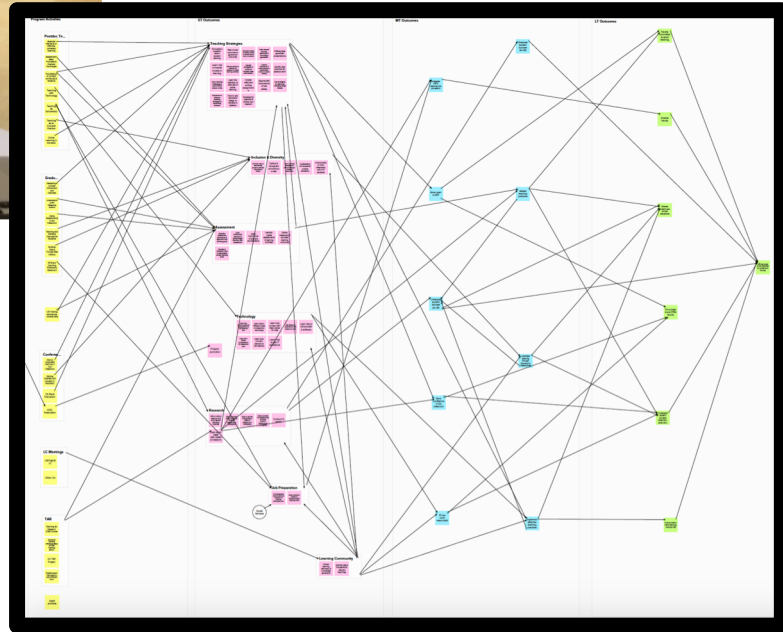
Cohort 1

- Cornell University
- Johns Hopkins University
- Northwestern University
- University at Buffalo
- Yale University

Cohort 2

- Drexel University
- Iowa State
- University of Maryland
- University of North Carolina
- University of Texas-Arlington

At each institution additional participants invited by program leaders -- local CIRTl program stakeholders (e.g., evaluators, center directors, program staff)



- Each local program mapped their activities and expected short-, medium-, and long-term program outcomes (under guidance of a systems evaluation consultant) using simplified version of the SEP

[additions to the SEP]

- Because of cohort model, intentionally created evaluation-focused “communities of practice” that meet in-person biannually at network meetings to exchange/collaborate
- Using a subset of Cohort 1 program leaders to assist in co-facilitating the evaluation intervention with Cohort 2 (early success!)

Leader experiences

Formative feedback reported in stakeholder surveys/interviews at each stage:

1. after program mapping
2. after evaluation planning
3. after sharing maps/plans with other program leaders at a network meeting

Grounded theory and general analytical process used by Auerbach et al. (2018)

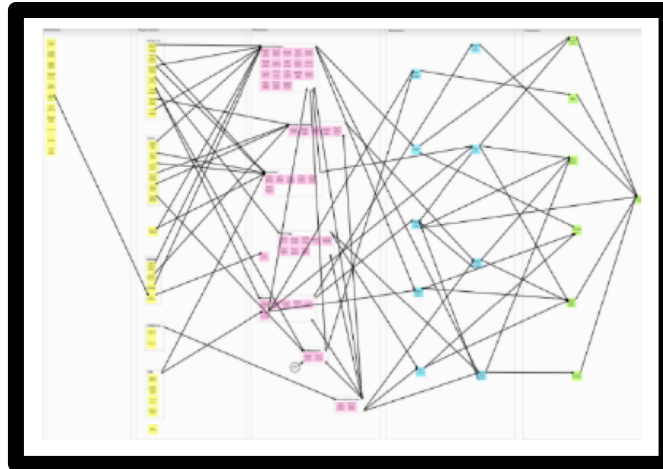
CIRTL Leader experiences

Theme	Freq.	Example Quote
Program	74	<i>“To identify gaps in our programs”</i>
Evaluation	41	<i>“Learning more about program assessment”</i>
Support/Affect	31	<i>“More confident knowing I can reach out to someone”</i>
Sharing	17	<i>“Create some common language across CIRTL leaders”</i>
General Work	15	<i>“Good for collaboration and productivity”</i>
CIRTL	4	<i>“Rectify CIRTL core goals with what current projects doing”</i>
Feedback	4	<i>“Only downside was the overall time commitment”</i>

Program outcomes

Analyzed program maps to identify “common outcomes” across maps ($n=6$)

Coded for CIRTL and non-CIRTL (emergent) themes



Program activities



Short-term outcomes



Medium-term outcomes



Long-term outcomes

Program outcomes

Most common outcome codes across institutions' program maps

*Code	Map 1	Map 2	Map 3	Map 4	Map 5	Map 6	Total
Career	33	13	14	4	22	4	90
EBT-1	14	17	16	10	13	5	75
LC	8	7	26	2	22	2	67
EBT-2	7	13	6	8	9	0	43
Culture	5	0	12	4	5	9	35

EBT-1 & EBT-2 relate to evidence-based teaching; LC relates to learning communities

*Example Outcome Codes

Example <i>CIRTL</i> Codes	Example <i>Non-CIRTL</i> Codes
Describe several <u>assessment</u> techniques and recognize their alignment with particular types of learning goals.	(<u>Career</u>) Improving job prospects, benefits for tenure and promotion.
Describe the scope of <u>diversity</u> in learning environments, of both students and instructors.	(<u>Attitudes</u>) Positive attitude and confidence related to teaching as a profession and/or teaching identity.
Describe several techniques and issues of establishing <u>learning communities</u> comprising a diverse group of learners.	(<u>Discipline</u>) Learn discipline-specific knowledge related to teaching & learning.

Benefits to programs and the local leaders

- Mapping products (program map, evaluation plans); can “see” their program
- Accountability partner (i.e., project evaluators checking in)
- Bringing key local people together to collaborate
- Creating a collaborative learning community of CIRTl leaders (different program structures, post-meeting collaboration around evaluation between two institutions)
- A way to advance collective impact - i.e., compare maps with other campus groups to see how their programs fit in the larger whole and other institutional priorities and goals

Benefits to program network

- Local leaders (10 of 38) supported/trained in evaluation and a common language for evaluation (evaluation capacity building in the network)
- Evaluation-focused community of practice formed (and growing); in part due to existing CIRTL infrastructure
- For a subset of CIRTL programs, an identification of common outcomes that can support network-wide evaluation
- The potential for ramp up effect where many campuses trained in systems evaluation, which then has a means to identify common metrics
- A test bed to figure out complicated metrics for a complicated collaborative reform initiative
- National ripple effect and potential large-scale effect on STEM reform; CIRTL members are often part of much more than CIRTL

Challenges and risks

- Time and energy required by each program leader to participate.
- How to incorporate multiple program stakeholders' perspectives into one program map.
- Multi-institutional evaluation projects: keeping the ball rolling on a project when the majority of work is facilitated remotely (i.e., online).

Lessons Learned

Reduce the mapping for large programs (i.e., map only a subset of their programs)

Go online first--have local leaders create their own program maps in the online program Mural (vs. paper and pen later digitized by evaluator)

Frequent check-ins move the project along (especially in evaluation planning and execution phases)

Semi-structured, informal cohort meetings allow for community building and cross-institution sharing (vs. rigid, highly structured set of activities)

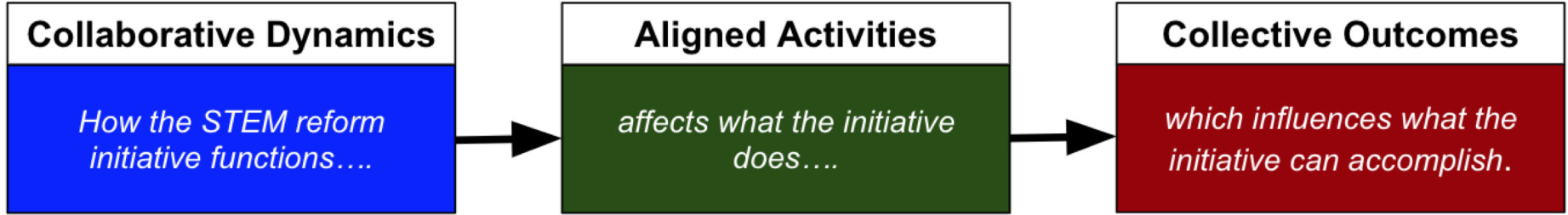
May be useful to have cohorts matched for a similar stage of program development (though pros/cons of each format)

Conclusions

Institution-level program leaders benefiting both in *program development* and *evaluation planning* from this evaluation intervention, also appreciate the *support and guidance* from evaluation consultants

Several activity outcomes are candidates for “*common measure*” instrument development (at least across the first 5 CIRTl institutions), but also significant institutional variation

Collaborative Dynamics Framework



- Institution-level program leaders invite additional stakeholders to engage in the evaluation process
- Cross-institutional sharing and collaboration

- Map of program activities
- Share activity maps with other program leaders

- Map of short-, medium-, and long-term outcomes
- Share outcomes and evaluation plans with other program leaders

mark.graham@yale.edu