# Teams Using Data Wisely

This module addresses the many issues related to the collection, analysis and effective use of data by educators as they work together to enhance learning for students. The creation of meaningful relationships, a spirit of inquiry, and a sense of collective efficacy are important foundational components of effective teams.

The following topics are discussed in the module text, linked resources and embedded videos from Ohio schools:

- Support for the effective use of data
- Cautions regarding the quantitative aspects of data
- Cautions regarding the qualitative aspects of data
- Teacher-based teams and the use of the 5-step process
- Decision making for teacher-based teams
- School-wide data and the building leadership team
- BLTs supporting collective efficacy and data use by TBTs
- Strategic data issues of the district leadership team
- DLT support for the evidence-based needs of buildings

The content in this module is intended to align closely with the content of the OLAC modules on Assessment and Creating Cultures Grounded in Data.

This module aligns with Ohio's Leadership Development Framework in the following areas:

- Area 1 Data and the Decision-Making Process
- Area 2 Focused Goal Setting Process
- Area 3 Instruction and the Learning Process
- Area 4 Resource Management Process

# Receiving Credit for this Module

The Ohio Leadership Advisory Council offers educators credit and contact hours for OLAC work. Teachers, principals, and superintendents who are working toward license renewal can receive university credit for completing OLAC modules from a number of Ohio universities. Pre-approval is required. For estimated contact hours for credit or to learn more about receiving credit for OLAC work, visit the Credit Corner.

# Table of Contents

# Introduction

Ohio schools and districts have a nested structure of collaborative leadership teams: from Teacher-based Teams (TBTs) to Building Leadership Teams (BLTs) to District Leadership Teams (DLTs). The teams lead a long-term mission to improve student learning statewide: with every educator, every school, and every district in the state. In these teams, *wise data use is a collaborative activity.*

Ohio's leadership teams rise to address this mission by using data wisely. That's what this module is about. The mission to improve instruction cannot end; it cannot be accomplished and forgotten. The *mission is continuous.* Why? Society changes. What educators need to teach changes. Indeed, everything that we call "education" is in flux. Leadership teams make sense of the flux i*n view of data about what is happening locally.*

No one else can do this work. It's up to us. This module is one support for the work.

# Teams Using Data Wisely

Education works because of important relationships-the relationships between students and teachers, the relationships among students, and the relationship between the people in the system and the meaningful ideas and skills on which the school hopes to focus everyone's attention. Education doesn't work using a factory model. Teachers are not machines, and students are not widgets.

Working with students requires care, attention, and lots of thoughtful improvisation. It's exhausting. Anyone who has taught knows it. But faculties can pull together to share the work and to consider the presenting issues. Together they will exercise better care and more careful attention. That's what *teams using data wisely* means in action. It's high-minded and it's hard work. But it's not difficult to grasp. Doing it well ("using data wisely") requires support, and that's where this module comes in.

This module drills down to issues of handling data in Ohio's collaborative learning teams. It will help teams and team members:

- take ownership of data and decision making based on data;
- ask and answer good questions with data;
- use *their* data to draw reasonable inferences about *their* students; and
- use *their* inferences to make wise decisions about *their* students.

# Why Teams Matter

Wise data use notably involves collaboration. This is a big change from the past, when teachers mostly acted solo (see, e.g., Hargreaves & Fullan, 2012).

The ultimate goal of this system of collaboration is the improvement of teaching and learning. Teams exist, in fact, so that educators (who are mostly teachers) can collaborate on the development and improvement of teaching. Sure, but what about "raising test scores"?

Neither learning nor teaching is primarily *about* testing. The act of learning is its own enterprise. An even more sobering thought: most learning does not happen in schools and is really tested only by life!

In schools, though, the learning that people do is supported by *teaching*. Schools are thus a unique and potentially useful part of life. The development and improvement of teaching realizes that potential. If we're not developing and improving teaching, that

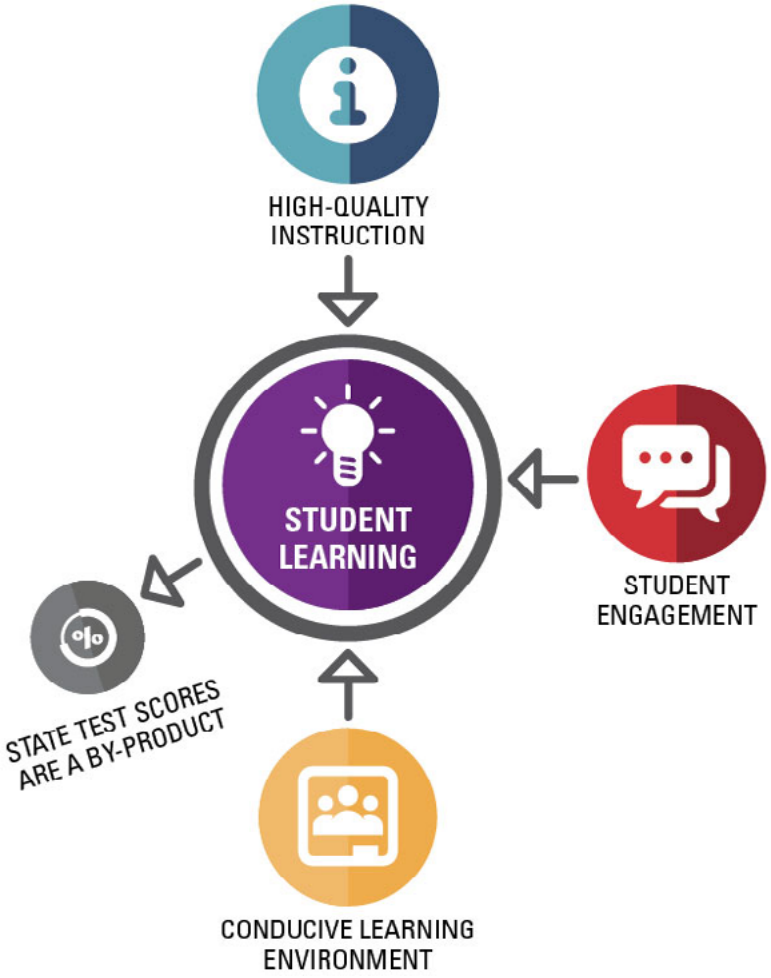
potential isn't realized. To examine this idea in a bit more depth explore the Foundational Concept, *Understanding Teacher Influence and Instructional Improvement*.

## Foundational Concept: Understanding Teacher Influence and Instructional Improvement

Go to the Appendix for more information about understanding teacher influence and instructional improvement.

By contrast, genuine learning is not supported by the practices often used to "raise test scores:" tweaking the way tests are given and putting students through lots of test preparation routines (Hursh, 2008). In fact, those practices often undermine learning, whereas better instruction facilitates learning. Teachers understand this reality very, very well (Cameron, 2015; Johnson, 2005). So do teams that use data wisely! Testing is part of this wise use, and so is using test scores: but test scores may not even be the most important part.

The figure below illustrates the centrality of learning to the whole process. It's a simplification, of course. What it says, though, is that student learning results when students are engaged in a process involving high-quality instruction in an environment that is conducive to learning. When students learn well, acceptable test scores result. The diagram suggests the relative importance of the different data that help educators determine whether or not the process is working as it should. Data about student learning, the quality of instruction, the engagement of students, and the conduciveness of the learning environment all are more important to understanding and improving learning than data about state test scores.




The diagram has five circles and four arrows. Student learning is in the center in the largest of the circles. Three equal-sized circles have arrows pointing to the circle in the center to indicate that they impact student learning. The words in those three circles are “high-quality instruction,” “student engagement,” and “conducive learning environment,” respectively. The smallest circle has an arrow leading to it from the large circle labeled, “student learning.” This shows that state test scores are a by-product of student learning.

# What Teams Do: The Big Picture

Another answer to “how to raise test scores,” one that this module addresses, operates on a somewhat different plan. That answer goes something like this:

> As a team of educators, we can use data to help us make better decisions about how to help students learn.

It doesn’t sound like much, but it’s actually critical. Notice too that the emphasis isn’t really on data! The emphasis is on what members of TBTs, BLTs, and DLTs say to one another and the actions they take as a team as they consider “the data.” That’s the hard part, not the data.

So, *teams using data wisely* refers to changing professional practice carefully to better support the learning and the school experience of all students.

The table on the next page provides some examples.

| What Team? | What Data? | What Discussion? | What Actions? |
|---|---|---|---|
| 3rd grade TBT composed of four general education teachers and one intervention specialist | Rubric scores from student essays along with examples of high-scoring, middle-scoring, and low-scoring essays | Are all five teachers interpreting the rubric criteria in the same way and achieving similar ratings of essays? | When the next essay is assigned, the team will arrange for two teachers to read each essay and talk about their ratings<br><br>Everyone on the team will share with their students three examples of essays with high ratings and explain how those essays met the rubric criteria. |
| Middle school BLT | Grade-by-grade comparisons of teachers' use of questions eliciting responses at differing depths of knowledge | How can questions at all grade levels be framed to elicit strategic thinking and extended thinking?<br><br>What professional development might help teachers feel more comfortable using questions eliciting strategic and extended thinking? | Provide school-wide PD in which the 4th-grade TBT helps other teachers learn how to develop questions to elicit strategic and extended thinking.<br><br>Create a practice profile that literacy coaches can use to help teachers school-wide develop questions that elicit strategic and extended thinking. |

| What Team? | What Data? | What Discussion? | What Actions? |
|---|---|---|---|
| Suburban district DLT | School-by-school discipline referrals by race | Why are discipline referrals so much higher at some schools than at others? | Over a one-year period implement PBIS at the high school.<br><br>Add to the BLT agenda at Downtown Elementary School a discussion of the overuse of discipline referrals. |

# Teams Create Synergies

Work in teams to make decisions based on data is not just one more thing that beleaguered teachers and administrators *have to do*. This work is relevant to the most important things teachers and administrators are already doing! Teams, however, enable *collective* thinking and *collective* action. The whole is more than the sum of the parts. Collaborating to organize data and make instructional judgments improves both individual practice and the work of the whole organization.

## A Deeper Dive: Related OLAC Modules

For the wider context, review two OLAC modules:

- Creating Cultures Grounded in Data
- Assessment

The focus of Teams Using Data Wisely is specifically on what Ohio leadership teams can do with data to improve teaching and learning.

# Important Definitions

- “Professional practice” refers to anything that educators do as part of their work: providing instruction, sharing leadership, cultivating community relationships, developing or refining curriculum, and arranging pupil services.
- “Teams” means DLTs, BLTs, and TBTs.
- “Data” means systematically gathered and systematically analyzed qualitative information (in words or pictures) and quantitative information (in numbers).
- “Wise” means judging well the sorts of data needed for a particular use and making good inferences and decisions from the data.

This module considers wise use of data by teams. First it looks at key data issues for teams in general. Then it considers key data issues for TBTs, key data issues for BLTs, and key data issues for DLTs.

# Key Data Issues for Teams in General

## Owning the Data

The big issue for teams in general is taking ownership of data. Ownership looks like this: *"We want to use these data to make the right next decision."* The alternative is to punish others with data or to be punished by data: *"These data show that your school gets a grade of 'F'."* For more examples, see the table below.

| Owning the Data | Being Punished by the Data |
|---|---|
| Let's look more closely at how we teach proportions. [TBT] | Get your 7th-grade math scores up! |
| Let's do an equity audit to get a clearer picture of the experience of our English learners. [BLT or DLT] | We got an "F" on gap-closing. We'll never get out of this hole. |
| Let's see how we help students connect the ideas of proportions, fractions, and decimals. [TBT] | Our 7th grade math proficiency increased by 2%!! Yay! |
| Let's reassess the way our reading program deals with comprehension. [DLT or BLT] | Our 3rd-grade reading scores declined. We need to do better next time with reading comprehension. |

Some teams in some districts are already on the path to owning their data, but it's not a well-trodden path (Mandinach & Gummer, 2016). In fact, most teams are just getting started when it comes to using data wisely. *This module is for those teams, especially.* And it's not an easy or short journey once they do get started.

## Owning What the Data Mean

The whole idea of data strikes many educators—and many other people too-as intimidating. Data are supposed to tell us something, but in fact, data rarely speak for themselves.

Team members must bring something to the data. But what?

Teams bring members' insights and ideas and judgment to the data-that's the crux of "data analysis" in this context! It's not a long list of specialized techniques known and loved only by researchers. This module will describe the few actual data-analysis techniques needed. Far more important is the ***quality of insights, ideas, and judgments that team members bring to the data.***

> *[One study] noted that it was not until the end of the third year of working together that teachers were able to understand and use effective data collection and analysis techniques in a way that supported the desired form of socio-constructed knowledge associated with collaborative inquiry.*
>
> (De Luca, Shulha, Luhanga, Christou, & Klinger, 2015, p. 655)

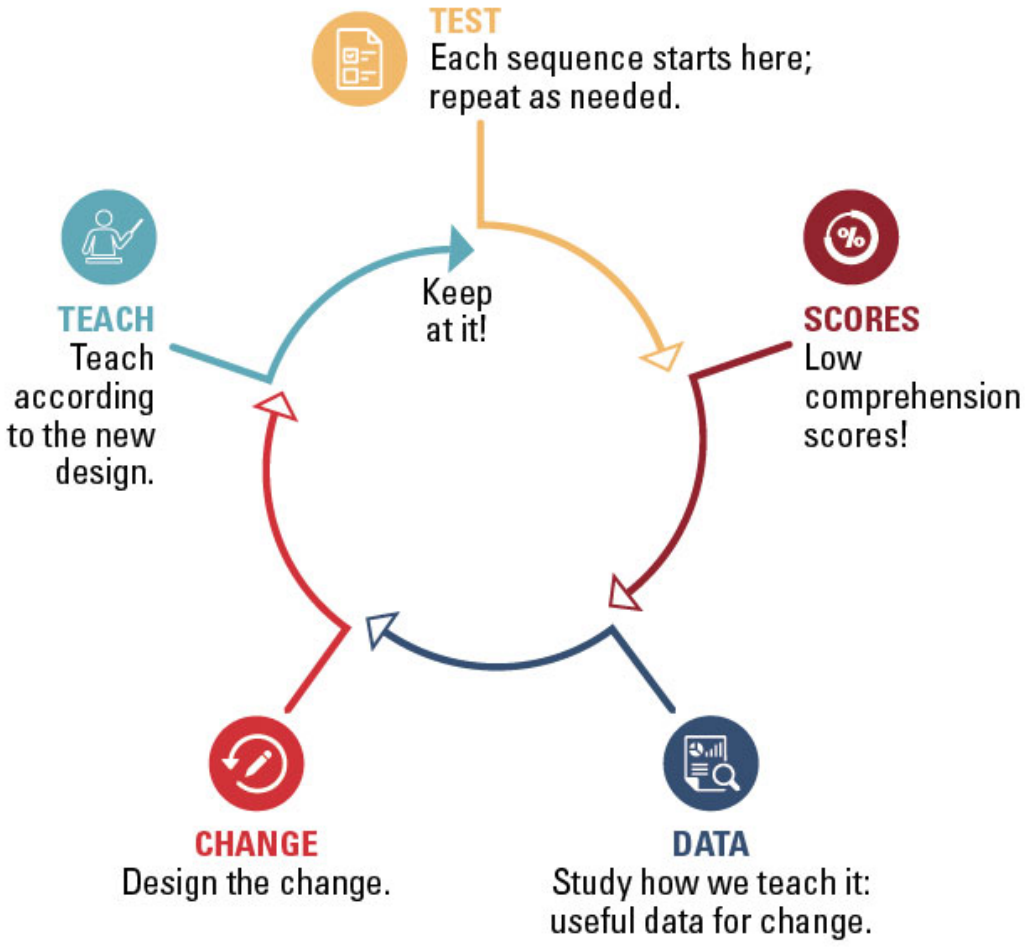
Here's an example. Let's say our math scores are low, probably because our students don't seem to understand the concept of proportional reasoning. We need to concentrate on teaching that better. Everyone resolves to do better!

Often the conversation ends there. That's unfortunate because figuring out how to teach (something) better to our students is where the real work with data comes in (Mandinach & Gummer, 2016). It's where the real work of teaching (and also learning) comes in. The perception of a problem might begin with low test scores, but that's not what changes teaching. And our intention to change is not what changes teaching either.

> *Data have no meaning. Meaning is imposed by interpretation.*
>
> (Love, Stiles, Mundry, & DiRanna, 2008, p. 5)

# An Example

The diagram below presents an example. It depicts a situation in which a team of educators (let's say a BLT) recognizes that reading comprehension scores are low. The sequence starts with test 1. Once the BLT reviews scores from the benchmark test, the members realize that scores across the school are, on average, below the proficient level. The BLT members decide to do something. They decide first to study how teachers are currently teaching comprehension and then to investigate other approaches-especially approaches that meet two criteria: (1) there's relatively strong evidence supporting their use and (2) the teachers believe they can use the approaches without much additional training. The BLT designs a new school-specific approach based on what they learn about evidence-based practices and practices that fit within teachers' comfort zone. The teachers use the new approach, first in a pilot test, and then for real. The school gives the next benchmark test. There's some gain-but not much. The BLT decides that the teachers need to keep using the approach they've designed. Eventually (and especially once the teachers are using the approach well), reading comprehension improves, and test scores go up.



As the example shows, the “hard data” don’t speak for themselves. The BLT speaks in consideration of the hard data. The people, not the data, make decisions, and the decisions involve actions. Once the people perform these actions for a long enough time, they look at data again. They ask, “Did the actions achieve their intended outcomes?” “If not, why not and what should we do about it?” “If so, how can we expand on our success?”

The process yields incremental improvements at first. Over the long term, not only do the increments accumulate, but the teams following such a plan-and the schools and district-develop a substantially enhanced capacity for improvement. In other words, the entire organization gets good at collaborating over data for the sake of better teaching and learning. This professional work shares something with science: a spirit of inquiry (*see the related Foundational Concept, The Spirit of Inquiry for Ohio’s Collaborative Teams*).

## A Deeper Dive: The OIP Embeds a Cycle, Too

The diagram above shows a capacity-building cycle interpreted at the BLT level of “granularity.” It’s close to the action where educ ators do their best work: the real work of helping students learn. It fits with the Ohio Improvement Process-where sometimes the “grains” are larger-focusing on the work of entire districts or even of the state as a whole. For the graphic representing the OIP from the large-grained perspective see:

http://www.ohioleadership.org/oipqs_recom/OIPQS_Visual_1_OIP_Visual_and_Summary_of_Stages.doc

To recap, here’s what ownership of data looks like:

- We’re involved.
- We interpret data.
- We plan.
- We do something.

- We interpret data again.
- We take the next step with care.

In this part of the module we'll be looking at the ownership issue-the really hard data issue-through these lenses:

- leadership and research,
- spirit of inquiry,
- issues and sources of data expertise, and
- cautions.

# Leadership and Research

## Foundational Concept: The Spirit of Inquiry for Ohio's Collaborative Teams

See the appendix for more information about the spirit of inquiry for Ohio's collaborative teams.

Under the principles of the OIP, leadership is a set of practices shared by leadership teams: TBTs, BLTs, and DLTs. Together, the teams and their members provide leadership. Data becomes a tool for their collective exercise of leadership. When this happens, educators stop being the victims of data and other people's questions and start asking their own questions and benefiting from their own data. It doesn't turn practitioners into researchers, but it does animate them with the spirit of inquiry. See the above Foundational Concept, *The Spirit of Inquiry for Ohio's Collaborative Teams* to learn more about the spirit of inquiry.

# Spirit of Inquiry: An Assessment Tool for Teams

Here's a tool that can help any team determine its already existing preparation for using data for collaborative inquiry. The team members might answer the questions individually and examine their scores to determine overall ratings and variation across their individual ratings. Or the team members might work together to reach consensus about the rating for each item.

The tool uses a scale from 1 to 5.

1 = I strongly disagree.

2 = I disagree.

3 = I neither agree nor disagree.

4 = I agree.

5 = I strongly agree

## *Spirit of Collaborative Inquiry Assessment Tool*

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Without good questions, teams can't get good answers. | | | | | | |
| I know how to ask really good questions. | | | | | | |
| It's good to have trustworthy help in developing questions pertinent to instructional practice. | | | | | | |
| I know someone in my school or district who is good with data. | | | | | | |
| It would be fun to work with other teachers on a team to improve our teaching. | | | | | | |

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| I understand the difference between formative and summative assessment. | | | | | | |
| I understand the difference between qualitative and quantitative data. | | | | | | |
| My team would be open to using interview data. | | | | | | |
| My team would be open to using survey data. | | | | | | |
| My team would be open to using classroom observations to collect data. | | | | | | |
| TOTALS | | | | | | |

It's simple. Add up the 1's, 2's, 3's, 4's and 5's to create a total score. A total score of 30 indicates overall neutrality when it comes to a collaborative team's spirit of inquiry. A team with this score is positioned to move forward with inquiry. A team with a score of 40 or more is prone to have the spirit of inquiry already activated.

A team with a score below 30 might want to spend some time working to increase its capacity for moving forward with collaborative inquiry. Here are some strategies they might use:

- Observe a team that the principal identifies as using data well.
- Talk about why members of the team gave low ratings to particular items on the tool and what actions they might take to increase their ratings.
- Ask a process coach (for instance, a consultant from a State Support Team) to observe one or two of their meetings to help them think about the process they've been using.

# Issues and Sources of Data Expertise

*Accessing data expertise* is important to teams. But it's a huge mistake to assume that having data somehow enables better decisions on its own. During the planning of this module, an Ohio educator now in a statewide leadership role, confessed:

> My staff got scared away before they even got into the process of collecting data. We were kings and queens of just admiring data, but we never knew what to do with it. I'm just being honest: we didn't know what to do. We would just look at it and look at it.

It's no wonder. Teachers and administrators deal with people from morning to night. Interactions are fast and furious, and most decisions are made on the spot. It has to be this way because that's what life is like (see, e.g., Schütz, 1967). No sane person would ask educators to consult data on the thousand decisions made every day! That's not what "data-driven decision making" could sensibly mean. See the table below.

## Driven Insane by Data

| Question | Sane? | Insane? | It Depends! |
|---|---|---|---|
| 1. Can the student go to the restroom? | | | |
| 2. Should we provide a recess period in the morning? | | | |
| 3. Should I pay a visit to the student's home? | | | |
| 4. Will I take courses for a master's degree? | | | |
| 5. What do I do with the student's answer? | | | |

| Question | Sane? | Insane? | It Depends! |
|---|---|---|---|
| 6. Who caused the fight? | | | |
| 7. Should I be assigning homework? | | | |
| 8. What homework will I assign? | | | |

The point of the table might be to draw attention to the “it depends” column! Much depends on the context in which the question occurs, it would seem. A leadership team might have decided that restroom policies are an issue (question 1). Addressing the issue with data might be a good idea. A teacher who knew that a particular student needed frequent trips to the restroom would, in a sense, be acting with data-but in a trivial sense. Question 6 is very familiar: but it is often the wrong question to begin with: “What caused the fight?” might be better. Questions 7 and 8 are a pair. Some teachers address question 8 with little forethought, and many with the “data” from the lesson they just taught.

Teams cannot and should not believe that they have to use data to guide everything they do. “Data-driven decision making” involves the use of data for strategically and tactically selected purposes. What are some of these purposes? (Remember: strategies are “grand plans” and tactics are “action steps.”)

## Strategic Purposes

Determining an appropriate balance between academic time on task and “brain breaks”

Identifying a school-wide plan for 0.2083teaching writing

Deciding how to engage parents from all parts of the community

Establishing a process for mentoring new teachers

## Tactical Purposes

Planning a common lesson on proportional reasoning

Creating a list of higher-level questions to guide a discussion about the Civil War

Developing a rubric for evaluating writing assignments

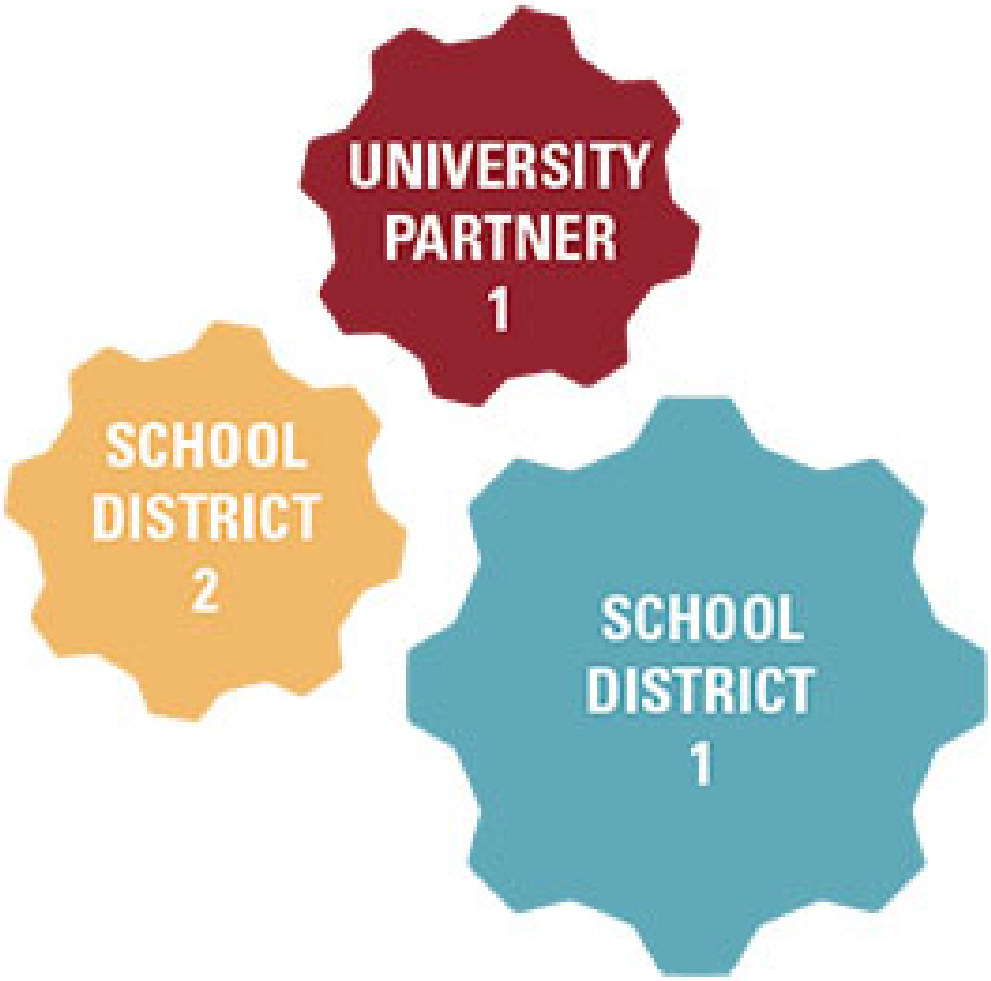
Writing a language arts unit on drawing inferences from text

Data can be used to help teams take on all of these projects and an infinite number like them. How? The team can survey students or interview them. The team can give a formative assessment to measure students' knowledge and skills. The team can create a draft of a lesson or unit plan and gather formative feedback from other educators and parents. The members of a team can all teach using an agreed-upon strategy and then observe one another using a formal observation tool they design (for instance, a rubric, checklist, or practice profile). The team can review the relevant literature about an evidence-based instructional strategy and discuss the implications for their own practice.

# Finding Data Experts

Leadership teams don't always have access to the help with data they might need. Pointing teams toward possible experts is one of the types of supports that a principal or central office administrator could provide. Here are some of the things they might do:

- identify teachers who like data and help them take on the role of local data expert;
- cultivate district staff who specialize in supporting the spirit of inquiry among collaborative teams, especially TBTs and BLTs;
- provide professional development to foster an appreciation for data collection, analysis, and interpretation among a relatively broad group of educators (see the Foundational Concept, *Handling Data in Ohio Leadership Teams*);
- identify consultants (for instance, from SSTs or ESCs) who can provide assistance to collaborative teams (including, in this case, the DLT); and
- broker contacts with helpful third-party consultants (e.g., carefully selected evaluators, practice-oriented researchers, and relevant non-profit organizations).



A systematic approach to bringing needed expertise to bear on educational problem-solving involves establishing, maintaining, and making strategic use of “Networked Improvement Communities” (Bryk et al., 2015). These communities create an on-going alliance between teachers, principals, central office administrators, researchers, and technical experts who work together to design and test innovations in ways that meet local needs. The “NICs” use a Plan-Do-Study-Act cycle of inquiry very much like the steps of the Ohio Improvement Process. The diagram to the left illustrates some of the partners who might join together to form a NIC. Other partners, such as SSTs, might also be members of the NIC.

# Cautions

There is an awful lot to learn about framing good questions, assessment tools, data gathering, data analysis, interpretation, and developing actionable plans (recommendations from the team to itself). And there is an awful lot for outside data experts to learn so they can become effective at supporting Ohio's collaborative teams.

Before leaving the key issues for teams in general, then, let's look at some of the predictable trouble-spots. The following pages explain the issues briefly and the Deeper Dive points to many relevant details, and some useful tools.

## Deeper Dive: Ideas Behind the Cautions

Here are some sources for further information about the issues discussed in the module's section on "Cautions." All these ideas are related to one another, by the way. Exploring the links can get you started on an interesting adventure. It will lead to a deeper understanding of various tools that can support inquiry among TBTs, BLTs, and DLTs.

**What is research:**
https://www.youtube.com/watch?v=v50ct9xJVKE

**Measurement error:**
https://www.edglossary.org/measurement-error/
https://en.wikipedia.org/wiki/Observational_error
http://www.proftesting.com/blog/2016/10/13/measurement-error-relationship-reliability/
https://www.socialresearchmethods.net/kb/measerr.php

**Sample size:**
https://www.socialresearchmethods.net/kb/sampterm.php
https://www.sciencebuddies.org/science-fair-projects/references/sample-size-surveys
https://en.wikipedia.org/wiki/Sample_size_determination
http://www.calculator.net/sample-size-calculator.html

**Reliability:**
https://en.wikipedia.org/wiki/Reliability_(statistics)
https://en.wikipedia.org/wiki/Internal_consistency
https://www.simplypsychology.org/reliability.html
https://www.socialresearchmethods.net/kb/reliablt.php

**Validity:**
https://en.wikipedia.org/wiki/Validity_(statistics)
https://www.edglossary.org/test-bias/
https://fcit.usf.edu/assessment/basic/basicc.html

**Significance level:**
https://en.wikipedia.org/wiki/Statistical_significance
https://hbr.org/2016/02/a-refresher-on-statistical-significance
https://www.medcalc.org/calc/comparison_of_means.php

**Qualitative data:**
https://www.youtube.com/watch?v=GbyYVS7_Las
https://en.wikipedia.org/wiki/Qualitative_research
https://en.wikipedia.org/wiki/Interview_(research)
http://salmapatel.co.uk/academia/coding-qualitative-research

# Cautions: Quantitative Ideas

## The unwcertainty of being certain: *measurement error*.

Every test score and every result from a survey is an estimate. So, when we look at a result it's good to take the "error band" into consideration. The error band around a specific result shows the range in which a true result is likely to fall. The lesson is ***don't pay attention, in general, to small differences.***

What does this mean practically? If the average percentage score of 30 students one month is 38% and it's 40% the next month *for the same 30 students on the same test*, the small difference might well result from measurement error rather than from actual learning. Check it out before celebrating!

## Representing a large group with a small one: *sampling*.

"To sample" means to take a small portion as a representative of a large thing. For instance, "Would you like to sample our soup?" Yes, you would, and when you take a spoonful, you infer that the spoonful is like the whole pot. It's a good bet with soup. A small sample won't work with surveys (or tests).

What's the lesson? ***Make sure there are enough participants to actually answer your question.*** Here's something you can remember: the smaller the whole group, the larger the proportion of them need to be in the sample in order to represent the whole group reasonably well.

When a team conducts a survey of students in a school, it should work very hard to get all students to respond to the survey. For a whole group of 50, accurate judgments require you to get 45 students to complete surveys! And for a whole group of 300, about 170 would need to complete surveys.

## Getting the same result again and again (sort of): *reliability*.

You can *rely* on a ruler to give the same result every time. An inch is the same every time. But if your measurement tool is a survey or a test, no matter how good it is, it won't yield *exactly* the same results for the same person every time.

What does this mean practically for teams? Poorly designed assessment tools are unreliable, and they won't tell the team what it wants to know. What's the lesson? ***Select or design good (reliable) instruments: most teams need help with this*** (see the Deeper Dive on the previous page)!

## The extent to which an assessment tool (aka "instrument") measures what it claims to measure: *validity.*

Here's a curious cooking fact. Americans use volume to measure ingredients in recipes! The choice is a poor one. Why? Volume measure misrepresents the proportions because volume isn't a valid comparison for ingredients. Weight is the valid measure to judge the amounts of different ingredients.

Similar mistakes are common with surveys, tests, and other assessment tools. They might not measure what we think they do. So what? Using an invalid instrument is practically guaranteed to produce a wrong answer! What's the lesson? Again, it's ***get good help with instrument design*** (reliability and validity are related, but different: see the Deeper Dive on the previous page).

## The odds of being wrong: *significance level.*

Do you buy lottery tickets? People have lots of reasons to play, but a sober calculation of the odds of winning doesn't come into it. The odds of winning the "Powerball" lottery are more than a hundred million to one!

If we want to compare test results across years or compare results for different groups, a question of odds also comes into play. What are the odds that a calculated difference flags a real difference and isn't just a bunch of noise (aka error)? We can compute the odds that observed differences are real differences ("significantly different"). Usually odds of 19 to 20 (that is, a 95% likelihood) that apparent differences are real (that is, that they are not due to chance) is regarded as an acceptable significance level. With those odds of it not raining, you'd feel fairly confident about having a picnic or mowing the lawn.

What's the lesson? Collaborative teams should be careful when analyzing quantitative data. Small changes might not mean anything. Alternately, teams can ***get good help analyzing quantitative data to discover real differences (for example gains between pre- and post-tests) that are not due to chance.***

# Cautions: Qualitative Ideas

What are “qualitative data”? Qualitative data are words and sentences: and the ideas and meanings revealed in words and sentences. We all use words. We all make meaning with words. We use words to teach. Students use words to learn. We’re surrounded by qualitative data.

A different set of cautions applies to qualitative data. The discussion of cautions about qualitative data actually starts off with a Big Encouragement! It’s a caution about something not done.

According to the advisers to this project, many of Ohio’s collaborative teams are ignoring this treasure trove of information about teaching and learning. Quantitative data provide an important part of the overall picture, but qualitative data are also essential.

Here are some typical modes of qualitative data collection:

- document analysis,
- observations,
- closed-ended question interviews (aka structured interviews),
- open-ended question interviews (aka semi-structured interviews), and
- video recordings.

If few teams are now using these methods, most teams will need help-possibly even more help than with quantitative data. But this caveat just reinforces the point that teams can benefit from help with data of all sorts, qualitative and quantitative.

In some ways qualitative data are closer to reality than quantitative data-they feel more real. They certainly can give a richer and deeper picture of things. But their closeness to reality has some downsides. For one thing qualitative data are messier than quantitative data. They are harder to analyze, and our pre-existing beliefs are more likely to influence them. Some short clips from observation notes illustrate the point. What might be wrong with each of the descriptions in the four squares below?

Sam was misbehaving, and Mrs. R. scolded him.

Mrs. R. was in a bad mood.

Mrs. R. was nicely dressed in a new pantsuit.

At 10:01 am Sam yawned; at 10:02 am Sam mumbled to himself; at 10:05 Sam got out of his seat; at 10:07 Mrs. R. said, "Sam, sit down."

Here's a quick critique:

- **Sam was misbehaving, and Mrs. R. scolded him.**
  "Misbehaving" doesn't really tell us what Sam was doing. And "scolding" doesn't really tell us what Mrs. R. said or how she said it.
- **Mrs. R. was in a bad mood.**
  In this case, the observer is making an inference about Mrs. R.'s internal state of mind based on external signs. That inference probably relates to the way the observer acts when he or she is in a bad mood. It probably doesn't have a whole lot to do with how Mrs. R. acts when she's in a bad mood.
- **Mrs. R. was nicely dressed in a new pantsuit.**
  This statement is likely to be biased. First, the choice to comment on how Mrs. R. was dressed may reflect the observer's belief that how a person is dressed tells us something about the person's character or behavior. Second, "nicely dressed" reflects the observer's criteria, not necessarily criteria that everyone would agree on. Third, the statement that the pantsuit is "new" is an inference. Unless the observer is close

friends with Mrs. R., he or she probably doesn't know when or how Mrs. R. obtained the pantsuit.

- **At 10:01 am Sam yawned; at 10:02 am Sam mumbled to himself...**
  This running record tries to use objective language to describe what is happening (almost from minute to minute). It is less likely to be biased than the other statements describing what's going on. But if it's part of a larger observation record (for instance, a running record from a 30-minute evaluation), the final set of observation notes will eventually be rather long and complicated. That's not a problem in and of itself. But figuring out a reasonable way to analyze so much data might prove challenging.

# Key Data Issues for Teacher-Based Teams

## Encouragement for Engaging the Spirit of Inquiry

Engaging the spirit of inquiry, as presented in the Foundational Concept, may be new territory for many Ohio educators, as it is for many educators nationwide. In fact, what Ohio's collaborative teams are doing is pioneering work!

This module now goes into more detail related specifically to particular teams: TBTs, BLTs, and DLTs. It starts with TBTs.

## Using the Ohio Improvement Process (OIP) 5-Step Process

Let's start at the beginning . . .

> The purpose of TBTs is to work continuously to strengthen teaching and learning for all students through collaborative planning based on the collective interpretation of formative assessment data. (OLAC Module: Teacher-based Teams: What Districts Need to Know)

A TBT is a group of teachers and, at times, other educators (e.g., related services personnel, paraprofessionals) who meet regularly to hold discussions and make decisions to influence their instruction and students' learning for the better. The Ohio Improvement Process (OIP) embeds a 5-step process for TBT work. The five steps are a concept, however, not a directive or a checklist. This work is not a compliance ritual, but the very heart of educators' professionalism.

## The 5-Step Concept

The OIP 5-steps describe a process, or better yet, an idea. This means that the actual use of the process will vary somewhat from team to team, and across schools and districts. The different circumstances of use mean that it cannot be used in the same way everywhere. If TBTs are *working continuously (and collaboratively with data) to strengthen teaching and learning* they are obligated to translate the idea behind the 5-step process in ways that help the team make sense of what they are looking at. And, in general, teams are free to look at anything that dovetails with district and school goals. Of course, in each specific instance, the logic of the actual work in which a TBT is engaged will determine what the team will consider at each of its meetings.

Should every 45-to-60-minute TBT meeting address all five steps? Or does that not matter? Depending on circumstances, the answers can vary. BLTs and principals should support TBTs as they tailor 5-step processes that make sense in their specific situations for their specific missions. Here's a simple statement of the five steps:

- Collecting and charting relevant data;
- Analyzing student work specific to the data;
- Establishing shared expectations for implementing specific effective changes in the classroom;
- Implementing changes consistently across all classrooms; and
- Collecting, charting, and analyzing post data.

## What exactly *is* the 5-step idea if it's not a checklist?

The process can be conceptualized like this:

*TBTs get data, talk about it sensibly, try to make good decisions about how to*

*teach, put their good decisions into action, and see what happens. Then they repeat the cycle.*

It's not rocket science. Anyone could devise this idea: get data, use data, do it again. Some people call it the **Plan, Do, Study, Act** cycle (see e.g., Deming, 1986). The idea may be simple, but the reality of doing it isn't simple.

# Why The 5-Step Process Isn't Simple

First, the collaboration of teachers in a focused improvement process is hardly traditional. What's traditional is the isolation of teachers: close the door and the classroom is yours. Those days are over. Second, teachers aren't trained to gather or analyze data, especially not as a group project.

That is the key data issue for TBTs. *Most teams talk about data, but few do much about it*. Teams are too often left to flounder… and then to *founder*.

Why "data"? It's a good question. A typical answer is that discussing data is better than discussing ("mere") personal opinions (see, e.g., Love et al., 2008; Mandinach & Gummer, 2016). Perhaps, but professional judgments aren't banished from conversations about data-they better not be: because professional judgment is the responsible party here. The data aren't in charge! It might be better to refer to this realm of thought and action as decision-driven data use. The content below shows some related ideas and points to video resources to help you learn about them.

## Understanding the 5-Step Process: Related Ideas

- Improvement Science
- Implementation Science
- Action Research
- Plan-Do-Study-Act

## Whose Data Are They?

The data under discussion by TBTs come from the teachers on the team (and perhaps others) and from their students. These data bring something about students and their teachers to the attention of TBT members. Not only that, the 5-step process intends that those data be the focus of conversation. Data are important for this reason. Data represent teaching and learning with *these particular students and these particular teachers*. Good data *describe*-to those who really care about it-what is going on in classrooms. Genuine engagement of the team with these descriptions-and taking action to improve based on these descriptions-is the whole point and purpose of TBTs.

# What Can TBT Members Do On Monday?

Get relevant data.

There are two main parts to "getting data." First, TBTs need a tool for *getting data*. Second, they need to use the tool and actually get the data. The first part is nearly always more challenging than the second.

TBTs can...

| Get a Tool | Use the Tool to Get the Data |
|---|---|
| A lesson observation form. | Observe team member's teaching of a lesson that the team developed collaboratively. |
| An assignment description and grading rubric. | Give students a task and gather the work students produce. |
| A test. | Give students a test. |
| A survey (by the way, an exit slip is a short survey). | Give the survey to students. |
| A short set of interview questions. | Interview parents. |

So far as getting data goes, this short list covers almost all bases for TBT data-gathering efforts: (1) inspecting student work, (2) observing teaching, (3) giving tests, (4) giving surveys, and (5) conducting interviews. Teams can also consider using videos (e.g., of a collaboratively developed lesson). And they can also pool their efforts to search for resources, with the results constituting data.

# Get The Data

"Getting data" is the first step of TBT work with data—once decisions have been made wisely about why to get the data. The first section of this module, on issues for all teams, goes into some of the complications involved with wise choices about getting data:

- teams need good (valid and reliable) assessment tools (aka instruments);
- questions have to be written with great care; and
- the assessment tools (tests, interview protocols, surveys) have to be administered with similar care.

## Do Something Meaningful with the Data

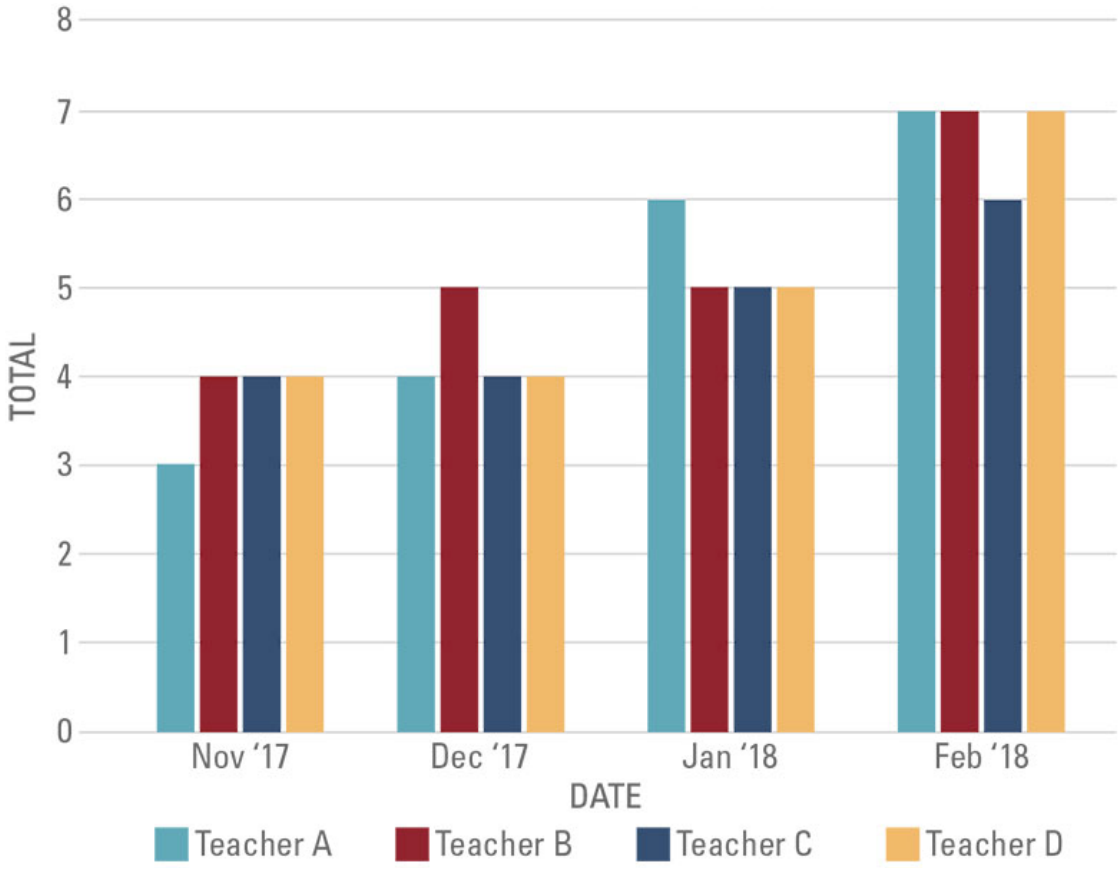
Once the team has the data, what can it actually do with it, aside from staring at it? With the 5-step process, TBTs are supposed to "analyze" data-but few teachers have experience with data analysis.

*Every TBT can start somewhere*. Though it would benefit from appropriate support, this sort of "data analysis" doesn't require the expertise of a highly trained researcher or statistician. (The Foundational Concept about data displays and the Deeper Dive for ideas related to the cautions can help TBTs make decisions about the help they might want to seek.)

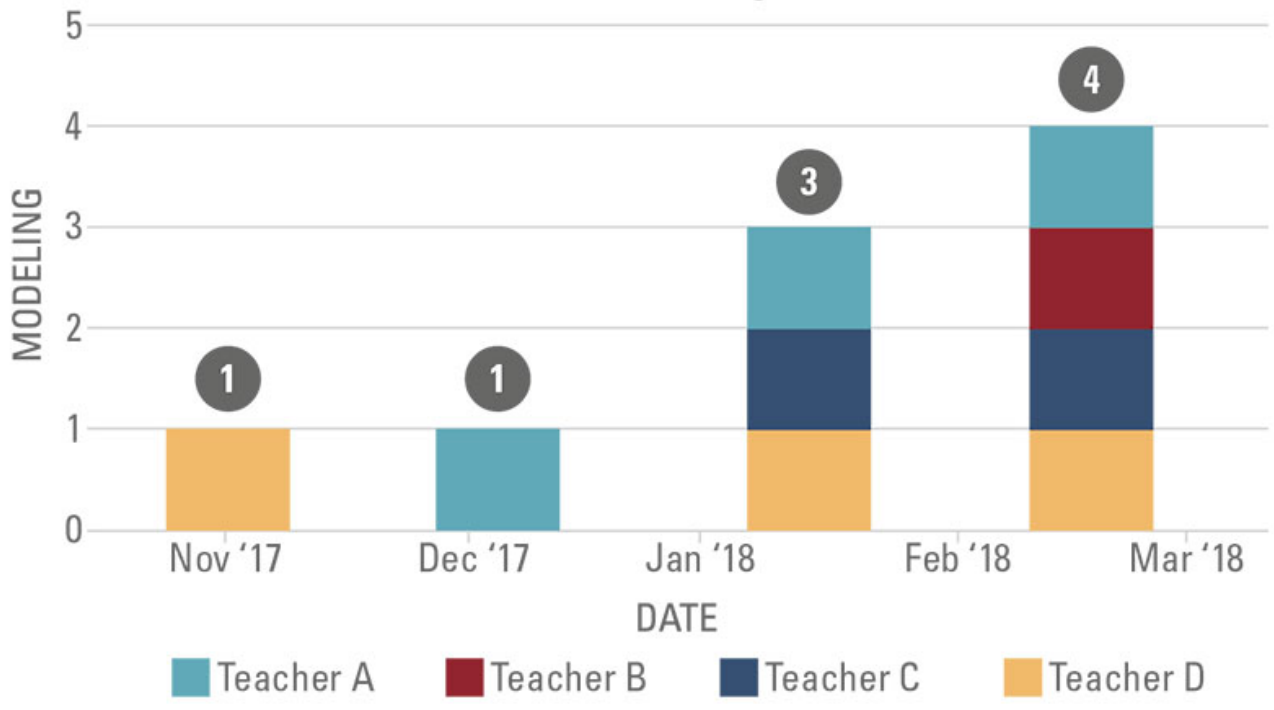
*Displaying data is where to start*. With most data that TBTs use, "analysis" is about finding a good way to display the data. Here are some examples of displays of teacher implementation data. They come from a TBT where teachers were practicing the use of direct instruction and observing one another to collect data on fidelity of implementation. Below is the full set of data on which the displays are based.

Note that the bar graph shows which teachers used how many direct instruction practices during each month's observation. The major storyline is that, over time, the teachers learned to implement more and more of the practices used for providing direct instruction. The stacked bar graph shows their use of the practice of modeling over the four months. Perhaps after the TBT reviewed the data for the first two months (November and December), the teachers realized that modeling was difficult to incorporate and therefore that they needed to be more intentional about using modeling in every direct-instruction (DI) lesson.

## Number of DI Practices




## Modeling



# Teacher Data

| Date | Teacher | Anticipatory Set | Objectives | Input | Modeling | Guided Practice | Independent Practice | Closure | Total |
|---|---|---|---|---|---|---|---|---|---|
| 11/1/17 | Teacher A | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 3 |
| 11/1/17 | Teacher B | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| 11/1/17 | Teacher C | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| 11/1/17 | Teacher D | 0 | 1 | 1 | 1 | | 1 | 0 | 4 |
| 12/4/17 | Teacher A | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 4 |
| 12/4/17 | Teacher B | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 5 |
| 12/4/17 | Teacher C | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| 12/4/17 | Teacher D | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 4 |
| 1/13/18 | Teacher A | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| 1/13/18 | Teacher B | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 5 |
| 1/13/18 | Teacher C | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 5 |
| 1/13/18 | Teacher D | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 5 |
| 2/17/18 | Teacher A | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 2/17/18 | Teacher B | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 2/17/18 | Teacher C | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 6 |
| 2/17/18 | Teacher D | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |

Let's look at another example. Perhaps the teachers in a TBT focusing on Algebra instruction at a large high school gave their 20 lowest-performing students a math anxiety survey to figure out if these students might be particularly uncomfortable doing math. The survey had 10 questions with responses on a 1-5 Likert scale (numerical, or quantitative, data). And let's say all 20 students completed the survey. Now what?

- Team members might circulate the 20 completed survey forms (all of them) for all the TBT members to review before the meeting. If the team simply scans the responses, they will indeed learn something about the data. This approach seems simple enough, but it also seems a bit cumbersome.
- Team members might, in fact, want a *better organized* way of reviewing the data. So, one team member could enter the data in a spreadsheet. Then the team could review all the responses at a glance, referring to a single document (instead of looking through all 20 completed forms).
- Maybe team members want even more. For instance, they may want the average response for each item and for total scores. Since the data are already in a spreadsheet, it's easy for a team member who uses spreadsheets to calculate total scores and calculate averages.

- If the team wanted to compare responses from females and males, that too (and much else) can be done using a spreadsheet (e.g., standard deviations, correlations, and the reliability of all 10 items or of subsets of items).

# Dealing With Words As Data

Numbers (quantitative data) are what comes to most people's minds when they hear the word "data." But TBTs should also embrace qualitative data: words. Humans are talkers- it might be one of their defining characteristics. And, for qualitative analysis, words (not numbers) are the data. Rarely does any treatment of educator teams dealing with data consider the usefulness of words as data, perhaps because professional development on using data wisely is so rare. But words are data.

## Words are Data

Why does the word "data" mean *numbers* to so many educators? One reason may be our cultural fixation on test scores: these are the numbers that everyone wants to improve. Another reason is that numbers seem to be scientific. E=mc2: the famous formula is a number. Numbers formulate laws ("the laws of physics") that are believed to be universally true.

But in everyday life, words represent what people see, think, and feel. Teaching happens with words. Classrooms run on words. Words may not be universally true, but they are used universally.

Ohio educators can embrace qualitative data. When teams need to know what people (students, teachers, parents) see, think, and feel, examining their words is critical. Those words are data for such an examination. Letting people talk allows them to represent in their own way what they see, think, and feel. If we care about that (and we do care), we'll listen.

Interviews are typical sources of data based on words. How might teams with little experience of qualitative data take on the analysis of interview data?

- Team members (some or all of them) could read the responses to the interview questions. And then re-read them, perhaps several times. Then they could discuss their impressions.
- To organize the data a bit more, a subgroup could read the interviews several times and summarize the responses. In this summary, an average response is not helpful. Instead of an average, what's needed is a characterization of the most **common responses.** Depending on what the team might want to know, it might make more

sense to tease out the most unusual responses.

- Still more organization might be helpful if the subgroup wanted to make some inferences about the major storyline in the data. These inferences are called **“themes.”** Themes take the common responses and derive meanings associated with groups of responses. *The meanings are expressed as sentences*. This step is most useful when there is a lot of qualitative data.

For points two (common responses) and three (themes) above, the essential “analysis” involves grouping similar things together. But sometimes we also want to identify differences based on certain criteria. For instance, as with quantitative data, we might want to look for patterns across subgroups. Other patterns that might be of interest relate to the sequence of events, possible causes and effects, or close associations of one idea with another idea. All of these analyses involve searching for patterns. Because we search for patterns in a lot of what we do, these analysis methods will feel familiar.

# Decisions, Decisions

Perhaps a team has collected the data on a 10-item survey and analyzed the results for 20 students. And let's say the team has concluded that overall, girls are more positive in their responses than boys. We'll assume that the survey is well-designed, without ambiguous questions, worded simply, and given responsively (e.g., read aloud to students with reading challenges).

> ***Data have no meaning. Meaning is imposed by interpretation.***
>
> (Love et al., 2008, p. 5).

It's odd to say, but the decision the team will make has less to do with the data than with members' knowledge of what teaching and learning are like in their school. In other words, TBT members must impose something meaningful onto the data based on what they see in the data and what they know about the context.

The interpretation might be that instructional methods need to change (somehow) in order to appeal to boys. Another interpretation might be that classroom climate needs to be understood better as it relates to male and female students. Still another

interpretation is that more information is needed. Perhaps interviewing *students* can provide the information needed to decide what is going on. Additional data can help the team follow its hunches.

Each alternative raised by team members represents a possible decision. Although each alternative takes the data into account, the main focus of decision-making is not the data. The main focus is what the data say in light of what team members *already know* about the context.

## Short-sighted Attempts to Raise Test Scores

There are no guarantees that any decision—even a wise one—will "produce results." Positive outcomes are, of course, desirable. But they are like happiness. Running straight at positive outcomes, like running straight at happiness, is usually short-sighted. And because it's short-sighted, the straight-line approach to something subtle and delicate and complex (student learning in this case) is likely to put one further away from it, not closer to it. Running straight at improved test scores, for instance, has often entailed cheating (see e.g., Amrein-Beardsley, Berliner, & Rideau, 2010). In the worst case, this approach corrupts the entire enterprise of schooling.

Discussing alternatives in consideration of data builds an instructional conversation. An instructional conversation is serious and interesting, and sometimes difficult. But, over time, participation in instructional conversations builds the capacity for and eventually the habit of discussing instruction in consideration of data. And that habit is where the improvement of instruction and learning comes from. TBTs that use data wisely work on their own teaching. And that good and collaborative work benefits student learning above and beyond ordinary, isolated, practice.

# Key Data Issues For Building Leadership Teams

## The BLT Mission

Because the BLT looks at data from all TBTs, its members can figure out patterns and respond to them (e.g., all of the TBTs want to learn more about questioning techniques, the 5th-grade math teachers are getting better results than the 4th-grade math teachers). This knowledge can help the BLT plan capacity-building efforts (e.g., giving everyone coaching on questioning skills, asking the 4th-grade math teachers to begin talking about teaching strategies with the 5th-grade math teachers). Building Leadership Teams exert schoolwide influence and build their own improvement capacity partly through scaffolding the improved capacity of TBTs.

Key issues for BLTs include data processes relevant to:

- gathering and analyzing schoolwide data,
- working with TBTs' data about teaching practices,
- coordinating and supporting TBTs' data capacities,
- using data to foster collective efficacy and purposeful school climate, and
- providing data to inform DLT planning.

The work of improving teaching and learning-at any level of the education system-requires collaboration. The reason should be obvious: the capacity required is larger than one individual can muster or sustain. BLT members collaborate mainly on behalf of building (school) leadership. Discussion of each key issue just listed follows.

### Working with TBTs on Teaching Practice

BLTs often look at test scores, but test scores of all sorts have been over-emphasized for many decades (e.g., Hursh, 2008). Turning more attention to teaching practices makes sense for TBTs, and a re-focus on teaching practice by TBTs will require scaffolding from BLTs. What can BLTs do to build the scaffolds?

- share information about effective teaching practices with everyone in the school,
- encourage teachers to share effective teaching practices within and across TBTs,
- provide coaching to help TBTs engage in productive instructional conversations,
- coach TBTs in the effective use of data,
- encourage regular collaborative (peer) observation,
- provide useful tools for peer observation,
- model the use of surveys and interview techniques,
- provide professional development focusing on the data collection and analysis skills that are needed for TBT work,
- introduce alternative models for collaborative inquiry, and embody the spirit of inquiry for the school and cultivate it widely (including among students).

Testing student knowledge and reviewing test scores is undeniably useful. Overuse, though, makes it counterproductive. Even short-cycle assessments can be used too often or misused. Testing is best when it is part of a systematic attempt to improve teaching practice. In other words, work on teaching practice takes priority in the wise use of data.

# Gathering And Analyzing Schoolwide Data

It's no longer recommended that TBTs focus their attention on analyses of state achievement and graduation test data. These data are too far removed, both in time and in level of detail, from actual instruction.

Such reviews, however, are appropriate for BLTs. Reviews of this sort include disaggregation of data by subgroups relevant to the locality. For such reviews, a BLT member proficient with spreadsheets might prepare materials based on the team's consensus about what the team needs to examine and discuss. Depending on the focus, the review materials might include:

- displays of the most recent annual results,
- five-year trend data (unless policy changes render comparison dubious), and
- data that offer a drilled-down focus on specific subgroups, subjects, or within-subject domains.

Accountability demands that BLTs review such data, identify needs, make plans to address the needs, and carry out the plans. Like TBTs, BLTs also collect data on their own implementation efforts, which have a distinct schoolwide flavor. The actual data tools, too, can be much the same as for TBTs: tests, surveys, interviews, observations, and teacher products of various sorts.

The real differences between the BLT mission and the TBT mission involve its scope and complexity. BLT members ideally bring considerably evolved professional skills, insights, and experiences to the table. The increased scope and complexity of the mission demand such background of members.

Although responsibility for learning schoolwide embeds attention on achievement and/or graduation test scores, the scope and complexity of building leadership extends much further. Other schoolwide responsibilities that call for data gathering, analysis, planning, and implementation include:

- reducing schoolwide manifestations of inequity (as manifest in race, gender, social class, and ability differences);
- addressing TBT implementation issues;
- providing support for the development of teachers' collective efficacy;

- eliciting and responding to students' and parents' views (e.g., of school climate, of school operations); and
- identifying, implementing, and monitoring schoolwide innovations (e.g., PBIS, Writing-Across-the-Curriculum).

Whereas TBTs *benefit* from basic data expertise, the scope and mission of BLTs *requires* additional data expertise. The schoolwide scope has a large data footprint.

The associated complexity moves the necessary discussion in BLTs beyond simple inspection of data and measures of central tendency (for quantitative data) and summaries of short interviews (for qualitative data), which are reasonable for TBTs. For BLTs, more formal data analysis helps reduce the data load, and it helps more clearly surface underlying trends and issues. Understanding the applicable associations within the data and determining which differences are due to chance and which are not can be critical to BLT decision-making.

## Significance Testing, Correlational Methods, and Controlled Experiments

These are statistical techniques that BLTs would do well to know about and use—from time to time. Here we overview them very briefly (see the "Cautions" Deeper Dive for additional perspectives).

**Significance Testing.** For BLTs, the key use of this technique is to assess how much of a difference in test scores (from year to year, positive or negative) is needed so that we can be sure it's not due to chance. This information isn't reported: it must be calculated by those interested in finding out. In fact, everyone involved should be interested! If we don't know, we can be duped.

**Correlational Methods.** Two or more measures (achievement, attitudes, socioeconomic status) can be correlated. In fact, a set of measures (for instance, social class, quality of instruction, and curriculum) can predict a key measure of interest (for instance, student test scores). But these influences vary from place to place. And if a BLT wants to reduce the influence, say of social class or race, on a school's achievement results, then correlational methods would prove useful: or, actually, essential.

**Controlled Experiments.** Leadership teams plan action and then want to

discover if the action “worked.” Without significance testing, they can’t really judge. But they also can’t judge simply by giving a pre- and post-test. If there are positive results that are statistically significant, that’s good, but it’s not proof that the action worked. For that, the BLT needs a controlled experiment. Controlled experiments carry out the action with one group and compare results with a different but comparable group. Groups formed by random assignment (that is, with students randomly assigned to groups that either get the treatment or not) is the hallmark of control, but alternatives exist.

These are just some of the ideas relevant to what BLTs might do. Actually testing significance, measuring associations, and running controlled experiments requires some expertise—*and learning by doing* can include help from outside the BLT.

# Working With TBTS' Implementation Data

The BLT's interest in the progress of instructional practice is a key part of the team's schoolwide responsibility. Indeed, a building leadership team exists to exercise instructional leadership. Thus, the realm of "implementation data" is of prime concern. BLTs' primary source of implementation data is from TBTs.

That expression–*implementation data*–needs some unpacking, however. First, anything that teachers do with a bearing on instruction falls into this category. With guidance from BLTs, TBTs select the instructional practices they want to learn and implement. Second, implementation of a particular instructional practice cannot be focused primarily on compliance. A focus on mere compliance destroys the spirit of collaboration. Instead, implementation must focus on the refinement of technique. The difference is subtle, but extremely important. The table below shows how the two ideas differ. It's critical to remember, however, that a TBT always works with a somewhat restricted set of techniques: district and school goals and strategies establish parameters specifying the set of techniques that members of TBTs work to refine.

| | **Compliance** | **Refinement of Technique** |
|---|---|---|
| Source of the requirement to change | External to the TBT | Internal to the TBT |
| Operational definition of the practice | A prescribed "script" that teachers enact | A flexible set of procedures that TBTs design and teachers learn to use |
| Motives for compliance | A system of external rewards and sanctions | Improved efficacy (personal and collective) |
| Response from teachers | Defensiveness | Pride in new accomplishments |

## Coordinating and Supporting TBTs' Data Capacities

- What can BLTs do to (1) coordinate and (2) support TBTs' data capacities? The two functions are related because the steps taken reinforce each other:
- observe TBT meetings,

- coach data discussions in TBT meetings,
- coach instructional discussions in TBT meetings,
- demonstrate to TBTs the BLT's appropriate use of TBT implementation data,
- establish a support system for TBTs at the school, and
- scaffold TBTs' use of the support system.

The chief data-relevant concept is that TBTs focus their discussions on information about teaching and learning. Here's the curious fact, however. A focus on data, embedded in an inquisitive framework, inevitably produces more questions than answers.

Those questions will, sooner or later, require additional capacities to secure answers. This is the juncture that ought to elicit just-in-time professional development for whatever the TBT requires in the way of enhanced data-analysis skills.

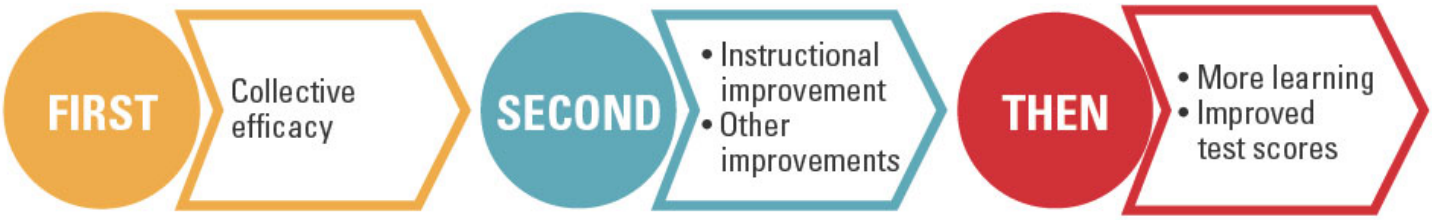
The skills needed are fairly predictable:

- asking good questions about their own teaching practice;
- developing protocols for observing one another's teaching;
- writing better items for surveys or tests;
- calculating and using measures of central tendency;
- categorizing data from survey comment fields;
- developing an instrument to measure something (e.g., questioning skills); and
- improving instrument reliability.

# Using Data To Foster Collective Efficacy

Instruction is the realm of professional action for teachers. Instruction is what educators do in the name of learning. True, instruction sometimes misfires. Sometimes, and in some places, the misfiring becomes routine. In such cases, low test scores are evidence of the routine misfiring. Accountability supports and sanctions then come into play. Teachers and administrators are embarrassed and upset. Test scores need to go up. Desperately. Desperation of this sort, though, does not lead to instructional improvement.

What's the alternative? Using data to foster collective efficacy. Collective efficacy is the cause of instructional improvement schoolwide. And instructional improvement is one cause of improved learning, which-schoolwide-will manifest as higher test scores. Here's a diagram showing the associations:




Collective efficacy is a cultural phenomenon, however. A BLT, the school's instructional leadership team, cultivates collective efficacy over the long term. Such cultivation, in fact, needs to be perpetual.

# Using Data to Inform DLT Planning

BLTs are middle-management units. This just means they occupy the ground between classroom teachers and district leadership. They have a larger data footprint than TBTs, but not so large as DLTs.

And just as TBTs are sources of data for the BLT, BLTs are sources of data for DLTs. Indeed, the BLT position in this regard is *interesting*. It gives the school a voice with respect to influencing district priorities.

This use of data, though, rides atop a great deal of discussion and reflection-of a strategic nature-first at BLT meetings and then at DLT meetings. The questions and issues are legion:

- What are the school's major struggles?
- What resources have been brought to bear?
- Where is it going well, and where not so well?
- What are the theories about *why?*
- What are the relevant data that address *why?*
- What additional support is thought to be needed?
- What capacities does the BLT have and what capacities does it lack?
- What might the district do to help?
- How does all this compare with the experience of other schools in the district?

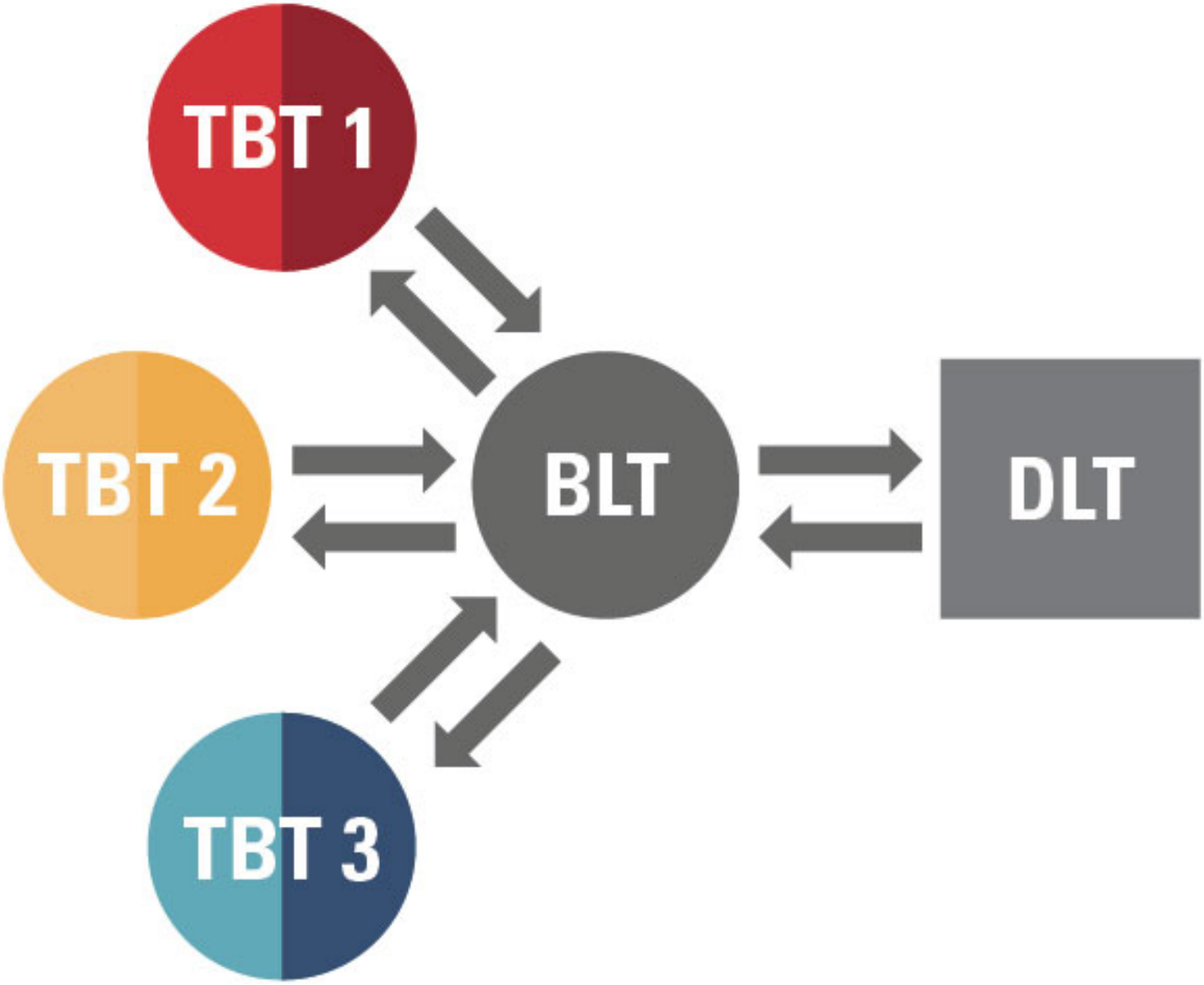
This is interesting work, and a hallmark of leadership in action. To make wise decisions in consideration of data, including decisions about what to emphasize in reports to the DLT, BLTs must focus attention on their most pressing issues. In fact, the practice of using data well depends on the practice of maintaining a focused set of goals and strategies. As this module has noted before, data that drift off target-away from focused goals and strategies-draw attention away from what leadership teams most need to talk about and do.

Of course, the DLT may-given its even larger data footprint-expect prescribed data-based reports of various sorts from the BLT. And those standing requests will inevitably structure some of the BLT's data gathering and analysis efforts. But it's important that the BLT reserve energy and time for its own agenda. The agenda will be consistent with the school and district improvement plan, and when necessary, the BLT can provide information to help the district revise the existing plan. Iterations of this sort are a key part

of the spirit of inquiry undergirding any effective cycle of improvement.

## Bi-directional Nature of the Information Flow

The diagram below illustrates the connection between the DLT and BLT, and between TBTs and the BLT. By now, it's a very familiar picture. It's included here just as a reminder of the bi-directional nature of the information flow.



# Key Data Issues for District Leadership Teams

## DLTs and District Scale

The extreme range of size in Ohio means that the concept of the District Leadership Team (DLT) is likely to need adjustment in very large districts. Why? The complexity of the operation and the scope of all data operations (e.g., financial, human resource, transportation) is so vast and the related issues so complex that a single team above the level of BLTs would be overwhelmed in large districts. Few of the things that need to happen could be accomplished, in other words, by a single team. In fact, a single DLT would probably contribute to communication bottlenecks rather than facilitate communication districtwide.

Ohio districts, in fact, range in size from enrolling fewer than 400 students to enrolling over 50,000. The average Ohio district enrolls about 2,000 students (Ohio Department of Education, 2018). So, the scale of the largest districts in the state are almost two orders of magnitude larger than the smallest and more than one order of magnitude larger than the average!

One alternative in very large districts might be to organize multiple DLTs around a high school and its feeder schools. The conception of the DLT is that it represents the portfolio of district interests–not that a district must operate just one such team. Another alternative is that a large district might subdivide a large and representative DLT into several collaborative subgroups focused on specific aspects of improvement and reporting to the DLT.

Key concerns of DLTs (however organized) related to data use include:

- articulation of strategic data issues (including the selection of strategic indicators),
- review of BLT evidence-based needs (including their data-capacity needs),
- maintenance of DLT data capacity, and
- cautious use of achievement and graduation test data.

# Articulation Of Strategic Data Issues

> *The structural and operational changes in the district hold the potential to cultivate a learning organization rooted in knowledge sharing and creation with a continuous goal of improving student learning... the teachers, principals, and senior administrators all invested in transforming the culture of their district. But it is noteworthy that it was the district leadership who purposefully and strategically engaged in deliberate actions to reform their district.*
>
> (Hannay, Ben Jaafar, & Earl, 2013, p. 77)

In theory at least, school districts are independent local entities. Certainly, the state education agency (SEA) and others influence Ohio's local school districts: but Ohio's local school districts retain an actual autonomy rare in educational systems among the world's developed nations.

This fact means that local school districts have the authority to determine their own data needs with respect to district and school improvement. DLTs have a role in exercising this authority.

As with BLTs, DLTs should not fixate on achievement and graduation test scores. They should inspect and discuss trends cyclically. Attempts to secure rapid increases in test scores routinely miscarry, rarely amounting to real improvement. Real improvement takes real work, especially work to build capacity. And building capacity takes time. Searching for a quick fix sets the wrong tone for an entire district!

The DLT plays a *strategic role* with respect to data. Micromanagement in the operation of BLTs and TBTs wastes DLTs' time and energy and cultivates learned helplessness (Hannay et al., 2013). The DLT's special mission is strategic.

What does that mean? "Strategy" refers to large-scope and long-term planning and action. The DLT sets districtwide goals related to improvement. On this view, its most critical data function is to identify initiatives that it judges will contribute, system-wide,

to instructional improvement. Clearly, the DLT will use data in the course of exercising its judgment for this effort.

## Districtwide Math Improvement

As winter was coming to an end, the Centerville City School District Leadership Team was holding its third meeting of the year. “We need to improve middle school math achievement,” reported the curriculum director. “I’ve been watching the state test scores, and they’ve been down for two years in a row.”

“Let’s be careful, here,” said the middle school principal. “You know the tests changed last year. Of course scores are down.”

The superintendent jumped in, “Well you’re both right in one way, and you’re both wrong in one way. You’re right because one of our major goals is to improve mathematics learning through equitable opportunities to learn, supportive grading practices, and improved instruction. But you’re wrong to focus primarily on state test scores. Those scores can be part of the picture, but they don’t tell us much about progress with our strategic initiatives.”

“Okay, I hear you,” said the middle school principal. “I was feeling a little defensive. Could you remind me: if we’re not looking at state test scores, what are we looking at?”

“Sure,” said the superintendent. “I’ll review our three primary indicators.”

# Strategic Indicators

The vignette in the call-out box on the previous page tells us something about the change process taking place in Centerville. It seems like the DLT has set goals and selected a focused set of strategies related to the improvement of math learning. It also seems like the DLT has selected some strategic indicators. What do we mean by that term?

> A strategic indicator is a measure that can be used routinely to monitor the progress of an organization toward a strategic goal.

## Long-term goal: To improve math learning district-wide

The table below shows a portion of Centerville's strategic plan that indicates its strategic goals, strategic indicators, and frequency of measurement of each strategic indicator.

| Strategic Goals | Strategic Indicators | Frequency of Measurement |
|---|---|---|
| To increase equity of access to mathematics learning opportunities | % of African American students in high-level math sections | Once a year, after the fall census count |
| To increase district-wide use of a grading system combining mastery learning and standards-based grading | % of teachers by school using mastery learning; % of teachers by school using standards-based grading | Twice a year, keyed to the roll-out of the two innovations |
| To improve math instruction | Average scores by school on the Centerville Standards-Based Math Teaching Rubric | Twice a year, in December and June |

In addition to the indicators listed on the table, Centerville's DLT will also review state test scores annually and scores from benchmark tests twice a year. Most of the work to monitor the implementation of strategies, the achievement of outcomes, and the linkage between the two (i.e., implementation and achievement) will happen at each school's BLT.

### Strengths Can Point to Strategies

Building on strengths is a major "rule of thumb" for expanding capacity. Strengths show where we already have capacity. BLTs, for instance, often work to discover their schools' best teachers and build on their strengths (e.g., by redeploying them part-time as instructional coaches and ensuring their presence on TBTs and the BLT). DLTs also have a responsibility to think strategically about how to make use of local assets to expand districtwide capacity.

## Review of BLT Evidence-based Needs (Including Data-capacity Needs)

DLTs exercise active care for their districts' BLTs by gathering data about the functioning of BLTs and TBTs. Such information helps DLTs determine how to interact with BLTs. For example, members of the DLT might play a participant-observer or coaching role-sitting in on BLT meetings and asking reflective questions afterwards. These are not the only ways for DLTs to know if the district's BLTs are appropriately engaged with their school improvement responsibilities, of course. But they go beyond assumptions and rumor-and they can help the district model the wise use of data.

When BLTs are appropriately engaged with school improvement, DLTs will recognize the fact based on data. Of course, DLTs' own data capacity and leadership capacity must be equal to the task.

# DLT's Data Capacity

Ideally, the DLT will exercise a comparatively robust capacity for collecting, analyzing, interpreting, and reporting data. Multiple members will be familiar with and understand such procedures as survey design, test construction and reliability, significance testing, and research design. Ideally, at least one member will also understand methods of qualitative inquiry. Some DLTs rely for data-based decision-making on those of their members who are not intimidated by serious research and evaluation work; others can seek expertise elsewhere (e.g., from other districts, SSTs, or university partners).

## The District's Research and Evaluation Mission

DLTs must develop the research, evaluation, and data capacity of the team to support the work of improving teaching and learning districtwide. And most districts employ personnel with these skills. Matching the capacity to the need is the responsibility of the DLT. How can the capacity be developed?

- Senior district leaders must recognize the need to build research, evaluation, and data capacity. They must allocate resources to this mission.
- DLTs can exploit connections with higher education. Some faculty members in colleges of education would welcome the chance to collaborate with a district in an instructional improvement project.
- Leadership can call on state support teams (SSTs) and educational service centers (ESCs) to help support the effort of building capacity for working with data.
- The DLT can sponsor a professional development series that addresses the skills of research and evaluation for instructional improvement in the district context (open to anyone in the district, if possible).
- Finally, the DLT might initiate a study related to an instructional improvement strategy, perhaps using a controlled experiment or a multiple regression technique; the first project might not be very good, but the point is to build capacity rather than to get everything right from the get-go.

# DLTs' Cautious Use of Achievement and Graduation Test Data

Every district officially identified for improvement is identified on the basis of low test scores. And educators in every such district would like to see test scores increase so that the official designation is changed or removed.

Test scores ultimately help districts monitor the outcomes of their efforts over the mid- to long-term, but this module advises against *desperation* over test scores. The hard work of school improvement is guaranteed to address the issue . . . over time, not overnight.

DLTs need to foster this sort of caution among BLT and TBT members, and among all teachers and principals and central office staff. If the district (via its teams) builds collective efficacy, a more educationally productive culture will take shape. Such a culture-in the spirit of inquiry- takes a long view of its work, its identity, and its purposes.

# Using Data Wisely

Teams using data wisely:

- address presenting issues (e.g., low test scores) *in the spirit of inquiry;*
- *refuse to place blame* (on families, students, or each other) for the presenting issues;
- understand that *instruction is the center of attention;*
- *ask good questions* about instructional practice;
- get the data they need to address their questions about instructional practice;
- *apply their collective wisdom* (ideas and insights) to the data they gather;
- *work on their collective capacity* to gather and analyze data; and
- grasp their *teamwork as the key to the district's capacity for improvement*.

Here at the end let's acknowledge that *data-driven decision making* is a slogan based on a couple of misconceptions. It's important for teams to shed these misconceptions.

One is that educators *don't bother* much with data. To the contrary, educators generate and use a lot of data: and have done so for a very long time. Another is that the data themselves necessarily "drive" good decisions. It's not true—good judgment and insight drive good decisions: data can help, and that's all. Good data plus bad judgment is guaranteed to yield bad decisions!

In the collaborative work of Ohio's leadership teams, good data—especially information about teaching practice-support educators' observations, experiences, insights, and ideas. Good data are critically important for one reason: they support honest inquiry, which is the core of professional practice.

# References


Amrein-Beardsley, A., Berliner, D. C., & Rideau, S. (2010). Cheating in the first, second, and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives, 18*(14). Retrieved from https://epaa.asu.edu/ojs/article/view/714/841

Cameron, J. (2015). Give me a 3, tell me I'm effective, and leave me alone. In K. Sturges (Ed.), *Neoliberalizing education reform: America's quest for profitable market-colonies and the undoing of public good* (pp. 185-212). Boston, MA: Sense Publishers.

Cochran-Smith, M. (1991). Learning to teach against the grain. *Harvard Educational Review, 61,* 279-310.

Deming, W. E. (1986). *Out of the crisis.* Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

De Luca, C., Shulha, J., Luhanga, U., Christou, T. M., & Klinger, D. A. (2015). Collaborative inquiry as a professional learning structure for educators: A scoping review. *Professional Development in Education, 41*(4), 640-670.

Hannay, L., Ben Jaafar, S., & Earl, L. (2013). A case study of district leadership using knowledge management for educational change. *Journal of Organizational Change Management, 26*(1), 64-82.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school.* New York, NY: Teachers College Press.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York, NY: Routledge.

Hursh, D. (2008). *High-stakes testing and the decline of teaching and learning: The real crisis in education.* Plymouth, UK: Rowman & Littlefield.

Institute of Education Sciences. (2018). *The what works clearinghouse* [homepage]. Retrieved from https://ies.ed.gov/ncee/wwc/

Johnson, T. S. (2005). A lesson from Hans. *English Education, 38*(3), 160-166.

Love, N., Stiles, K. E., Mundry, S., & DiRanna, K. (2008). *The data coach's guide to improving learning for all students: Unleashing the power of collaborative inquiry.* Thousand Oaks, CA: Corwin. Pre-Assessment Instrument (Data Module)

Mandinach, E. B., & Gummer, E. S. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. San Francisco, CA: WestEd.

Ravitch, D. (2011). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.

Schütz, A. (1932). *The phenomenology of the social world*. (G. Walsh & F. Lenhert, Trans.). Evanston, IL: Northwestern University Press.

What Works Clearinghouse. (2018). *What we do* [webpage]. Retrieved from https://ies.ed.gov/ncee/wwc/WhatWeDo

# Appendix: Foundational Concepts

## Understanding Teacher Influence and Instructional Improvement

First, what we do in classrooms influences what and how students learn, as well as *how much* they learn. Why else would we bother doing what we do (teaching, collaborating, leading)? Teaching is hard work because we want students actually to learn.

Second, we realize that students learn many things, at different rates, and in many ways. Whatever influence teachers have works on something that is already in play with students because students are born to learn! In formal educational settings, however, teachers have a special concern: to help student think with words and numbers; language arts, social studies, math, and science are typical tools for this activity.

Third, the mission of teaching is to help students become better and better at thinking with words and numbers in all those realms (called “academics”). We teach academics not for their own sake, but for the sake of thinking.

Notice that these three points are a logical chain: from an assumption (teachers have influence) to interactions with intellectually active students (they are always learning) to a result (students thinking better with words and numbers). In short, the idea and the fact of improvement is built into this logical chain and into what teachers do!

There’s a lot of empirical proof that this is the case. The effect size (on student learning) for an average teacher is +.20 (Hattie, 2008). So, on average, teachers do have an effect on learning and students learn more! It’s good news. But as individuals and as a team, we can do even better (and we want to do even better) than average. And the way we get better is by teaching better. If our discussions and decisions improved our effect size from average to above average (say, an effect size of .30 or .40), we’ll have done good work.

“Effect size” isn’t a complicated or subtle calculation. It’s just one way to quantify (and then talk about) the size of an impact. Note that +.20 isn’t a very strong effect. The fact that ordinary teaching has an effect of that magnitude tells us that there is room for improvement in *every* school and *every* classroom. So, teaching is important; and it has the potential to improve everywhere.

Still, a misguided fixation on test scores is hard to stop. Indeed, one common (and unfortunate) answer to “how to raise scores” is to adopt a product or a branded practice, especially one said to be backed by research. Many products have shown

some effectiveness in relationship to test scores according to the Institute of Education Sciences's *What Works Clearinghouse* (2018). Using these products is what is usually meant by "evidence-based practice."

But using a product is just a beginning. No product does anything on its own, even if excellent studies show it to be somewhat effective. *All* implementation of products and practices requires the effort of teachers and students.

## The Spirit of Inquiry for Ohio's Collaborative Teams

Teaching can be endlessly fascinating, or it can be terminally deadening (see, e.g., Cameron, 2015). What will happen on any day? Today? Tomorrow? All next week? Staying curious about what's going on with teaching and learning is essential to good teaching. And our curiosity shows us just how much data we already have at our finger tips. Things happen with our students and with us, and we take note. All of this is part of the spirit of inquiry because it is *grounded in curiosity* (Cochran-Smith, 1991).

For exactly this reason, members of a great many TBTs and BLTs are already set to bring the interpretation of data into their collaborative work. The next step is to agree to see what happens next as they work on an instructional practice *together*. They become curious together about the next right decision. That decision is informed by data-often produced through data collection and analysis efforts they organize themselves (Mandinach & Gummer, 2016).

> Annual performance data aren't close enough to classroom practice to be much help. It's the wrong information, and having teachers fixate on the wrong information as if it were the right information is demoralizing. It undermines the spirit of inquiry.

Working together on a project to improve instruction, team members need to construct a common experience. Think about it. They are no longer working in isolation; they are trying to use the same practice and collect information about it (Bryk, Gomez, Grunow, & LeMahieu, 2015). As the team makes meaning from the information it collects, it builds the common experience the team needs (Love et al., 2008).

They might, for example, observe in one another's classrooms and then discuss what they thought they saw (their observations). Those "observations" are data. *Lots of tools* are available to make observations more consistent from person to person. Using such a tool would help move the discussion from "my opinion of what I saw" to something *shared more widely* by members of the team. The team *decides* they want to look at as an aspect of their teaching and *chooses* a tool to help them do it.

When teams get to this point (and it's not so easy to do this!), they are developing a shared *spirit of inquiry*. Here's the surprising thing. The spirit of inquiry reflects a team's willingness to let go of the idea that professionals are people who *already* know the answers.

They've asked a question about the issue that is important to them and their students. They've gathered data to help them make the next right decision. And they keep on in that way. The spirit of inquiry becomes a *disposition* that breathes life into the team's practice (see, e.g., Bryk et al., 2015; DeLuca et al., 2015; Love et al., 2008; Mandinach & Gummer, 2016).

A lot can get in the way, however. Ellen Mandinach and Edith Gummer (2016) explain one commonly encountered impediment to the wise use of data:

> Practices that suggest that annual test scores can be disaggregated by strands in the disciplines for individuals or groups of students misinform teachers about the appropriateness of data use in the classroom. The use of these summative test results is unfortunately what people often think of when assessments are mentioned, forgetting that there are many other sources of student performance data and other data that are much more informative to the instructional decision-making process. It is also part of the reason that data use has gotten a bit of a bad reputation, because it is being conflated with the overemphasis on testing and the use of wrong test results to make decisions. (pp. 60-61)

It's a well-known fact that many, and perhaps most teachers and school administrators are anxious about data or find the whole idea of "data-based decision making" threatening or distasteful (e.g., Dunn, Airola, Lo, & Garrison, 2013; Mandinach & Gummer, 2016). Part of the problem is that data have been used in the past in a spirit of blame. The

spirit of inquiry, however, is contrary to the spirit of blame. With a spirit of inquiry, Ohio leadership teams can collect, analyze, interpret, and use data to answer the questions *they* pose.

## References


Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.

Cameron, J. (2015). Give me a three, tell me I'm effective, and leave me alone. In K. Sturges (Ed.), *Neoliberalizing education reform* (pp. 185-212). Boston: Sense Publishers.

Cochran-Smith, M. (1991). Learning to teach against the grain. *Harvard Educational Review*, 61, 279-310.

De Luca, C., Shulha, J., Luhanga, U., Christou, T. M., & Klinger, D. A. (2015). Collaborative inquiry as a professional learning structure for educators: A scoping review. *Professional Development in Education, 41*(4), 640-670.

Dunn, K. E., Airola, D. T., Lo, W.-J., Garrison, M. (2013). What teachers think about what they can do with data: Development and validation of the data driven decision-making efficacy and anxiety inventory. *Contemporary Educational Psychology, 38*, 87-98.

Love, N., Stiles, K. E., Mundry, S., & DiRanna, K. (2008). *The data coach's guide to improving learning for all students: Unleashing the power of collaborative inquiry*. Thousand Oaks, CA: Corwin.

Mandinach, E. B., & Gummer, E. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. New York, NY, and San Francisco, CA: Teachers College Press and WestEd.


# Handling Data in Ohio Leadership Teams

Leadership teams are foundational to the Ohio Improvement Process. The teams meet to discuss issues, plan initiatives, and exert follow-through on the initiatives. At least in concept, teams lead their schools and districts in school improvement.

Some of what teams do involves handling data: accessing it, collecting it, analyzing it, and interpreting it. Team members apply their collective insights and judgments informed by data to determine a course of action.

Little guidance exists to help teams handle data. It's a foundational concept in need of elaboration. This Foundational Concept is just a beginning.

## Data Analysis and Data Displays in Leadership Teams

In TBTs, BLTs, and DLTs data analysis involves looking at data displays and discussing what the team sees. The team can't waste time sifting through reams of data or puzzling out how best to present the data to one another. Instead, a team member or a small group can prepare the necessary data displays to bring to the whole team.

These displays can take many forms. The discussion here considers displays for quantitative data (responses to survey items, rubric ratings, test scores, and so on) and qualitative data (comments on surveys, interview data, and some observation data).

## Quantitative Data Displays

The most basic display is called a frequency table. Let's say the team has given a short survey of 10 items to 100 students and that 88 of the students responded. The survey used the familiar 1-to-5 response scale, with 1 being "strongly disagree" and 5 being "strongly agree." Here's what a frequency table for one of the 10 items might look like:

### I really like to read.

| response category | frequency | percentage |
|---|---|---|
| 1 (strongly disagree) | 7 | 8% |
| 2 (disagree) | 18 | 20% |
| 3 (don't agree or disagree) | 37 | 42% |
| 4 (agree) | 20 | 23% |
| 5 (strongly agree) | 6 | 7% |
| Missing | 12 | (not included) |

The display could be presented on a slide, and it could serve many purposes depending on what the team needs. (For calculating the percentage, the missing responses aren't included. In other words, to find the percentage, the frequency is divided by the total number of people who actually responded, in this case 88 not 100.)

Cross-tabulations (cross-tabs) are a variation on frequency tables. Let's say the team is curious about the differences between female and male students. Cross-tabulations report frequencies for the two groups, like this:

## I really like to read.

| response category | females | males | total | total percentage |
|---|---|---|---|---|
| 1 (strongly disagree) | 2 | 5 | 7 | 8% |
| 2 (disagree) | 7 | 11 | 8 | 20% |
| 3 (don't agree or disagree) | 22 | 15 | 37 | 42% |
| 4 (agree) | 14 | 6 | 20 | 23% |
| 5 (strongly agree) | 5 | 1 | 6 | 7% |
| missing | 4 | 8 | 12 | (not included) |

Team members who prepare the display might swap out "total percentage" with a different calculation, perhaps the percentage of the total in each response category that were females (see below).

## I really like to read.

| response category | females | males | total | percentage of females in the category |
|---|---|---|---|---|
| 1 (strongly disagree) | 2 | 5 | 7 | 29% |
| 2 (disagree) | 7 | 11 | 18 | 39% |

| response category | females | males | total | percentage of females in the category |
|---|---|---|---|---|
| 3 (don't agree or disagree) | 22 | 15 | 37 | 60% |
| 4 (agree) | 14 | 6 | 20 | 70% |
| 5 (strongly agree) | 5 | 1 | 6 | 83% |
| Missing | 4 | 8 | 12 | (not included) |

Note that, for dramatic effect, this example stereotypes males as much less enthusiastic readers.

Another way to present the same information is to report averages ("means," in the language of statistics). The advantage of this approach is that means can be reported in much less space.

To summarize the information in the frequency display and the cross-tabs displays above, we can be more efficient by showing the means.

| Items | Means | | |
|---|---|---|---|
| | all | females | Males |
| Item Ten: I really like to read. | 3.00 | 3.26 | 2.73 |

In fact, we could present data summarizing responses to all 10 items in just a little more space than we needed for presenting the cross-tabs for just one item. So, why not use means all of the time? What's the disadvantage of this approach to displaying data?

When we use means, we lose the detail about particular responses: how many were low, how many in the middle, and how many high. Still, we can format a table of means to include data that will help us make comparisons. For example, we can present means for several different subgroups in one display. For example, we might divide groups of students by classroom, by proficiency level, by gender (as in the above example) and

so forth. We could then calculate and display means for each subgroup. A team could address lots of different questions using this approach.

Displays of means are also useful for tracking changes over time. Let's say the TBT adopts a program to help boys engage better with reading and decides to use the 10-item survey four times over the course of the year to measure their attitudes towards reading (as well as the attitudes of the girls). Again, let's keep our focus on just one item. Here's what the display for the data after all four administrations might look like:

| **Items** | **Means** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | First | | Second | | Third | | Fourth | |
| | F | M | F | M | F | M | F | M |
| Item Ten: I really like to read. | 3.26 | 2.73 | 3.30 | 2.72 | 3.32 | 2.89 | 3.35 | 2.98 |

This is actually a relatively complicated display for a leadership team! Often, the displays they develop and use are simpler. But this one is useful for examining trends and perhaps for helping make instructional decisions. What are the helpful features of this display?

- It presents outcome data for a year-long effort.
- It provides four measures over the course of the year.

Two conditions would make the display even more useful:

- if data come from an instrument with established reliability and validity) and
- if the item for which it displays data is highly associated with overall scores on the 10-item scale.

But the display doesn't show everything. In particular, it lacks information about variability in ratings. What a shame to lose that information!

In addition to means, those preparing data displays can include standard deviations. Think of the mean. It's the average, but it's produced by combining all of the scores—and those scores reflect a range: low, middle, and high. The standard deviation is a measure of how far those scores are from the average. Here's what a display with both means and standard deviations (SD) might look like:

| Items | Means (SDs in parentheses) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | First | | Second | | Third | | Fourth | |
| | F | M | F | M | F | M | F | M |
| Item Ten: I really like to read. | 3.26 | 2.73 | 3.30 | 2.72 | 3.32 | 2.89 | 3.35 | 2.98 |
| | (.96) | (.89) | (.97) | (.92) | (.96) | (1.2) | (.99) | (1.3) |

In this case, the ratings show improvement overall, with larger improvement for the boys. But for the boys, even though scores increased overall, the standard deviation also grew notably larger over time. This means that for some of the boys, there may have been no improvement at all, and for some, attitudes may actually have deteriorated. The increasing variability in scores tells us to dig deeper to figure out what might be going on.

These examples barely scratch the surface. Lots of other data analysis methods and data displays might also be useful. And they would address some of the other issues a team might want to understand. What are some of these issues?

- Is the difference we see in reading attitudes of males and females real? How large a difference would it have to be to be real?
- What do other subgroup comparisons look like (for instance, comparisons based on race, comparisons based on poverty)?
- What do the students whose attitudes remained low over the course of the year have in common?
- Is the relationship between reading proficiency and reading attitudes strong or weak?

## Qualitative Data Displays

It's too bad that most talk about data concerns numerical data: survey responses and test scores (of various sorts). It's only half the picture. Qualitative data comprise the other half.

Qualitative data tend, however, to be more unruly than numbers. We cannot just "add up" words or calculate an average.

Nevertheless, gathering data through the words of students, teachers, and parents can tell us a great deal about what they think, believe, or want. Their words can be data when teams treat them as data. Qualitative data come to teams in three basic ways: (1) as brief comments in response to open-ended items on surveys, (2) longer comments from

interviews, and (3) notes from observations.

Although no one can use numerical procedures (like addition) on words, people have been summarizing verbal material for thousands of years. Teams can use various options for boiling down and displaying data that start out as words.

For some examples, let's assume—as part of a long-term project related to fostering community engagement—a BLT has surveyed 250 parents to learn more about their opinions about the school and that the BLT has also conducted interviews with 25 of the parents. The team now has written responses to the surveys and transcriptions of the recorded interviews.

The survey included an item that asked:

> What would you most like to see the school do in the community?

The interview, consisted of five questions, one of which was:

> How would your family use community tutoring services if the school district could provide them?

The survey drew comments from 43 people who chose to respond to the open-ended item. In the interview, all 25 people provided comments in response to the question about how families would use community tutoring services.

How might a member of the BLT prepare these data to display them to the team?

The BLT member might just give the team all 43 comments from the survey. Of course, reading all of the comments might take too long and wouldn't tell the team much about trends.

Alternately, the BLT member might group the responses into the ideas that were mentioned most often, and list those on a table such as the one below.

| Category | Example | Number of mentions |
|---|---|---|
| Provide tutoring (especially at the library) | Math tutoring, maybe at the library or mall | 12 comments |
| Expand hours for after-school care | 5 PM is too early to end after-school care! | 7 comments |
| Improve website | website info out of date | 6 comments |

Clustering survey comments is good practice for summarizing interview material. Working with more extensive interview material—as with the transcripts produced in this example—can be time-consuming. Of course, if it's the only way to answer certain kinds of questions, it's probably worth the effort.

In this case, the interviews were used as a way to get a deeper understanding of what different families meant when they indicated their preferences for the various community tutoring services the school district might provide. So, data analysis would involve looking for patterns in the insights offered by parents. Organizing their insights into categories helps educators spot patterns. A display that looks for patterns based on family type (for instance, families with young children, families with children of various ages, families with mostly older children) might look something like the grid below. Color coding might help those analyzing the data identify possible patterns.

## How would your family use community tutoring services if the school district could provide them?

| Type of family | Insights | Number of comments revealing a similar insight |
|---|---|---|
| Family with mostly young children | Headstart and church groups are already offering tutoring and enrichment programs for young children | 5 |

| Type of family | Insights | Number of comments revealing a similar insight |
|---|---|---|
| | Coming to the school may be an issue for some families. An off-school location might help draw more families. | 3 |
| | Tutoring of preschool students would be especially helpful to prepare them for kindergarten | 2 |
| Family with children across all grade levels | Tutoring of elementary school and middle-school students would be especially helpful | 4 |
| | To make it possible for students to participate in tutoring, they would need school-bus transportation | 2 |
| | I'd love to see a program that would show my high school children how to help their younger sibs with their homework | 1 |
| | Tutoring ought to take place during the school day. My children have chores and activities after school. | 1 |

| Type of family | Insights | Number of comments revealing a similar insight |
|---|---|---|
| Families with middle school and high school students | Coming to the school may be an issue for some families. An off-school location might help draw more families. | 4 |
| | Middle school students really need to have access to after-school tutoring | 3 |
| Families with mostly older children | The community college offers tutoring services for high school students | 3 |
| | High school students definitely need tutoring that prepares them for college-entry tests | 3 |
| | What about peer tutoring? I like that idea. | 1 |

Note that the color coding points to one likely pattern, namely that families of all types believe that tutoring services would be beneficial. Other possible patterns include the insight that various agencies also provide tutoring and that various types of peer tutoring might be helpful. The remaining insights also raise issues that relate to how students might access tutoring services—where the services are located, how students can get there, and when they're provided. Like quantitative data, these qualitative data give educators a lot of information on which to base decisions about services that would help the district make better connections with families.

## Making Data Displays for the Team

Making data displays in advance of team meetings will save time and it will represent a major contribution to data analysis. There's lots more to learn, but it's clear that a team member with a feel for graphic organizers and infographics would be a big asset to a team.


OLAC
Ohio Leadership
Advisory Council