A Deeper Dive into Plan-Do-Study-Act Cycles and Measures

Marilyn Strutchens
Auburn University
strutme@auburn.edu

Michele Iiams
University of North Dakota
michele.iiams@und.edu

Ruthmae Sears
University of South Florida
ruthmaesears@usf.edu

Mark Ellis
California State University Fullerton
mellis@fullerton.edu

This talk explores more deeply the structures and measures of the Plan-Do-Study-Act (PDSA), a core principle of improvement science. We began this discussion with the challenges that we faced in moving from a traditional research paradigm to using improvement science, specifically the struggles we experienced in terms of starting small with the goal to ramp up rapidly. Lewis (2015) sums up the transitions that we had to make:

> Improvement science… treats variation in implementation and setting as important sources of information and provides tools to grasp and learn from variation (in both positive and negative directions) in order to redesign both the intervention and the system. (p. 55)

And,

> Improvement science assumes scale-up occurs through integration of basic knowledge with the "system of profound knowledge," such as knowledge about how to build shared ownership of improvement, to detect and learn from variations in practice, to build and share knowledge among practitioners, to motivate frontline innovators, and so forth. (p. 55)

The PDSA cycle (Lewis, 2015), Figure 1, is an essential tool of improvement science,

> a process for rapid cycles of learning from practice, coupled with three fundamental questions that drive improvement work: (1) What are we trying to accomplish? (2) How will we know that a change is an improvement? and (3) What change can we make that will result in improvement? (pp. 54-55)

Lawler, B. R., Ronau, R. N., & Mohr-Schroeder, M. J. (Eds.). (2016). *Proceedings of the fifth annual Mathematics Teacher Education Partnership conference*. Washington, DC: Association of Public Land-grant Universities.

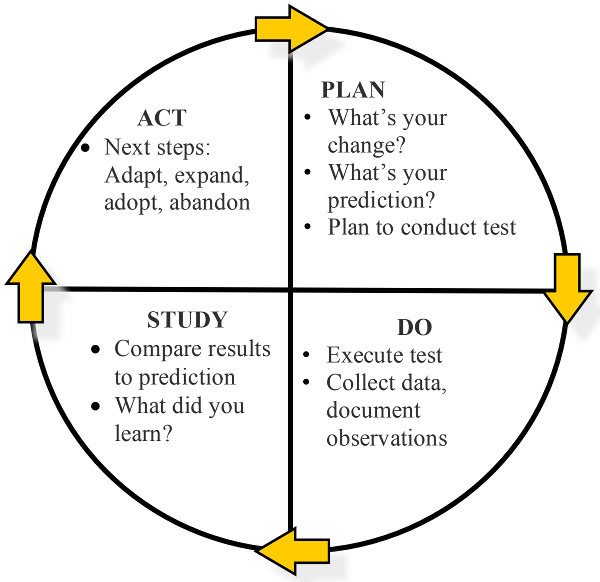

Figure 1. Plan-Do-Study-Act model.

In the "Plan" stage we articulate the change we intend to implement and record predictions about what we expect will happen. As, "Do' suggests, in this stage, we attempt the change and document what happens. The "Study" stage is where we compare the actual results to the predictions we made. Next, we "Act", deciding what to do next. Do we adapt, adopt, expand, or abandon the change idea? In improvement science PDSAs are used (Kawar, Mejia, Bennett, & Dolle, 2015) as

- the key mechanism by which we learn,
- a way to test and revise theories at an appropriate scale,
- a way to gain information by doing SOMETHING (even if it's small) rather than obsessing over getting it "right" from the start,
- a common approach that disciplines our efforts so we are efficient.

The predictions made in the Plan phase make explicit our understanding of the system we are working within and how we think our change idea will impact that system. The gap between our predictions and the actual results is where our learning happens. When what we predict comes true, we only have confirmation—suggesting that there is no gap in our understanding of how our system operates. If we are unable to explain why a test succeeded, then we may still have a gap in our knowledge, and we may wish to repeat the test to solidify our understanding of the system. When our predictions are wrong, we have exposed a gap in our knowledge providing an opportunity and a target to dig in to understand more about why things are the way they are.

The change ideas to be examined in PDSA cycles are taken from the driver diagram. The driver diagram is developed by establishing a clear aim, a specific statement of what we are trying to accomplish, and identifying primary drivers or factors which directly impact the aim. The change ideas are the actions identified as appropriate for affecting the primary drivers so that we move closer to our aim (Bryk et al., 2015). Whether or not our change is actually an improvement is determined by collecting and analyzing data.

This approach is different from the traditional approach often taken in education reform. The traditional approach involves choosing a change idea to implement, delaying implementation until the idea has been “perfected,” and then enacting the idea system-wide. When the change does not improve the system we have often done more harm than good to the whole system and have missed the opportunity to gain a deeper understanding of our system and the change idea from the failure (Kawar et al., 2015). Improvement science implies that testing change ideas should start small and slowly expand to learn along the way. Consequently, when we fully implement a change idea, we have gathered knowledge about not just the effectiveness of the change but also how to get the change to happen in various contexts.

In initial PDSA cycles the purpose is to determine how to get the change idea to work. The next cycles focus on learning how to get the change to work across multiple contexts and determining the support process needed to enact the change system wide. Finally, the change idea is integrated into the system (Kawar et al., 2015). Although, this process may seem long to address one small change, particularly in a complex system, PDSAs are focused cycles designed to move through the testing of change ideas relatively quickly. Moreover, PDSAs can be run in parallel so several changes can be tested on different parts of our driver diagram at the same time.

The scale at which we choose to test a change idea depends on several factors: (1) How confident are we that the change idea will lead to improvement?; (2) What is the cost of failure?; and (3) How resistant to change is our system? Low confidence in the change, with a high cost of failure and a system resistant to change suggests that we run a very small scale test. Whereas, a large-scale test might be appropriate when there is high confidence in the change, a low cost of failure and a system indifferent to change (Kawar et al., 2015). As we progress through PDSA cycles confidence in the change idea grows and the cost of failure decreases, consequently our system becomes more ready for change and we can implement the idea system wide.

PDSA FORM WITH PROMPTS

Test Title:		**Date:**			
Tester:		**Cycle#**:		**Driver:**	
What change idea is being tested?					
What is the goal of the test?*					

*Identify your overall goal: To make something work better? Learn now a new innovation works? Learn how to text in a new context? Learn how to spread or implement?

1. PLAN				**3. STUDY**
Questions: Questions you have about what will happen. What do you want to learn?	**Predictions:** Make a prediction for each question. Not optional.	**Data:** Data you'll collect to test predictions		**What were the results?** Comment on your predictions in the rows below.
			→	
			→	
			→	
			→	
Details: Describe the who/what/when/where of the test. Include your data collection plan.				**What did you learn?**

2. DO (Briefly describe what happened during the test, surprises, difficulty getting data, obstacles, successes, etc.)	**4. ACT** (Describe modifications and/or decisions for the next cycle; what will you do next?)

Figure 2. PDSA Form with Prompts. By Lawrence Morales and Alicia Grunow, ©2016 by the Carnegie Foundation for the Advancement of Teaching. Licensed for public use under the Creative Commons license CC-BY-NC-SA 2.0.

The PDSA form, Figure 2, provided to us by the Carnegie Foundation, is an effective tool for keeping a record of our progression through the testing process. Completion of the form focuses our thinking. We are reminded how the planned cycle fits into our overall driver diagram, our PDSA steps are clearly outlined, and a concise record of the results are available as the next cycle is planned.

Measures

Ideally, data should only be garnered for a clear purpose. Data can be used for accountability, research, and improvement (Bryk et al., 2015). If used for accountability, the data is examined to identify problematic or exceptional performers, with the possibility that the outcome may result in problematic performers being terminated. When data is used for research, theories are developed and relationships among variables may also be examined. Data used for improvement purposes seek to develop as well as evaluate change in practices, with the outcomes of data documenting the nature of change and relevant processes. As a network improvement community, the Clinical Experience RAC used data to improve clinical experiences to place a greater focus on the mathematics teaching practices (NCTM, 2014)

The Clinical Experience RAC of the Mathematics Teacher Education Partnership (MTE-P) uses a balanced set of measures, namely: outcome measures, process measures, and balancing measures, consistent with improvement science (Bryk et al., 2015; Lewis, 2015). Outcome measures consider how the system performs, and the overall result. Process measures evaluate whether the various parts of the system are performing as anticipated, and the balancing measures monitor adverse effects to other parts of the system as change is implemented. On a driver diagram, the outcome measures, which are often lagging indicators, provide insight into the extent the aim statement is achieved; the process measures examine secondary drivers, which are early indicators as to whether the proposed change is improving the system; and the balancing measures consider items that are not identified as primary or secondary drivers for the diagram to ensure the change is not resulting in unintended consequences. Using a balanced set of measures provides a holistic view of the implementation and sustainability of a change idea.

Clinical Experiences Research Action Cluster's (CERAC) Measures

The Clinical Experiences RAC is sub-divided into three groups: pair-placement, methods, and co-planning and co-teaching (CPCT). In the pair placement group, two teacher candidates are paired with one mentor teacher during the student teaching experience. The methods group focuses on preparing teacher candidates while they are enrolled in the mathematics education methods course, and the CPCT group encourages the teacher candidate and mentor teacher to teach lessons together. The aim statement for the RAC indicates teacher candidates should use the NCTM (2014) eight teaching practices at least once a week for the duration of their field experiences. The secondary drivers indicate that there is a need to: increase the amount of mentor teachers who are informed of national standards and current reform initiatives; attend to the teaching practices within methods courses; encourage teacher candidates to engage in self-assessment and reflect on the extent their enacted lesson embodies the teaching practices; facilitate collaborative meetings to discuss beliefs, complexities and challenges; and develop infrastructures to support teacher candidates' needs.

Common instruments (*MCOP^2*, *Mathematics Teaching Practices Survey*, and the *MTE-P Completer Survey*) are used across the three groups within the RAC to gather data for the outcome measures. The MCOP^2 measures K-12 classrooms instructional practices' alignment with national standards documents (Gleason, Livers, & Zelkowski, 2015). The Mathematics Teaching Practices Survey is a checklist tool used to identify whether any of the NCTM (2014) eight standards are addressed by teacher candidates during each day of their field experiences. The MTE-P Completer Survey asks teacher candidates to share their perspectives about the extent their teacher education program prepared them to be effective teachers. Therefore, the instruments are used explicitly to gather data to determine whether teacher candidates are taught to use the mathematics teaching practices in their teacher education program, and the extent to which they actually use the practices.

Each group within the RAC created focused instruments to align with the group objectives and to gain insight into the process measures. For example, the methods group created modules about the teaching practices and assesses the extent to which teacher candidates deemed the modules to be effective. The methods group also uses pre- and post-methods-course questionnaires. The data garnered provide insight into the nature of the methods course and the extent it prepares teacher candidates to address national standards. Similarly, the paired placement Sub-RAC uses surveys (of the mentor teacher and university supervisor) and focus groups to monitor the complexities of the field experiences and to refine the infrastructure. The CPCT group uses professional development, just-in-time, and exit surveys to monitor the extent to which teacher candidates and mentors are informed about the national standards and CPCT, to gain insight into personal reflections and self-assessments, and to learn about the nature of the clinical experiences. Hence, the process measures provide actionable data that can be used to reduce the likelihood that the change efforts are cultivating unintended negative results.

To gather balancing measures faculty members are encouraged to converse with mentor teachers, teacher candidates and supervisors, to discuss other variables that are also being affected due to the implementation of change ideas related to clinical experiences. Considering that time is valuable, and can have implications on how a reform idea is introduced and sustained, our balancing measures focus on out-of-class planning time. To garner these data, a weekly online survey could be sent to all concerned parties that asks about the amount of out-of-class time being used to prepare for enacted lessons.

Employing a balanced set of measures provides insight into the extent the clinical experiences RAC's overall goal was achieved, and whether the changes examined through PDSA cycles are having desirable effects. Considering that improvement science relies on PDSA cycles as a systematic process to gather data, changes can be made to the idea during iterations of the PDSA cycles to ensure the overall system is improved. Well-crafted measures and thoughtful implementation of PDSA cycles allow for the robust research design of improvement science to embrace variation in implementation and setting as important sources of information, and to learn from this variation to improve both the interventions and hopefully the system.

References

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

Gleason, J., Livers, S. D., & Zelkowski, J. (2015). *Mathematics classroom observation protocol for practices: Descriptors manual*. Retrieved from jgleason.people.ua.edu/uploads/3/8/3/4/38349129/mcop2_descriptors.pdf

Kawar, A., Mejia, E., Bennett, B. & Dolle, J. (2015). *Disciplined inquiry & testing strategies: Learning lab session II* [PowerPoint document]. Retrieved from Carnegie NIC Design Learning Lab: sites.google.com/a/carnegiehub.org/learning-labs-2015/home

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54-61.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.