



Carnegie Foundation
for the Advancement of Teaching

Measurement for Improvement

Melissa Chabran

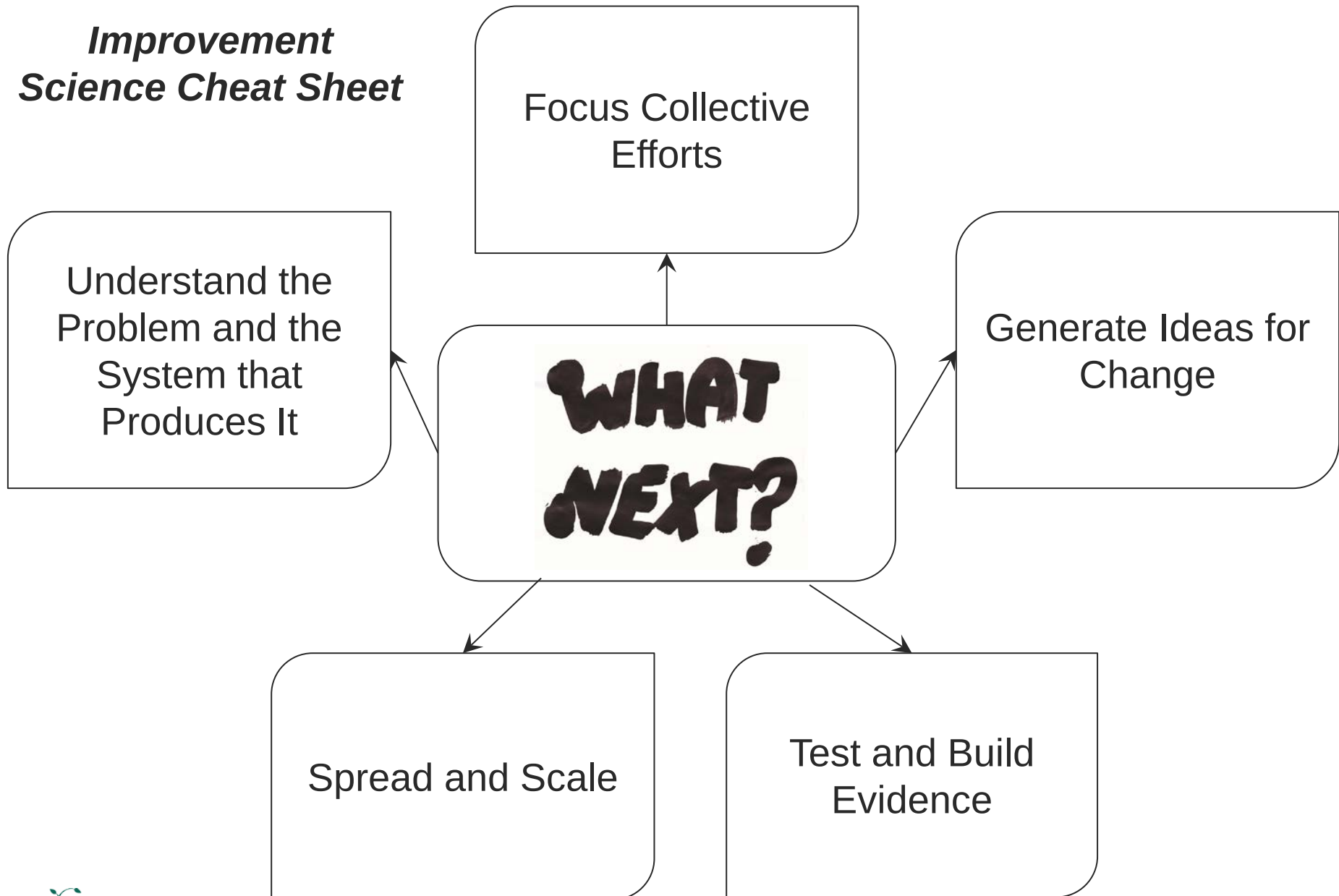
Associate, Networked Improvement Science
Carnegie Foundation for the Advancement of Teaching

Jon Norman

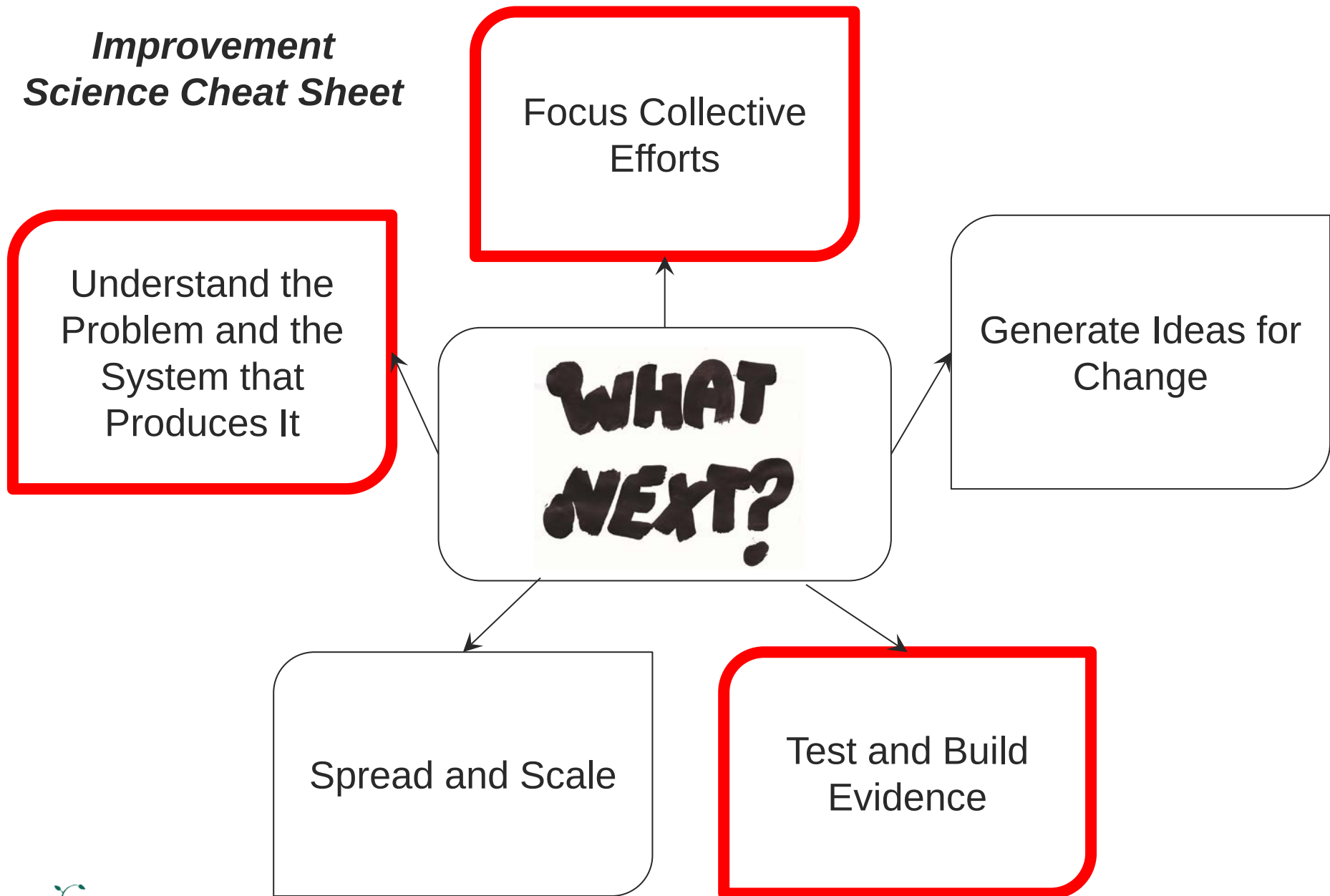
Associate, Improvement Analytics & Measurement
Carnegie Foundation for the Advancement of Teaching

April 4, 2018

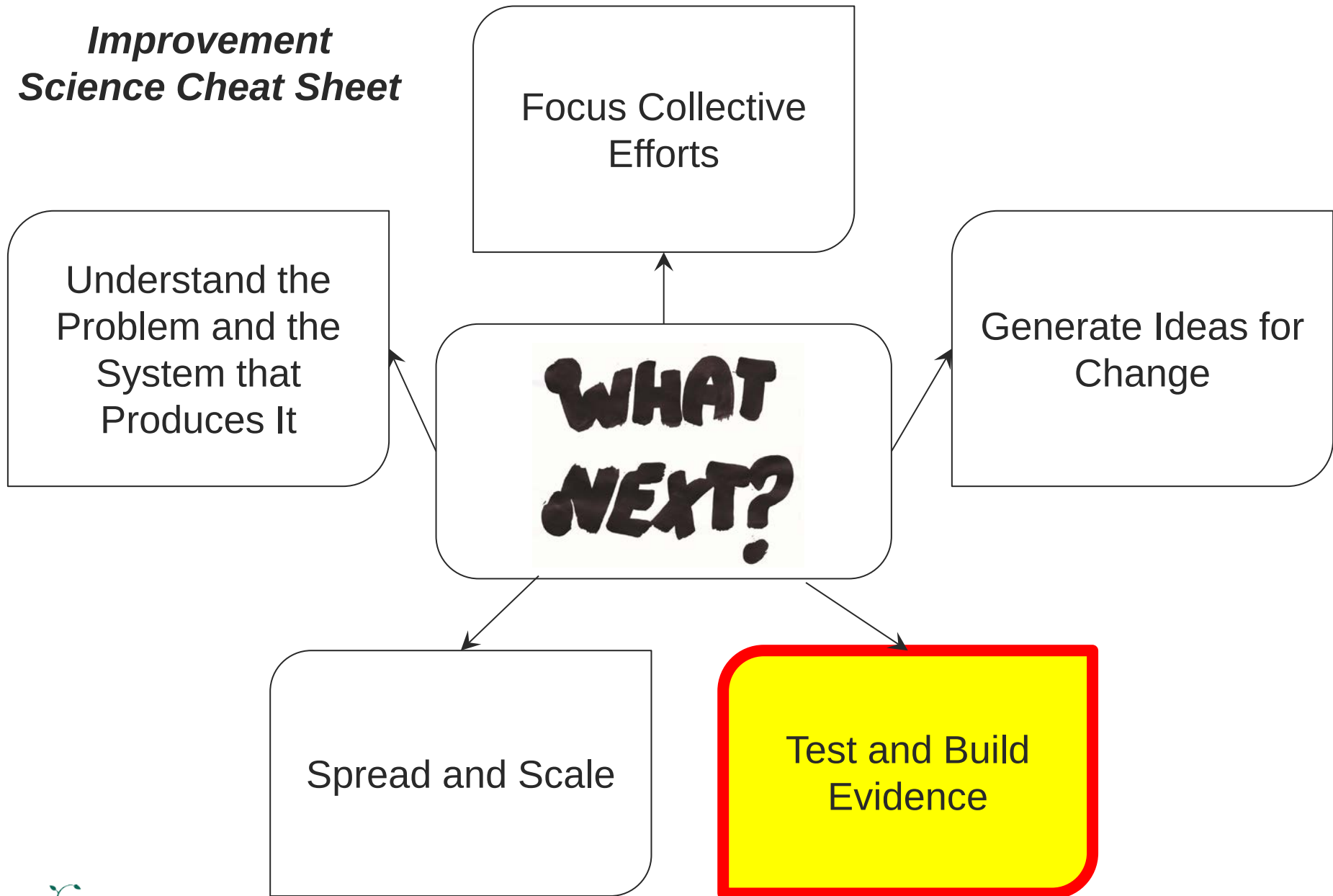
Improvement Science Cheat Sheet



Improvement Science Cheat Sheet



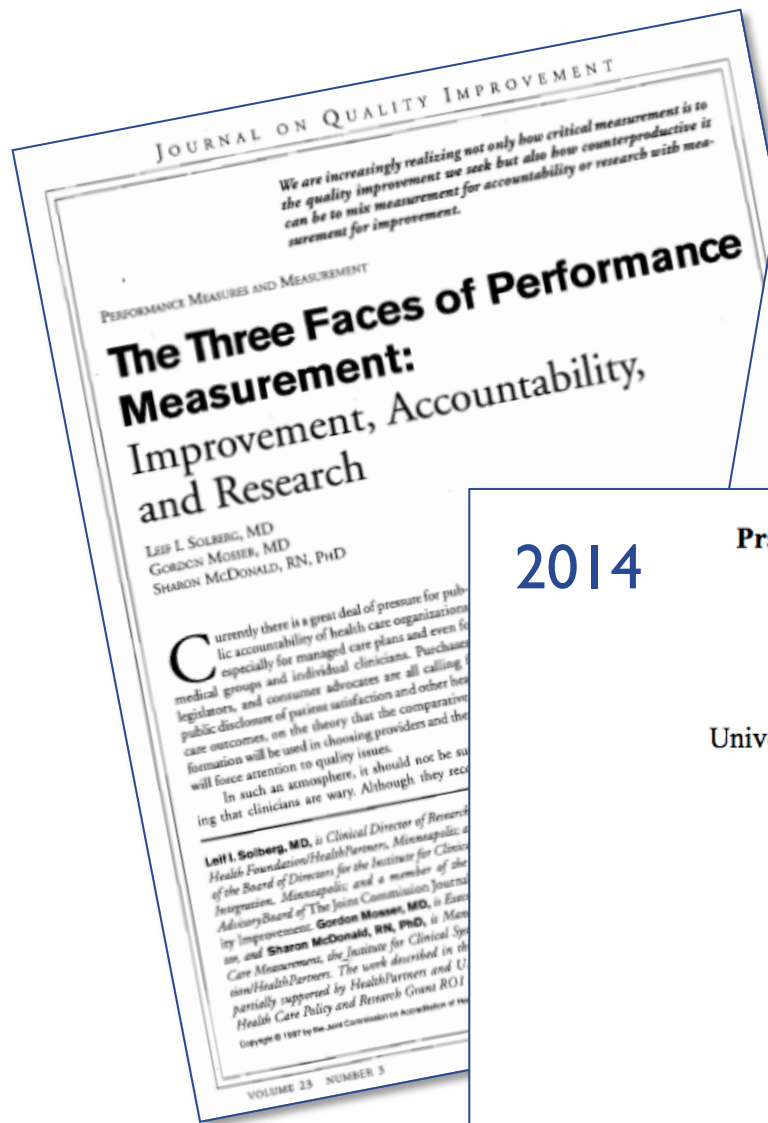
Improvement Science Cheat Sheet



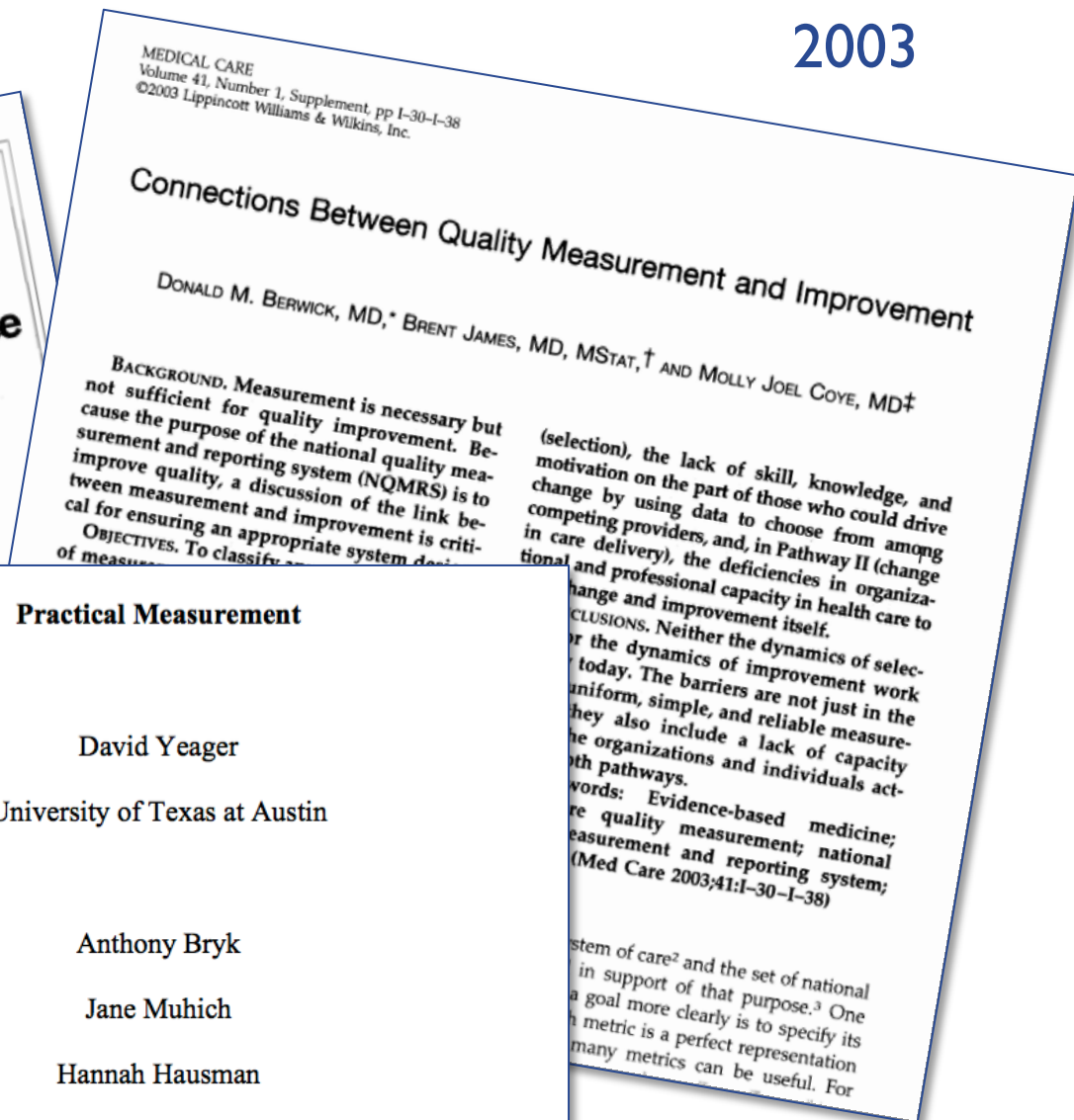
Session Objectives

- 1. The Purpose of Measurement:** Explore how measurement for improvement is different than measurement used for accountability or research.
- 2. System of Measures:** Motivate the need for a “system of measures” for improvement that includes outcome, driver, process and balancing measures.
- 3. Case:** Show how a measurement system supports improvement efforts.

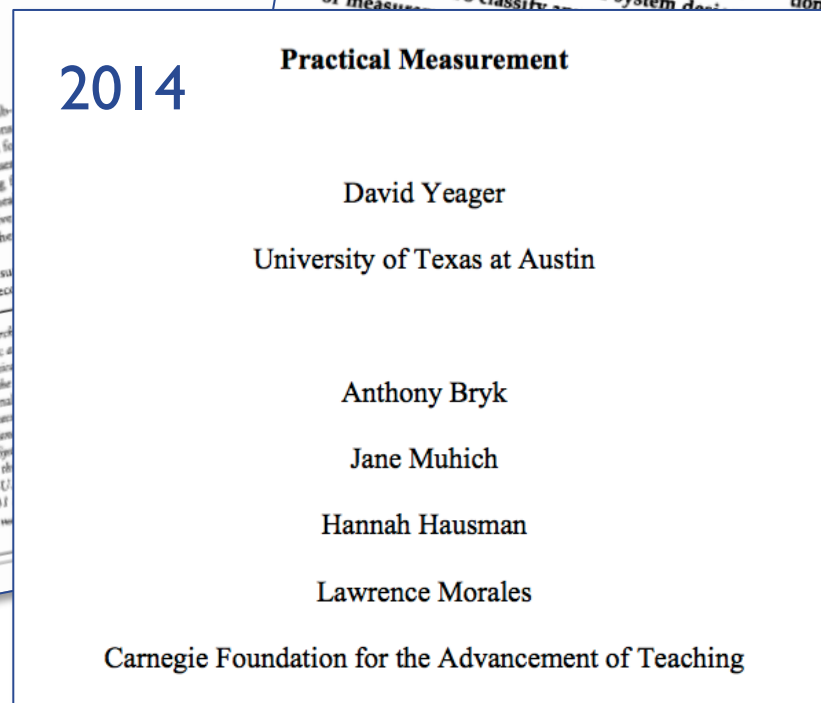
1997



2003



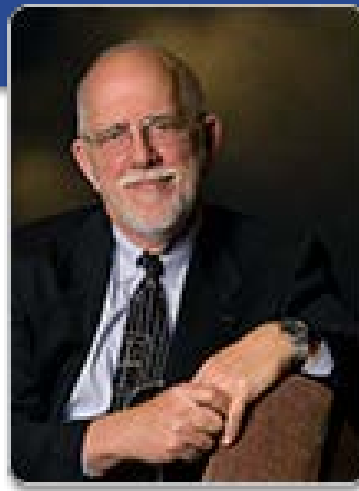
2014



Lessons from Healthcare

“We are increasingly realizing how critical measurement is for the improvement we seek, yet how counterproductive it can sometimes be to mix measurement for accountability with measurement for improvement. Considered one by one, measurement for each purpose can be good and very important. If done poorly, it can be bad. If the measurements are mixed together in inappropriate ways, they can indeed be harmful or destructive, with the mixed purposes interfering with one another.”

Leif I. Solberg, MD



To provide usable information for improvement, we need to consider:

What is measured:

Needs to be **closely aligned** to the actual work and **specific** to the processes and outcomes you hope to change.

How & when it is measured:

Needs to be **embedded** in the daily workflow.

Must produce data accessible in a **timely** manner.

Social processes shaping *use*:

Requires transparency, **trust**, low stakes, and the **safety** to take risks.

Language needs to be meaningful to those doing the work

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual variables
<u>What?</u>	Outcomes and processes that are the object of change	End of the line outcomes	Latent variables

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual variables
<u>What?</u>	Outcomes and processes that are the object of change	End of the line outcomes	Latent variables
<u>How often?</u>	Frequently as practice occurs	Usually collected once a year (after the fact)	Typically once or twice per study (after the fact)
<u>Testing your theory</u>	Sequential tests	No theory to test	One large test
<u>Sample size</u>	“Just enough” data, small sequential samples	Obtain 100% of available, relevant data	“Just in case” data

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual variables
<u>What?</u>	Outcomes and processes that are the object of change	End of the line outcomes	Latent variables
<u>How often?</u>	Frequently as practice occurs	Usually collected once a year (after the fact)	Typically once or twice per study (after the fact)
<u>Testing your theory</u>	Sequential tests	No theory to test	One large test
<u>Sample size</u>	“Just enough” data, small sequential samples	Obtain 100% of available, relevant data	“Just in case” data
<u>Social Conditions of Use?</u>	Data shared in a low-stakes, safe environment conducive to change.	Publically available. Formal collection process to assure appearances of neutrality and objectivity.	Meets scientific standards that are held in the field.

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual variables
<u>What?</u>	Key limitation for improvement: Does not illuminate WHY the outcomes occur or what should be done to change them	End of the line outcomes	Latent variables
<u>How?</u>		Usually collected once a year (after the fact)	Typically once or twice per study (after the fact)
<u>Testing your theory?</u>		No theory to test	One large test
<u>Sample size</u>		Obtain 100% of available, relevant data	“Just in case” data
<u>Social Conditions of Use?</u>		Publically available. Formal collection process to assure appearances of neutrality and objectivity.	Meets scientific standards that are held in the field.

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools,	Develop and test theories about the relationships between conceptual variables
<u>What?</u>	Outcomes and that are the object		Latent variables
<u>How often?</u>	Frequently as pra		Typically once or twice per study (after the fact)
<u>Testing your theory</u>	Sequential		One large test
<u>Sample size</u>	"Just enough" of sequential samples	relevant data	"Just in case" data
<u>Social Conditions of Use?</u>	Data shared in a low-stakes, safe environment conducive to change.	Publically available. Formal collection process to assure appearances of neutrality and objectivity.	Meets scientific standards that are held in the field.

Key limitation for improvement:

IMPRACTICAL to administer; not designed to inform changes in practice

ASPECT	IMPROVEMENT	ACCOUNTABILITY	RESEARCH
<u>Why?</u>	Develop and evaluate changes in practice	Identify exemplary or problematic performers (teachers, schools,	Develop and test theories about the relationships between conceptual
<u>What?</u>	Outcomes and processes that are the object of change	End of the	les
<u>How often?</u>	Frequently as practice occurs	Usually a year (a	twice per fact)
<u>Testing your theory</u>	Sequential tests	No t	st
<u>Sample size</u>	"Just enough" data, small sequential samples	Obtain 10 relevant data	data
<u>Social Conditions of Use?</u>	Data shared in a low-stakes, safe environment conducive to change.	Publically available. Formal collection process to assure appearances of neutrality and objectivity.	Meets scientific standards that are held in the field.

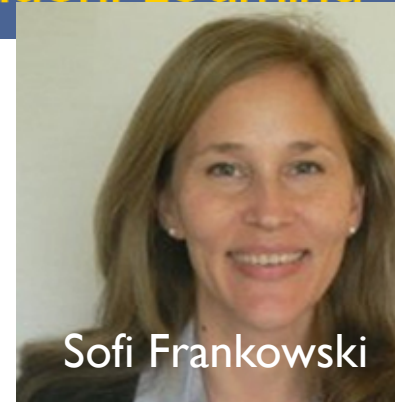
Measurement for improvement supports the ongoing refinement of knowing what works for whom and under what conditions



Carnegie Foundation
for the Advancement of Teaching

A System of Measures to Inform Improvement

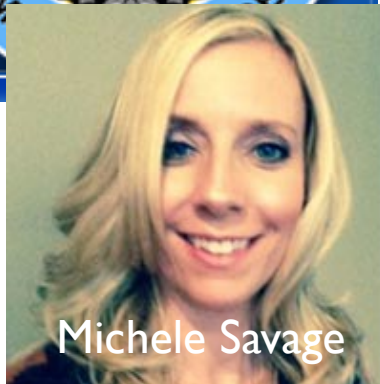
Schools That Lead – Shue-Medill – Milford Central Academy



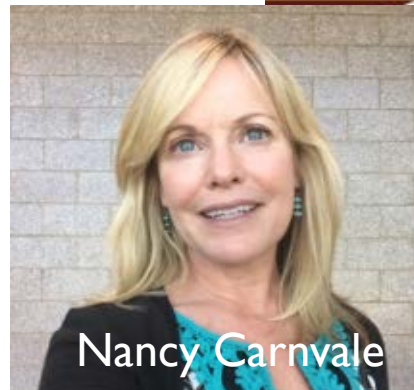
Sofi Frankowski



Dana
Diesel Wallace



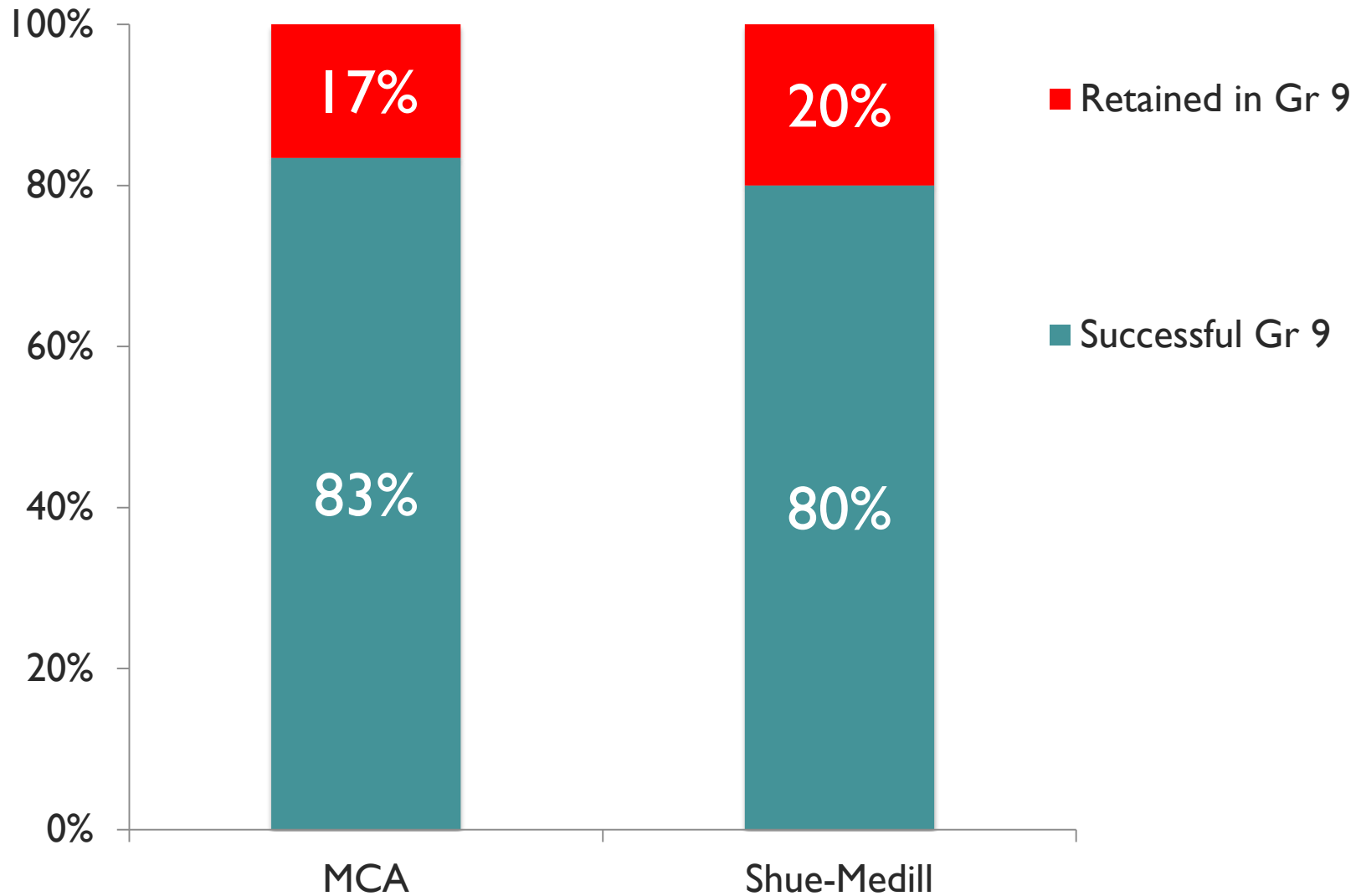
Michele Savage



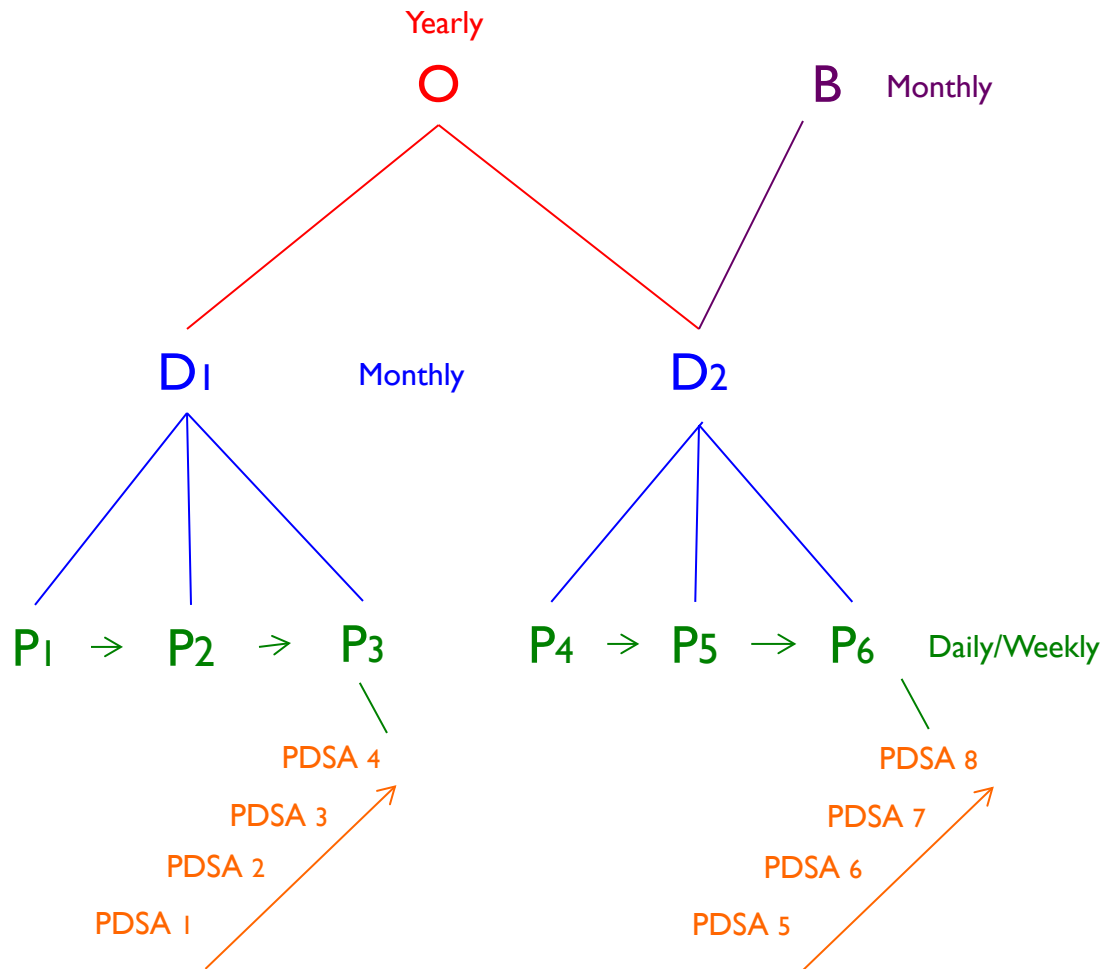
Nancy Carnvale



Proportion of Graduates Retained in Grade 9 (2015)



System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

B = Balancing Measures

PDSA = Learning Cycle Measure

Visual adapted from
Improvement Science Consulting

System of Feedback

Yearly



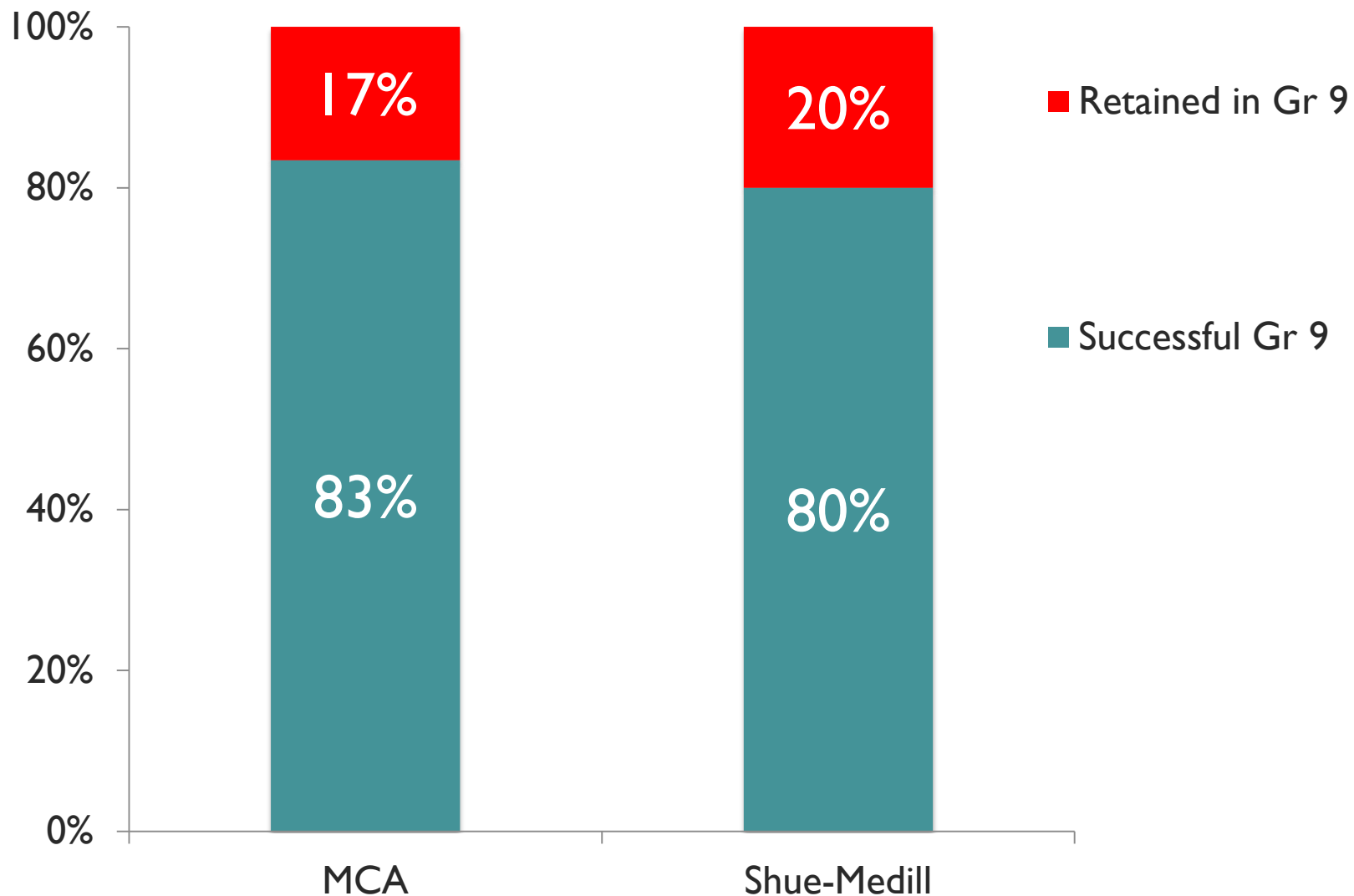
How is the system
performing?
What is the result?

Measure Types

O = Outcome Measures

Visual adapted from
Improvement Science Consulting

Proportion of Graduates Retained in Grade 9 (2015)



System of Feedback: STL-Delaware Case

Aim:

O: % of Graduates Repeating Grade 9

Measure Types

O = Outcome Measures

System of Feedback: STL-Delaware Case

Aim:

O (Lagging): % of Graduates Repeating Grade 9

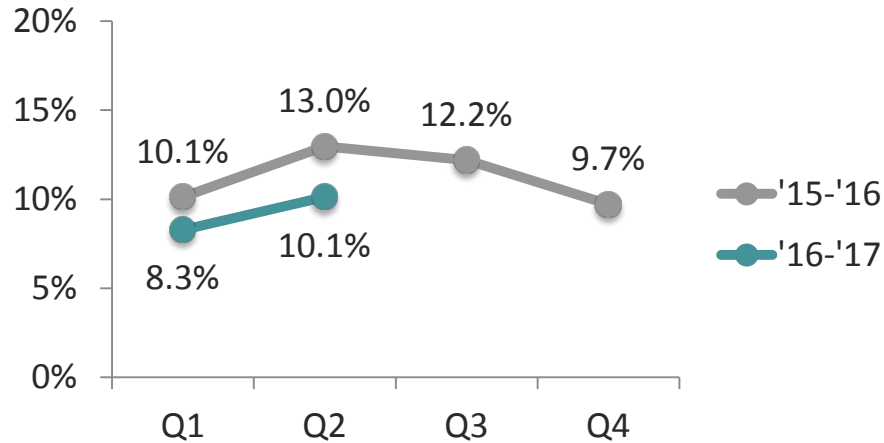
O (Leading): % of Students on Watch Lists

Measure Types

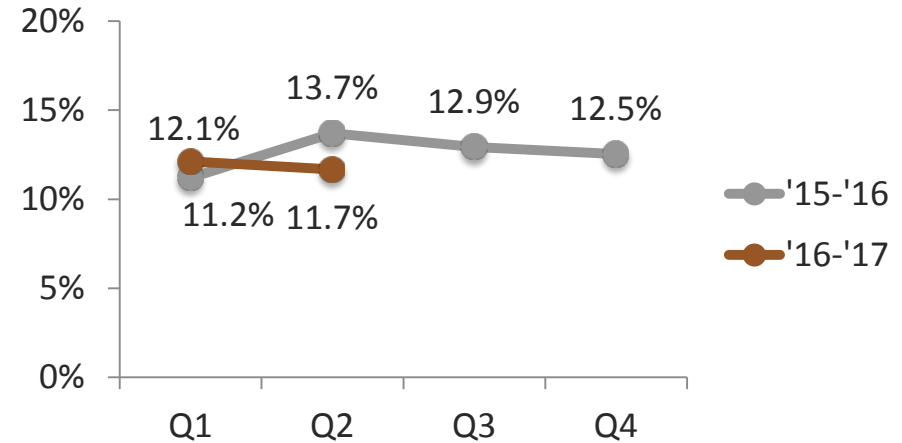
O = Outcome Measures

Leading Outcome Measures: 4 Early Warning Indicators

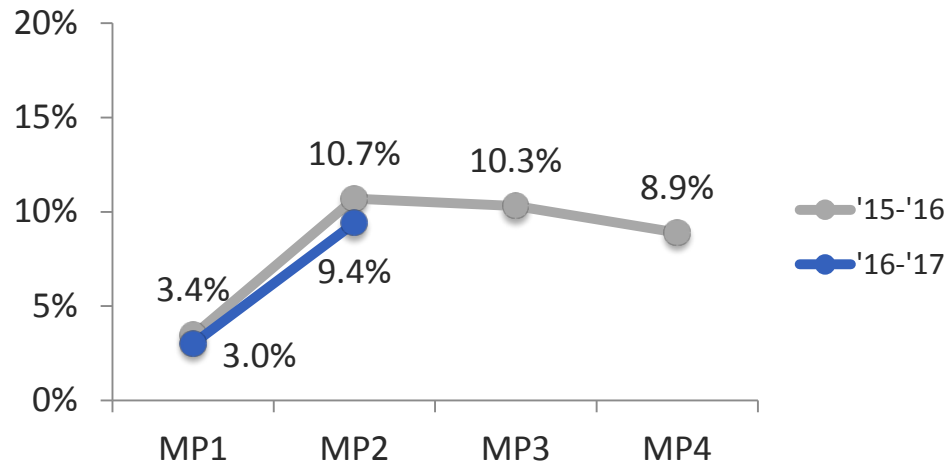
Percent Students on Math Watch List, 15-16 and 16-17



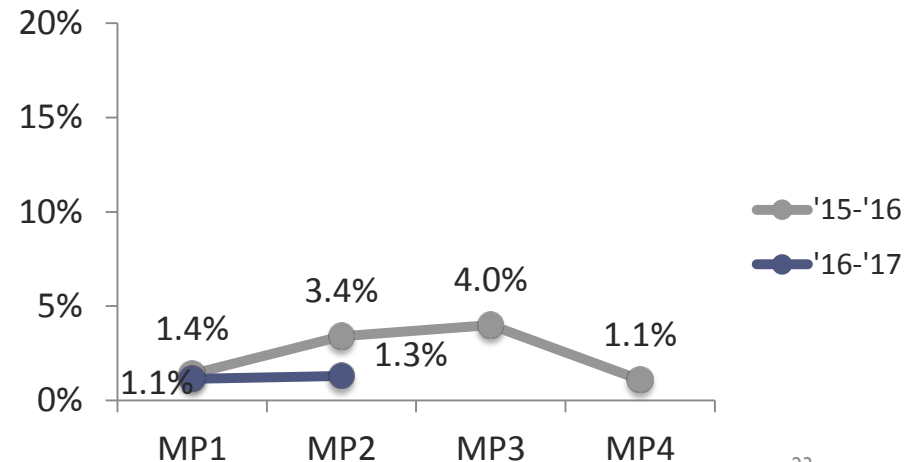
Percent Students on ELA Watch List, 15-16 and 16-17



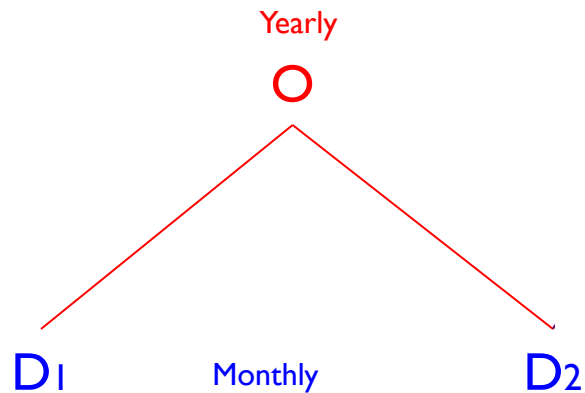
Percent Students on Attendance Watch List, 15-16 and 16-17



Percent Students on Discipline Watch List, 15-16 and 16-17



System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

Are we making progress
on the intermediate
outcomes?

Visual adapted from
Improvement Science Consulting

System of Feedback: STL-Delaware Case

Aim:

○ (Lagging): % of Graduates Repeating Grade 9

○ (Leading): % of Students on Watch Lists

Measure Types

○ = Outcome Measures

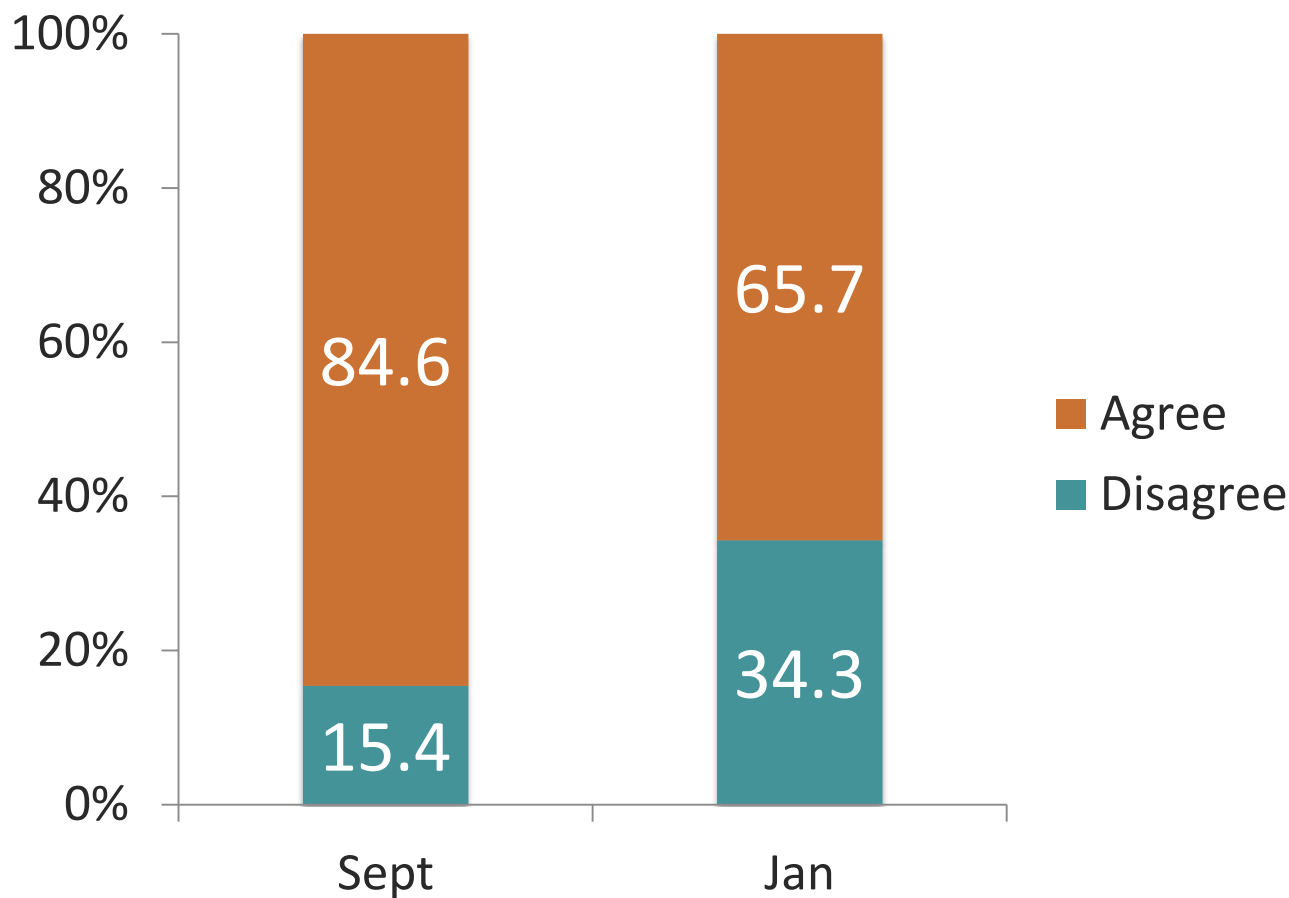
D = Driver Measures

D₁: Instructional Practices

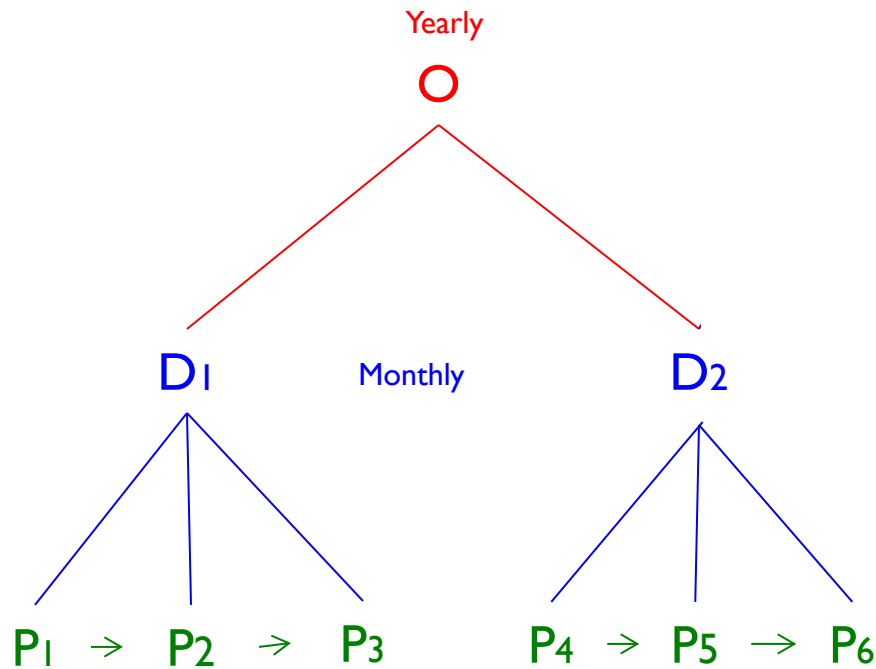
D₂: Student Mindsets

Driver Measure: Growth Mindset Student Survey Responses in Class A, September and January

Being a "math person" or not is something about you that you really can't change.
Some people are good at math and other people aren't.



System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

Are the parts of your system performing as planned?

Visual adapted from
Improvement Science Consulting

System of Feedback: STL-Delaware Case

Aim:

○ (Lagging): % of Graduates Repeating Grade 9

○ (Leading): % of Students on Watch Lists

D₁: Instructional Practices

D₂: Student Mindsets

Retesting Process

→ P₂ → P₃

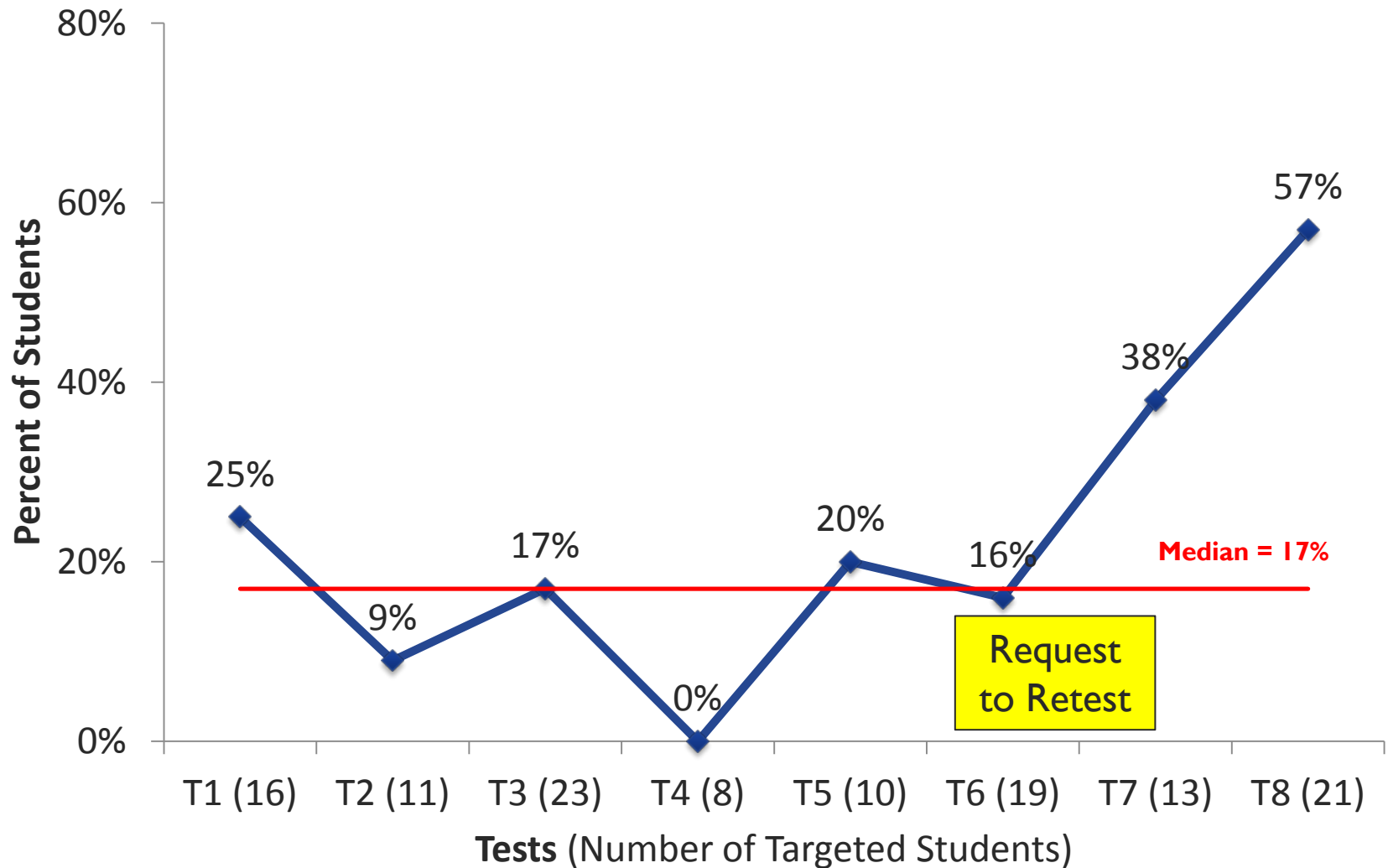
Measure Types

○ = Outcome Measures

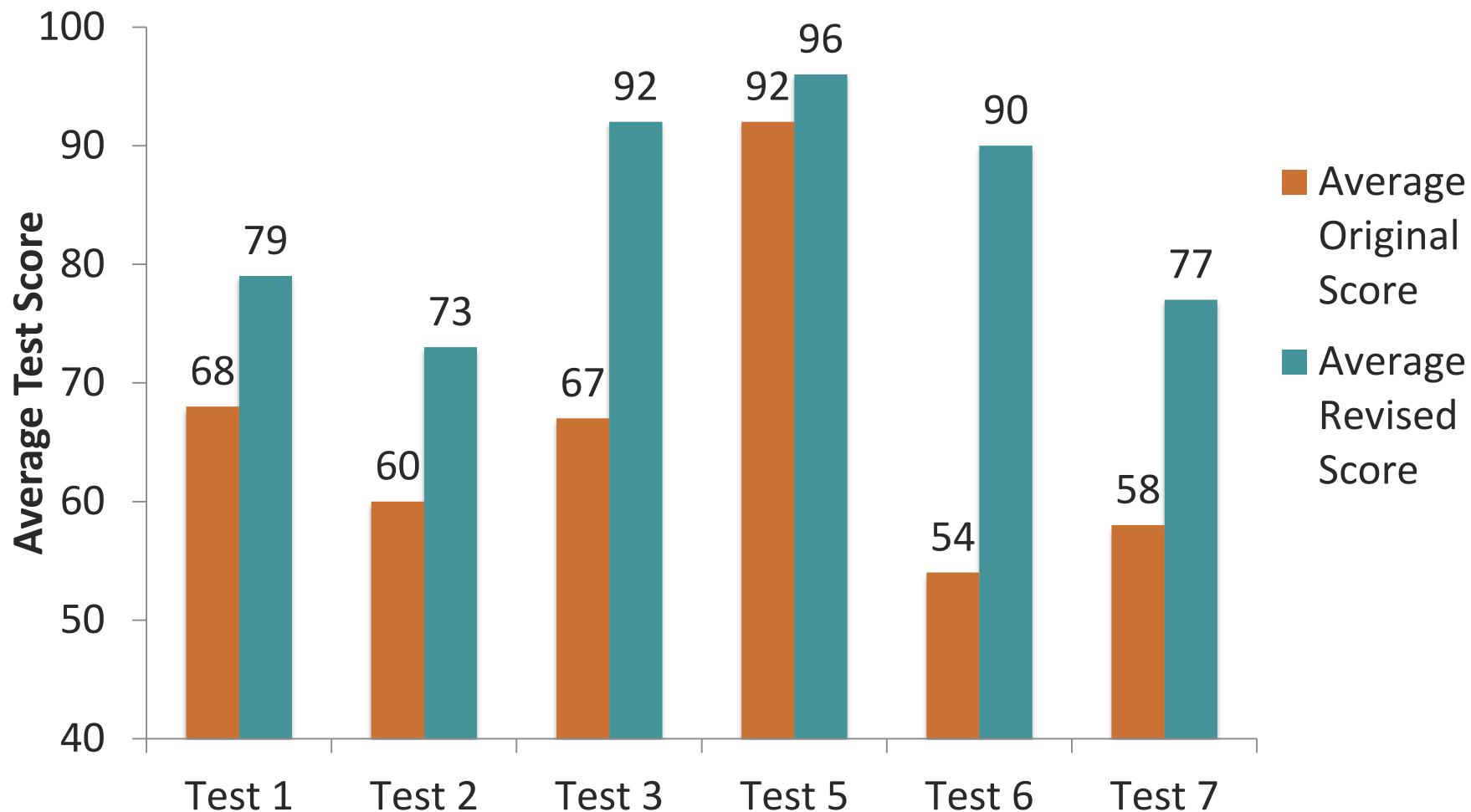
D = Driver Measures

P = Process Measures

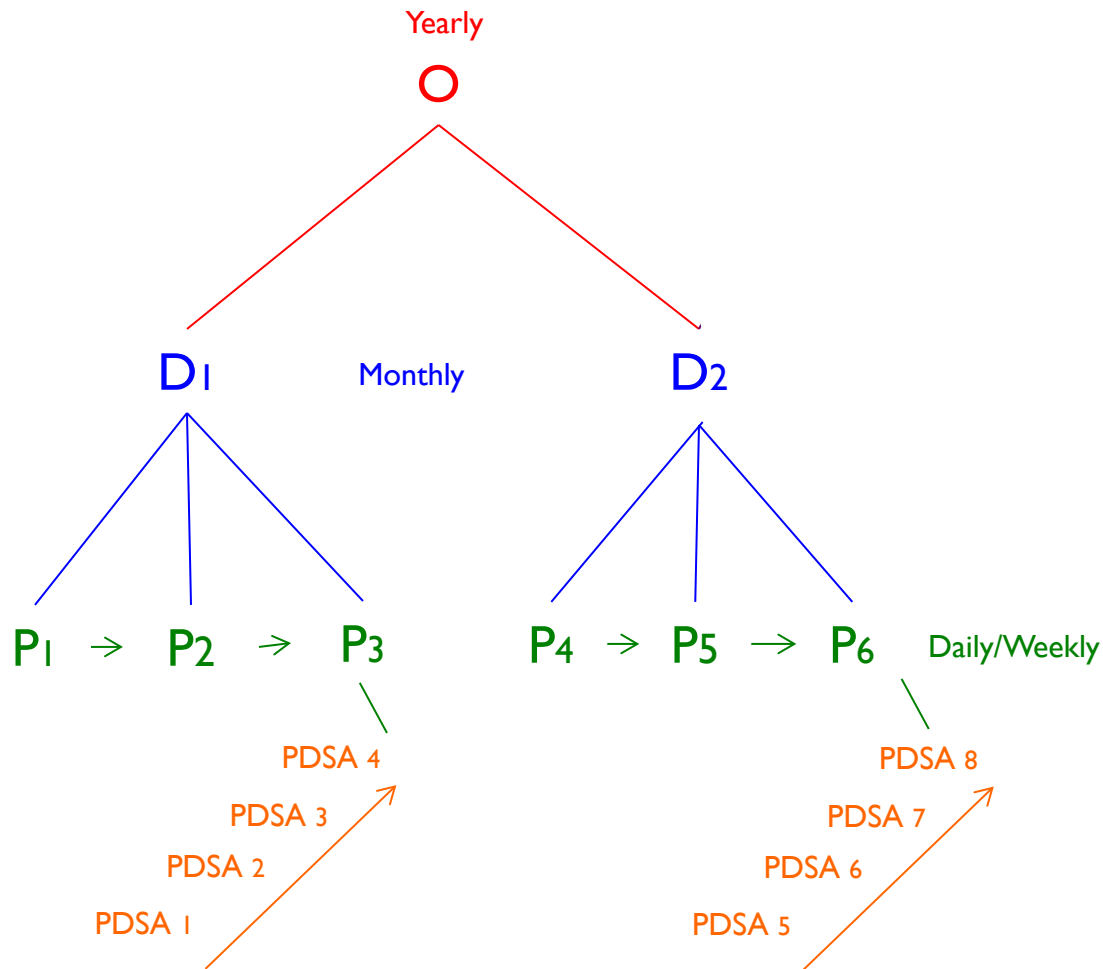
Process Measure: Percent of Targeted Students Revising Tests, Brandy Cooper, Grade 6 Math



**Process Measure:
Average Original and Revised Grades
for Students Revising Tests, Brandy Cooper, Grade 6 Math**



System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

PDSA = Learning Cycle Measure

Adapted from
Improvement Science Consulting

System of Feedback: STL-Delaware Case

Aim:

○ (Lagging): % of Graduates Repeating Grade 9

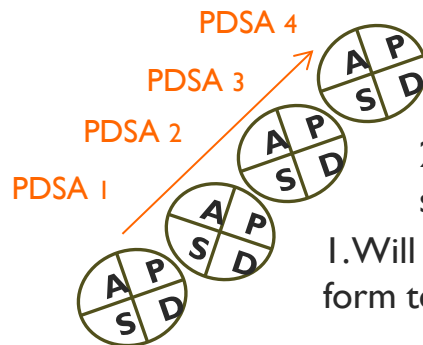
○ (Leading): % of Students on Watch Lists

D_1 : Instructional Practices

D_2 : Student Mindsets

Retesting Process

→ P_2 → P_3



2. How can I get more students to use form?
1. Will students use a form to make revisions?

Forms picked up

Reasons why students don't use form

Measure Types

○ = Outcome Measures

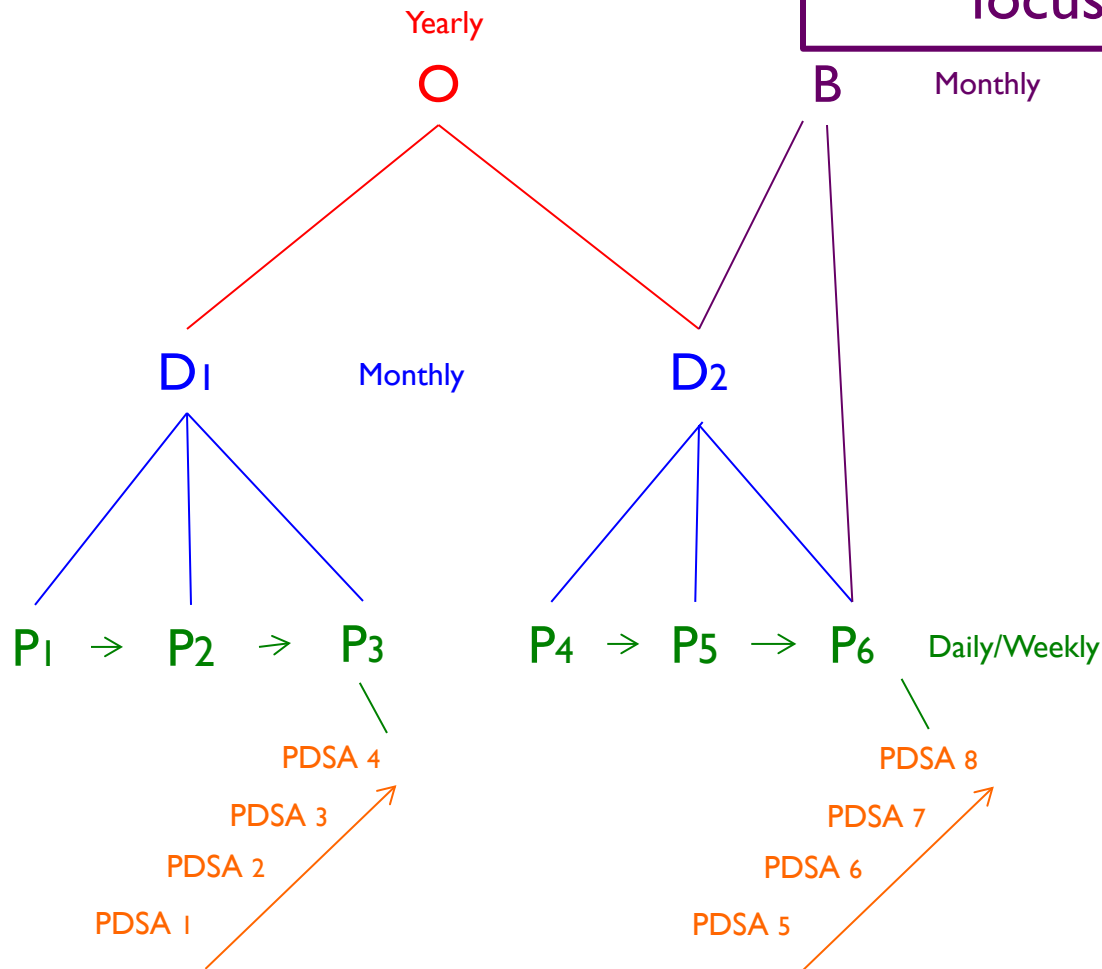
D = Driver Measures

P = Process Measures

PDSA = Learning Cycle Measure

System of Feedback

What's happening to the parts of the system we aren't currently focused on?



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

B = Balance Measures

PDSA = Learning Cycle Measure

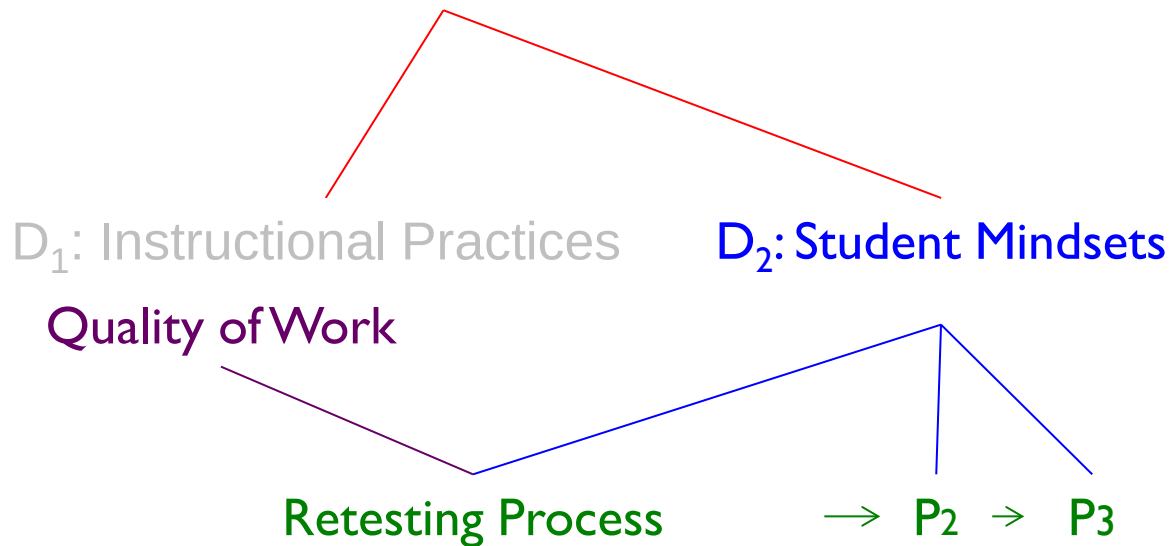
Visual adapted from
Improvement Science Consulting

System of Feedback: STL-Delaware Case

Aim:

O (Lagging): % of Graduates Repeating Grade 9

O (Leading): % of Students on Watch Lists



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

B = Balance Measures

PDSA = Learning Cycle Measure

Adapted from Schools That Lead

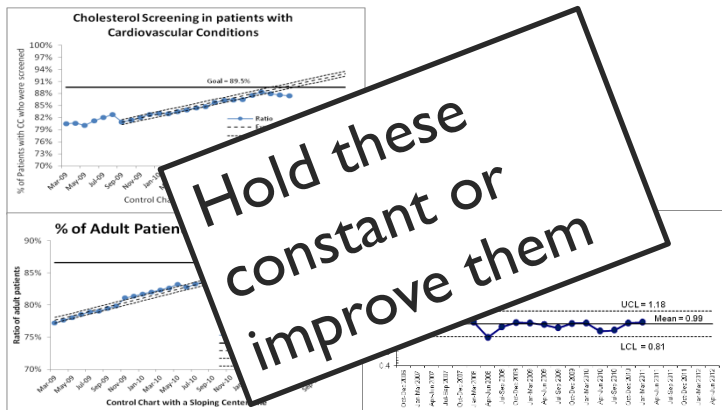
Outcome Measure(s)



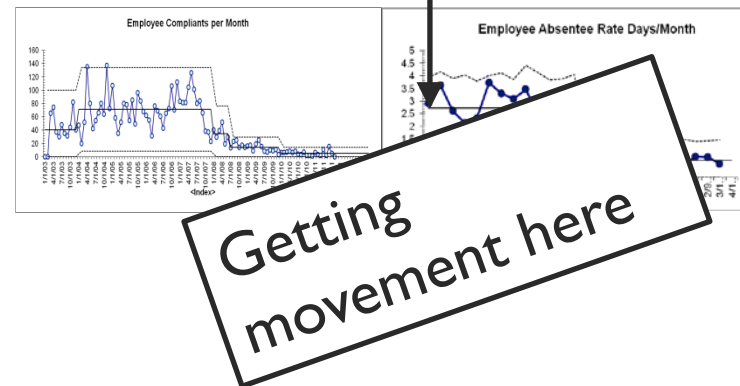
Driver Measures



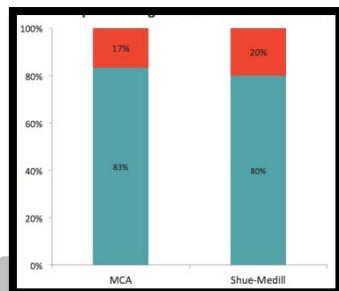
Balancing Measure(s)



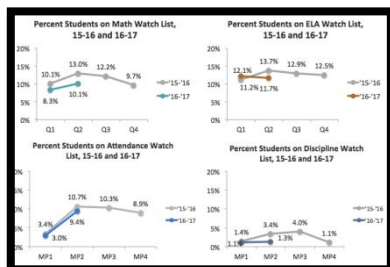
Process Measures



Lagging Outcome



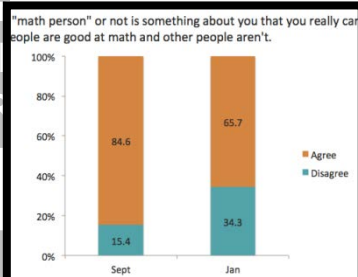
By May 2017, halve the proportion (from 1 in 5 to 1 in 10) of students who get held back in 9th grade



Leading Outcome

Primary Drivers

Driver



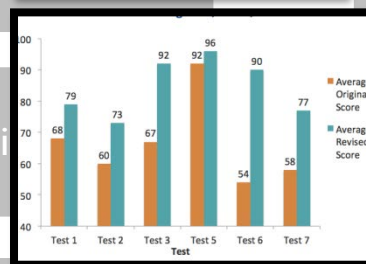
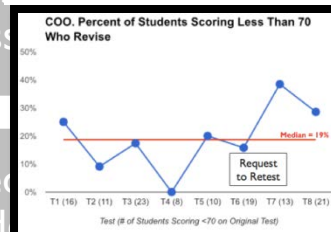
Instructional Practices & Mindsets

Student Behavioral Supports

Others...

Secondary Drivers

Process



Change Ideas

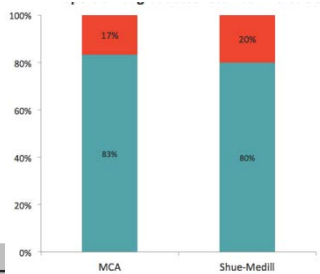
Request to Retest Form

Etc.

Balancing

Quality of First Attempt

Charter. Summer 2015



Common network measure. First survey September 2015

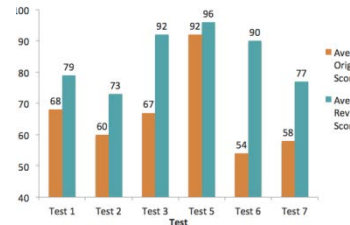
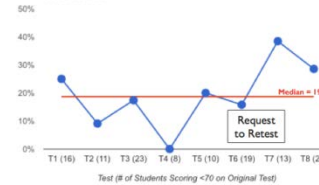
"math person" or not is something about you that you really care
people are good at math and other people aren't.



Secondary

Systematic collection Dec 2016

COO. Percent of Students Scoring Less Than 70
Who Revise



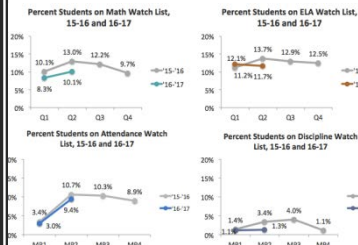
Change Ideas

Request to
Retest Form

Etc.

By May 2017,
halve the
proportion
(from 1 in 5 to 1
in 10) of
students who

Started tracking early 2016



Instructional
Practices &
Mindsets

Student
Behavioral
Supports

Others...

Quality of First
Attempt

In the Future?



Carnegie Foundation
for the Advancement of Teaching

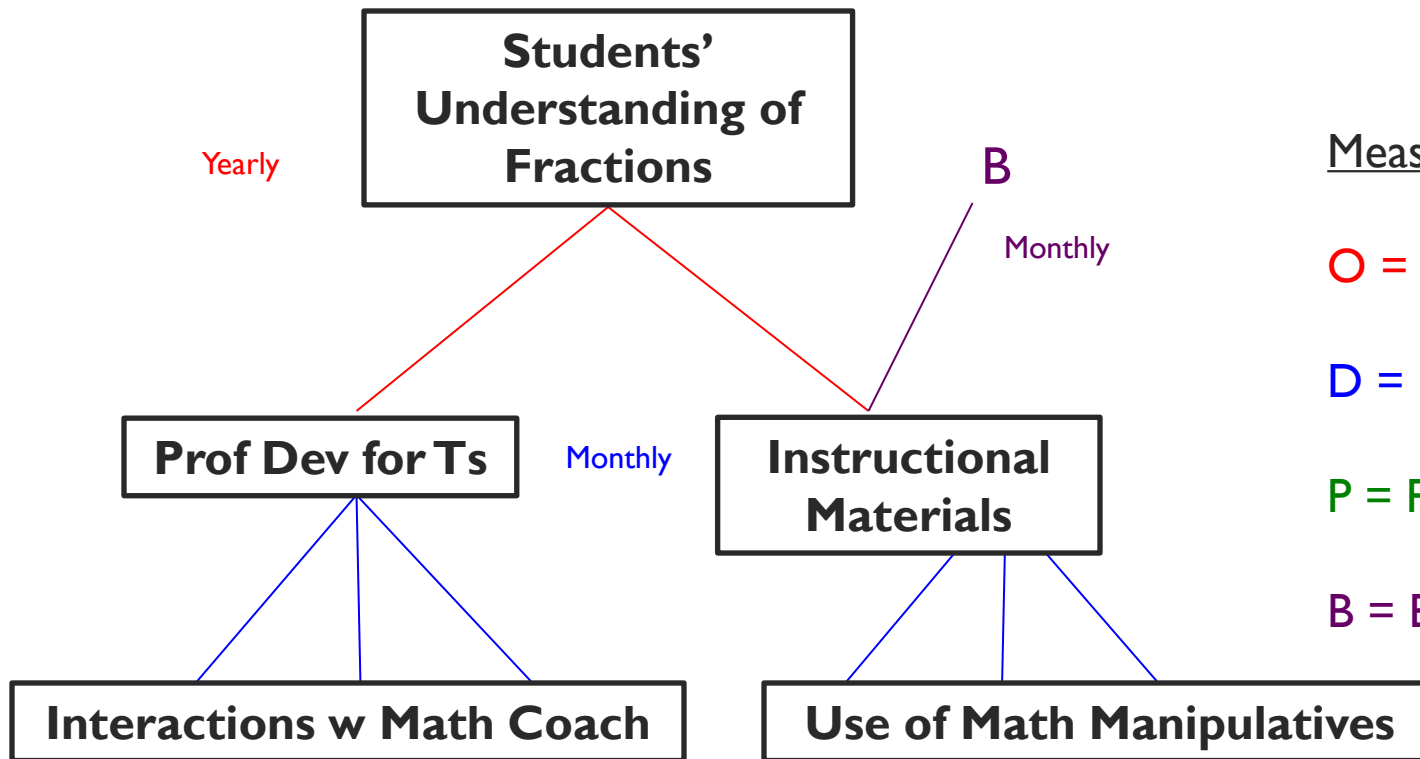
Activity

Activity:

Developing a Set of Measures

- **Scenario:** You are part of an improvement team at an elementary school. The students at your school are struggling with fractions. You launch an improvement project with your school team to tackle this problem.
- **Goal:** The goal of the project is to increase students' understanding of fractions.
- **Your Task:** Develop a system of measures for this project:
 - Outcome Measures: 1-2 measures
 - Process or Driver Measures: 2-3 measures
 - Balancing Measures: 1 measure

FRACTIONS MEASUREMENT SYSTEM



Measure Types

O = Outcome Measures

D = Driver Measures

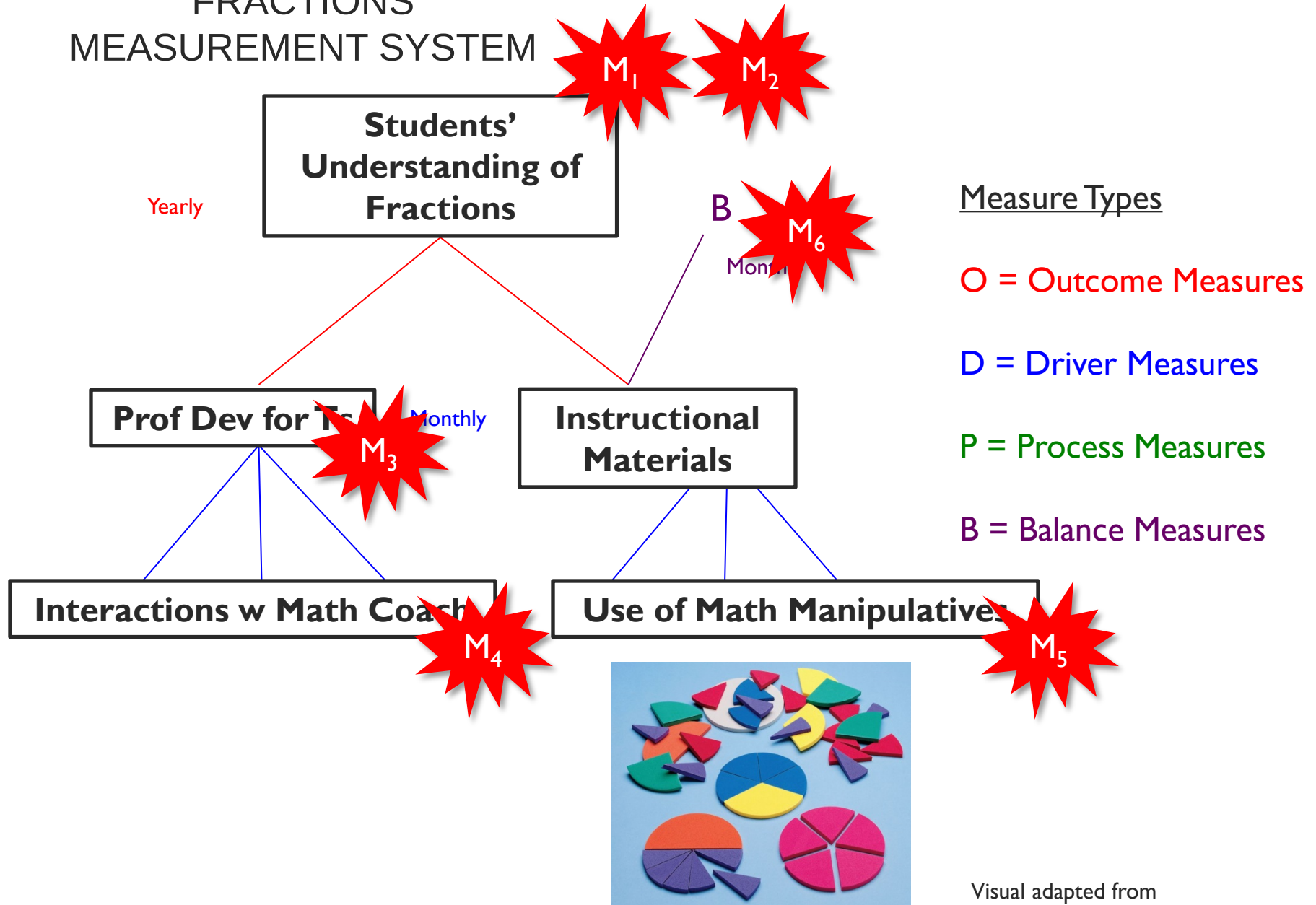
P = Process Measures

B = Balance Measures



Visual adapted from
Improvement Science Consulting

FRACTIONS MEASUREMENT SYSTEM



Visual adapted from
Improvement Science Consulting

Key Takeaways

- Data help us understand our systems and whether changes we make lead to improvement
- Measurement for improvement is different than measurement for accountability and research
- A family or system of measures, which includes outcome, driver, process, and balancing measures is needed to guide improvement efforts



Carnegie Foundation for the Advancement of Teaching

© 2017 Carnegie Foundation for the Advancement of Teaching

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License:

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

