

CPRE

CONSORTIUM *for* POLICY
RESEARCH *in* EDUCATION

Leading Improvement in Challenging Times

An introduction to tools & skills for educational
leaders to adapt in times of crisis

Jonathan A. Supovitz

John D'Auria



The Consortium for Policy Research in Education is a leading hub for innovative and collaborative research that transforms education policy and practice through knowledge.

About the Consortium for Policy Research in Education

At the Consortium for Policy Research in Education (CPRE), we push the boundaries of knowledge to inform, shape, and lead the education community. Headquartered at the University of Pennsylvania Graduate School of Education, our nationwide consortium of experts is committed to the advancement of educational policy and practice through rigorous research and evidence-based resources, designed to improve student achievement and move education forward.

Our Mission & Vision

CPRE is a leading hub for innovative and collaborative research that transforms education policy and practice through knowledge. Our experts understand that knowledge shapes progress and that research and evaluation of education programs and policies are essential to improving the field. CPRE contributes to the production and dissemination of high-quality research, and fosters conversations that create opportunities to advance the profession.



This guide is a companion to a training video called [Leading Improvement in Challenging Times](#). In the video series we use the story of a school leadership team planning for school opening in the fall of 2020 to illustrate the key elements of the Leadership for Learning Framework. By examining the interactions and decisions of the leadership team, illuminated by expert commentary, we highlight the essential conditions, improvement processes, and leadership skills that make up the Leadership for Learning Framework.

The Leading Improvement in Challenging Times training video and guide was funded by **The Harold W. McGraw, Jr. Family Foundation**, a not-for-profit foundation focused on supporting education, youth services, community health, and medical research.

Leading Improvement in Challenging Times

An introduction to tools & skills for educational
leaders to adapt in times of crisis

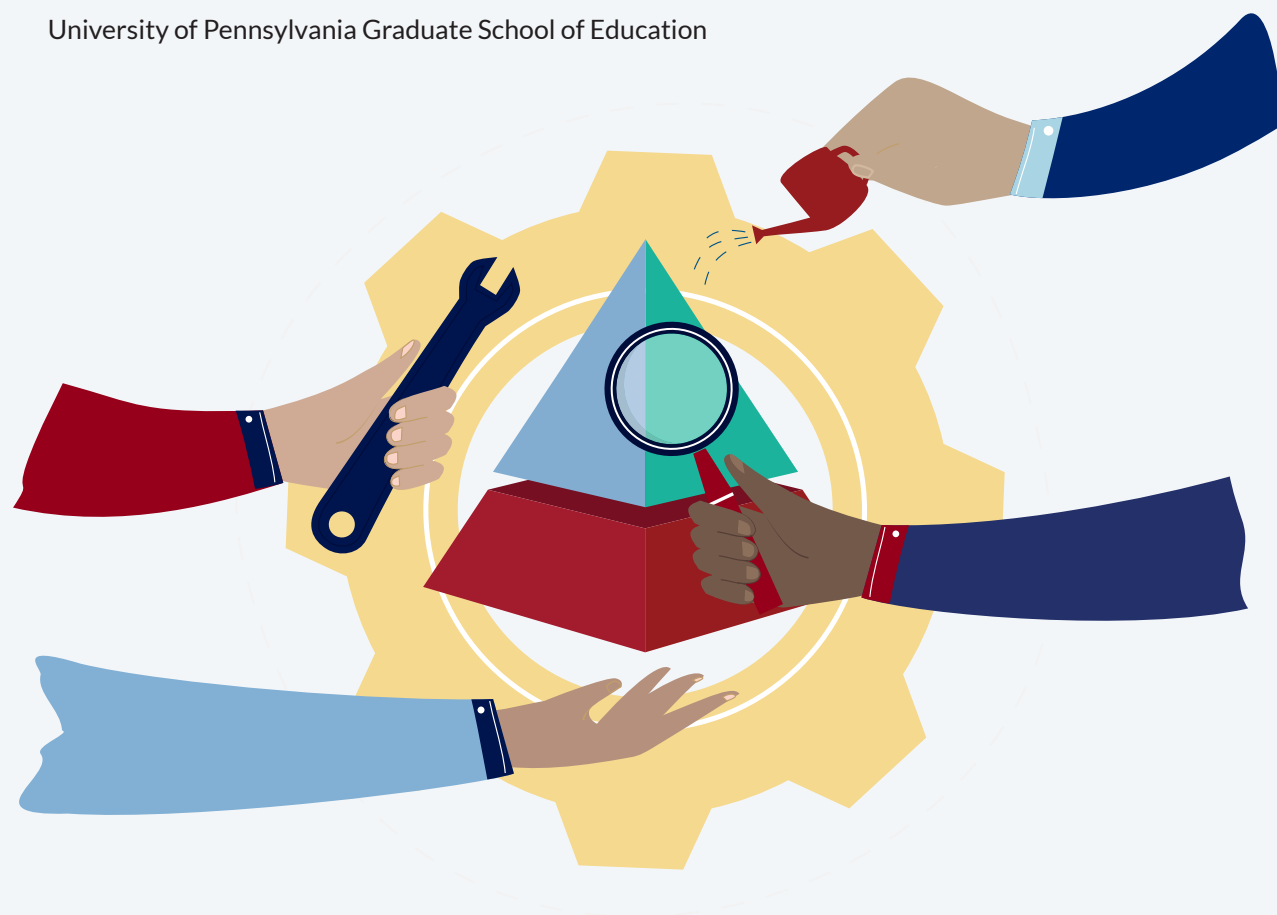
Jonathan A. Supovitz (jons@upenn.edu)

University of Pennsylvania Graduate School of Education

John D'Auria (dauriaj@upenn.edu)

William James College

University of Pennsylvania Graduate School of Education



Author Biography



Jonathan Supovitz, Ed.D.

Jonathan Supovitz is Professor of Education Policy and of Teaching, Learning, and Leadership at the University of Pennsylvania Graduate School of Education Penn GSE. He also is the Executive Director of the Consortium for Policy Research in Education (CPRE). Dr. Supovitz conducts research on how education organizations use different forms of evidence to inquire about the quality and effect of their systems to support the improvement of teaching and learning in schools. Dr. Supovitz also leads the evidence-based leadership strand of Penn's mid-career leadership program and teaches courses on how current and future leaders can develop an inquiry frame of thinking about continuous improvement and the skills necessary to compile, analyze, and act upon various forms of evidence.

While studying policy analysis at Duke University, Dr. Supovitz first focused on education leadership and policy. Before earning his doctorate at Harvard's Graduate School of Education, he gained middle and high school teaching experience in Queretaro, Mexico, and Boston, Massachusetts. His dissertation at Harvard focused on the classroom and accountability uses of portfolio assessment in an urban school district. Upon completing his degree, Dr. Supovitz worked as a research associate at Horizon Research in Chapel Hill, North Carolina, where he directed the evaluation of the New Jersey Statewide Systemic Initiative and evaluated the effectiveness of electronic "netcourses" for teacher enhancement. He joined the Consortium for Policy Research in Education (CPRE) in 1997 as a senior researcher and the faculty at Penn GSE in 2005.

He has published findings from a number of educational studies, including multiple studies of programmatic effectiveness, studies of educational leadership, research on the development of instructional practice communities in schools, an examination of the equitability of different forms of student assessment, the use of technology for evaluative data collection, and the relationship between data use, professional development, teacher and leadership practice, and student achievement. His current research focuses on how schools and districts use different forms of data to support the improvement of teaching and learning. He also leads the Evidence-Based Leadership strand of the Mid-Career Leadership Program at the University of Pennsylvania.



John D'Auria, Ed.D.

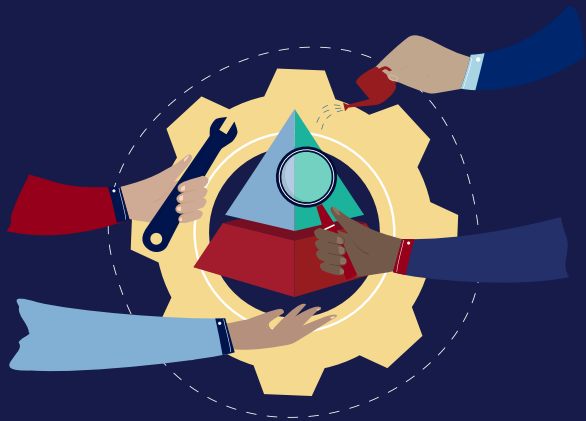
Dr. D'Auria's research focuses on the ways in which the assumptions that people hold about intelligence significantly influence their learning. He co-authored *School Systems That Learn* with Dr. Paul Ash (Corwin Press, 2012) and is the author of *Ten Lessons in Leadership and Learning* (2010), a resource geared toward new and experienced leaders. Dr. D'Auria authored a curriculum for aspiring school leaders called "The DNA of Leadership," which became a cornerstone for the Leadership Licensure Program sponsored by the Massachusetts School Administrators Association. Additionally, Dr. D'Auria co-authored *How To Bring Vision to School Improvement* (Research for Better Teaching, 1993) with Dr. Jon Saphier. He is a frequent speaker at national and regional educational conferences and has served as an executive coach to a wide variety of educational leaders across the country.

In addition to his current position as a member of the core faculty in Organizational and Leadership Psychology at William James College, Dr. D'Auria is an adjunct professor for the Mid-Career Doctoral Program in Educational Leadership Ed.D. at Penn GSE.

Table of Contents

Overview	6
About the Guide to Leading Improvement in Challenging Times	7
The Learning for Leadership Framework.....	8
Essential Conditions.....	9
The power of conditions	9
Critical conditions for engaging in improvement	9
Psychological Safety.....	9
Trust.....	10
Accountability	11
Creating a culture of mutual learning	11
The Fractal Nature of Schooling	12
Improvement Process.....	13
Understanding the problem.....	13
Involving Multiple Perspectives	14
Recognizing the System around the Problem.....	16
Plan-Do-Study-Act Cycles.....	18
Design, Measurement, and Analysis.....	19
Design	19
Measurement.....	19
Analysis	20
Leadership Skills	21
Navigating power differentials	21
Addressing non-discussables.....	21
Listening in stereo	22
Managing emotions	23
Curiosity in the face of criticism and wrong sounding ideas.....	23
Balancing inquiry with advocacy.....	24
The discomfort of Disagreement	24
Leading Improvement in Challenging Times	25
Appendix A: Study Designs for Intervention Studies	27
Appendix B: Quantitative Analyses for Intervention Studies	28

Guide to Leading Improvement in Challenging Times



This guide is a companion to a training video called [Leading Improvement in Challenging Time](#). In the video series we use the story of a school leadership team planning for school opening in the fall of 2020 to illustrate the key elements of the Leadership for Learning Framework. By examining the interactions and decisions of the leadership team, illuminated by expert commentary, we highlight the essential conditions, improvement processes, and leadership skills that make up the Leadership for Learning Framework.

Overview

In a time of global warming, where weather events are increasingly severe, communities across the United States have become used to intense storms that at any time can knock out the power, down the trees, and flood the levees. But nothing in modern times has compared to the cataclysm of March 2020, when the world turned upside down, inside out, and left-side right all at the same time. As the coronavirus pandemic swept across the nation, schools shuttered their doors, and students, teachers, and leaders quarantined in their homes.

The spring was a struggle for school systems across the country as they confronted a host of challenges and makeshift conditions. While schools commonly had their own computers and educational technology infrastructure, districts were not equipped for every student to have their own device and Internet at home, and teachers were not prepared to provide education solely through bits and bytes. As a consequence, many schools in the United States barely got to instruction, as it quickly became clear that their first priority was to provide more basic social services. Districts had to figure out how to distribute tens of thousands of meals to students and families and to provide other basic socio-emotional and psychological services that are now regularly delivered through schools. The circumstances laid bare two important disparities within the education system; first, between districts with different levels of demand for these kinds of services; and second between the array of services that education systems are expected to provide and the outdated conception that schools are only responsible for teaching and learning.

Now, educational leaders are faced with a fundamentally new landscape. Some schools have tentatively gone back to in-person schooling, many have continued to be solely online, while many others are some hybrid of face to face and online schooling. Planned solutions today may not meet the needs of tomorrow, as the menace of coronavirus resurgence is looming. Consequently, we can't rely on the education system that have prevailed for generations. We need new solutions to new problems. The way to face these challenging times it to be flexible and adaptive and lead for learning.

These circumstances challenge school leaders to continually question and rethink the basic tenets of the educational experience for students and faculties. How will we reconstitute and improve upon interpersonal educational experience in an online or hybrid environment? What are the affordances and limitations of technology? How do we recognize, rethink, and rectify the systemic inequities that are embedded within in the education system? How do we support the socio-emotional needs of students and faculty members in a distance environment? There are so many questions to answer and so little experience to draw upon. To meet these challenges, schools must do what they ask students to do every day: learn.

About the Guide to Leading Improvement in Challenge Times

This booklet introduces educational leaders to the Leadership for Learning Framework. The Framework has three main components.

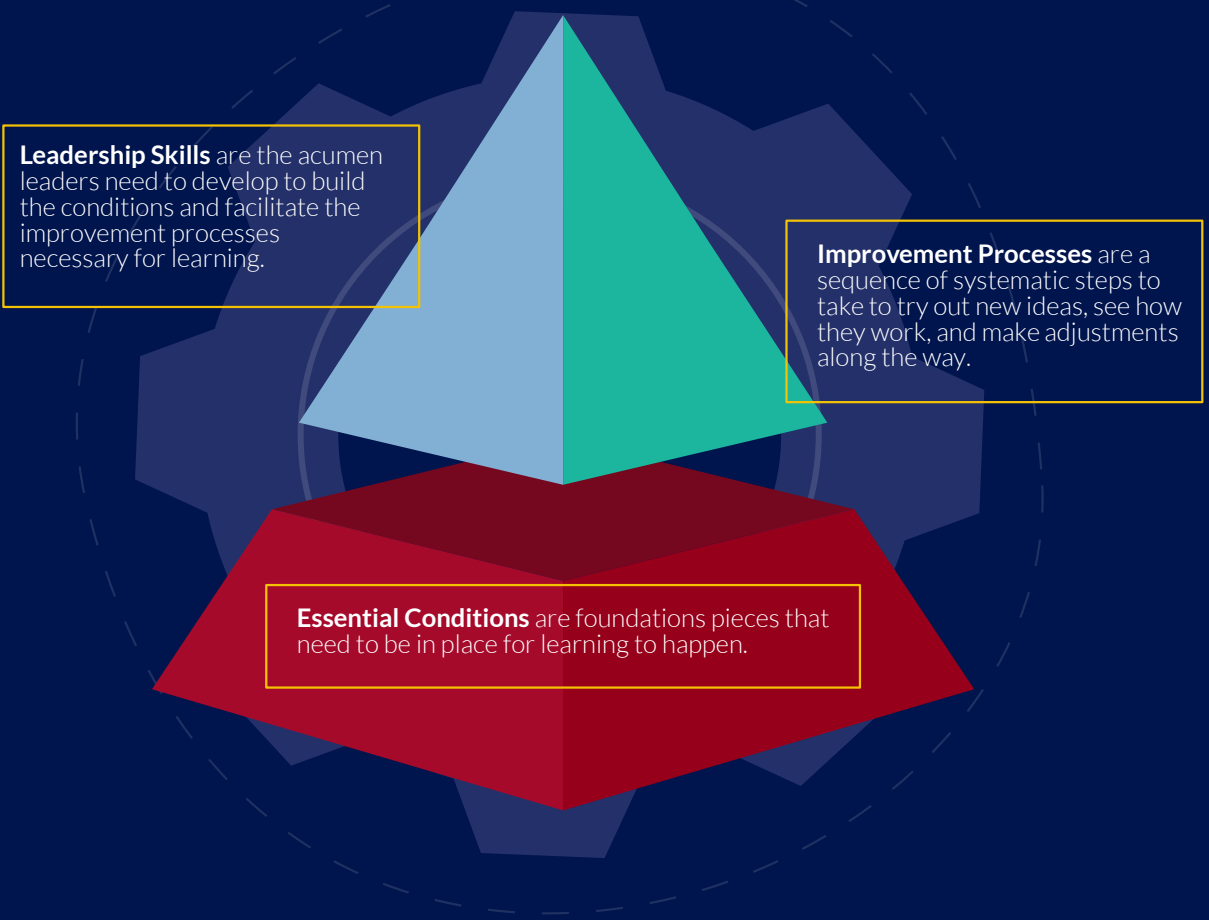
- A set of **essential conditions** that schools must have in place to engage with these challenging times.
- Systematic **improvement processes** that allows teams to identify the core problems that impede school effectiveness, and engage in a process of developing hypothesized solutions, testing those hypotheses, learning from the results, and iterating that learning into a more refined approach.
- A collection of **leadership skills** with which to navigate the challenges of uncertainty and guide the organization towards learning and improvement.

Herein we provide a brief overview of the essential conditions, improvement processes, and leadership skills that make up the Leadership for Learning Framework. As you read them, consider a few points:

- While we describe the subcomponents of the framework individually, their real power comes from integrating them situationally. Improvement processes are only likely to happen if they are nested within conducive conditions and nurtured by skillful leadership. Similarly, setting conditions and employing leadership skills are useful only insofar as they are applied for meaningful purposes.
- Understanding the concepts within each of the essential conditions, improvement processes, and leadership skills are important, but only the first step. Naming and describing them provides you with touchstones to recognize them in situations in your setting, but practicing them consistently and constructively takes persistent practice and reflection.
- The troika of conditions, processes and skills are useful whether you are developing your own change efforts or implementing existing programs. Regardless of what you are trying to do, every adjustment has to fit your own setting and work for your community.
- The Leadership for Learning Framework is well-aligned with using a distributed leadership approach in schools. A distinct advantage of distributed leadership is that it provides an organizing principle for selectively involving more members of the school community in the improvement process and, in doing so, gaining both more diverse perspectives into the underlying causes of challenging problems and a shared commitment to the solutions that emerge. Distributed leadership can help to inform who should participate in the essential components of continuous improvement. In this way, distributed leadership is an essential companion to the continuous improvement processes that are increasingly recognized as the ways to make headway on impediments to consequential school improvement. For more information see our report on [Meaningful & Sustainable School Improvement with Distributed Leadership](#).

The Learning for Leadership Framework

An introduction to tools & skills for educational leaders to adapt in times of crisis.



Essential Conditions

- Psychological Safety & Trust
- Creating a Learning Culture
- Accountability
- The Fractal Nature of Schooling

Improvement Processes

Understanding the Problem

- Involving Multiple Perspectives
- Recognizing the System around the Problem

Engaging in Improvement

- Plan-Do-Study-Act Cycles
- Design, Measurement, and Analysis

Leadership Skills

- Navigating Power Differentials
- Addressing Non-Discussables
- Listening in Stereo
- Managing Emotions
- Curiosity in the Face of Criticism
- Balancing Inquiry and Advocacy

Essential Conditions

In this section we provide an overview of the essential organizational conditions that form a basis for organizational members to engage in respectful and constructive discussions about the challenges they face.

The power of conditions

Though we like to think we fully control our interactions with others, we don't. Our interactions are highly influenced by a variety of context-dependent conditions. These conditions, mostly taken for granted, are part of every organization's *social structure*. Social structure includes a range of things that mediate our interactions, including the social norms that govern expected behavior, the organizational routines which guide much of our activity, the agendas and protocols we choose to use, and even the language we use to communicate, to name just a few. These conditions do not simply influence our interactions; we could go so far as to say they *largely define* how we interact with one another, because they do so much to constrain the bounds of our conduct and efforts. For this reason, increasing our awareness of the social structures within which we operate is integral to leadership practice and its effectiveness.

To appreciate the extent to which social structures define our interactions, let's go back to the last time you performed a required evaluation of a teacher's practice. If you reflect on this experience, you will come to appreciate that the practice of evaluating the teaching takes form as it unfolds in the interactions between the supervisor and the teacher, but these interactions are also fundamentally defined by taken-for-granted aspects of the situation. Most obviously, the teacher evaluation protocol you used not only shaped what was paid attention to but also what the teacher expected the supervisor to attend to. Moreover, as a leader conducts a post-observation interview/debrief with the teacher, items from that protocol were likely used to negotiate understanding of what was noticed and what it said about the quality of the teaching. At the same time, the interactions were very likely fundamentally shaped by a set of norms that neither you nor the teacher explicitly named – perhaps something as simple as: “Begin the debrief by describing what you saw, rather than making a value judgment.” More complex norms could also guide whether a teacher feels safe to illuminate where she thought the lesson could be improved, and even guide both of your definitions of what improvement might look like. This is how the conditions define practice by shaping how we interact with one another.

Critical conditions for engaging in improvement

While there are a host of conditions that influence our social and professional interactions, we think five are especially pertinent for leaders to create the environment where people can engage deeply with the challenges of substantive improvement. Below we provide a brief overview of the essential conditions.

Psychological Safety

The foremost essential condition for high quality interactions amongst adults within professional settings is *psychological safety*. According to Harvard Business School professor Amy Edmondson, an expert on teams in organizations, psychological safety is the belief that one will not be punished or humiliated for speaking up with ideas, questions, concerns, or mistakes.¹ The degree to which

¹ Edmondson, A. (2012). *Teaming: How Organizations Learn, Innovate, and Compete in the Knowledge Economy*. San Francisco: Jossey-Bass Pfeiffer.



educators are able to share vulnerabilities, acknowledge mistakes, respectfully disagree, and challenge the thinking of colleagues as well as those with more status and power provides insight into the perceived level of psychological safety within an organization. One of the key ways that a leader can contribute to psychological safety is when he or she chooses to use influence more than authority to gain educator *commitment*. Influence is more about engagement, collaboration, and building the trust and commitment to improve rather than using one's authority and expecting compliance to a new set of expectations.

This leadership emphasis was captured vividly in the words of Alfred P. Sloan, the CEO of General Motors in the 1940s and 1950s: "I never give orders. I sell my ideas to my associates if I can. I accept their judgment if they convince me, as they frequently do, that I am wrong. I prefer to appeal to the intelligence of a man rather than attempt to exercise authority over him."²

Trust

At their core, the dynamics of schooling are based on interdependent social exchanges, whether they be amongst adults, amongst students, or amongst adults and students. Relational trust is the oil that facilitates these social exchanges. Bryk and Schneider illuminated the importance of relational trust in school improvement efforts.³ Trust impacts the quality of relationships between students and teachers, teachers and administrators, and educators and parents. In turn, the quality of those relationships shapes communications – how open people are to feedback, how willing they are to share their ideas and perspectives, and the respect and personal regard one feels and is willing to give to others through careful and deep listening.

Consequently, trust is a linchpin for developing a healthy and vital school culture and moving a school forward. Without sufficient trust, improvement efforts often stall. As Bryk and Schneider summarize their work, "Strong relational trust also makes it more likely that reform initiatives will diffuse broadly across the school because trust reduces the sense of risk associated with change. When school professionals trust one another and sense support from parents, they feel safe to experiment with new practices."⁴ Additionally, when we examine turnaround efforts in schools we see an accelerated agenda of change. When those efforts are unsuccessful, it is often not a function of an ineffective strategy but an insufficient investment in building trust within the community.⁵



2 Sloan, A.P. (1925). *Industrial Digest and Commodities and Finance*, (1925), Vol. 4. p. 16

3 Bryk, A.S., and Schneider, B (2002), *Trust in Schools: a core resource for improvement*. N.Y.: Russell Sage Foundation, New York

4 Bryk, Anthony S., and Barbara Schneider. "Trust in schools: a core resource for school reform." *Educational Leadership*, Mar. 2003, pp. 40-44.

5 University of Chicago. "Lack Of Trust Leads To Dysfunctional School Systems." *ScienceDaily*. ScienceDaily, 27 August 2008. <www.sciencedaily.com/releases/2008/08/080827164035.htm>

Accountability

How to balance accountability with psychological safety is one of the challenges school leaders face. Supporting continuous improvement requires a safe zone where educators can learn from mistakes and setbacks. At the same time, standards of quality must also be maintained. Edmondson suggests that a balance can be achieved if the central focus of accountability is on learning. This means our work, our interactions, our reflections need to lead to learning over time, and if we fail to learn from our work, we are not meeting an important standard. Similar to scientists and medical researchers, it is not a failed experiment or lesson that leads to accountability, but an inability to learn from that failure. In a learning organization, lessons will fall flat, meetings will be unproductive, initiatives will miss their targets. What is key, however, is not simply the data that emerges from that work but how we gain insight into future directions and growth by learning from what has not worked. Those who fail to learn over time must be held accountable. Additionally, disruptive behaviors that are reckless as well as behaviors that violate established group and behavioral norms also must be addressed, but again it is not simply a single violation that is at the heart of the matter but the inability to learn from one's mistakes and from the feedback one receives related to the gap between our intentions and our impact. This is how we can balance safety with the need for accountability. If we get the balance right, we avoid making people overly anxious or allowing our comfort needs to impede the need to take on creative risks and continually improve.

Creating a culture of mutual learning (instead of a culture of blame)

“Leadership is the ability to guide others without force into a direction or decision that leaves them still feeling empowered and accomplished.”

– Lisa Hanson, CEO

Social scientist Chris Argyris describes how we all grow up developing ways to approach stressful situations. Often those approaches or mental models involve a set of rules that influence our actions and help us interpret the actions of others.⁶ When educators and educational leaders tackle difficult issues, particularly around school reform initiatives, those mental models are often on full display. Argyris' work showed that a common approach in the face of stress involves typical behaviors to help us remain in “unilateral control.” Typically, we try to maximize winning and minimize losing, suppress negative feelings, and be as rational as possible. Argyris points out that the purpose of these behaviors is to avoid vulnerability, risk, and embarrassment. Additionally, in order to protect ourselves from failure and the appearance of incompetence, we often resort to blaming others, deflecting any responsibility away from ourselves. Teachers often blame administrators, students, or parents when school initiatives fail. Administrators often blame teachers or parents or central office to explain lack of progress. This protective set of strategies ultimately arrests learning. Argyris points out that we can learn new strategies and update our mental models. In particular, we can become adept at a mutual learning⁷ approach. One of the key aspects of a mutual learning framework is the idea that *I might be contributing to the problem*. If a group of educators trying to address what appears to be an intractable problem grounded their discussions in the assumption that each of them might be contributing to the problem, defensiveness and blame would be reduced and the potential for learning would increase. Too often, we look outward initially rather than inward and, consequently, we do not own our part of the problem. Imagine if educators addressing a challenging issue such as low attendance or poor growth in mathematical understanding examined a wide range of factors

6 Argyris, C. “Good Communication That Blocks Learning.” Harvard Business Review, July-August, 1994, pp. 78-85

7 Schwarz, R. M. (1994). The skilled facilitator : practical wisdom for developing effective groups. San Francisco: Jossey-Bass Publishers.

including a discussion of, “How might I (the teachers, the administrators, the support staff) be contributing to this problem?” This is not an easy mindset to achieve and it involves a combination of all three factors: sufficient psychological safety, relational trust, and a mental model of mutual learning. When all of these factors are in place, the environment is rich for learning.

The Fractal Nature of Schooling

I’ve yet to see a school where the learning curves of the youngsters are off the chart upward while the learning curves of the adults are off the chart downward, or a school where the learning curves of the adults were steep upward and those of the students were not. Teachers and students go hand in hand as learners - or they don’t go at all.

-Roland Barth, *Learning by Heart*

Barth’s observation is a keen one, but school improvement efforts often miss this essential link between adult learning and student learning. We are not going to improve our schools for our students if we do not support the learning of the adults. Educators need to continuously adapt and tinker with their pedagogical and curricular strategies in order to improve their effectiveness with students. This vital connection between adult learning and student learning illuminates the fractal nature of schools.⁸ A fractal is a term developed by mathematicians that refers to a self-similar pattern; that is, individual or small segments of the entire design resemble the whole. Nature abounds with fractals. Broccoli’s shape is a fractal. If one breaks off a small floret from a broccoli plant it looks like a miniature version of the entire edible portion of the plant.

One example of an educational fractal is how a teacher shapes a classroom culture that either encourages or discourages students to leave their comfort zone, practice and learn new skills, and try out important ideas. If that culture is harsh and shaped more by fear than encouragement, the quality of student engagement in learning will diminish. Similarly (in a fractal manner), the adult educators in effective schools must constantly experiment with new approaches in order to improve results. How enthusiastically and effectively educators will embrace leaving their comfort zones in order to try new approaches is linked to the culture and climate established by the principal. If the culture is threatening and intimidating, educators will shy away from robust attempts to discover new strategies and approaches to teaching within their classrooms. Principals in turn are impacted by the culture established by the superintendent and district administrators. Central office leaders who effectively develop and manage principals also understand that they must encourage and support a principal’s learning and provide encouragement for thoughtful experimentation with leadership strategies, if the district schools are going to improve.

The fractal nature of schools also implies that negative effects can also be distributed throughout the system. In a district where the superintendent blames the principal for low-test scores, that pattern of blame will often replicate itself in the way a principal addresses weak scores within the classrooms of their teachers. Moreover, the pattern of a principal blaming the teachers may too often lead to teachers blaming students or colleagues or parents for their students’ lack of progress. Because of the fractal nature of schools, the pattern of blaming others for the gaps will be reinforced within the system and momentum and energy will be lost, thus slowing down improvement efforts.

While no system will be perfectly aligned all the time, the way each segment influences the whole system is a dynamic that needs to be monitored regularly. Acknowledging the connection between adult learning and student learning increases the ways that schools can improve.

⁸ Ash, P, D’Auria, J, *Schools Systems That Learn*, Corwin Press, 2012

Improvement Processes

In this section we describe the key elements of continuous improvement. Continuous improvement is the essential means for learning inside of organizations: efforts to deliberately and systematically introduce change into schools, districts, and even classrooms, and then to monitor what happens and learn from the results are essential for learning. First we examine the importance of spending time on diagnosing a problem or challenge, and then we describe a strategy for deliberately introducing change and building a system to capture the effects of the change through improvement cycles. While changes may or may not improve things, the real benefit is that introducing improvement cycles will help you from running around in circles.

Understanding the problem

“If I were given one hour to save the planet, I would spend 59 minutes defining the problem and one minute resolving it.”

– Albert Einstein

Problem diagnosis involves defining the underlying source of an issue or dilemma. The key to successful diagnosis is to distinguish between the symptoms of a problem and its underlying causes. This is important because to be effective, solutions need address the problem itself rather than its offshoots or consequences. This happens all the time. Who hasn't looked back on a decision they have made and realized that a faulty assumption led them to choose one path and not another with major consequences. A good portion of the likelihood of the success of response is related to an appropriate diagnosis. Further, the potential consequences of a misdiagnoses are substantial, because addressing the symptoms of a problem rather than the source will create frustration, waste energy, and fail to alleviate the problem.

Examples of misdiagnoses are replete in education. One only has to review typical school improvement plans to see abundant examples. As one illustration, it is common to see schools identify gender or racial gaps in student achievement as the problem. Achievement gaps in education are endemic and are very complex issues to disentangle and meaningfully address. However, the stated strategies in improvement plans are often woefully under-conceptualized to realistically address the difficult adaptive challenge of reducing achievement gaps. Strategies for solving the achievement gap problem often include approaches such as more frequent assessments to inform teachers, more individualized instruction, computer programs that target students skill levels, or after-school tutoring. These are all well-intentioned initiatives, but will they really chip away at an entrenched problem like the achievement gap? And why are these even the right things to do, as opposed to other equally well-intentioned strategies? Do they really address the core issues that underlie differences in student performance?

These are the kinds of questions that spending time trying to understand the root causes of a problem can inform. Upon reflection, for example, we might realize that all the strategies for addressing the achievement gap problem described above are all educator-derived solutions for what are perceived to be educational problems. But is this what parents would say contributes to performance differences? Would the school's psychologist or guidance counselor have a different take on the problem? What about the early grade teachers who see students arrive with differing readiness levels and where gaps are already discernible? Are the possible explanations for achievement gaps different in different subject areas? All of these questions point to the advantages of bringing a broader array of people to the table to contribute to the understanding of the source of the problem.

Involving Multiple Perspectives

The reason to involve multiple perspectives in problem diagnosis is that different people will have different conceptions of what is the underlying cause of a problem. A key aspect of strong diagnosis is involving a range of people with different perspectives about the problem – especially those who are closest to the source of the problem and those who deal with the consequences of the problem on a regular basis. Those experiencing or affected by the problem – whether they are faculty members, students, or even parents and community members – may have varying interpretations of cause. The advantage to involving a more diverse group of people in discussing the source of the problem is that you will get many different conceptions of what might be underlying the problem. Involving a diversity of perspectives in the problem definition process will increase the likelihood that a broader range of ideas will be put forward, and that people from different backgrounds and different perspectives will be able to push on these ideas and bring out considerations that might not otherwise surface.

Inviting people with different perspectives to take a leadership role in the definition of the problem often produces a very different diagnosis than if this is done by an individual leader or a leadership team who are often more distant from the problem and its consequences on the ground.

“When people are brought together to solve problems in groups, they bring different information, opinions and perspectives.... People who are different from one another in race, gender, and other dimensions bring unique information and experiences to bear on the task at hand.”

– Katherine Phillips, Professor of Leadership

The importance of group diversity in decision making is well grounded in research. A central focus of the work of Katherine Phillips, a professor of leadership at Columbia Business School, has been to compare the quality of decisions of homogeneous and diverse groups on different dimensions. Phillips’ own studies and her synthesis of decades of research have led her to conclude that diversity matters in multiple ways. Most obviously, diversity of expertise is essential to addressing challenging problems. This is why, in schools, we want to have educators who have multiple kinds of expertise; subject matter knowledge, pedagogical knowledge, and knowledge of child psychology are all important types of expertise that help in the education of children.

In addressing adaptive challenges, Phillips has found that social diversity matters too. Hers and other research shows that diverse groups (including gender, racial/ethnic, nationality, and class differences) make better decisions than homogeneous ones. As Phillips explains, “When people are brought together to solve problems in groups, they bring different information, opinions and perspectives... People who are different from one another in race, gender, and other dimensions bring unique information and experiences to bear on the task at hand.”⁹ Interestingly, Phillips has also found that diverse groups are less confident in their decisions than are groups that are more similar to each other because diverse groups are less likely to reinforce each others’ views.

Bringing together people from different backgrounds, with different experiences, expertise, and perspectives, should be actively sought out. The experiences offered by front-line educators can illuminate the dynamics of what is supporting or limiting learning and provide opportunities to explore root causes and underlying conditions that are at the core of either a problem, an impediment to growth, or a desired improvement. If, for example, a high school principal wanted to engage the faculty in increasing minority student participation in honors classes, seeking out and listening carefully to the perspectives of those

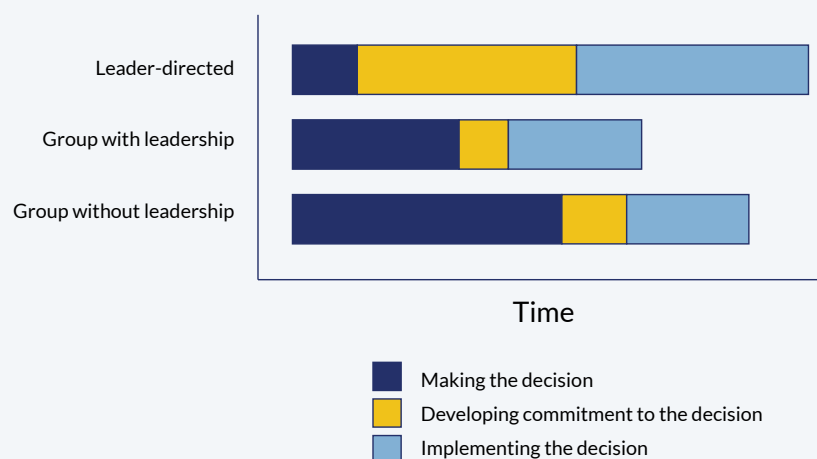
⁹ Phillips, K. W. (2014). How diversity makes us smarter. *Scientific American*, 311(4), 43-47.

teaching both honors classes and as well as non-honors classes would inform what teachers perceive as necessary prerequisites to success, as well what are seen as the qualities of high achievement. Involving students in both types of classes might give insights into the barriers from both vantage points. This process might also surface considerations about the relationship of expectations to student achievement, the influence of unconscious or hidden biases, the role of mindset on learning, and the impact of signals on student motivation. By including a range of perspectives, and creating the conditions and norms that allow for the exchange of multiple and differing perspectives from a wide group of constituencies, this approach would create the most accurate portrait of the underlying factors contributing to limited participation of certain groups of students in honors courses. While this is a more involved process, it is much more likely to produce a meaningful analysis of the problem.

Involving more people in the problem diagnosis stage is also valuable because many of the same people involved in the diagnosis process will also likely be those playing leading roles in the solution design and enactment. There are at least three reasons that many of the same people involved in the problem diagnosis should also take leadership roles in the solution design. First, as part of the diagnosis process, these folks understand the problem more intimately. Second, and perhaps even more importantly, many of these are probably going to be these same people whose commitment will be needed to address the problem. Third, and more pragmatically, involving those in the solution design who will be required to enact the decision is simply more efficient.

Consider the common situation in schools where change is introduced from outside and school faculty are expected to implement the change. They don't know exactly why the change was introduced or the rationale behind this particular reform. Further, it may conflict with current practices which have their own logic behind them. As psychologist Robert Evans argues, people are generally conservative when it comes to change, and we cling to the patterns represented by our routines. When we are asked to change, we often are not adequately provided with the rationale of why it is important and beneficial to change before we are told what we are supposed to change.¹⁰ A distributed leadership approach to diagnosis and design helps to alleviate this problem by involving those who are expected to implement an approach in the problem-definition and solution-strategizing processes. Involving the solution implementers in the process gives them more ownership of the reform implementation as they engage with its implications for their particular context. *Engagement brings with it more ownership and commitment, and commitment deepens implementation.*

Decision Time Line for individual and group decision-making



Source: Sashkin & Morris, 1984

¹⁰ Evans, R. (1996). *The Human Side of School Change: Reform, Resistance, and the Real-Life Problems of Innovation*. San Francisco, CA: Jossey-Bass Publishers.

Research indicates that a distributed leadership approach leads to more effectual implementation of decisions in cases where a group is relied on to enact the decision. When considering decision efficiency, Sashkin & Morris (1984) distinguished between the time it takes for individuals and groups to make decisions and the time it takes to enact the decisions. They argued that it is much more efficient for individual leaders to make decisions relative to groups. But this does not consider the time it takes to actually implement the decision. While it is always more efficient for leaders to make decisions alone, they still must gain the commitment of others to implement the decision. When combining the time it takes to explain a decision and gain the commitment of others to implement a decision, individual decision making actually takes longer than group decision making. While leaders can make decisions quickly, they still must gain the commitment of others before implementing the decision.

Recognizing the System around the Problem

Education issues are often interconnected to other parts of the education ecosystem. As people better understand problems and challenges, they often become more aware of the implications and connections between possible approaches and other parts of their system. In schools and districts, this might mean recognizing how a reform that emphasizes teaching for understanding might create conflict between curriculum coverage and pacing guidelines, or how attempts to build more support for certain students might affect the scheduling for other activities, or how a reform of the professional learning community structure might influence the time that PLC members have to address other issues.

One of the Carnegie Foundation for the Advancement of Teaching's Six Core Principles of Improvement is to "See the system that produces the current outcomes."¹¹ This statement reminds us of two things. First, it is the current way of doing things that is producing the current set of outcomes, and if we want to improve outcomes we have to do things differently. Second, the statement encourages us to increase our awareness of the connection between what we hope to change and how this might have an influence on other parts of the system.

Engaging in Improvement

We are all trying to make things better in our worlds. District leaders, school leaders, and teachers are all trying to find better ways of doing things. How we think about this process is the first step towards *engaging in improvement*.

***"Without continual growth and progress,
such words as improvement, achievement,
and success have no meaning."***

– Benjamin Franklin

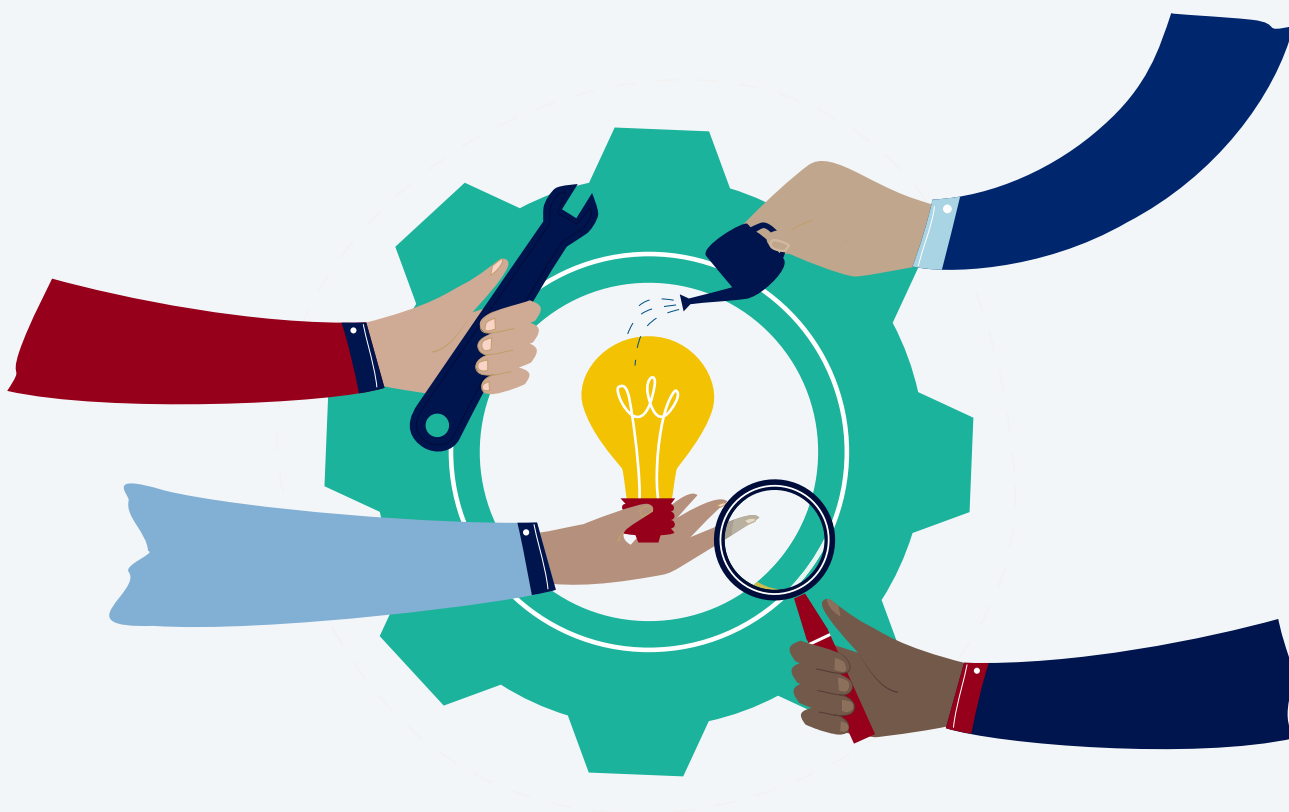
Lee Shulman, a renowned Stanford professor and researcher, called the improvement process *disciplined inquiry*, which he described as a systematic process of knowing that "uses the principles of discovery and verification that constitute the ground rules for creating and testing knowledge in a field." (p. 6).

There are essentially two types of investigations that practitioners can engage in. The first type is an investigation of the way the world currently is, and the second is an investigation of an effort to change the world. Examining the way a school or school system currently operates and the way members experience

¹¹ <https://www.carnegiefoundation.org/our-ideas/six-core-principles-improvement/>

the current system are called *descriptive studies*, because we are trying to describe existing phenomena or how things currently work. Descriptive studies can be incredibly valuable. For example, let's say we want to understand how students feel about online learning, or how students experience transitions from elementary to middle school, or their sense of engagement with their classes and school; these are the basis for all descriptive studies that provide us with insight into the way the world currently works.

The second type of investigation is an *intervention study*, which is an examination of an intervention into the natural world. Every time you are trying out a new idea – introducing a change into your school or district, or reorganizing a current practice – you are essentially doing an intervention into the way the world currently works. When you think of it, you are intervening or experimenting all the time, whether it is small-scale tinkering or major changes. The question, then, is how do you know the changes you introduce are improving things, making them worse, or not really having much effect at all? The way to know is to adopt a framework of disciplined inquiry.

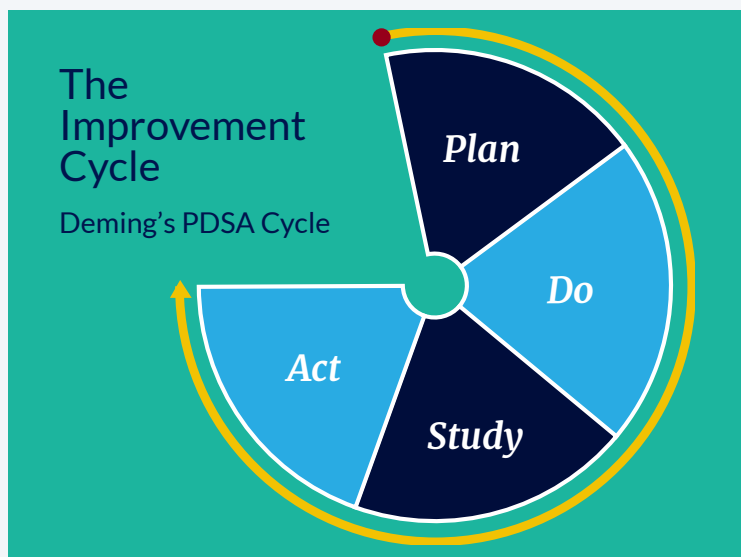


Plan-Do-Study-Act Cycles

Plan-Do-Study-Act (PDSA) cycles are disciplined inquiry in action. Although there are different models of this cyclical improvement process, perhaps the best known is the Plan-Do-Study-Act process popularized by engineer W. Edward Demings, who infused this method of quality control and continuous improvement into the post-World War II Japanese industry revitalization in the 1950s and 1960s. Demings' process was explicitly related to the scientific method of hypothesis, experiment, and evaluation, in which local knowledge is built by developing and testing a series of hypotheses that lead to ongoing improvement. Thus the entire PDSA cycle is appropriate for interventions into the natural world (i.e. the current way things work), while only the study component of the PDSA cycle is needed for descriptive studies of the natural world. Often, descriptive studies lead to a deeper understanding of how things currently work and spark ideas for interventions into the natural world that can use the entire PDSA cycle.

An important benefit of the PDSA cycle as a strategy for practitioners is that it sequentially combines the entire improvement process in one framework. As the name indicates, the process has four elements:

- **Plan:** Just like the phrase “look before you leap,” a little planning can go a long way. Part of the planning process is to use a root cause analysis, as described earlier. As previously noted, bringing in multiple perspectives can help inform what is the real problem you are trying to address and hence what should you really be trying to change.



Planning occurs in two levels at the same time. First, planning involves thinking through the actual change you plan to adopt. What exactly is the change, reform, or intervention that you intend to introduce and how will you go about introducing the change? What sequence of steps will you follow?

The second part of the planning process is to plan for a way to assess the impact of the planned change. This focuses on hypothesizing exactly what will be different after you introduce the change and what evidence you will use to assess the difference. We will talk about this more in the section on design and measurement below.

During the planning process it can be helpful to identify a *focusing question* (or research question), which helps you to target and clarify what is the central emphasis of your inquiry.

- **Do:** The “do” part of the PDSA cycle is to actually enact the change that you have planned. This also includes collecting the data you have planned to collect and organizing it in preparation for analysis.
- **Study:** The “study” component of the PDSA cycle is to analyze the data you have collected and assess the impact of the planned change. Did it meet its goals? Why or why not? Should you revise the reform and iterate again? Are you ready for wider enactment? The study part of the PDSA cycle is a central aspect of disciplined inquiry, which we will talk about in greater detail below.

- Act: The “act” part of the PDSA cycle is to either make an adjustment and iterate through the improvement cycle, or spread the change more widely because you feel confident in its efficacy. Even if you spread the change more widely, it can still be valuable to collect and analyze data for ongoing improvement.

**“Strive for continuous improvement,
instead of perfection.”**

– Kim Collins

Design, Measurement, and Analysis

Design, measurement, and analysis are all integral parts of the disciplined inquiry process. This process is essential to learning, but you can also get bogged down in its complexities. So just as you are learning how to gain experience and master the essential conditions and leadership skills of the Leadership for Learning Framework, so too should you continually try to improve on your practice of design, measurement, and analysis of things you try.

Design and measurement typically come during the planning part of the PDSA cycle, while the analysis component typically comes during the study part of the PDSA cycle.

Design

The Importance of Contrast. When you are making an adjustment to current practice or introducing a new way of doing things, the question you have to ask about the effectiveness of the change is: *compared to what?* Is it compared to the current way of doing things?¹² Is it compared to some alternative idea you have? Either way, the only way to determine if the change is an improvement is to compare it to something. The idea of contrast is integral to the design of a study because it is in the design that we assure a contrast from which to compare our change against.

In designing an intervention study, the best way to create contrast is to create some sort of comparison. The comparison could be against the previous way of doing things, the comparison could be against a group that was not doing the intervention (i.e. a comparison or control group), or the comparison could be against a group doing something different than the intervention you are investigating. Designs have many wrinkles. For those interested in more details about different designs, we have included a table at the end of the report to provide more information on study design considerations.

Measurement

Measurement refers to what type(s) of data you will collect to either capture the phenomena you are seeking to understand (in a descriptive study) or to assess the influence of the change you are introducing (in an intervention study). There are many types of data and they come in many forms; you can use existing data or collect your own data. Overall, it is most straightforward to categorize four types of data that you might use:

¹² The current way of doing things is sometimes referred to as ‘business as usual.’

1. **Existing indicators** are data that are already collected by your school or district, including attendance data, behavior data, running records of early grade students' reading levels, and test data like iready scores, MAP (measures of academic progress) scores, or other test score data. The advantage of existing indicators is that they already exist. The disadvantage is that they are unlikely to be aligned (or sensitive to) what you are trying to investigate, and they may represent an effect that is very removed from what you are examining.
2. **Survey data** include both close-ended and open-ended responses. Close-ended responses are often on some kind of scale (like a frequency scale or an agreement scale), which allow respondents to rate their views or practices. Open-ended responses are useful for having people respond in their own words.
3. **Interview Data** can be collected by casually talking to people, or through more formal interviews with a protocol of pre-planned common questions you will ask consistently to a targeted group of people.
4. **Observation Data** are collected by observing and taking notes on people in a natural setting. These could be classroom observations, meeting observations, or observations of the ways that students interact with each other in person or online. As with interviews, observational data can be informally collected or collected with the use of a structured protocol that specifies what to look for during an observation.

There are several important things to note about the measures you have available to you or the data you might collect:

- Existing data are probably least aligned with the change you are trying to introduce, so may not reflect nor be sensitive to your efforts.
- The other three types of data (surveys, interviews, observations) are more customizable, but take more effort to collect.
- Data can pile up quickly and take time to analyze, so think about the consequences of collecting too much data that subsequently make you feel obligated to analyze them.
- Just as multiple perspectives helps with problem diagnosis, collecting data from different sources can give you a different views about a phenomenon or intervention you are examining.

Analysis

Once you have collected data, it's time to make sense of them. Data analysis typically means boiling down the collected data into a handful of key findings. Analyzing data that come from interviews, observations, and open-ended survey items are usually qualitative. The goal with qualitative data are to try to find patterns or perceptive observations that provide you with insights into key messages from the data.

“What good is the warmth of summer, without the cold of winter to give it sweetness.”

— John Steinbeck

Analyzing data that come from existing indicators, observations, or close-ended survey data are usually quantitative. Quantitative data that are numerical can be represented in tables or graphs to help you identify trends or patterns. When analyzing data from intervention studies which compare groups or compare an intervention to “business as usual,” researchers use statistical tests of significance to see if the differences are larger than would just happen by chance. Appendix B provides a chart that you can follow to do basic statistical tests.

Here are a few additional key points about data analysis:

- While data analysis can be quite complex, remember that you are trying to gain insight from the data you have collected and you don’t want to get too bogged down in the intricacies of analysis.
- Many schools and systems have people on staff that are adept at analyzing data—make use of them!
- Analyzing a mix of data sources from different perspectives can provide you with a more robust picture of your efforts.

Leadership Skills

The following leadership skills can powerfully and positively impact educator interactions and maximize leaders’ ability to facilitate improvement.

Navigating power differentials

Adding to the complexity that emerges from a conversation involving philosophical differences and diverse values, are the communication challenges that stem from power differentials. “Can I be forthright and say what I am really thinking to the principal who supervises and evaluates me?” Besides the traditional boundary lines that make open and honest communication difficult, there are multiple, less visible but challenging divisions that occur around status. Can the novice teacher challenge the ideas of a veteran teacher? Can a teaching assistant disagree with her teacher colleague? These divisions also occur in many schools around departments and entire groups. Can a special educator openly disagree with the approach of a regular classroom teacher without hearing a comment like, “She has no idea what it is like to teach a classroom full of students when she has only five students at a time”? Addressing power issues that create boundary lines blocking open communication requires courageous and strategic leadership. For many of us, it requires unlearning what we have observed and adhered to throughout our work lives. Alfred P. Sloan once again models this when he suggests, “Gentlemen, I take it we are all in complete agreement on the decision here. Then, I propose we postpone further discussion of this matter until the next meeting to give ourselves time to develop disagreement, and perhaps gain some understanding of what the decision is all about.”¹³

Addressing non-discussables

Addressing school issues openly and transparently can bump into non-discussables.¹⁴ Roland Barth writes, “Non-discussables are subjects sufficiently important that they are talked about frequently but are so laden with anxiety and fearfulness that these conversations take place only in the parking lot, the rest

¹³ Sloan, A.P. quoted in: “Alfred Sloan, Guru,” economist.com, Jan. 30, 2009.

¹⁴ Barth, R. (2002). *The Culture Builder*. Educational Leadership, Volume 59 (8). Pages 6-11.



rooms, the playground, the car pool, or the dinner table at home. Fear abounds that open discussion of these incendiary issues at a faculty meeting, for example—will cause a meltdown.”¹⁵ Issues that pertain to racial relationships or the poor performance of a leader or a department are examples of potential non-discussables. A common non-discussable is the unwillingness of staff and the administration to critique their own behavior and motivation and discuss their contributions to a particular issue.¹⁶ Typically, faculty might be comfortable critiquing leadership but leave their own behaviors unexamined. Leaders, too, often leave their own contributions out of the discussion of a problem for fear of appearing weak. Leaders can also be leery of directly raising concerns about faculty performance and attitudes. In an attempt to keep morale positive, leaders may choose to communicate indirectly or stay silent about their concerns related to faculty performance. These limited and filtered exchanges between educators and building leaders, while capable of producing “middling commitment and morale”¹⁷ and, in some cases, reasonable productivity, fall short of the excellence and higher standards that many educators seek.

Listening in stereo

This is the ability of a leader to listen carefully to both the content of the conversation and how it is expressed. No one would argue with the idea that problem solving requires valid data. Where people diverge is in what counts as valid data. Peter Block in *Flawless Consulting* notes that data encompasses both objective data (facts about situations and events) and personal data. Block writes that personal data are also “facts,” “but they concern how individuals feel about what is happening to them and around them. If people feel they will not get a fair shake, it is a ‘fact’ that they feel that way, and it is also a ‘fact’ that this belief will have an effect on their behavior. To ignore this kind of ‘fact’ is to throw away data that may be crucial to any problem-

¹⁵ Ibid

¹⁶ Argyris, C. Good Communication That Blocks Learning, *HBR*, July-August 1994, p.85

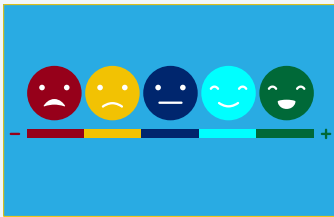
¹⁷ Ibid, p.85

solving effort.”¹⁸ When we discuss listening in stereo we are emphasizing that hearing and acknowledging the feelings embedded in communication are invaluable parts of data collection and trust building. Emotions are data, and overlooking the data communicated through peoples’ affect often produces blind spots that can negatively impact understanding and decision making.

“Leadership and learning are indispensable to each other.”

– John F. Kennedy

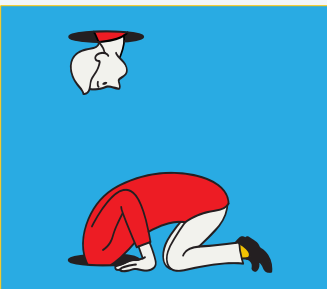
Managing emotions



The conflict and discomfort that emerges from these disagreements often brings out difficult emotions to manage. Underlying frustration, anger, and annoyance often get expressed as sarcasm, personal attacks, or silence, none of which move the conversation forward. Understanding the emotions expressed at meetings is no small challenge. People express their feelings differently. Some do so directly; others indirectly. Indirect expressions can be both verbal and nonverbal. This range of expression makes navigating and understanding emotions complicated. It is for these reasons that engaging a staff to present their perspectives on a topic, issue, or problem – while appearing attractive – is often a choice viewed by educational leaders as risky and challenging. Despite the risks, however, moving forward in this area provides leaders with access to valid data – data related to the “facts” of how individuals feel about what is happening to them and around them (Block, p.18).¹⁹

In sum, seeking the perspectives of the staff and creating a culture that values open and honest self-examination is an essential aspect of our model of distributed leadership. However, simply reaching out for the ideas and perspectives of educators will not generate creative solutions. Often what it creates is dissonance and a tendency to avoid critical self-examination. Learning to manage and pay attention to the interactions that occur, anticipate disagreements, and fully explore the thoughts and feelings of constituents, will generate more ideas, deeper commitment, and potential breakthroughs than approaches that avoid the messiness of conflict.

Curiosity in the face of criticism and wrong sounding ideas



Modern leaders encounter a great deal of conflict and disagreement. Managing these differences in a way that produces insight and better decision making requires that leaders navigate the turbulence of difficult conversations. Stone, Patton and Heen emphasize the importance of taking a learning stance when approaching conversations where the other person’s perspective is contrary to one’s own values and perspective.²⁰ A learning stance involves becoming interested in the other’s story. Moving from certainty about one’s own point of view to curiosity about how someone else thinks differently is a powerful skill that can lead to new insights and understanding. The insights from Stone et al derive from their involvement in complex and challenging negotiations and mediation. Their research showed that curiosity needs to be authentic, i.e. a genuine quest to understand where the other person is coming from. It is not something that simply flows from questions or scripts.

¹⁸ Block, Peter. *Flawless Consulting, Enhanced Edition: A Guide to Getting Your Expertise Used* (p. 18). Wiley. Kindle Edition.

¹⁹ Block, P. (2011). *Flawless Consulting*. San Francisco, CA: Jossey-Bass.

²⁰ Stone, D., Patton, B., & Heen, S. (1999). *Difficult conversations: how to discuss what matters most*. New York, N.Y.: Viking.

Rather, curiosity derives from one's genuine interest in learning about someone else's perspective. Authentic curiosity often produces additional data and the increase in information often surfaces missing and critical aspects of an issue.

Balancing inquiry with advocacy



In his work on the qualities that make companies into learning organizations, MIT's Peter Senge notes that leaders are often too quick to jump from an inquiry stance to one of advocating for an idea or action.²¹ When we are inquiring, we are exploring an idea or potential action, gathering information, asking questions, and trying to understand the merits and shortcomings of the idea. When we become advocates, we take on a stake in the outcome and, in doing so, our position becomes more defensive because we have, at least psychologically, taken some ownership of the idea's success or failure.

Senge found that advocacy leads to fewer questions, an understatement of risk, and imbalanced judgment of success. He advises that leaders retain their inquiry as long as possible, resist becoming premature advocates, and even when we choose a course of action to remember that we are testing an informed hypothesis and remain open to revisiting it if it does not go as planned. This will help us to avoid the defensive ownership that comes from too quickly advocating for an idea.

The Discomfort of Disagreement

While exciting and informative, seeking out diverse perspectives from a wide variety of constituencies also creates conflict. Rarely do people agree when it comes to unpacking complex educational issues. Disagreements abound about how to approach, for example, literacy instruction, student discipline, or effective parent engagement. And while some may experience open disagreements as a natural and normal part of a healthy exchange, others view disagreements as inappropriate and a form of disrespect. Consequently, there is often discomfort when disagreements emerge. People react to this discomfort in a variety of ways including avoidance, accommodation, or competition.²² Curiosity and collaboration are unfortunately not the most consistent choices that people make when confronted with opposing ideas, but these approaches often provide a means to bridge the divide that emerges when values, ideas, and perspectives clash.

21 Senge, P. M. (1992). Mental Models. In *Planning Review*, 20(2), 4.9-10,44.

22 Anderson, D.L. (2017). *Cases and Exercises in Organization Development & Change*. Thousand Oaks: Sage Publications



Leading Improvement in Challenging Times

The legendary Harvard Business School professor Chris Argyris, who studied patterns of organizational learning for over 50 years, coined the distinction between single loop and double loop organizational learning. Single loop learning, Argyris observed, was the most common type of organizational response to a challenge, which involved incremental adjustments in response to a problem. Argyris wrote that single loop learning, under the best of conditions, can help organizations make small improvements, but would not generate great leaps of progress. Double loop learning is much more difficult. It requires people to question the underlying assumptions about organizational processes. But double loop learning was the most likely way to shift the paradigm to produce breakthroughs in thinking. In his work with organizations, Argyris saw relatively few examples of double loop learning, which he concluded was mostly due to the protective behaviors and defensive routines that organizations adopted and which impeded the frank questioning of underlying beliefs that were necessary to achieve significant improvements. We see the tackling of difficult educational challenges using distributed leadership as a way of creating the conditions for double loop learning.

Efforts to achieve meaningful school improvement require three essential ingredients which we have described in this booklet. The first ingredient is an organizational culture in which learning and change are the norm, where learning and improvement are a regular part of what you do. The second ingredient is an improvement approach that provides a disciplined process to iteratively develop and test hypotheses to identify root causes and then design potential strategies to overcome challenges that impede progress. The third ingredient is skilled leaders who can lead develop organizational conditions and facilitate the improvement and learning process.

Embracing this process can be daunting for leaders since there are risks associated with the more inclusive process that distributed leadership entails. As we've described in the booklet, some of the potential risks

include challenges to the authority of leaders, the discomfort of disagreeing openly with colleagues and superiors, the possible surfacing of uncomfortable and awkward feelings and issues, and managing one's own emotions and the emotions of others who bring passion and conviction to their work. This is akin to the defensiveness that Argyris thought got in the way of double loop learning. In polite circumstances, these are all things to be avoided. But this is the crux of the issue – because these are the very circumstances where the truly meaningful issues come to the surface.

This presents school leaders with a double-edged dilemma. One edge promises the candid interactions amongst stakeholders that produce the insights necessary to understand and make progress on difficult problems that impede school improvement. On the other edge sits the discomfort, awkwardness, and painful emotions that may surface when discussing sensitive issues. Yet these two things go hand in hand. It is exactly at this fulcrum where important insights emerge. We learn when we are just outside our comfort zone.

The way out of this dilemma requires that leaders create the conditions and develop the skills to manage the more contentious discussions that are necessary to hash out, enact, and revise meaningful improvement plans. School leaders and faculty don't have a lot of training or experience operating in the uncomfortable space of disagreement that produces double loop learning. The key idea that we have offered on how to create the foundation for this work is to establish a psychologically safe space solidified by trust and embedded in a culture of mutual learning. However, psychological safety doesn't mean providing comfort. It means making people more comfortable with discomfort. Breakthroughs come in such a crucible. People have to be willing to feel comfortable enough with the discomfort of disagreement, manage the associated emotions that often emerge in such an arena, and not shy away from the challenging conversations. This work also requires having faith that hanging in there will get you to a better place. Skilled leaders recognize this and seek to enter and encourage these kinds of conversations because this is where the real learning comes that fuels meaningful and sustainable improvement.

Appendix A - Study Designs for Intervention Studies

Campbell and Stanley (1963) devised a really useful way to categorize studies that evaluate the impact of a reform, change, or intervention (which Campbell and Stanley called a treatment). The table below encapsulates their categorization. In general, the designs are in ascending order of rigor.

Study Type	Study Design	Study Description	Strengths and Weaknesses
Case study	X O	Introduction of a treatment, followed by a measure of the impact of the treatment (i.e. posttest).	The weakest of designs. Without a comparison group, we do not know whether it was the treatment or something else that produced the posttest results.
One group pretest-posttest	O X O	A treatment with a pretest and a posttest.	Better than a case study, but again, without a comparison group there is no way to determine whether it was the introduction of the treatment or something else that produced any change in the posttest.
Static group comparison	X O O	Two groups, one receives the treatment and the other does not (control group); both groups receive a posttest.	This design introduces a comparison group, but we do not know if the two groups are equivalent and therefore if we see differences in the posttests, we cannot attribute it to the treatment or to the non-equivalence between the two groups.
Static group comparison (variation)	X O Y O	Two groups, each receives a different treatment; both groups receive a posttest. (Note, there could be a third, control, group, or even multiple groups receiving different treatments).	A variation of the design above, and similarly, if we see differences in the posttests, we cannot attribute it to the treatment or to the non-equivalence between the two groups.
Non-equivalent control group	O X O O O	Two groups, each receiving; one receiving a treatment and the other does not (control group). Both groups receive a pretest and posttest.	The introduction of a pre-test or control variables for both groups allows for a fairer comparison between the two groups, as it allows us to 'adjust' for any observable differences. However, this does not eliminate the possibility that the treatment group has unobservable differences (most importantly potential differences in motivation to participate in the treatment) between itself and the control group.
Non-equivalent comparison group	O X O O Y O	Similar to the non-equivalent control group, but instead of a control group, this is a comparison between two different treatments.	A variation of the design above, but instead of a control group, the contrast is between two different treatments. The concern about non-equivalence between groups remains.
Time series	OOO O X O OOO	Multiple pre-treatment measures allow for a trendline to be drawn before the application of a treatment. Any difference in the trend that occurs after the treatment can be seen as the effect of the treatment.	This is a strong design. Its weakness is that, since there is no external comparison, some external event co-incident with the treatment could have caused the change in the post-treatment slope, rather than the treatment itself.
Experiment	R O X O R O O	Randomizing people to receive the treatment or to be in the control or comparison group. Both groups receive a pretest and, while only the treatment group receives the intervention. Both groups receive a posttest.	Typically viewed as the strongest design from which to determine whether the treatment caused the posttest impact. Random assignment assures that there are no differences between those who receive the treatment and those who don't (given adequate sample sizes). Therefore, there any difference between the two groups can be attributed to the treatment.

Study Design: X represents a change or improvement idea (i.e. treatment); O represents some measure to assess the change (i.e. a pretest, other covariates or control variables and/or posttest); R represents random assignment.

From Campbell, D.T. and Stanley, J.C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Appendix B – Quantitative Analyses for Intervention Studies

Below is a decision tree that can help you decide what kind of statistical analysis you should use in different situations when conducting quantitative analyses of intervention designs. Dependent variables refer to the outcome or variable that you seek to change. Independent variables are inputs or the number of groups amongst which you are looking for differences. A key distinction that influences what type of statistical test to use is whether either the dependent variable or independent variable(s) are categorical (i.e two or more groups) or continuous.

