

CARNEGIE FOUNDATION SUMMIT
on
IMPROVEMENT IN EDUCATION 2020

Measurement for Improvement

Jon Norman, Associate, Managing Director, Evidence and Analytics, Carnegie Foundation

Dave Sherer, Associate, Evidence and Analytics, Carnegie Foundation

This session was
created for
on-demand viewing.

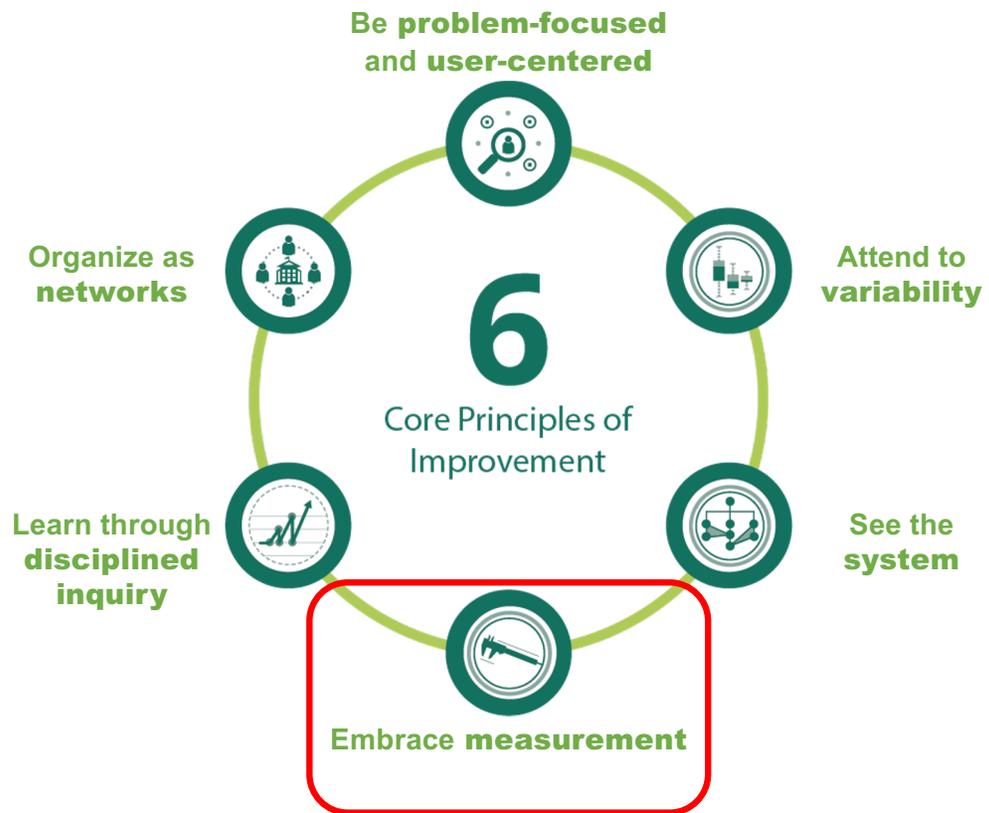


#CarnegieSummit2020

@CarnegieFdn

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.



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.**”

Leif I. Solberg, MD



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

Lief Solberg, Gordon Mosser and Sharon McDonald *Journal on Quality Improvement* vol. 23, no. 3, (March 1997), 135-147.

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

Lief Solberg, Gordon Mosser and Sharon McDonald *Journal on Quality Improvement* vol. 23, no. 3, (March 1997), 135-147.

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

Lief Solberg, Gordon Mosser and Sharon McDonald *Journal on Quality Improvement* vol. 23, no. 3, (March 1997), 135-147.

ASPECT	ACCOUNTABILITY	RESEARCH	IMPROVEMENT
<u>Why?</u>	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual	Develop and evaluate changes in practice
<u>What?</u>	End of the line outcomes		processes that are the of change
<u>How often?</u>	Usually collected once a year (after the fact)	Type	as practice occurs
<u>Testing your theory</u>	No theory to test		ential tests
<u>Sample size</u>	Obtain 100% of available, relevant data		data, small sequential samples
<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.	Data shared in a low-stakes, safe environment conducive to change.

Key limitation for improvement:
Does not illuminate WHY the outcomes occur or what should be done to change them

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

Key limitation for improvement:

IMPRACTICAL to administer; not designed to inform changes in practice

ASPECT	ACCOUNTABILITY	RESEARCH	IMPROVEMENT
<u>Why?</u>	Identify exemplary or problematic performers (teachers, schools, districts)	Develop and test theories about the relationships between conceptual variables	Develop and evaluate changes in practice
<u>What?</u>	End	ables	Outcomes and processes that are the object of change
<u>How often?</u>	Usually	vice per study (fact)	Frequently as practice occurs
<u>Testing your theory</u>		test	Sequential tests
<u>Sample size</u>	Obta relevant data	" data	"Just enough" data, small sequential samples
<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.	Data shared in a low-stakes, safe environment conducive to change.

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

Reflection Questions

- What types of measures have you encountered most in your career?
- Have you ever had the experience of accountability or research measures being used for improvement purposes?
- What are some examples from your context of measures that can support improvement?



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



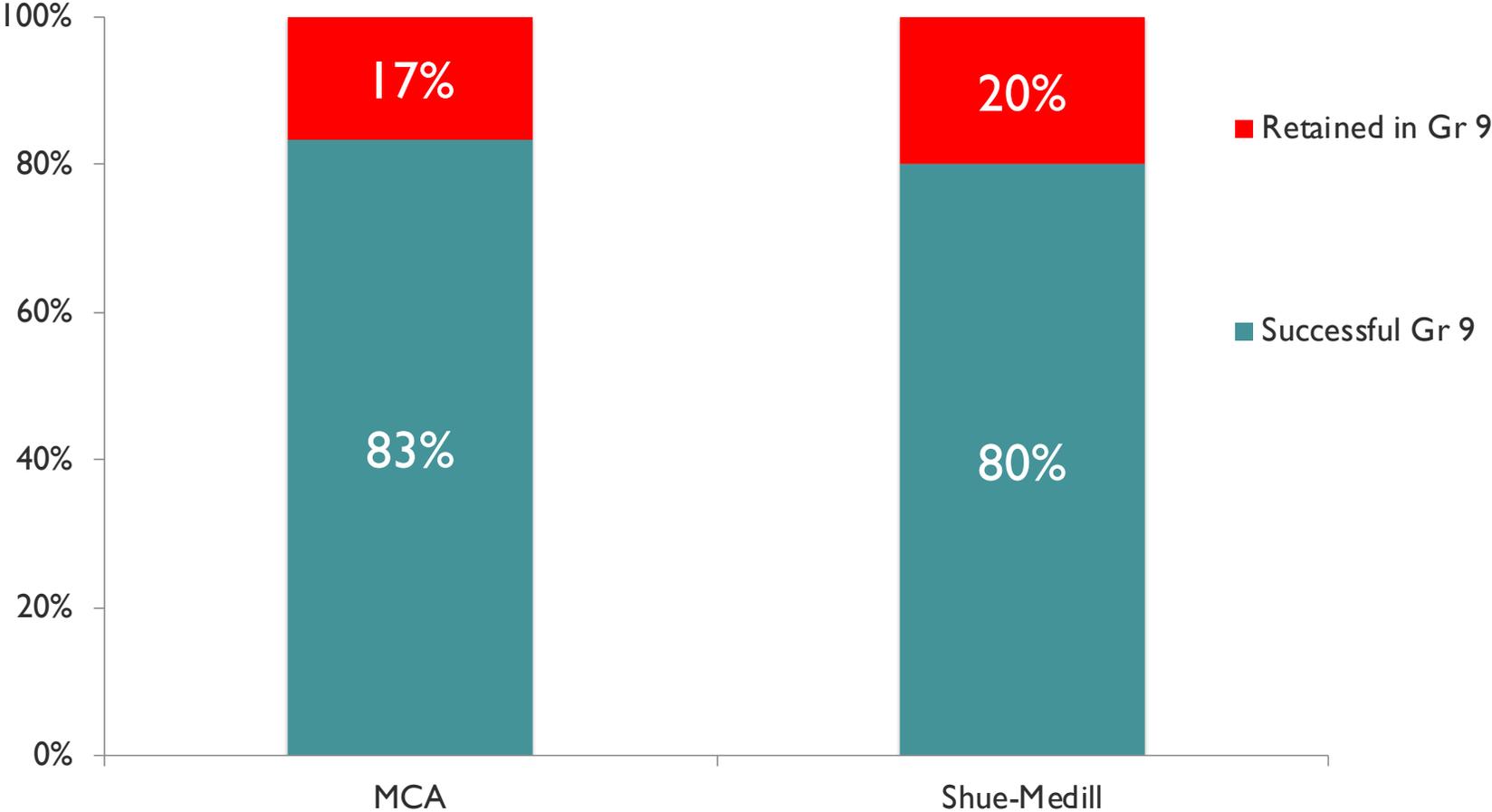
Michele Savage



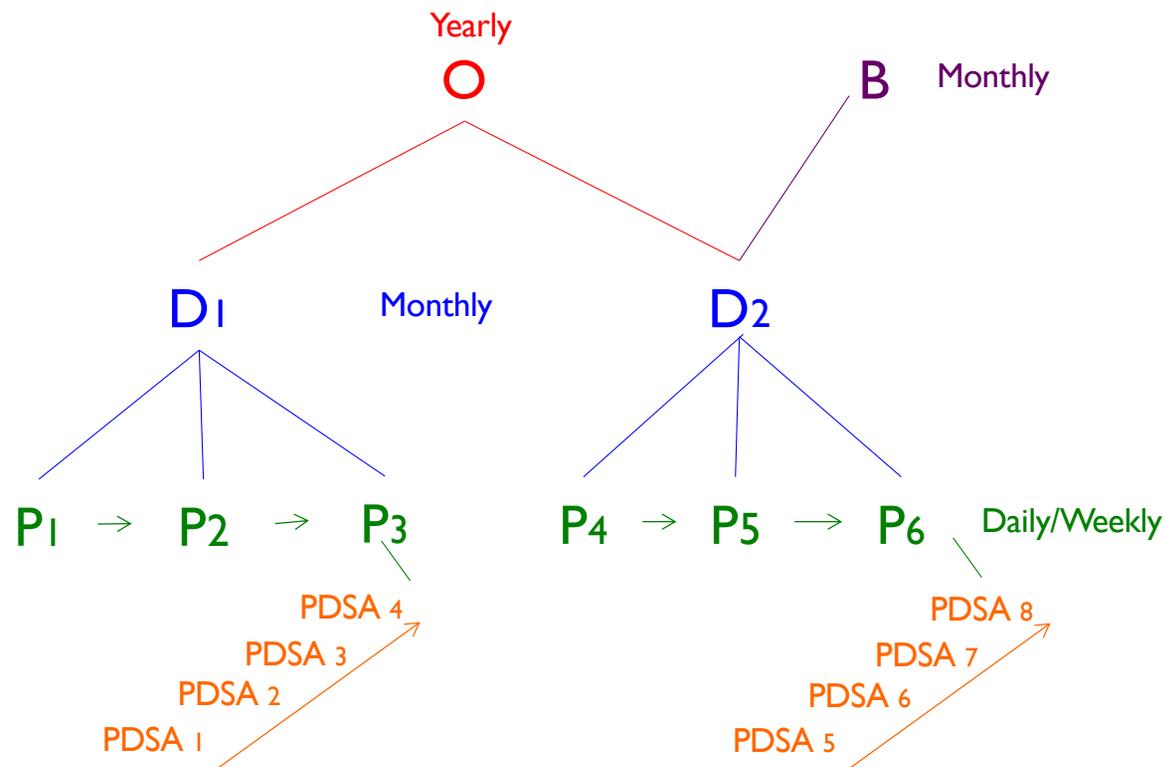
Nancy Carnvale



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

○ = Outcome Measures

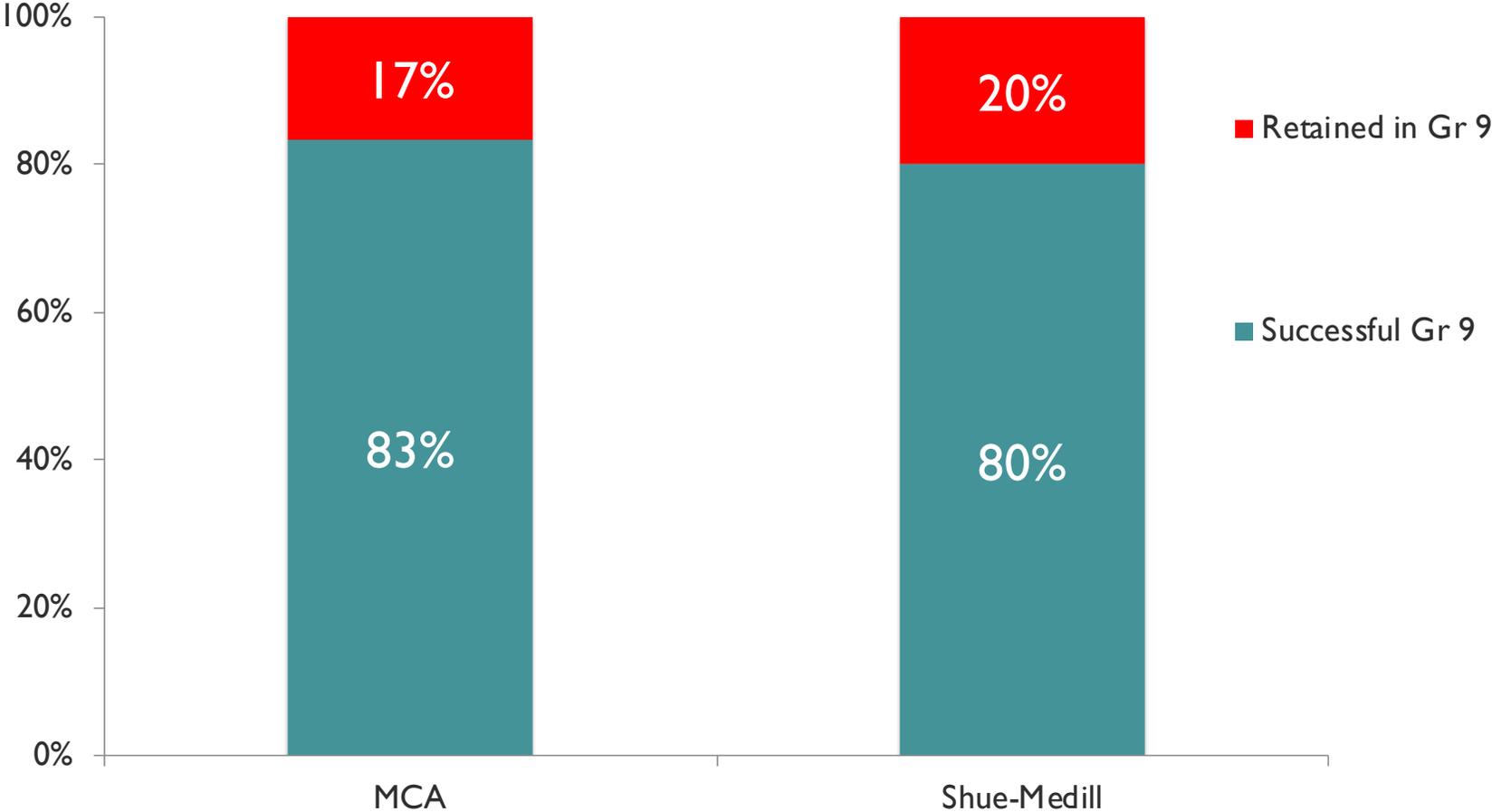
Can you think of an outcome measure in your context?

Visual adapted from
Improvement Science Consulting

© 2018 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/>



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

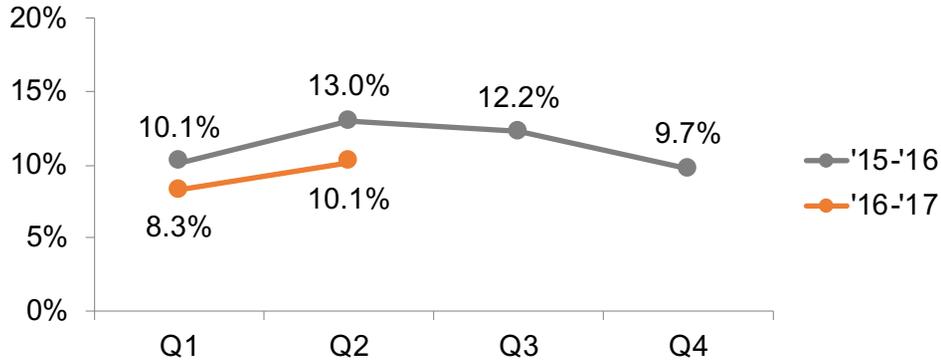
O (Leading): % of Students on Watch Lists

Measure Types

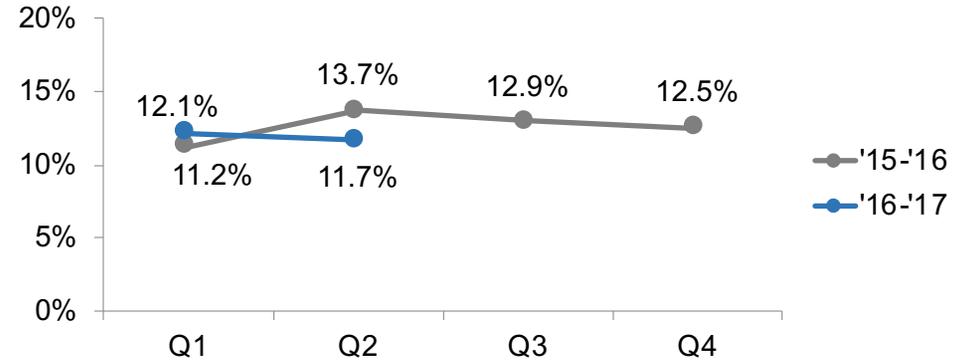
O = Outcome Measures

Leading Outcome Measures: 4 Early Warning Indicators

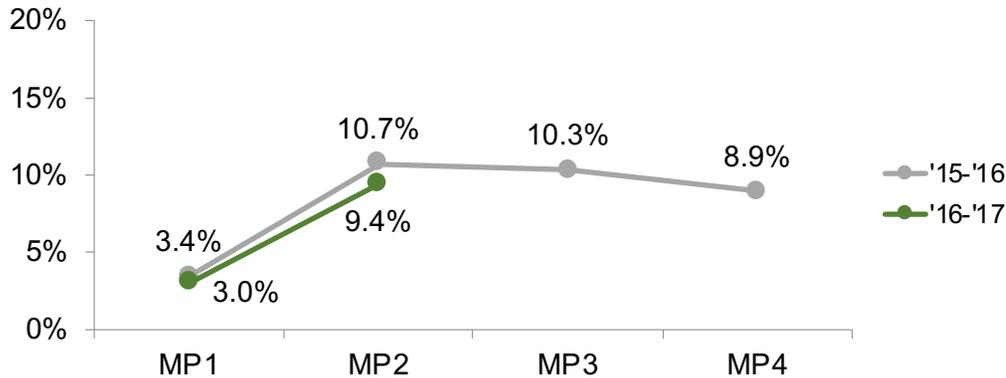
Math Watch List, 15-16 and 16-17



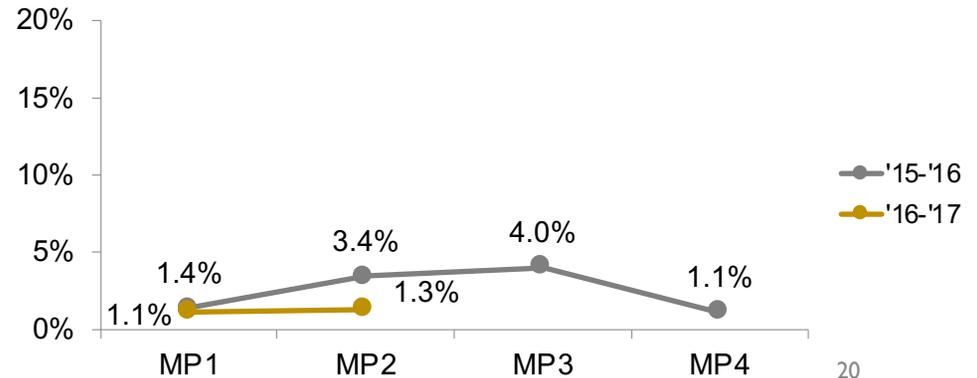
ELA Watch List, 15-16 and 16-17



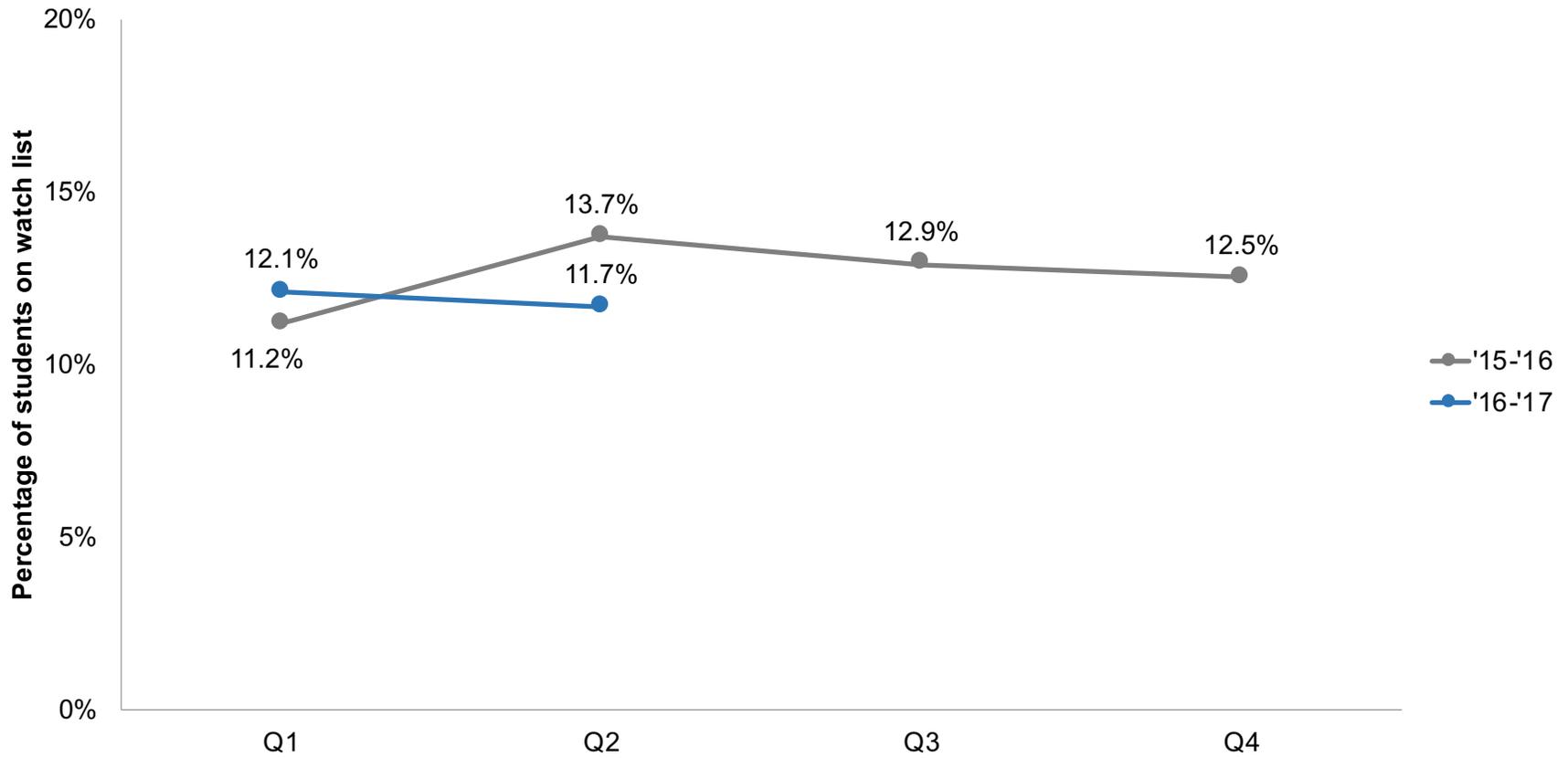
Attendance Watch List, 15-16 and 16-17



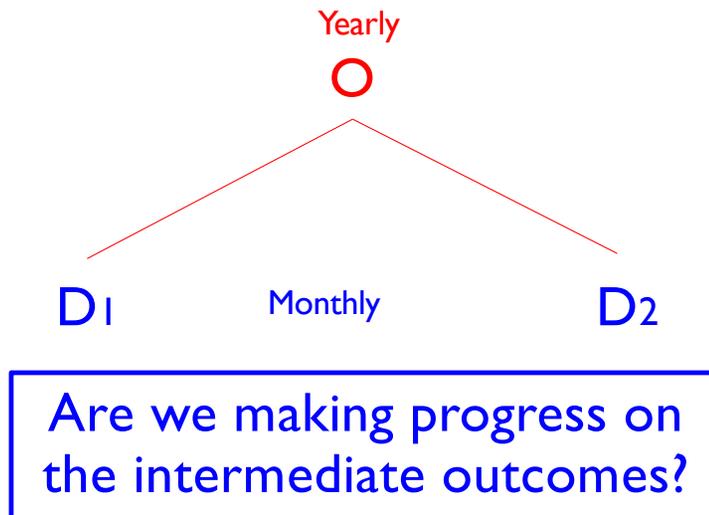
Discipline Watch List, 15-16 and 16-17



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



System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

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

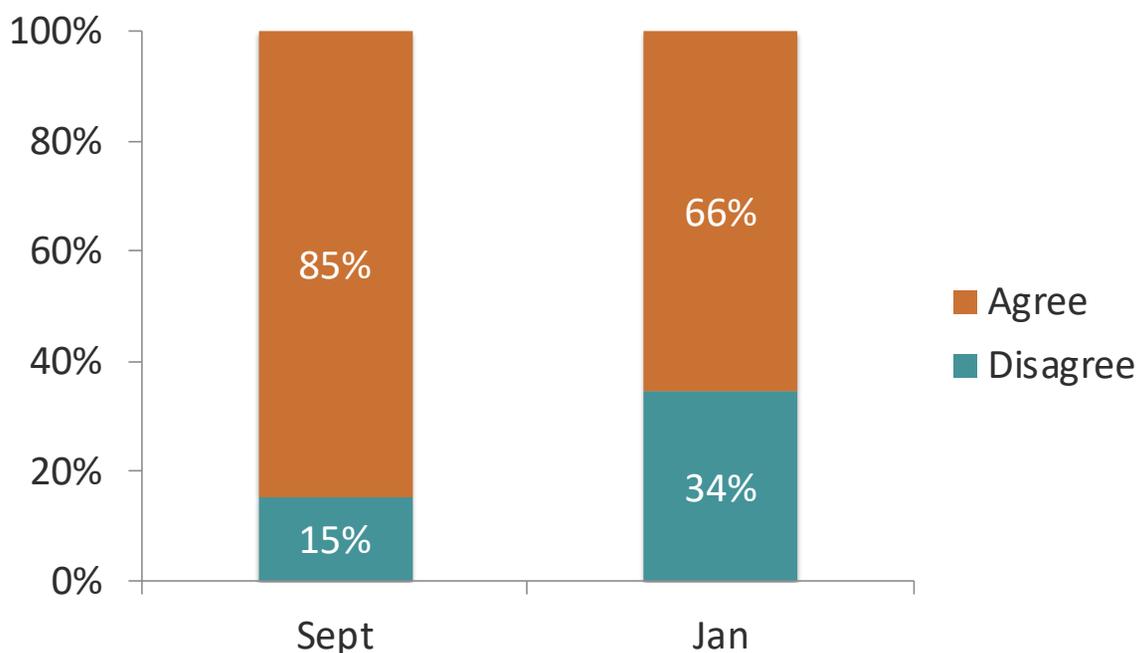
Measure Types

○ = Outcome Measures

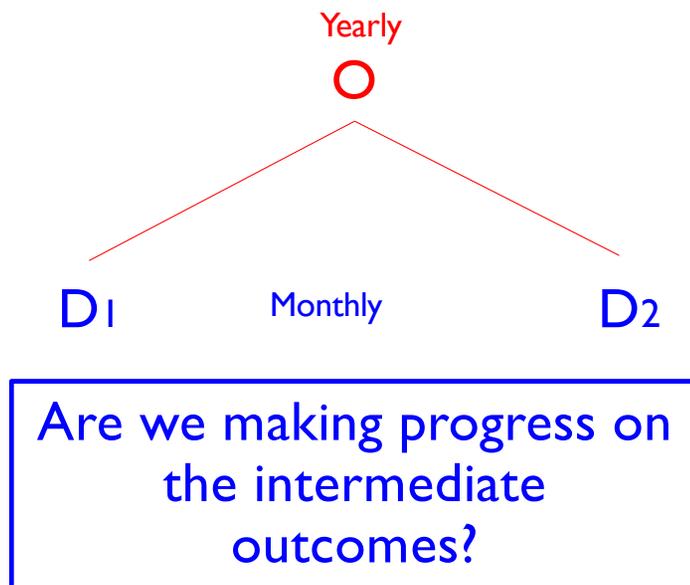
D = Driver Measures

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



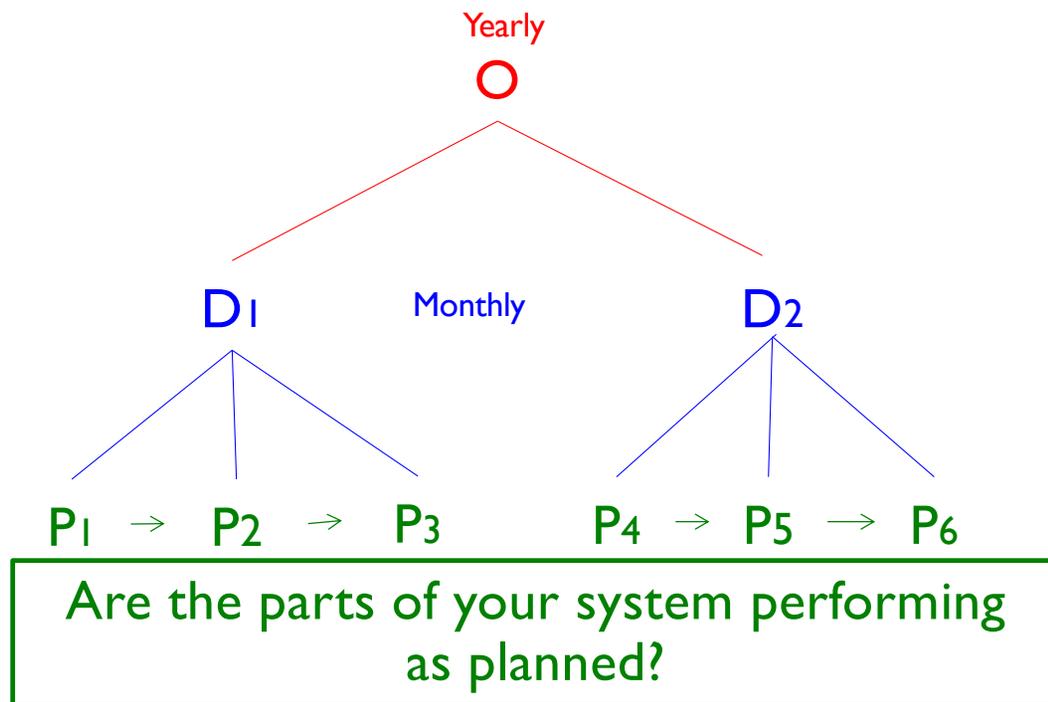
Measure Types

O = Outcome Measures

D = Driver Measures

Can you think of a driver measure in your context?

System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

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

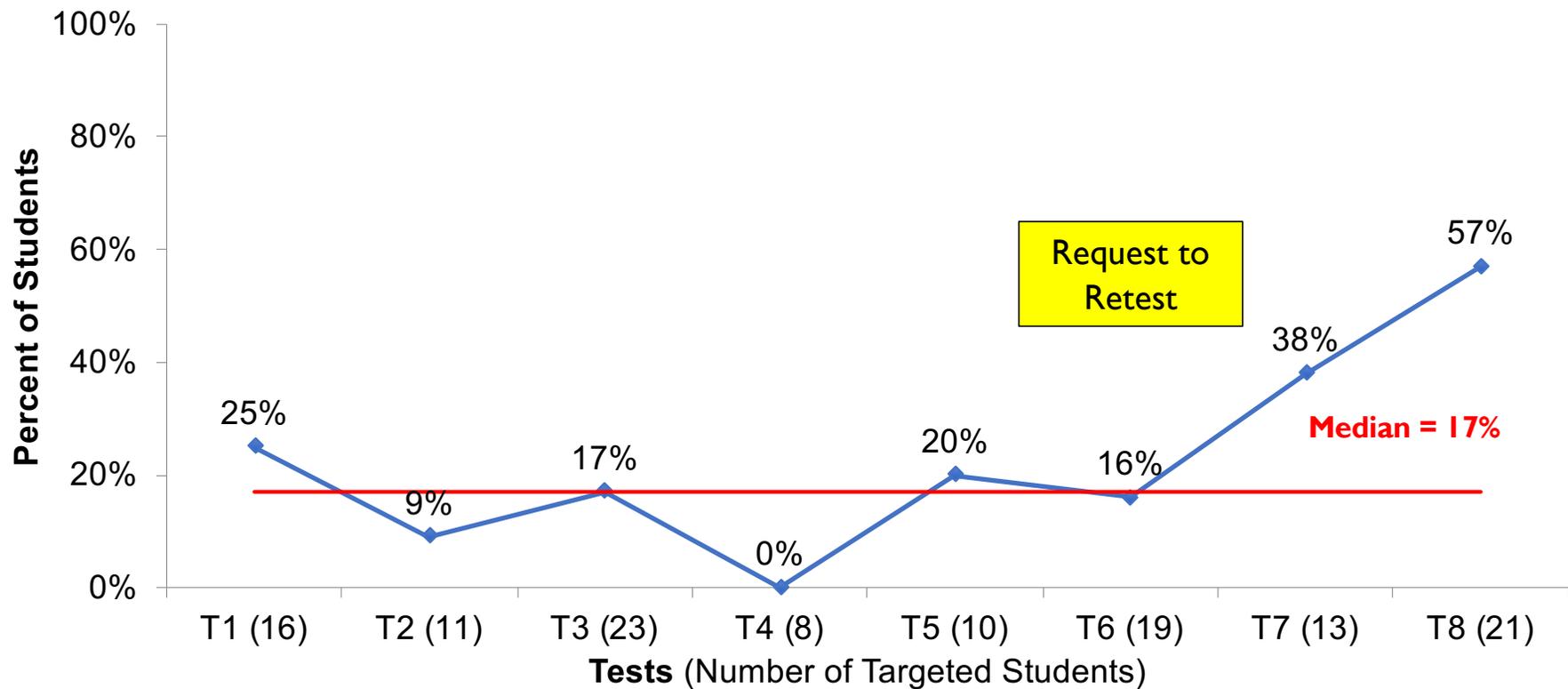
Measure Types

○ = Outcome Measures

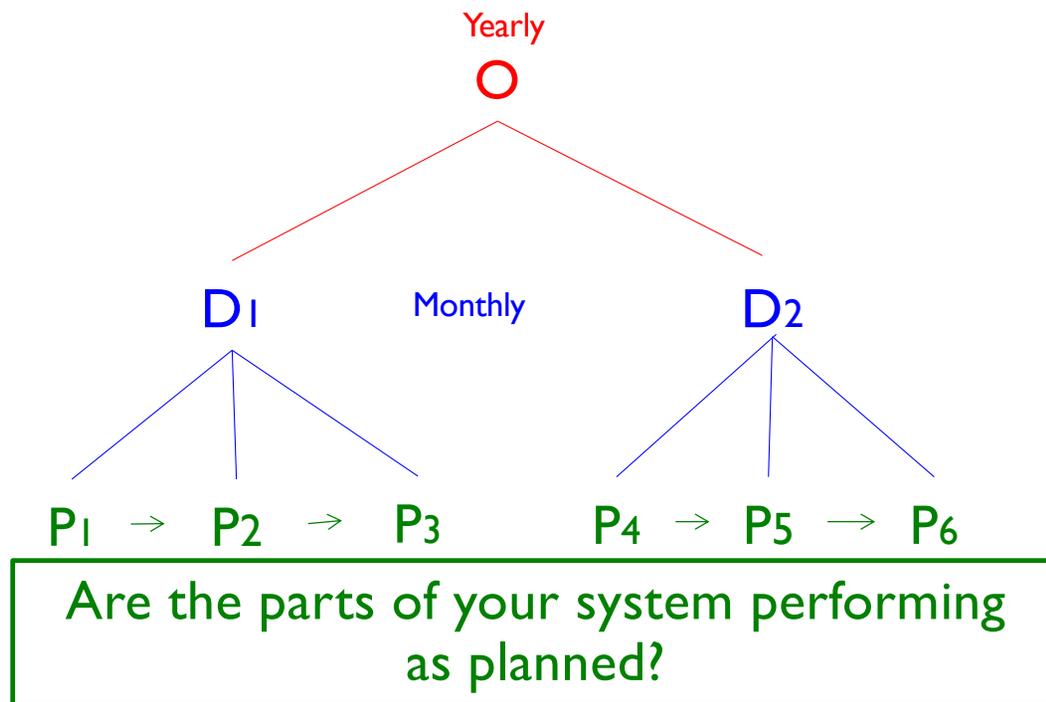
D = Driver Measures

P = Process Measures

**Process Measure:
Percent of Targeted Students Revising Tests,
Classroom A, Grade 6 Math**



System of Feedback



Measure Types

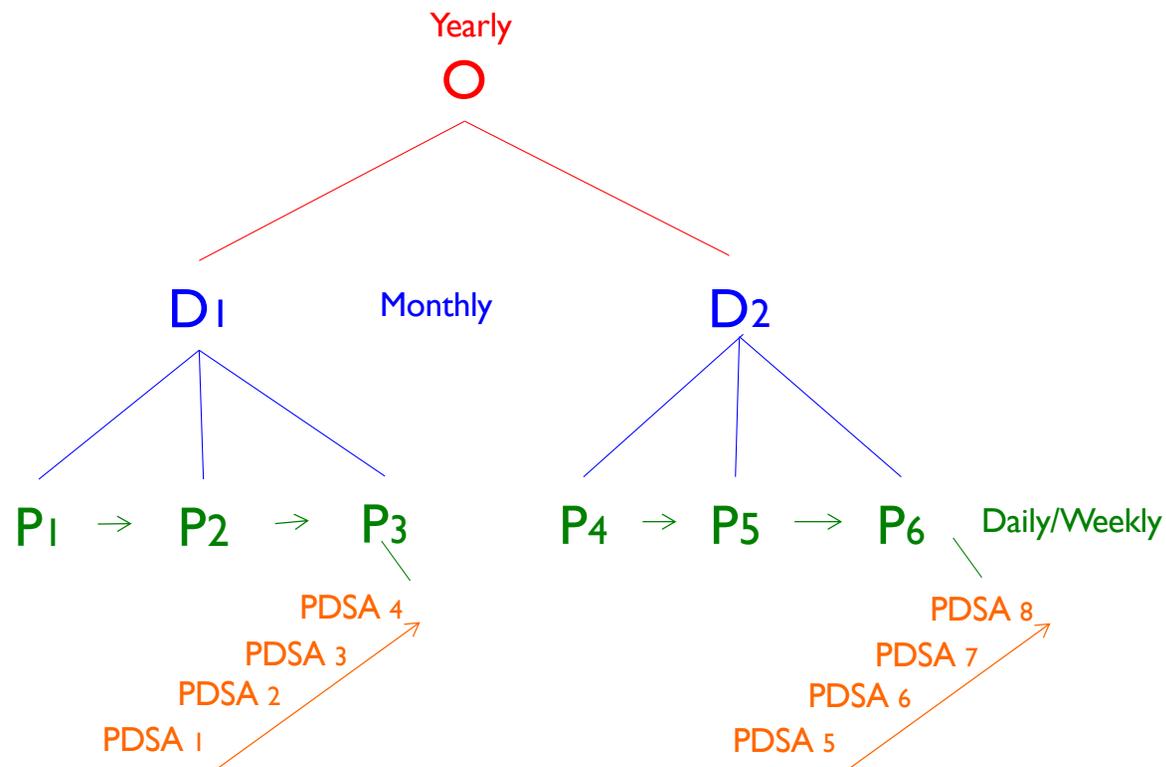
O = Outcome Measures

D = Driver Measures

P = Process Measures

Can you think of a process measure in your context?

System of Feedback



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

PDSA = Learning Cycle Measure

Adapted from
Improvement Science Consulting

© 2018 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/>



System of Feedback: STL-Delaware Case

Aim:

○ (Lagging): % of Graduates Repeating Grade 9

○ (Leading): % of Students on Watch Lists

Measure Types

○ = Outcome Measures

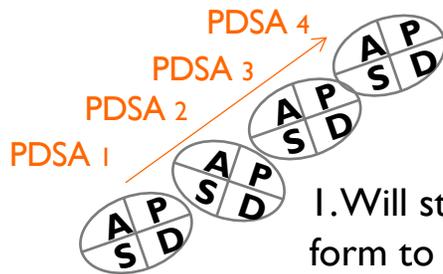
D = Driver Measures

D₁: Instructional Practices

D₂: Student Mindsets

Retesting Process

→ P₂ → P₃



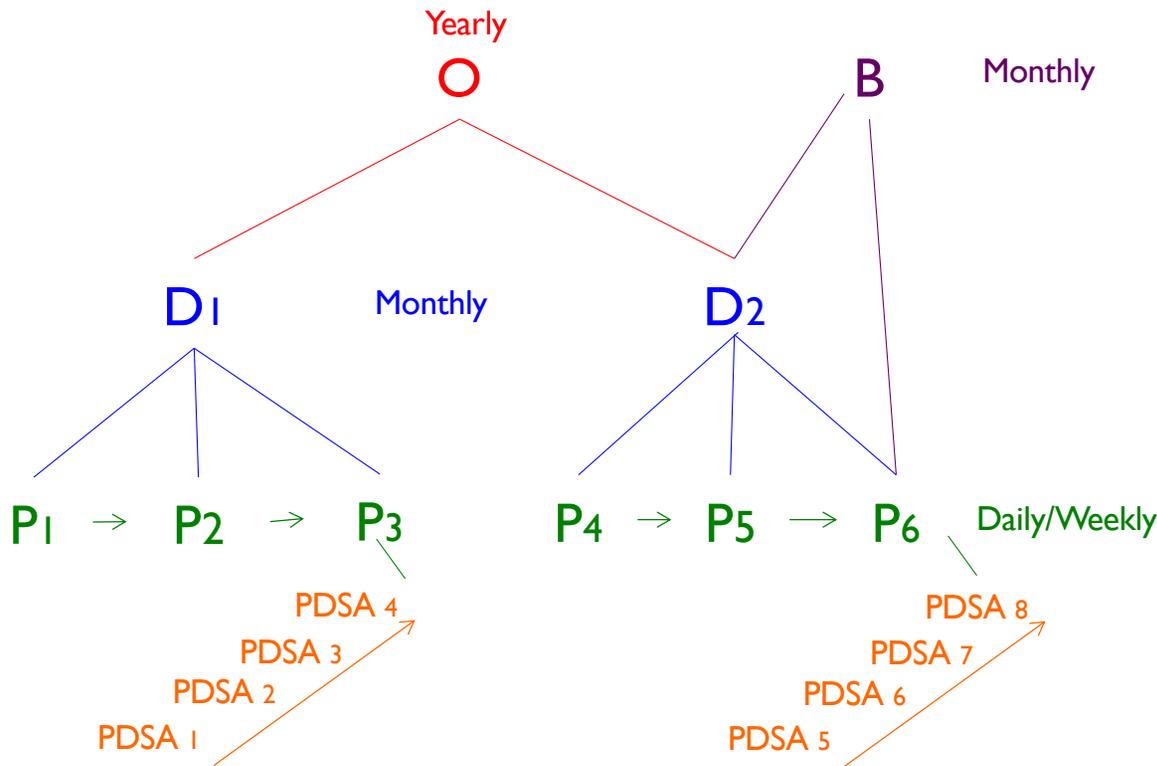
1. Will students use a form to make revisions?
2. How can I get more students to use form?

Forms picked up
Reasons why students don't use form

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

D₁: Instructional Practices

D₂: Student Mindsets

Retesting Process

→ P₂ → P₃

Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

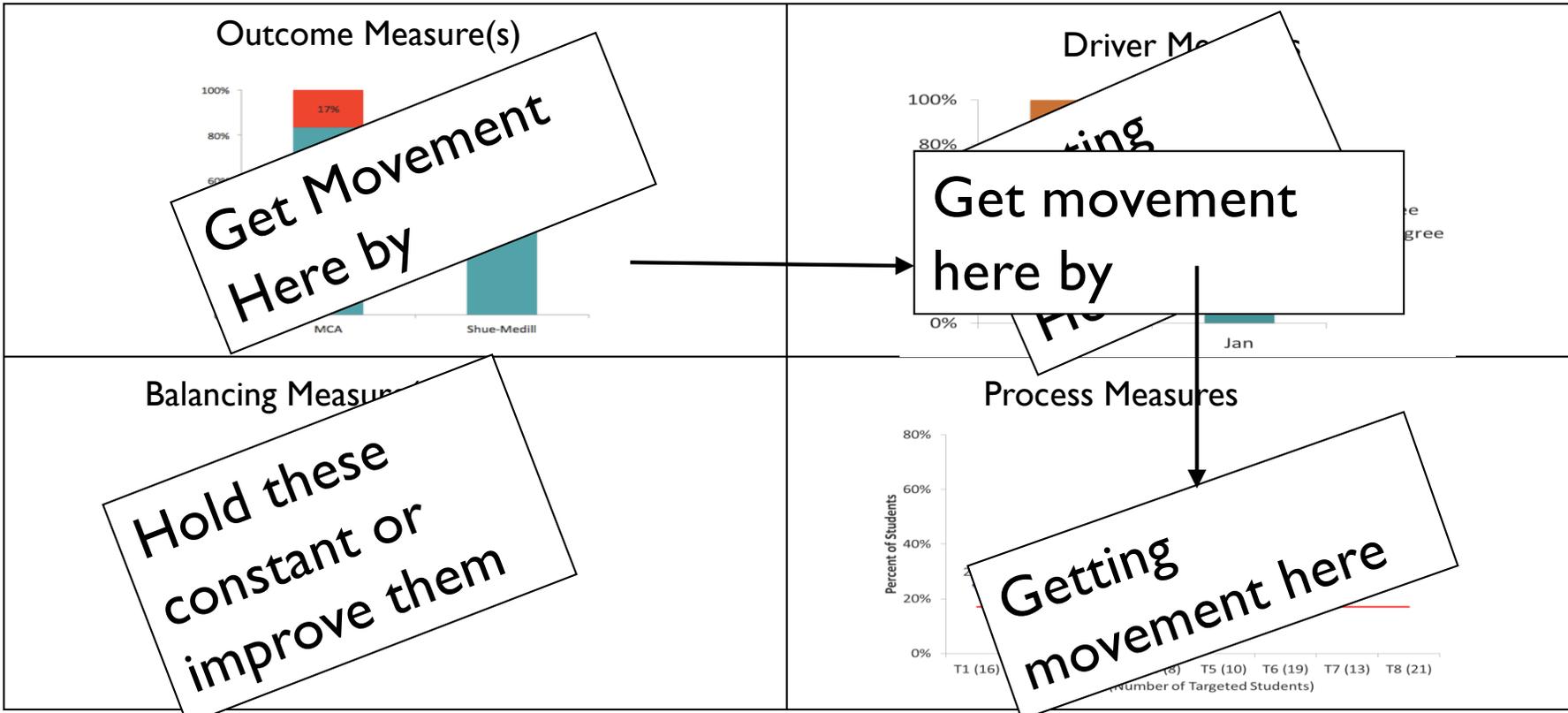
B = Balance Measures

PDSA = Learning Cycle Measure

Adapted from Schools That Lead

© 2018 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/>







Carnegie Foundation
for the Advancement of Teaching

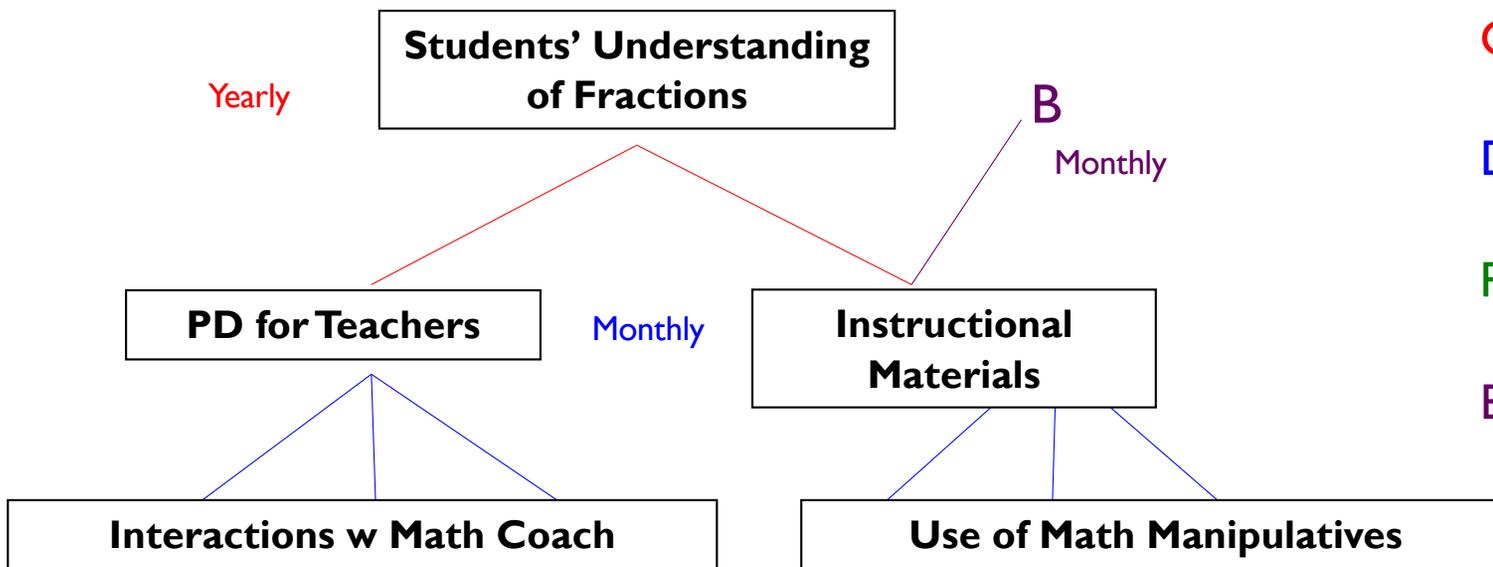
Activity

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

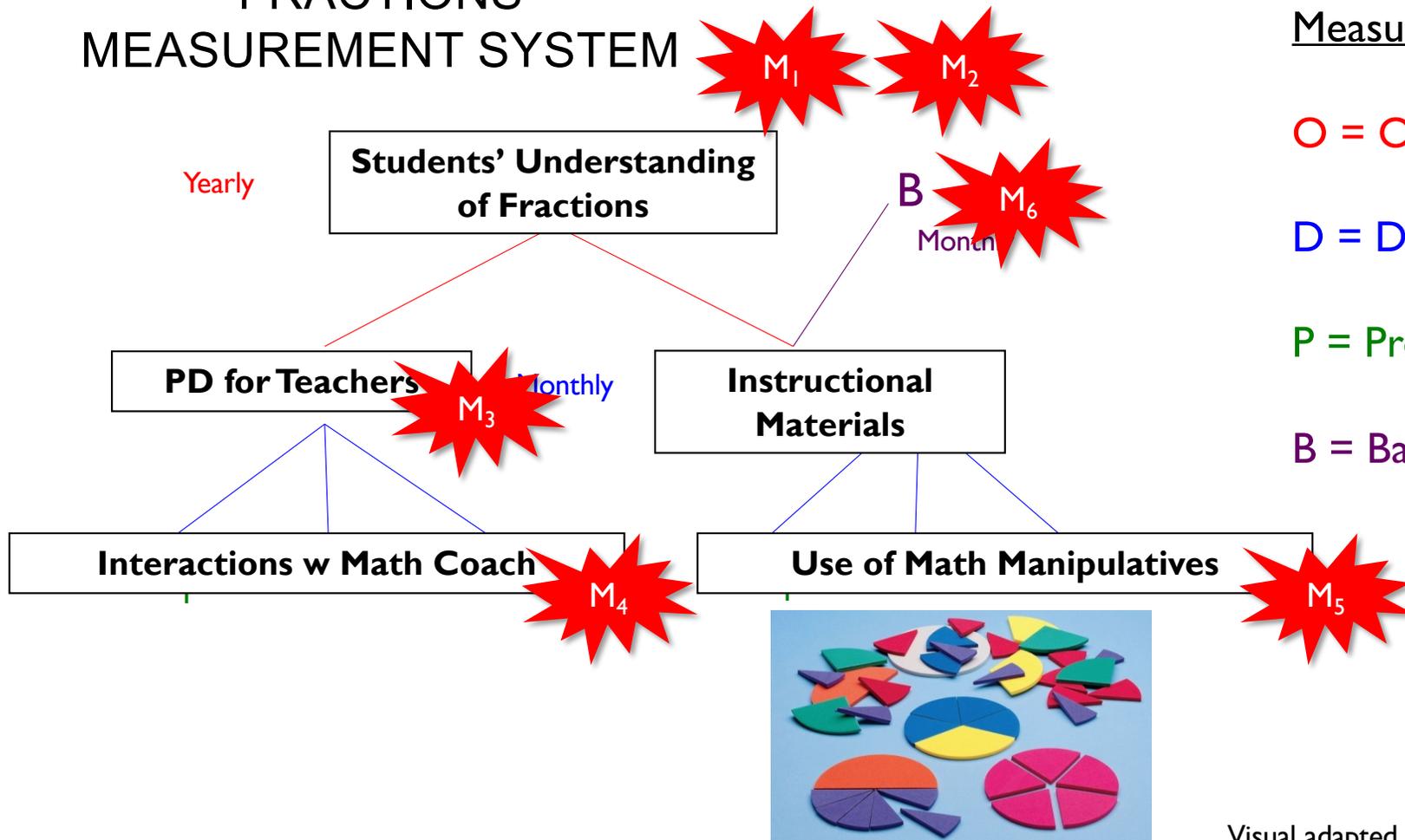
O = Outcome Measures

D = Driver Measures

P = Process Measures

B = Balance Measures

FRACTIONS MEASUREMENT SYSTEM



Measure Types

O = Outcome Measures

D = Driver Measures

P = Process Measures

B = Balance Measures

Yearly

Monthly

Monthly

Some ideas....

- **Outcome measures**
 - Students' understanding of fractions
 - School-created interim assessments (leading)
 - Yearly summative assessments (lagging)
- **Driver measures**
 - Professional development for teachers
 - Teacher survey (satisfaction with math PD including coaching)
 - Instructional materials
 - Teacher survey (effectiveness of math manipulatives)

More ideas...

- **Process measures**
 - Interactions with math coach
 - % of teachers that meet with a math coach each week
 - Use of math manipulatives
 - % of teachers that teach fraction lesson using recommended manipulatives each week
- **Balancing measures**
 - Parent satisfaction with math program
 - Student Performance in ELA

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 SUMMIT **2020**
on
IMPROVEMENT IN EDUCATION
SAN FRANCISCO

Help us improve!

To submit your session evaluation:

1. Navigate to the specific **session page** in the Summit mobile app or online web browser version
2. **Scroll to the bottom of the page** to find the link to the **session evaluation survey**



Carnegie Foundation
for the Advancement of Teaching

© 2019 Carnegie Foundation for the Advancement of Teaching.
All rights reserved.

CARNEGIE FOUNDATION SUMMIT
on
IMPROVEMENT IN EDUCATION 2020
SAN FRANCISCO

Help us improve!

To submit your session evaluation:

1. Navigate to the specific session page in the Summit mobile app or online web browser version
2. Scroll to the bottom of the page to find the link to the session evaluation survey

