Instructional Gemba Walk:
A 6 Step Observation Process
Focusing on Instructional Improvement
(Stephen Cousins)

In order to nurture a continuous improvement mindset and to focus our energy on high yield improvement strategies, we have to implement processes into our work that allow us to reflect on the current state of our work. It is in our nature to improve. We are wired to continuously adapt to an ever-changing environment so that our lives become more meaningful, more comfortable, and more satisfying. It is not so much that we have to create the incentive to change, it is more about putting processes in place that allow us to take that journey together without falling over our own feet.

Too often, we allow ourselves to get sidetracked to unproductive areas of professional growth. We focus on professional practices that do not provide the level of improvement we desire. Instead, we put our efforts into doing things that only improve learning at the margins. I worked at a golf course during the summer months when I was in college. One of my jobs was to mow the fairways. There are few things in life more immediately rewarding than mowing the grass. The results are instantaneous, and the improvement obvious. Well-manicured fairways make a golf course look great, and to some extent, help improve the game of a golfer. But if I want to significantly improve the game of a golfer, focusing on the fairways is a low yield strategy. Whether the fairways get mowed every day or every third day doesn't really matter. What matters most (high yield) is whether the golfer can find the right grove in his/her swing to square the club to the ball on impact. To improve, a golfer has to spend time focused the golf swing, not complaining to the superintendent of the course about the length of the grass on the fairways.

Similarly, we can spend way too much time aligning curriculum, creating pacing guides, and discussing the structure of the instructional day, instead of focusing on improving our professional practice. It isn't that we shouldn't take care of these other issues; we should. In the long term, however, they will not yield the kind of improvement we desire.

We can also get stuck in behavior and thinking cycles that develop ruts so deep that steering out of them becomes a Herculean task. To move forward, we need to make sure we build the type of processes that allow us the space to take advantage of our own innate curiosity. This requires purposeful, scheduled, and strategic moments of reflection guided by protocols that lead us to reasonable predictions about how to improve.

The Instructional Gemba Walk is a way to create the space we need to reflect on our practice in a meaningful and deliberate manner. It uses data we collect to chart a path of improvement utilizing protocols that filter bias and keep us focused on what we can control. There are six steps to the process: Observe and Record, Share and

Challenge, Determine Patterns, Look for Opportunities, Use Root Cause Analysis, and Apply the Improvement Kata. I have provided a description of each step following.

Step 1: Gemba Walk - Observe and Record

Gemba is a Japanese word that means "at the site." When Gemba is used in conjunction with process improvement methodologies, it refers to the act of making observations of a particular process in action. Gemba Walks evolved out of the quality improvement procedures developed by Japanese automakers. Subsequently, manufacturers, health care systems, and technology industries have adopted them.

Rudimentarily, Gemba walks involve stakeholders observing a process with a goal of identifying waste (inefficiencies and ineffectiveness), and therefore, areas for possible improvement. There are many forms of waste, but simply put, waste can be described as the gap between what we intend to happen and what actually happens. For example, in a classroom, the teacher's intention is that 100% of the students will be actively learning 100% of the time. In reality, this is nearly impossible to achieve. The gap between his/her intention (100%) and the observed reality (the percentage of students actively learning) represents the potential for improvement.

In the realm of instructional leadership practice, going to Gemba requires that principals and teachers visit a classroom to watch students and teachers engage in activities that immerse them in curriculum content. The observations create an articulate, objective data set that describes the current condition of instructional practices and the effectiveness of these instructional practices. Instructional leaders and practitioners can use the observational data set to strategically plan for improvement.

For Gemba Walks to me most effective, it is helpful to narrow the observation parameters. Classroom instruction is a very complex and messy process. The larger community in which the students live, including a student's background and family community, funding cycles and government regulation, school board policies and strategic planning, and a myriad of other variables that I cannot account for within this short paragraph influence it. However, there is considerable influence over the learning process afforded teachers and principals. The instructional activities that are used every moment, in every classroom, in every district, are within teachers' and principals' locus of control. We need to limit the focus of our Gemba Walks to only the issues we can control; it is here that we have the ability to improve the system.

It is helpful to focus Gemba Walks on the most impactful areas of instructional practice. By zeroing in on high impact areas of practice, we create the greatest opportunity for leveraging individual, cohort and organizational improvement. Gemba Walks require precious time within the instructional day, so it makes little sense to spend that time looking for improvement in processes that reflect low yield improvement opportunities.

Rubrics designed to articulate the critical attributes of effective instructional and leadership practices are very helpful in leveraging accurate observational data from those participating in Gemba Walks. Rubrics provide consistent language, clarify meaning and intent, remind us of critical attributes for success, and help develop inter-rater reliability. Front-loading Gemba Walks with rubric research and adoption can improve the data gathering process immensely.

The guidelines for a Gemba Walk are fairly simple:

1. Spend enough time observing to see a complete process. In a classroom, this will require observations of 10 - 20 minutes. Quick processes, like scale setting at the beginning of a lesson, can often be observed in 10 minutes. Longer processes such as individual practice combined with feedback loops and discussion, might take 20 minutes to complete.

2. Take notes during the observation, but only record what is actually heard or seen. This is difficult for many educators to do. We are so familiar with instructional practices that we will often add unobserved attributes to our data. For example, if a teacher prompted students to engage in conversation with a partner in relation to some procedure or content, we might be inclined to observe, "that students participated in substantive conversation." The use of the word "substantive" places a value on the content of the students' exchange. Also, "students participated" does not tell the approximate percentage and may suggest that all participated when far fewer did. Unless we heard all the students participating in conversation that was substantive, what we actually observed was "X number of students participating in conversation."

   Some simple guidelines for taking notes:
   a. Use an interval of 2-3 minutes
   b. Watch what the students are doing and saying
   c. Record only what you actually see or hear
   d. Avoid value-laden words
   e. Do not "fill-in" processes
   f. Do not assume
   g. Do not infer

3. Without interfering with the teacher's instruction, ask some students to explain what they are doing and why.

4. Watch for the gap between what the teacher intends to have happen and what actually happens. This can only be achieved by watching what the students actually do and say.

Step 2: Share and Challenge

The next step requires the observing group to visually display and explain their observations so that all group members can see and hear what is meant by the observational data. This can be done in a number of ways (i.e. white board, poster paper, etc.). It is important that all participate and an exhaustive list of observations is recorded. This can be accomplished by having each person share one observation at a time. The person speaking should explain what he/she saw or heard, and the group should display this visually. The process should follow in a round-robin format until every person's list is exhausted. Don't record duplications, but make sure that these duplications are clarified to check for understanding.

The recorded observations must be as factual as possible, which means eliminating judgmental and value-laden language.

Some guidelines for recording and explaining data:

1. Use non-judgmental wording
2. Verbalize and write from the students' perspective not the teacher's
   Good – "20 of 22 students participated in partner share."
   Avoid - "The teacher prompted the students to participate in partner share."
3. Make it visible for all group members to see
4. Note repeated observations; make it visible
5. Make sure everyone in the group understands what is meant by the observation; check for understanding

The observations must be as objective as possible. In spite of our best efforts when recording our observations in the classroom and our attempts to be precise when explaining our observations to the group, it is inevitable that some errors will slip through and make it to our visual display of data. We can improve our data set by zero defecting our observations.

Value-laden words like "substantive", "active," "deep," "critical," "rigorous," "engaging" and "robust" should be challenged to make sure that the observer had visible or auditory evidence of it occurring. As stated above, we often add these attributes to observations, even when they are not observable, because we have a habit of using these words to describe what instruction should be. They are part of our lexicon. Most of us have been exposed to these words repeatedly through our teacher preparation courses, staff meetings and professional development, and so we use them out of habit and often times without understanding what they mean in terms of student learning within a specific, observed environment. In other words, we know from our training that teachers need to create activities that cause students to interact in substantive conversation about processes and content. When we observe a teacher implementing an activity with a group of students, we have a tendency to add to our observation description ***what should be happening*** rather than ***what is actually happening***.

In their book, ***Made to Stick: Why Some Ideas Survive and Others Die***, Chip Heath and Dan Heath write about this phenomenon –the curse of knowledge. Simply put, when we become very knowledgeable about something, it becomes harder for us imagine not knowing it. This "curse of knowledge" can creep into our observations in the form of "filling in" a process with a missing step or "assuming" that something will or did happen.

Some guidelines for zero defecting data:

1. Constantly restate the rule - "only what was actually seen or heard."
2. Challenge inferences, assumptions and value statements
3. Require evidence through direct observation
4. Beware the "curse of knowledge"

Step 3: Determine Patterns

Once we have presented, recorded and challenged all our observations, it is important to look for patterns. The group should scan the recorded observational data and sort the observations into categories. The question running in the background should be "what patterns of instruction are we seeing?" Remember, we are looking at our observational data through a lens with a specific focus.

In light of this step, a word of caution should be noted. In the classroom Gemba Walk, participants spend a short period of time in any one classroom. The short duration of each observation does not allow the team to make evaluative judgments about a teacher's work. We cannot assume that what we are seeing is a pattern of behavior or a simple anomaly that represents a very small error within the broader spectrum of instruction.

Said another way, we all have moments when we do not perform at our best, and because these moments can be fleeting, addressing them is not worth our effort. We want to focus on patterns of behavior that reflect habits of instruction. Every process used to produce an outcome has a level of error. Since nothing can be done to perfection, there is always an ***acceptable error rate*** and the ***opportunity for improvement*** at the same time. The acceptable error rate is determined by the limit in resources, technology, and expertise that are out of our locus of control and for which we rely on others for change. Trying to fix errors within this bandwidth can tax the system to the point of paralysis with little gain in improvement; "the juice is not worth the squeeze." Outside the acceptable error rate and within our locus of control is a broad spectrum of opportunity for improvement represented by patterns of reoccurring inefficiencies or ineffectiveness and waste.

When using the Gemba Walk to support the practice of a single teacher, multiple Gemba Walks must be completed before looking for patterns. When using Gemba Walks to look at instructional practice from the cohort level (i.e. building, grade level, subject area), we must visit multiple classrooms to gather our data set.

Some guidelines for sorting the observational data:

1. Try to limit the sorting to 3-5 categories
2. Align to your rubric
3. Check for understanding

Step 4: Look for Opportunities

Given the patterns of instructional practice presented in the observational data, what **opportunities** might exist for the building, cohort or individual to improve? This might present itself in the form of a pattern of deficiency in an important area of instructional practice. For example, a Gemba Walk might reveal a pattern in which 80% of the observed activities lacked meaningful learning activities for most of the students. Participants might see a pattern of fact-based questioning devoid of critical thinking that is directed at a small percentage of students within the classroom. Certainly, the teachers doing the instruction do not intend to implement learning activities that are not meaningful or interactive. If this pattern persists, we can assume the teachers do not know how to provide this kind of instruction; if they did, they would. By targeting this pattern, we can improve instruction for a majority of students and increase teacher and student achievement.

Similarly, the observational data might reveal anomalies within the system that point to possible **solutions** to patterns of ineffective or inefficient instruction. These anomalies represent as pockets of exemplary practice that could be standardized and utilized by other teachers, and they represent a great resource to help us engage in rapid improvement cycles. For example, if our observational data revealed a pattern of instructional activities for which 80% were not interactive and meaningful, then 20% of our observational data might reveal anomalies that are the seeds for growing solutions.

During this step, fashion a **problem statement** based on the observational data and the patterns the group has identified. A problem statement should focus on improvement by recognizing an area of inefficiency and/or ineffectiveness. For example, if the observational data reveals a pattern of teachers using activities that do not require interaction among students and that are meaningful to students, then the problem statement might become - "We do not involve our students in enough meaningful and interactive activities."

Some guidelines for questions about opportunity and solutions:

1. Apply a standard of 90/90/90 – are 90% of the teachers at 90% proficiency 90% of the time. Where this standard is not met, we have an opportunity for significant improvement.
2. Ask "What opportunities do we see?' and "Are there solutions present within our observational data?"
3. Fashion a **problem statement** based on the patterns of instruction. For example, "Our instructional strategies are not causing active learning to occur."

Step 5: Use Root Cause Analysis

Once we have created a problem statement, based on our Gemba Walk and study of the observational data, we want to find the root cause of this problem. In the analysis, it might be helpful to think of the observational data as the symptoms being manifest by the problem. The data tells us what it looks like and what it sounds like, which can help identify the problem. Once we have narrowed down the problem, we can use the data set and patterns to determine what is causing the problem to occur. For example, if I observe that I have a sore throat, runny nose, and fever (observational data), then I can determine that I am ill (problem). By determining the root cause of my illness, viral or bacterial, I can start to formulate what course of treatment (solution) would provide the greatest gains (improvement) in health.

For example, although Galen Rupp was one of the most celebrated high school and college distance runners ever to step on a track in the United States, his coach observed that in his distance of choice, the 10,000 meters, his opponents in international races often out-kicked him over the last 400 meters. In particular, although he had the lead in the 2008 Olympic 10,000 meter race with two laps to go, he finished out of the top three, and therefore, did not achieve his goal of an Olympic medal. This also occurred in the 2009 World Championships. Galen's "fading" over the last 400 meters presented as a symptom to his coach. His coach determined that if he did not understand the underlying problem and its root cause, Galen would never medal in the Olympics or the World Championship.

Subsequently, Galen's coach observed the performance of the most successful 10,000-meter runners in the Olympics and the World Championships over the past decade. In doing so, his coach noted a pattern among those who won Olympic and World Championship medals. Almost all were able to run the 10,000 meters in under 27 minutes ***and*** run the last 400 meters of that race in under 53 seconds. A root cause analysis revealed that, although Galen was capable of running under 27minutes for 10,000 meters, he lacked the leg speed and strength to muster a closing 400 meters in under 53 seconds. Determining the root cause helped Galen and his coach revamp his training program to improve his leg speed over 400 meters by two seconds while increasing his overall strength. In the 2012 Olympic 10,000 meter run, Galen closed the final 400 meters in 52 seconds and won the silver medal.

This example illustrates that it is not enough to determine the problem. Many things could have caused Galen to fade over the last 400 meters of a championship 10,000-meter race. He could have been ill with a virus that weekend his ability to respond. He could have been suffering from allergies. His girlfriend might have dumped him five minutes before the race and the resulting emotional state depleted his resolve. By looking for patterns within observational data his coach was able to determine the root cause of the problem and provide the correct strategy for improvement.

Two tools can be effective in helping a team get to the root cause of a problem. One is called the **5 Whys,** and the other is a **Fishbone Diagram**. The 5 Whys can be used for simple problems, and the fishbone can help with more complex problems.

Using the 5 Whys protocol requires that we ask the following, "Why does our current condition (problem statement) exist? What is causing it?" Once answered, we continue asking and answering the questions until we repeat a cycle of five.

For example, a Gemba walk might produce the following problem statement: our students were involved in active learning only 20% of the time. By applying the 5 Whys, we might be able to determine the root cause of the problem.

Question 1: Why were our students involved in active learning only 20% of the time?

Answer: Because the activities the teachers used required very little student-to-student interaction or interaction with the curricular content and the whole class.

Question 2: Why did the activities the teachers used require very little student-to-student interaction or interaction with the curricular content and the whole class?

Answer: The teachers provided mostly direct instruction and called on students individually to answer questions.

Question 3: Why did the teachers provide mostly direct instruction and call on students individually to answer questions?

Answer: The teachers did not seem to understand how to create opportunities that required all students to be involved in learning.

Question 4: Why didn't the teachers seem to understand how to recreate opportunities that required all students to be involved in learning?

Answer: Professional development opportunities within the district are mostly voluntary, and less than 30% of the teaching staff has been exposed to strategies that require all students to be involved in learning activities.

Question 5: Why has fewer than 30% of the teaching staff been exposed to strategies that require all students to be involved in learning activities.

Answer: Because we have not made this focus a priority.

From the example above, one can begin to see pieces of the root cause exposed and the seeds of a potential solution. It becomes apparent that the teachers seem to lack

the requisite skills to create active learning environments for the students, and the district has never made training within this area a priority for a majority of the staff.

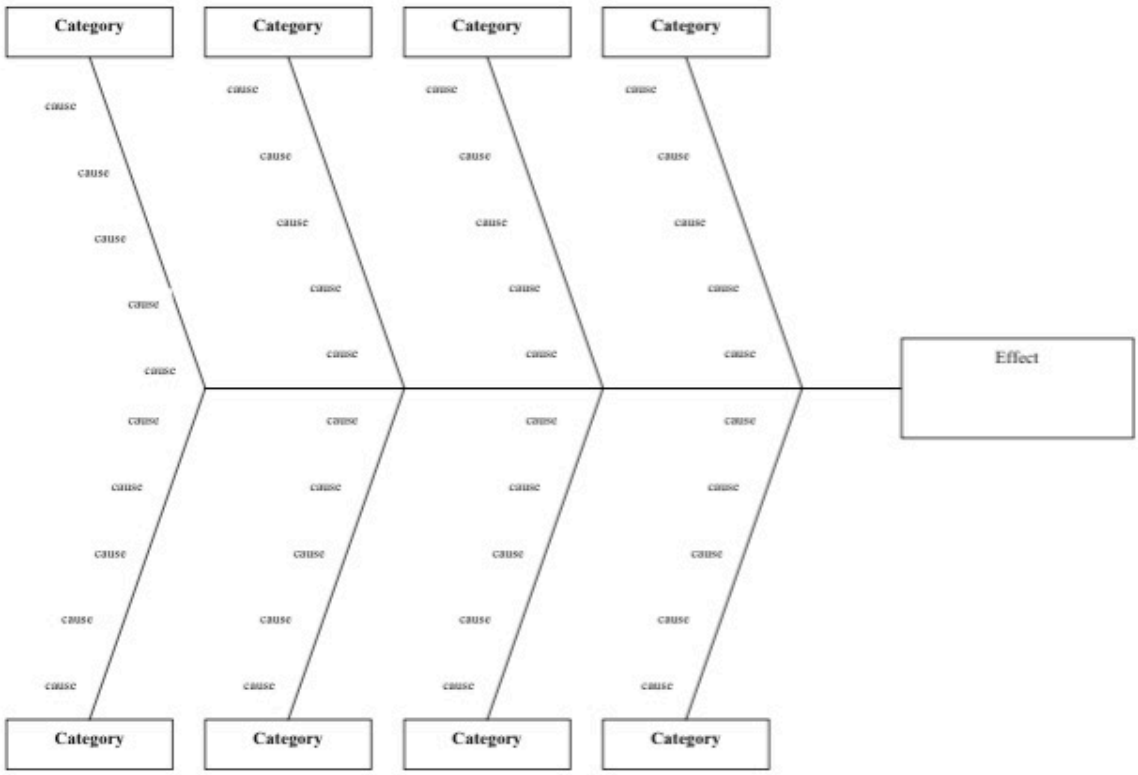
It is important to note that using a root cause analysis carries with it the inherent danger of finger pointing. Once we determine a root cause, it is a natural human tendency to assign blame. We don't do this because we are malevolent or unkind beings, we just do it because it is in our nature to affix a problem to someone's realm of responsibility. The purpose of a root cause analysis, however, is not to assign blame to any person or group; the purpose is to find solutions to what is keeping us from improving our practice. It should be of no surprise that our professional practice can be improved – I have never met anyone in my decades long career who claimed to have solved all life's instructional or leadership problems. This isn't about being effective – it is about improving. It isn't about labeling teachers and principals as ineffective – it is about helping them become better. Galen Rupp was already one of the best distance runners in the world when he refocused his training on building his finishing speed.

Some guidelines for using the 5 Whys protocol:

1. Always go to the fifth level of questioning. Sometimes, a root cause seems to reveal itself before we reach the fifth why. Avoid the tendency to stop the questioning when this happens. Push on to the fifth cycle.
2. If working with a group while using the tool (this can be used by individuals to look for root cause), work from a consensus format at each level of why.
3. Be as articulate as possible when answering each level of why.

The Fishbone Diagram can be used for more complex problems. This protocol is called a 'fishbone' because it resembles the bones of a fish skeleton. A template is below:

Title:



The basic idea is one of identifying the root causes of some effect that we want to improve. The effect is essentially the problem statement created by the team after moving through the first four steps of the instructional Gemba process. Once we identify the effect we want to improve, then the team brainstorms possible causes of this effect. The causes should have a direct connection to the observational data that was collected and shared by the team.

The team should focus on causes that are within the team's locus of control. There are a lot of variables that impact how well a student will learn. We know that some of these variables land outside the influence of the school's environment. These are things that we have little or no control over, and therefore, it is not productive spending time trying to resolve them. It doesn't mean that they do not exist, and it doesn't mean that they will not impact how a student learns. It just means that we cannot control them.

Fortunately, we do control a great many of the variables that influence a student's learning; if not, schools would have no legitimate mission and educators would have

no legitimate professional practice. During an instructional Gemba, our focus must be on variables that fall within the realm of our own professional practice. This we have total control over, and it is here that we can exercise big improvements within the system.

For example, if a team of teachers and principals was using the fishbone protocol to determine why students were involved in active learning only 20% of the time, the group might have a tendency to list variables outside its locus of control. Parents are too overwhelmed with work or do not care enough about education to support the school. Students begin elementary school with limited school readiness and a lack of basic skills. We are a quick fix society so students and parents do not have the drive strength to meet the demands of the tough standards for learning mandated by our state governments. The department of education keeps throwing tests at us and disrupting the flow of teaching.

All these things are real – they are not figments of our collective imagination. All get in the way of our progress. But they always have to some degree – this isn't new ground. Professional practice has always been the way to balance the odds back in our favor. When the kids are with the teacher, in the moment, learning happens. Teachers and principals cause learning to happen. To improve, like Galen Rupp, we need to focus on what we can control within our own professional practice.

Once the team has exhausted the possibilities of the root cause of the problem, then the team must determine which root causes it is going to attack. It would be too difficult to address all the root causes of a complex problem, so focusing on one or two causes that have the potential for high yield return is usually best. The team should make its thinking visible by circling or highlighting the causes to be addressed immediately.

The next step is to determine if any other root causes could be eliminated by addressing those identified for immediate action. Often, by attacking one or two, high yield root causes we can mitigate the impact of other root causes. Again, the group should make its thinking visible by highlighting these root causes. Finally, any causes left over should be put into a parking lot to be addressed at a later time. We do not want to just push them to the side. Using a continuous improvement mind set requires that we commit the team to revisiting any unaddressed causes in the future.

Step 6: Apply the Improvement Kata

Once the root cause analysis is completed, the team would follow-up with another protocol for implementing strategies for improvement (i.e. A3, Kata Coaching Model, Value Stream Map, etc.). The Improvement Kata is covered in another publication, "The Improvement Kata: Managing Improvement Cycles One Step at a Time."