Center for Assessment

RILS 2021: Design Innovation in Educational Assessment Systems

The National Center for the Improvement of Educational Assessment

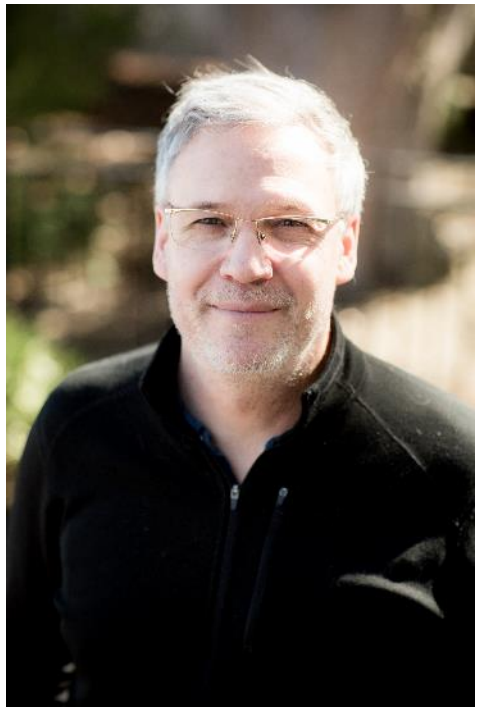
September 2021

Welcome to RILS 2021

Purpose of RILS 2021:

Discuss the different aspects of the design innovation process for assessment systems including:

- exploring the process for designing an innovative assessment system,
- providing an overview of current innovations in assessment systems, and
- delving into the attributes necessary for the design of an innovative assessment system.



RILS Sessions

Session	Date	Presenter(s)
Session 2: Understanding the Problems for Design Innovation – *Root Cause Analysis*	Monday, September 20, 2021 1:00-2:30	Juan D'Brot and Chris Brandt Guests
Session 3: Leveraging Community for Design Innovation - *Engaging Stakeholders*	Monday, September 20, 2021 3:00-4:30	Carla Evans Guests
Session 4: Planning for Design Innovation – *Assessment Systems and Theory of Action*	Thursday, September 23, 2021 1:00-2:30	Erika Landl Nathan Dadey
Session 5: Exploring the Design Innovation Process – *Iteration in Assessment System Design*	Thursday, September 23, 2021 3:00-4:30	Brian Gong Guests
Session 6: Exploring the IADA Innovation Process –*Challenges and Opportunities*	Friday, September 24, 2021 1:00-2:30	Scott Marion and Carla Evans Chris Domaleski

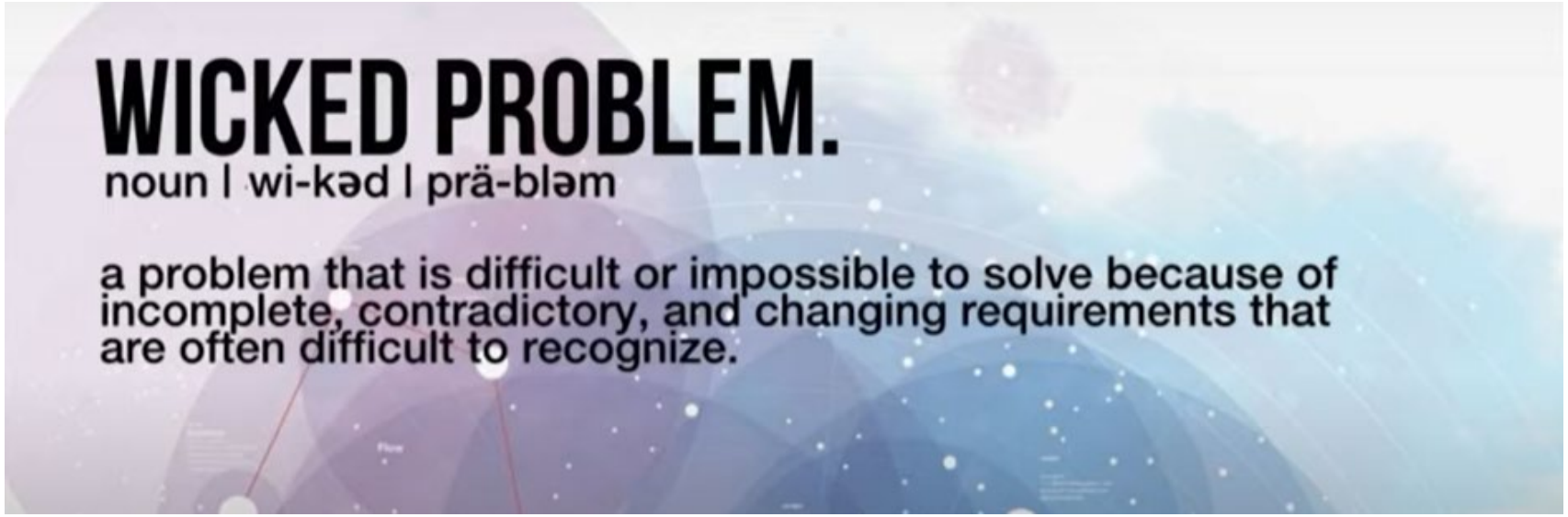
Design Innovation – What is it?

Ideology

+

Process

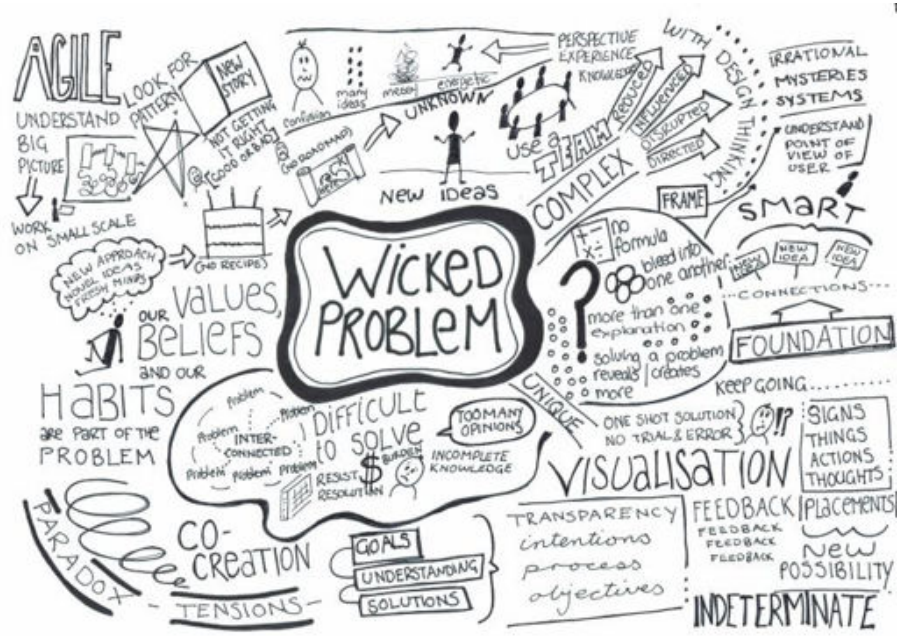
To solve "wicked" problems in a user-centric way



Assessment Systems as a *Wicked Problem*

How to appropriately assess students?

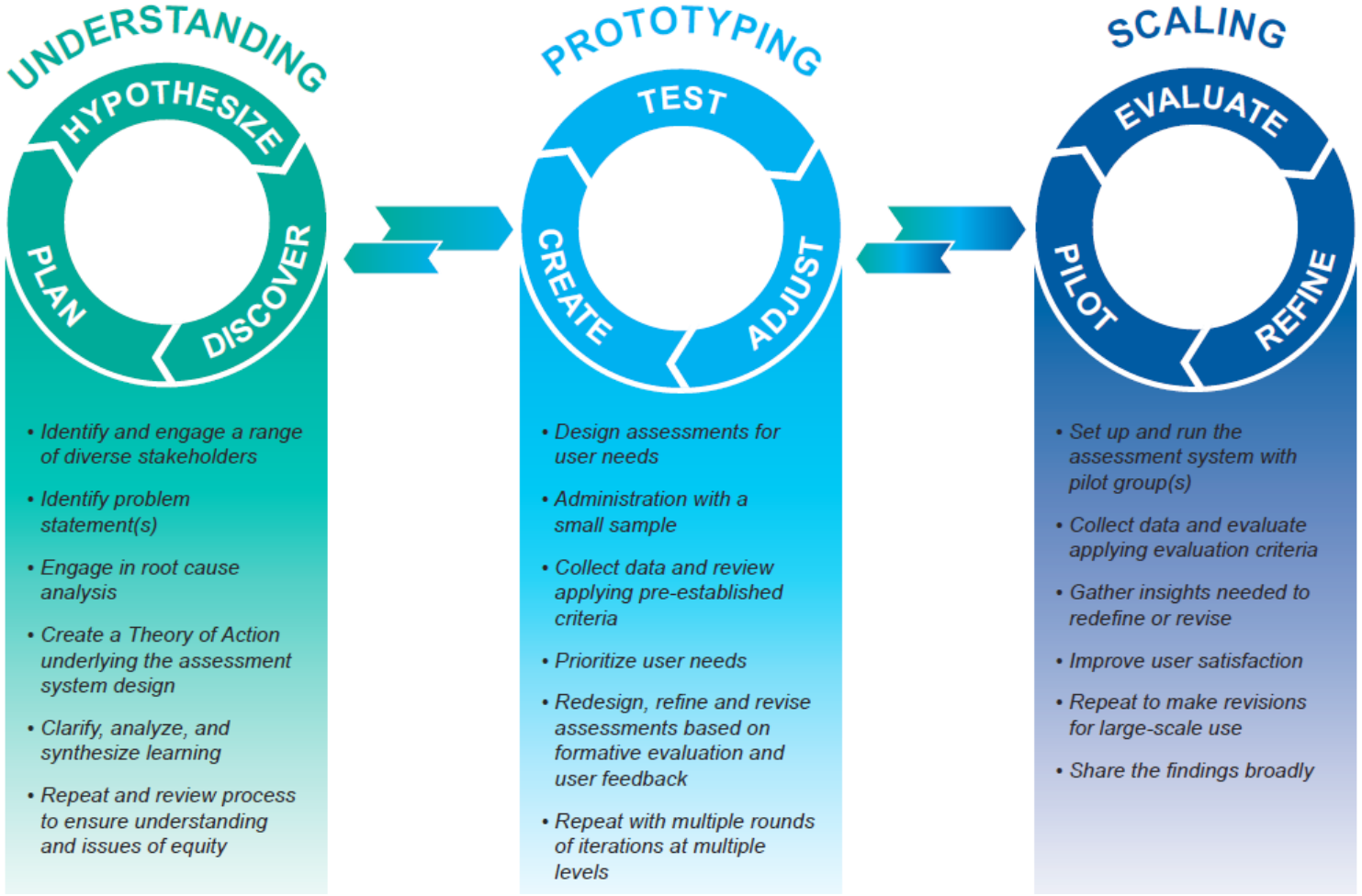
- *What does this mean?*
- *What is the end result*
- *Will it work for all students?*
- *Will it work in all content areas?*
- *Will it work in all environments?*
- *How does the solution of one aspect impact the next situation?*



One Model for the

DESIGN INNOVATION PROCESS

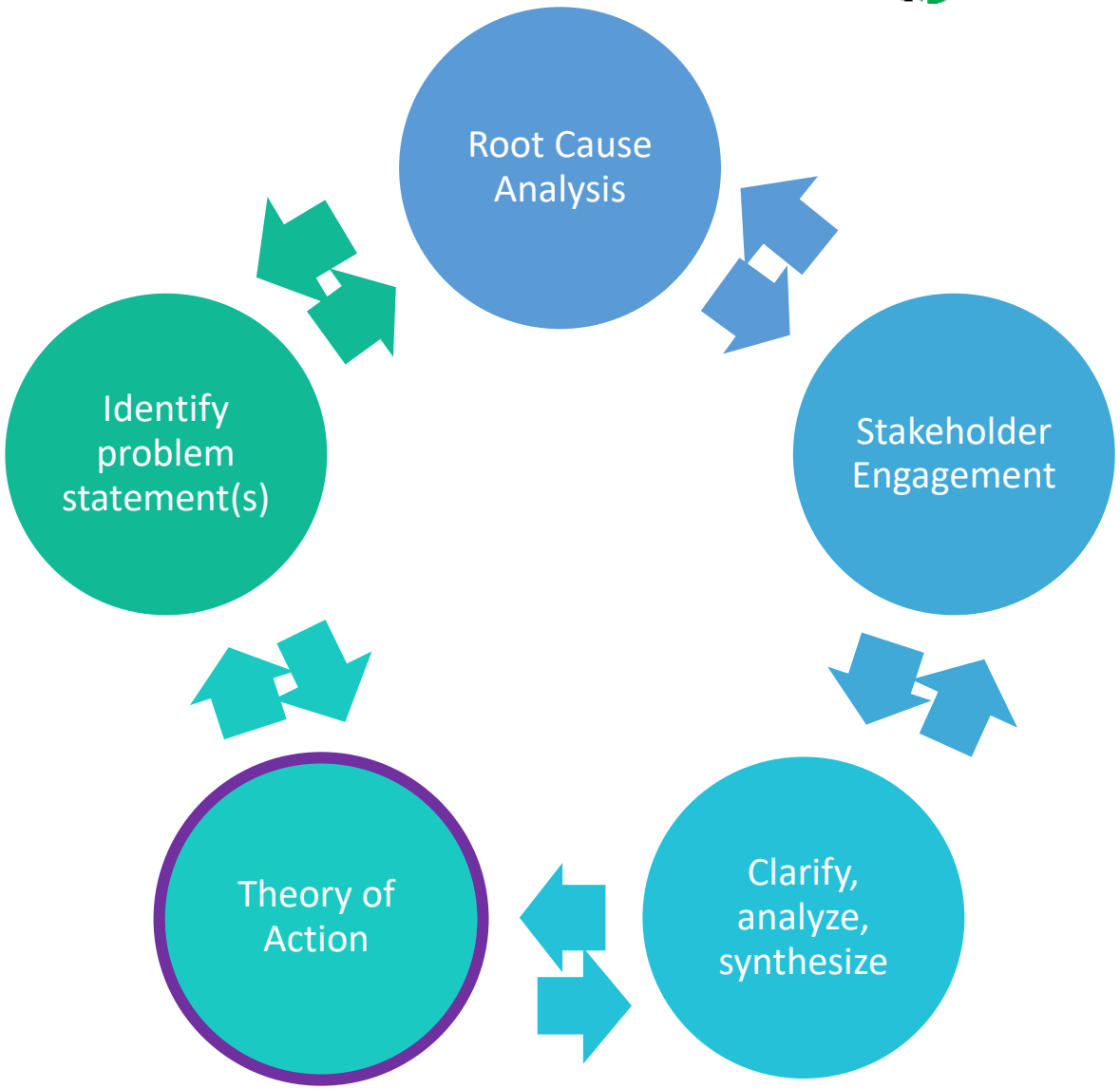
Center for Assessment



"Understanding" Stage

- *Identify and engage a range of diverse stakeholders*
- *Identify problem statement(s)*
- *Engage in root cause analysis*
- *Create a Theory of Action underlying the assessment system design*
- *Clarify, analyze, and synthesize learning*
- *Repeat and review process to ensure understanding and issues of equity*

Hypothesize
Discover
Plan



Session 4 Focus

"Design adds value faster than it adds costs." -- Joel Spolsky, web programmer, writer, and creator of Trello

Planning for Design

- Defining Theory of Action
- Characteristics of a Balanced Assessment System
- Questions within a Theory of Action for a Balanced Assessment System
- Using the Template
- Q&A / Closure

Planning for Design – *Theory of Action for a Balanced Assessment System*

Erika Landl

Nathan Dadey

Center for Assessment

Leveraging Theories of Action to Support Design Innovation in Assessment Systems

Nathan Dadey & Erika Landl, with contributions from Brian Gong

The National Center for the Improvement of Educational Assessment

1

Theory of Action

A brief review of theory of action within the context of assessment systems

2

Dimensions of Innovation

Defining kinds of innovation within a balanced assessment system

3

Example Theories of Action

Worked, hypothetical examples of assessment systems with innovative components

4

Q&A

Question and answer session facilitated by Scott Marion

1. Theory of Action

Would program logic by any other name smell as sweet?

-Shakespeare, probably

Design innovation is complex.

Managing this complexity is critical.

A theory of action is *the* approach to help manage this complexity.

A Theory of Action is

A **logical argument** that connects the **goals** of a system to its **component parts**

By describing the **actions and conditions** that lead to the goals

as well as the **rationales**, **assumptions** and **evidence** that support and justify the connections within the system

For a deep dive in the context of general programs, see Patton (2008), particularly chapter 10. For consideration of theory of action in the context of assessment, see the Center's Theory of Action Template.

The Importance of a Theory of Action

A theory of action makes **design of a system explicit** and in doing so:

Acts as a roadmap for design and a touchstone for iterative design

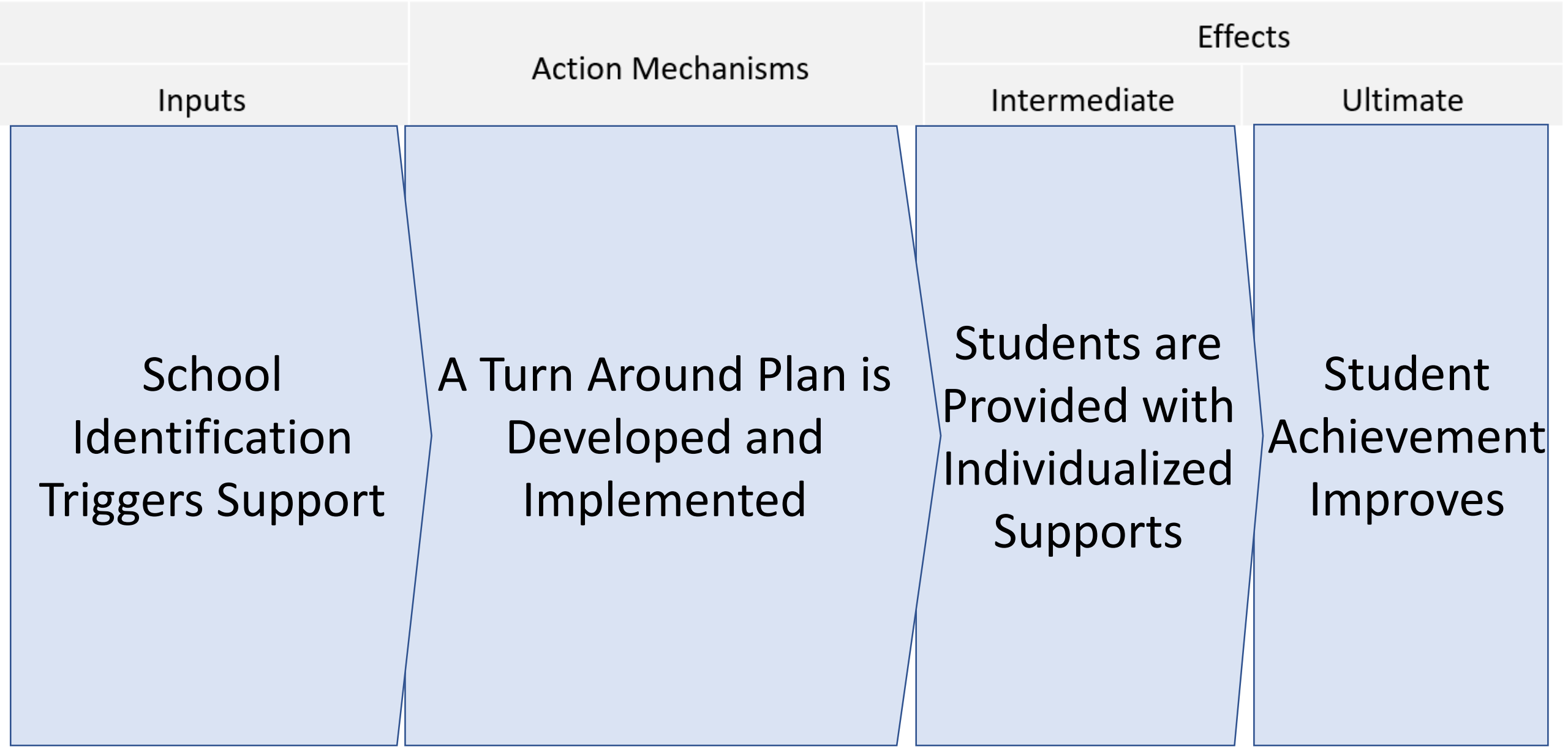
Provides shared language & understanding

Supports the investigation of problem areas

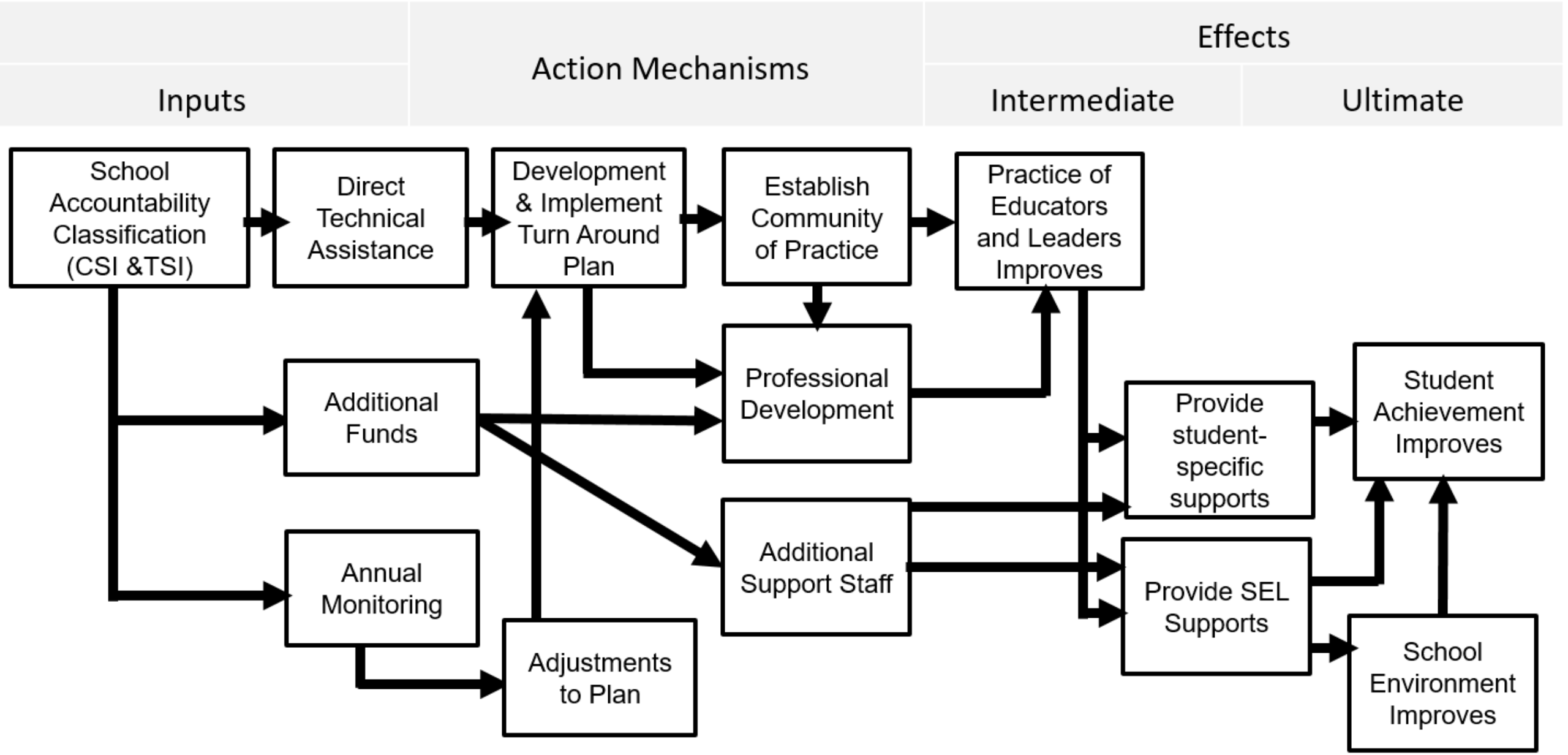
A Simple Logic Model Framing

- There are a number of ways to visualize and structure a theory of action, including Logic Models (e.g., Frechtling, 2007, W.K. Kellog Foundation, 1998) or Driver Diagrams (e.g., Bennett & Provost, 2015).
- We present theories of action graphically. Others use formats like tables (e.g., SCILLSS, 2017, p. 5). There is no *one* correct format and each application is tailored by the developer.

A Simplified, Hypothetical State Accountability Plan



One Level to Detail Down



Implications for Innovation Design?

Inputs	Action Mechanisms	Effects	
		Intermediate	Ultimate

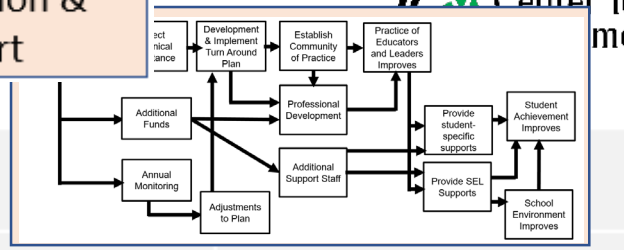
So what does this mean for design innovation?

Implications

- There are **multiple levels of complexity**
- Innovation design requires us to both
  - **Drill down** into specific parts of the theory of action, and
  - **Attend to the whole.**
- While also being clear about **what** is being innovated on:
  - Assessments, whether in part of whole, and
  - The use of that assessment information[1].

[1]This focus on linking assessments explicitly to the program they function within draws from the work that ties score use to interpretation (e.g., Cronbach's work throughout the 70's and 80's; Kane, 2006, p. 53; and more recently, Bennett, Kane & Bridgeman, 2011).

ESSA School Identification & Support



Center for ment

This theory of action is just one of many that are operating in an educational context.

Intermediate

Ultimate

School Accountability Classification (CSI &TSI)

Direct Technical Assistance

Development & Implement Turn Around Plan

Establish Community of Practice

Practice of Educators and Leaders Improves

Additional Funds

Professional Development

Provide student-specific supports

Student Achievement Improves

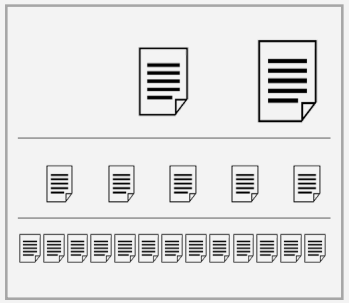
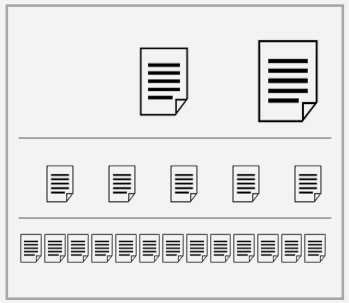
Annual Monitoring

Additional Support Staff

Provide SEL Supports

Adjustments to Plan

School Environment Improves

Levels		Uses
State.	Statewide Accountability Assessment 	ESSA School Identification & Support
District.	District-Wide Middle of Year Interim (e.g., Benchmark) 	District Resource Allocation
Classroom.	End of Unit & Mid-Unit Check in Assessments, Weekly Exit Tickets, Daily Conversations  Instructional Unit . . .  Instructional Unit	Formative Assessment Cycle for Tailored Instruction

Quarter 1	Quarter 2	Quarter 3	Quarter 4

Levels	Uses
State. Statewide Accountability Assessment 	ESSA School Identification & Support
District. District-Wide Middle of Year Interim (e.g., Benchmark)	District Resource Allocation
Classroom. End of Unit & Mid-Unit Check in Assessments, Teacher Created Daily Formative	Formative Assessment Cycle for Tailored Instruction
Instructional Unit	Instructional Unit

What are we trying to innovative?

Quarter 1	Quarter 2	Quarter 3	Quarter 4

Levels

Uses

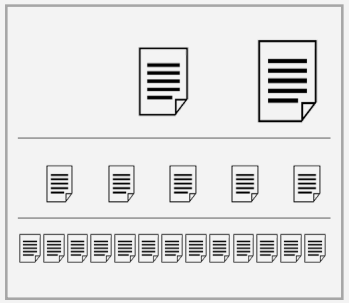
State. Statewide Accountability Assessment

ESSA School Identification & Support

District. District-Wide Middle of Year Interim (e.g.

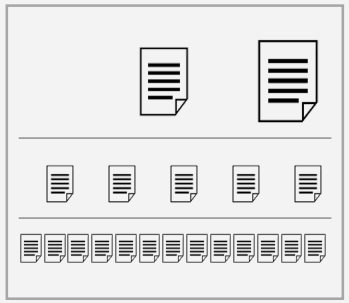
Is it, for example, an aspect of the state test a small grain-size?

Classroom. End of Unit & Mid-Unit Check in Assessments, Weekly Exit Tickets, Daily Conversations



Instructional Unit

. . .



Instructional Unit

Formative Assessment Cycle for Tailored Instruction

Quarter 1	Quarter 2	Quarter 3	Quarter 4

Levels

Uses

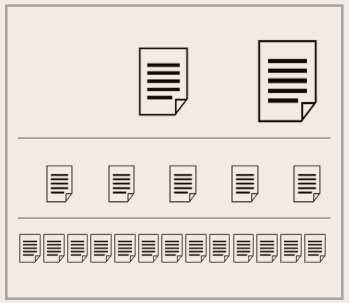
State. Statewide Accountability Assessment

ESSA School Identification & Support

District. District-Wide Middle of Year Interim (e.g.

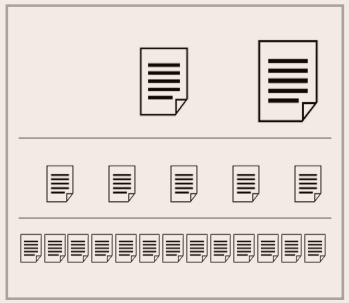
Or is it some much larger aspect of the state assessment that has implications both for the state level use, but also implications for other levels.

Classroom. End of Unit & Mid-Unit Check-In Assessments, Weekly Exit Tickets, Daily Conversations



Instructional Unit

. . .



Instructional Unit

Formative Assessment Cycle for Tailored Instruction

Quarter 1	Quarter 2	Quarter 3	Quarter 4

Levels	Uses
State. Statewide Accountability Assessment	ESSA School Identification & Support
District. District-Wide Middle of Year Interim (e.g.	Allocation
Classroom. End of Unit & Mid-Unit Check in Assessments, Weekly Exit Tickets, Daily Conversations	Formative Assessment Cycle for Tailored Instruction

Or is the innovation crossing all levels?

Instructional Unit

. . .

Instructional Unit

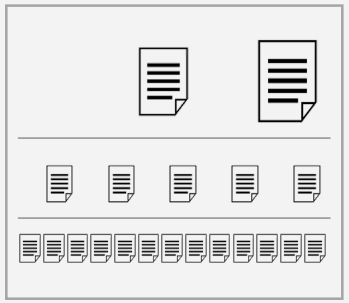
Quarter 1	Quarter 2	Quarter 3	Quarter 4

Levels

State. Statewide Accountability Assessment

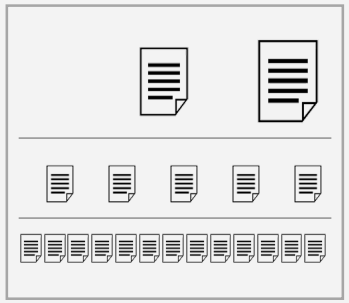
District. District-Wide Middle of Year Interim (e.g., Benchmark)

Classroom. End of Unit & Mid-Unit Check in Assessments, Weekly Exit Tickets, Daily Conversations



Instructional Unit

. . .



Instructional Unit

Quarter 1	Quarter 2	Quarter 3	Quarter 4

Uses

ESSA School Identification & Support

Ultimately, our theory of action can and should acknowledge other levels.

Instruction

Theory of Action in terms of Balanced Assessment Systems

- This framing is rooted in a **balanced assessment system** approach.
- This framing helps shift our focus to include **the information provided to the assessments** and the **use of that information** within the theory of action (see the Appendix for more detail).

Knowing What Students Know (2001), crystallized the appeal for balanced systems of assessment (see p. 253 – 257). Recent work has shown that building these systems faces a number of barriers (e.g., Marion et al., 2019a; Marion et al., 2019b; Conley, 2018). These challenges are also considered in a 2018 Special Issue of EM:IP.

2. Defining Dimensions of Innovation.

Multiplicative Roles of TOA in Design Innovation

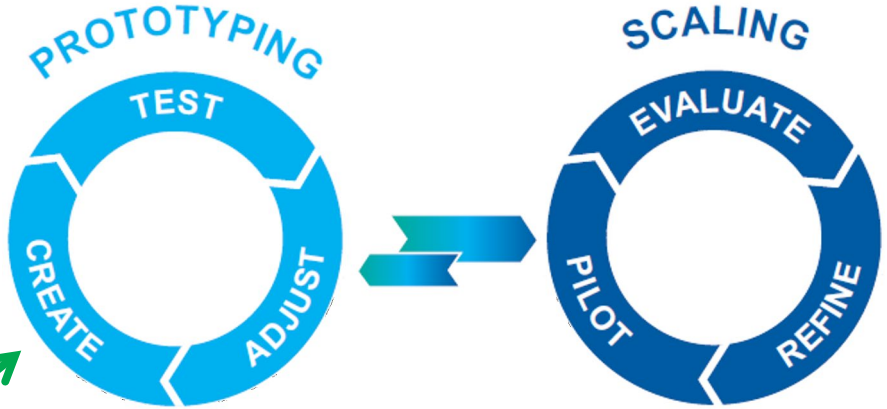
Center for Assessment

Root Cause Analysis

Identify problem statement(s)

Stakeholder Engagement

Theory of Action



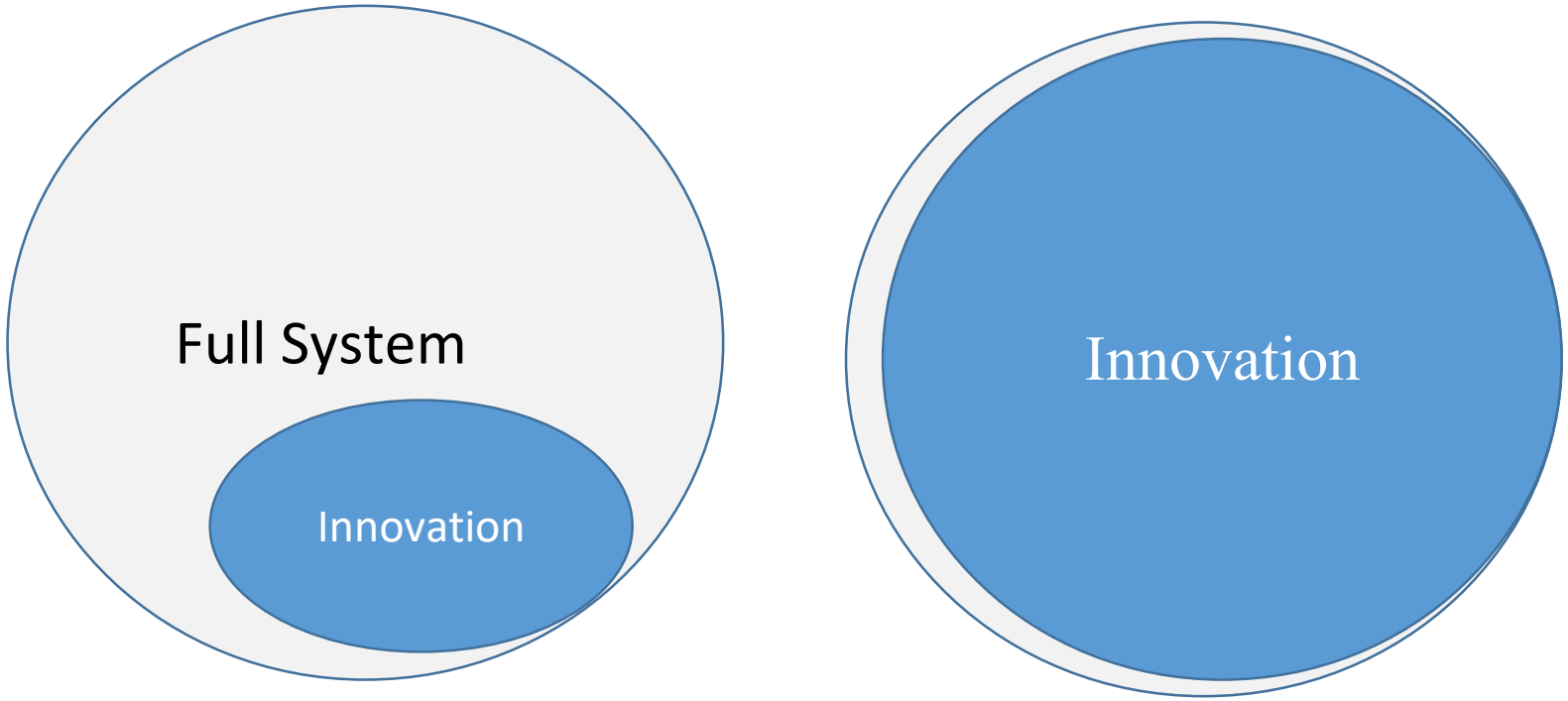
Serves to highlight the differences between typical programs and more innovative programs which supports communication

Implications of Layered Design of TOA on Innovation

- When it comes to design innovation in a balanced assessment system understanding how/where the smaller of theories of action support or potentially constrain the innovation is key.
- Must understand the interplay among the components, as they currently exist, and also understand where and how a proposed innovation may impact what is happening within and across different levels of the system.
- The impact could be significant and broad or minor and isolated depending on the dimension of change necessitated.

System vs Component Level Innovation

How much are we innovating?

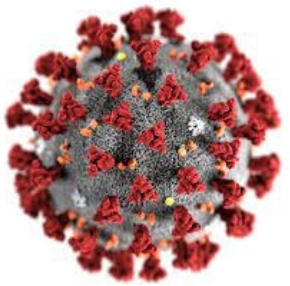


Dimensions of Innovation

- The dimension of innovation represents the key element(s) that you are looking to modify in order to address an existing problem or improve upon a current existing state.
  - Can be broad or granular.
  - Can be simple or complex
  - Can impact the entire assessment system, (across levels), a level within the system, or a specific process/test within a level.

- In some cases an innovation results from implementing an existing solution at a different level of a system. (e.g., using a locally scored performance assessments as an element of the state summative assessment).

Identifying the Focus of Innovation in BAS



The *what* of innovation may stem from:

- root cause analysis
- advances in technology
- an "event" by
  - establishing new requirements or flexibilities (federal or state laws)
  - shining a light on deficiencies in an existing systems
  - reflecting a shift in values or priorities

ESSA
Every Student Succeeds Act

Where/What is the innovation?	Example
Test Score	Type (NCE, Scaled Score, Growth Score), Interpretation (Criterion/Norm Reference), Estimation (Multidimensional IRT), Use (e.g., Instruction & Accountability), Reporting
Test Items or Tasks	Type (TEI, OE) ; Scoring (Machine, AI, Human)
Test Form	Design (Length, Content, Representation) ; Administration (standardized/individual); Mode of Delivery (paper-pencil/computer/CAT)
Test Development/ Evaluation	Who is involved (community, representation); How quality is evaluated/endorsed (e.g., peer review);
Construct Definition	Traditional, Anti-racist
Theory of Learning	Learning progressions; Learning expectations/ appropriate demonstrations of learning
Infrastructure	Design of Data and Learning Management systems
Role of State/District	Partner and Resource Provider, Auditor

Why useful to highlight the key dimension(s) underlying a proposed innovation?

- Serves to highlight/focus articulation of the theory of action (e.g., inputs, interactions, assumptions) on those things that not well understood or have not yet been played out within the system.

- And, in doing so, clarifies one's understanding of how/where the innovation will have an impact on the assessment system.

Problem	Source	Hypothesized Solution (Based on RCA)	What is the innovation?	Level of Impact
Persistent gaps in achievement outcomes across student groups	State	Engage stakeholders to review and modify the content standards to ensure they outline expectations are appropriate/fair for all students	Construct or Domain Specification (process and content)	All
Persistent gaps in achievement outcomes across student groups	District	Provide educators in schools serving predominantly low income and historically disadvantaged students with access to high quality educational resources including curriculum embedded assessments that provide ongoing feedback and support progress monitoring.	District role Curriculum and assessment design (high quality, aligned, relevant, accessible).	District, classroom

3. Examples & Tools

Unfortunately, no one can be told what a Theory Of Action is. You'll have to see it for yourself.

-Morpheus, probably

Example 1: District has not been effective in supporting the success of it's lowest performing students.

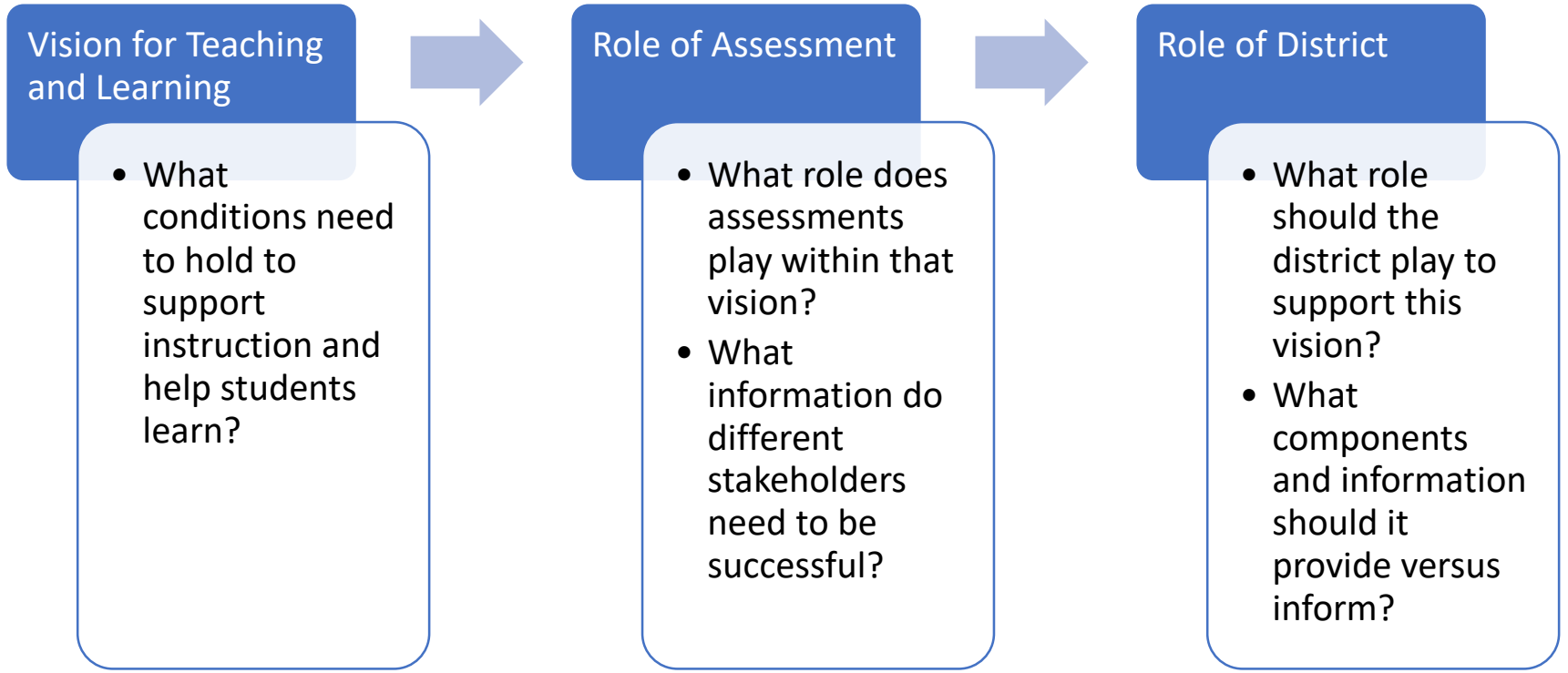
Example 2: The summative assessment does not provide educators with information in time for within year instructional shifts.

Example 1

District assessment system is incoherent and not effective in supporting the success of all students.

The Challenge for Example 1

- District has not been effective in supporting the achievement and growth of its lowest performing students.



Center for Assessment

This work is licensed under a Creative Commons Attribution 4.0 International License.

Part 2: Assessment Vision: The assessment information we value, irrespective of what we have)

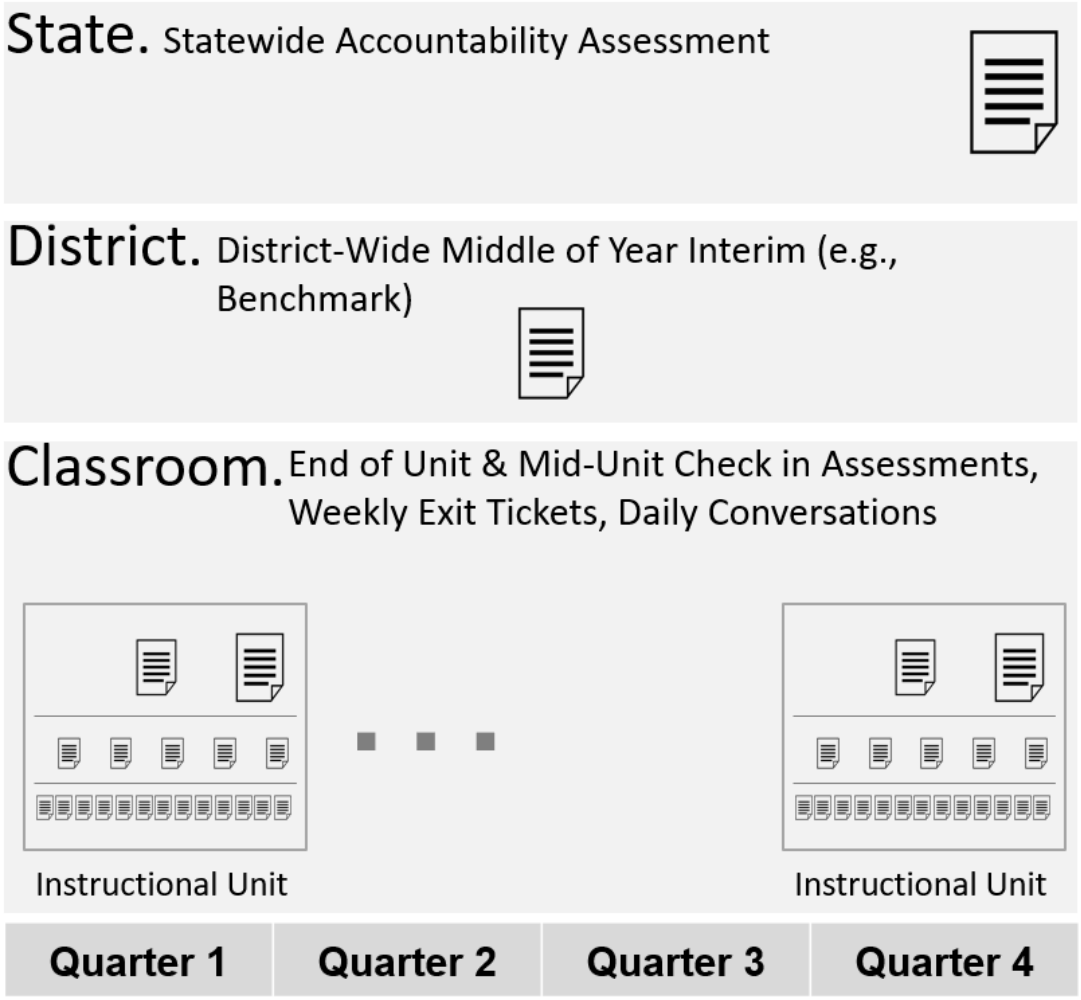
After thinking about what students, teachers, and principals should do that promote student learning, it is important to think about the kinds of assessment information that support those activities. What type of information would each group of people find most useful in their work? What, specifically, would help them make better decisions about their next steps, instruction, planning, or guidance?

Activity for Part 2: *In each row of the table, briefly respond to each question about what information from an assessment could help benefit them. The responses to the questions should be related to the response from Part 1. Please note that assessment information does not have to be the results of a test, but can include a process of gathering evidence.*

	What assessment information would be most helpful for each question?	
Student	*Consider your description of **student** success and how students learn.*	
	Which kind of assessment information do **students need** given your vision for teaching and learning and why? **How should students use** assessment information given your vision for teaching and learning?	
Teacher	*Consider your beliefs regarding how a **teacher** best facilitates student learning*	
	What type(s) of assessment information do **teachers need** to facilitate student learning and why? **How should teachers use** each type of assessment information?	(scores, student work, comparisons, informal observation) (remediation, regrouping)
Principal	*Consider what you believe **principals** should do to support high quality instruction.*	
	What assessment information do **principals need** and why (e.g., support teachers)? **How should principals use** each type of assessment information?	
Local Education Leaders	*Consider the behaviors you believe **district leaders** should engage in to facilitate teachers and students.*	
	What assessment information do **district leaders need** and why? **How should district leaders use** each type of assessment information?	
	***Others'** needs and uses (please add as many rows as necessary)*	
Others		

[2019: Reidy Interactive Lecture Series: Interim Assessment Toolkit]

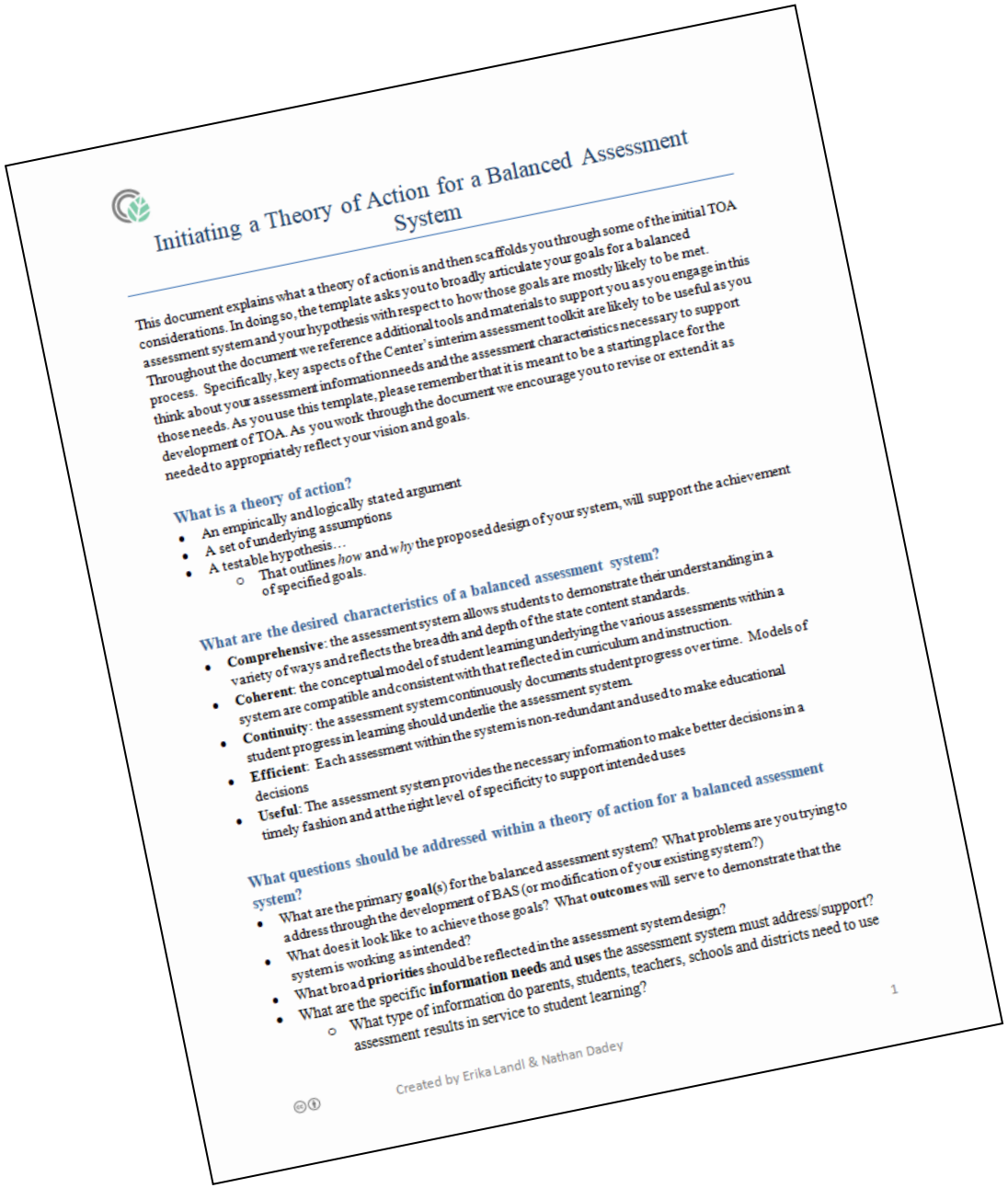
Identifying Sources of Information



- What sources/tools/processes will provide the needed information and support the intended use?
- What already exists? How is it working?
- What are the gaps?

Defining the Parameters of Innovation

The Why (RCA)	Hypothesized Solution	Nature of innovation
• Constant changes to assessment and curriculum • District expectations for student performance are unclear and defined differently across schools. • High numbers of transient students • Lack of assessment literacy • Educator use/modification of pre-existing tools that do not align to the curriculum	Providing schools with resources that serve to clarify and demonstrate the expectations of the standards (e.g., rubrics/exemplars) and common district-developed interim assessments that allow for monitoring of student progress relative to those expectations will foster collaboration and consistency within and across educators and schools.	Role of the district in defining and measuring expectations for student performance. Establishing resources that provide for coherence across schools in the absence of a common curriculum or pacing.



Initiating a Theory of Action for a Balanced Assessment System

This document explains what a theory of action is and then scaffolds you through some of the initial TOA considerations. In doing so, the template asks you to broadly articulate your goals for a balanced assessment system and your hypothesis with respect to how those goals are mostly likely to be met. Throughout the document we reference additional tools and materials to support you as you engage in this process. Specifically, key aspects of the Center's interim assessment toolkit are likely to be useful as you think about your assessment information needs and the assessment characteristics necessary to support those needs. As you use this template, please remember that it is meant to be a starting place for the development of TOA. As you work through the document we encourage you to revise or extend it as needed to appropriately reflect your vision and goals.

What is a theory of action?

- An empirically and logically stated argument
- A set of underlying assumptions
- A testable hypothesis…
  - That outlines *how* and *why* the proposed design of your system, will support the achievement of specified goals.

What are the desired characteristics of a balanced assessment system?

- **Comprehensive**: the assessment system allows students to demonstrate their understanding in a variety of ways and reflects the breadth and depth of the state content standards.
- **Coherent**: the conceptual model of student learning underlying the various assessments within a system are compatible and consistent with that reflected in curriculum and instruction.
- **Continuity**: the assessment system continuously documents student progress over time. Models of student progress in learning should underlie the assessment system.
- **Efficient**: Each assessment within the system is non-redundant and used to make educational decisions
- **Useful**: The assessment system provides the necessary information to make better decisions in a timely fashion and at the right level of specificity to support intended uses

What questions should be addressed within a theory of action for a balanced assessment system?

- What are the primary **goal(s)** for the balanced assessment system? What problems are you trying to address through the development of BAS (or modification of your existing system?)
- What does it look like to achieve those goals? What **outcomes** will serve to demonstrate that the system is working as intended?
- What broad **priorities** should be reflected in the assessment system design?
- What are the specific **information needs** and **uses** the assessment system must address/support?
  - What type of information do parents, students, teachers, schools and districts need to use assessment results in service to student learning?

1

Created by Erika Landl & Nathan Dadey

Center for Assessment

6 sections

I. Goals
II. Outcomes
III. Design Priorities
IV. Information Needs and Uses
V. System Components
VI. Conditions and Inputs

The Center's Theory of Action Template is available online. Please use and modify as you see fit.

Addressing Innovation in the TOA

How/where will this innovation be represented in the District's theory of action?

- Specification of the characteristics of system components
- Identification of assumptions underlying the revised system working as intended/specification of potential solutions.
  - Quality and utility of district-provided resources
  - Assessment literacy
- Conditions and inputs necessary to ensure the system works as intended.

Example 2

District assessment system is incoherent and not effective in supporting the success of all students.

Skip to Q&A

Hypothetical Example High Level Theory Of Action for a Through-Year Assessment Program (Math)

Assumptions

If the state provides:	Then:	So that:
• Assessment scores in early fall, winter and spring based through course administration • Assessment literacy training on use of assessment data	• Educators will adjust instruction to meet the needs of students • Administrators will examine trends to allocate additional support	• Students achievement improves

This topic is the subject of an upcoming online convening hosted by the Center: Claims and Evidence for Through Year Assessments: What We Know and What We Need to Know.

Framing Assumptions

- Educators will value and engage with state provided (a) professional development and (b) assessment information
- The provided assessment evidence can be used by educators to adjust instruction
- Adjusting instruction is sufficient to improve student learning and resulting student achievement

Which of these assumptions are both critical and risky? Which of these do we need to investigate?

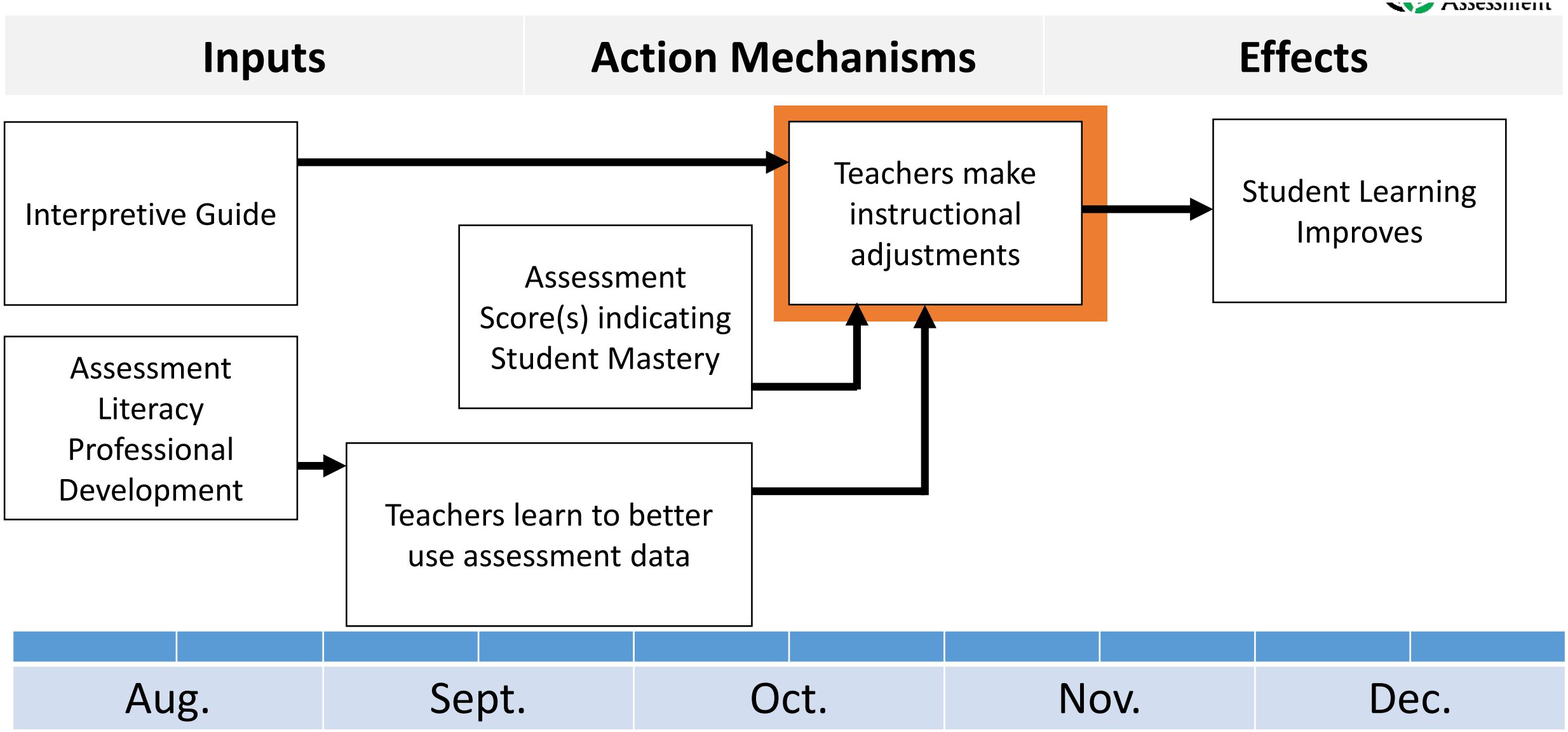
Hypothetical Example *High* Level Theory Of Action for a Through-Year Assessment Program (Math)

Assumptions

If the state provides:	Then:	So that:
• Assessment scores in early fall, winter and spring based through course administration • Assessment literacy training on use of assessment data	• Educators will adjust instruction to meet the needs of students • Administrators will examine trends to allocate additional support	• Student learning improves

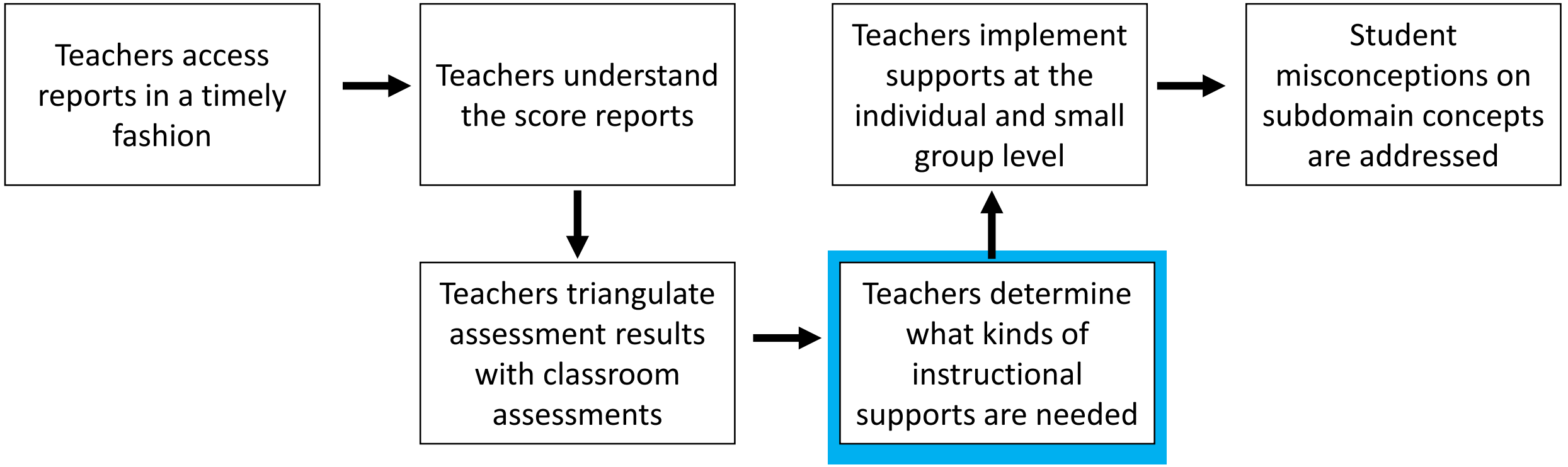
We will drill down into these areas.

Example *Slightly More Detailed* Logic Model



Teachers make instructional adjustments

Supporting Detail

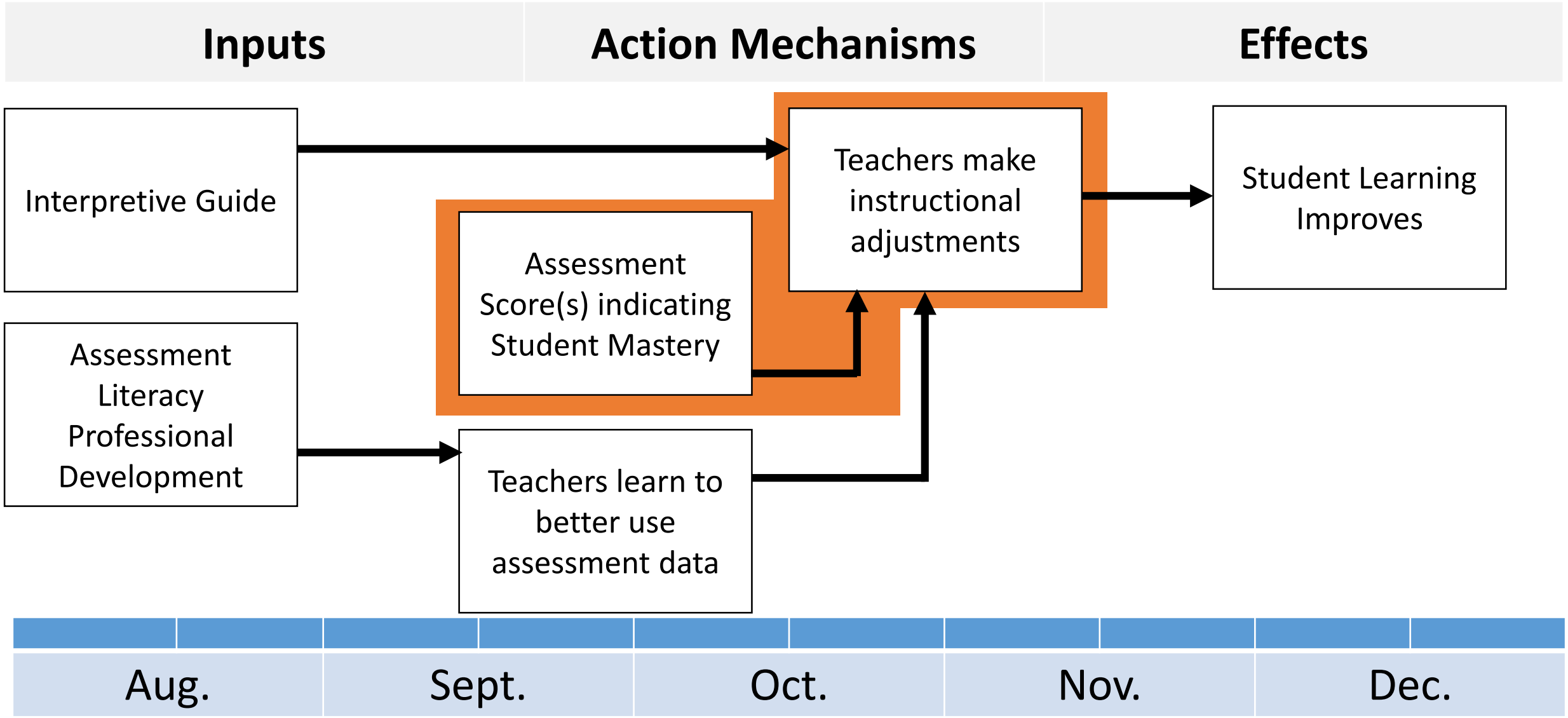


Teachers determine what kinds of instructional supports are needed

Specific Assumptions

- Teachers have access to, and understand, a variety of instructional supports connected to the math subdomains
- Teachers can use assessment evidence to select amongst the various kinds of support
- The assessment evidence is needed to select amongst these kinds of support

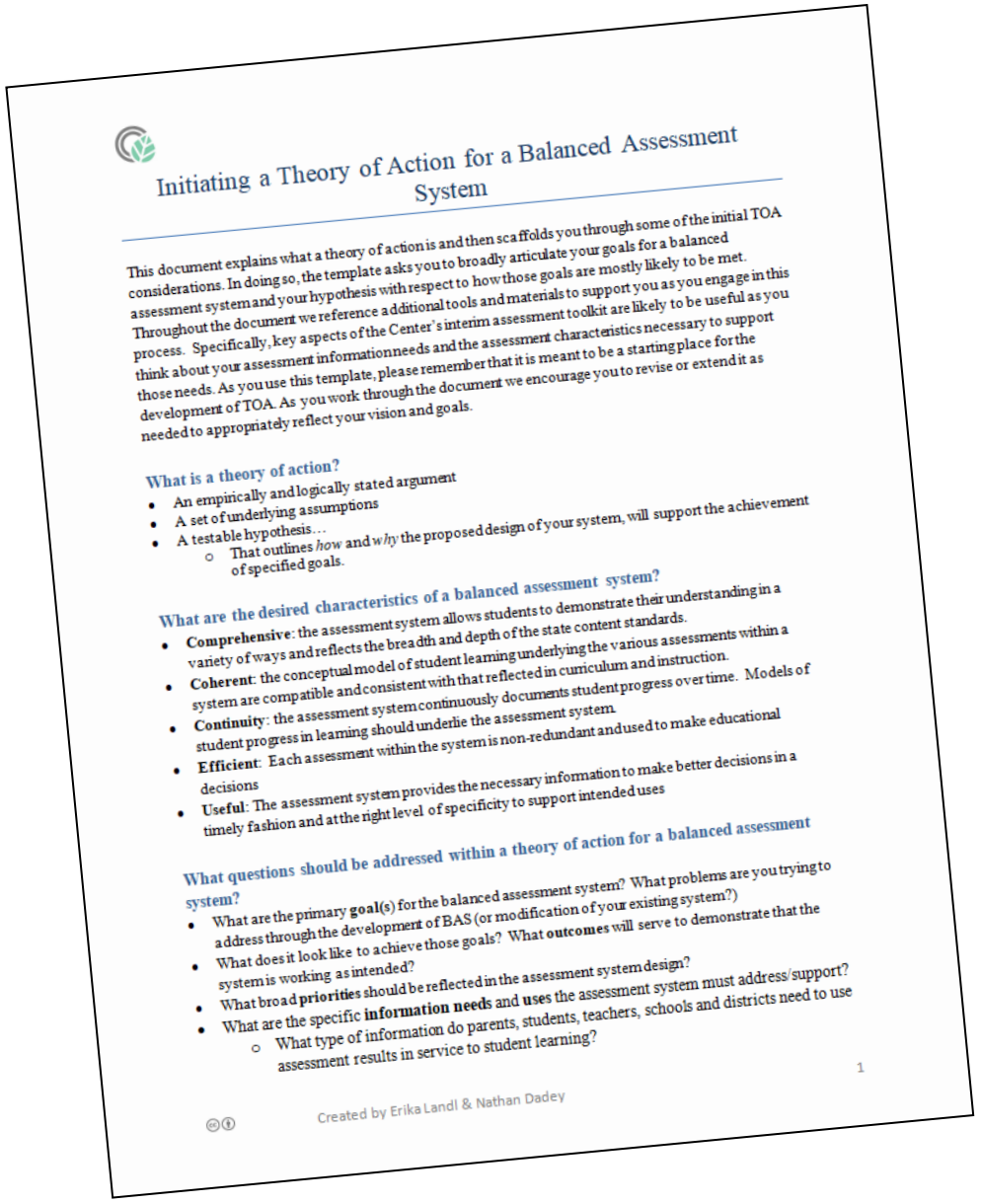
Example *Slightly More Detailed* Logic Model



4. Q&A

Wisdom from Scott

Facilitator: Scott Marion



Initiating a Theory of Action for a Balanced Assessment System

This document explains what a theory of action is and then scaffolds you through some of the initial TOA considerations. In doing so, the template asks you to broadly articulate your goals for a balanced assessment system and your hypothesis with respect to how those goals are mostly likely to be met. Throughout the document we reference additional tools and materials to support you as you engage in this process. Specifically, key aspects of the Center's interim assessment toolkit are likely to be useful as you think about your assessment information needs and the assessment characteristics necessary to support those needs. As you use this template, please remember that it is meant to be a starting place for the development of TOA. As you work through the document we encourage you to revise or extend it as needed to appropriately reflect your vision and goals.

What is a theory of action?

- An empirically and logically stated argument
- A set of underlying assumptions
- A testable hypothesis…
  - That outlines *how* and *why* the proposed design of your system, will support the achievement of specified goals.

What are the desired characteristics of a balanced assessment system?

- **Comprehensive**: the assessment system allows students to demonstrate their understanding in a variety of ways and reflects the breadth and depth of the state content standards.
- **Coherent**: the conceptual model of student learning underlying the various assessments within a system are compatible and consistent with that reflected in curriculum and instruction.
- **Continuity**: the assessment system continuously documents student progress over time. Models of student progress in learning should underlie the assessment system.
- **Efficient**: Each assessment within the system is non-redundant and used to make educational decisions
- **Useful**: The assessment system provides the necessary information to make better decisions in a timely fashion and at the right level of specificity to support intended uses

What questions should be addressed within a theory of action for a balanced assessment system?

- What are the primary **goal(s)** for the balanced assessment system? What problems are you trying to address through the development of BAS (or modification of your existing system?)
- What does it look like to achieve those goals? What **outcomes** will serve to demonstrate that the system is working as intended?
- What broad **priorities** should be reflected in the assessment system design?
- What are the specific **information needs** and **uses** the assessment system must address/support?
  - What type of information do parents, students, teachers, schools and districts need to use assessment results in service to student learning?

1

Created by Erika Landl & Nathan Dadey

All materials shared here, including the Theory of Action Template, can be found on a shared Google Drive folder & will be posted to www.nciea.org shortly.

Appendix: Additional References

A Balanced Assessment System is:

Center for Assessment

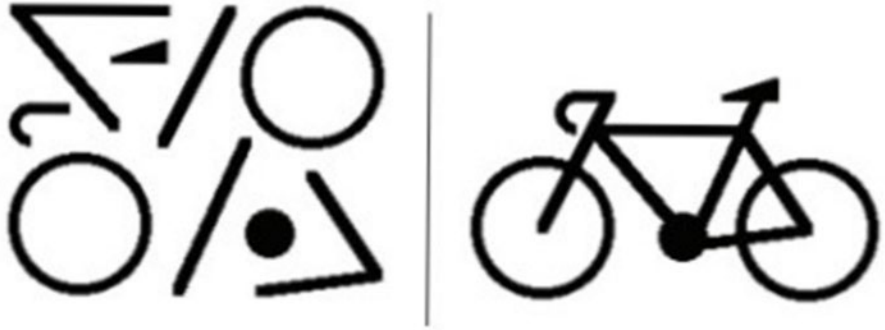
Multiple assessments with potentially **different designs**, sponsored by different people, who are at **different levels of control**

Coordinated by a **common theory of learning**

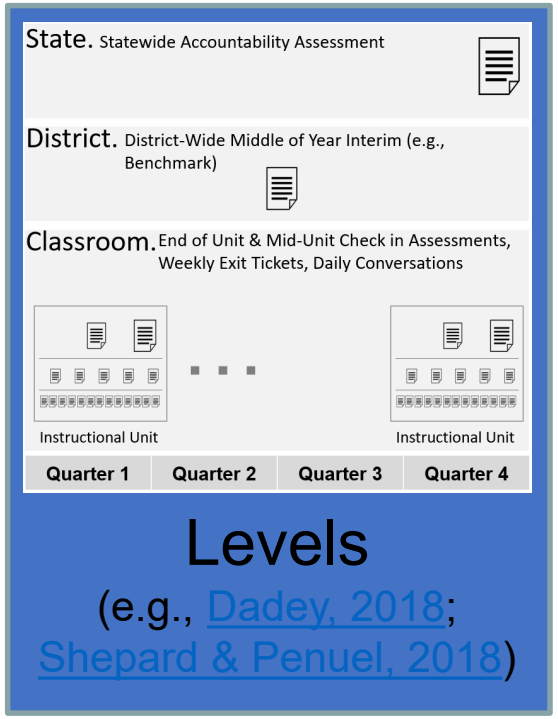
Working together to meet a **specific use or uses**

In Other Words…

- to have the desired impact:
  - *"to provide information and data that informs policy, programs, and individual teachers and learners in a coherent and coordinated manner….and improve student learning and school capacity" (Gong, 2010).*
- a balanced assessment system cannot be a bunch of assessments that are designed and implemented independently of one another.

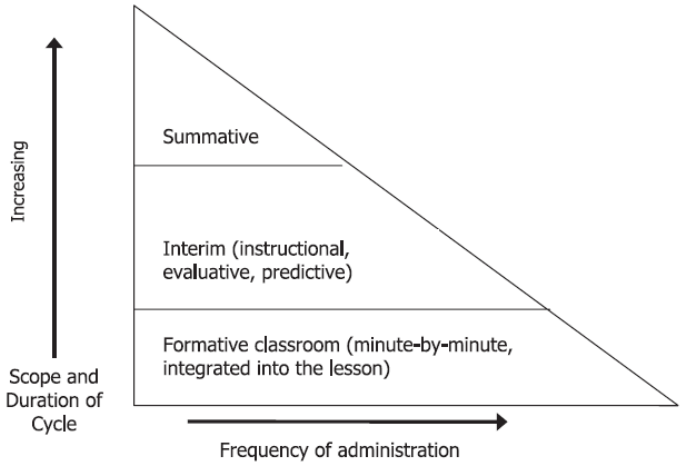


Conceptualizing the Components of a Balanced Assessment System



Levels

(e.g., Dadey, 2018; Shepard & Penuel, 2018)



Type or Tier

(e.g., Perie, Marion & Gong, 2009; Sigman & Mancuso, 2017)

- Monitoring
- Classroom
- OTL

Purpose

(e.g., NRC, 2014)

To a large degree, these conceptualizations (or others!) shape how we design systems as well as ToAs.

Regardless of how the system is conceptualized, the "overall" theory of action can and must connect all of the parts together for it to be truly balanced.